



THE WILLIAM DAVIDSON INSTITUTE  
AT THE UNIVERSITY OF MICHIGAN BUSINESS SCHOOL

***Local Protectionism and Regional Specialization:  
Evidence from China's Industries***

*By: Chong-En Bai, Yingjuan Du, Zhigang Tao, and Sarah Y. Tong*

William Davidson Working Paper Number 565  
May 2003

# Local Protectionism and Regional Specialization: Evidence from China's Industries\*

Chong-En Bai<sup>†</sup>, Yingjuan Du, Zhigang Tao<sup>‡</sup> and Sarah Y. Tong<sup>§</sup>

April 2003

## Abstract

This paper uses a dynamic panel estimation method to investigate the determinants of regional specialization in China's industries, paying particular attention to local protectionism. Less geographic concentration is found in industries where the past tax-plus-profit margins and the shares of state ownership are high, reflecting stronger local government protection of these industries. The evidence also supports the scale-economies theory of regional specialization. Finally, the overall time trend of regional specialization of China's industries is found to have reversed an early drop in the mid 1980s, and registered a significant increase in the later years.

*Key Words:* local protectionism, regional specialization, scale economy, external economy.

*JEL Classification Number:* F1, R12, P2

---

\*We are grateful to James E. Rauch and two anonymous referees for their valuable suggestions, to Songnian Chen, David D. Li, Barry Naughton, Ivan Png, Shang-Jin Wei, Richard Wong, Lijing Zhu, and participants of the 2003 AEA annual meetings for their useful comments, to Zhimin Cha, Xiaohong Chen, Xuewen Zhou, and the China Statistics Bureau for their generous help in securing some of the data, and to the University of Hong Kong and the Hong Kong Institute of Economics & Business Strategy for their financial support.

<sup>†</sup>School of Economics & Finance, The University of Hong Kong, Pokfulam Road, Hong Kong. Tel: (852) 2859-1036. Fax: (852) 2548-1152. baic@hku.hk.

<sup>‡</sup>School of Business, The University of Hong Kong.

<sup>§</sup>Hong Kong Institute of Economics & Business Strategy, The University of Hong Kong.

# 1 Introduction

Trade facilitates specialization, which in turn leads to more gains from trade. To understand the pattern of trade among geographic units, one needs to investigate the determinants of international/regional specialization. For this reason, the study of geographic concentration in production has been an important area of research in both international and regional economics. Much of the empirical literature on this topic, however, is carried out using sub-national data, and hence focusing on the regional specialization of economic activities (see Hanson (2001), and Overman, Redding, and Venables (2001) for most recent surveys). Such an approach has two advantages. One is that comparable data is more readily available for sub-national units and the other is that it avoids the difficulty of controlling institutional differences across countries in international studies (Davis, Weinstein, Bradford and Shimpo (1997), Bacchetta, Rose, and van Wincoop (2001), and O'Connell and Wei (2002)).

A number of theories have been proposed to account for international/regional specialization of economic activities. One theory emphasizes the disparity in resource endowments across geographic units (Ohlin, 1933). Second, for industries that enjoy increasing returns to scale, there is a natural tendency to have production clustered in a few places as opposed to scattered in many places (Krugman, 1991). Third, even for industries that exhibit constant or decreasing returns to scale, it is possible that a firm's cost of production (or its ability to introduce new products and services) is reduced (or enhanced) by the presence in the same region of other firms in the same industry. Such spillover effects or external economies could then lead to the geographic concentration of production (Marshall, 1920).

While the benefits of trade and specialization are well understood, a pre-condition for realizing these benefits — namely, free flow of goods and services across regions and countries — is not always satisfied due to possible protectionism at both international and sub-national levels. Protectionism creates barriers to trade, making trade more difficult and specialization less beneficial. Therefore, protectionism should have a significant effect on the degree of specialization.

The relationship between protectionism and specialization, however, has not received its deserved attention in the existing literature. As far as we know, there exists no

systematic study, especially empirical study, of the relationship in the literature. The lack of study on this issue in the international context is probably due to the data problem discussed earlier, whereas the deficiency of sub-national studies on the issue is partly because in many countries, such as the United States, interregional trade barriers are prohibited by the national government and therefore local protectionism is not a factor. The case of China is different and it provides us with a unique opportunity to study the role of local protectionism in regional specialization. China's economic reform since 1978 has introduced fiscal decentralization, which provided the local governments with a strong incentive to protect their tax base by shielding local firms and industries from interregional competition. The local governments also have incentives to protect state-owned enterprises under their administration, which are their base of political power, their source of private benefits as well as fiscal revenue. Meanwhile, there was no promulgation in the early years of economic reform, and no effective implementation in the later years, of central-government policies that prohibit interregional trade barriers. Therefore, local protectionism is an important factor in China's regional specialization.

There is considerable controversy about the degree of local protectionism in China. Young (2000) provides anecdotal evidence on the rise of local protectionism in China during the reform era especially in the 1980s. He also presents statistical evidence of declining regional specialization based on the evolution of the five sectors in the socialist measure of national income (agriculture, industry, construction, transport, and commerce) and on the evolution of the three sectors in GDP accounting (primary, secondary, and tertiary). Naughton (1999), on the other hand, uses data from the input-output tables among Chinese provinces in 1992, and finds evidence consistent with increasing regional specialization between 1987 and 1992. A systematic study on local protectionism as a determinant of regional specialization and a further investigation on the time trend of regional specialization in China would shed useful light on this controversy.

We construct a panel data set of 32 two-digit industries in 29 Chinese regions<sup>1</sup> over the period of 13 years between 1985 and 1997. Our data on regional specialization are more disaggregated than those used by Young (2000), and cover a longer and more

---

<sup>1</sup>The sample includes 29 provinces, autonomous regions, and municipalities directly under the central government. Hainan gained the status of a province in 1988. However, its data are included in Guangdong province in this study.

recent time period than those used by Naughton (1999). Using the data, we study not only the overall time trend but also the determinants of regional specialization. We pay particular attention to the role of local protectionism, in addition to the traditional theories of regional specialization. Specifically, it is conjectured that local governments tend to protect industries that yielded high profit and tax in the past, thereby reducing the geographic concentration in those industries. Local protectionism is also expected to be significant for industries with large shares of state ownership.

Because the relocation of industrial activities is a slow process, their distribution across regions should be strongly influenced by its historical pattern as well as the factors that we just discussed. To accommodate this consideration, we estimate a dynamic panel structure in which the lagged values of the dependent variable, namely, the degree of regional specialization, appear on the right hand side of the equation together with other explanatory variables. A procedure developed by Arellano and Bond (1991) is used for the estimation.

Our empirical findings generally support our hypotheses about local protectionism. Other things being equal, regional specialization is found to be low for industries that yielded high profit and tax in the past, and for industries with large shares of state ownership. Our study also lends strong support to the scale-economies theory and weak support to the external-economies theory of regional specialization. Because we cannot find a satisfactory measure of the degree of reliance on immobile resources of the industries, we do not test the resource-endowment theory. Finally, the overall time trend of China's regional specialization of industrial production has reversed an early drop in the mid 1980s, and registered a significant increase in the later years. This finding contributes to the settling of the debate over the time trend of local protectionism in China.

The remainder of the paper is organized as follows. In Section 2, we discuss the theories of regional specialization in more detail and develop hypotheses based on them. In Section 3, we construct some variables for the testing of various hypotheses. Descriptive statistics of key variables are offered and compared with some of the findings in the existing literature. Section 4 presents econometric testing of the hypotheses and assesses the relevance of various theories in the context of China. The paper concludes with Section 5.

## 2 Theories and Hypotheses

Regional specialization of industrial production within a country shares many common features with international specialization and has received considerable attention in the study of international trade and regional economics.

The first theory of regional specialization is a natural extension of the resource-endowment theory of international specialization (Ohlin, 1933). Different regions are endowed with different sets of natural, physical, and human resources. When trade among different regions is possible, each region specializes in producing a subset of goods and services. The pattern of specialization is determined by the comparative advantages of regions implied jointly by resource endowments and technological capabilities. However, it is important to note that this theory is based on a crucial assumption that factors of production are immobile. To summarize the above discussion, we have:

*Hypothesis 1: Industries with heavy employment of immobile resources are geographically concentrated.*

The second theory of regional specialization comes from the scale-economies theory of international trade (Krugman, 1991). In an industry where there is a significant fixed cost of production or a decreasing average variable cost of production, a firm would enjoy a low average cost of production by producing a large volume of goods and services, which in turn enhances the firm's competitiveness and increases the demand for its products. The positive feedback eventually leads to a high concentration of production. Hence we have:

*Hypothesis 2: Geographic concentration is more likely in industries that exhibit increasing returns to scale.*

The third theory of regional specialization is that of external economies (Marshall, 1920). There are three main channels through which the presence in the same region of other firms in the same industry may exert positive spillover effects: a cluster of an industry attracts specialized suppliers, it allows labor-market pooling, and it helps foster knowledge spillover (Enright, 1990; Krugman and Obstfeld, 2000). The first two channels imply that a firm's cost of production is reduced by being among a cluster

of other firms in the same industry, whereas the third channel suggests that a firm is more likely to develop new products and services by being among the cluster. Under each of the three channels, a positive feedback arises, which eventually leads to regional specialization of industrial production. For empirical tests of the theory, please see influential studies by Rauch (1993), Dumais, Ellison, and Glaeser (1997), and Rosenthal and Stranger (2001). We summarize the above discussion with the following hypothesis.

