创业与管理学院

School of Entrepreneurship and Management

SHANGHAITECH SEM WORKING PAPER SERIES

No. 2018-007

Clustering, Growth, and Inequality in China

Di Guo

The University of Hong Kong

Kun Jiang

The University of Roehampton

Chenggang Xu

CKGSB, HKU, Tsinghua

Xiyi Yang

ShanghaiTech University

September 2017

http://ssrn.com/abstract= 3159741

Clustering, Growth, and Inequality in China*

Di Guo, The University of Hong Kong
Kun Jiang, The University of Roehampton
Chenggang Xu, CKGSB, HKU, Tsinghua
Xiyi Yang, ShanghaiTech University

Abstract

This study examines the effects of China's industrial clusters on its economic growth and urban–rural income inequality. A density-based index (DBI) is developed to capture the unique features of such clusters in China. From a county-level DBI panel data constructed based on firm-level and county-level datasets, we find that strong clusters and entrepreneurial clusters substantially enhance economic growth. Moreover, entrepreneurial clusters reduce urban–rural income inequality by increasing the income of rural residents. Identification issues are carefully addressed by deploying an instrumental variable approach.

Keywords: China, clustering, geography, growth and inequality, institutions
JEL Classification: O14, O15, R11

---

* Corresponding author: Chenggang Xu, CKGSB, Tsinghua University, The University of Hong Kong, Pokfulam Rd., Hong Kong Island, Hong Kong (e-mail: cgxu@ckgsb.edu.cn). Di Guo, Kun Jiang, and Chenggang Xu acknowledge the financial support from the RGC Theme-based Research Scheme (TRS) (Project #: T31-717 *112*-R). Chenggang Xu acknowledges the financial support from the RGC General Research Fund (Project #: 756711) and CKGSB research support.

1. Introduction

This study investigates the co-existence of clustering, economic growth, and income inequality in China. The economic reforms in China transformed the world's largest developing country from one of the poorest nations into a major power (World Bank, 2013). The emergence of industrial clusters in numerous towns, mostly along China's coastal areas, is among the most striking developments throughout the said reforms. Various Chinese townships became national or international production centers for certain products because of the clustering of a large number of small entrepreneurial firms.[1] Considering the rise of industrial clusters as one of the major engines of China's growth is a warranted assertion (Fleisher et al., 2010; Sonobe and Otsuka, 2006b, Long and Zhang, 2011).[2]

However, along with China's record-breaking growth was the rapid increase in inequality. China has become one of the least equal economies in the world (Sicular, 2013), a status that might threaten the country's social stability and economic sustainability. The national Gini coefficient of household per capita income increased from 0.38 in 1988 to 0.49 in 2007 (Li et al., 2013), to 0.53–0.55 (Xie and Zhou, 2014) or even 0.61 in 2010 (Gan, 2013).[3] The urban–rural income gap remains a dominant component of the overall inequality (Li et al., 2013). Meanwhile, regional disparity, particularly inland–coastal disparity, remains an important dimension of the increased inequality during the reform era (Chen and Fleisher, 1996; Kanbur and Zhang, 1999).

The contrast between economic growth and the escalating inequality since the industrialization raises challenging questions (Kuznets, 1963). How does economic growth or development affect inequality and vice versa? Does a Kuznets curve (1963) explain such contrast in China? Do different kinds of institutions lead to alternative trends between growth and inequality (e.g., a la Acemoglu and Robinson, 2015)?

This paper studies how industrial clusters interact with economic growth and inequality. First, we determine whether clustering associated with different strengths and ownership structures, affects regional economic growth. Second, we investigate whether such clustering influences local urban–rural income inequality, and if so, through which channel. We are particularly interested in the impacts of clustered entrepreneurial firms. Based on available data, the best statistical proxy of entrepreneurial firms is non-state-owned firms (non-state firms in short) as discussed in the literature (Che and Qian, 1998a; Xu, 2011; Long and Zhang, 2011).[4]

[1] For instance, one-third of the world's socks, 40% of the world's neckties, and 60% of China's cashmere sweaters were produced in the towns of Datang, Shengzhou, and Puyuan, respectively (Xu and Zhang, 2009).

[2] According to Long and Zhang (2011), 62% of the growth of the number of firms in China from 1995 to 2004 was caused by the rise of these clusters; 14% of China's total industrial GDP growth during the same period was attributed to these firms within clusters.

[3] Although Gini coefficient estimations vary depending on the different data sources used, almost all studies show the same rising inequality trend in China over the past three decades. For instance, based on household data from nine provinces, Benjamin et al. (2008) find that the overall Gini coefficient of China in 2006 exceeded 0.50.

[4] For the same reason, our discussions designate clusters composed mainly of non-state firms as entrepreneurial clusters. However, when merely presenting statistics, we simply use the term non-state firms or non-state clusters when referring to the data.

To capture the distinctive features of the industrial clusters in China, we first define clustering measurements. Confined by China's institutional constraints to entrepreneurship and factor mobility, the country's industrial clusters differ from those in market economies in various crucial aspects. Labor and capital in market economies are mobile, and land is traded freely in the market. Under the assumption of mobility of factors, which captures the essence of a market economy, industrial agglomeration is an equilibrium outcome of the co-location decisions of firms in the spatial equilibrium model (Ellison and Glaeser, 1997; Glaeser and Gottlieb, 2009) or the new economic geography model (Krugman, 1991b). Industrialization and urbanization often co-occur in market economies.

However, at the onset of China's reform in the end of the 1970s, all input factors were completely controlled by the government, restricting the freedom of entrepreneurial firms in choosing their locations. Location decisions of state-owned firms are made by the government, and non-business factors often play essential roles. The relaxation of the controls has been gradual and partial. Particularly, one of the top constraints (which is still in place today) is the prohibition of free trading agricultural land when the land is to be used for non-agricultural purposes. The second important restriction is the *Hukou* system (residence registration system), which especially restricts the labor mobility of peasants, particularly their movement from rural to urban areas (Meng, 2000; Wang, 2005; Au and Henderson, 2006a). Although the restriction of the *Hukou* system has been reduced gradually since the mid-1990s, the relaxation is partial, and its effects are long lasting and substantial. The third constraint involves entrepreneurs and rural areas in general and peasants and small-medium enterprises (SMEs) in particular that have limited access to formal financial resources (Lin and Li, 2001). Consequent to these restrictions, a substantial proportion of entrepreneurial industrial firms in China are clustered in officially defined rural towns, many of which could be regarded as urban areas by general economic geography standards, especially in the 1980s and 1990s (Oi, 1999). Given the location restrictions and financial constraints, firms agglomerate into this kind of clusters are typically highly related in technologies and are small in size.

In contrast to the clusters of entrepreneurial firms are the industrial cities or agglomerations of highly specialized gigantic state-owned enterprises (SOEs), such as Kelamayi of Xinjiang Province, Yinchuan of Ningxia Province, and Panzhihua of Yunnan Province. These agglomerations of SOEs were established by central planning regimes, and market forces were minimally involved. Some large cities have agglomerations of a mixture of large SOEs and numerous entrepreneurial firms including SMEs also exist, such as Shanghai, Beijing, Shezhen, and Tianjin.

To capture the distinctive features of the industrial clusters created and developed under the institutional restrictions of China's economic reform and reduce the noise in the data created by industrial agglomerations of the specialized large SOEs, we construct a density-based index (DBI) of clusters according to the density of firms of each industry within a geographical location. Compared to measurements used in the existing literature, our measurement is more consistent with the observations documented by other sources on China's economic activities, such as satellite photos of night lights. Applying our DBI measurements, we construct a panel of county-level cluster indices that measure the existence, strength, and ownership structure of industrial clusters in each county. The indices are based on the firm-level panel dataset from the Above-Scale Industrial Firm Panel (ASIFP)[5] between 1998 and 2007.

Based on our county-level DBI cluster indices, we find that industrial clustering in China significantly enhances local economic growth. Significantly higher economic growth is present in counties with entrepreneurial clusters and with strong clusters (measured by the contribution of the clusters to national outputs or establishments). More interestingly, we discover that entrepreneurial clusters substantially reduce urban–rural income inequality, and this outcome is driven by the increased income of rural residents in the counties where these clusters are located.

Two-stage estimations are deployed to address identification concerns. Regarding economic growth, we use per capita mining outputs in the region and land sales revenues of the county governments as instrumental variables (IVs) of clusters. Mine-rich regions are typically dominated by large companies, smaller businesses are often marginalized, and entrepreneurship is often weakened (Chinitz, 1961; Rosenthal and Strange, 2003; and Glaeser et al., 2010, 2015). More importantly, such situation affects ownership given the state ownership of mining rights in China. Therefore, provinces with higher per capita mining outputs have weaker industrial clusters and fewer private firms in clusters. Furthermore, because the mine richness of a region is determined geologically, the importance of mining in a region is unrelated to regional economic growth. Thus, we believe the per capita mining output of a province is exogenous.

The government's land revenue is relevant to the strength and ownership structure of clusters because in most localities, industrial cluster development and entrepreneurial activities depend heavily on the government's land conversion (mostly through selling) from agricultural to non-agricultural use.[6] Moreover, land prices are higher in regions where clusters are stronger and entrepreneurial activities are more intensive. Therefore, in counties where the government obtains more revenue from the land, clusters are like to be stronger and may contain more privately owned firms. Additionally, we believe that without entrepreneurial activities in clusters, the local government's land revenue is not independently related to local economic growth because localities with weak entrepreneurial activities have low demand for land. The first stage regression results are consistent with our intuitions on the relevance of the two IVs. The Sargan test results also confirm that the two IVs are exogenous. Our second-stage regression results confirm our findings.

With respect to urban–rural inequality and the income of rural residents, we use Buddhist temples at the county level and the inflow of migrant labor at the provincial level as two IVs to identify the effects of clustering. We regard the existence of a historical Buddhist temple in a county as exogenous. In addition, we expect places

[5] ASIFP is composed of virtually all manufacturing firms in China with annual sales of RMB 5 million (US$ 750,000) or more between 1998 and 2007. The database provides detailed financial information and other firm-specific information, including location, industry, age, and ownership structure.

[6] According to the law, nationalization of collectively owned agricultural land is necessary before the land can be used for industrial or commercial usage; or traded in the market for non-agricultural purposes.

wherein Buddhist culture is popular to be likely to foster greater cluster development. Such supposition is because a Buddhist temple can be used as a proxy of the social capital as Buddhism emphasizes social harmony (Pan et al., 2012), which might mitigate contract enforcement in localities via its facilitating social networks.

Regarding our second IV (the inflow of migrant labor), we postulate that provinces with more developed industrial clusters would hire more immigrant workers from other provinces. Moreover, because as inequality in our paper is measured at the county level, provincial level immigrant measurement should not have a direct relationship with county level inequality. Our first stage regression results confirm that the two IVs are relevant and exogenous. Our second-stage regression results also confirm our findings on clustering and urban–rural inequality.

Our discoveries contribute to the literature on economic development and inequality. Various studies report a negative relationship between growth and inequality (e.g., Alesina and Rodrik, 1994; Persson and Tabellini, 1994; Knowles, 2001), particularly in developing economies. Nevertheless, such relationship can be positive in high- and mid-income nations (e.g., Forbes, 2000; Barro, 2000; Banerjee and Duflo, 2003). Development and inequality are posited to be determined by institutions because a good institution simultaneously protects private property rights, promotes development, and reduces inequality (Acemoglu et al., 2002, 2005; Engermann and Sokoloff, 1997, 2000; Easterly, 2007). Indeed Benjamin, Brandt, and John Giles (2005, 2011) find that inequality impedes China's rural development, and Chinese villages with higher initial inequality grow slower and have the most protracted non-agricultural development. Accordingly, we find entrepreneurial clusters stimulate economic development and reduce local inequality, indicating that different kinds of institutions drive different trends similar to the argument of Acemoglu and Robinson (2015).

Moreover, our study provides a methodological contribution and new evidence for the research on economic geography and urban economics. The usual measurements of regional specialization (Krugman, 1991a; Porter, 1990) or industry concentration (Ellison and Glaeser, 1997) are based on market economy assumptions. Given the substantial institutional constraints to mobility and trade in China, applying these indices directly may not be the most suitable approach for our purpose because they mix entrepreneurial industrial clusters with the SOE-dominated agglomerations inherited from the central planning system[7]. The indices of clustering we constructed capture institutionally constrained entrepreneurial clusters in China.

Furthermore, in relation to the economic geography and urban economics literature, the current study provides new evidence on the effects of clustering on economic growth. No empirical consensus is available on the effects of clusters on growth in prior studies. Glaeser et al. (1992) find that employment and wage growths are positively correlated to the clusters of diverse industries in the United States. Nevertheless, a cross-country study of Henderson (2003) fails to detect growth-promoting effects from agglomeration. Our findings, which indicate that the effects of clusters on growth

[7] When the purpose is to measure agglomeration itself without considering the type and formation of clusters (through markets or through bureaucracies' planning), applying the Ellison–Glaeser index to China can still produce informative results, as Lu and Tao (2009).

depend on their ownership type, provide an innovative dimension in studying the effects of clusters on economic growth. Industrial clusters are nested in small cities. Thus, our findings complement the study of Au and Henderson (2006b), which finds that most of China's cities are sub-optimally small because of the country's institutional constraints.

Finally, our work contributes to the literature on industrial clusters and income inequality in China. Our new evidence that industrial clusters drive China's growth is consistent with previous findings (Long and Zhang, 2011; Nee and Opper, 2012; Sonobe and Otsuka, 2014). Yet, this paper also analyzes the effects of entrepreneurial industrial clusters on growth and inequality. From 1984 to 2005, urban–rural income disparity almost doubled (Sicular et al. 2007), and inequality in rural China is related to impeding institutions, such as the *Hukou* system (Whyte, 2010). Moreover, inter-regional income difference increases over time (Fujita and Hu, 2001; Kanbur and Zhang, 1999, 2005). The main sources of persistent regional inequality include structural and long–term factors (Candelaria et al., 2013) or policy, such as fiscal decentralization (Li et al., 2013). Our findings, that different types of clusters contribute divergent trends in the relationship between growth and inequalities, complement those of previous studies.

The rest of the paper is organized as follows. Section 2 introduces the institutional background of industrial clusters in China. Section 3 discusses the related literature. Section 4 introduces the data, sample and constructs the density-based indices. Section 5 describes the empirical findings on clustering, economic growth, and urban–rural inequality. Section 6 discusses the identification issues. Section 7 concludes this study.

2. Industrial Clusters in China

2.1 Characteristics and emergence of China's entrepreneurial industrial clusters

The definition for the term "cluster" or "geographical agglomeration" can vary depending on the purpose of a study. In general, it refers to the co-location choice of a group of firms. In the literature, clustering is discussed in two ways: regional specialization and urbanization. Clustering driven by regional specialization indicates a group of firms in the same industry or related industries that co-locate within a geographical territory (Porter, 1990; Krugman, 1991a). Conversely, clustering driven by urbanization denotes a group of firms from diverse industries that cluster together to form modern cities (Henderson, 1974). In both literatures, the central condition for the "clustering" decision or formation in a market economy is factor mobility: labor and capital are perfectly mobile and the land is freely tradeable in the market, whereas market prices of the factors affect the decision on co-location. In addition, industrialization and urbanization develop closely with each other.

