DEPARTMENT OF ECONOMICS

COLLEGE OF BUSINESS AND ECONOMICS

UNIVERSITY OF CANTERBURY

CHRISTCHURCH, NEW ZEALAND

Industrial Agglomeration, Geographic Innovation and Total Factor Productivity: The Case of Taiwan

by Chia-Lin Chang and Les Oxley

WORKING PAPER

No: 14/2008

Department of Economics
College of Business and Economics
University of Canterbury
Private Bag 4800, Christchurch
New Zealand

WORKING PAPER No. 14/2008

Industrial Agglomeration, R&D and Total Factor Productivity:

The Case of Taiwan

by Chia-Lin Chang [1†], Les Oxley [2]

July 2008

Abstract

The paper analyses the impact of geographic innovation on Total Factor Productivity (TFP) in Taiwan. Using 242 four-digit standard industrial classification (SIC) industries in Taiwan in 2001, we compute TFP by estimating Translog production functions with K, L, E and M inputs, and measure the geographic innovative activity using both Krugman's Gini coefficients and the location Herfindahl index. We also consider the geographic innovation variable as an endogenous variable and use 2SLS to obtain a consistent, albeit inefficient, estimator. The empirical results show a significantly positive effect of geographic innovation, as well as R&D expenditure, on TFP. These results are robust for the Gini coefficients and location Herfindahl index, when industry characteristics and heteroskedasticity are controlled. Moreover, according to the endogeneity of geographic innovation, the Hausman test shows that the geographic innovation variable should be treated as endogenous, which supports the modern theory of industrial clustering about innovation spillovers within clusters.

Keywords: Industry agglomeration; Geographic innovation; Total factor productivity; Cluster; Research and Development

JEL classifications: O32, O33, L60, R12

[1] Department of Applied Economics, National Chung Hsing University, 250 Kuo Kuang Road, Taichung 402, Taiwan.

[2] Department of Economics, University of Canterbury, Private Bag 4800, Christchurch 8140, New Zealand.

[†] Author for correspondence: *Email: changchialin@nchu.edu.tw*

1. Introduction

Porter, [24] popularised the idea that agglomeration (or clusters) affects industrial performance and global competition, however, it is still widely debated whether geographical location affects national competitiveness. Classic theories argue that industrial agglomeration provides firms with easy access to critical resources, lower transport costs, access to customers, and a specialized and skilled labour pool [21, 22].

Using a two-region model, and assuming immobility of farmers and free mobility of manufacturing workers, [20] concluded that agglomeration tends to emerge when economies of scale create more profit than the offsetting transportation costs, or when transportation cost alone is sufficiently low,. Following Krugman, [10] argued that resources which are critical to a firm or an industry should not be limited to natural resources, but should include all resources, such as human capital, when they are not perfectly mobile. They even suggested that all industries are at least slightly agglomerated, and attributed the agglomeration to cost advantages.

More recent studies for example, [18, 1, 2, 12, 13, 14], and [3], have emphasized that spatially-mediated knowledge spillovers are likely to play a crucial, if not dominant, role in industrial agglomeration (or clusters). As [3] observed, if the ability to receive knowledge spillovers is influenced by distance from the knowledge source, then geographic concentration should be controlled, especially in industries. From a managerial perspective, spillovers within clusters are normally generated by informal exchange of information, such as labour turnover, industrial events, or even from using the same suppliers [24, 27, 29]. [9] and [26] suggested that firms within the same cluster may also benefit from joint-bidding, scaled contract tender, or joint marketing. In addition, firms may also benefit from accessibility to public goods, such as research resources and infrastructure. [26] also stated that, due to easy access to skilled labour as well as diverse suppliers and input, clusters have become the main source of innovation. Thus, in explaining why industry technology varies across industries, we also need to explain, and control for the geographic concentration of innovation.

Although previous evidence supports the idea that geographic concentration is important for spillovers , research in the industrial organization literature linking the underlying degree of concentration of economic activity within a geographic context to industrial performance, is still rare. Most research typically explains the agglomeration economy across different industries [15, 16, 11, 28]. Few have demonstrated the degree of geographic concentration as generating greater industrial performance. According to [24], while traditional thinking on innovation focuses on *internal* factors such as technology, the *external* factors are usually ignored,. If innovation arises within the same cluster, then one might expect a positive impact of concentration of innovation on industrial productivity. Consequently, the geographic concentration of innovation within a cluster can be affected by the geographic concentration of production. As the impact of geographic innovation on productivity may vary hugely across industries, it may be positive in some high-tech industries, and negative in others.

The paper examines whether geographical concentration of innovation is a spur to industrial productivity and establishes the following outcomes. Firstly, industry agglomeration augments knowledge spillovers within the cluster, and thereby creates greater opportunity for innovation. Secondly, the agglomeration of innovation may lead to an increase in industry Total Factor Productivity (hereafter, *TFP*). We examine the effect of agglomeration of innovation on productivity by using the four-digit standard industrial classification (hereafter, SIC) manufacturing industries for Taiwan in 2001.

The remainder of the paper is organized as follows: Section 2 describes industry agglomeration in Taiwan. Section 3 presents the theoretical and empirical framework, while the data and variable description are described in Section 4. Estimation results are presented in Section 5, and Section 6 concludes.

2. Industrial agglomeration in Taiwan

In Taiwan, industry agglomeration can be directly linked to public policy, where targets actively promote industrial and technological upgrading. Overall, we can identify three types of industrial agglomeration: (1) Industrial zones; (2) Export processing zones; and (3) Science-based industrial parks.

2.1 Industrial zones

Since the 1970s, the concept of an ‘industrial zone’ has been directly linked to Taiwan's industrial policy. More recently, the *Industrial Development Bureau* has focused on incentives aimed at encouraging investment which might lead to industry clustering, which in turn can promote local economic development and an environment that emphasizes high added-value production. This will hopefully lead to the emergence of strategic industries, i.e., those that are expected to benefit economic development in a significant way.

Figure 1 shows the industrial zones in Taiwan. According to the current administrative districts in Taiwan, there are twenty-five “Counties or Cities”, which include Taiwan Kinmen County and Lienchiang County, which comprise a small archipelago of islands administered by Taiwan. Figure 1 shows only twenty-three of these “Counties or Cities”. Each “County or City” can be divided into smaller geographic districts, including County-Adm. City, Jhen, Siang, or District. Table 1 shows the number of administrative districts across Counties and Cities in Taiwan.