*Hypothesis 3: The degree of regional specialization is higher for industries that enjoy significant external economies.*

Underlying each of the above three traditional theories of regional specialization is the assumption of free flow of goods and services across regions. If each region were an isolated island, then there would not be any specialization in industrial production across the regions, even if there were scale economies, or external economies, or significant disparity in resource endowment. In general, there is interregional trade in goods and services, the ease of which, however, depends on the severity of local protectionism, among other factors. Therefore, the degree of regional specialization depends on the extent of local protectionism.

It should be pointed out that local governments in almost every country, whether it is economically developed or still developing, have the incentive to protect their local industries. This is because local governments rely on their local industries for tax revenue. In addition, local industries, when profitable, offer stable employment opportunity for the local people, which is crucial in elections in economically developed economies and for social stability in transition economies (Bai, Li, Tao, and Wang, 2000). To ensure a solid tax base and maintain employment, local governments can erect various barriers of trade to protect local industries from interregional competition. This problem is similar to the protectionism in international trade. Compared with international trade among countries, however, it should be easier to ensure smooth interregional trade as the national government does have authority over local governments. Indeed, in the United States, the constitution prohibits interstate tariffs. This has greatly facilitated interregional trade of goods and services and led to regional specialization of industrial production.

During China's economic transition in the past two decades, anecdotal evidence

suggests that there is substantial flow of goods and services across regions, though local protectionism has been from time to time a serious problem. The main force behind local protectionism arises from some mismatch in the economic policies during the reform era. Prior to the economic reform in 1978, China had a highly centralized fiscal system. All the tax revenue collected had to go first to the central government. The planning commission of the central government had the authority to decide the expenditure of the local governments and allocate revenue from the central pool (Qian, 2000). Such a system delinked tax revenue and expenditure at the level of local governments, and provided little incentive for local protection or even local production. Since 1978, fiscal decentralization has been introduced, which allows the local governments to retain a percentage of the revenue collected and therefore provides them with a strong incentive to protect local industries. What is lacking in the fiscal reform is the promulgation in the early years and effective implementation in the later years of a policy that prohibits barriers to interregional trade.

It is difficult to directly measure the extent of local protectionism in China. Protection is not carried out by imposing tariffs or setting quota on inter-regional trade, rather by administrative decrees that are designed ostensibly for other purposes. For example, the Shanghai government protected the local automobile industry by adopting environmental regulations that were tailored to the technical specifications of locally produced passenger cars, effectively shutting out cars produced in other regions. These means of protection are idiosyncratic in nature and it is difficult to develop a measure of them that can be applied to all industries.

To develop testable hypotheses regarding local protectionism, we take an indirect approach by looking into the benefits that the local governments can derive from protecting local industries from interregional competition and inferring which industries the local governments would like to protect. First of all, as pointed out earlier, the local governments rely on local industries for tax revenue. It is thus conjectured that the local governments want to protect industries that have high tax margins. For state-owned enterprises, which remain significant in China despite two decades of economic reform, the local governments care about their profits as well. Furthermore, due to the lack of rule of law, even profits of privately owned enterprises are subject to some degrees of expropriation, in the form of ad hoc taxes and fees, by the local governments. We thus

summarize our above discussion with the following hypothesis.

*Hypothesis 4: Geographic concentration is low for those industries that had high tax-plus-profit margins in the past.*

Note that it takes time for the local governments to learn which industries have high tax-plus-profit margins, and hence the use of lagged tax-plus-profit margins in the above hypothesis. It should be emphasized that, though the local governments also care about the employment created by the local industries, their pursuit for the tax-plus-profit margins is overriding. This is because supporting money-losing industries just for the sake of maintaining local employment is not sustainable. It is only when the local industries are profitable, then the created employment is stable, consequently providing long-lasting benefits to the local governments. Increasing evidence suggests that local governments in China are anxious to get rid of money-losing enterprises so that they do not have to deal with the eventual loss of employment from these enterprises (Kung, 1999). We therefore subsume the local governments' benefits from local employment in the hypothesis on the tax-plus-profit margins.

Next, it is an undeniable fact that the local governments in China derive much more benefits from the state-owned enterprises than from other types of enterprises. As the local governments/officials hold the right to appoint the chief executives of state-owned enterprises, they have many more ways of milking the state-owned enterprises as compared with other enterprises. For example, local government officials can arrange employment for their relatives, friends, and political supporters in the state-owned enterprises. Local governments can also divert money from the state-owned enterprises, even publicly listed ones, to public works at best and personal uses at worst. Advertising and sponsorship by state-owned enterprises in government-led activities are considered politically correct and actively encouraged. Given the special benefits from the state-owned enterprises, the local governments have stronger incentive to protect them. Therefore, we have:

*Hypothesis 5: Regional specialization is low for industries with high shares of state ownership.*

### 3 Data and Measurement

In this section, we first discuss how to construct a measure of regional specialization of industrial production. Then we define and measure other variables that will be used for testing the hypotheses discussed in the previous section. Finally, we present some summary statistics. In particular, we discuss the time trend of regional specialization in China, and comment on the related work by Naughton (1999) and Young (2000).

#### 3.1 A measure of regional specialization

One way to measure regional specialization is to quantify the interregional trade patterns resulting from specialization. This approach is widely adopted for studying division of labor and specialization in the global economy. Compared with trade among different countries, however, data on interregional trade within a country are difficult to come by. Hence, in this paper, we take a more direct approach to measuring the degree of regional specialization: namely, mapping out the geographic distribution of production activities in each industry and normalizing it by that of overall production activities.

Output data of 32 industries in 29 Chinese regions are obtained from: the *China Statistical Yearbook* for 1985–1987, the *China Statistical Yearbook on Industrial Economy* for 1988–1994 and 1997, the *China Industrial Census* for 1995, and the China Statistics Bureau for 1996.<sup>2</sup> With the output data, we then construct a measure of regional specialization called *Hoover coefficient of localization* (1936). It is based on the location quotient with respect to output, which is defined as

$$L_{ij} = \frac{OUTPUT_{ij}}{OUTPUT_i} \bigg/ \frac{OUTPUT_j}{OUTPUT}$$

where  $OUTPUT_{ij}$  is output of industry  $i$  in region  $j$ ,  $OUTPUT_j$  is total output in region

---

<sup>2</sup>While the data are obtained from different statistical yearbooks, they are all compiled by the same China Statistics Bureau and are supposed to follow a common set of statistical criteria. In general, the most detailed industry-by-region data are provided by the *China Statistical Yearbook* in the early years, but by the *China Statistical Yearbook on Industrial Economy* in the later years. For 1995 and 1996, there was no publication of the *China Statistical Yearbook on Industrial Economy*, the reason being the publication of the *China Industrial Census* of 1995 (private communications with officials at the China Statistics Bureau). The data for 1995 are thus from the *China Industrial Census*, while those for 1996 are kindly provided by the China Statistics Bureau. For a summary of data sources, please refer to the Data Appendix.

$j$ ,  $OUTPUT_i$  is total output of industry  $i$ , and  $OUTPUT$  is total industrial output of China.<sup>3</sup> If  $L_{ij}$  is larger than one, then region  $j$  has a higher percentage of industry  $i$  than of total industrial output. Similarly, if  $L_{ij}$  is smaller than one, then region  $j$  has a lower percentage of industry  $i$  than of total industrial output. Given the location quotients of industry  $i$  for all regions  $j = 1, \dots, R$ , we rank regions by their location quotients in descending order and get a sequence of regions. Then, following that sequence, we calculate the cumulative percentage of output in industry  $i$  over the regions ( $y$ -axis) and the cumulative percentage of output in all industries over the regions ( $x$ -axis), and thus plot the localization curve for industry  $i$ . If the industry is evenly distributed across regions, then the location quotient will be equal to one for all regions, and the localization curve will be the 45-degree line. If the industry is more regionally concentrated, then the localization curve will be more concave. Analogous to the Gini coefficient for income distribution, the Hoover coefficient of localization (henceforth denoted by *HOOVER*) is defined as the area between the 45-degree line and the localization curve divided by the entire triangular area in which the localization curve is contained. Thus the Hoover coefficient is between 0 and 1, and the higher its value the more localized is the industry.

A Hoover coefficient of localization can also be constructed from the employment data. In fact, this approach is used in a number of studies on the regional specialization of economic activities in the United States (Kim, 1995; Ellison and Glaeser, 1997; Dumais, Ellison, and Glaeser, 1997). In this paper, we construct a Hoover coefficient using output data instead of employment data, for two reasons. First, there are fewer employment data than output data. Employment data of 32 industries in 29 Chinese regions can be obtained from China Industrial Census for 1985, from the China Statistical Yearbook on Industrial Economy for 1988-1994 and 1997, and from China Statistics Bureau for 1996. The employment data for other years are disaggregated only to the level of industry, not the level of industry by region, which makes it impossible to calculate the Hoover coefficient. The lack of employment data for three (i.e., 1986, 1987 and 1995) out of the thirteen years would pose significant challenges when dynamic panel data estimation

---

<sup>3</sup>When constructing the Hoover coefficient of localization, we generally use output values in current prices. Unfortunately, for 1985 and 1986, output values in current prices are not available; instead their values in 1980 constant prices are used. This may cause discrepancy in calculating the numerator of the location quotient. But we believe the discrepancy to be small, as the prices of various products in the same industries are expected to move closely together.

procedures (e.g., the Arellano-Bond procedure) are used in the econometric analysis. Second, employment data may suffer from the surplus labor problem that is particularly prevalent in state-owned enterprises (SOEs). As the extent of the surplus labor problem varies across regions and industries, the Hoover coefficient obtained through employment data will be biased. Indeed, we find that, while the overall correlation between the Hoover coefficient calculated using output data and that using employment data is more than 95%, it is decreasing slowly and consistently over time, from 96.89% in 1988 to 92.39% in 1997, which indicates increasing surplus labor problem in SOEs and growing biases of the Hoover coefficient calculated from the employment data.