However, the case of China is different. When private ownership of a business was illegal at the early stages of China's reforms, township–village enterprises (TVEs) emerged as a leading form of adaptation to the weak legal protection of property rights for capturing market opportunities (Weitzman and Xu, 1994; Che and Qian, 1998; Mukherjee and Zhang, 2007). TVEs are vaguely defined as collectively owned cooperatives as all township or village residents that "set up" the TVE own the firm collectively (Weitzman and Xu, 1994). The community government of the township or village "represents" the communal collective owners and is the de facto executive owner of the TVE (Byrd and Lin, 1990). All TVEs maintain an intimate relationship with local governments (Qian and Xu, 1993; Chang and Wang, 1994; Che and Qian, 1998). TVEs were a major engine of China's growth in general and of rapid rural industrialization in particular. The TVEs' share of the gross domestic product (GDP) increased from 14.3% in 1980 to 37.5% in 1995 (Xu, 2011). In the mid-1990s, TVE development peaked with employment reaching 61 million in 1995. Meanwhile, TVEs also played a key role in fostering entrepreneurship in China. A large proportion of entrepreneurial industrial clusters in today's China can be traced to locations with concentration of TVEs.

With the rapid development of TVEs and private firms since the early 1990s, both political and legal resistances to private ownership were gradually relaxed. (Private property ownership was formally recognized by the Chinese constitution in 2004.) Since the late 1990s, numerous TVEs have become privatized (Xu 2011). Such firms are becoming increasingly specialized and clustered together. With the concentration of a vast number of small specialized firms, many townships have become national or international production centers of specific products. In Zhejiang Province, the Songxia Township produces 350 million umbrellas annually, the Qiaotou Township supplies 70% of the buttons for clothing made in China (Hessler, 2007), and the Puyuan Township produces over 500 million cashmere sweaters per year (Ruan and Zhang, 2009). Many of these clusters -consist of privatized TVEs or their spin offs.

In addition to property rights, entrepreneurs face three other types of institutional constraints that prevent them from choosing locations freely for setting up their businesses. First, according to the constitution, urban land is state owned, whereas rural land is collectively owned by villages and is not tradable for non-agricultural usage. In reality, land has not been tradable before the mid-1990s. Afterwards, peasants, individually or collectively, remain prohibited from trading "their" land for non-agricultural purposes. The only way for peasants to use their collectively owned land beyond agriculture activities without losing land ownership is to establish industrial firms within their villages or towns.[8] This restriction is part of the reasons that a large proportion of China's industrialization process was implemented in officially defined rural areas through township–village enterprises (TVEs).

Second, the *Hukou* system restricts citizens' mobility, particularly the movement of peasants from rural to urban areas. *Hukou* is a household registration system that officially identifies a person as a resident of a specific area and the social welfare that such person may enjoy. Under the *Hukou* system, individuals are classified as "rural" or "urban" citizens and "local" or "non-local" residents. Converting one's registration status from a "rural" to an "urban" resident is rigidly controlled. A peasant who seeks to move from a rural to an urban area and take up a non-agricultural job used to be required the approval, which involved extremely difficult bureaucratic processes. Meanwhile, people working outside the geographical area of their *Hukou* identities (i.e.

[8] Land ownership restriction was somewhat relaxed in the recent 15 years, such that non-local entrepreneurs can lease a piece of "collectively owned" land to develop rural industrial firms by recruiting local peasants who collectively "own" the land. Nevertheless, developing real estate for urban residences before nationalization is strictly forbidden by the Constitution.

“non-local” citizen) are rendered unqualified for local social welfare, including housing, health care, and education benefits (Cai, 2000; Au and Henderson, 2006a). Whyte (2010) characterizes the *Hukou* system as “socialist serfdom.” Although the *Hukou* system has relaxed over time such that peasant migrants were allowed to work in cities as lower-level citizens, moving businesses to urban areas remains difficult because of various discrimination policies imposed on rural residents.

Third, the capital market in China is highly underdeveloped (Allen et al., 2005). China’s banking system is particularly biased against lending to private enterprises. Although the share of the private sector on the national total GDP soared to 50% in 2009, the short-term bank loans issued to the private sector was only 4.9% of the national total (Guo et al., 2014).

Because of the aforementioned institutional restrictions, the development and features of China’s industrial clusters differ from the concepts of “clustering” or “geographical agglomeration” defined in existing studies. First, industrial clusters in China tend to be defined by administrative boundaries. As discussed, TVEs and the subsequent privatized firms were the bases for various entrepreneurial industrial clusters today. These clusters emerged from the struggles and maneuvers of entrepreneurs and local governments under institutional restrictions. They are concentrated within the administrative boundaries of certain local governments, which facilitate and protect the interests of private firms. Second, in association with the *Hukou* system, “rural” and “urban” are terms that describe social status rather than reality. Therefore, a large percentage of employees of these firms are officially defined as peasants, although they do not perform any agricultural work (in cities they are called peasant-workers). Third, under institutional constraints, many clusters in China consist of numerous small private firms working closely with each other. Highly specialized clustering effectively decomposes the production process into many small steps, which weakens both technical and capital barriers to entry (Huang et al. 2008; Xu and Zhang, 2009; Long and Zhang, 2011; Ruan and Zhang, 2009)[9].

The rapid growth of entrepreneurial industrial clusters since the 1980s is an essential element in transforming China from a centrally planned and state-ownership dominated economy into a market economy.

3. Related Literature: Industrial Clusters, Growth, and Inequality

Both anecdotal and systematic evidence show that in general clustering areas are more productive than other locations (Ciccone and Hall, 1996). Such observation can be explained by the following: 1) the advantage of a large labor pool or large production/consumption varieties associated with these areas (Marshall, 1890; Krugman, 1991a; Rotemberg and Saloner, 2000; Combes and Duranton, 2006), 2) the reduced transportation costs across firms when firms are located closely to suppliers or customers (Krugman, 1991a, 1991b; Fujita et al., 1999), as emphasized by new economic geography (NEG) theory, and 3) information spillovers within a cluster

[9] Ruan and Zhang (2009) documents a typical case where nearly 12,000 small firms and over 70,000 people engaged in the production of cashmere sweaters in the Puyuan Township of Zheijiang, which is the largest production or trading center for cashmere sweaters in China.

because closely located firms exchange information effectively, particularly through informal channels. Different explanations involve urban variety (Jacobs, 1969) and/or regional specialization (Marshall, 1890; Arrow, 1962; Romer, 1986). Factor mobility is the key assumption in all of these explanations.

However, the manner in which clustering or geographical agglomeration affects regional economic growth remains a contested issue. First, NEG theory suggests that economic convergence occurs when per capita growth rate is likely to be related inversely to the starting level of output or income per person, resulting in the faster growth of the poorer rather than the wealthier regions (Solow, 1956). This convergence is based chiefly on the assumption that a perfect market exists for production resources, which have diminishing returns to scale (Baumol, 1986; Barro and Sala-i-Martin, 1992). When firms cluster together, the diminishing returns of the clustering is expected with the intensive competition and increasing prices of production inputs.

By contrast, endogenous growth theory explains growth through the Schumpeterian processes of creative destruction with a focus on entrepreneurship and innovation (Aghion and Howitt, 1992). Under this theoretical framework, the convergence or divergence across regions depends on how various regions adopt to new technologies or innovation. Aghion et al. (1999) emphasize that the role of organizational change in the production process (which specifies the way in which workers or organizations interact and learn from each other) may be crucial in determining productivity, and thus economic growth. A major advantage of clustering is the strong cross-firm spillovers that include the externalities generated by sharing knowledge, innovation, and entrepreneurial culture. Some empirical studies find a strong positive relationship between clustering and regional innovation (Saxenian, 1994; Audretsch and Feldman, 1996; Delgado et al., 2010; Kerr, 2010) and entrepreneurial activities (Chatterji et al., 2013; Delgado et al., 2010).

Overall, empirical evidence on how clustering influences regional economic growth is mixed. Based on US data, Glaeser et al. (1992) and Delgado et al. (2014) find that employment and wage growth are positively correlated to clusters composed of firms from diverse industries but not from regional specialization. Conversely, Henderson (2003), using a panel data for 70 countries from 1975 to 2000, failed to observe growth-promoting effects from agglomeration in any means. Brülhart and Sbergami (2009) find that agglomeration boosts the GDP growth of European Union countries but only up to a certain economic development level.

The relationship between growth and inequality has been a debated subject since Kuznets (1955). Kuznets regarded inequality as necessary and transitional in the urbanization or growth process (Kuznets, 1955). He finds that the relationship between industrialization and inequality in the US is an inverted U-shape between 1770 and 1970 (Kuznets, 1963). This result is described as the evolution of income distribution in the transition from a rural to an industrial economy. Income inequality should increase during the early stages of development (because of urbanization and industrialization) and decrease later (because industries would have already attracted a large fraction of the rural labor force).

However, recent findings challenge this theory because high income inequality is found being associated with the deceleration of economic growth in developed countries since the 1970s (Alesina and Rodrik, 1994; Persson and Tabellini, 1994; Perotti, 1996). The general unanswered question is how industrialization or economic growth affects inequality (for a detailed survey, see Aghion et al., 1999).

Along with its rapid economic growth, inequality increased substantially in the past three decades in China. Although the Gini coefficients vary depending on the different data sources of various studies, the rising trend of inequality in China during those decades remains evident. Li et al. (2013) claim that the national Gini coefficient of household per capita income increased from 0.38 in 1988 to 0.49 in 2007. Gan (2013) reports that the same coefficient reached 0.61 in 2010. Benjamin et al. (2008) indicate that the overall Gini coefficient of China exceeds 0.50.

The urban–rural disparity is a dominant component of China's overall inequality. Using provincial-level data from 1983 to 1995, Kanbur and Zhang (1999) prove that urban–rural inequality accounted for over 70% of the overall inequality. Using household survey data of two Chinese provinces, Yang (1999) finds that the urban–rural inequality explains 82% and virtually all the overall inequality changes in Jiangsu and Sichuan Provinces, respectively, from 1986 to 1994. A recent study shows that the contribution of the urban–rural divide in terms of both disposable income and wealth to the overall national income inequality rose from 37% in 1988 to 54% in 2007 (Li et al., 2013).

4. Data and the construction of DBI

4.1 Data Sources

A panel of approximately 2,815 Chinese counties (including county-level city districts) is constructed by combining several data sets from 1998 to 2007. First, our key explanatory variables, including the existence and features of clusters, are constructed based on data in the Above-Scale Industrial Firm Panel (ASIFP) from 1998 to 2007, which covers all state- and non-state-owned industrial firms with annual sales of 5 million RMB or above, including information on industry, location, age, ownership, financial information, and ownership at firm level. The enterprises included in this database account for 90% of the total sales of all industrial firms in China.[10] However, the ASIFP excludes firms with annual sales under 5 million RMB, almost all of which are non-state-owned firms. Thus, the potential biasness of our findings should be underestimating the impacts of entrepreneurial clusters (private firms). As a robustness check, we apply the same methodology to identify clusters using the Chinese Economic Census data in 2004, which includes industrial firms of all sizes. The identified clustering patterns are qualitatively consistent with those identified based on the ASIFP.

Second, data on county-level per capita GDP, per capita household income, and other general county-level economic and demographic variables (e.g., total GDP, rural and urban populations, and investment in fixed assets), are all from the China Socio-

[10] In the first Chinese Economic Census conducted in 2004, the amount of the total sales for all industrial firms was 218 billion RMB, whereas that of the total sales for all the ASIFP firms was 196 billion RMB.

Economic Development Statistical Database.[11] Per capita household income statistics include rural household per capita net income and urban household per capita disposable income in the said database. Rural household net income is defined as the total family income excluding the family business expenses, depreciation of productive fixed assets, taxes, and land contract fees. Urban household disposable income is the total family income minus personal income tax and expenditures on social security. The panel is unbalanced because of missing data for certain counties and years.

Third, we control some county-specific variables in the empirical investigation. Public finance data, including various county-government expenditures, are from the *National Prefecture and County Public Finance Statistical Yearbooks* for the same period. Furthermore, we construct a panel to indicate whether a county is officially designated as a "National Poor County."[12] Each year, these designated poor counties receive sizeable amounts of fiscal transfers from the central government. This subsidy may affect the local income and the urban–rural inequality that we aim to investigate. The list of the poor counties is obtained from the official website of the State Council. Finally, to differentiate the effect of clusters from that of Special Economic Zones (SEZs), we construct a panel that indicates the existence and number of provincial-level SEZs in each county from 1998 to 2007. The list of SEZs is obtained from the website of the Ministry of Commerce of the People's Republic of China.

All the above data are deflated to 1998 price levels when applicable. During our sample period, some counties changed their names or judiciary boundaries. New counties were established, while some existing counties combined to form larger ones or were elevated into cities. We identify the changes and convert the corresponding county codes into a benchmark system. China also modified its industry coding system in 2002 (from GB/T 4754-1994 to GB/T 4754-2002). The four-digit industry codes that have become either more disaggregated or more aggregated after 2002 were tracked, and the aggregated codes are used to group the industries from 1998 to 2007.

All variables used in this study are defined in Appendix 1 (Table A.1).

4.2 Measuring Industrial Clusters in China

Constructing clustering indices to capture entrepreneurial clusters in China is a major challenge. As discussed, a key assumption in the literature on clustering or geographical agglomeration is that factors and firms can move and choose their location freely.

Under the standard free market assumption, the clustering indices constructed in existing studies focus either on regional specialization or on the inter-connectedness of industries or firms within a locality (Porter, 1990; Krugman, 1991a; Glaeser et al., 1992). Most studies on regional specialization apply the Herfindahl–Hirschman index

[11] Most of the data in this database are official information from prefectural- or county-level yearbooks. Thus, the reported household income data generally cover the incomes of local residents and exclude those of migrant workers.

[12] Two rounds of the poverty reduction program (known as the 8-7 Plan) of the government were conducted in China during 1986– 1993 and 1994– 2000, aiming to promote local economic development through targeted public investments with fiscal transfer. In 1986 and 1994, the Poverty Reduction and Development Team supervised by the State Council (*Guo-wu-yuan Fu-pin Kai-fa Ling-dao Xiao-zu*) published two National Poor County lists. The 1994 list was modified further in 2006 and 2012. As of 2012, there were 592 national level poor counties in the list.

(HHI), Gini coefficient (Gini), or location quotient (LQ) measure to capture the entire distribution of regional output across industries and gauge the degree of regional specialization in a few industries.[13] The Krugman dissimilarity index (Krugman, 1991a) focuses on the deviation of a region's industry structure from the average industry structure of a regional reference group to reveal a region's comparative advantages. Following Krugman and Venables (1995), studies on industrial clusters emphasize the "inter-connectedness" of industries and firms in a region. The assumption is that manufacturing firms benefit from being located in a region where they have access to suppliers of specialized inputs. This argument is supported by empirical evidence (Feser and Bergman, 2000; and Porter, 2003). Based on the revealed comparative advantage in product export, Hausmann and Klinger (2007) construct a proximity measure for all four-digit Standard International Trade Classification products. Long and Zhang (2012) and Long, Zhang and Zhang (2015) employ this proximity index to measure clusters in China.

