Industrial Agglomeration and Use of the Internet*

Chia-Lin Chang
Department of Applied Economics
Department of Finance
National Chung Hsing University
Taichung, Taiwan

Michael McAleer
Department of Quantitative Finance
National Tsing Hua University, Taiwan
and
Econometric Institute, Erasmus School of Economics
Erasmus University Rotterdam
and
Tinbergen Institute, The Netherlands
and
Department of Quantitative Economics
Complutense University of Madrid, Spain

Yu-Chieh Wu
Department of Applied
Economics National Chung Hsing
University Taichung, Taiwan

EI2015-24

August 2015

***** For financial support, the first author wishes to thank the National Science Council, Taiwan, and the second author is grateful to the National Science Council, Taiwan and the Australian Research Council.

Abstract

Taiwan has been hailed as a world leader in the development of global innovation and industrial clusters for the past decade. This paper investigates the effects of industrial agglomeration on the use of the internet and internet intensity for Taiwan manufacturing firms, and analyses whether the relationships between industrial agglomeration and total expenditure on internet usage for industries are substitutes or complements. The sample observations are based on 153,081 manufacturing plants, and covers 26 2-digit industry categories and 358 geographical townships in Taiwan. The Heckman selection model is used to adjust for sample selectivity for unobservable data for firms that use the internet. The empirical results from two-stage estimation show that: (1) for the industry overall, a higher degree of industrial agglomeration will not affect the probability that firms will use the internet, but will affect the total expenditure on internet usage; and (2) for 2-digit industries, industrial agglomeration generally decreases the total expenditure on internet usage, which suggests that industrial agglomeration and total expenditure on internet usage are substitutes.

Keywords: Industrial agglomeration and clusters, Global innovation, Internet penetration, Manufacturing firms, Sample selection, Incidental truncation.

JEL: D22, L60.

1. Introduction

With the arrival of the Internet, its usage by business enterprises has continued to increase dramatically. Furthermore, the proliferation of Internet technology has as a result enhanced the development of electronic commerce and online shopping. Internet technology has replaced long-distance non-electronic communications (such as communications and business travel), and has reduced the costs of relaying information over long distances, thereby making it much easier for businesses to communicate with each other.

As an important case in point, Taiwan's overall industrial Internet penetration rate (that is, the proportion of medium-sized enterprises that use the Internet) has increased from 62% in 2002, to 79% in 2003, and to 94.3% in 2010. Taiwan has been hailed as a leader in the development of global innovation and industrial clusters for the past decade. According to reports prepared by the Institute for Information Industry in 2008[1], 2009[2] and 2010, the growth of the Internet has been the fastest in the manufacturing industry and distribution services. The industries with the highest Internet usage include banking and insurance, accommodation and catering.

As internet usage continues to develop and information is exchanged increasingly rapidly, the management information systems of businesses are becoming more complete, to the extent that firms can use the Internet to communicate and share information with other enterprises, both directly and in real time. It is for this reason, among others, that businesses have lower costs of communicating and collecting

[1] See http://www.find.org.tw/market_info.aspx?n_ID=7068
[2] See http://www.find.org.tw/market_info.aspx?n_ID=7095

information. Owing to the increased convenience that the Internet has brought in enabling firms to communicate with each other and in reducing the costs of transportation, as well as an abundance of resources that has further hastened the exchange of information, the "distance" factor is clearly no longer as important as it was in the past.

According to the 2009-2013 Global Competitiveness Report compiled by the World Economic Forum, Switzerland, the state of cluster development for Taiwanese industry was ranked first in the world for three consecutive years from 2006 to 2008. Not surprisingly, Taiwan has been hailed as a leading model for the development of global innovation and industrial clusters. Despite its ranking falling to 6 and 3 in the following two years, the state of its cluster development enabled Taiwan to receive a score of 5.5 (of a possible 7) in 2014, thereby regaining its leading position in the world. As for the pattern of spatial distribution of Taiwan's industrial clusters, the northern region is characterized by "electronics technology industrial clusters", the central region by "precision machinery industrial clusters", and the southern region by "electrical machinery industrial clusters". Each of the industrial clusters is clearly well developed (for further information, see Schwab and Sala-i-Martin, 2009, 2010, 2011, 2012).

Many scholars have focused on R&D and new technology (Audretsch and Feldman, 1996; Bertschek and Fryges, 2002; Chang and Oxley, 2009), while others have examined the relationship between Internet usage and urbanization economics (Forman et al., 2005a, b, c). However, there has also been research undertaken on the relationship between Internet usage and industrial agglomeration. Moreover, in respect of the total expenditure on internet usage, actual figures are observed only if the firm uses the internet, which leads to the problem of sample selection. The purpose of this paper is to incorporate the effect of sample correction, examine whether a relationship exists between agglomeration and Internet usage, and evaluate the factors that determine the extent of Internet influence.

The remainder of the paper is as follows. The literature on the influence of the factors related to Internet usage is reviewed in Section 2, the selection bias model and Heckman's two-step efficient estimation method are presented in Section 3, a description of the sample and the variables to be used are in Section 4, the empirical results are presented in Section 5, and the Conclusion is given in Section 6.

2. Literature Review

Forman et al. (2005a) proposed three related theories for the relationship between internet technology and urban agglomeration, namely: (1) global village theory, (2) urban density theory, and (3) industry composition theory. The global village theory suggests that the new network technologies would help break down the barriers between individuals and groups. As the suppliers and consumers of these manufacturers located in villages or small towns were likely themselves to be located in relatively faraway places, when these companies used the Internet the geographical barriers between manufacturers could be broken, thereby reducing transaction costs and reaping greater benefits. In other words, a manufacturer located in a village or a small city will gain the maximum benefit as a result of using the internet technology. Internet technology can make up for the disadvantages faced by manufacturers due to their being located far from the city's center of economic activity, and for this reason a substitution relationship exists between the adoption of internet technology and urban agglomeration.

The urban density theory suggest that, as the density and scale of urbanization increase, the costs borne by manufacturers using internet technology will be reduced. As the use of new technology often requires specialized technical skills, manufacturers will need to hire additional staff and purchase more equipment and software. Moreover, in urbanized areas there tends to be greater internet-related basic infrastructure and a larger labor market, so that the costs of using internet technology in cities will be lower. In other words, if the manufacturer is located in the city center, a reduction in the cost of using internet technology will increase Internet usage, so that a complementary relationship exists between the adoption of internet technology and urban agglomeration.

Industry composition suggests that when the density and scale of urban areas increase, the benefits that manufacturers derive from using the Internet will also increase. Before network technology began to be used widely, manufacturers had already decided where to locate their activities, and large numbers of manufacturers that used information-intensive technology industry tended to agglomerate in a certain area. Such firms were inclined to locate their operations in urban areas, so that the demand for the Internet was greater in these built-up areas. In other words, the demand for the Internet increased with the scale of urbanization. For this reason, a complementary relationship exists between the usage of Internet technology and urban agglomeration.

Forman et al. (2005a) used U.S. data to examine the relationship between internet penetration and urbanization, and find that when the number of manufacturers in leading industries in urban areas increases, this will cause Internet usage in such regions to increase. This indicates that the use of the Internet will be enhanced as the scale of urbanization increases, that is, a complementary relationship exists between Internet usage and urban agglomeration. Forman et al. (2005b) subsequently compared the influence of the location of enterprises and industrial agglomeration on Internet usage for the information usage and information-producing manufacturing industries. They found that in the areas in which manufacturers are located, the larger is the scale of industrial agglomeration, the more frequently will manufacturers use the Internet. A similar result using U.S. business data in Kolko (1999) also indicated a complementary relationship between the Internet usage rate and the scale of urbanization.

An alternative investigation on information technology-related manufacturing industry in the U.S. (computer and peripheral parts manufacturing, semiconductors and other components manufacturing) and information technology-related service industries (software publishing, computer systems design and related services) by Kauffman and Kumar (2007) tested three hypotheses: (1) Internet usage reduces market linkages; (2) the effects of Internet usage on market linkages are equal for IT-related industries and information technology-related service industries; and (3) the effects of these market linkages in urban and non-urban areas are equal. Their results indicate that Internet usage will lead to a reduction in market linkages and that the Internet effect will be less pronounced in urban areas than in rural areas. However, the effects of Internet usage in terms of the extent of its impact on IT-related manufacturing and information technology-related services are not significantly different.

Galliano and Roux (2008) used French manufacturers' sample survey data for the year 2002 to examine the behaviour of firms in the e-commerce industry in terms of their use of "Information and Communications Technology (ICT)." Their empirical research indicated that, for those manufacturers located in non-urban areas, the extent to which they used the Internet was lower than that for their counterparts in urban areas.

Moreover, for those industries for which there was a higher degree of agglomeration, the less frequently that manufacturers used the Internet led to a substitution relationship between Internet usage and agglomeration.

Lal (1999) used survey data for the year 1994 to investigate the factors affecting the manufacturers' use of the Internet for the Indian manufacturing industry. Based on the extent to which the sampled firms used IT technology (IT), Lal (1999) grouped the manufacturers into: (1) manufacturers without technology, (2) manufacturers with a low level of technology, (3) manufacturers with a medium level of technology, and (4) manufacturers with a high level of technology. Furthermore, Lal (1999) referred to four categories of factors that affected Internet usage, namely: (1) the characteristics of entrepreneurs, which included the managers' qualifications and their ability to understand R & D, and the degree of importance they attached to product quality and market share, (2) international orientation (the extent to which products were imported and exported), (3) human capital, and (4) the manufacturers' scale of operations. The empirical results showed that the education of managers, the scale of the manufacturers' operations and R & D had a significant and positive impact on the use of the Internet. Moreover, Lal (1999) emphasized that the rapid growth of Internet technology and information technology had increased the demand for skilled labor in developing countries, thereby making small and medium-sized enterprises more globally competitive.