An alternative measure of localization is developed by Ellison and Glaeser (1997). Their index controls for differences in the size distribution of plants by factoring in the Herfindah index of the concerned industry. Since we do not have data on China's industries that allow us to compute or estimate the Herfindah indices for the industries in the sample period, we do not use the Ellison-Glaeser index, and will control for the scale difference among the industries by including the average firm size of an industry as an explanatory variable for the degree of localization. Recall Hypothesis 2 and see Section 3.2.B for details.

### **3.2 Other variables**

Next we turn to the challenge of finding variables for testing the hypotheses discussed in Section 2.

#### **A. Resources**

A crucial assumption for the resource-endowment theory of regional specialization is that certain resources are immobile or their transportation costs are exceedingly high. To test the hypothesis that resource-based industries tend to be localized, we therefore need to find an appropriate measure of those immobile resources.

In a study on regional specialization in the United States, Kim (1995) uses the cost of raw materials divided by total value added as the measure of resource intensity. Note that the measure is a ratio of the value of all material inputs to the industry's total value added. However, not all inputs are equally immobile; thus the measure used by Kim may not reflect the industry's true dependence on immobile resources. To illustrate

this point, consider China's electronics industry, which is dominated by low-value-added OEMs using expensive inputs such as embedded chips. According to Kim's measure, the resource dependence rate is very high, but the inputs involved, such as the embedded chips, are highly mobile. Furthermore, for the case of China, there is another drawback with Kim's measure, namely, the raw materials are often under government price control and therefore the measure of resource intensity could be significantly undervalued. Take for example China's tobacco industry. The prices for the raw materials are kept low due to the government policy of supporting industrial development at the expense of rural development (the price scissors phenomenon studied by Sah and Stiglitz, 1984). Thus Kim's measure of resource intensity would be low for China's tobacco industry, though the actual degree of resource dependence is very high.

To demonstrate the potential problems with the use of Kim's measure for China, we compute the raw-material intensity, defined as the difference between output and value added divided by value added, for the year 1992. As shown in Table 1, highly resource-based industries such as coal mining & processing, ferrous & nonferrous metals mining & processing, nonmetal minerals mining & processing, petroleum & natural gas extraction, and tobacco processing have low values of Kim's measure. In contrast, less resource-based industries such as food processing & production, garments & other fiber products, and textile have high values of Kim's measure. We conclude that this measure is not appropriate for Chinese industries.

One alternative measure for an industry's dependence on resources is the energy consumption intensity, which is defined as the ratio of total energy consumption to total output. The rationale for this measure is based on the observation that coal is the most important energy source for industries in China, and freight transportation of coal in China has been expensive.<sup>4</sup> However, this measure is not without its own problems, because more diverse energy sources are used in more recent years and some new energy sources are more mobile than coal. In addition, there has been significant efficiency improvement in energy consumption, which implies the decreasing importance of energy resources in the overall resource endowments. On balance, we still believe that energy consumption intensity is a more appropriate measurement for an industry's dependence on immobile resources than the measure used by Kim (1995). Unfortunately, except for

---

<sup>4</sup>Similar measures have been used in Rosenthal and Stranger (2001).

the year of 1995, industry-level data on energy consumption are based on various non-standard industry classification systems which are significantly different from the one our data are based on. Given that dynamic panel data estimation methods will be employed to test the various hypotheses of regional specialization, the lack of usable energy consumption data for all but one year prevents us from testing the specific hypothesis on resource-endowment theory of regional specialization. This sacrifice is, however, justifiable, as the focus of this paper is to examine the impact of local protectionism on regional specialization.<sup>5</sup> We will discuss the potential missing variable bias in Section 4.

## B. Scale economies

To test the hypothesis that industries characterized by increasing returns to scale should be geographically concentrated, we use average firm size in an industry as a measure of scale economies. To be consistent with our measure of geographic concentration, we use output data to calculate the average firm size. Data on output and number of firms at the industry level are obtained from the *China Statistical Yearbook on Industrial Economy* for 1985–1994 and 1996–1997, and from the *China Industrial Census* for 1995.<sup>6</sup> A price deflator<sup>7</sup> for industrial output is constructed and used to obtain the output value in constant terms. A panel data set of the average firm size in an industry, defined as the total output in an industry divided by the total number of firms in the industry and denoted by *SCALE*, are readily constructed across 32 industries for 1985–1997.

## C. External economies

External economies, through enhanced supply of specialized inputs, labor-market pooling and knowledge spillover, are generally difficult to measure directly. In this

---

<sup>5</sup>The resource-endowment theory of regional specialization has been extensively investigated in the existing literature. In their study on the geographic concentration in the U.S. industries, Dumais, Ellison and Glaeser (1997) use more recent and disaggregated data than Kim (1995) and find significant shifts in industrial activity across regions, which suggests increasing irrelevance of the resource-endowment theory. Presumably, as transportation becomes less costly, the key assumption for the resource-endowment theory — immobile resources — may no longer hold in both developed and developing economies.

<sup>6</sup>Though the *China Statistical Yearbook* provides the most updated industry-by-region data for the early years of the sample as stated in footnote 2, the 1993 *China Statistical Yearbook on Industrial Economy* gives the historical data on industry-level output and number of firms, which are suitable for constructing the variable of average firm size and are subsequently used due to their consistency.

<sup>7</sup>The price deflator for year  $t$  is calculated as  $IND\_curt \div (IND\_cur1978 \times INDEX_t)$ , where  $IND\_curt$  denotes gross domestic industrial product in current price for year  $t$ , and  $INDEX_t$  is the index of gross domestic industrial product for year  $t$  in comparable prices (1978 = 100). For consistency, all data needed for calculating the price deflator for the period of 1985–1997 are from the *China Statistical Yearbook*.

study, we use the share of engineers and technicians in an industry's employment as a proxy for the external economies. We believe that this variable offers proxy for both labor market pooling and knowledge spillover, following the arguments and analyses in Dumais, Ellison, Glaeser (1997) and Rosenthal Stranger (2001).<sup>8</sup> Data on the number of engineers and technicians and total employment, both in large and medium enterprises, are obtained for the 32 industries from the *China Industrial Census* for 1985, from the China Statistics Bureau for 1987-1989, and from the *China Statistical Yearbook on Science and Technology* for 1990-1997. A panel data set of the share of engineers and technicians in an industry's employment, defined as the number of engineers and technicians divided by the total employment and denoted by *ET*, are readily constructed across 32 industries for 1985-1997.<sup>9</sup>

#### D. Tax-plus-profit margin

Prior to the economic reform in 1978, most firms were state owned and both their profits and tax payments were counted as government revenue. In fact, the official statistics only reported tax plus profit as a combined item and did not report their separate figures for the early years in our sample period. Data on *tax plus profit* and *sales* for the 32 industries are obtained from the *China Industrial Census* for 1995, and the *China Statistical Yearbook on Industrial Economy* for 1985-1994 and 1996-1997. A panel data set of the tax-plus-profit margin, defined as tax plus profit divided by sales and denoted by *TPM*, is constructed across the 32 industries and for the period of 1985-1997.

---

<sup>8</sup>Dumais, Ellison, Glaeser (1997) use firm-level data from the U.S. Census Bureau's Longitudinal Research Database to construct proxies for the three channels of external economies of regional specialization, while Rosenthal and Stranger (2001) rely extensively on the data from the U.S. Bureau of Economic Analysis. In our study on regional specialization of industrial activities in China, however, we are severely constrained by data availability. For most of the U.S. Bureau of Economic Analysis data, there are simply no counterparts in China's economic statistics. Furthermore, the China Statistics Bureau makes its census data available only at the industry level. We therefore focus on the variable of the share of engineers and technicians, which is similar to the percentage of workers with Doctorates, Master's degrees and Bachelor's degrees used in Rosenthal and Stranger to proxy for labor market pooling, and also similar to the percentage of the employment with the college degree used by Dumais, Ellison and Glaeser (1997) as an interaction variable for several proxies of knowledge spillover.

<sup>9</sup>While entering data from the *China Statistical Yearbook on Science and Technology*, we noticed there was a significant change in the terminology of various statistical variables therein for the year of 1993, whereas the data both before and after 1993 followed the same terminology. Private conversation with officials in the China Statistics Bureau revealed that, in 1993, a reform in the statistical criteria for various variables in this yearbook was attempted and subsequently aborted, hence the inconsistency of the data. Therefore, in our econometric analysis, for the year of 1993, we use the simple average of the shares of engineers and technicians in 1992 and 1994.