However, employing regional specialization or inter-connectedness measurements directly to identify entrepreneurial clustering in China may not be the most suitable method for our purpose. This is because restrictions on factor mobility and location decisions of firms may create biases and potential measurement errors. At the onset of the economic reform, all firms were owned or controlled by national or local governments and their locations were chosen by administrative decisions. As such, the concentration of heavy industries in certain areas of China was driven mostly by political concerns, including those related to national security. Today, in commanding heights sectors similar to Lenin's New Economic Policy as instructed by Deng (Deng, 1986), for industries including finance, energy, mining, railway, airlines, and communication, state ownership still dominates. Most of the 69 Chinese enterprises among the world's top 500 in 2011 are SOEs. The regions with these giant SOEs are specialized, with high specialization scores measured by HHI, Gini, or LQ. For instance, oil refining and processing SOEs contributed 24.5% to the local industrial output in Daqing City in 2007. Four SOEs in particular dominated this industry, which accounted for 89.39% of the industrial outputs. Changchun City is highly specialized in manufacturing transportation equipment, which contributed to 68.26% of the industrial output there in 2007, with 79.62% of the outputs coming from 13 gigantic SOEs.

As discussed in Section 2, the entrepreneurial clusters in China are characterized by the emergence of numerous "specialty towns," each of which produces a particular type of product, such as sweaters, socks, furniture, and electric appliances. Each cluster consists of large number of small firms and workshops. Outputs of many clusters comprise significant shares of national or global markets. These observations suggest

[13] For instance, Glaeser et al. (1992) focus on the contribution of a region's top five largest industries to the local economy to reveal the extent to which a given region is specialized or diversified, regardless of how the economic structure of the country as a whole evolves. The Gini measures how far away a country or region is from an equal distribution in which each industry produces the same share of output or value added. Midelfart–Knarvik et al. (2000) use Gini to explore industrial location changes in terms of spatial concentration in Europe. LQ is an analytical statistic that measures a region's industrial specialization relative to a larger geographic unit (usually the nation). Glaeser et al. (1992) apply LQ as a specialization measure of an industry in a city and test its effect on city-industry employment growth. Porter (2003) utilizes LQ as an important criterion in defining traded industries that form clusters.

that in addition to specialization, density of firms in an industry within a locality is one of the most important features of entrepreneurial clustering in China. Hence, we construct a density-based index (DBI) to measure entrepreneurial clusters.[14] The DBI counts the number of firms in the same industry within a locality, e.g. a county. For the DBI at the county level, we denote the number of firms in industry $\mathrm{j} \in \{1, 2, \ldots J\}$ in county $\mathrm{i} \in \{1,2, \ldots I\}$ at time t as $fn_{j,i,t}$. We define county i as "a county with an α cluster of industry j" if the number of firms in this county is among the top α percentile of all counties in this industry at time t in the nation. Formally, we define

$$c_{j,i,t} = \begin{cases} 1, \ if \ \ fn_{j,i,t} \geq (100-\alpha) \ \ percentile \ \ of \{fn_{j,1,t}, fn_{j,2,t}, \ldots, fn_{j,I,t}\} \\ \qquad\qquad 0, \ otherwise \end{cases};$$

And

$$C_{i,t} = \sum_{j \in C_{it}} c_{j,i,t}.$$

In this paper, we focus on the top five percentile county level clusters. To simplify the expression, for the remainder of this paper, the term cluster means α = 5, and we will omit mentioning this unless a definition is specified.[15] The variable $Cluster_{it} = \begin{cases} 1, \ if \ \ C_{i,t} \neq \emptyset \\ 0, \ otherwise \end{cases}$ captures the existence of the DBI cluster in county i in all industries in year t.

Identifying the strength of clusters is essential because counting the number of firms does not capture the sizes of firms. We define the output[16] strength of each cluster in industry j located in county i at time t, by the ratio of its output contribution over the national average level in industry j. Specifically, the output strength of each cluster ji at time t is defined as $S\_national\_V_{jit} = \frac{S\_V_{jit}}{\frac{1}{I}\sum_{i=1}^{I} S\_V_{jit}}, \ where \quad S\_V_{jit} = \frac{Output_{jit}}{Output_{jt}}$ is the output share of the cluster ji in the national total output of industry j in year t, whereas $\frac{1}{I}\sum_{i=1}^{I} S\_V_{jit}$ is the average S_V_{jit}. $S\_national\_V_{jit} > 1$, if output strength of cluster ji at time t is stronger than the national average; otherwise, $S\_national\_V_{jit} \leq 1$. Similarly, we define the establishment strength of each cluster ji at time t as $S\_national\_E_{jit} = \frac{S\_E_{jit}}{\frac{1}{I}\sum_{i=1}^{I} S\_E_{jit}}$, where $S\_E_{jit} = \frac{Establishment_{jit}}{Establishment_{jt}}$ is cluster ji's share in the national total of establishments in industry j, and $\frac{1}{I}\sum_{i=1}^{I} S\_E_{jit}$ is the national average number of establishments in industry j.

To measure cross industry aggregate strength of clustering in each county i based on the strength ratios defined above, we construct the overall strengths of clusters in the following. The overall output strength of the clusters in county i is defined as the weighted average of the strength of each cluster (if any):

[14] Ciccon and Hall (1996) study clusters by measuring labor intensity and physical capital.

[15] Our results stay robust when we assign other values for α, such as 3 and 8.

[16] The output and establishment data used in calculating cluster strength indices are based on the firm-level data from the ASIFP. For our county level cluster measurement, the output of a given county in a given year is the aggregated output of all firms located in the county in that year.

$$Str\_national\_V_{it} = \begin{cases} \dfrac{\sum_{j\in C_{it}} output_{jit} S\_national\_V_{jit}}{\sum_{j\in C_{it}} output_{jit}}, & if \ \ C_{it} \neq \emptyset \\ 0 \ , & if \ \ C_{it} = \emptyset \end{cases}.$$

Similarly, the overall establishment strength of the clusters in county *i* is defined as

$$Str\_national\_E_{it} = \begin{cases} \dfrac{\sum_{j\in C_{it}} establishment_{jit} S\_national\_E_{jit}}{\sum_{j\in C_{it}} establishment_{jit}}, if \ \ C_{it} \neq \emptyset \\ 0 \ , & if \ \ C_{it} = \end{cases}$$

The ownership structure of clusters in county *i* is measured by the share of non-state firms[17] in the total outputs of the clusters in that county,

$$Nonstat\_V_{it} = \begin{cases} \dfrac{\sum_{x\in X_{it}} output_{it}}{output_{it}}, & if \ \ C_{it} \neq \emptyset \\ 0 \ , & if \ \ C_{it} = \emptyset \end{cases},$$

where $\boldsymbol{X_{it}}$ is the set of non-state firms in county *i* at year *t*. The share of non-state firms in the total establishments of the clusters in that county is expressed as

$$Nonstat\_E_{it} = \begin{cases} \dfrac{\sum_{x\in X_{it}} establishment_{it}}{establishment_{it}}, & if \ \ C_{it} \neq \emptyset \\ 0 \ , & if \ \ C_{it} = \emptyset \end{cases}.$$

Table 1 presents the summary statistics of DBI clusters. Taking year 2007 as an example, among 2,734 counties, 739 (about 27%) have clusters. The total number of industrial clusters is 2,213, which accounted for 5% of the 44,175 county-industry observations. These clusters contributed 38% of the national total outputs and 37% of total employment. By average, the output of a county with clusters is about 6.5 times that of a county without. In addition, the number of establishments in an average cluster county is about 5.9 times that of a non-cluster county. On average, more than 83% of firms in the clusters are non-state owned, and more than 80% of the clusters' outputs are from non-state firms.

Table 1 also shows the dynamics of cluster development in Chinese counties. In our sample period (1998–2007), 294 counties (about 10% of all the counties), always have some clusters in at least one industry. Conversely, 1,576 counties (about 56% of all the counties) have never developed any industrial cluster. A total of 317 counties initially did not have clusters in 1998, but they developed industrial clusters by 2007. Out of these counties, 52 are from Shandong and 25 are from Fujian. By contrast, 292

[17] Non-state-owned firms refer to firms where state capital constitutes less than 50% of the total paid-in capital.

counties had clusters operating in 1998, but clusters in these counties disappeared by 2007. Most of these backward developments occurred in inland areas.

Table 2 presents comparisons between provincial level DBI cluster indices and some standard cluster indices. All measurement results are based on the 2007 ASIFP data. Measured by DBI cluster count, the top five provinces are Zhejiang, Jiangsu, Guangdong, Shandong and Shanghai. Measured by DBI output strength, the top five provinces (in descending order) are Shanghai, Tianjin, Zhejiang, Shandong, and Jiangsu. Finally, when measured by DBI firm (establishment) strength, the top five provinces are Zhejiang, Shanghai, Jiangsu, Shandong, and Tianjin. The strongest five top-ranked provincial regions in DBI non-state cluster indices (in descending order) (measured either by output volume or by establishments) are Zhejiang, Shanghai, Jiangsu, Shandong, and Tianjin. The common feature shared by all the results above is that the top provinces are all coastal regions. Note that this common feature is consistent with the general perceptions on spatial distributions of entrepreneurial clusters in China.[18]

In contrast, applying standard measurements to Chinese data tends to capture lots of agglomerations of highly specialized large SOEs, many of which are located in interior regions. For example, measured by HHI the top 5 provincial regions will be Xinjiang, Shanxi, Hainan, Jilin, Gansu; or Xinjiang, Qinghai, Gansu, Yunnan, Hainan if measured by Gini; or Tibet, Ningxia, Xinjiang, Qinghai, Yunnan if measured by LQ; and Shanxi, Tibet, Xinjiang, Qinghai, Yunnan if measured by Krugman Index. Except for Tibet (which is one of the most under-developed regions in China), all these provinces are known for a concentration of SOEs and their weakness in terms of entrepreneurial activities. Some of the provincial level cluster measurements shown in Long and Zhang (2011), which applies the Industrial Proximity measurement (Hausmann and Klinger, 2007), are closer to ours; but some other results reflect noises. The top five provincial regions by their measurement are Tibet, Beijing, Jilin, Zhejiang, and Ningxia. We further compare the average weights of SOEs in the regions defined by different clustering measurements as shown in Table A.2. The comparison further confirms that standard clustering measurements tend to capture the specialization or concentration of SOEs under the centrally planning economy.

Figure 1 illustrates how standard approaches and DBI capture China's clusters geographically by showing maps of cluster counts at the prefectural level. The map shows that the DBI clusters are concentrated heavily along coastal line regions, which is highly consistent with a satellite night vision of China (Figure 2). By contrast, coastal line regions are not captured adequately by standard clustering measurements.

Table 3(a) presents the summary statistics of the dependent variables for counties with and without DBI clusters. Compared with other counties, on average, counties with clusters grow faster, and their urban and rural residents are more equal. Between 1998 and 2007, the average growth rate of the counties with clusters is 1.3% higher than other counties, whereas the urban–rural per capita income ratio in counties with clusters is lower by about 20% than other counties.[19] Table 3(b) provides the summary

---

[18] For instance, studies that document clusters in Wenzhou City (Huang et al., 2008), Tongxiang County (Ruan and Zhang, 2009), and Wuxing County (Sonobe et al., 2002).

[19] These summaries are based on county level aggregate data. The individual level data obtained from a household survey, the 2007 Chinese Household Income Project (CHIP) database, confirms these summaries. The CHIP data covers 92 rural counties and 38 of them have DBI clusters (much higher than the population). We find that in counties with clusters the individuals' average non-agricultural income is about 8% (1,398 vs. 1,293 in Table A.3) higher than that in counties without clusters, and the income for individuals engaged in non-agricultural activities in counties with clusters is about 20%

statistics of the characteristics of counties with and without DBI clusters. On the one hand, counties with clusters have higher per capita GDP on average, are more industrialized, and have more private firms than other counties. On the other hand, government and education expenditures in counties with clusters are only half of those in other counties.

5. Baseline Estimations and Findings

In this section we present our baseline regressions, which investigate the relationships between clustering and growth (Section 5.1) and between clustering and inequality (Section 5.2).

5.1 Clustering and regional economic growth

Here, we examine how clusters are related to local economic growth. Our hypothesis is that clusters should be positively and significantly associated with economic growth. Concretely, we expect counties should have higher growth rates than others if they have clusters, or if they have stronger clusters, if they have entrepreneurial clusters. We test this hypothesis by estimating a type of Barro model (Barro, 2000):

$$\ln\left(\frac{p.c.GDP_{it+1}}{p.c.GDP_{it}}\right) = \alpha + \beta Cluster_{it} + \gamma \ln(p.c.GDP_{it}) + \mu[\ln(p.c.GDP_{it})]^{2} + \tau \ln(CPI_{pt}) + \boldsymbol{\delta Z_{it}} + \varepsilon_{it}, \quad (1)$$

where $p.c.GDP_{it}$ is the per capita GDP of county *i* at year *t*, and the dependent variable represents the annual growth rate of per capita GDP. Our major explanatory variables are $Cluster_{it}$, which is a dummy variable that equals to one if at least one cluster operates in county *i* in year *t* and zero otherwise; the strength of clusters ($Str\_national\_V_{it}$, and $Str\_national\_E_{it}$); and the ownership structure of clusters ($Nonstate\_V_{it}$ and $Nonstate\_E_{it}$).

The initial level of per capita GDP controls for any convergence or divergence effect, and the square term of per capita GDP controls the speed of convergence or divergence. Provincial level consumer price index (CPI) is included to control for the inflation effects. $\boldsymbol{Z_{it}}$ is a vector of other county-level control variables, which include *fraction of industrial outputs to total GDP* for capturing the local economic structure; *fraction of non-state owned firms* and *fraction of micro firms*[20] for controlling the effect of privatization and small businesses; *fraction of education expenditure to GDP; fraction of investment in fixed assets to GDP* (including investment in infrastructure, renovation, and real estate among others) for controlling the local investment in human and physical capital; and *fraction of government expenditure to GDP* to control for

(42.76% vs. 35.58% in Table A.3) more than that in counties without clusters. Moreover, in counties with clusters, more than half of the local residents engage in non-agricultural activities, which is almost twice the number of such residents in counties without clusters. However, the sample of CHIP covers only about 3% of the counties in the nation. Given the small sample size and the sampling bias of the survey in selecting counties, we are unable to run regressions by using this dataset.

[20] The fractions of non-state-owned and micro firms are derived from firm-level data from the ASIFP. For instance, for any county during the sample period, we calculate the total number of non-state-owned firms or micro firms and divide it by the total number of firms in the county in that year to obtain the fraction.

government expenditure. All the above ratios are in log form. The *number of SEZs* within a county is included to control the effects of SEZs because those often involve intensive and extensive government efforts. We also control for the national governments poverty eradication policy, as this policy involves fiscal transfers aiming to boost the growth and individuals' income for identified "Poor Counties". For this purpose, we include a dummy variable *Poor,* which is equal to one if the county is identified by the State Council as a National Poor County and zero otherwise.