INSERT FIGURE 1 ABOUT HERE

INSERT TABLE 1 ABOUT HERE

Table 2 provides an overview of the area, number of industrial parks, and number of plants for each geographic area (“City or County”). As shown in Tables 1 and 2, “County or City” covers seven cities and sixteen counties (see also Figure 1), and each “City or County” is grouped from a number of County-Adm. City, Jhen, Siang, or District. In the paper, as shown in Table 1, the total number of geographic areas is 359.

INSERT TABLE 2 ABOUT HERE

Table 3 provides an overview of the number of plants by two-digit SIC industry and 23 geographic “City or County” areas. The left-hand column presents two-digit SIC industries (for a list of the industry names see the Appendix), and the top row denotes “City or County”. The figures in the table are the shares of the number of plants in each 2-digit industry by geographic area. As can be seen from Table 3, the phenomenon of industry agglomeration seems to exist in all the 2-digit SIC industries. Taipei county has the largest share of industry clusters for example, industry (26), “Audio & Video products”; industry (27),“Electronic parts and components” and industry (28) “Electric machinery and parts”. Taichung county accomodates many traditional industries for example industry (12) “Leather & Fur Products”; industry (13) “Wood & Bamboo Products”; industry (14) “Furniture & Fixtures” and industry (25) “Machinery & Equipment industries” . Chunghua county has concentrations in the textile and apparel accessories sectors and transportation equipment. Taoyuan county, which is closest to Taipei county, has agglomerations in audio & video products, electronic parts and components;, basic chemicals and chemical products. Moreover, every 2-digit industry tends to agglomerate in three main areas; Taipei county, Taichung country, and Changhua country.

NSERT TABLE 3 ABOUT HERE

2.2 Export processing zones

Export Processing Zones, hereafter EPZs, were pioneered in the 1960s. The first EPZ was established at Kaohsiung in 1966, with two more created at Nantze and Taichung in 1969. Firms located within EPZs have received assistance via zero-tariff rates on imported inputs thereby improving the cost competitiveness of their exports. Moreover, EPZs create an upstream, mid-stream and downstream industry link between, R&D, order-taking and production, and repackaging, storage, and delivery. This creates cost-efficient and highly competitive industrial clusters. By 2007 other EPZs had been created at Chengkung, Pingtung. Table 4 shows the number of firms in EPZs by 2-digit industry, as well as the land area for each EPZ in 2007.

INSERT TABLE 4 ABOUT HERE

2.3 Science-based parks

Hsinchu Science Park, hereafter *HSP*, is the most famous example of the geographic agglomeration of firms in the high-tech sector. Based on the concept of the industrial zone, Taiwan's science parks are exclusively devoted to high-tech industries and are home to many world-renowned companies. The *HSP* was established in 1980 and includes parts of Hsinchu and Taoyuang counties. It focuses on six main high-tech industries, including Integrated Circuits, PC/Peripherals, Telecommunication, Optoelectronics, Precision Machinery, and Biotechnology. By the end of 2006, the *HSP* had a total of 391 high-tech companies and. Table 5 shows the number of high-tech firms in the HSP from 1983 to 2006.

INSERT TABLE 5 ABOUT HERE

The Southern Taiwan Science Park, hereafter *STSP*, includes Tainan Science Park and Kaohsiung Science Park. The idea to establish the *STSP* was pioneered by the government's economic revitalization project of 1993, while the *STSP* Development Plan received government approval in 1995. The Southern Taiwan Science Park focuses on Ptoelectronics, Telecommunication and Precision Machinery. By the end of 2006 there were 100 high-tech companies in the Park. Table 6 shows the number of high-tech firms in the *STSP* from 1988 to 2006.

INSERT TABLE 6 ABOUT HERE

3. Theoretical framework and empirical analysis

We adopt a three-step empirical strategy to estimate the impact of geographic concentration of innovation on *TFP*. The first step involves estimation of a Translog production function for each two-digit SIC industry. We use the parameter estimates to compute each firm's *TFP* in a given industry, and then average the firm's *TFP* in a given four-digit SIC industry. In the second step, we construct a geographic concentration indicator in a given four-digit SIC industry. In the third step, we construct an empirical model to examine the effect of industry agglomeration on *TFP*.

3.1. Computation of TFP

We obtain our measure of the *TFP* of Taiwanese firms by estimating a production function, and linking sales (our measure of firm output Q) to inputs X. For industry i operating in the manufacturing industry, we write:

(1) $Q_i = F(X_{1i}, X_{2i}, X_{3i}, X_{4i})$

where X_1, X_2, X_3 and X_4, denote, respectively, capital, labor, energy and materials (also generally referred to as K, L, E and M inputs). In order to conduct the empirical analysis we need to specify a functional form for F, which we wish to keep as flexible as possible. Therefore, we assume a Translog specification which is usually considered a reasonable second-order approximation of an arbitrary production function (see, for example, [4], [5], [8], [7]). We rewrite (1) as:

(2) $\ln Q = \beta_0 + \Sigma_j \beta_j . \ln X_{ji} + \frac{1}{2} [\Sigma_j \Sigma_k \delta_{jk.} (\ln X_{ji})(\ln X_{ki})] + \varepsilon_i$

where, ε_{it} a transitional error term.

Under the usual symmetry assumption (that is, $\delta_{jk} = \delta_{kj}$, $\forall\ j, k$), we can also compute input shares, for $k = 1, 2, 3, 4$:

(3) $S_{ki} = \partial ln\ Q_i / \partial ln\ X_{ki} = \hat{\beta}_k + \Sigma_j \hat{\delta}_{kj} . \ln X_{ji}$ with $j = 1, 2, 3, 4$.

Returns to scale are then defined as the sum of input shares over k = 1, 2, 3, 4:

(4) $\text{RTS}_i = \Sigma_k S_{ki}$

Finally, we can compute *TFP* for industry *i* as [19]:

(5) $TFP_i = (RTS_i - 1).\ \Sigma_k (S_{ki}.X_{ki})/RTS_i$

3.2. Measures of geographic concentration of plants

Due to the lack of information on the actual spatial distance (in miles) between the centroids of "County or City" in the data set, in this paper we use the two most popular indicators to measure geographical concentration: the geographic Herfindahl index and the geographic Gini coefficient. Both measures of geographic concentration can be calculated using any geographic unit, or parcel [6]. In our case, the County-Adm. City, Jhen, Siang, or District is the geographic unit. As shown in Table 1, the geographic area totals 359.