Bertschek and Fryges (2002) used sample survey data for German companies in both the services and manufacturing industry sectors for the year 2000, and examined the factors affecting the degree to which manufacturers decided to use B2B (business-to-business) Internet technology. They categorized the intensity of Internet technology usage by manufacturers according to whether they had: (1) not used B2B Internet technology, (2) used B2B internet technology, and (3) extensively used B2B Internet technology. They used factors which had been deemed in the literature to have affected the manufacturer's adoption of new technologies, including the scale of the manufacturer's operations, the age of manufacturing plants, human capital, and international competitive pressure, as well as variables that had not been considered in the literature, such as electronic data interchange (EDI), which can be regarded as a precursor to B2B electronic commerce, and the bandwagon effect or herd behaviour.

Bertschek and Fryges (2002) found that the scale of manufacturers' operations, the quality of staff, and the degree of openness to international markets had a significant and positive impact on the extent to which manufacturers used B2B Internet technology. Moreover, they found that: (1) the probability that manufacturers with a history of using EDI technology would also use B2B technology extensively in the future was extremely high; and (2) the greater was the Internet technology usage by other manufacturers within the same industry, the greater was the likelihood that the manufacturers would also use new technologies.

Giunta and Trivieri (2007) examined the factors determining the use of information technology (IT) by small manufacturing enterprises in Italy's manufacturing industry. Using sample survey data for 17,000 small and medium-sized firms for the period July 2001 to February 2002, and by focusing on the extent to which the manufacturers used IT, they categorized the manufacturers into those that had zero, low, medium, and high use of IT. They found that the factors that significantly affected the manufacturer's use of IT included the scale of the manufacturer's operations, the geographical location of the plant, the training provided by the manufacturers for their employees, the extent to which they engaged in R&D, the amount of outsourcing that took place, and the extent of cooperation with other manufacturers.

Galliano et al. (2011) used survey data on French manufacturers for 2001 and 2002, and discovered that using the Internet to co-ordinate and monitor the company's branch network within particular sectors was an important factor affecting the manufacturer's use of information and communications network technology. Therefore, the distance between the enterprise's head office and branch units, and the geographical dispersion of the enterprise's branch units, significantly affected the extent to which manufacturers used information and communications network technology. In addition, greater was the usage of Internet technology by enterprises within the same industry or geographical area, the greater was the contagion effect arising from the Internet technology, with a significant positive impact on the extent to which enterprises used the Internet. These empirical results lend support to the theories advanced by Mansfield (1963a, 1963b) and Saloner and Sheppard (1995).

As indicated in the literature review, much research has focused on the problems associated with Internet penetration related to urbanization, but few studies have examined the relationship between industrial agglomeration and the extent to which firms use the Internet. For this reason, this paper will focus on the issue of Internet use and industrial agglomeration.

3. Heckman Selection Model

In order to correct the inherent problem of selection bias, as discussed above, this paper uses the Heckman selection model (see Lewis 1974; Heckman 1976, 1979; Greene, 2003a), which assumes that there exists an underlying regression relationship, namely:

$$y_i = \boldsymbol{x}_i'\beta + u_{yi} \ , \quad i = 1,2,\dots,n \qquad (1)$$
$$u_{yi} \sim N\left(0, \sigma_y^2\right).$$

However, the dependent variable, y_i, is not always observed. Rather, the dependent variable for observation i is observed if $\boldsymbol{\omega}_i'\gamma + u_{2i} > 0$, as $\boldsymbol{\omega}_i'$ is the vector of variables that determines whether dependent variable, y_i, is observed or unobserved (that is, selected or not selected). Therefore, the selection equation can be written as:

$$z_i^* = \boldsymbol{\omega}_i'\gamma + u_{zi} \ , \ i = 1,2,\dots,n \qquad (2)$$
$$u_{zi} \sim N(0,1),$$
$$corr\left(u_{yi}, u_{zi}\right) = \rho.$$

When $\rho \neq 0$, Ordinary Least Squares (OLS) estimation applied to equation (1) yields biased estimates. As z_i^* is latent, it is more convenient to specify a binary variable, z_i, that identifies the observations for which the dependent is observed ($z_i^* \neq 0$) or not observed ($z_i^* = 0$). Thus, we reformulate the selection mechanism as follows:

$$z_i = \boldsymbol{\omega}_i'\gamma + u_{zi} = 1, \text{ if } z_i^* > 0,$$
$$z_i = \boldsymbol{\omega}_i'\gamma + u_{zi} = 0, \text{ otherwise}, \qquad (3)$$
$$prob(z_i = 1|\omega_i) = \Phi(\boldsymbol{\omega}_i'\gamma),$$
$$prob(z_i = 0|\omega_i) = 1 - \Phi(\boldsymbol{\omega}_i'\gamma),$$

where $\Phi(\cdot)$ is the standard normal cdf, and the regression model is given as follows:

$$y_i = \boldsymbol{x}_i'\beta + u_{yi} \text{ , observed only if } \quad z_i = 1,$$

$$(u_{zi,}, u_{yi}) \sim \text{bivariate normal } [0,0,1,\sigma_y^2,\rho].$$

The mean and variance of the incidentally truncated (or sample selection) bivariate normal distribution are given as equations (4) and (5)[3]:

$$\begin{aligned} E[y_i|z_i = 1] = E[y_i|u_{zi} > -\boldsymbol{\omega}_i'\gamma] \\ = \boldsymbol{x}_i'\beta + E\left[u_{yi}\middle|u_{zi} > -\boldsymbol{\omega}_i'\gamma\right] \\ = \boldsymbol{x}_i'\beta + \rho\sigma_y\lambda_i(\alpha_z) \\ = \boldsymbol{x}_i'\beta + \beta_\lambda\lambda_i(\alpha_z), \end{aligned} \tag{4}$$

$$Var[y_i|z_i = 1] = \sigma_y^2[1 - \rho^2\delta_i(\alpha_z)], \tag{5}$$

where $\alpha_z = -\boldsymbol{\omega}_i'\gamma/\sigma_z$, $\lambda_i(\alpha_z) = \emptyset(\alpha_z)/[1 - \Phi(\alpha_z)]$, $\delta_i(\alpha_z) = \lambda_i(\alpha_z)[\lambda_i(\alpha_z) - \alpha_z]$, and $0 < \delta_i < 1$. $\lambda_i(\alpha_z)$ is called the inverse Mill's ratio, $\emptyset(\cdot)$ is the standard normal pdf, and $\Phi(\cdot)$is the standard normal cdf. Thus, the regression with the observed data can be written as:

$$\begin{aligned} y_i|z_i = 1 = E[y_i|z_i^* > 0] + \upsilon_i \\ = \boldsymbol{x}_i'\beta + \beta_\lambda\lambda_i(\alpha_z) + \upsilon_i \end{aligned} \tag{6}$$

where the disturbance υ_i is heteroscedastic.

OLS regression of y_i on **x** and λ would give a consistent estimator, but if λ is

[3] The theorem of moments of the incidentally truncated bivariate normal distribution is given in Green (2003b, p. 781).

omitted, then there will be a specification error of an omitted variable (see Green, 2003). The marginal effect of the regressors on y_i in equation (6) is given as:

$$\frac{\partial E[y_i|z_i^*>0]}{\partial x_{ik}} = \beta_k - \gamma_k \left(\frac{\rho\sigma_y}{\sigma_z}\right) \delta_i(\alpha_z) \qquad (7)$$

where $\delta_i(\alpha_z) = \lambda_i(\alpha_z)[\lambda_i(\alpha_z) - \alpha_z]$, $0 < \delta_i < 1$.

The full marginal effect of the regressors on y_i in the observed sample consists of two parts: (i) the direct effect, which is β_k, and (ii) the indirect effect, which is $\gamma_\mathrm{k} \left(\frac{\rho\sigma_\varepsilon}{\sigma_u}\right) \delta_i(\alpha_\mathrm{u})$. Suppose that ρ is positive and $\mathrm{E}[y_i]$ is greater when $\mathrm{z}_i^* > 0$ than otherwise. Since $0 < \delta_i < 1$, for a particular independent variable, if it appears in the probability that $\mathrm{z}_i^* > 0$, then it will influence y_i through λ_i, and then reduce the marginal effect (see Green, 2003b, p.783).

The parameters of the sample selection model can be estimated by maximum likelihood (for details, see Maddala, 1983). However, Heckman's (1979) two-step estimation procedure is typically used. The first step estimates the selection equation by maximum likelihood to obtain an estimate of γ in equation (3), and to compute $\widehat{\lambda_i} = \emptyset(\boldsymbol{\omega}_i'\hat{\gamma})/\Phi(\boldsymbol{\omega}_i'\hat{\gamma})$ and $\widehat{\delta_i} = \hat{\lambda}_i(\hat{\lambda}_i - \boldsymbol{\omega}_i'\hat{\gamma})$. The second step estimates the regression equation by OLS to obtain estimates of β and $\beta_\lambda = \rho\sigma_y$. Green (2003a) provides the statistical proof for consistent estimators of the individual parameters ρ and σ_y^2.

4. Data and Variables

In this paper, we use census data for Taiwan's manufacturing industries obtained from the Directorate-General of Budget, Accounting and Statistics (DGBAS) for 2006. The entities surveyed include enterprises and establishments, these concepts being very similar to firms and plants, respectively. As manufacturing enterprises in Taiwan's manufacturing sector account for only 3.1% of all manufacturing enterprises, which is an exceedingly low ratio, we simply refer to plants as manufacturers. In order to reflect the use of the Internet by manufacturers from a geographical dimension, we adopt the establishment as the focus of the current research. If we were to adopt the enterprise instead, the scope of coverage would likely not be limited to just one location, and it would not be possible for this unit to reflect accurately the use of the Internet in a spatial context. Therefore, the sample comprises a total of 153,081 manufacturers, with 26 units at the 2-digit level and 212 units at the 4-digit level. The scope of coverage includes the island of Taiwan and the Penghu archipelago, for a total of 358 urban and rural areas. The 26 industries associated with the 2-digit code and the numbers of firms are given in Table 1, for traditional, technology-intensive and basic industries.