County and year fixed effects are controlled in the panel analysis to address county- and time-specific effects. As each county's observations in the panel data are auto-correlated, following Peterson (2009) we group the standard errors within counties. Furthermore, as a robustness check, utilizing the fact that certain provinces have high concentration of industrial clusters, we conduct estimations by grouping standard errors at the provincial level. The results are robust, as shown in Appendix 2 (Table A.2.1–Table A.2.3).

Our major results are shown in Table 4. Column (1) of the table shows that *Cluster* is insignificant, indicating that when we pool all clusters together without differentiating for ownership and strength, existence of clustering alone has no significant impact on growth. However, Columns (4) and (5) demonstrate that entrepreneurial clusters, i.e. clusters dominated by non-state firms (measured by *Nonstate_V* or *Nonstate_E*) are positively and significantly associated with growth. A 1% increase in the contribution of the non-state sector (either *Nonstate_V* or *Nonstate_E*) will result in approximately 1.5% increase in the per capita GDP growth. Moreover, Columns (2) and (3) indicate that the strengths of clusters (*Str_national_V* and *Str_national_E*) are positively and significantly associated with growth. A 1% increase in the clusters' output strength will result in a 1% increase of per capita GDP growth.

For all regressions, both initial levels of per capita GDP and the squared term of per capital GDP are significantly and negatively associated with growth, suggesting a mean convergence in the economic growth among Chinese regions, and the convergence's acceleration over time. However, the dummy variable *Poor*, defined by poverty-relief-fund receiving poor counties, is negative and significant in all regressions. This finding suggests that these counties are not converging to average counties, i.e., the regional income inequality gap between the officially designated poor counties and the rest of the counties is widening. This might suggest the ineffectiveness of the national poverty-relief program[21].

Regarding the structure of the local economies, estimated coefficients of *fraction of industrial output* and *fraction of non-state firms* are all significant and positive, thereby implying that counties with higher levels of industrialization and more non-state firms experience higher economic growth rates. Furthermore, expenditures for education and fixed investments are significantly and positively correlated to local

[21] Our result is not necessarily in contradiction to the discoveries that fiscal transfers to poor counties improved those economies significantly (Meng, 2013), because without the transfers, the cross region inequality gap might increase even faster.

economic growth, indicating that regions that invest more in human and physical capitals experience higher economic growth. Lastly, local government expenditure is positively correlated to high growth, which denotes a positive correlation between fiscal revenue and growth rate.

Up to now, what we have demonstrated is the associations between entrepreneurial clusters and economic growth. In Section 6, we deploy the instrument variable approach to address identification problems.

5.2. Clustering and urban–rural income inequality

In this subsection, we examine the relationship between clustering and urban–rural income inequality, and explore how clustering affects household income in urban and rural areas. We use urban–rural household per capita income ratio as proxy for urban–rural inequality. In our sample, the county-level urban–rural income ratio increased from 2.08 to 2.69 (on average) from 1998 to 2007. Our summary statistics indicates that by average counties with DBI clusters have smaller urban-rural income gap than others. Thus, our hypothesis is that clusters, measured by different DBIs, should be negatively and significantly associated with urban-rural income ratio. Our baseline regression model of urban–rural inequality vs. clustering is the following equation.

$$\ln(ratio_{it}) = \alpha + \beta Cluster_{it} + \boldsymbol{\delta W_{it}} + \varepsilon_{it}, \qquad (2)$$

where $ratio_{it}$ refers to the urban–rural household income inequality measured by the ratio of urban household per capita income over the rural household per capita income in county i in year t. The explanatory and major control variables $\boldsymbol{W_{it}}$ are the same as those in Equation (1) except for the inclusion of the total GDP of the county and dropping the square term of the per capita GDP and CPI.

The estimation results are reported in Table 5. Columns (1) indicates that by pooling all clusters together without differentiating for the ownership structure of clusters, cluster is negatively but insignificantly associated with local urban–rural inequality. The correlation becomes statistically significant when we focus on non-state clusters, as shown in Columns (4) and (5). These findings indicate that an increase of 1% in the number and outputs of non-state-owned firms within clusters is associated with a 3% reduction in the urban–rural inequality in the county.

However, when focusing on strength of clusters in terms of output, Column (2), or in terms of number of firms, Column (3), the statistical relationship between clusters and inequality disappears. These five regression results suggest that non-state firms in clusters are associated with the reduction of urban–rural inequality, but clusters dominated by other type of firms do not show the same association.

Furthermore, in all five columns, investments both in fixed assets and in education are positively and significantly correlated to urban–rural inequality. These outcomes indicate that growth enhancing efforts, such as investments, FDI and exporting, are associated with widening the income gap between urban and rural residents from 1998 to 2007.

A positive association between growth enhancing efforts and the widening of inequality may be unsurprising. However, negative associations between entrepreneurial clusters and inequality seems challenge a popular view that a market economy dominated by privately owned firms tends to worsen inequality, and the view that one of the benefits of maintaining state-owned firms is that it contributes to containing inequality. To understand why the development of non-state firms in clusters might reduce inequality, we subsequently explore its mechanism. Noticing that entrepreneurial clusters observed today are located mostly in areas with a concentration of TVEs in the 1990s, it is likely that entrepreneurial clusters employ more rural laborers, which contributes to the rural residents' income and reduces urban–rural inequality. Our baseline regression models for testing this hypothesis are as follows.

$$\ln(rural\ \ household\ \ income_{it}) = \alpha + \beta Nonstate_{it} + \boldsymbol{\delta W_{it}} + \varepsilon_{it}, \qquad (3)$$

$$\ln(urban\ \ household\ \ income_{it}) = \alpha + \beta Nonstate_{it} + \boldsymbol{\delta W_{it}} + \varepsilon_{it}, \qquad (4)$$

where $Nonstate_{it}$ is the non-state cluster measurement, including $Nonstate\_V_{it}$ for output and $Nonstate\_E_{it}$ for number of establishments.

The results are summarized in Table 6. Panel A shows that the development of non-state firms in clusters measured by both output and establishment is significantly and positively associated with the rural household per capita income, suggesting that local rural residents' income is positively associated with entrepreneurial clusters. Everything else being equal, a 1% increase in the non-state firms within clusters is associated with 4% increase of the rural household per capita income. By contrast, Panel B indicates that the development of non-state sectors in the clusters is not statistically related to urban household income. This finding suggests a potential mechanism through which entrepreneurial clusters reduce urban–rural income inequality.

Table 6 also shows that the share of the private sector, the total GDP, the investments in fixed assets, and the education expenditure are all significantly and positively correlated to rural household income. These results suggest that rural household income is higher in areas where the non-state sector is more developed and where investments in fixed assets and in human capital are larger. Our regressions also controlled for some competing government policies, including 1) the policy of special economic zones or development zones for attracting FDI and developing export-oriented industries (number of SEZ in regressions); 2) the poverty eradication policy by subsidizing officially identified or recognized national Poor Counties ("Poor" in regressions); and 3) the administrative expenditure in general. Interestingly, government policies on special economic zones and on poverty eradication do not demonstrate any statistical significance on rural household income. However, the expenditures on administration seem to have a negative effect on rural household income but positively affect its urban counterpart. This outcome seems consistent with another finding in this table that "Poor" is positively and significantly correlated to urban household income.[22]

[22] This finding might indicate some abuse of poverty eradication funds. Although highly interesting, a further investigation on this issue is beyond this paper.

In summary, our baseline regressions on the relationship between inequality and clustering indicate that entrepreneurial clusters seem to reduce the income gap between rural and urban residents (Table 5), and that the reduction of inequality is achieved chiefly by increasing the income of rural residents (Table 6).

6. Identification and Robustness Checks

Our baseline regressions demonstrate strong statistical associations between clustering (measured by strength and non-state) and economic growth, and between entrepreneurial clusters and urban–rural inequality. However, the causalities for the relationships are yet to be established because the clustering might be a result rather than a cause of economic growth or rural household income increase. Moreover, the existence of clusters or the features of clustering might coincide with other unobservable variables that might influence local economic growth or the local residents' income. To address identification concerns, we employ two-stage estimations.

6.1 Identifying the effects of clustering on growth

To address the aforementioned identification concerns, we employ the two-staged Least Square estimation procedure to identify clustering effects on growth. In particular, we employ two instrumental variables (IVs) to identify the effects of strength and ownership structure of clusters on growth. The first IV, "P.C.Mining output," is the per capita mining output in each province obtained from the *China Mining Yearbook* (2001–2007). This instrumental variable satisfies the two conditions of exogeneity and relevance.

We believe P.C.Mining output is relevant to strength and the ownership structure of clusters, e.g., *Str_national_V* and *Nonstate_V*, because we expect provinces with higher per capita mining outputs to have weaker industrial clusters and a smaller share of the private sector in the clusters. This anticipation is based on an observation that mine-rich regions are often dominated by large companies due to large fixed investments and high returns to scale in the mining industry. Thus, smaller businesses are often crowd out and entrepreneurship is often depressed as argued by Chinitz (1961)[23] and as supported by empirical evidence (e.g., Rosenthal and Strange, 2003; and Glaeser et al., 2010, 2015). Moreover, given the state-ownership of mining rights in China, mine-rich regions should have higher shares of the state sector, which in general would affect all industries in those regions.

We also believe that the per capita mining output of a province is exogenous because the mine-richness of a region is geologically determined and by itself is unrelated to regional economic growth. Given our panel data is at the county level,

[23] Chinitz (1961) also argues that when a region is dominated by large mining companies, the culture of entrepreneurship is weak because the executives of large companies in regions with large mining companies are less likely to transfer entrepreneurial knowledge to the next generations. Moreover, in such regions, the financial and labor constraints for entrepreneurial firms may be severe, because both financial institutions and labor may easily access large firms with low levels of risks and uncertainty. Furthermore, large companies are more likely to internalize supplies or source them outside the region to enjoy low costs, which consequently depresses the local supply development of small entrepreneurial firms.

whereas this particular IV is at the provincial level. In general, provincial per capita mining output is not directly correlated to county-level economic growth.[24]

The second instrumental variable is the county government's revenue from land sales. The data are collected from the *National Prefecture and County Public Finance Statistical Yearbook*s. We argue that government land revenue is relevant to the strength and ownership structure of the clusters. In China, land is ultimately state-owned, such that rural residents must yield their land ownership to the government for the land to be traded in the market for industrial or other commercial usage (see Section 2 of this paper). Thus, developing industrial clusters and entrepreneurial activities rely heavily on the government's conversion of land from agricultural to industrial use, e.g., selling land. Moreover, land prices are higher in regions where clusters are stronger and entrepreneurial activities are more intensive. Consequently, in counties where the government obtains more revenue from land, clusters are stronger and have more privately owned firms.

We also argue that the local government's revenue from land is not independently related to local economic growth. This is because in localities without active entrepreneurial activities, without clusters, the demand for land and the price of land will be low. The scatter diagram of Figure 3 shows the relationship between land revenue ratio and growth rate at county level data in 2006. Here, land revenue ratio is the ratio of the county government's land sale revenue over its total fiscal revenue. The data is from the *National Prefecture and County Public Finance Statistical Yearbooks*. The figure indicates no clear statistical relationship exists between the land revenue ratio and local economic growth.

Table 7 presents the regression results of the two-stage estimations for economic growth when cluster strength and ownership structure are instrumented. Panel A shows the results of the first-stage estimations, confirming that the two variables are qualified IVs. The findings show that as expected government's land sales revenue is significantly and positively associated with the strength of clusters (the contribution of the clusters to the national economy) and the ownership structure of clusters (the weights of non-state firms within the clusters). Per capita mining output is, also as expected, significantly and negatively associated with the strength of clusters and the ownership structure of clusters. Finally, the Sargan test results confirm that the two IVs are statistically exogenous in the cases of the three variables except for *Str_national_V*.

Panel B of Table 7 reports the second-stage estimation results. Consistent with our baseline OLS regression results, the instrumented strength of the clusters and the non-state ownership clusters remain significantly and positively associated with economic growth. These outcomes confirm that stronger clusters and entrepreneurial clusters drive local economic growth.

6.2 Identifying the effects of clustering on urban–rural inequality

To identify whether the development of the non-state sector within the clusters

[24] Indeed, even in the case of dealing with the relationship between provincial level mining and provincial economic growth, evidence provided in the literature (e.g. Glaeser et al., 2015) still supports our conjecture.

reduces urban–rural inequality, we employ the two-staged Least Square estimation procedure with two instrumental variables, Buddhist temple at the county level and the inflow of migrant labor at the provincial level.

The existence of Buddhist temples in a county is basically historically given. Here, we use this data as a proxy of the impact of Buddhism, or social capital, in the locality. An argument is that Buddhism emphasizes social harmony and compassion and discourages direct confrontation (Pan et al., 2012). These Buddhist principles may facilitate social networks and relationships. On the one hand, it is well documented that contract enforcement in China is very weak (Long and Zhang, 2012). On the other hand, industrial clusters in China typically consist of numerous small private firms from similar or related industries that work closely with each other (Section 2). These small firms have to seek self-enforcing contracts by relying on the trust within networks (Long and Zhang, 2012). Therefore, we expect that places where Buddhist culture prevails tend to foster more cluster development. The variable Buddhist temple is constructed as a dummy ***Buddhism*,** which is equal to one if a county has a Buddhist temple, otherwise, zero.[25] Data are collected from the "Spatial Explorer on Religion Dataset", which covers information on the spatial distribution of major religious sites in China at the township level between 1995 and 2004 in a panel data fashion. Considering that most Buddhist temples do not play roles in re-distributing wealth to the poor, we do not expect this IV, *Buddhism*, to affect the urban–rural inequality directly without going through channels such as clusters.

Our second instrumental variable, *Inflow_Migrant,* measures the percentage of immigrant labor over the total employment in each province. Immigrant data are taken from the *China Population Yearbook* (1999–2007). It is known that industrial clusters employ large numbers of immigrants from other provinces. Therefore, we expect that provinces with large inflow of migrant workers from other provinces have more developed industrial clusters. However, provincial level immigrant measurement should not have a direct relationship with county level inequality. Table 8 reports our two-stage estimations on the clusters' impacts on urban–rural inequality. Dealing with the identification problem raised by our baseline OLS results, the focus here is the non-state clusters. Panel A reports the outcomes of the first-stage estimations. The findings show that both of the IVs are significantly and positively associated with the development of the non-state clusters. That is, these IVs are relevant for our purpose. Moreover, the Sargan test results for over-identifying restrictions suggest that the null hypothesis cannot be rejected, confirming that the two IVs are exogenous to urban–rural inequality.

Panel B of Table 8 shows the outcome of the second-stage estimation. Consistent with our baseline OLS regression results, these findings show that the development of the non-state sector within the clusters remains significantly and negatively correlated with urban–rural inequality after the independent variables are instrumented by the two IVs. The results confirm that clusters with a developed non-state sector reduce urban–rural income inequality.

For identifying the mechanisms through which the non-state sector in clusters

[25] Due to the data constraints, we take a three-year lag and use the religion data from 1995-2004.

reduces urban–rural inequality, we conduct two-stage estimations to identify the effects of clustering on rural per capita income. The IVs are the same as those in Table 8, and all other data are the same as those in Table 6. Panel A of Table 9 presents the first-stage estimation results; similar to the findings in Panel A of Table 8, it confirms the relevance and the exogeneity of the two IVs. Second-stage estimations in Panel B of Table 9 confirm that the development of a non-state sector in the clusters (measured by *Nonstate_V* and *Nonstate_E*) is significantly and positively correlated with rural household per capita income after these independent variables are instrumented. The estimations verify that clusters with more non-state firms lead to higher income of rural residents. To summarize, our instrumented estimations confirm that entrepreneurial clusters reduce urban–rural income inequality (Table 8) by increasing the income of rural residents (Table 9).