For brevity, the lower case for each industry is suppressed in the formula. The location Herfindahl indicator for a given industry is then defined as[1]:

$$GHHI = \sum_{k=1}^{m} \left(\frac{j_k}{\sum_{k=1}^{m} j_k} \right)^2$$

where j_k denotes the number of plants in a given geographic area, for a given industry (k is a certain geographic area), and m is a sufficiently large number of geographic areas (m= 359 in our data set). When a geographic area is located by only one plant in a given industry, the index has a maximum value of 1, (or 10,000, when the market shares are measured in percentage terms). The value declines with increases in an industry that is not geographically concentrated in geographic area m, and increases with rising inequality among any given number of geographic areas.

[1] The Herfindahl indexes can be estimated through the numbers employed (see [23, 17, 25, 6]).

The Geographic Gini coefficient, as first proposed by [20]), can be estimated through numerical integration of the area inside the Lorenz curve in the graph of cumulative employment of jobs, sorted according to decreasing geographic area, for any given industry. In calculating the Geographic Gini coefficient we use the number of plants and follow a measure suggested by [23] and [6].

$$Gnni = 1 + \frac{1}{m} - \frac{2}{m\sum_{k=1}^{m} j_k}\left(\sum_{k=1}^{m} r_k j_k\right)$$

where j_k denotes the number of firms in a given geographic area for a given industry, m is a sufficiently large number of geographic areas, r_k denotes the rank of the number of firms in a geographic area when the geographic area is sorted in decreasing order of the number of plants. The closer the Gini coefficient is to 1, the more geographically concentrated the industry would be; alternatively an industry which is not geographically concentrated would have a coefficient of 0.

3.3 Measures of geographic concentration of innovation

Both measures of geographic concentration can be taken as measures of geographic concentration of innovation. We use $GHHI_{R\&D}$ to represent the location Herfindahl indicator of innovation, and $Gnini_{R\&D}$ to represent the geographic Gini coefficient of innovation. The location Herfindahl index of R&D geographic concentration for a given industry, is given by the formula:

$$GHHI_{\mathrm{R\&D}} = \sum_{k=1}^{m}\left(\frac{R_k}{\sum_{k=1}^{m} R_k}\right)^2$$

where R_k denotes the number of plants which have recorded their own R&D expenditures in a given geographic area for a given industry (k is certain geographic area), and m is a sufficiently large number of geographic areas (m = 359 in our case).

The Geographic *Gini* coefficient of innovation for a given industry is given by the formula:

$$Gnni_{\mathrm{R\&D}} = 1 + \frac{1}{\mathrm{m}} - \frac{2}{\mathrm{m}\sum_{\mathrm{k=1}}^{\mathrm{m}} \mathrm{R_k}} \left(\sum_{\mathrm{k=1}}^{\mathrm{m}} \mathrm{r_k R_k} \right)$$

where R_k denotes the number of plants with a record of R&D expenditure in a given geographic area for a given industry, m is a sufficiently large number of geographic areas, r_k denotes the rank of the number of plants with recorded R&D expenditures in a geographic area, when the geographic areas are sorted in decreasing order of numbers of plants.

3.4 Estimation procedure

In order to investigate the impact of geographic innovation on industry total factor productivity, we consider the following linear model which is a function of industry R&D input and geographic concentration of innovation such that:

(6) $$TFP_{\mathrm{i}} = \theta_0 + \theta_1 Gini_{\mathrm{R\&D}} + \theta_2 RD_{\mathrm{i}} + e_{\mathrm{i}}$$

Where RD_{i} is R&D input in industry i, e_{i} is an error term

Given the discussion in Section 1 above, innovation activity could grow more rapidly within clusters [3]. Therefore, we should consider the variable of industry agglomeration of innovation, $Gini_{R\&D}$, to be endogenous and use two stage least squares (hereafter, 2SLS) to obtain consistent, though inefficient, estimators. Thus, we first estimate equation (7)

(7) $$Gini_{\mathrm{R\&Di}} = \gamma_1 + \gamma_2 Gini_{\mathrm{i}} + \gamma_3 RD_{\mathrm{i}} + \varepsilon_{\mathrm{i}}$$

where $Gini_{\mathrm{i}}$ is the geographic Gini coefficients in industry i and ε_{i} is an error term.

We can obtain the fitted value, $\hat{Gini}_{\mathrm{R\&Di}}$, from the reduced form equation (7), and use them as an explanatory variable in equation (6) replacing $Gini_{R\&Di}$, such that

$$TFP_{\mathrm{i}} = \theta_0 + \theta_1 RD_{\mathrm{i}} + \theta_2 \hat{Gini}_{\mathrm{R\&Di}} + e_{\mathrm{i}} \tag{8}$$

The data arise from a Census such that the population is large and involves all Taiwan manufacturing industries. Therefore, we also account for potential heteroskedasticity in the data by 'robustifying' standard errors using the White correction. The robust standard errors are typically slightly larger than their asymptotic counterparts. The models also include an industry-specific effect using a set of three-digit SIC industry dummy variables. The resulting coefficient estimates are 'proper' 2SLS estimates, but the reported standard errors are not correct in the two-step regression process as they are based on an improper covariance matrix of the error term σ^2. Therefore, we use the econometric software package, Stata 10 to compute the IV estimates and their correct standard errors. The data used in the paper are described in the following section.

4 The DGBAS data

We use data provided by the Directorate General of Budget, Accounting and Statistics (DGBAS) of Taiwan's Executive Yuan. The DGBAS data are from a large survey conducted every five years by the DGBAS. In this paper, the data cover 153,923 plants for all manufacturing industries in 2001. It should be noted that the Tobacco industry has only 8 plants, and hence was deleted from the data set. As a result, the sample has 153,915 plants.

Table 7 provides a classification of the 153,915 observations by two-digit SIC manufacturing industry. Our empirical model will be based at the industry level, therefore, we aggregate or average the original observations in a given four-digit SIC for each variable (as described in Section 3).

INSERT TABLE 7 ABOUT HERE

The DGBAS data also provides information on plants' sales, net value of fixed assets in operation at the end of the year, total sum of gross wages, number of employees, energy expenditures, total expenditures on raw materials, and R&D expenditures. This information is used to construct the dependent variable, total factor productivity *(TFP).* The DGBAS census data also provide detailed geographic information on plants' city codes which allows us to measure both geographical concentration indicators..

Finally, the DGBAS data allows us to define innovating plants on the basis of their innovation expenditures. In our paper, we define a plant that has reported R&D expenditures as an "innovating plant". However, the proportion of innovating plants remains fairly small in every 2-digit industry, except in high-tech industries such as industry (26) "Audio & video products" and industry (27) "Electronic parts & components" (see also Table 7).