Since there are different ways of calculating industrial concentration in the literature, we use two of the more common indices to measure the degree of industrial concentration, namely the Herfindahl-Hirschman index (hereafter *HHI*) and the concentration ratio for the top four firms (*CR*4). The concept of the degree of industrial concentration is further extended to the estimation of industrial agglomeration, in which case we use the Geographical Herfindahl-Hirschman index (*GHHI*) as a proxy variable for industrial agglomeration. The formulae for the degree of industrial concentration and the geographical concentration index are given below.

(1) Herfindahl-Hirschman Index (*HHI*): The degree of industry concentration is used to measure the extent of the competition faced by an industry. The HHI for industry j is

calculated as follows:

$$HHI_j = \sum_{i=1}^{n} S_{ij}^2,\ 0 \le HHI \le 1,$$

where s_{ij} is the market share of firm *i* in industry j, $i = 1,2,\dots n$, and *n* is the number of firms in industry *j*.

The *HHI* is obtained by dividing the individual manufacturer's sales by the total sales of the industry in order to arrive at each manufacturer's market share, which is then squared. The advantage of *HHI* is that the manufacturer's market share serves as a weight, with smaller (larger) manufacturers being given smaller (larger) weights. The lower is *HHI* , the lower is the degree of concentration in the industry; and the higher is the value, the higher is the degree of industrial concentration.

(2) Concentration Ratio for the Top Four Firms (*CR*4): *CR*4 is the weighted average of the market shares of the top four firms in an industry. The formula for calculating the index for industry *j* is as follows:

$$CR4_j = \sum_{i=1}^{4} S_{ij},\ 0 \le CR4_j \le 1,$$

where s_{ij} is the market share of firm *i* in industry j. The higher is *CR*4, the higher is the degree of industrial concentration (see Bain, 1968).

(3) Geographical Herfindahl-Hirschman Index (*GHHI*): This is the Herfindahl-Hirschman Index (*HHI*) for industrial market concentration, together with a

geographical concept that reflects how firms are dispersed within a particular area. The formula for calculating the index is as follows:

$$GHHI_{jk} = \sum_{k=1}^{M} S_{jk}^2,\ \ 0 \le GHHI_{jk} \le 1,$$

where s_{jk} is the ratio of the number of firms in industry j in region k to the total number of firms in industry j, $k = 1,2, \ldots M$, and M is the number of regions.

When $GHHI_{jk}$ is close to 1, this means that the firms within the industry are more geographically concentrated, and when $GHHI_{jk}$ is close to 0, this means that the firms within the industry are more geographically dispersed. The advantage of $GHHI_{jk}$ is its simplicity of calculation, whereas its shortcomings include the following: (1) as it is necessary to obtain the market share of an industry for each firm, it is not easy to acquire the data; (2) if $GHHI_{jk}$ is not part of a neighborhood messaging system, it is not possible to reveal the differences brought about by being either closer or more distant, or to reflect the spatial correlation for different economic activities; thus, all one can do is indicate that economic activities are unevenly distributed; (3) $GHHI_{jk}$ can only reveal the spatial concentration for a single industry, without taking into consideration the spatial distribution characteristics for all industries as a whole.

In accordance with the literature, we select those factors influencing manufacturers' use of the Internet, including industrial characteristics (concentration), manufacturers' characteristics (scale of operations, manufacturers' organization, manufacturers' export intensity), geographical concentration of industry, geographical location, and the contagion effect for internet technology within the same region. Other explanatory variables include the manufacturer's size (size), with the number of staff hired by firms (staff + employees) representing the size of the manufacturer. The export rate (export_rate), namely the ratio of the manufacturer's export revenue to total revenue, is used to measure the extent to which manufacturers export their products.

The geographical locations (area_city) are divided into county and city categories. When area_city = 1, this means that the manufacturers are located in the following cities: Keelung, Hsinchu, Taichung, Chiayi, Tainan, Taipei, or Kaohsiung. When area_city=0, this means that the manufacturers are located in the following counties: Taipei, Yilan, Taoyuan, Hsinchu, Miaoli, Taichung, Changhua, Nantou, Yunlin, Chiayi, Tainan, Kaohsiung, Pingtung, Taitung, Hualien, or Penghu.

The group with independent operations is a control variable for firm characteristics. When group=1, this indicates that the manufacturer is an independent operating unit. When group=0, this refers to the manufacturer having branches (subsidiaries). Computer expenditure 1 (computer1) refers to the manufacturer having incurred expenses, as well as capital expenditure on investment in computer equipment. Computer expenditure 2 (computer2) refers to the total expenditure on computer equipment by other manufacturers within the same industry and same area after deducting the expenditure on computer equipment by the manufacturer. The computer2 variable is used to measure the contagion effect for the Internet technology within a certain area. Table 2 shows the variable definitions, and Table 3 represents the descriptive statistics of the explanatory variables.

As explained in Section 3, we use the Heckman two-stage estimation method to estimate the parameters of the sample selection model, which is specified as:

$$y_i = \beta_0 + \beta_1 HHI_{ji} + \beta_2 export\_rate_i + \beta_3 GHHI_{jki} + \beta_4 city_i + \beta_5 computer1_i + \beta_6 computer2_{jki} + \beta_\lambda \lambda_i + \varepsilon_{yi}, \quad (8)$$

where y_i is the ratio of total expenditure on internet use to total sales of firm *i* (intensity of internet use), and ε_{yi} is the disturbance. HHI_{ji} is the Herfindahl-Hirschman Index for the industry *j* to which firm *i* belongs, $\mathrm{export\_rate}_i$ is export intensity for firm *i*, $GHHI_{jki}$ is the Geographical Herfindahl-Hirschman Index for industry j in region *k* in which firm *i* is located, $city_i$ is a dummy variable indicating the firm's geographical location ($\mathrm{city}_i = 1$, if firm *i* is located in the city, $\mathrm{city}_i = 0$, otherwise), $computer1_i$ is the cost of buying the computer equipment for firm *i*, and $computer2_{jki}$ is the total cost of computer equipment for industry *j* in region *k*, excluding that of firm *i* itself. The variable "$\mathrm{computer2}_{jki}$" captures the contagion effect for the Internet technology in the same area and industry.

The coefficient λ_i is estimated from the selection model, which is given as:

$$z_i = \gamma_0 + \gamma_1 HHI_{ji} + \gamma_2 export\_rate_i + \gamma_3 GHHI_{jki} + \gamma_4 city_i + \gamma_5 size_i + \gamma_6 group_i + \varepsilon_{zi}, \quad (9)$$

where z_i is a binary variable, that is, $z_i = 1$ if firm *i* reports use of the Internet, $z_i = 0$, otherwise, and ε_{zi} is the error term. The explanatory variables to determine whether the dependent variable, z_i, is observed or unobserved, include industry characteristics (HHI_{ji}), export intensity ($export\_rate_i$), geographical concentration of the industry ($GHHI_{jki}$), geographical location ($city_i$), firm's characteristics ($size_i$), and firm's organization ($group_i$).

Table 4 shows the correlation coefficients for each variable. In addition to the correlation coefficient between $export\_rate_i$ and (HHI_i and $CR4_{ji}$) and $size_i$ being greater than 0.1, the correlation coefficients between each of the other variables are less than 0.1, reflecting the low degree of correlation between the variables. In the next section, we report the empirical results based on Heckman two-stage estimation.

5. Empirical Results

Column 2 of Tables 5 and 6 report the Heckman two-stage estimation for equation (8), which estimates the factors affecting the extent to which manufacturers use the Internet after correcting for sample bias. Table 5 reports the results with *HHI* as the proxy variable for the degree of industrial concentration, while Table 6 reports the results with *CR4* as the proxy variable for the degree of industrial concentration. Column 3 of both Tables 5 and 6 give the estimates for the select equation (9), which is estimated as a probit model.

In order to enhance efficiency in estimation, we also use bootstrapping methods to estimate the variances. The standard deviations, with and without bootstrapping, are reported in Tables 5 and 6. The 2-digit industry dummies are included in the empirical model to control for heterogeneity but, for reasons of space, we do not report each coefficient estimates of the 2-digit industries. The empirical result show that, regardless of whether the bootstrapping method is used, a non-zero Mill's lambda (β_λ) rejects the null hypothesis that β_λ is equal to zero at the 1% level of significance, indicating that sample selection bias should be taken into account. In order to make the empirical results more straightforward, we present the results for the whole manufacturing industry and then the results for individual 2-digit industries.

For the whole industry, we first summarize the results of the selection-corrected equation of the firm's internet use for the factors influencing the extent to which manufacturers use the Internet, and the marginal effects of the explanatory variables. Then we summarize the results of the selection equation for the factors determining the manufacturers' use of the Internet.

Regression model with selection corrected for all industries:

The coefficient of HHI_{ji} is positive but insignificant in Column 2 of Table 5, while the coefficient of $CR4_{ji}$ is positive and significant in Column 2 of Table 6. These results indicate that a higher degree of industrial concentration increases a firms' expenditure on internet use. The coefficients of $export\_rate_i$ are positive but insignificant in Column 2 of both Tables 5 and 6, indicating that export intensity has no statisitical impact on the expenditure of firms on internet use.