## E. Share of SOEs

The share of state-owned enterprises (SOEs) in an industry can be measured in several dimensions. It could be the percentage of output, or sales, or employment of the state-owned enterprises in the industry's overall figures. Unexpectedly, data for calculating the share of SOEs are the most difficult to come by. Publicly available sources such as the statistical yearbooks contain data mostly at the aggregated levels, or data collected with changing statistical criteria partly due to the decreasing role of state ownership. In an earlier version of this paper (Bai, Du, Tao and Tong, 2002), we constructed a panel data of SOE employment share from the various sources,<sup>10</sup> and found that the resulting SOE employment share had a significant drop between 1992 and 1993, which was not easily explained. Private conversation with officials in the China Statistics Bureau revealed that more detailed and consistent data on the share of state-owned enterprises were collected, but they have only been made available to researchers for internal use. In preparing this version of the paper, we collaborated with researchers in the Development Research Center of the State Council, People's Republic of China, and had access to a set of panel data on the share of SOEs in industrial output (*SSOE*) for the 32 industries and over the period of 1985-1997.

### 3.3 Summary statistics

As described in Section 3.1, Hoover coefficients of localization are calculated using output data for the 32 two-digit industries over the period of 1985-1997. One way of examining the Hoover coefficients is to trace the time trend of all industries as a whole. As shown in Figure 1, the simple average across all industries was 0.313 in 1985. It went down slightly till 1987 and then rose steadily to 0.343 in 1997. The trend is similar for the

---

<sup>10</sup>For 1986-1988, data on SOE employment and COE (collectively-owned enterprises) employment, and their combined share in the total employment are available from the *China Statistical Yearbook*. The above data are used to calculate the total employment in an industry. The share of SOE employment, defined as SOE employment divided by total employment, follows immediately. For 1993-1994, SOE employment is not provided, but it can be calculated from SOE value added and SOE productivity (defined as SOE value added divided by SOE employment), both of which are available from the *China Statistical Yearbook*. As the total employment in an industry is also available, the share of SOE employment can be calculated. Finally, for 1988-1992, 1995 and 1997, data on both SOE employment and total employment are available from, respectively, *China Statistical Yearbook*, *China Industrial Census*, and *China Statistical Yearbook on Industrial Economy*. Hence the share of SOE employment can be calculated, in some cases through indirect ways that may have resulted in significant inconsistency in the data.

weighted average. The weighted (by the output values) average across all industries was 0.256 in 1985. It decreased to 0.250 in 1988 and then increased for all later years to 0.304 in 1997. The aggregate coefficients clearly indicate that, over the 13-year period of 1985–1997, regional specialization of Chinese industries increased quite substantially. Our results are in sharp contrast to those in Young (2000) but are consistent with those in Naughton (1999).

Another way of examining the Hoover coefficients of localization is to compare the cross-time averages for various industries. As shown in Table 2, there are large variations in the Hoover coefficients across industries, ranging from 0.146 (metal products) to 0.847 (logging & transport of timber & bamboo). Mining industries, which depend heavily on resources, are more localized than manufacturing industries: the average Hoover coefficient for mining industries over the 13-year period is 0.613, while that for manufacturing is only 0.273.<sup>11</sup> Even within manufacturing industries, there exist significant differences. Tobacco processing is the most localized, followed by cultural, educational & sports goods, and electronics & telecommunications. Metal products, raw chemical materials & chemical products, and ordinary machinery & special purposes equipment are the three least localized industries.

Regarding the measure for the external-economies theory (share of engineers and technicians) and that for the scale-economies theory (average firm size), there are also large variations across industries. Electronics & telecommunications has the highest share of engineers and technicians (13.99% of the total employment), followed by instruments, meters, cultural & clerical machinery (12.63%), and medical & pharmaceutical products (10.85%). On the other hand, coal mining & processing has the lowest share of engineers and technicians (2.73%), with garments & other fiber products and textile having slightly higher shares, 2.85% and 3.22% respectively, of engineers and technicians. Next we examine the variation in average firm size across industries. Petroleum & natural gas extraction has the largest firm size among the industries: RMB5,103,000 per firm, which is five standard deviations more than the mean value (RMB216,000 per firm). Industries with the second and third largest firm sizes are tobacco processing

---

<sup>11</sup>The figures are obtained by taking simple averages of the relevant data from Table 2. Mining industries include industries 6 to 12; manufacturing industries, 13 to 42.

(RMB687,000 per firm) and chemical fibers (RMB184,000 per firm), respectively. The two industries with the smallest firm size (RMB6,000 per firm) are furniture manufacturing, and timber processing, bamboo, cane, palm fiber & straw products.

Finally, we discuss the variables used for testing the hypotheses on local protectionism, *TPM* (tax-plus-profit margin) and *SSOE* (share of state-owned enterprises in industrial output). As shown in Figure 2, the weighted average of *TPMs* across all industries first underwent a dramatic decrease from 21.0% in 1985 to 11.6% in 1990, and then declined gradually to 9.1% in 1997. This is a result of the economic reform that began in late 1978. Between 1949 and 1978, the Chinese economy was characterized by a system of central planning. Two important manifestations of central planning were the lack of competition and the suppression of factor prices, both of which implied high profit margins for industrial production. The economic reform since 1978, however, has unleashed forces that have increased product market competition and raised the factor prices, resulting in lower profit margins for industrial production. The phasing-out of central planning has made it easier for both local governments and private entrepreneurs to enter various industries, increasing the competitive pressure in the product market. Meanwhile, the restrictions on prices of various inputs have gradually been eliminated, resulting in higher and more volatile market prices, which increase the cost for most industrial production. The stable profit margins since 1991 signal the maturing of the competitive markets in China.<sup>12</sup> The time trend of the share of state-owned enterprises in industrial output (namely, *SSOE*) is shown in Figure 3, and it clearly confirms the commonly held perception that the state ownership has declined substantially during the reform era. The cross-industry weighted average of *SSOE* decreased from 73.11% in 1985 to 68.00% in 1990, and then plunged to 40.92% in 1997.

Besides the clear-cut time trends of *TPM* and *SSOE*, Table 2 shows that there are also significant differences in these variables across industries. The average tax-plus-profit margin across industries was 13.7%. The industry with the highest *TPM* was tobacco processing (56.1%). Electric power, steam & hot water production & supply, and petroleum refining, coking, & gas production & supply came next, at 22.7% and 19.9% respectively. The industries with the lowest *TPMs* were coal mining & processing

---

<sup>12</sup>Although the discussion here is about profits, tax plus profit should follow the same trend as that of profits.

(2.0%), food processing & production (5.9%), and leather, furs, down, & related products (6.0%). Meanwhile, the average share of *SSOE* across industries was 57.38%. Industries with the highest *SSOEs* were petroleum & natural gas extraction (98.24%), tobacco processing (97.37%), and logging & transport of timber & bamboo (96.55%), whereas those with the lowest *SSOEs* were furniture manufacturing (10.95%), garments & other fiber products (11.73%), and plastic products (19.45%).

## 4 Regression Analysis

In this section, we carry out econometric tests of our hypotheses. As discussed in Sections 3.1 and 3.2, a panel data set for 32 industries and 13 years (1985–1997) has been constructed for the following variables: Hoover coefficient of localization (*HOOVER*), tax-plus-profit margin (*TPM*), the share of SOEs (*SSOE*), average firm size (*SCALE*), and the share of engineers and technicians in industry employment (*ET*), with the 1986 data of *ET* missing.

In setting up the model to estimate, we note that the geographical distribution of the industries may adjust slowly and it may depend on its historical pattern as well as the factors that we discussed in Section 2. To account for the possible influence of the history, we consider the following dynamic panel structure:

$$\begin{aligned} HOOVER_{it} = & \delta_1 HOOVER_{i,t-1} + \delta_2 HOOVER_{i,t-2} + \beta_1 TPM_{i,t-l} \\ & + \beta_2 SSOE_{it} + \beta_3 SCALE_{it} + \beta_4 ET_{it} + \alpha_i + \varepsilon_{it}. \end{aligned} \quad (1)$$

where  $\alpha_i$  is the industry-specific effect,  $\varepsilon_{it}$  is the error term, and  $l$  is the years of lag for variable *TPM* used in the regression. We consider cases where  $l = 1, 2,$  or  $3$ . The reason we use lagged value of *TPM* is because it takes time for the local government to identify which industry brings it more tax revenue and therefore is worth protecting. To maintain sufficient sample size, we do not consider lags of more than three years.

The use of the lagged value of *TPM* also mitigates the potential endogeneity problem associated with *TPM*. If *HOOVER* and *TPM* are both endogenously affected by a common factor and the factor is not controlled for, then *TPM* is correlated with

the error term in the regression and the OLS estimate of the equation will be biased. For example, consider an industry where there is exactly one firm in every region that serves local demand for a product that cannot be easily traded across regions. That industry will have low regional specialization and high profit margin due to its local monopoly. However, the reason for the low degree of specialization is not because of local government protection; rather both the low degree of specialization and the high profit margin are functions of the aforementioned characteristic of the industry. Such an endogeneity problem disappears with the inclusion of the industry-specific effects if the common factor mentioned above is time-invariant. If the common factor is time-variant but its time-varying component is not auto-correlated, then the use of one-year lag of  $TPM$  together with the industry-specific effects will solve the endogeneity problem. If the common factor is time-variant but its time-varying component only has first-degree auto-correlation, then the use of two-year lag of  $TPM$  together with the industry specific effects will solve the endogeneity problem. Similar arguments can be made for higher-degree of auto-correlation. If first differencing across time periods is performed to the equations to eliminate the industry-specific effects, then longer lag of  $TPM$  is needed to solve the endogeneity problem for a given degree of auto-correlation.