Table 8 provides summary statistics for all the explanatory variables, except for the control variables.

INSERT TABLE 8 ABOUT HERE

5. Results

5.1. Estimation of TFP values

As explained in Section 3.1, we estimate a Translog production function with *K*, *L*, *E* and *M* inputs, and use the production function estimates to compute *RTS* and *TFP*. The estimated values of *RTS* and *TFP* for each industry are given in Table 9. The table shows high *TFP* values in modernising traditional industries, such as industry (10) "Textile Mill Products", as well as reasonably high values of *TFP* in high-tech industries (27) "Electronic parts and components" and industry (30) "Precision Instruments". If we choose to define *TFP* as the part of productivity which is not explained by the conventional *K, L, E and M* inputs, then these results are sensible:

TFP should be high in traditional industries that are upgrading their technological levels and in high-tech industries.

A second important result observed in Table 9 is that in every 2-digit industry *RTS* is close to one, which is consistent with the classical idea of a constant returns-to-scale technology. Therefore, assuming a production function with constant returns to scale in every industry would in the context , be a reasonable approximation.

INSERT TABLE 9 ABOUT HERE

5.2. Estimation of the effect geographical concentration of R&D on TFP

We use the econometric software package Stata 10 to compute the Instrumental Variable estimates and their standard errors. For four-digit SIC cross sectional data, we also present the robust instrumental variables standard errors by using White's heteroskedasticity correction such that the overall Wald chi-squared test is also based on the robust estimators. In virtually all cases, the asymptotic standard errors are smaller than their robust counterparts. Each model also included a set of 3-digit SIC dummy variables.

The estimation results are presented as Tables 10 and 11. Table 10 provides the results of the reduced form equation, or first stage regression, and Table 11 the results of two stage least squares (or instrumental variables). The second column of Table 10 presents the results of estimation when *Gini* is the indicator of geographic concentration, and the right-hand column the results when the geographic Herfindhal indicator is chosen. For brevity, we do not present the estimates for the dummy variables in Tables 10 and 11.

INSERT TABLE 10 ABOUT HERE

Firstly, in the *TFP* equation, the variable *Gini* is not included because *Gini* affects $Gini_{R\&D}$, but does not affect *TFP*. In order to use 2SLS to estimate the *TFP* equation in the reduced form for $Gini_{R\&D}$, the variable *Gini* should be significant for 2SLS to be useful. Table 10 reveals a statistically significant correlation between the Gini coefficient and the Gini coefficient of innovation, and the results from the *GHHI* indicator are consistent with *Gini*. Therefore, we can rely on the 2SLS estimates for the *TFP* equation.

Table 11 presents the regression results using 2SLS. In Table 11, the $Gini_{R\&D}$ variable has a significant and positive effect. The coefficient of the R&D variable is also significantly positive, and R&D is more significant than geographic innovation. Similarly, the results with the *GHHI* indicator also strongly support our hypothesis. These results suggest that higher geographic innovation can positively influence *TFP*.

INSERT TABLE 11 ABOUT HERE

Finally, we use the Hausman procedure to test for the endogeneity of the $\mathrm{Gini}_{\mathrm{R\&D}}$ variable. The null hypothesis for the Hausman test is that there is no correlation between the *Gini* (*GHHI*) variable and the error tem. In other words, if the null hypothesis is not rejected, the *Gini* (*GHHI*) variable is exogenous. We first estimate the reduced form for $Gini_{R\&D}$ (c.f. equation (7) in Section 3) by OLS, obtain the residuals, $\hat{\varepsilon}$, include $\hat{\varepsilon}$ as an explanatory variable in equation (6), and then estimate the auxiliary regression by OLS.

Table 12 presents the result of the Hausman test. We focus on the key variable $\hat{\varepsilon}$ which has a significant effect at the 1% level for the *Gini* indicator, and the 10% level for the *GHHI* indicator. These results suggest that $Gini_{R\&D}$ ($GHHI_{R\&D}$) are strongly correlated with the residuals, so there is strong evidence to suggest that the geographic innovation variable, $Gini_{R\&D}$ ($GHHI_{R\&D}$), should be treated as endogenous.

INSERT TABLE 12 ABOUT HERE

6. Conclusions

The objectives of this paper were to examine the effects of geographic innovation on total factor productivity (*TFP*) at the industry level in Taiwan. In order to do so, we used a number of 242 4-digit SIC industries in 2001 and aggregated from 153,915 manufacturing plants in Taiwan. We computed *TFP* at the 4-digit SIC level by estimating a Translog production function with conventional *K, L, E* and *M* inputs.

To measure the extent to which manufacturing in specific industries is concentrated geographically and the extent to which innovative activity tends to cluster spatially, we used Krugman's Gini coefficients and the location Herfindahl indicator for the geographic concentration of innovative activity and for the location of manufacturing.

Based on the modern theory of industrial clustering which emphasizes that knowledge spillovers could be transferred more easily within clusters [3], we considered the geographic innovation variable, $Gini_{R\&D}$ ($GHHI_{R\&D}$) to be endogenous, and used two stage least squares (2SLS) to investigate the effects of geographic innovation on *TFP*.

The results showed a significantly positive effect of geographic innovation on *TFP*. This result was quite robust across both Krugman's geographic Gini indicator and geographic Herfindahl index, when industry characteristics and heteroskedasticity were controlled. Moreover, the endogeneity of the geographic concentration of innovations has been assessed using the Hausman test, and the empirical results showed strong support for treating the $Gini_{R\&D}$ variable as endogenous.

References

[1] Z. J. Acs, , D. Audretsch, M. P. Feldman, Real effects of academic research: comment, Ameican Economic, 82 (1992) 363-367.

[2] Z. J. Acs, D. Audretsch, M. P. Feldman, R&D spillovers and recipient firm size, Review of Economics and Statistics, 76 (1994) 336-40.

[3] D. B. Audretsch, M. P. Feldman, R&D Spillovers and the Geography of Innovation and Production, American Economic Review, 86 (1996) 630-640.

[4] R. Beason, D.E. Weinstein, Growth, economies of scale, and targeting in Japan (1995-1990), Review of Economics and Statistics 78 (1996) 286-295.

[5] E. Berndt, L. Christensen, The translog function and the substitution of equipment, structures, and labor in US manufacturing, 1929-1968, Journal of Econometrics, 1 (1973) 81-114.