The coefficients of $GHHI_{jki}$ are negative and significant in Column 2 of both Tables 5 and 6, indicating that the lower is the level of the industrial agglomeration, the greater is the extent to which manufacturers will use the Internet. The coefficient of $city_i$ has positive and significant effects in Column 2 of both Tables 5 and 6.

The coefficient of $computer1_i$ has a positive but insignificant effect in Column 2 of both Tables 5 and 6, which indicates that the manufacturers' expenditure on computer equipment has no statistical impact on the expenditure of firms on internet use. The coefficient of $computer2_{jki}$ has a positive but insignificant effect, with bootstrap standard deviations, in Column 2 of both Tables 5 and 6. These results indicate that the manufacturers' expenditure on computer equipment within the same industry and region has no statistical impact on the expenditure of firms on internet use.

The marginal effects of equations (7) and (8) are reported in Table 7. Column 2 in Table 7 gives the industrial marginal effects with HHI_{ji} as the proxy variable for the degree of industrial concentration, while Column 3 gives the industrial marginal effects with $CR4_{ji}$ as the proxy variable.

For HHI_{ji}, the marginal effect is -0.0902 in Column 2 and -0.007 in Column 3 in Table 7. For example, -0.0902 means that when the degree of industrial concentration rate is increased by 1 unit, the extent to which manufacturers use the Internet is reduced by 0.0902%. Thus, the lower is the degree of industrial concentration, the greater is the extent to which manufacturers use the Internet. Not surprisingly, there are differences between the marginal effects of HHI_{ji} and $CR4_{ji}$ on the extent to which manufacturers use the Internet, as HHI_{ji} takes into account all firms in an industry, using manufacturer's market share as weights, while $CR4_{ji}$ considers only the weighted average of the market shares of the top four firms in an industry. These empirical findings of industrial concentration agree with Galliano and Roux (2008) and Galliano et al. (2011), who used French manufacturing industry data.

For $export\_rate_i$, the marginal effect are 0.2708 and 0.2963 for Columns 2 and 3 in Table 7 where, for example, 0.2708 means that when the export intensity is increased by 1 unit, the extent to which the manufacturers use the Internet will increase by 0.2708%.

For $GHHI_{jki}$, the marginal effects are -0.0245 and -0.0133 for Columns 2 and 3 in Table 7 where, for example, when the industrial agglomeration is reduced by 1 unit, the extent to which manufacturers use the Internet will increase by 0.0245%. Thus, there exists a substitution relationship between the extent to which manufacturers use the Internet and the level of industrial agglomeration. This empirical result accords with those obtained by Kauffman and Kumar (2007), who used U.S. information technology-related manufacturing and service industry data, and Galliano and Roux (2008), who used French manufacturing data. This result also confirms that the popularity of the Internet is such that the distance factor is no longer so important, so the Internet seems to have overcome the problem of the distance between manufacturers.

It worth noting that, associated with the dummy variable, $city_i$, the marginal effects are -0.0051 and -0.0062 for Columns 2 and 3 in Table 7 so that, for example, manufacturers who are located in the city areas will use the Internet -0.0051% less than those located in non-urban areas. This results also confirms the empirical findings of Forman et al. (2005) and Kolko (1999), so that a complementary relationship exists between Internet usage and urbanization.

Returning to Column 3 of Tables 5 and 6 for the probit results, as given in equation (9), the probit model estimates the factors relating to whether manufacturers will use the Internet for their business.

The empirical results show that, regardless of whether *HHI* or *CR4* is used as the proxy variable for the degree of industrial concentration, the coefficients of HHI_{ji} and $CR4_{ji}$ are negative and significant at the 1% level of significance in Column 3 of both Tables 5 and 6. These results indicate that the greater is the competition faced by manufacturers to increase their ability to compete with other manufacturers, the more likely they will be to use the Internet for business.

Export intensity is also an important factor that affects the manufacturers' use of the Internet. The coefficients of $export\_rate_i$ are positive and significant at 1% level of significance in Column 3 of both Tables 5 and 6. This is not surprising as the greater is the reliance of manufacturers on exports, the greater is their export intensity and the need to use the Internet for communicating with their foreign customers.

The coefficient of the geographical location, $city_i$ in Column 3 of both Tables 5 and 6 has a negative and significant effect on manufacturers' use of the Internet for their business. This result suggests that manufacturers who are located in non-urban areas will be more likely to use the Internet for business than those located in city areas. However, this result is in contrast with the empirical results of the coefficient of $city_i$ in Column 2 of Tables 5 and 6, which suggests that manufacturers who are located in city areas will spend more on Internet use than firms in non-urban areas.

The coefficient of the manufacturer's scale of operations, $size_i$ has a positive and significant effect on the manufacturers' use of the Internet for their business. It is not surprising that larger firms will be more likely to use the Internet for business. Moreover, the positive and significant coefficient of $group_i$ suggests that manufacturers with independent operations will be more likely to use the Internet for business than those who have subsidiaries (or branches). It is not surprising that, as Taiwan consists largely of manufacturers with independent operations, the likelihood of such manufacturers using the Internet is relatively high.

While the impact of the degree of industrial agglomeration on the manufacturers' use of the Internet is not significant in Column 3 of both Tables 5 and 6, the effect on the extent to which manufacturers use the Internet is significant and negative in Column 2 of both Tables 5 and 6. Therefore, the extent of industrial agglomeration does not affect whether manufacturers will use the Internet, but it does affect the extent to which manufacturers will use the Internet when they already do so.

Regression model with selection corrected for 2-digit industries:

In this section we report the Heckman two-stage estimation with *HHI* as the proxy variable for the degree of industrial concentration and the marginal effects for 2-digit industries in Tables 8 and 9, respectively. A nonzero Mill's lambda (β_λ) rejects the null hypothesis that $\beta_\lambda = 0$ at the 1% level of significance for (08) Food, (09) Beverages, (22) Plastic Products, (28) Electrical Equipment, (29) Machinery and Equipment, (30) Motor Vehicles and Parts, and (32) Furniture. However, as the industries are different, the empirical results for the individual industries based on the 2-digit level classifications also vary. For individual 2-digit industries, we first discuss the results of the selection-corrected equation regarding the extent to which manufacturers use the Internet, then the results of the selection equation for the factors that determine whether manufacturers use the Internet, followed by a summary of the marginal effects.

The effect of the degree of industrial agglomeration ($GHHI_{jki}$) on the extent to which manufacturers use the Internet vary across the 2-digit industries. For traditional industries, such as (08) Food, (12) Wearing Apparel and Clothing Accessories, (13) Leather, Fur and Related Products, (32) Furniture, technology-intensive industries, such as (28) Electrical Equipment, (30) Motor Vehicles and Parts, (31) Other Transport

Equipment, and basic industries, such as (24) Basic Metal, the lower is the level of industrial agglomeration, the greater is the extent to which manufacturers will use the Internet. However, only two traditional industries, such as (16) Printing and Reproduction of Recorded Media, and basic industries, such as (20) Medical Goods, show the higher is the degree of industrial agglomeration, the greater is the extent to which manufacturers will use the Internet.

The effect of the degree of industrial concentration (HHI_{ji}) in terms of the extent to which manufacturers use the Internet also differ across the 2-digit industries. In the case of traditional industries, such as (08) Food, (13) Leather, Fur and Related Products, technology-intensive industries, such as (26) Electronic Parts and Components, and basic industries, such as (25) Fabricated Metal Products, show the higher is the degree of industrial concentration, the greater is the extent to which manufacturers will use the Internet. On the contrary, traditional industries, such as (32) Furniture, (33) Manufacturing Not Elsewhere Classified, and technology-intensive industries, such as (28) Electrical Equipment, (29) Machinery and Equipment, (30) Motor Vehicles and Parts, (31) Other Transport Equipment, show the lower is the degree of industrial concentration, the greater is the extent to which manufacturers will use the Internet.

The variable, $export\_rate_i$, shows a positive and significant influence on the extent to which manufacturers use the Internet for traditional industries, such as (09) Beverages, (33) Manufacturing Not Elsewhere Classified, technology-intensive industries, such as (26) Electronic Parts and Components, Machinery and Equipment, (30) Motor Vehicles and Parts, and basic industries, such as (18) Chemical Material, (19) Chemical Products, (25) Fabricated Metal Products. However, only basic industries, such as (24) Basic Metal, show a significant and negative effect of the extent to which manufacturers use the Internet.

The effect of the geographic location, $city_i$, shows manufacturers that are located in non-urban areas will use the Internet to a greater extent than those located in the city areas for traditional industries, such as (08) Food Manufacturing, (09) Beverages. On the contrary, traditional industries, such as (15) Pulp, Paper and Paper Products, and technology-intensive industries, such as (31) Other Transport Equipment, show manufacturers that are located in city areas will use the Internet to a greater extent than those located in non-urban areas.

Manufacturers' expenditure on computer equipment, $\mathrm{computer1}_i$, has no statistical impact on the expenditures of firms on internet use for most of the 2-digit industries, except for traditional industries, such as (16) Printing and Reproduction of Recorded Media, technology-intensive industries, such as (30) Motor Vehicles and Parts, (31) Other Transport Equipment, and basic industries, such as (21)Rubber Products, (22) Plastic Products, (25) Fabricated Metal Products.

Similarly, $\mathrm{computer2}_{jki}$, which captures the contagion effects for Internet technology in the same area, has no statistical impact on the expenditures of firms on internet use for most 2-digit industries, except for traditional industries, such as (13) Leather, Fur and Related Products, and technology-intensive industries, such as (29) Machinery and Equipment and (31) Other Transport Equipment.

The probit model given in equation (9), which estimates the factors that determine whether manufacturers adopt the Internet for their business across the 2-digit industries, are given in Table 8. The coefficient estimates will now be discussed.