It is not straightforward to estimate equation (1) because the lagged dependent variable is correlated with the error term  $\varepsilon_{it}$  even if it is assumed that  $\varepsilon_{it}$  is not itself auto-correlated (Greene, 2000). We use a procedure developed by Arellano and Bond (1991) to estimate the equation.<sup>13</sup> According to the procedure, the industry-specific effects are removed by first-differencing equation (1) across time periods. The resulting equation is then estimated by using the lagged levels of the dependent variable and the pre-determined variables and the difference of the exogenous variables as instrumental variables. We treat the lag values of  $TPM$  as pre-determined variables when applying the Arellano-Bond procedure.

Table 3 summarizes the estimation results of equation (1). In the table, the coefficient of  $TPM$  is always negative regardless of the number of lag years used. It is statistically significant at the 5% level or the 1% level, respectively, when  $TPM$  is lagged for one year or two years. These results strongly support Hypothesis 4 and imply that local governments have stronger incentives to protect industries that have brought them more

---

<sup>13</sup>We thank James E. Rauch for suggesting this procedure to us.

revenue.

The coefficient of *SSOE* is always negative regardless of the number of lag years used for *TPM*. When *TPM* is lagged for three year, the coefficient of *SSOE* is statistically significant at the 5% level. These results suggest that local governments have stronger incentives to protect industries that have a larger share of state-owned enterprises, supporting Hypothesis 5.

The coefficient of *SCALE* is always positive and statistically significant at the 1% level regardless of the number of lag years used for the value of *TPM*. These results strongly support Hypothesis 2 and imply that there is a higher degree of regional specialization in industries where the average size of the firms is larger.

The coefficient of *ET* is also always positive in all cases but is not statistically significant. These results offer weak support to Hypothesis 3 that the degree of regional specialization is higher in industries with larger shares of engineers and technicians.

The coefficient of the one-year lag value of *HOOVER* is positive and highly significant in all cases. So is the coefficient of its two-year lag value, with lower level of statistical significance. We therefore conclude that the degree of regional specialization of an industry is strongly influenced by its history, and that its adjustment in response to changing characteristics of the industry is slow.

To check the robustness of our results, we estimate some variations of equation (1). We add one or two lagged values of *TPM* on the right hand side of equation (1); that is, we estimate

$$\begin{aligned}
 HOOVER_{it} = & \delta_1 HOOVER_{i,t-1} + \delta_2 HOOVER_{i,t-2} + \sum_{l=l_1}^{l=l_2} \beta_l TPM_{i,t-l} \\
 & + \gamma_1 SSOE_{it} + \gamma_2 SCALE_{it} + \gamma_3 ET_{it} + \alpha_i + \varepsilon_{it},
 \end{aligned} \tag{2}$$

where  $l_1 = 1$  or  $2$ ,  $l_2 = 2$  or  $3$ , and  $l_1 \neq l_2$ . The estimation results are summarized in Table 4. The results are remarkably consistent with those in Table 3. The coefficient of the two-year lag value of *TPM* is negative and statistically significant in every case. The coefficient of the one-year lag value is always negative as well even though it is not statistically significant. The coefficient of the three-year lag is positive in all cases, however, its magnitude is always dominated by the magnitude of the coefficient of the two-year lag. The effect of *SSOE* is always negative and it is also statistically significant

in two of the three cases in the table. The effect of *SCALE* is always positive and it is statistically significant in all cases. The effect of *ET* is also always positive but not statistically significant. Finally, the coefficients of the one-year lag and the two-year lag values of *HOOVER* are always positive and statistically significant.<sup>14</sup>

We do not include any proxy for immobile resources in our regressions because we cannot find any suitable one for which panel data is available for our sample period. As the importance of resources declines because of technological advances, their effect on regional specialization may have become less significant. Even if they were still important, the lack of a proxy for them in our regressions would not, we believe, compromise our main conclusion about local protectionism. The reasons are as follows. In China, resources are mainly controlled by government monopoly. Therefore, one should expect a positive correlation between resource reliance and the share of state ownership. In fact, four of the five highly resource-based industries that we discussed in Section 3.2 have high shares of industrial output produced by SOEs. According to Table 2, the SSOE ranks of coal mining & processing, ferrous and nonferrous metals mining & processing, non-metal minerals mining & processing, petroleum & natural gas extraction, and tobacco processing industries are 7, 12, 25, 1, and 2, respectively among the 32 industries. One should also expect a positive correlation between resource reliance and tax-plus-profit margin. The TPM ranks of the highly resource-based industries are 32, 12, 9, 6, and 1, respectively, which are relatively high with the exception of the notoriously inefficient coal mining & processing industry. The positive correlation between resource reliance and each of the two variables related to local protectionism implies that the missing of the former in the regression should result in upward biases in the estimated coefficients of the two protectionism-related variables. Had a proxy for immobile resources been included, the estimated coefficients of the two protectionism-related variables should be more negative and our results about local protectionism should become stronger.

In summary, the results from estimating various specifications of the determinants of the degree of regional specialization are remarkably consistent. They support our hypotheses that the degree of regional specialization is lower in industries where the

---

<sup>14</sup>We also estimated equations in which only the one-year lag value of *HOOVER* is included together with *SSOE*, *SCALE*, *ET*, and the lagged values of *TPM*, on the right hand side. The results are consistent with the findings reported here and are available upon request. However, since the two-year lag value of *HOOVER* has statistically significant coefficients in the regressions we have performed in this section, there does not seem to be a valid reason to drop it from the equations.

tax-plus-profit margin and the share of state ownership are higher, suggesting that the local governments have stronger incentives to protect these industries. They also offer strong support to the scale-economies theory and weak support to the external-economies theory of regional specialization.

## 5 Conclusion

Although protectionism is an important determinant of trade and specialization, there has been no systematic empirical study on this issue in the literature. This paper attempts to fill this void. We construct a panel data of 32 two-digit industries in 29 Chinese regions over a period of 13 years (1985–1997) and use a dynamic panel estimation method to investigate the determinants of regional specialization in China’s industries, paying particular attention to local protectionism. We find that the degree of regional specialization is lower for industries with higher profit-plus-tax margins in the past and for industries with larger shares of state ownership, reflecting stronger incentives for local governments to protect these industries. There are also evidence supporting the scale-economies theory and, weakly so, the external-economies theory of regional specialization.

Despite the evidence for the role of local protectionism, the overall time trend of regional specialization of industrial production in China has reversed an early drop and registered a significant increase in the later years of the reform era. This finding is in contrast to that in Young (2000) but is consistent with that in Naughton (1999). Since our data are more disaggregated than those used by Young (2000) and span a longer and more recent time period than those used by Naughton (1999), our finding contributes to the settling of the debate about the time trend of local protectionism in China.

## References:

- Arellano, Manuel, and Bond, Stephen, "Some Tests of Specification for Panel Data: Monte Carlo Evidence and an Application to Employment Equations." *Review of Economic Studies*, 58, 2:277-297, April, 1991.
- Bacchetta, P., Rose, A. K., and van Wincoop, E., "Intranational Economics and International Economics." *Journal of International Economics*, 55, 1:1, 2001.
- Bai, C. E., Du, Y. J., Tao, Z. G., and Tong, Y. T., "Protectionism and Regional Specialization: Evidence from China's Industries." mimeo, The University of Hong Kong, 2002.
- Bai C. E., Li, D. D., Tao, Z. G., and Wang, Y. J., "A Multitask Theory of State Enterprise Reform." *Journal of Comparative Economics*, 28, 4:716-738, Dec. 2000.
- Davis, D. R., Weinstein, D. E., Bradford, S. C., and Shimpo, K., "Using International and Japanese Regional Data to Determine When the Factor Abundance Theory of Trade Works." *American Economic Review*, 87, 3:421-446, 1997.
- Dumais, G., Ellison, G., and Glaeser, E., "Geographic Concentration as a Dynamic Process." NBER working paper no. 6270, 1997.
- Ellison, G., and Glaeser, E., "Geographic Concentration in U.S. Manufacturing Industries: A Dartboard Approach." *Journal of Political Economy*, 105, 5:889-927, Oct. 1997.
- Enright, M., *Geographic Concentration and Industrial Organization*. Ph.D. thesis, Harvard University, 1990.
- Greene, W. H., *Econometric Analysis*, 4th edition. Prentice Hall, 2000.
- Hanson, G. H., "Scale Economies and the Geographic Concentration of Industry." *Journal of Economic Geography*, 1, 255-276, 2001.
- Hoover, E.M., "The Measurement of Industrial Localization." *Review of Economics and Statistics*, 18, 162-171, 1936.

Kim, S., “Expansion of Markets and the Geographic Distribution of Economic Activities: The Trends in U.S. Regional Manufacturing Structure, 1860–1987.” *Quarterly Journal of Economics*, 110, 881–908, 1995.

Krugman, P., and Obstfeld, M., *International Economics: Theory and Practice*, 5th edition. Reading, MA: Addison-Wesley, 2000.

Krugman, P., “Increasing Returns and Economic Geography.” *Journal of Political Economy*, 99, 483–499, 1991.

Kung, James Kai-Sing, “The Evolution of Property Rights in Village Enterprises: The Case of Wuxi County.” In Jean C. Oi and Andrew G. Walder (Eds), *Property Rights and Economic Reform in China*. Stanford: Stanford University Press, chapter 5, 1999.