7. Conclusion

In this study, we construct an industrial cluster measurement, the DBI, which captures institutionally constrained industrial clusters, particularly entrepreneurial clusters in China. Combining both firm- and county-level data, we create a county-level DBI cluster panel, and find that stronger clusters, particularly entrepreneurial clusters, enhance economic growth. Moreover, entrepreneurial clusters reduce inequality by increasing rural residence income, which is qualitatively and substantially different from SOEs' impacts on income inequality. To our knowledge, this is the first study of this kind in the literature.

We carefully address identification concerns through the two-stage least squares approach. For investigating the effects of clustering on economic growth, we use local per capita mining outputs and the land sales revenue of county governments as the IVs. With respect to urban–rural inequality and the income of rural residents, we use the Buddhist temple at the county level and the inflow of migrant labor at provincial level as the IVs.

This study contributes to the literature on development economics, growth, inequality, and economic geography. Several challenging questions arising from our discoveries require further research. One of our major findings is that entrepreneurial industrial clusters push economic growth and reduce urban–rural inequality effectively. However, such clusters are only concentrated in limited regions, and entrepreneurial clusters are non-existent in many provinces. Why is this so? What barriers prevent the occurrence of entrepreneurial clusters? Evidence on the simultaneous existence of the strong growth-enhancing effect and inequality-reducing effect of entrepreneurial industrial clusters indicates a possibility that some kind of Schumpeterian growth mechanism (e.g. Aghion, 2002) is at work. Determining the mechanism and addressing why particular mechanisms are more prevalent than others in certain areas require much more research. These future research directions promise further contribution to the literature on economic development and institutions (e.g. Acemoglu et al., 2002, 2005; Engermann and Sokoloff, 1997, 2000; Easterly, 2007), and knowledge on growth and inequality, economic geography, urban economics, and the economic development of China.

References

Acemoglu, Daron, Simon Johnson, and James A. Robinson, 2002. "Reversal of fortune: Geography and institutions in the making of the modem world income distribution." *Quarterly Journal of Economics* 118: 1231-1294.

Acemoglu, Daron, Simon Johnson, and James A. Robinson, 2005. "Institutions as a fundamental cause of long-run growth." *Handbook of Economic Growth* 1: 385-472.

Acemoglu, Daron, and James A. Robinson. "The rise and decline of general laws of capitalism." *The Journal of Economic Perspectives* 29.1 (2015): 3-28.

Allen, Franklin, Jun Qian, and Meijun Qian. 2005. "Law, Finance, and Economic Growth in China." *Journal of Financial Economics* 77: 57-116.

Aghion, Philippe, Eve Caroli, and Cecilia Garcia-Penalosa, 1999. "Inequality and Economic Growth: The Perspective of the New Growth Theories." *Journal of Economic Literature* 37: 1615-1660

Aghion, Philippe, and Peter Howitt, 1992. "A Model of Growth through Creative Destruction." *Econometrica* 60(2): 323-351.

Aghion, Philippe, Peter Howitt, and Giovanni L. Violante, 2002. "General purpose technology and wage inequality." *Journal of Economic Growth 7*(4): 315-345.

Alesina, Alberto, and Dani Rodrik, 1994. "Distributive Politics and Economic Growth." *The Quarterly Journal of Economics* 109(2): 465-490.

Arrow, Kenneth J., 1962. "The Economic Implications of Learning by Doing." *The Review of Economic Studies* 29(3): 155-173.

Au, Chun-Chung, and Vernon Henderson, 2006a. "How migration restrictions limit agglomeration and productivity in China." *Journal of Development Economics* 80(2):350-388.

Au, Chun-Chung, and Vernon Henderson, 2006a. "Are Chinese cities too small?". *The Review of Economic Studies*, *73*(3), pp.549-576.

Audretsch, David, and Maryann Feldman, 1996. "R&D Spillovers and the Geography of Innovation and Production." *American Economic Review* 86(3): 630-640.

Banerjee, Abhijit, and Esther Duflo, 2003. "Inequality and Growth: What Can the Data Say?" *Journal of Economic Growth* 8: 267-299.

Barro, Robert, and Xavier Sala-i-Martin, 1992. "Convergence." *Journal of Political Economy* 100(2): 223-251.

Barro, Robert J, 2000. "Inequality and Growth in a Panel of Countries." *Journal of Economic Growth* 5(1): 5-32.

Baumol, William, 1986. "Productivity Growth, Convergence, and Welfare: What the Long-run Data Show." *American Economic Review* 76(5): 1072-1085.

Benjamin, Dwayne, Loren Brandt, and John Giles, 2005. "The Evolution of Income Inequality in Rural China." *Economic Development and Cultural Change* 53(4): 769-824.

Benjamin, Dwayne, Loren Brandt, John Giles, and Sangui Wang, 2008. "Income Inequality During China's Economic Transition." In *China's Great Economic Transformation*, edited by Loren Brandt and Thomas Rawski. Cambridge: Cambridge University Press.

Benjamin, D., Brandt, L, & Giles, J. (2011). "Did higher inequality impede growth in rural China?" The *Economic Journal, 121*, 1281-1309.

Bosker, Maarten, Steven Brakman, Harry Garretsen, and Marc Schramm, 2012. "Relaxing Hukou: Increased labor mobility and China's economic geography."*Journal of Urban Economics* 72(2): 252-266.

Brülhart, Marius, and Federica Sbergami, 2009. "Agglomeration and growth: Cross-country evidence." *Journal of Urban Economics* 65(1): 48-63.

Byrd, William, and Qinsong Lin, 1990. *China's Rural Industry: Structure, Development, and Reform*. New York: Oxford University Press.

Cai, Fang, 2000. *Zhongguo Liudong Renkou Wenti (China's Floating Population)*. Zhengzhou: Henan Renmin Chubanshe (Henen People's Publishing House).

Candelaria, C, M Daly, G Hale (2013), "Persistence of Regional Inequality in China," FEDERAL RESERVE BANK OF SAN FRANCISCO.

Chang, Chun, and Yijiang Wang, 1994. "The nature of the township-village enterprise." *Journal of Comparative Economics* 19(3): 434–452.

Chatterji, Aaron, Edward Glaeser and William Kerr, 2013. "Clusters of Entrepreneurship and Innovation." NBER Chapters, in: *Innovation Policy and the Economy* 14: 129-166.

Che, Jiahua, and Yingyi Qian, 1998a. "Insecure Property Rights and Government Ownership of Firms." *Quarterly Journal of Economics,* 113: 467-496. 1998.

Che, Jiahua, and Yingyi Qian, 1998b. "Institutional environment, community government, and corporate governance: Understanding China's township-village enterprises." *Journal of Law, Economics, and Organization* 14(1): 1–23.

Chen, Jian, and Belton Fleisher, 1996. "Regional Income Inequality and Economic Growth in China." *Journal of Comparative Economics* 22(2): 141-164.

Chinitz, Benjamin. 1961. "Contrasts in agglomeration: New york and pittsburgh." *The American Economic Review* (1961): 279-289.

Ciccone, Antonio, and Robert Hall, 1996. "Productivity and the Density of Economic Activity." *American Economic Review* 86(1): 54-70.

Combes, Pierre-Philippe, and Giles Duranton, 2006. "Labour pooling, labour poaching, and spatial clustering." *Regional Science and Urban Economics* 36(1): 1-28.

Delgado, Mercedes., Michael Porter, and Scott Stern, 2010. "Cluster and entrepreneurship." *Journal of Economic Geography*: 1-24.

Delgado, Mercedes; Michael Porter, and Scott Stern, 2014. "Cluster, convergence, and economic performance." *Research Policy* 43: 1785-1799.

Deng, Xiaoping, 1986 [1994]. "Reform is the only way for China to develop its productive forces," a speech on August 28, 1985, in *Selected Works of Deng Xiaoping,* Volume III. Beijing: Foreign Languages Press. 1994.

Easterly, William, 2007. "Inequality does cause underdevelopment: Insights from a new instrument." *Journal of Development Economics* 84 (2): 755-776.

Ellison, Glenn, and Edward Glaeser, 1997. "Geographic Concentration in U.S. Manufacturing Industries: A Dartboard Approach." *Journal of Political Economy* 105(5): 889-892.

Engerman, Stanley, and Kenneth Sokoloff, 1997. "Factor endowments, institutions, and differential paths of growth among new world economies." *How Latin America Fell Behind*: 260-304.

Engerman, Stanley, and Kenneth Sokoloff, 2000. "History lessons: Institutions, factors endowments, and paths of development in the new world." *The Journal of Economic Perspectives*: 217-232.

Feser, Edward, and Edward Bergman, 2000. "National Industry Cluster Template: A Framework for Regional Cluster Analysis." *Regional Studies* 34(1): 1-20.

Fleisher, Belton, Dinghuan Hu, William McGuire and Xiaobo Zhang, 2010. "The evolution of an industrial cluster in China." *China Economic Review* 4: 1-14.

Forbes, Kristin, 2000. "A Reassessment of the Relationship between Inequality and Growth." *American economic review* 90(4): 869-887.

Fujita, Masahisa, Paul Krugman, and Anthony Venables, 1999. *The Spatial Economy*. Cambridge MA: MIT Press.

Fujita, Masahisa, and Dapeng Hu, 2001. "Regional disparity in China 1985–1994: the effects of globalization and economic liberalization." *The Annals of Regional Science* 35.1: 3-37.

Gan, Li, 2013, "Income Inequality and Consumption in China," mimeo, Texas A&M University.

Glaeser, Edward, Hedi Kallal, Jose Scheinkman and Andrei Shleifer, 1992. "Growth in cities." *Journal of Political Economy* 100(6): 1126-1152.

Glaeser, Edward, and Joshua Gottlieb, 2009. "The Wealth of Cities: Agglomeration Economies and Spatial Equilibrium in the United States." *Journal of Economic Literature* 47(4): 983-1028.

Glaeser, Edward L., Stuart S. Rosenthal, and William C. Strange. 2010 "Urban economics and entrepreneurship." *Journal of Urban Economics* 67 (1): 1-14.

Glaeser, Edward., Sari Pekkala Kerr, and William Kerr. 2015. "Entrepreneurship and urban growth: An empirical assessment with historical mines." *Review of Economics and Statistics* 97(2): 498-520.

Guo, Di, Kun Jiang, Byung-Yeon Kim, and Chenggang Xu, 2014. “Political economy of private firms in China.” *Journal of Comparative Economics* 42: 286-303.

Hausmann, Ricardo, and Bailey Klinger, 2007. “The structure of the product space and the evolution of comparative advantage.” Working Paper, Harvard University.

Henderson, Vernon, 1974. "Optimum City Size: The External Diseconomy Question." *Journal of Political Economy* 82(2): 373-388.

Henderson, Vernon, 2003. "Urbanization and Economic Development." *Annals of Economics and Finance* 4(2): 275-341.

Hessler, Peter, 2007. “Boomtowns.” *National Geographic*, June: 88-117.

Huang, Zuhui, Xiaobo Zhang, and Yunwei Zhu, 2008. “The Role of Clustering in Rural Industrialization: A Case Study of Wenzhou’s Footwear Industry.” *China Economic Review* 19: 409-420.

Jacobs, Jane, 1969. *The Economy of Cities*. Random House: New York

Kanbur, Ravi, and Xiaobo Zhang, 1999. "Which Regional Inequality? The Evolution of Urban-rural and Inland-Coastal Inequality in China from 1983 to 1995." *Journal of Comparative Economics* 27(4): 686-701.

Kanbur, Ravi, and Xiaobo Zhang, 2005. "Fifty Years of Regional Inequality in China: A Journey through Central Planning, Reform, and Openness." *Review of Development Economics*, 9(1): 87-106.

Kerr, William, 2010. "Breakthrough inventions and migrating clusters of innovation." *Journal of Urban Economics* 67(1): 46-60.

Knowles, Stephen, 2001. “Inequality and Economic Growth: The empirical Relationship Reconsidered in the Light of Comparable Data.” *Journal of Development Studies* 41(1): 135-159.

Krugman, Paul, 1991a. *Geography and Trade*. Cambridge MA: MIT Press.

Krugman, Paul, 1991b. “Increasing returns and economic geography.” *Journal of Political Economy* 99(3): 484-499.

Krugman, Paul, 1993. “Lessons of Massachusetts for EMU.” In *Adjustment and Growth in the European Monetary Union* edited by Francisco Torres and Francesco Giavazzi. Cambridge: Cambridge University Press.

Krugman, P. (1998), “What’s new about the new economic geography?”, Oxford

Review of Economic Policy, 14, p. 7-17.

Krugman, Paul, and Anthony Venables, 1995. "Globalization and the Inequality of Nations." *The Quarterly Journal of Economics* 110(4): 857-880.

Kuznets, Simon, 1955. "Economic Growth and Income Inequality." *American Economic Review* 45(1): 1–28.

Kuznets, Simon, 1963. "Quantitative aspects of the economic growth of nations, VIII: The distribution of income by size." *Economic Development and Cultural Change* 11(2): 1–92.

Li, David D, 1996. "A theory of ambiguous property rights in transition economies: The case of the Chinese non-state sector." *Journal of Comparative Economics* 23(1): 1–19.

Li, Ming, Ran Tao, and Dali Yang, 2013. " Decentralization, Inequality and Poverty Relief in China." University of Chicago.

Li, Shi, Hiroshi Sato, and Terry Sicular (ed.), 2013. *Rising Inequality in China*. Cambridge: Cambridge University Press.

Lin, Justin Yifu, Fang Cai, and Zhou Li, 1998. "Competition, Policy Burdens, and State-Owned Enterprise Reform." *American Economic Review* 88(2): 442-427.

Lin, Justin Yifu, and Yongjun Li, 2001. "Promoting the Growth of Medium and Small-sized Enterprises through the Development of Medium and Small-sized Financial Institutions." *Economic Research Journal* 1: 10-18.

Long, Cheryl, and Xiaobo Zhang, 2012. "Patterns of China's industrialization: Concentration, specialization, and clustering." *China Economic Review* 23(3): 593-612.

Long, Cheryl, and Xiaobo Zhang, 2011. "Cluster-based industrialization in China: Financing and performance."J*ournal of International Economics* 84(1): 112-123.

Lu, Jiangyong, and Zhigan Tao, 2009. "Trends and determinants of China's industrial agglomeration." *Journal of Urban Economics*, *65*(2): 167-180.

Marshall, Alfred, 1890. *Principles of Economics: An Introductory Volume.* London: Macmillan.

Meng, Lingsheng, 2013. "Evaluating China's poverty alleviation program: A regression discontinuity approach." *Journal of Public Economics* 101: 1-11.

Meng, Xin., 2000. *Labour market reform in China*. Cambridge University Press.

Meng, Xin. 2012. "Labor market outcomes and reforms in China." *The Journal of Economic Perspectives* 26(4): 75-101.