The effect of the degree of industrial agglomeration, $GHHI_{jki}$), on whether manufacturers will use the Internet, shows differences across the 2-digit industries. For traditional industries, such as (8) Food, (11) Textiles Mills, (13) Leather, Fur and Related Products, (14) Wood and Bamboo Products, technology-intensive industries, such as (29) Machinery and Equipment, (31) Other Transport Equipment, and basic industries, such as (25) Fabricated Metal Products, when the degree of industrial agglomeration is high, manufacturers will be more inclined to use the Internet. For traditional industries, such as (15) Pulp, Paper and Paper Products, (16) Printing and Reproduction of Recorded Media, (32) Furniture, (33) Manufacturing Not Elsewhere Classified, technology-intensive industries, such as (26) Electronic Parts and Components, (30) Motor Vehicles and Parts, and basic industries, such as (22) Plastic Products, when the degree of industrial agglomeration is high, manufacturers will be less inclined to use the Internet.

However, industrial agglomeration will not affect whether manufacturers use the Internet for most basic industries, such as (18) Chemical Material, (19) Chemical Products, (20) Medical Goods, (21) Rubber Products, (24) Basic Metal, traditional industries, such as (9) Beverages, (12) Wearing Apparel and Clothing Accessories, (23) Non-metallic Mineral Product, and technology-intensive industries, such as (27) Computers, Electronic and Optical Products, (28) Electrical Equipment.

The effect of the degree of industrial concentration, HHI_{ji}, on whether manufacturers will use the Internet shows differences across the 2-digit industries. For traditional industries, such as (11) Textiles Mills, (15) Pulp, Paper and Paper Products, (23) Non-metallic Mineral Products, (32) Furniture, technology-intensive industries, such as (29)

Machinery and Equipment, and basic industries, such as (22) Plastic Products, when the degree of industrial concentration increases, manufacturers will be more inclined to use the Internet. On the contrary, for traditional industries, such as (08) Food, (12) Wearing Apparel and Clothing Accessories, (13) Leather, Fur and Related Products, and basic industries, such as (25) Fabricated Metal Products, when the degree of industrial concentration decreases, manufacturers will be more likely to use the Internet.

The effect of $export\_rate_i$ is important on the manufacturers' decision to use the Internet for many 2-digit industries. For traditional industries, such as (14) Wood and Bamboo Products, (15) Pulp, Paper and Paper Products, (16) Printing and Reproduction of Recorded Media, technology-intensive industries, such as (26) Electronic Parts and Components, (30) Motor Vehicles and Parts, and basic industries, such as (20) Medical Goods, (22) Plastic Products, when the degree of export intensity increases, manufacturers will be more likely to use the Internet. On the contrary, for basic industries, such as (18) Chemical Material, (19) Chemical Products, (21) Rubber Products, when the degree of export intensity increases, manufacturers will be less likely to use the Internet.

The coefficient of $size_i$ has a positive effect on the manufacturers' decision to use the Internet for most 2-digit industries, whereas the coefficient of $group_i$ has a positive and significant effect on the decision to use the Internet for most 2-digit industries.

In the following, we will present the total marginal effects of each explanatory variable on the extent to which manufacturers use the Internet for the individual 2-digit industries in Table 9. Of these 26 industries, seven 2-digit industries significantly reject the null hypothesis that β_λ = 0 at the 10% level of significance, with bootstrapping standard deviations, namely (08) Food, (09) Beverages, (22) Plastic Products, (28) Electrical Equipment, (29) Machinery and Equipment, (30) Motor Vehicles and Parts, and (32) Furniture. As these industries are affected by the problem of sample selection bias, it is necessary to correct for such bias.

In the following paragraphs, we present the marginal effects, as given in equations (7) and (8). In terms of industrial agglomeration, $GHHI_{jki}$, for traditional industries, the largest value is 2.3761 for (09) Beverages, while the smallest is -1.4581 for (32) Furniture; for technology-intensive industries, the largest value is 5.5503 for (27) Plastic Products, while the smallest is -12.6278 for (30) Motor Vehicles and Parts; for basic industries, the largest value is 21.886 for (20) Medical Goods, while the smallest is -1.3668 for (21) Rubber Products.

Regarding the marginal effects of industrial concentration, HHI_{ji}, for traditional industries, the largest is 0.1812 for (13) Leather, Fur and Related Products, while the smallest is -0.1393 for (08) Food; for technology-intensive industries, the largest value is 0.2549 for (26) Electronic Parts and Components, while the smallest is -0.2781 for (29) Machinery and Equipment; for basic industries, the largest value is 2.3671 for (22) Plastic Products, while the smallest is -0.2068 for (24) Basic Metal.

For the marginal effects of export intensity, $export\_rate_i$, for traditional industries, the largest value is 0.5523 for (08) Food, while the smallest is -0.0095 for (13) Leather, Fur and Related Products; for technology-intensive industries, the largest is 0.4583 for (27) Plastic Products, while the smallest is 0.0221 for (26) Electronic Parts and Components; for basic industries the largest is 0.5053 for (21) Rubber Products, while the smallest is 0.0393 for (19) Chemical Products.

Regarding the marginal effects of geographic location, $city_i$, for traditional industries, the largest value is 0.0266 for (08) Food, while the smallest is -0.0018 for (11) Textiles Mills; for technology-intensive industries, the largest is 0.0527 for (26) Electronic Parts and Components, while the smallest is -0.0249 for (27) Plastic Products; for basic industries, the largest is 0.0578 for (21) Rubber Products, while the smallest is -0.0216 for (24) Basic Metal.

For the marginal effects of the manufacturer's scale of operations, $size_i$, for traditional industries, the largest value is 0.0029 for (09) Beverages; for technology-intensive industries, the largest is 0.0002 for (27) Plastic Products and (28) Electrical Equipment; for basic industries, the largest is 0.0015 for (22) Plastic Products.

With respect to the marginal effects of manufacturers' expenditure on computer equipment, $computer1_i$, for traditional industries, the largest value is 17.4643 for (11) Textiles Mills, while the smallest is -0.0075 for (13) Leather, Fur and Related Products; for technology-intensive industries, the largest is 6.2498 for (31) Other Transport Equipment, while the smallest is -5.6547 for (30) Motor Vehicles and Parts; for basic industries, the largest is 139.043 for (24) Basic Metal, while the smallest is -5.4236 for (21) Rubber Products.

Regarding the marginal effects of the manufacturers' expenditure on computer equipment within the same industry and region, $computer2_{jki}$, 0.0045 for (15) Pulp, Paper and Paper Products, 0.0025 for (27) Plastic Products, and 0.0008 for (24) Basic Metal, are the largest values for the traditional, technology-intensive, and basic industries, respectively.

6. Conclusion

Taiwan has long been hailed as a world leader in the development of global innovation and industrial clusters. In this paper, we used Taiwanese manufacturing census data compiled by the Directorate-General of Budget, Accounting and Statistics of the Executive Yuan for the year 2006, to examine the factors influencing the extent to which manufacturers use the Internet. When we consider total expenditure on internet usage, an actual figure is observed only if the firm uses the Internet, which leads to the problem of sample selection. In order to correct the problem of selection bias, this paper uses the Heckman selection model and two-stage estimation procedure to obtain estimates of the parameters of the sample selection model.

In order to improve the effectiveness of estimation, we use the bootstrapping approach to estimate the sample variances. The empirical results show that, regardless of whether bootstrapping is used, the Mill's lambda test statistic rejects the null hypothesis that β_{λ} = 0 at the 1% level of significance for the aggregated full industry, and 7 of 26 industries reject the null hypothesis that β_{λ} = 0at the 10% level of significance. Therefore, the problem of sample selection bias needs to be corrected.

The primary conclusions of the empirical analysis are as follows:

(1) The manufacturer's decision to use the Internet is influenced by five factors, namely the degree of industrial concentration, export intensity, geographical location, manufacturer's size of operations, and the independence of operations: (1) As Taiwan largely consists of manufacturers with independent operations, it is not

surprising that the likelihood of such manufacturers using the Internet is relatively high, with the manufacturers' independence of operations having the greatest impact; (2) The manufacturers' export intensity indicates that a greater reliance of manufacturers on exports, the greater is the export intensity, and the greater the need to use the Internet to communicate with overseas customers; (3) The degree of industrial concentration is such that, the greater is the competition faced by manufacturers, the more that they will be inclined to use the Internet to increase their ability to compete with other manufacturers; (4) Manufacturers who are located in mon-urban areas would be more likely to use the Internet for business than those located in city areas; and (5) larger firms would be more likely to use the Internet for business than smaller firms, even though the impact of the degree of industrial agglomeration on manufacturers' use of the Internet is not significant.

(2) The extent to which manufacturers' use of the Internet is primarily influenced by three factors, namely the degree of industrial agglomeration, geographical location, and the contagion effect. While the impact of the degree of industrial agglomeration on the manufacturers' use of the Internet is not significant, the effect on the extent to which manufacturers use the Internet is significant and negative. Therefore, the extent of industrial agglomeration does not affect whether manufacturers will use the Internet, but it does affect the extent to which manufacturers who already use the Internet will continue to do so. The results suggest that there exists a substitution relationship between the agglomeration of localization and the extent to which manufacturers use the Internet, indicating that Internet technology hasthe importance of the "distance" factor.

(3) The industrial agglomeration variable shows a negative marginal effect on the

extent to which manufacturers use the Internet, indicating there exists a substitution relationship between the two. Such results confirm the findings in Kauffman and Kumar (2007), who used U.S. information technology-related manufacturing and service industry data, and Galliano and Roux (2008), who used French manufacturing data.

(4) The more competitive is the industry, the more will manufacturers need to use the Internet to communicate and trade with other entities to increase their competitiveness. The empirical findings agree with those of Galliano and Roux (2008) and Galliano et al. (2011), who used French manufacturing industry data.