Marshall, A., *Principles of Economics*. New York: Macmillan, 1920.

Naughton, B., “How Much Can Regional Integration Do to Unify China’s Market?” mimeo, 1999.

O’Connell, P. G. J., and Wei, S. J., “‘The Bigger They Are, the Harder They Fall’: Retail Price Differences across U.S. Cities.” *Journal of International Economics* 56, 1:21–53, 2002.

Ohlin, B., *Interregional and International Trade*. Cambridge, MA: Harvard University Press, 1933.

Overman, H. G., Redding, S. J., and Venables, A. J., “The Economic Geography of Trade, Production, and Income: A Survey of Empirics.” CEPR discussion paper no. 2978, 2001.

Qian, Y. Y., “The Process of China’s Market Transition (1978–1998): The Evolutionary, Historical, and Comparative Perspectives.” *Journal of Institutional and Theoretical Economics*, 156, 1:151–171, Mar. 2000.

Rauch, J. E., “Productivity Gains from Geographic Concentration of Human Capital: Evidence from Cities.” *Journal of Urban Economics*, 34, 3: 380-400, 1993.

Rosenthal, S. S., and Stranger, W. C., “The Determinants of Agglomeration.” *Journal of Urban Economics*, 50, 2: 191-229, 2001.

Sah, R. K., and Stiglitz, J. E., “The Economics of Price Scissors.” *American Economic Review*, 74, 1:125–138, 1984.

Young, A., “The Razor’s Edge: Distortions and Incremental Reform in the People’s Republic of China.” *Quarterly Journal of Economics*, 115, 4:1091–1135, Nov. 2000.

**Table 1: Raw-Material Intensity for Chinese Industries in 1992**

Industry Code	Industry Name	Raw Material Intensity
13	Food Processing & Production	5.3252
25	Petroleum Refining, Coking, & Gas Production & Supply	3.9547
17	Textile Industry	3.8828
19	Leather, Furs, Down & Related Products	3.5629
30	Plastic Products	3.3184
18	Garments & Other Fiber Products	3.2795
20	Timber Processing, Bamboo, Cane, Palm Fiber & Straw Products	3.2598
22	Papermaking & Paper Products	3.2004
41	Electronics & Telecommunications	3.1740
32	Smelting & Pressing of Ferrous & Nonferrous Metals	2.9839
34	Metal Products	2.9658
26	Raw Chemical Materials & Chemical Products	2.9256
37	Transportation Equipment	2.8912
21	Furniture Manufacturing	2.7851
40	Electric Equipment & Machinery	2.7743
24	Cultural, Educational & Sports Goods	2.6981
28	Chemical Fibers	2.6257
23	Printing & Medium Reproduction	2.5793
35	Ordinary Machinery & Special Purposes Equipment	2.5335
29	Rubber Products	2.4215
27	Medical & Pharmaceutical Products	2.2479
44	Electricity Power, Steam & Hot Water Production & Supply	2.0085
31	Nonmetal Mineral Products	1.9918
8	Ferrous & Nonferrous Metals Mining & Processing	1.9464
15	Beverage Production	1.8854
42	Instruments, Meters, Cultural & Clerical Machinery	1.8376
6	Coal Mining & Processing	1.7866
46	Tap Water Production & Supply	1.7219
10	Nonmetal Minerals Mining & Processing	1.3631
7	Petroleum & Natural Gas Extraction	1.2121
16	Tobacco Processing	0.7984
12	Logging & Transport of Timber & Bamboo	0.7494
All		2.6958

*Note: Following Kim's measurement, the Raw-Material Intensity is defined as the ratio of (Output-Value added) to Value added. Data on both output and value-added are obtained from China Statistical Yearbook on Industrial Economy.*

Figure 1: Time Trend of Cross-Industry Average Hoover Coefficient of Localization