[6] M. L.Busch, E. Reinhardt, Industrial location and protection: The political and economic geography of U.S. nontariff barriers, American Journal of Political Science, 43 (1999) 1028-1050.

[7] C.-L. Chang, S. Robin, Public policy, innovation and total factor productivity: An application to Taiwan's manufacturing industry", to appear in Mathematics and Computers in Simulation (2008).

[8] M.W.L. Chan, D.C. Mountain, Economies of scale and the Törnqvist discrete measure of productivity growth, Review of Economics and Statistics, 70 (1983) 663-667.

[9] P. Cooke, Cluster as key determinants of economic growth: The example of biotechnology, cluster policies –cluster development?, Nordregio Report 2001:2.

[10] G. Ellison, E. Glaeser, The geographic concentration of industry: Does natural advantage explain agglomeration?, The American Economic Review, 89 (1999) 311-316.

[11] C. C. Fan, A. J. Scott, Industrial agglomeration and development: A survey of spatial economic issue in East Asia and a statistical analysis of Chinese regions, Economic Geography, 79 (2003) 295-319.

[12] M. P. Feldman, Knowledge complementarity and innovation, Small Business Economics, 6 (1994) 363-372.

[13] M. P. Feldman, The new economics of innovation, spillovers and agglomeration: A review of empirical studies, The Economics of Innovation and New Technology, 8 (1999) 5-25.

[14] M. P. Feldman, Location and innovation: The new economic geography of innovation, spillovers, and agglomeration, in G. Clark, M. Feldman and M. Gertler, (Eds.), Oxford Handbook of Economic Geography, Oxford University Press, 2000, pp. 373-394.

[15] M. Fledman, Y. Schreduer, Initial advantage: The origins of the geographic concentration of the pharmaceutical industry in the Mod-Atlantic region, Industrial and Cooperate Change, 5 (1996) 839-862.

[16] P. Guerrieri, C. Pietrobelli, (2004), Industrial districts' evolution and technological regimes: Italy and Taiwan, Technovation, 24 (2004) 899-914.

[17] K. B. Grier, M. C. Munger, B. E. Roberts, The determinants of industry political activity, 1978-1986, American Political Science Review 88 (1994) 911-926.

[18] A. B. Jaffe, M. Trajtenberg, R. Henderson, Geographic localization of knowledge spillovers as evidenced by patent, Quarterly Journal of Economics, 63(1993) 577-98.

[19] S.C. Kumbhakar, C.A.K Lovell (2000), *Stochastic Frontier Analysis*, Cambridge University Press, Cambridge, 2000.

[20] P. Krugman, Geography and Trade, Cambridge: MIT press, 1991.

[21] A. Marshall, The Principles of Economics, London: Macmillan Press, 1890.

[22] R. Oakey, High-Technology Industries and Agglomeration Economies, Silicon Landscapes, Chap. 7, Boston, Allen & Unwin, 1985.

[23] D. W. Pearce, The MIT Dictionary of Modern Economics, fourth ed., MOT Press, Cambridge, 1992.

[24] M. Porter, The Competitiveness of Nations, Macmillan Press, Cambridge, 1990.

[25] L.M. Salamon, J. J. Siegfried, Economic power and political influence: The impact of industry structure on public policy, American Political Science Review, 71 (1977) 1026-1043.

[26] T. Sonobe, K. Otsuka, Cluster-based industrial development: An East Asian model, Palgrave Macmillan, New York, 2006.

[27] C. Steinle, H. Schiele, When do industries cluster? A proposal on how to assess an industry's propensity to concentrate at a single region or nation, Research Policy, 31 (2002) 849-858

[28] M.J. Waits, The added value of the industry cluster approach to economic analysis, strategy development, and service delivery, Economic Development Quarterly, 14(2000) 35-50.

[29] H. Yamawaki, The evolution and structure of industrial clusters in Japan, Small Business Economics, 18 (2002) 121-140.

Figure 1 Industry zones in Taiwan

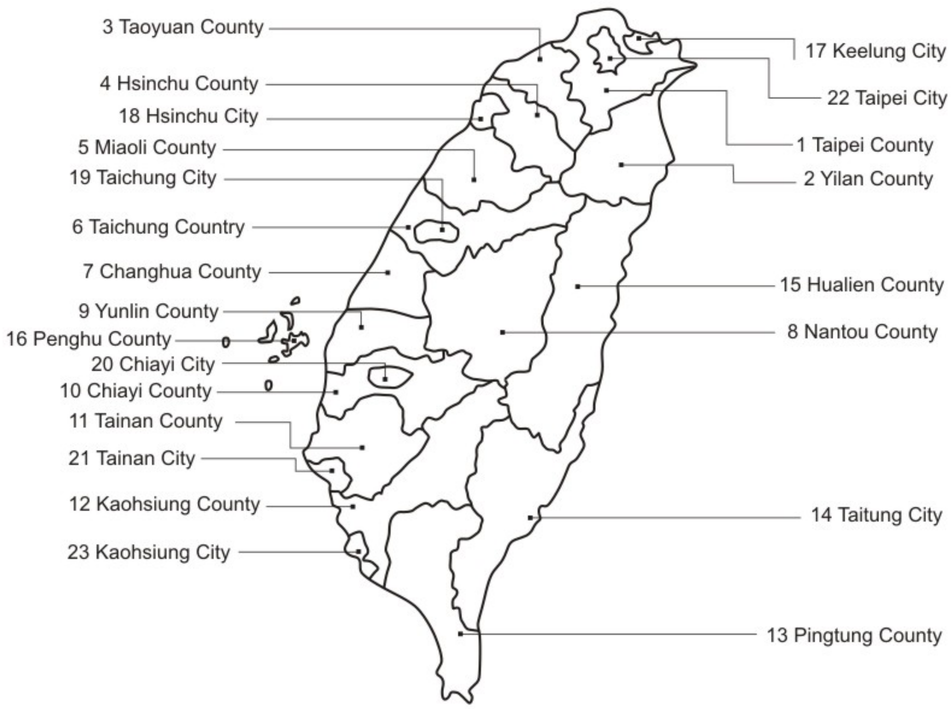

Table 1 the number of administrative district in Taiwan

County or City	County-Adm. City	Jhen	Siang	District	Total
1. Taipei County	10	4	15	-	29
2. Yilan County	1	3	8	-	12
3. Taoyuan County	4	2	7	-	13
4. Hsinchu County	1	3	9	-	13
5. Miaoli County	1	6	11	-	18
6. Taichung County	3	5	13	-	21
7. Changhua County	1	7	18	-	26
8. Nantou County	1	4	8	-	13
9. Yunlin County	1	5	14	-	20
10.Chiayi County	2	2	14	-	18
11. Tainan County	2	7	22	-	31
12. Kaohsiung County	1	3	23	-	27
13.Pingtung County	1	3	29	-	33
14.Taitung County	1	2	13	-	16
15. Hualien County	1	2	10	-	13
16. Penghu County	1	-	5	-	6
17. Keelung City	-	-	-	7	7
18. Hsinchu City	-	-	-	3	3

19. Taichung City	-	-	-	8	8
20. Chiayi City	-	-	-	2	2
21. Tainan City	-	-	-	7	7
22. Taipei City	-	-	-	12	12
23. Kaohsiung City	-	-	-	11	11
Total	32	58	219	50	359

Source : County and City Government, Taiwan.