(5) Export intensity has the greatest marginal effect on the extent to which manufacturers use the Internet, indicating that international competition has a relatively large influence on the extent of Internet usage. The second and third largest positive marginal effects on the extent to which manufacturers use the Internet are the manufacturers' expenditure on computer equipment and the contagion effect, though the magnitudes of both are relatively small.

(6) As the industries are different, the empirical results for the individual industries based on the 2-digit level classifications are quite varied. In terms of the degree of industrial agglomeration, (09) Beverages and (32) Furniture have the largest positive at 2.376 and smallest negative at -1.458 marginal effects on the extent to which the manufacturers use the Internet for traditional industries; (27) Plastic Products and (30) Motor Vehicles and Parts have yjr largest positive at 5.550 and smallest negative at -12.628 marginal effects for technology-intensive industries; and (20) Medical Goods and (21) Rubber Products have the largest positive at 21.886 and smallest negative at -1.367 marginal effects for basic industries.

(7) The marginal effects of localized agglomeration on the extent to which manufacturers use the Internet also vary. The largest positive and smallest negative values for traditional industries are 0.0266 for (08) Food, and -0.0018 for (11) Textiles Mills; the largest and smallest values for technology-intensive industries are 0.0527 for (26) Electronic Parts and Components, and -0.0249 for (27) Plastic Products; the largest and smallest values for basic industries are 0.0578 for (21) Rubber Products, and -0.0216 for (24) Basic Metal.

(8) Industries with a higher degree of export intensity and greater reliance on exports have a higher degree of Internet usage among those manufacturers that use the Internet. The empirical results indicate that exports of export-oriented industries such as (08) Food, (26) Electronic Parts and Components, and (22) Plastic Products have the largest marginal effects for traditional, technology-intensive and basic industries in Taiwan, respectively.

Table 1
Industry Associated 2-digit Codes and Numbers of Firms

	code	2-digit Industries	Numbers of Firms
Traditional Industries	08	Food	6,165
	09	Beverages	644
	11	Textiles Mills	6,439
	12	Wearing Apparel and Clothing Accessories	4,084
	13	Leather, Fur and Related Products	1,870
	14	Wood and Bamboo Products	2,849
	15	Pulp, Paper and Paper Products	3,605
	16	Printing and Reproduction of Recorded Media	9,439
	23	Non-metallic Mineral Products	3,677
	32	Furniture	2,849
	33	Manufacturing Not Elsewhere Classified	5,435
Technology-intensive Industries	26	Electronic Parts and Components	6,023
	27	Computers, Electronic and Optical Products	3,717
	28	Electrical Equipment	6,198
	29	Machinery and Equipment	18,545
	30	Motor Vehicles and Parts	3,580
	31	Other Transport Equipment	2,905
	34	Repair and Installation of Industrial Machinery and Equipment	3,907
Basic Industries	17	Petroleum and Coal Products	229
	18	Chemical Material	1,549
	19	Chemical Products	2,304
	20	Medical Goods	543
	21	Rubber Products	1,756
	22	Plastic Products	11,012
	24	Basic Metal	4,710
	25	Fabricated Metal Products	39,047
	Total	**All Manufacturing Industries**	**153,081**

Table 2

Variable Definitions

Variables	Description
Dependent variables	
y_i	The extent to which firm *i* uses the Internet = (Internet purchase amount + Internet sales amount) / total sales (unit: 100%)
z_i	z_i = 1, if firm *i* uses internet equipment for business information z_i = 0, otherwise
Independent variables	
HHI_{ji}	Herfindahl-Hirschman Index for industry *j* in which firm *i* belongs
$CR4_{ji}$	Top Four Firms Concentration Index for industry *j* in which firm *i* belongs
$export\_rate_i$	Export rate for firm *i* = export value / total sales
$GHHI_{jki}$	Geographic Herfindahl-Hirschman lndex for industry *j* in region *k* to which firm *i* is located
$size_i$	Firm size (total number of employees for firm *i*)
$computer1_i$	Total expenditure on computer equipment for firm *i* (unit: T$1000)
$computer2_{jki}$	Total expenditures on computer equipment for industry *j* in region *k*, excluding expenditure of firm *i* (unit: NT$1000)
$city_i$	$city_i = 1$, if firm *i* is located in the city $city_i = 0$, if firm *i* is located in the county
$group_i$	$group_i = 1$, if firm *i* has no subsidiary (branch) $group_i = 0$, otherwise

Table 3
Descriptive Statistics

Variables (unit)	Mean	Std. Dev.	Min	Max
y_i	1.9998	43.2231	0	7153.077
z_i	0.6069	0.4884	0	1
HHI_{ji}	0.0322	0.0656	0.0020	1
$CR4_{ji}$	0.2053	0.1683	0.0407	1
$export\_rate_i$	0.0709	0.1669	0	1
$GHHI_{jki}$	0.0031	0.0239	0	0.4752
$size_i$	16.7994	113.8733	0	17,040
$computer1_i$	0.0029	0.2871	0	99.2
$computer2_{jki}$	0.4011	6.4387	0	1264.754
$city_i$	0.1845	0.3879	0	1
$group_i$	0.9327	0.2505	0	1

Table 4

Correlation Coefficients

Variables	HHI_{ji}	$CR4_{ji}$	$GHHI_{jki}$	$export\_rate_i$	$city_i$	$computer1_i$	$computer2_{jki}$	$size_i$
HHI_{ji}	1							
$CR4_{ji}$	0.8518	1						
$GHHI_{jki}$	-0.0078	0.0011	1					
$export\_rate_i$	0.1558	0.178	0.0413	1				
$city_i$	0.0261	0.029	-0.0428	0.0093	1			
$computer1_i$	0.0028	0.0066	-0.0008	-0.0032	-0.0002	1		
$computer2_{jki}$	0.0077	0.0155	0.014	-0.0149	0.001	0.0401	1	
$size_i$	0.0803	0.0863	0	0.1729	0.0072	0.001	-0.0062	1

Table 5

Selection-corrected Internet Use Model (with HHI) for All Industries

Variables	Intensity of internet use (y_i)	Select (z_i)
HHI_{ji}	0.321 (3.660) [2.727]	-1.369 (0.065)*** [0.067]***
$export\_rate_i$	1.110 (1.284) [1.336]	3.807 (0.207)*** [0.057]***
$GHHI_{jki}$	-2.792 (1.057)*** [5.238]	0.051 (0.237) [0.201]
$city_i$	0.868 (0.523)* [0.378]**	-0.201 (0.013)*** [0.010]***
$computer1_i$	0.240 (51.880) [0.432]	-
$\mathrm{computer2_{jki}}$	0.068 (0.119) [0.019]***	-
$size_i$	-	0.003 (0.001)*** [0.0002]***
$group_i$	-	58.543 (16.397)*** [0.005]***
$constant$	2.702 (0.755)*** [0.881]***	-57.606 (16.400)***
Mills lambda (λ)	-7.404 (2.595)*** [2.187]***	
# of observations	153081	
# of censored observation	31924	
Wald Chi2(df)	2458.61(31)	

Notes: Bootstrap standard errors are in parentheses and standard errors without bootstrapping are in brackets. The asterisks ***, ** and * denote significance at the 1%, 5% and 10% levels, respectively. 2-digit industry dummies are included in the empirical model to control for heterogeneity, but are not reported in the tables for reasons of space.

Table 6

Selection-corrected Internet Use Model (with CR4) for All Industries

Variables	Intensity of internet use (y_i)	Select (z_i)
$CR4_{ji}$	4.214 (1.160)*** [1.240]***	-0.645 (0.028)*** [0.025]***
$export\_rate_i$	0.545 (1.143) [1.342]	3.813 (0.214)*** [0.057]***
$GHHI_{jki}$	-1.871 (1.064)* [5.247]	0.071 (0.203) [0.202]
$city_i$	0.917 (0.344)*** [0.377]**	-0.201 (0.011)*** [0.010]***
$computer1_i$	0.241 (55.104) [0.432]	-
$\mathrm{computer2}_{jki}$	0.069 (0.142) [0.019]***	-
$size_i$	-	0.004 (0.001)*** [0.0002]***
$group_i$	-	61.607 (22.335)*** [0.007]***
$constant$	1.912 (0.763)** [0.893]**	-60.585 (22.243)***
Mills lambda (λ)	-8.217 (2.444)*** [2.164]***	
# of observations	153081	
# of censored observation	31924	
Wald Chi2(df)	1976.69(31)	

Note: See footnotes to Table 5.