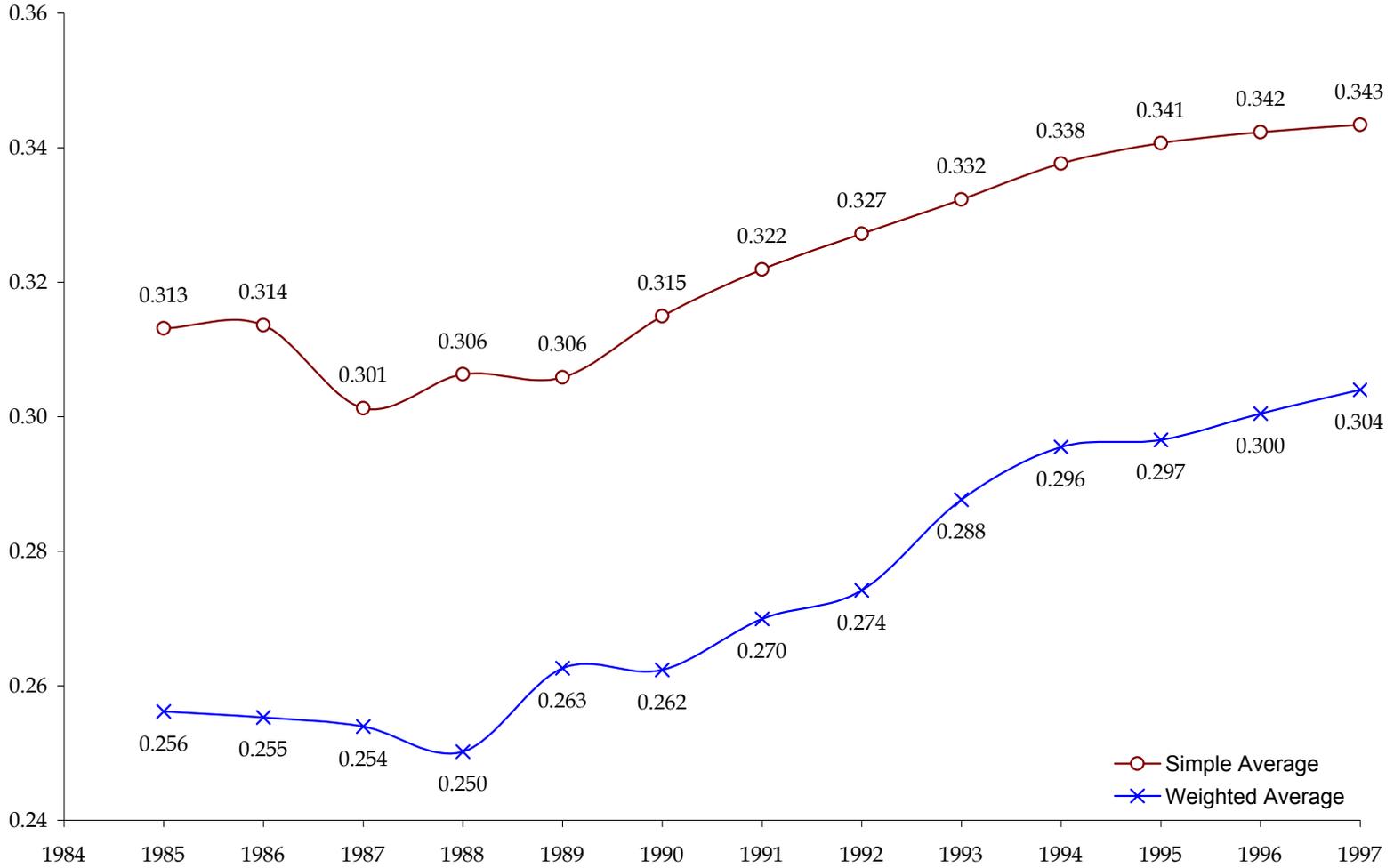
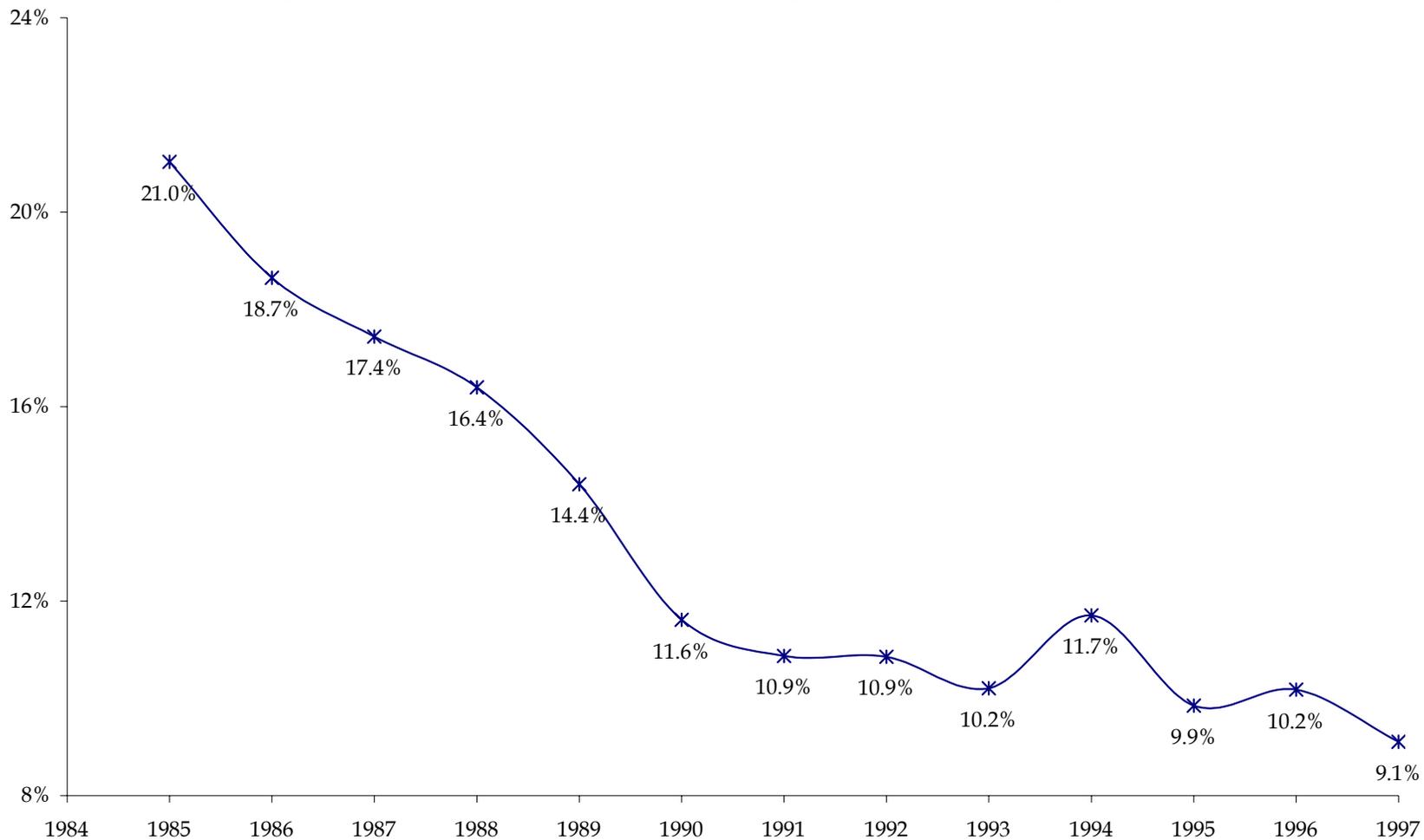
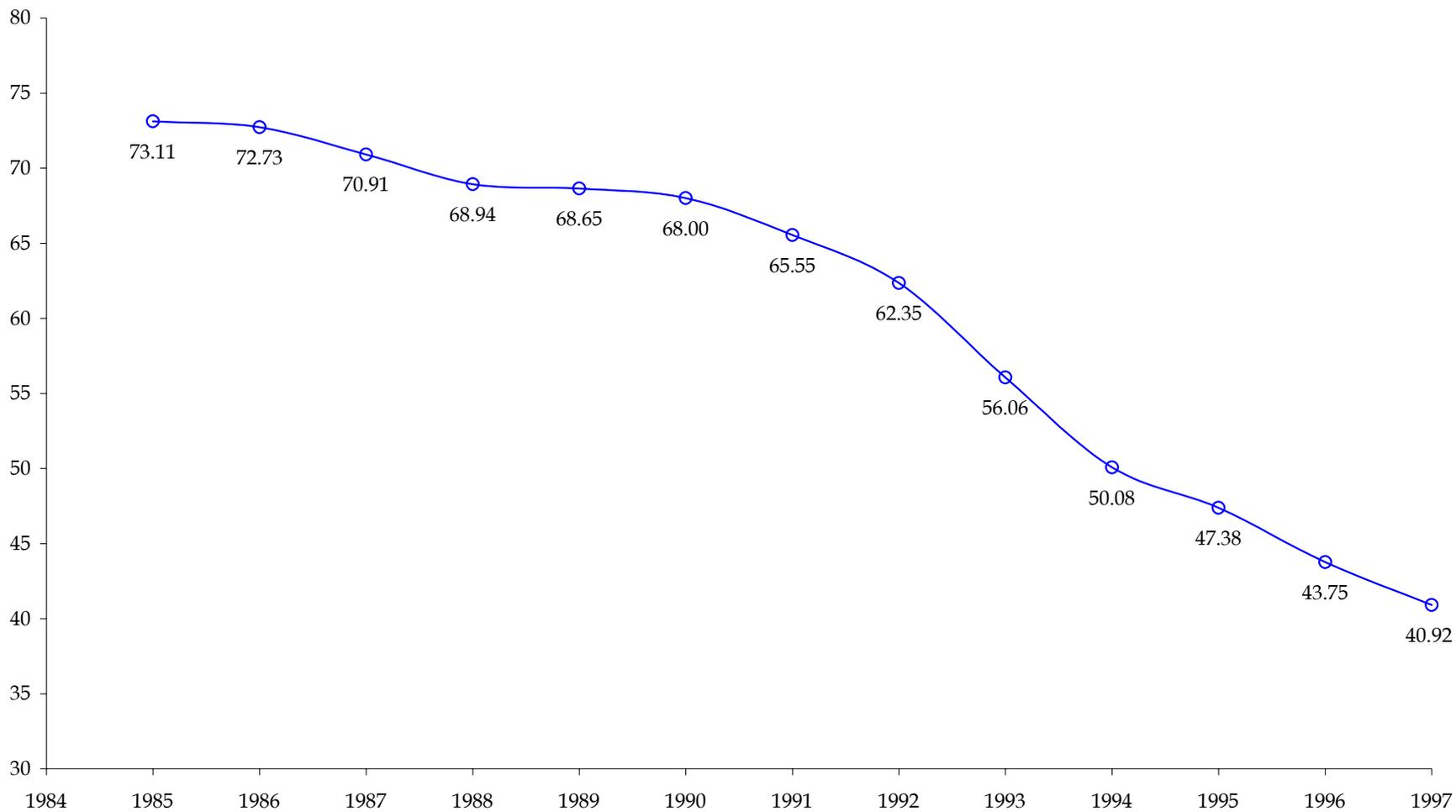


Figure 2: Time Trend of Cross-Industry Average Tax-Plus-Profit Margin (%)



Note: weighted averages are used where the weights are based on sales.

Figure 3: Time Trend of Cross-Industry Average Shares of SOEs (%)



Note: Weighted averages are used where the weights are based on output.

**Table 2: Mean Value and Rank of Main Variables**

Industry Code	Industry name	HOOVER	Rank	SCALE	Rank	ET	Rank	TPM	Rank	SSOE	Rank
6	Coal Mining & Processing	0.579	3	0.0024	15	2.73	32	2.0%	32	79.34	7
7	Petroleum & Natural Gas Extraction	0.744	2	0.5103	1	8.91	7	16.1%	6	98.24	1
8	Ferrous & Nonferrous Metals Mining & Processing	0.578	4	0.0018	19	4.94	18	14.1%	12	66.58	12
10	Nonmetal Minerals Mining & Processing	0.319	12	0.0007	30	4.19	25	15.4%	9	36.13	25
12	Logging & Transport of Timber & Bamboo	0.847	1	0.0040	10	4.38	24	15.0%	11	96.55	3
13	Food Processing & Production	0.199	24	0.0019	18	4.73	19	5.9%	31	68.23	11
15	Beverage Production	0.238	20	0.0018	22	6.31	13	18.8%	5	65.51	14
16	Tobacco Processing	0.537	5	0.0687	2	4.69	20	56.1%	1	97.37	2
17	Textile Industry	0.283	15	0.0047	8	3.22	30	7.4%	29	50.85	21
18	Garments & Other Fiber Products	0.277	17	0.0016	23	2.85	31	7.6%	28	11.73	31
19	Leather, Furs, Down & Related Products	0.279	16	0.0018	20	3.33	29	6.0%	30	22.06	29
20	Timber Processing, Bamboo, Cane, Palm Fiber & Straw Products	0.348	10	0.0006	31	4.58	22	7.7%	27	35.95	26
21	Furniture Manufacturing	0.219	22	0.0006	32	3.89	26	8.1%	26	10.95	32
22	Papermaking & Paper Products	0.197	25	0.0018	21	4.50	23	11.0%	20	51.00	19
23	Printing & Medium Reproduction	0.183	27	0.0008	29	3.65	28	13.8%	14	55.30	18
24	Cultural, Educational & Sports Goods	0.476	6	0.0014	24	3.87	27	12.1%	18	26.84	27
25	Petroleum Refining, Coking, & Gas Production & Supply	0.391	8	0.0148	4	10.18	4	19.9%	3	92.47	4
26	Raw Chemical Materials & Chemical Products	0.154	31	0.0036	13	7.96	10	13.1%	15	69.38	9
27	Medical & Pharmaceutical Products	0.159	29	0.0055	7	10.85	3	13.9%	13	69.10	10
28	Chemical Fibers	0.387	9	0.0184	3	6.90	12	15.4%	8	66.29	13
29	Rubber Products	0.192	26	0.0036	12	5.07	17	15.0%	10	56.91	17
30	Plastic Products	0.264	18	0.0014	25	5.96	15	8.8%	25	19.45	30
31	Nonmetal Mineral Products	0.161	28	0.0011	27	4.68	21	12.3%	16	44.47	23
32	Smelting & Pressing of Ferrous & Nonferrous Metals	0.332	11	0.0143	5	6.03	14	15.5%	7	79.05	8
34	Metal Products	0.146	32	0.0012	26	5.96	16	10.0%	24	22.28	28
35	Ordinary Machinery & Special Purposes Equipment	0.154	30	0.0022	16	8.18	9	10.8%	22	58.78	16
37	Transportation Equipment	0.302	14	0.0042	9	9.66	6	10.8%	21	64.68	15
40	Electric Equipment & Machinery	0.229	21	0.0032	14	8.52	8	11.5%	19	39.34	24
41	Electronics & Telecommunications	0.414	7	0.0068	6	13.99	1	10.2%	23	50.72	22
42	Instruments, Meters, Cultural & Clerical Machinery	0.308	13	0.0020	17	12.63	2	12.3%	17	50.88	20
44	Electricity Power, Steam & Hot Water Production & Supply	0.207	23	0.0039	11	9.93	5	22.7%	2	90.02	5
46	Tap Water Production & Supply	0.252	19	0.0008	28	7.21	11	19.8%	4	89.58	6
Mean		0.324		0.0216		6.39		13.7%		57.38	
Std. Dev.		0.174		0.0900		2.93		9.0%		25.68	

Unit of measurement: RMB 100,000,000 per firm for SCALE, number per 100 employees for ET, and percentages for TPM and SSOE.