Table 2: The area, number of industrial park and number of firm by geographic locations

City or County	Industrial Area[1] (hectares)	Number of industrial Park[1]	Number of plant[2]
1.Taipei County	2696.13	4	34709
2. Yilan County	610.72	2	1906
3. Taoyuan County	3131.38	7	13348
4. Hsinchu County	791.32	1	2342
5. Miaoli County	675.35	3	2939
6. Taichung County	1916.67	3	22591
7. Changhua County	676.33	6	17302
8. Nantou County	314.97	2	1832
9. Yunlin County	610.72	4	2451
10. Chiayi County	560.82	5	2518
11. Tainan County	2551.01	3	8842
12. Kaohsiung County	2411.88	6	6980
13. Pingtung County	652.46	3	2024
14. Taitung County	146.17	1	327
15. Hualien County	520.61	2	816
16. Penghu County	42.13	0	175
17. Keelung City	558.54	1	756
18. Hsinchu City	403.02	0	2741
19. Taichung City	657.50	2	6671
20. Chiayi City	223.09	0	1294
21. Tainan City	908.13	2	6130
22. Taipei City	452.4	1	10400
23. Kaohsiung City	906.7	1	4821

1 Source: Urban and Housing Development Department Council for Economic Planning and Development, Executive Yuan, Taiwan, 2007

2 Source: Directorate General of Budget, Accounting and Statistics (DGBAS) of Taiwan's Executive Yuan, 2001.

Table 3 the distribution of share of number of firm by two-digit SIC industry and geographical location in Taiwan unit:%

SIC 2 \County	1 Taipei	2 Yilan	3 Taoyuan	4 Hsinchu	5 Miaoli	6 Taichung	7 Changhua	8 Nantou	9 Yunlin	10 Chiayi	11 Tainan	12 Kaohsiung	13 Pingtung	14 Taitung	15 Hualien	16 Penghu	17 Keelung City	18 Hsinchu City	19 Taichung City	20 Chiayi City	21 Tainan City	22 Taipei City	23 Kaohsiung City	Total
8	8.82	3.45	6.67	1.76	2.30	6.45	10.50	2.51	8.87	6.87	6.95	5.89	4.41	0.90	1.12	0.70	0.75	1.78	2.95	1.81	2.51	8.09	3.92	**100**
10	22.28	0.61	18.55	0.76	1.05	7.26	26.06	0.46	1.87	0.78	6.59	0.97	0.55	0.03	0.02	0.00	0.06	0.08	1.23	0.20	2.35	7.71	0.53	**100**
11	26.11	6.21	5.71	0.57	3.15	4.60	13.30	0.50	4.18	2.47	5.84	2.06	1.31	0.07	0.18	0.02	0.33	0.85	1.93	0.72	6.15	11.64	2.10	**100**
12	13.87	0.55	2.88	0.18	2.69	32.18	15.70	2.10	2.46	1.32	5.80	2.65	0.59	0.00	0.05	0.00	0.00	0.37	4.93	0.55	6.30	3.74	1.10	**100**
13	7.83	3.60	6.28	2.10	4.39	22.33	8.71	6.25	3.14	2.83	3.75	6.61	3.81	0.43	0.82	0.12	0.15	1.61	4.72	2.59	2.13	2.83	2.96	**100**
14	14.08	0.98	4.42	1.29	2.30	23.93	12.28	2.11	2.15	4.92	5.68	5.08	1.99	0.28	0.22	0.03	0.06	2.43	4.80	2.15	2.21	4.26	2.34	**100**
15	23.54	0.97	9.15	1.52	4.84	11.14	11.22	2.20	1.80	1.80	5.41	3.74	0.86	0.13	0.10	0.00	0.08	1.94	4.81	0.63	4.26	7.71	2.14	**100**
16	28.75	0.84	4.74	0.49	0.95	7.17	4.78	0.87	0.71	0.62	2.05	1.48	1.10	0.34	0.42	0.06	0.42	1.65	8.35	1.29	5.67	21.48	5.75	**100**
17	12.80	0.98	11.46	3.17	3.54	10.67	8.05	1.65	1.40	3.48	7.50	8.54	1.40	0.00	0.43	0.00	0.18	0.98	2.38	0.55	2.01	15.00	3.84	**100**
18	17.29	0.89	11.81	2.70	1.47	9.35	5.65	1.44	1.71	1.95	7.09	5.03	1.51	0.00	0.14	0.00	0.31	1.34	4.14	0.89	4.21	17.19	3.90	**100**
19	14.19	5.88	5.54	2.42	4.50	7.96	5.19	0.69	6.57	3.11	5.19	8.30	3.11	1.04	5.54	0.00	0.00	1.04	4.84	1.73	0.69	6.92	5.54	**100**
20	29.11	0.40	7.49	1.61	1.04	12.51	17.75	1.73	1.10	1.15	3.40	5.01	0.69	0.00	0.17	0.00	0.06	2.42	3.17	0.46	2.77	4.96	3.00	**100**
21	26.27	0.51	7.10	1.14	1.28	18.24	12.17	0.70	1.00	1.54	8.93	3.34	0.65	0.01	0.07	0.02	0.06	1.69	3.43	0.51	6.49	3.42	1.43	**100**
22	19.08	2.76	7.05	2.95	9.34	6.20	5.35	2.76	2.43	2.36	4.72	5.21	2.69	0.80	8.18	0.54	0.40	4.79	1.67	0.66	1.11	6.32	2.64	**100**
23	17.83	0.86	10.20	1.56	1.59	14.68	11.51	0.89	1.19	1.19	7.53	8.92	1.43	0.18	0.27	0.00	0.10	1.54	3.98	1.06	3.89	4.48	5.12	**100**
24	18.09	1.21	5.98	1.26	1.79	17.38	18.94	1.53	1.39	1.60	5.48	6.45	1.75	0.42	0.56	0.15	0.60	1.13	4.13	0.99	3.02	3.58	2.57	**100**
25	26.30	0.84	9.43	1.27	1.32	21.18	6.69	0.55	0.66	0.95	4.85	4.72	0.53	0.07	0.22	0.01	0.30	2.08	6.57	0.82	3.92	3.47	3.23	**100**
26	44.37	0.34	13.75	3.41	1.15	3.77	1.73	0.20	0.20	0.53	1.48	1.34	0.31	0.00	0.00	0.00	1.06	3.72	2.43	0.34	1.26	16.13	2.49	**100**
27	35.75	0.67	23.72	4.44	1.37	5.42	1.62	0.41	0.69	0.22	2.10	2.75	0.85	0.09	0.02	0.00	0.52	2.95	1.75	0.07	1.30	9.34	3.94	**100**
28	30.58	0.66	9.68	2.19	1.77	12.89	4.07	0.55	0.79	0.69	5.26	3.71	1.12	0.03	0.15	0.07	0.94	3.89	3.87	0.35	4.75	8.24	3.75	**100**
29	14.42	0.68	9.15	1.42	0.68	16.10	18.09	0.53	0.50	1.20	9.97	7.20	1.32	0.07	0.07	0.53	3.07	0.72	3.90	0.60	5.70	4.08	7.20	**100**
30	23.22	0.53	5.80	2.53	1.06	7.34	4.54	0.47	0.79	1.16	17.99	2.85	0.32	0.00	0.05	0.00	0.58	1.85	4.64	0.53	12.88	10.87	2.85	**100**
31	24.29	1.10	6.66	0.84	1.38	17.07	10.70	1.58	1.15	1.99	6.70	4.17	1.36	0.06	0.32	0.17	0.28	1.32	4.52	0.56	6.72	8.45	2.77	**100**
Total	23.62	1.30	9.08	1.59	2.00	15.38	11.77	1.25	1.67	1.71	6.02	4.75	1.38	0.22	0.56	0.12	0.52	1.87	4.54	0.88	4.17	7.07	3.29	**100**