Table 7

Marginal Effects of the Internet Use Model for All Industries (unit: %)

Variables	Intensity of Internet Use (1)	Intensity of Internet Use (2)
$GHHI_{jki}$	-0.0245	-0.0133
HHI_{ji}	-0.0902	
$CR4_{ji}$		-0.0070
$export\_rate_i$	0.2708	0.2963
$city_i$	-0.0051	-0.0062
$size_i$	0.0002	0.0003
$computer1_i$	0.0024	0.0024
$computer2_{jki}$	0.0007	0.0007

Table 8. Selection-corrected Internet Use Model (with HHI) for 2-digit Industries

Variables	(8)		(9)		(11)		(12)		(13)		(14)		(15)	
	y_i	z_i	y_i	z_i	y_i	z_i	y_i	z_i	y_i	z_i	y_i	z_i	y_i	z_i
$GHHI_{jki}$	-15.98	22.91	-41.84	206.89	-8.21	10.75	-0.35	0.23	-11.30	35.24	-31.03	98.98	37.49	-193.30
	(3.49)***	(4.76)***	(17.83)**	(202.05)	(8.77)	(2.38)***	(0.20)*	(0.26)	(4.50)**	(17.95)**	(68.63)	(45.62)**	(33.70)	(33.33)***
HHI_{ji}	10.00	-8.14	-0.28	-3.84	-3.23	3.81	1.69	-1.80	17.47	-16.81	9.78	-10.51	-1.03	3.66
	(2.86)***	(0.99)***	(0.87)	(33.51)	(4.10)	(0.88)***	(1.56)	(0.56)***	(9.40)*	(4.99)***	(24.98)	(8.32)	(1.09)	(1.39)***
$export\_rate_i$	0.84	18.50	0.76	4.26	3.99	21.51	1.23	12.96	-0.27	17.69	7.31	676.48	0.40	916.90
	(1.20)	(396.24)	(0.37)**	(246.23)	(5.57)	(18.24)	(1.41)	(303.46)	(0.20)	(705.29)	(5.81)	(192.02)***	(0.43)	(270.73)***
$city_i$	-0.73	1.15	-0.22	244.99	-0.09	-0. 21	0.38	0.46	0.24	-0.12	0.60	-0.20	0.16	-0.55
	(0.22)***	(0.16)***	(0.07)***	(160.83)	(1.20)	(0.06)***	(0.40)	(0.06)***	(0.32)	(0.12)	(0.49)	(0.10)**	(0.09)*	(0.08)***
$size_i$		0.05		0.22		0.00004		0.003		0.01		0.01		0.01
		(0.02)***		(0.13)*		(0.002)		(0.002)		(0.02)		(0.004)*		(0.01)
$computer1_i$	24.42		-0.04		1746.43		87.35		-0.75		26.43		50.92	
	(18.16)		(11.90)		(1553.83)		(83.22)		(7.52)		(110.67)		(104.95)	
$computer2_{jki}$	0.02		-0.22		-5.02		0.07		-1.91		-0.89		0.45	
	(0.46)		(0.24)		(5.75)		(0.12)		(0.67)***		(0.67)		(0.97)	
$group_i$		91.68		313.64		7.09		12.05		25.35		14.07		16.86
		(35.31)***		(205.49)		(1.75)***		(3.17)***		(30.53)		(4.48)***		(7.04**
constant	0.74	-90.28	0.20	-312.06	-0. 05	-6.78	-0.41	-11.47	0.13	-24.39	0.05	-12.82	0.27	-15.63
	(0.15)***	(35.35)**	(0.08)***	(205.57)	(1.42)	(1.75)***	(0.47)	(3.16)***	(0.13)	(30.56)	(0.75)	(4.51)***	(0.20)	(7.08)**
# of observations	6165		644		6439		4084		1870		2849		3605	
# of censored	1081		106		1783		936		306		329		595	
Mills Lambda	-2.98		-1.36		-2.07		0.97		0.16		-0.51		-1.02	
	(1.06)***		(0.81)*		(2.57)		(0.89)		(0.61)		(1.52)		(0.81)	
Wald Chi2(ddl)	34.94(6)		12.71(6)		3.76(6)		7.39(6)		11.27(6)		12.82(6)		12.27(6)	

Table 8. Selection-corrected Internet Use Model (with HHI) for 2-digit Industries (cont.)

	(16)		(18)		(19)		(20)		(21)		(22)	
	y_i	z_i	y_i	z_i	y_i	z_i	y_i	z_i	y_i	z_i	y_i	z_i
$GHHI_{jki}$	24.58 (13.40)*	-40.82 (3.20)***	-174.60 (205.59)	12.05 (160.45)	86.17 (299.12)	9.18 (128.80)	2188.6 (1231.76)*	4.75 (154.72)	139.65 (346.79)	17.62 (38.59)	292.81 (334.88)	-33.14 (11.62)***
HHI_{ji}	-0.03 (1.81)	0.08 (3.21)	-3.73 (2.73)	0.51 (1.71)	8.43 (4.20)**	1.13 (1.82)	53.80 (44.68)	-0.22 (9.92)	0.12 (5.97)	-0.97 (0.91)	89.59 (82.23)	25.45 (9.47)***
$export\_rate_i$	4.46 (4.05)	1155.05 (504.82)**	4.85 (2.04)**	-3.46 (0.37)***	2.78 (1.53)*	-1.87 (0.25)***	7.47 (5.96)	1662.65 (722.46)**	4.08 (3.97)	-2.96 (0.21)***	-1.56 (0.98)	1.32 (0.63)**
$city_i$	-0.01 (0.05)	-0.13 (0.03)***	-0.52 (0.47)	0.07 (0.38)	-0.09 (0.34)	0.12 (0.15)	0.37 (2.32)	-0.24 (0.83)	2.11 (1.64)	-0.24 (0.17)	0.49 (0.42)	-0.31 (0.05)***
$size_i$		0.01 (0.003)***		0.06 (0.02)***		0.04 (0.02)*		0.02 (0.03)		0.12 (0.03)***		0.03 (0.01)***
$computer1_i$	130.53 (37.29)***		-21.04 (75.35)		-80.94 (165.92)		-40.09 (636.13)		-542.36 (260.34)**		380.61 (166.22)**	
$computer2_{jki}$	-0.02 (0.02)		-0.58 (1.26)		-0.11 (0.48)		0.03 (2.20)		-0.05 (1.02)		-0.11 (0.10)	
$group_i$		11.30 (1.73)***		218.90 (65.70)***		39.08 (16.68)**		18.50 (14.21)		304.01 (116.14)***		43.54 (9.94)***
$constant$	-0.08 (0.30)	-10.74 (1.74)***	1.46 (0.43)***	-216.90 (65.65)***	0.66 (0.40)*	-37.20 (16.79)**	-1.57 (2.33)	-17.10 (14.72)	0.48 (1.23)	-302.68 (116.18)***	2.02 (0.72)***	-42.42 (9.95)***
# of observations	9439		1549		2304		543		1756		11012	
# of censored observation	2790		455		499		142		249		1487	
Mills Lambda	0.02 (0.53)		1.05 (6.20)		0.63 (5.63)		-22.79 (19.08)		15.73 (11.45)		-7.49 (2.42)***	
Wald Chi2(ddl)	20.41(6)		8.20(6)		8.08(6)		11.80(6)		6.58(6)		10.16(6)	

Table 8. Selection-corrected Internet Use Model (with HHI) for 2-digit Industries (cont.)

	(23)		(24)		(25)		(26)		(27)		(28)	
	y_i	z_i	y_i	z_i	y_i	z_i	y_i	z_i	y_i	z_i	y_i	z_i
$GHHI_{jki}$	-2.25 (1.71)	0.35 (0.64)	-105.17 (55.66)*	-0.78 (3.43)	-57.28 (37.92)	15.17 (2.10)***	-81.58 (79.95)	-10.87 (2.43)***	535.71 (532.84)	3.88 (12.03)	-40.61 (7.48)***	2.74 (7.13)
HHI_{ji}	0.75 (1.98)	1.19 (0.62)*	-20.43 (12.75)	-0.05 (0.25)	16.35 (6.93)**	-4.71 (0.31)***	25.53 (8.95)***	0.18 (0.31)	-13.55 (19.57)	-0.38 (0.30)	-4.56 (1.72)***	-0.23 (1.16)
$export\_rate_i$	3.63 (2.25)	8.69 (388.00)	-8.31 (4.45)*	5.72 (359.95)	6.62 (2.76)**	68.80 (52.19)	6.00 (2.20)***	18.18 (1.61)***	7.98 (6.61)	7.59 (321.00)	0.41 (0.54)	11.06 (421.94)
$city_i$	-0.20 (0.25)	-0.25 (0.11)**	-0.27 (1.18)	-0.37 (0.08)***	0.27 (0.93)	-0.09 (0.02)***	5.29 (6.75)	0.09 (0.08)	-3.00 (3.97)	0.10 (0.10)	0.22 (0.18)	-0.05 (0.08)
$size_i$		0.03 (0.01)***		0.01 (0.01)		0.003 (0.001)***		0.0001 (0.0003)		0.003 (0.002)**		0.01 (0.01)
$computer1_i$	141.03 (207.19)		13904.32 (8486.87)		52.32 (12.26)***		271.39 (305.89)		7.21 (3027.91)		-13.46 (58.69)	
$computer2_{jki}$	-0.38 (0.59)		0.08 (4.28)		0.04 (0.06)		-0.03 (0.02)		0.25 (5.80)		-0.03 (0.04)	
$group_i$		70.24 (26.97)***		12.51 (12.36)		10.52 (0.99)***		6.99 (3.33)**		18.74 (6.13)***		22.70 (12.14)*
$constant$	0.57* (0.33)	-69.12 (26.99)***	1.43 (1.49)	-11.41 (12.42)	1.90 (0.57)***	-9.93 (0.99)***	0.17 (0.50)	-6.59 (3.33)**	6.12 (5.45)	-17.80 (6.16)***	1.61 (0.31)***	-21.69 (12.16)*
# of observations	3677		4710		39047		6023		3717		6198	
# of censored	684		861		8496		1558		716		1065	
Mills Lambda	-1.11 (1.71)		-9.27 (7.41)		-0.89 (1.37)		2.20 (5.61)		-9.90 (21.52)		-3.51 (1.12)***	
Wald Chi2(ddl)	6.46(6)		5.48(6)		60.43(6)		31.83(6)		5.18(6)		34.19(6)	

Table 8. Selection-corrected Internet Use Model (with HHI) for 2-digit Industries (cont.)