**Table 3: Estimation of Equation (1)**

Dependent variable: Hoover coefficient of localization

Explanatory Variables	1	2	3
Hoover			
Lag 1	0.5618 *** (0.0981)	0.5519 *** (0.0935)	0.6388 *** (0.1)
Lag 2	0.1124 * (0.0678)	0.1356 * (0.0711)	0.1366 * (0.0773)
TPM			
Lag 1	-0.0503 ** (0.024)		
Lag 2		-0.0861 *** (0.0304)	
Lag 3			-0.0068 (0.033)
SSOE	-0.0001 (0.0003)	-0.0002 (0.0003)	-0.0005 ** (0.0002)
SCALE	0.1007 *** (0.0227)	0.0986 *** (0.0159)	0.1129 *** (0.0215)
ET	0.0004 (0.0007)	0.0005 (0.0006)	0.0007 (0.0005)
CON	0.0010 (0.0012)	0.0002 (0.0012)	-0.0008 (0.0009)
# of obs.	306	306	280
Significance test			
Wald Chi <sup>2</sup>	310.37	267.29	508.19

Note: \*\*\*, \*\*, and \* denote 1%, 5%, and 10% significant level, respectively.  
The numbers in the parentheses are the standard errors.

**Table 4: Estimation of Equation (2)**

Dependent variable: Hoover coefficient of localization

Explanatory Variables	1	2	3
Hoover			
Lag 1	0.6413 *** (0.0995)	0.5586 *** (0.0963)	0.6377 *** (0.098)
Lag 2	0.1495 * (0.0776)	0.1286 * (0.0707)	0.1423 * (0.0779)
TPM			
Lag 1	-0.0032 (0.0249)	-0.0137 (0.0277)	
Lag 2	-0.0808 ** (0.0385)	-0.0591 * (0.0314)	-0.0856 ** (0.0331)
Lag 3	0.0428 (0.0391)		0.0477 (0.0376)
SSOE	-0.0004 ** (0.0002)	-0.0001 (0.0003)	-0.0005 ** (0.0002)
SCALE	0.0700 ** (0.0319)	0.1049 *** (0.0232)	0.0651 ** (0.0296)
ET	0.0005 (0.0006)	0.0005 (0.0006)	0.0005 (0.0006)
CON	-0.0010 (0.0009)	0.0005 (0.0012)	-0.0009 (0.0009)
# of obs.	280	306	280
Significance test			
Wald Chi <sup>2</sup>	554.68	331.51	446.13

Note: \*\*\*, \*\*, and \* denote 1%, 5%, and 10% significant level, respectively.  
The numbers in the parentheses are the standard errors.

### Data Appendix: Summary of data and measurement

	1985	1986	1987	1988-1989	1990-1994	1995	1996	1997
HOOVER coefficient of localization (output based)								
Output by industry by region	2	2	2	1	1	3	5	1
HOOVER coefficient of localization (employment based)								
Employment by industry by region	4	NA	NA	1	1	NA	5	1
SCALE (total output/number of firms/deflator)								
Total output by industry	1	1	1	1	1	3	1	1
Number of firms by industry	1	1	1	1	1	3	1	1
Price deflator	2	2	2	2	2	2	2	2
TPM (tax-plus-profit margin=(tax+profit)/sales)								
Tax & profit by industry	1	1	1	1	1	3	1	1
Sales by industry	1	1	1	1	1	3	1	1
SSOE (share of output produced by state-owned enterprises)								
Total output by industry	6	6	6	6	6	6	6	6
SOE output by industry	6	6	6	6	6	6	6	6
ET (share of engineers and technicians in employment)								
Total employment of large & medium enterprises by industry	4	NA	5	5	7	7	7	7
Number of engineers and technicians in large & medium enterprises by industry	4	NA	5	5	7	7	7	7

1. *China Statistical Yearbook on Industrial Economy*
2. *China Statistical Yearbook*
3. *China Industrial Census 1995*
4. *China Industrial Census 1985*

5. *China Statistics Bureau*
6. *Development Research Center, The State Council of PR China*
7. *China Statistical Yearbook on Science and Technology*

## DAVIDSON INSTITUTE WORKING PAPER SERIES - Most Recent Papers

The entire Working Paper Series may be downloaded free of charge at: [www.wdi.bus.umich.edu](http://www.wdi.bus.umich.edu)

CURRENT AS OF 5/1/03

<b>Publication</b>	<b>Authors</b>	<b>Date</b>
No. 565: Local Protectionism and Regional Specialization: Evidence from China's Industries	Chong-En Bai, Yingjuan Du, Zhigang Tao, Sarah Y. Tong	May 2003
No. 564: Corporate Governance and Market Valuation in China	Chong-En Bai, Qiao Liu, Joe Lu, Frank M. Song, and Junxi Zhang	May 2003
No. 563: Revenue Sharing and Control Rights in Team Production: Theories and Evidence From Joint Ventures	Chong-En Bai, Zhigang Tao, and Changqi Wu	May 2003
No. 562: Financial Dependence, Stock Market Liberalizations and Growth	Nandini Gupta and Kathy Yuan	May 2003
No. 561: Growth and Regional Inequality in China During the Reform Era	Derek Jones, Cheng Li and Owen	May 2003
No. 560: Choice of Ownership Structure and Firm Performance: Evidence from Estonia	Derek Jones, Panu Kalmi, Niels Mygind	May 2003
No. 559: Explaining Postcommunist Economic Performance	Lawrence P. King	May 2003
No. 558: Tax Structure and the FDI: The Deterrent Effects of Complexity and Uncertainty	Kelly Edmiston, Shannon Mudd and Neven Valev	Apr. 2003
No. 557: Provincial Protectionism	Konstantin Sonin	Apr. 2003
No. 556: Nominal and Real Convergence in Estonia: The Balassa-Samuelson (dis)connection	Balázs Égert	Apr. 2003
No. 555: Banks-Firms Nexus under the Currency Board: Empirical Evidence from Bulgaria	Nikolay Nenovsky, Evgeni Peev and Todor Yalamov	Apr. 2003
No. 554: To Steal or Not to Steal: Firm Attributes, Legal Environment, and Valuation	Art Durnev and E. Han Kim	Apr. 2003
No. 553: Corporate Stability and Economic Growth	Kathy S. He, Randall Morck and Bernard Yeung	Apr. 2003
No. 552: So Many Rocket Scientists, So Few Marketing Clerks: Occupational Mobility in Times of Rapid Technological Change	Nauro F. Campos and Aurelijus Dabušinskas	Mar. 2003
No. 551: Determinants of Interregional Mobility in Russia: Evidence from Panel Data	Yuri Andrienko and Sergei Guriev	Feb. 2003
No. 550: Gross Job Flows in Ukraine: Size, Ownership and Trade Effects	Jozef Konings, Olga Kupets and Hartmut Lehmann	Mar. 2003
No. 549: Technology Transfer through FDI in Top-10 Transition Countries: How Important are Direct Effects, Horizontal and Vertical Spillovers?	Jože P. Damijan, Mark Knell, Boris Majcen and Matija Rojec	Feb. 2003
No. 548: Does Foreign Direct Investment Increase the Productivity of Domestic Firms? In Search of Spillovers through Backward Linkages	Beata K. Smarzynska	Mar. 2003
No. 547: Re-employment Probabilities and Wage Offer Function for Russian Labor Market	Natalia V. Smirnova	Feb. 2003
No. 546: Democratization's Risk Premium: Partisan and Opportunistic Political Business Cycle Effects on Sovereign Ratings in Developing Countries	Steven Block, Burkhard N. Schrage and Paul M. Vaaler	Feb. 2003
No. 545: Structural Reforms and Competitiveness: Will Europe Overtake America?	Jan Svejnar	Feb. 2003
No. 544: Why the Rich May Favor Poor Protection of Property Rights	Konstantin Sonin	Dec. 2002
No. 543: Reinvested Earnings Bias, The "Five Percent" Rule and the Interpretation of the Balance of Payments – With an Application to Transition Economies	Josef C. Brada and Vladimír Tomšík	Feb. 2003
No. 542: The Impact of Ownership Reform in Chinese Industry, 1995-2001	Gary H. Jefferson, Su Jian, Jiang Yuan and Yu Xinhua	Feb. 2003
No. 541: Defensive and Strategic Restructuring of Firms during the Transition to a Market Economy	Domadenik, Janez Prašnikar and Jan Svejnar	Feb. 2003
No. 540: Tenuous Financial Stability	Neven T. Valev and John A. Carlson	Feb. 2003