Mukherjee, Anit, and Xiaobo Zhang, 2007. "Rural industrialization in China and India: Role of policies and institutions." *World Development* 35 (10): 1621–1634.

Nee, Victor, and Sonja Opper, 2012. "*Capitalism from below: Markets and institutional* *change in China*." Harvard University Press.

Oi, Jean. C., 1999. "*Rural China takes off: Institutional foundations of economic reform"*. University of California Press.

Pan, Yaotian, Julie A. Rowney, and Mark F. Peterson. 2012. "The structure of Chinese cultural traditions: An empirical study of business employees in China." *Management and Organization Review* 8(1): 77-95.

Perotti, Roberto, 1996. "Growth, Income Distribution, and Democracy: What the Data Say." *Journal of Economic Growth* 1(2): 149-187.

Persson, Torsten, and Guido Tabellini, 1994. "Is Inequality Harmful for Growth?" *American Economic Review* 84(3): 600-621.

Peterson, Mitchell A, 2009. "Estimating standard error in finance panel data sets: Compering approaches." *Review of Financial Studies* 22: 435-480.

Piore, Michael, and Charles Sabel, 1984. *The second industrial divide: possibilities for prosperity.* New York: Basic Books.

Porter, Michael, 1990. *The Competitive Advantage of Nations*. London: Macmillan.

Porter, Michael, 1998. "Clusters and the new economics of competition." *Harvard Business Review* 76 (6): 77–90.

Porter, Michael, 2003 "The Economic Performance of Regions." *Regional Studies* 37(6-7): 549-578.

Qian, Yingyi, and Chenggang Xu, 1993. "Why China's Economic Reforms Differ: The M-form Hierarchy and Entry/Expansion of the Non-State Sector." *The Economics of Transition* 1(2): 135-170.

Romer, Paul M, 1986. "Increasing Returns and Long-run Growth." *Journal of Political Economy* 94(5): 1002-1037.

Rotemberg, Julio, and Garth Saloner, 2000. "Competition and Human Capital Accumulation: A Theory of Interregional Specialization and Trade." *Regional Science and Urban Economics* 30(4): 373-404.

Rosenthal, Stuart S., and William C. Strange. 2003. "Geography, industrial organization, and agglomeration." *Review of Economics and Statistics* 85 (2): 377-393.

Ruan, Jianqing, and Xiaobo Zhang, 2009. "Finance and Cluster-Based Industrial Development in China." *Economic Development and Cultural Change* 58(1): 143-164.

Saxenian, AnnaLee, 1994. *Regional Advantage: Culture and Competition in Silicon Valley and Route 128*. Cambridge, MA: Harvard University Press.

Shleifer, Andrei, and Robert Vishny, 1994. "Politicians and Firms." *The Quarterly Journal of Economics* 109(4): 995-1025.

Sicular, Terry, 2013, "The Challenge of High Inequality in China," Poverty Reduction and Equity Department, The World Bank.

Solow, Robert M., 1956. "A contribution to the theory of economic growth." *The Quarterly Journal of Economics* 70(1): 65-94.

Sonobe, Tetsushi, Dinghuan Hu, and Keijiro Otsuka, 2002. "Process of cluster formation in China: A case study of a garment town." J*ournal of Development Studies* 39(1): 118-139.

Sonobe, Tetsushi, and Keijiro Otsuka, 2006a. "The Division of Labor and the Formation of Industrial Clusters in Taiwan." *Review of Development Economics* 10(1): 71–86.

Sonobe, Tetsushi, and Keijiro Otsuka, 2006b. *Cluster-based industrial development: An East Asia model*. New York: Palgrave MacMillan.

Sonobe, Tetsushi, and Keijiro Otsuka., 2014. "*Cluster-Based Industrial Development:: KAIZEN Management for MSE Growth in Developing Countries*." Springer.

Wang, Feiling, 2005. *Organizing through Division and Exclusion*. Stanford: Stanford University Press.

Weitzman, Martin, and Chenggang Xu, 1994. "Chinese township village enterprises as vaguely defined cooperatives." *Journal of Comparative Economics* 18(2): 121-145.

Whyte, Martin King (ed.), 2010, *One Country, Two Societies: Rural-Urban Inequality in Contemporary China*, Harvard University Press. 2010.

Wong, Bernard, 1978. "A comparative study of the assimilation of the Chinese in New York City and Lima, Peru." *Comparative Studies in Society and History* 20 (3): 335-358.

World Bank, 2013. "China: Poverty Alleviation through Community Participation." World Bank Official Website. (http://www.worldbank.org/en/results/2013/04/09/china-poverty-alleviation-through-community-participation.)

Xie, Yu, and Xiang Zhou. "Income inequality in today's China". Proceedings of the National Academy of Sciences of the United States of America 111.19 (2014): , 111, 19, 6928-2933.

Xu, Chenggang, 2011. "The Fundamental Institutions of China's Reform and Development." *Journal of Economic Literatures* 49(4): 1076-1151.

Xu, Chenggang, and Xiaobo Zhang, 2009. "The Evolution of Chinese Entrepreneurial Firms: Township-Village Enterprises Revisited." In *China's Economic Transformation* edited by Ronald Coase, forthcoming.

Yang, Dennis Tao, 1999. "Urban-biased policies and rising income inequality in China." *American Economic Review Papers and Proceedings*, 89(2): 306-310.

Table 1: Summary Statistics of DBI Clusters

Clusters over time

Year	Counties	County-industry observations	Counties with Clusters	Industrial clusters	Clusters' share in total output	Clusters' Share in total employment
1998	2700	41899	685	2024	0.3838	0.2797
1999	2747	41571	654	2037	0.3527	0.2546
2000	2746	40272	657	1958	0.3733	0.2821
2001	2673	38712	632	1896	0.3633	0.2760
2002	2726	39432	668	1931	0.3680	0.2905
2003	2787	40207	709	1971	0.3811	0.3124
2004	2790	41996	696	2118	0.4035	0.3650
2005	2790	41809	750	2125	0.3914	0.3636
2006	2793	43177	724	2157	0.3911	0.3748
2007	2734	44175	739	2213	0.3802	0.3668

Dynamics of Counties and Clusters

Counties always with clusters, 98-07	Counties always without clusters, 98-07	Counties without clusters in 1998, but with clusters in 2007	Counties with clusters in 1998, but without in 2007
294	1,576	317	292

Features of Clusters

	Str_national_V	Str_national_E	Nonstate_V	Nonstate_E
Mean	6.5391	5.9087	0.8016	0.8331
Median	3.9976	4.7278	0.9523	0.9444
Standard deviation	8.1344	3.6840	0.2922	0.2511
Min	0.1041	2.3384	0	0
Max	50.9959	23.4122	1	1
N	6914	6914	6914	6914

Table 2: Clustering Measurements at Provincial Level: DBI vs. Standard Approaches

Province	HHI	Gini	LQ	Krugman Index	Industrial Proximity	DBI Cluster#	DBI Strength_V	DBI_Strength_E	DBI_Nonstate_V	DBI_Nonstate_E
Beijing	0.0913	0.6668	2.5211	0.6478	0.22	64	3.8909	3.3761	0.5189	0.5199
Tianjin	0.0808	0.6372	1.7035	0.5681	0.208	57	6.4096	3.9931	0.5569	0.5491
Hebei	0.1205	0.6325	1.9691	0.5956	0.219	44	1.4906	1.3629	0.2108	0.2140
Shanxi	0.1605	0.8014	4.8040	1.0693	0.208	25	0.6417	1.1312	0.1842	0.1859
Inr Mongolia	0.0802	0.6689	2.6402	0.8504	0.214	11	0.6925	0.4150	0.0663	0.0807
Liaoning	0.0656	0.5983	1.3517	0.4519	0.205	145	2.4913	2.8489	0.4274	0.4351
Jilin	0.1527	0.7231	2.6820	0.8912	0.22	24	1.7743	1.5403	0.2677	0.3055
Heilongjiang	0.1217	0.7201	4.7328	0.9622	0.197	8	0.3400	0.2626	0.0473	0.0585
Shanghai	0.0871	0.5938	1.4276	0.4596	0.219	165	9.2155	5.3317	0.5761	0.6106
Jiangsu	0.0684	0.6208	1.2523	0.4317	0.21	284	5.5121	4.8527	0.6344	0.6375
Zhejiang	0.0530	0.5138	1.4781	0.4851	0.22	414	6.2603	7.5372	0.7646	0.7770
Anhui	0.0584	0.5520	1.2671	0.3973	0.211	18	0.5599	0.6983	0.1411	0.1394
Fujian	0.0503	0.4798	1.5538	0.4789	0.202	112	2.9643	3.8209	0.5252	0.5500
Jiangxi	0.0779	0.5705	1.8355	0.4831	0.206	19	0.8680	0.7392	0.1305	0.1483
Shandong	0.0506	0.5183	1.2188	0.3872	0.205	233	5.9145	4.6156	0.6733	0.6896
Henan	0.0560	0.5513	1.5212	0.5771	0.209	51	1.9954	1.3594	0.2078	0.2194
Hubei	0.0706	0.6004	1.3529	0.4179	0.216	27	1.0604	0.9416	0.1654	0.1818
Hunan	0.0556	0.5426	1.4514	0.4850	0.21	47	0.7942	1.5509	0.2114	0.2203
Guangdong	0.0839	0.5813	1.5243	0.5668	0.215	260	4.0450	3.3413	0.4271	0.4533
Guangxi	0.0791	0.6458	1.6292	0.6692	0.214	11	0.3128	0.4710	0.0824	0.0942
Hainan	0.1555	0.7286	3.6836	0.9652	0.207	2	0.2016	0.4371	0.0870	0.0870
Chongqing	0.1581	0.6954	2.7831	0.7327	0.197	17	2.2100	2.6696	0.3624	0.3684
Sichuan	0.0517	0.5339	1.3152	0.3757	0.202	52	0.5788	1.0106	0.1912	0.1978

Table 2: Clustering Measurements at Provincial Level: DBI vs. Standard Approaches (continued)

Province	HHI	Gini	LQ	Krugman Index	Industrial Proximity	DBI Cluster#	DBI Strength_V	DBI_Strength_E	DBI_Nonstate_V	DBI_Nonstate_E
Guizhou	0.1079	0.7017	2.3672	0.8837	0.196	6	0.4276	0.2974	0.0575	0.0583
Yunnan	0.1209	0.7382	5.0078	0.9727	0.197	12	0.3499	0.3980	0.0831	0.0897
Tibet	0.1330	0.6293	11.3831	1.0326	0.238	0	0.0000	0.0000	0.0000	0.0000
Shaanxi	0.0762	0.6401	2.5095	0.7317	0.192	18	0.4552	0.6820	0.0913	0.1064
Gansu	0.1394	0.7531	3.0550	0.9771	0.205	10	0.4165	0.3509	0.0671	0.0556
Qinghai	0.1296	0.7588	5.1840	1.0117	0.217	0	0.0000	0.0000	0.0000	0.0000
Ningxia	0.1173	0.7146	7.3161	0.8497	0.22	3	0.1385	0.4480	0.0769	0.0884
Xinjiang	0.1799	0.7727	7.1079	1.0229	0.199	5	0.1098	0.1739	0.0300	0.0350

Note: The HHI, Gini, LQ, Krugman-Index, and DBI measures are based on the 2007 ASIFP data. The Industrial Proximity is from Long and Zhang (2012). DBI Cluster # is the summation of county-level clusters; DBI Strength_V, DBI Strength_E, DBI Nonstate_V, and DBI Nonstate_E are the county averages of the corresponding indices that we use in the regressions of the paper.

Table 3(a): Summary Statistics of County Growth, Income, and Inequality

Variables	P.C. GDP growth	Urban household P.C. income	Rural household P.C. income	Urban–rural income ratio
Counties with Clusters				
Mean	0.135	9855.471	3999.749	2.244
Median	0.124	8803.	3672.771	2.127
Standard deviation	0.150	4098.133	1874.159	0.583
Min	−0.273	1585.087	643.471	1.097
Max	0.9137	31597	14845	6.910
N	4810	1353	4841	1274
Counties without Clusters				
Mean	0.122	6697.413	2628.195	2.741
Median	0.109	6297.533	2357.346	2.569
Standard deviation	0.143	2503.023	1647.037	0.933
Min	−0.273	388	435.097	1.097
Max	0.9137	22384.460	14845.000	6.910
N	13695	2625	13643	2446
Mean Difference	0.013***	3158.057***	1370.554***	-0.497***

Note: * = p < 0.1; ** = p < 0.05; *** = p < 0.01

Table 3(b): Summary Statistics of Other County Characteristics

Variables	P.C. GDP (thousand yuan)	GDP (billion yuan)	Fraction of industrial output	Fraction of non-state firms
Counties with clusters				
Mean	14.270	8.863	1.097	0.833
Median	10.084	6.112	0.845	0.904
Standard deviation	13.205	8.876	0.930	0.184
Min	0.133	0.026	0.016	0
Max	85.599	48.526	5.000	1
N	5544	5626	5626	6914
Counties without clusters				
Mean	6.923	2.788	0.523	0.640
Median	4.898	1.765	0.3350	0.708
Standard deviation	7.461	3.607	0.6577	0.287
Min	0.133	0.026	0.009	0
Max	85.599	48.526	5.000	1
N	16194	16367	16367	20572
Mean Difference	0.735***	60.743***	0.574***	0.194***

Variables	Fraction of micro firms	Fraction of education expenditure	Fraction of f. a. investment	Fraction of gov expenditure
Counties with clusters				
Mean	0.056	0.019	0.363	0.011
Median	0.033	0.015	0.312	0.008
Standard deviation	0.075	0.014	0.246	0.012
Min	0	0.004	0.007	0.003
Max	0.909	0.136	1.756	0.154
N	6914	5192	4528	5194
Counties without clusters				
Mean	0.079	0.034	0.376	0.025
Median	0.029	0.026	0.311	0.016
Standard deviation	0.142	0.026	0.275	0.026
Min	0	0.004	0.007	0.003
Max	1	0.136	1.756	0.154
N	20572	14920	13184	14918
Mean Difference	-0.022***	-0.016***	-0.013***	-0.014***