	(29)		(30)		(31)		(32)		(33)	
	y_i	z_i	y_i	z_i	y_i	z_i	y_i	z_i	y_i	z_i
$GHHI_{jki}$	4.78 (6.10)	5.31 (1.16)***	-1262.78 (363.88)***	-58.65 (32.68)*	-93.60 (40.92)**	23.42 (9.37)**	-145.81 (81.75)*	-40.96 (19.08)**	-1.40 (24.87)	-19.64 (4.93)***
HHI_{ji}	-27.81 (2.28)***	7.32 (1.49)***	-1.70 (7.72)	0.90 (0.78)	-5.77 (8.66)	-0.67 (1.42)	-76.10 (42.31)*	15.18 (5.89)***	-21.96 (8.42)***	0.83 (1.43)
$export\_rate_i$	5.19 (0.82)***	99.56 (330.80)	25.11 (5.61)***	2270.49 (797.45)***	1.17 (1.11)	6.68 (231.68)	-2.66 (4.95)	524.98 (416.88)	1.36 (0.74)*	47.78 (647.64)
$city_i$	0.13 (0.21)	-0.27 (0.04)***	1.29 (1.94)	-0.36 (0.10)***	1.03 (0.60)*	-0.35 (0.09)***	0.47 (1.03)	-0.09 (0.11)	0.46 (0.36)	0.03 (0.06)
$size_i$		0.02 (0.01)***		0.01 (0.01)		0.01 (0.01)		0.01 (0.01)**		0.004 (0.004)
$computer1_i$	49.00 (47.17)		-565.47 (296.11)*		624.98 (339.93)*		34.54 (83.23)		-0.72 (76.01)	
$computer2_{jki}$	-0.13 (0.06)**		0.05 (0.46)		-0.14* (0.08)		-0.45 (0.79)		0.83 (1.09)	
$group_i$		25.93 (7.17)***		27.25 (13.43)**		24.66 (11.50)**		12.47 (2.35)***		9.96 (4.19)**
$constant$	1.55 (0.15)***	-25.33 (7.19)***	2.97 (0.92)***	-26.49 (13.48)**	1.21 (0.72)*	-23.84 (11.55)**	4.25 (2.30)*	-11.45 (2.38)***	1.67 (0.38)***	-9.06 (4.20)**
# of observations	18545		3580		2905		2849		5435	
# of censored	3076		686		521		367		780	
Mills Lambda	-0.87 (0.46)*		6.03 (2.59)**		-2.43 (1.60)		-14.30 (8.60)*		-0.53 (1.26)	
Wald Chi2(ddl)	169.23(6)		49.09(6)		28.81(6)		10.74(6)		44.93(6)	

Note: For reasons of space, we do not present (17) Petroleum and Coal Products Industry and (34) Repair and Installation of Industrial Machinery and Equipment in Tables 8 and 9. Moreover, some coefficients of the explanatory variables were not available for the corrected regression model.

Table 9

Marginal Effects of the Internet Use Model (with HHI) for 2-digit Industries (unit: %)

	Marginal Effects											
	(8)	(9)	(11)	(12)	(13)	(14)	(15)	(16)	(19)	(20)	(21)	(22)
$GHHI_{jki}$	0.5136	2.3761	-0.0431	-0.0042	-0.1265	-0.3103	0.3749	0.2458	0.8053	21.886	-1.3668	1.0123
HHI_{ji}	-0.1393	-0.0547	-0.0184	0.0223	0.1812	0.0978	-0.0103	-0.0003	0.0774	0.538	0.1534	2.3671
$export\_rate_i$	0.5523	0.0651	0.118	-0.0263	-0.0095	0.0731	0.0040	0.0446	0.0393	0.0747	0.5053	0.0608
$city_i$	0.0266	-	-0.0018	0.0026	0.0024	0.0060	0.0016	-0.0001	-0.0016	0.0037	0.0578	-0.0136
$size_i$	0.0014	0.0029	0	0	0	0	0	0	-0.0002	0	-0.018	0.0015
$computer1_i$	0.2442	-0.0004	17.4643	0.8735	-0.0075	0.2643	0.5092	1.3053	-0.8094	-0.4009	-5.4236	3.8061
$computer2_{jki}$	0.0002	-0.0022	-0.0502	0.0007	-0.0191	-0.0089	0.0045	-0.0002	-0.0011	0.0003	-0.0005	-0.0011
$group_i$	-	-	0.102	-0.0982	-0.0359	0	0	0	-0.2325	0	-	-

	Marginal Effects										
	(23)	(24)	(25)	(26)	(27)	(28)	(29)	(30)	(31)	(32)	(33)
$GHHI_{jki}$	-0.0187	-1.089	-0.5687	-0.7931	5.5503	-0.3503	0.0478	-12.6278	-0.6135	-1.4581	-0.0140
HHI_{ji}	0.0205	-0.2068	0.1622	0.2549	-0.1543	-0.0502	-0.2781	-0.017	-0.0669	-0.7610	-0.2196
$export\_rate_i$	0.1307	0.1907	0.0846	0.0221	0.4583	0.2295	0.0519	0.2511	0.1037	-0.0266	0.0136
$city_i$	-0.0047	-0.0216	0.0026	0.0527	-0.0249	0.0012	0.0013	0.0129	0.0053	0.0047	0.0046
$size_i$	0.0003	0.0003	0	0	0.0002	0.0002	0	0	0.0001	0	0
$computer1_i$	1.4103	139.0432	0.5232	2.7139	0.0721	-0.1346	0.4900	-5.6547	6.2498	0.3454	-0.0072
$computer2_{jki}$	-0.0038	0.0008	0.0004	-0.0003	0.0025	-0.0003	-0.0013	0.0005	-0.0014	-0.0045	0.0083
$group_i$	-	1.0142	0.0675	-0.0957	1.5564	0.7173	0.1314	0	0.5535	0	0.0108

Note: For the (18) Chemical Material industry, the marginal effect is not available.

References

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

Bain, J.S. (1968), *Industrial Organization*. New York: John Wiley & Sons, Inc.

Bertschek, I. and H. Fryges (2002), "The Adoption of Business-to-Business E-Commerce Empirical Evidence for German Companies", *ZEV Discussion Paper*, pp. 02-05.

Chang, C.-L. and L. Oxley. (2009), "Industrial Agglomeration, Geographic Innovation and Total Factor Productivity: The Case of Taiwan", *Mathematics and Computers in Simulation,* 79(9), 2787-2796.

Forman, C., A. Goldfarb, and S. Greenstein (2005a), "Geographic Location and the Diffusion of Internet Technology", *Electronic Commerce Research and Applications*, 4(1), 1-13.

Forman, C., Goldfarb, A., and S. Greenstein (2005b), How Do Industry Features Influence the Role of Locationon internet adoption?, *Journal of the Association for Information Systems*, 6(12), 383-408.

Forman, C, Goldfarb, A., and S. Greenstein (2005c) How Did Location Affect the Adoption of the CommercialInternet? Global Village vs. Urban Density, *Journal of Urban Economics*, 58(3), 389-420.

Galliano, D. and P. Roux (2008), "Organisational Motives and Spatial Effects in Internet Adoption and Intensity of Use: Evidence from French Industrial Firms", *Annals of Regional Science*, 42(2), 425-448.

Galliano, D., P. Roux, and N. Soulie (2011), "ICT Intensity of Use and the Geography of Firms", *Environment and Planning A,* 43(1), 67-86.

Giunta, A., and F. Trivieri (2007), "Understanding the Determinants of Information

Technology Adoption", *Applied Economics*, 39(10), 1325-1334.

Greene, W. (2003), "Sample Selectin Bias as a Specification Error: Comment", Econometrica, 49(3), 795-798

Greene, W. (2003), *Econometric Analysis*, New York: Macmillan.

Heckman, J.J. (1976), "The Common Sttructure of Statistical Models of Truncation, Sample Selection and Limited Dependent Variables and a Simple Estimator for Such Models", *Annals of Economic and Social Measurement*, 5(4), 475–492.

Heckman, J.J. (1979), "Sample Selection Bias as a Specification Error", *Econometrica*, 47(1), 153-161.

Kauffman, R., and A. Kumar (2007), "The New Economic Geography of IT Industries The Impacts of the Internet on Their Market Linkages and Agglomeration", *Proceedings of 28th Annual International Conference on Information Systems (ICIS)*, Montreal, Canada.

Kolko, J. (1999), "The Death of Cities? The Death of Distance Evidence from the Geography of Commercial Internet Usage", Mimeo, Harvard University.

Lal, K. (1999), "Determinants of the Adoption of Information Technology: A Case Study of Electrical and Electronic Goods Manufacturing Firms in India", *Research Policy*, 28(7), 667-680.

Lewis, H.G. (1974), "Comments on Selectivity Biases in Wage Comparisons", *Journal of Political Economy*, 82, 1145–1155.

Maddala, G.S. (1983), *Limited-Dependent and Qualitative Variables in Economics*, New York: Cambridge University Press.

Mansfield, E. (1963a), The Speed of Response of Firms to New Techniques, *Quarterly Journal of Economics*, 77(2), 290-311.

Mansfield, E. (1963b), "Intrafirm Rates of Diffusion of an Innovation", *Review of*

Economics and Statistics, 45(4), 348–359.

Saloner, R. and J.M. Sheppard (1995), “Adoption of Technologies with Network Effects: An Empirical Examination of the Adoption of Automated Teller Machines”, *Rand Journal of Economics*, 26(3), 479-501.

Schwab, K. and X. Sala-i-Martin (2009), *The Global Competitiveness Report, 2009-2010*, Geneva, Switzerland: World Economic Forum.

Schwab, K. and X. Sala-i-Martin (2010), *The Global Competitiveness Report, 2010-2011*, Geneva, Switzerland: World Economic Forum.

Schwab, K. and X. Sala-i-Martin (2011), *The Global Competitiveness Report, 2011-2012*, Geneva, Switzerland: World Economic Forum.

Schwab, K. and X. Sala-i-Martin (2012), *The Global Competitiveness Report, 2012-2013*, Geneva, Switzerland: World Economic Forum.