Note: The name of SIC 2-digit industry are shown in the appendix. The figures represent the share of number of firm in each 2-digit industry by location.

Among 23 locations, 17-23 are cities, 1-16 are counties. The shadow means the four larges share in the 23 locations.

Table 4 the number of firms in EPZs by two-digit industry in 2007

SIC Code	Industry \ EPZs	Nantze EPZ	Kaohsiung EPZ	Taichung EPZ	Chengkung EPZ	Pingtung EPZ	Others EPZ	Total
8	Food Manufacturing	0	0	1	1	0	0	2
10	Textile Mill Products	0	0	0	0	0	0	0
11	Wearing Apparel & Accessories	3	6	0	0	0	0	9
12	Leather & Fur Products	0	0	0	0	0	0	0
13	Wood & Bamboo Products	0	0	0	1	0	0	1
14	Furniture & Fixtures	0	0	0	0	0	0	0
15	Pulp, Paper & Paper Products	0	0	0	1	0	0	1
16	Printing Processing	1	0	0	3	0	0	4
17	Basic Chemical Matter Manufacturing	0	0	0	5	0	0	5
18	Chemical Products	4	3	0	1	0	1	9
19	Petroleum & Coal Products	0	0	0	0	0	0	0
20	Rubber Products Manufacturing	0	0	0	0	0	0	0
21	Plastic Products Manufacturing	6	3	0	3	2	0	14
22	Non-Metallic Mineral Products	1	0	0	1	0	0	2
23	Basic Metal Industries	0	2	0	0	0	0	2
24	Fabricated Metal Products	5	6	2	5	0	1	19
25	Machinery & Equipment	4	2	3	3	0	2	14
26	Audio & video products	0	0	0	0	0	0	0
27	Electronic parts and components	30	33	19	5	0	16	103
28	Electric machinery and parts	8	16	10	1	1	4	40
29	Transportation Equipment	0	2	0	0	0	1	3
30	Precision Instruments	0	0	0	0	0	0	0
31	Miscellaneous Industrial Products	4	2	5	0	0	0	11
	Total	66	76	40	29	3	25	239
	Land Area (hectare)	97.8	72.0	26.2	177.0	124.1	339.4	

Source: the website of the Export Processing Zone Administration, MOEA.
Note: Others EPZ include Chengkung Logistic EPZ, Linkuang EPZ, Kaushsiung Software Science-Based industrial park, Yulin Silk EPZ.

Table 5 Number of Firm in HSP from 1983 to 2006,

Year	Number of Firms
1983	37
1989	105
1996	203
2001	312
2006	391

Source: HSP administration

Table 6 Number of Firm in STSP from 1998 to 2006

Year	Number of Firms
1998	2
2000	13
2002	33
2004	73
2006	101

Source: STSP administration

Table 7: Breakdown of number of plants by 2-digit industry

2-digit industry		Number of Firm		Innovating firms (%)	Innovating intensity (%)
8	Food Manufacturing	Frequency	%	3.43	0.31
10	Textile Mill Products	6566	4.27	2.82	0.15
11	Wearing Apparel & Accessories	4570	2.97	1.88	0.06
12	Leather & Fur Products	2191	1.42	2.37	0.16
13	Wood & Bamboo Products	3282	2.13	0.46	0.01
14	Furniture & Fixtures	3168	2.06	1.93	0.05
15	Pulp, Paper & Paper Products	3824	2.48	1.70	0.09
16	Printing Processing	8729	5.67	0.74	0.06
17	Basic Chemical Matter Manufacturing	1640	1.07	9.82	0.59
18	Chemical Products	2921	1.90	10.58	3.10
19	Petroleum & Coal Products	289	0.19	4.50	0.36
20	Rubber Products Manufacturing	1735	1.13	3.98	0.19
21	Plastic Products Manufacturing	12061	7.84	2.29	0.12
22	Non-Metallic Mineral Products	4243	2.76	2.97	0.19
23	Basic Metal Industries	5961	3.87	2.40	0.11
24	Fabricated Metal Products	28614	18.59	1.60	0.08
25	Machinery & Equipment	28186	18.31	2.59	0.35
26	Audio & video products	3577	2.32	14.71	3.32
27	Electronic parts and components	5384	3.50	12.11	5.48
28	Electric machinery and parts	7242	4.71	4.97	0.38
29	Transportation Equipment	6303	4.10	5.19	0.37
30	Precision Instruments	1923	1.25	7.18	1.23
31	Miscellaneous Industrial Products	4820	3.13	2.80	0.20
	Total manufacturing	153915	100.00	3.36	0.74