Note: * = $p < 0.1$; ** = $p < 0.05$; *** = $p < 0.01$

Table 4: Clustering and Economic Growth

	Economic growth				
	(1)	(2)	(3)	(4)	(5)
Cluster	0.006☆				
	(1.454)				
Str_national_V		0.010***			
		(3.034)			
Str_national_E			0.005*		
			(1.877)		
Nonstate_V				0.015**	
				(2.154)	
Nonstate_E					0.014**
					(2.060)
p.c.GDP	-0.137***	-0.139***	-0.138***	-0.137***	-0.137***
	(-12.725)	(-12.776)	(-12.727)	(-12.751)	(-12.749)
p.c.GDP square	-0.012***	-0.012***	-0.012***	-0.012***	-0.012***
	(-5.694)	(-5.837)	(-5.719)	(-5.712)	(-5.713)
CPI	-0.213*	-0.207*	-0.209*	-0.214*	-0.213*
	(-1.805)	(-1.746)	(-1.771)	(-1.806)	(-1.801)
Fraction of industrial output	0.030***	0.029***	0.030***	0.030***	0.030***
	(9.706)	(9.340)	(9.595)	(9.661)	(9.657)
Fraction of non-state firms	0.022*	0.023*	0.023*	0.022*	0.022*
	(1.873)	(1.918)	(1.874)	(1.817)	(1.793)
Fraction of micro firms	0.023	0.023	0.022	0.023	0.022
	(1.536)	(1.547)	(1.475)	(1.515)	(1.496)
Fraction of edu expenditure	0.021***	0.021***	0.021***	0.021***	0.021***
	(2.926)	(2.922)	(2.930)	(2.952)	(2.942)
Fraction of f. a. investment	0.023***	0.023***	0.023***	0.023***	0.023***
	(9.806)	(9.884)	(9.811)	(9.811)	(9.823)
Fraction of gov expenditure	0.018***	0.018***	0.018***	0.018***	0.018***
	(3.248)	(3.269)	(3.246)	(3.233)	(3.239)
number of SEZ	-0.013	-0.013	-0.013	-0.013	-0.013
	(-1.026)	(-1.013)	(-1.029)	(-1.018)	(-1.015)
Poor	-0.137***	-0.139***	-0.138***	-0.137***	-0.137***
	(-12.725)	(-12.776)	(-12.727)	(-12.751)	(-12.749)
Constant	0.469***	0.466***	0.468***	0.469***	0.469***
	(17.317)	(17.179)	(17.282)	(17.304)	(17.309)
Time fixed effects	Yes	Yes	Yes	Yes	Yes
County fixed effects	Yes	Yes	Yes	Yes	Yes
Clustered standard error at county-level	Yes	Yes	Yes	Yes	Yes
N	15875	15875	15875	15875	15875
R-sqaure	0.372	0.372	0.372	0.372	0.372

Note: Values in parentheses are t- statistics; * = p < 0.1; ** = p < 0.05; *** = p < 0.01; ☆ = p < 0.15

Table 5: Clustering and Urban-Rural Household Income Inequality

	Urban-Rural Inequality				
	(1)	(2)	(3)	(4)	(5)
Cluster	-0.011				
	(-1.287)				
Str_national_V		-0.001			
		(-0.104)			
Str_national_E			-0.007		
			(-1.421)		
Nonstate_V				-0.033**	
				(-2.149)	
Nonstate_E					-0.032**
					(-2.167)
p.c.GDP	0.043	0.044	0.044	0.043	0.043
	(1.111)	(1.156)	(1.130)	(1.132)	(1.129)
GDP	0.035	0.033	0.035	0.035	0.035
	(0.832)	(0.800)	(0.824)	(0.831)	(0.837)
Fraction of industrial output	0.014	0.013	0.015	0.015	0.015
	(1.493)	(1.376)	(1.537)	(1.600)	(1.602)
Fraction of non-state firms	0.082	0.081	0.082	0.084	0.085
	(1.559)	(1.545)	(1.552)	(1.590)	(1.596)
Fraction of micro firms	-0.062	-0.066	-0.060	-0.061	-0.059
	(-1.125)	(-1.199)	(-1.071)	(-1.092)	(-1.059)
Fraction of edu expenditure	0.042**	0.042**	0.041**	0.041**	0.041**
	(2.194)	(2.195)	(2.190)	(2.171)	(2.184)
Fraction of f. a. investment	0.024***	0.024***	0.024***	0.024***	0.024***
	(2.866)	(2.850)	(2.860)	(2.874)	(2.864)
Fraction of gov expenditure	-0.021	-0.021	-0.021	-0.020	-0.020
	(-1.365)	(-1.378)	(-1.357)	(-1.330)	(-1.338)
number of SEZ	-0.009	-0.010	-0.009	-0.009	-0.009
	(-0.402)	(-0.423)	(-0.387)	(-0.381)	(-0.391)
Poor	0.043	0.044	0.044	0.043	0.043
	(1.111)	(1.156)	(1.130)	(1.132)	(1.129)
Constant	0.520	0.543	0.524	0.523	0.520
	(0.933)	(0.993)	(0.933)	(0.936)	(0.931)
Time fixed effects	Yes	Yes	Yes	Yes	Yes
County fixed effects	Yes	Yes	Yes	Yes	Yes
Clustered standard error at county level	Yes	Yes	Yes	Yes	Yes
N	3363	3363	3363	3363	3363
R-square	0.246	0.246	0.246	0.247	0.247

Note: Values in parentheses are t- statistics; * = p < 0.1; ** = p < 0.05; *** = p < 0.01.

Table 6: Entrepreneurial Clusters and Household Per Capita Income

	Panel A: Rural household income		Panel B: Urban household income	
	(1)	(2)	(3)	(4)
Nonstate_V	0.043***		-0.004	
	(4.306)		(-0.243)	
Nonstate_E		0.042***		-0.010
		(4.012)		(-0.648)
p.c.GDP	0.008	0.008	0.158***	0.158***
	(0.976)	(0.964)	(2.894)	(2.880)
GDP	0.122***	0.122***	-0.060	-0.059
	(4.989)	(4.983)	(-1.301)	(-1.286)
Fraction of industrial output	0.008	0.008	0.027***	0.028***
	(1.132)	(1.127)	(3.036)	(3.083)
Fraction of non-state firms	0.081**	0.080**	0.063	0.064
	(2.438)	(2.410)	(1.633)	(1.645)
Fraction of micro firms	-0.066*	-0.067*	-0.065	-0.064
	(-1.902)	(-1.934)	(-1.228)	(-1.195)
Fraction of edu expenditure	0.033***	0.033***	-0.026	-0.026
	(2.726)	(2.710)	(-1.147)	(-1.149)
Fraction of f. a. investment	0.020***	0.020***	0.041***	0.041***
	(4.860)	(4.863)	(6.739)	(6.746)
Fraction of gov expenditure	-0.042***	-0.042***	0.025*	0.025*
	(-3.140)	(-3.133)	(1.795)	(1.807)
number of SEZ	0.013	0.013	-0.009	-0.009
	(0.239)	(0.243)	(-0.593)	(-0.585)
Poor	0.008	0.008	0.158***	0.158***
	(0.976)	(0.964)	(2.894)	(2.880)
Constant	6.567***	6.567***	9.937***	9.931***
	(21.843)	(21.818)	(16.430)	(16.379)
Time fixed effects	Yes	Yes	Yes	Yes
County fixed effects	Yes	Yes	Yes	Yes
Clustered standard error at county-level	Yes	Yes	Yes	Yes
N	13646	13646	3513	3513
R-sqaure	0.455	0.455	0.798	0.798

Note: Values in parentheses are t- statistics; * = p < 0.1; ** = p < 0.05; *** = p < 0.01.

Table 7: Two-stage estimations on clustering and economic growth

	Panel A: First-stage estimations				Panel B: Second-stage estimations			
	(1)	(2)	(3)	(4)	(1)	(2)	(3)	(4)
	Str_national_V	Str_national_E	Nonstate_V	Nonstate_E	Economic growth	Economic growth	Economic growth	Economic growth
Land Revenue	0.128***	0.099**	0.044***	0.037**				
	(2.520)	(2.140)	(2.500)	(2.129)				
P.C. Mining outputs	-0.003*	-0.004**	-0.002***	-0.002***				
	(-1.807)	(-2.021)	(-2.667)	(-2.890)				
Str_national_V					0.470***			
					(2.636)			
Str_national_E						0.606***		
						(2.747)		
Nonstate_V							1.325***	
							(3.319)	
Nonstate_E								1.470***
								(3.275)
Time Fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
County Fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
N	9283	9283	9283	9283	9283	9283	9283	9283
P-Values of F-tests	0.015	0.012	0.001	0.002	0.000	0.00	0.000	0.000
Underidentification test p-value	0.016	0.012	0.002	0.002				
Sargan's test p-value	0.005	0.169	0.182	0.363				

Note: For convenience, we do not present all the control variables in the table. The control variables for both 1st and 2nd stage estimations are the same as those shown in Table 4. The values in parentheses are t- statistics; * = p < 0.1; ** = p < 0.05; *** = p < 0.01.

Table 8: Two-stage Estimations on Entrepreneurial Clusters and Urban-Rural Income Inequality

	Panel A: 1st stage estimations		Panel B: 2nd stage estimations	
	(1) Nonstate_V	(2) Nonstate_E	(1) Urban-rural Inequality	(2) Urban-rural Inequality
Buddhism	0.077*	0.088**		
	(1.827)	(2.079)		
Inflow_Migrant	0.011**	0.012**		
	(1.993)	(1.992)		
Nonstate_V			−0.564*	
			(−1.791)	
Nonstate_E				−0.511*
				(−1.890)
Time Fixed Effects	Yes	Yes	Yes	Yes
County Fixed Effects	Yes	Yes	Yes	Yes
N	3049	3049	3049	3049
P-value of F-test	0.046	0.030	0.000	0.000
Underidentification test p-value	0.105	0.088		
Sargan's test p-value	0.520	0.436		

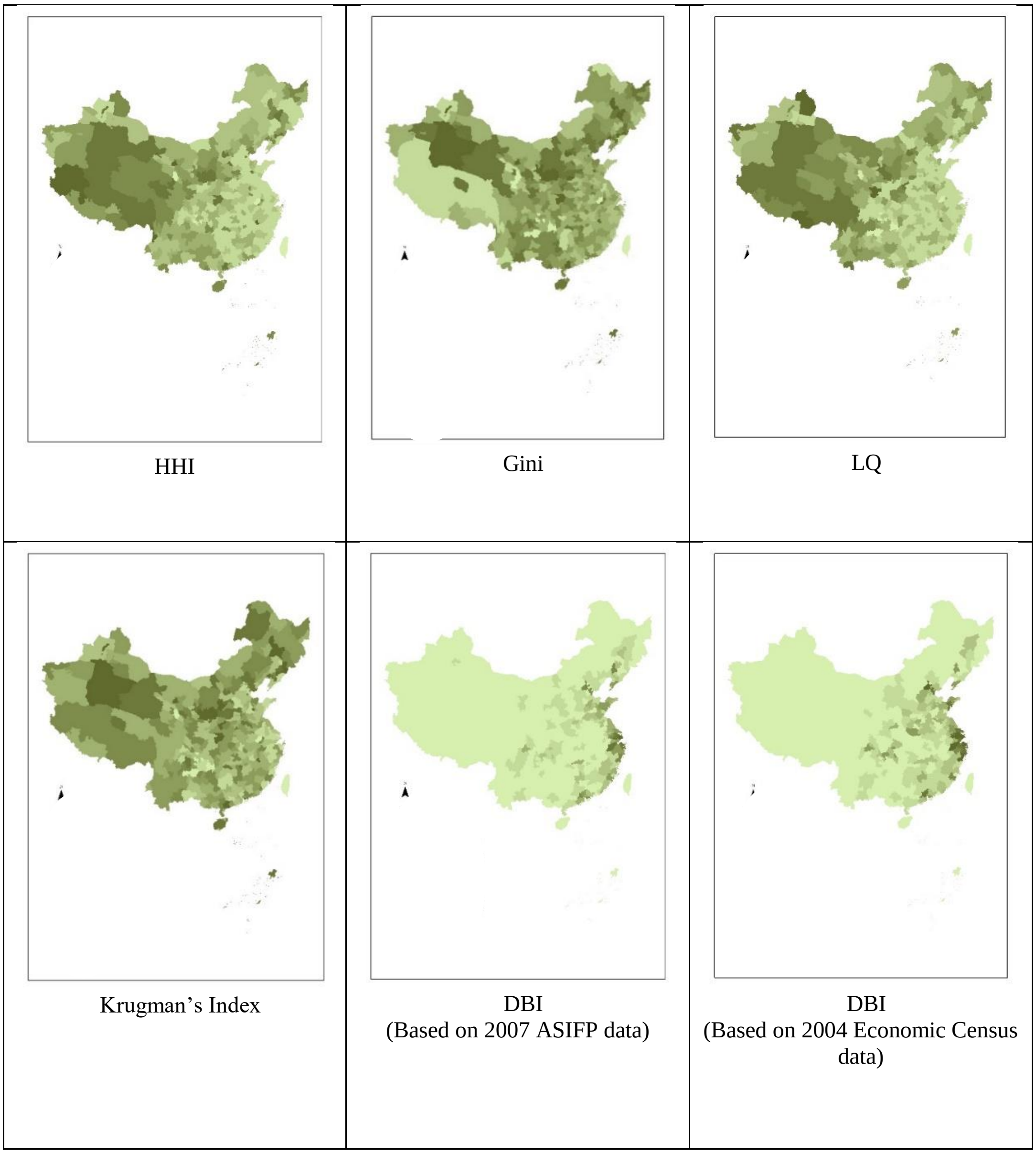
Note: For convenience, we do not present all the control variables in the table. The control variables for both 1st and 2nd stage estimations are the same as those shown in Table 5. The values in parentheses are t- statistics; * = $p < 0.1$; ** = $p < 0.05$; *** = $p < 0.01$.

Table 9: Two-stage Estimations on Entrepreneurial Clusters and Rural Household Per Capita Income

	Panel A: 1st stage estimation		Panel B: 2nd stage estimation	
	(1) Nonstate_V	(2) Nonstate_E	(1) Rural income	(2) Rural income
Buddhism	0.017^	0.014		
	(1.150)	(0.924)		
Inflow_Migrant	0.017***	0.018***		
	(5.069)	(5.350)		
Nonstate_V			1.408***	
			(4.430)	
Nonstate_E				1.298***
				(4.565)
Time Fixed Effects	Yes	Yes	Yes	Yes
County Fixed Effects	Yes	Yes	Yes	Yes
N	12734	12734	12734	12734
P-value of F-tests	0.000	0.000	0.000	0.000
Underidentification test p-value	0.000	0.000		
Sargan's test p-value	0.998	0.811		

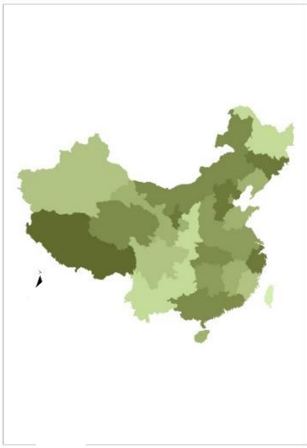
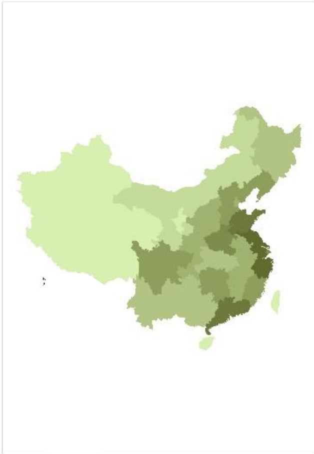
Note: For convenience, we do not present all the control variables in the table. The control variables for both 1st and 2nd stage estimations are the same as those shown in Table 6. The values in parentheses are t- statistics; * = $p < 0.1$; ** = $p < 0.05$; *** = $p < 0.01$, ^ =$p < 0.15$.

Figure 1.a: Prefecture Level Cluster Counts Measured by Standard Approaches and DBI
(Darker areas represent a higher level of clustering)



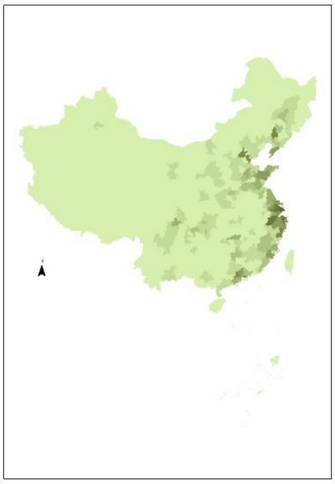
* The HHI, Gini, LQ, Krugman-Index are calculated from the ASIFP 2007 data.

Figure 1.b: Provincial Level Cluster Counts Measured by Industrial Proximity and DBI (Darker areas represent a higher level of clustering)

 Industrial Proximity (Based on 2004 Economic Census data)	 DBI (Based on 2004 Economic Census data)

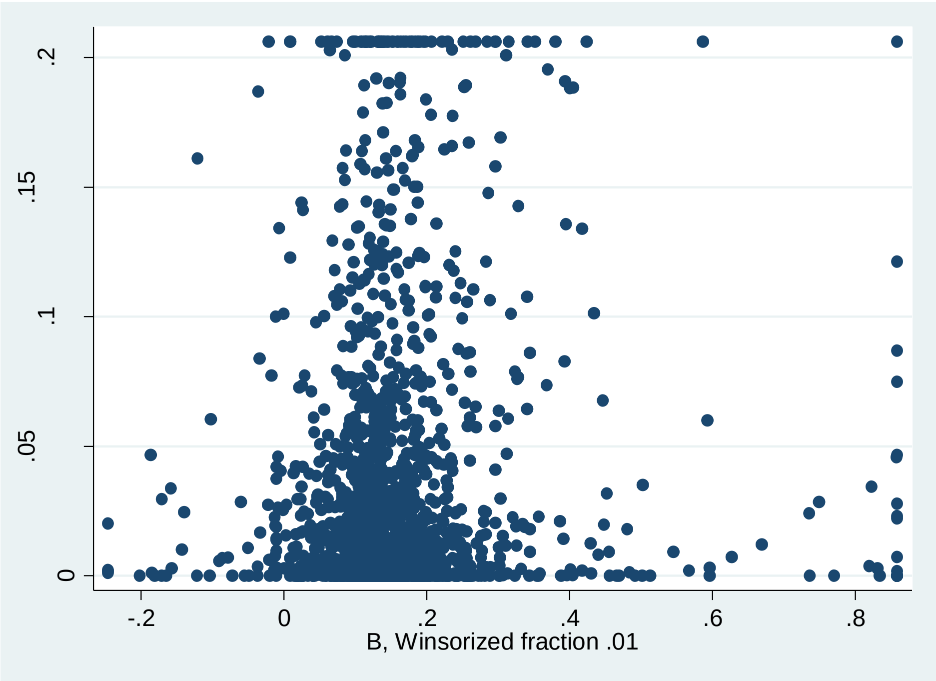
* The industrial proximity is from Long and Zhang (2012).

Figure 2: Satellite Night Vision of China vs. Prefecture-Level DBI Cluster Counts

 Satellite Photo of China at Night*	 Prefecture-level Cluster Number (Based on 2007 ASIFP data)

* The Satellite Night Vision is taken from the NASA website, and the brightness represents light at night (http://earthobservatory.nasa.gov/NaturalHazards/view.php?id=79790)

Figure 3. Relationship between Land Revenue Ratio and Economic Growth in 2006



Appendix

Table A.1 Variable Definitions

Variables	Definitions
Dependent Variables	
Economic growth	Growth rate of per capita GDP in a given county in a given year in logarithm form: $\ln\left(\frac{p.c.GDP_{it+1}}{p.c.GDP_{it}}\right)$; $p.c.GDP_{it}$ is the per capita GDP of county i in year t
Urban–rural inequality	Ratio of urban household income over rural household income of a given county in a given year in logarithm form: $\ln(ratio_{it})$; $ratio_{it} = per\ capita\ urban\ household\ income_{it} / per\ capita\ rural\ household\ income_{it}$
Urban household income	Per capita urban household income in a given county in a given year
Rural household income	Per capita rural household income in a given county in a given year
Independent Variables	
Cluster	A dummy variable that equals one if county has any industrial cluster in year *t* and equal zero if otherwise.
Str_national_V	Relative contribution of the local cluster to the national output of a corresponding industry
Str_national_E	Relative contribution of the local cluster to the national establishment number of a corresponding industry
Nonstate_V	Contribution of non-state firms to the local cluster(s) in outputs
Nonstate_E	Contribution of non-state firms to the local cluster(s) in establishment number
Control Variables	
p.c. GDP	Per Capita GDP of a given county in a given year
GDP	Total GDP of a given county in a given year
CPI	Provincial-level Consumer Price Index of a given county in a given year
Fraction of industrial output	Ratio of total industrial output over total GDP of a given county in a given year
Fraction of non-state firms	Ratio of the number of non-state-owned firms over the total number of firms in a given country in a given year
Fraction of micro firms	Ratio of the number of micro firms over the total number of firms in a given county in a given year
Fraction of edu expenditure	Ratio of fiscal expenditure in education over total fiscal expenditure of a given county in a given year
Fraction of f. a. investment	Ratio of total investment in fixed assets over total GDP of a given county in a given year
Table A.1 Variable Definitions (continued)	
Fraction of gov expenditure	Ratio of administration expenditure over total fiscal expenditure in a given county in a given year
Poor	A dummy variable that equals one if a county is identified as a national designated poor county in a given year and equals zero if otherwise.
Instrumental Variables	
Mining	The per capita mining outputs of a province in a given year
Land Revenue	The revenue gained from land sales by the county-level

	government in each year.
Buddhism	A dummy variable that is equal to one if a county has a Buddhism temple included in the "Spatial Explorer on Religion dataset" in a given year.
Inflow_Migrant	The number of migrant labors from other provinces as a percentage of the total employment in a given province in each year

Table A.2 Weights of SOEs in Regions Defined by Different Clustering Measurements

Clustering measurements	Identified counties	SOE number /total firm number	SOE output / total output	SOE employment/ total employment
DBI	Clustered	1.63%	12.49%	9.12%
	Non-clustered	5.38%	32.4%	29.69%
CR5 Glaeser et al. (1992)	Clustered	5.62%	44.04%	39.58%
	Non-clustered	2.21%	11.83%	10.8%
GINI Midelfart-Knarvik et al. (2000)	Clustered	3.93%	30.88%	27.67%
	Non-clustered	2.08%	9.16%	9.1%
LQ Glaeser et al (1992) & Porter (2003)	Clustered	4.25%	39.12%	32.9%
	Non-clustered	2.26%	13.06%	11.37%
KDI Krugman (1993)	Clustered	3.22%	29.44%	24.36%
	Non-clustered	2.24%	9.2%	9.62%

Note: County-level firm number, total output, and employment are calculated based on CEC 2004.

Table A2.1: Clustering and Growth (Clustered Standard Error at Provincial-Level)

	Economic growth				
	(1)	(2)	(3)	(4)	(5)
Cluster	0.006☆				
	(1.526)				
Str_national_V		0.010***			
		(2.801)			
Str_national_E			0.005☆		
			(1.612)		
Nonstate_V				0.015**	
				(2.097)	
Nonstate_E					0.014**
					(2.075)
p.c.GDP	-0.137***	-0.139***	-0.138***	-0.137***	-0.137***
	(-4.349)	(-4.352)	(-4.335)	(-4.357)	(-4.353)
p.c.GDP square	-0.012**	-0.012**	-0.012**	-0.012**	-0.012**
	(-2.473)	(-2.499)	(-2.472)	(-2.471)	(-2.471)
CPI	-0.213	-0.207	-0.209	-0.214	-0.213
	(-0.774)	(-0.748)	(-0.755)	(-0.772)	(-0.768)
Fraction of industrial output	0.030***	0.029***	0.030***	0.030***	0.030***
	(5.798)	(5.489)	(5.711)	(5.780)	(5.737)
Fraction of non-state firms	0.022	0.023	0.023	0.022	0.022
	(0.838)	(0.859)	(0.836)	(0.811)	(0.798)
Fraction of micro firms	0.023	0.023	0.022	0.023	0.022
	(1.238)	(1.231)	(1.183)	(1.206)	(1.193)
Fraction of edu expenditure	0.021	0.021	0.021	0.021	0.021
	(1.526)	(1.517)	(1.526)	(1.541)	(1.535)
Fraction of f. a. investment	0.023***	0.023***	0.023***	0.023***	0.023***
	(3.815)	(3.827)	(3.806)	(3.812)	(3.812)
Fraction of gov expenditure	0.018**	0.018**	0.018**	0.018**	0.018**
	(2.366)	(2.370)	(2.364)	(2.362)	(2.362)
Number of SEZ	-0.013	-0.013	-0.013	-0.013	-0.013
	(-1.401)	(-1.370)	(-1.400)	(-1.394)	(-1.389)
Poor	-0.137***	-0.139***	-0.138***	-0.137***	-0.137***
	(-4.349)	(-4.352)	(-4.335)	(-4.357)	(-4.353)
Constant	0.469***	0.466***	0.468***	0.469***	0.469***
	(6.701)	(6.634)	(6.679)	(6.703)	(6.699)
Time fixed effects	Yes	Yes	Yes	Yes	Yes
County fixed effects	Yes	Yes	Yes	Yes	Yes
Clustered standard error at provincial-level	Yes	Yes	Yes	Yes	Yes
N	15875	15875	15875	15875	15875
R-sqaure	0.372	0.372	0.372	0.372	0.372

Note: Values in parentheses are t- statistics; * = p < 0.1; ** = p < 0.05; *** = p < 0.01; ☆ = p < 0.15

Table A2.2: Clustering and Inequality (Clustered Standard Error at Provincial Level)

	Urban-Rural Inequality				
	(1)	(2)	(3)	(4)	(5)
Cluster	-0.011				
	(-1.094)				
Str_national_V		-0.001			
		(-0.099)			
Str_national_E			-0.007		
			(-1.066)		
Nonstate_V				-0.033☆	
				(-1.494)	
Nonstate_E					-0.032☆
					(-1.491)
p.c.GDP	0.043	0.044	0.044	0.043	0.043
	(0.744)	(0.765)	(0.764)	(0.750)	(0.752)
GDP	0.035	0.033	0.035	0.035	0.035
	(0.807)	(0.764)	(0.802)	(0.797)	(0.803)
Fraction of industrial output	0.014**	0.013*	0.015**	0.015**	0.015**
	(2.226)	(2.058)	(2.207)	(2.335)	(2.278)
Fraction of non-state firms	0.082	0.081	0.082	0.084	0.085
	(1.187)	(1.177)	(1.180)	(1.208)	(1.212)
Fraction of micro firms	-0.062	-0.066	-0.060	-0.061	-0.059
	(-0.706)	(-0.740)	(-0.676)	(-0.690)	(-0.668)
Fraction of edu expenditure	0.042	0.042	0.041	0.041	0.041
	(1.620)	(1.617)	(1.623)	(1.603)	(1.611)
Fraction of f. a. investment	0.024**	0.024**	0.024**	0.024**	0.024**
	(2.698)	(2.670)	(2.675)	(2.688)	(2.671)
Fraction of gov expenditure	-0.021	-0.021	-0.021	-0.020	-0.020
	(-1.203)	(-1.216)	(-1.197)	(-1.179)	(-1.181)
number of SEZ	-0.009	-0.010	-0.009	-0.009	-0.009
	(-0.293)	(-0.311)	(-0.282)	(-0.275)	(-0.282)
Poor	0.043	0.044	0.044	0.043	0.043
	(0.744)	(0.765)	(0.764)	(0.750)	(0.752)
Constant	0.520	0.543	0.524	0.523	0.520
	(0.824)	(0.865)	(0.827)	(0.820)	(0.816)
Time fixed effects	Yes	Yes	Yes	Yes	Yes
County fixed effects	Yes	Yes	Yes	Yes	Yes
Clustered standard error at provincial level	Yes	Yes	Yes	Yes	Yes
N	3363	3363	3363	3363	3363
R-square	0.246	0.246	0.246	0.246	0.247

Note: Values in parentheses are t- statistics; * = p < 0.1; ** = p < 0.05; *** = p < 0.01; ☆ = p < 0.15

Table A2.3: Entrepreneurial Clusters and Household Per Capita Income (Clustered Standard Error at Provincial Level)

	Panel A: Rural household income		Panel B: Urban household income	
	(1)	(2)	(1)	(2)
Nonstate_V	0.043***		-0.004	
	(4.064)		(-0.263)	
Nonstate_E		0.042***		-0.010
		(4.153)		(-0.663)
p.c.GDP	0.008	0.008	0.158***	0.158***
	(0.379)	(0.376)	(3.579)	(3.578)
GDP	0.122**	0.122**	-0.060	-0.059
	(2.405)	(2.404)	(-1.430)	(-1.420)
Fraction of industrial output	0.008	0.008	0.027***	0.028***
	(0.462)	(0.460)	(3.277)	(3.261)
Fraction of non-state firms	0.081	0.080	0.063	0.064
	(1.237)	(1.220)	(1.152)	(1.166)
Fraction of micro firms	-0.066	-0.067	-0.065	-0.064
	(-1.371)	(-1.401)	(-0.754)	(-0.736)
Fraction of edu expenditure	0.033	0.033	-0.026	-0.026
	(0.689)	(0.685)	(-1.263)	(-1.272)
Fraction of f. a. investment	0.020*	0.020*	0.041***	0.041***
	(1.865)	(1.865)	(4.319)	(4.319)
Fraction of gov expenditure	-0.042	-0.042	0.025*	0.025*
	(-0.722)	(-0.721)	(1.832)	(1.845)
number of SEZ	-0.002	-0.002	0.041	0.041
	(-0.106)	(-0.106)	(0.838)	(0.840)
Poor	0.013	0.013	-0.009	-0.009
	(0.404)	(0.411)	(-0.469)	(-0.459)
Constant	6.567***	6.567***	9.937***	9.931***
	(10.282)	(10.275)	(17.887)	(17.909)
Time fixed effects	Yes	Yes	Yes	Yes
County fixed effects	Yes	Yes	Yes	Yes
Clustered standard error at provincial-level	Yes	Yes	Yes	Yes
N	13646	13646	3513	3513
R-sqaure	0.455	0.455	0.798	0.798

Note: Values in parentheses are t- statistics; * = p < 0.1; ** = p < 0.05; *** = p < 0.01.

Table A.3. Clustering and Non-agricultural Activity Engagements of Rural Residents

	Counties with clusters		Counties without clusters		Difference between the two groups	
	Obs.	Mean	Obs.	Mean	Difference	p-value
Income from Non-agricultural activities (yuan)	43	1,398	38	1,293	105	0.0764
Ratio of residents (above 16) engaged in non-agricultural jobs in general	43	42.76%	38	35.58%	7.18%	0.0140
Ratio of private business owners in people with non-agricultural jobs	43	4.39%	38	2.94%	1.45%	0.0469
Ratio of residents engaged in non-agricultural jobs in home county	43	50.74%	38	27.39%	23.35%	0.0000

Note: the information shown in this table is based on income and personal characteristics information from Chinese Household Income Project (CHIP) 2007 data.