Table 8: Description of Variables

Variable	Description	Mean	Std. Error
TFP	Total Factor Productivity of four-digit SIC industry in 2001	0.15	0.32
Gini	Gini coefficient of four-digit SIC industry number of firms across 359 geographic area, weighted by total number of firms for the industry in 2001.	0.89	0.06
$Gini_{\mathrm{R\&D}}$	Gini coefficient of four-digit SIC industry number of innovation firms across 359 geographic area, weighted by total number of innovation firms for the industry in 2001.	0.92	0.24
GHHI	Herfindahl index of four-digit SIC industry number of firms across 359 geographic area in 2001 (*GHHI* are measured in percentage terms).	539.4	950.8
$GHHI_{\mathrm{R\&D}}$	Herfindahl index of four-digit SIC industry number of innovation firms across 359 geographic area in 2001($GHHI_{\mathrm{R\&D}}$ are measured in percentage terms)	1870.5	2304.1
RD	Log of industry expenditures on research and development in 2001	5.52	2.46

Table 9: Summary statistics on computed *RTS* and *TFP*

Industry		RTS		TFP	
		Mean	*Std. Dev.*	*Mean*	*Std. Dev.*
8	Food Manufacturing	1.01	(0.03)	0.07	(0.17)
10	Textile Mill Products	1.11	(0.04)	0.41	(0.24)
11	Wearing Apparel & Accessories	0.95	(0.05)	-0.10	(0.26)
12	Leather & Fur Products	0.92	(0.04)	-0.34	(0.39)
13	Wood & Bamboo Products	1.07	(0.07)	0.13	(0.21)
14	Furniture & Fixtures	1.16	(0.08)	0.48	(0.38)
15	Pulp, Paper & Paper Products	1.00	(0.04)	-0.02	(0.16)
16	Printing Processing	0.97	(0.29)	-0.03	(0.07)
17	Basic Chemical Matter Manufacturing	1.01	(0.02)	0.05	(0.09)
18	Chemical Products	0.90	(0.03)	-0.43	(0.31)
19	Petroleum & Coal Products	1.26	(0.19)	1.56	(1.89.)
20	Rubber Products Manufacturing	1.13	(0.05)	0.51	(0.38)
21	Plastic Products Manufacturing	1.03	(0.03)	0.07	(0.12)
22	Non-Metallic Mineral Products	1.09	(0.05)	0.28	(0.33)
23	Basic Metal Industries	1.05	(0.02)	0.20	(0.21)
24	Fabricated Metal Products	1.04	(0.06)	0.06	(0.20)
25	Machinery & Equipment	1.06	(0.06)	0.14	(0.20)
26	Audio & video products	1.06	(0.03)	0.27	(0.22)
27	Electronic parts and components	1.17	(0.02)	0.76	(0.48)
28	Electric machinery and parts	1.02	(0.05)	0.08	(0.22)
29	Transportation Equipment	0.94	(0.03)	-0.02	(0.21)
30	Precision Instruments	1.10	(0.06)	0.40	(0.40)
31	Miscellaneous Industrial Products	1.03	(0.04)	0.12	(0.20)

All indicators (*RTS* and *TFP*) are computed using the parameters of a Translog production function, as described in Equations (2). *RTS* and *TFP*, as defined by Equations (4) and (5) respectively, vary inside a given 2-digit industry.

Table 10 Regression results estimating reduced form

Variables	$Gini_{R\&D}$	$GHHI_{R\&D}$
Gini	0.118 (0.022)*** [0.028]***	---
GHHI	--	1.680 (0.313***) (0.372)***
RD	-0.001 (0.0006)** [0.0007]**	-196.70 (99.4)** (123.1)
Constant	0.885 (0.025)*** [0.028]***	3290.70 (1950.4)* (444.2)***
F-statistic	342.08	76.72
Adjusted R^2	0.393	0.321
Sample size	224	224

Numbers in parentheses are standard errors, while numbers in brackets are the white robust standard errors; * significant at 10%; *** significant at 1%

Models include a set of 3-digits industries dummies

Table 11 2SLS Regression results estimating total factor productivity of four-digit SIC industry

Variables	*TFP*	*TFP*
$Gini_{R\&D}$	5.771 (1.720)*** [3.198]*	--
$GHHI_{R\&D}$	--	0.00004 (0.00001)*** [0.00001]***
RD	0.032 (0.006)*** [0.011]***	0.035 (0.006)*** [0.010]***
Constant	-5.823 (1.727)*** [3.217]*	-0.219 (0.118)* [0.083]***
Wald Chi-square	42009	1.5e+07
Adjusted R^2	0.849	0.880
Sample size	224	224

Numbers in parentheses are standard errors, while numbers in brackets are the white robust standard errors; * significant at 10%; *** significant at 1%

Models include a set of 3-digits industries dummies.

Table 12 the Hausman test for endogenous regressor

Variable	*TFP*	*TFP*
$Gini_{R\&D}$	5.771 (2.003)***	--
$GHHI_{R\&D}$	--	0.00004 (0.00001)***
RD	0.032 (0.007)***	0.035 (0.007)***
$\hat{\varepsilon}$	-6.26 (2.223)***	-0.00002 (0.00001)*
Constant	-5.823 (2.010)***	-0.219 (0.150)
F-statistic	9.77	10.48
Adjusted R^2	0.797	0.809
Sample size	224	224

Numbers in parentheses are standard errors; * significant at 10%; *** significant at 1%.

Models include a set of 3-digits industries dummies.

Appendix

2-digit SIC code and industry

2-digit SIC Code	2-digit SIC Industry
8	Food Manufacturing
10	Textile Mill Products
11	Wearing Apparel & Accessories
12	Leather & Fur Products
13	Wood & Bamboo Products
14	Furniture & Fixtures
15	Pulp, Paper & Paper Products
16	Printing Processing
17	Basic Chemical Matter Manufacturing
18	Chemical Products
19	Petroleum & Coal Products
20	Rubber Products Manufacturing
21	Plastic Products Manufacturing
22	Non-Metallic Mineral Products
23	Basic Metal Industries
24	Fabricated Metal Products
25	Machinery & Equipment
26	Audio & video products
27	Electronic parts and components
28	Electric machinery and parts
29	Transportation Equipment
30	Precision Instruments
31	Miscellaneous Industrial Products