A Statistical Analysis of Industrial Penetration and Internet Intensity in Taiwan*

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

Michael McAleer
Department of Finance
Asia University, Taiwan
and
Discipline of Business Analytics
University of Sydney Business School, Australia
and
Econometric Institute, Erasmus School of Economics
Erasmus University Rotterdam, The Netherlands
and
Department of Quantitative Economics
Complutense University of Madrid, Spain
and
Institute of Advanced Sciences
Yokohama National University, Japan

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

EI2018-04

Revised: January 2018

***** For financial support, the first author wishes to thank the National Science Council, Ministry of Science and Technology (MOST), Taiwan, and the second author acknowledges the Australian Research Council and the National Science Council, Ministry of Science and Technology (MOST), Taiwan..

Abstract

This paper investigates the effect of industrial penetration (geographic concentration of industries) and internet intensity (the proportion of enterprises that use the internet) for Taiwan manufacturing firms, and analyses whether the relationships are substitutes or complements. The sample observations are based on 153,081 manufacturing plants, and covers 26 two-digit industry categories and 358 geographical townships in Taiwan. The Heckman selection model is used to accommodate sample selectivity for unobservable data for firms that use the internet. The empirical results from two-stage estimation show that: (1) a higher degree of industrial penetration will not affect the probability that firms will use the internet, but will affect the total expenditure on internet intensity; (2) for two-digit SIC industries, industrial penetration generally decreases the total expenditure on internet intensity; and (3) industrial penetration and internet intensity are substitutes.

Keywords: Industrial penetration, Internet intensity, Sample selection, Incidental truncation.

JEL: D22, L60.

1. Introduction

With the arrival of the internet era, internet intensity by business enterprises has continued to increase in recent years. Furthermore, the proliferation of internet technology has 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 thereby reduced the costs of relaying information over long distances, making it easier for businesses to communicate with each other over long distances. Taiwan's overall industrial internet intensity (that is, the proportion of enterprises that use the internet) has increased from 62% in 2002, to 79% in 2003, and to 94.3% in 2010. 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 increasingly rapidly in the manufacturing industry and distribution services. As internet intensity continues to develop and information is exchanged increasingly rapidly, the management information systems of businesses are becoming increasingly 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 that businesses have lowered their costs of communicating and collecting information. Because of the increased convenience that the internet has brought in enabling firms to communicate with each other and in reducing the cost of transportation, as well as an abundance of resources that has further speeded up the exchange of information, the "distance" factor is clearly no longer as important as it was in the past. online purchases

According to the 2009-2013 Global Competitiveness Report compiled by the

[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

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, with Taiwan being hailed as a model for the development of global innovation and industrial clusters. Despite its ranking falling to 6th and 3rd in the following two years, the state of its cluster development enabled Taiwan to receive a score of 5.5 (out of a possible maximum of 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 well-developed (Schwab and Sala-i-Martin, 2009, 2010, 2011, 2012).

In previous research literature, many scholars have focus on R&D and new technology (Audretsch and Feldman, 1996；Bertscheck and Fryges, 2002; Chang and Oxley, 2009) and also sone scholars have examined the relationship between internet intensity and urbanization economics (Forman et al. (2005a, b, c), as well as a link between computers and productivity (Atrostic and Nguyen, 2005) but there have been quite few researches on the relationship between internet intensity and industrial penetration. Moreover, when we consider that the total expenditure on internet intensity, an actual figure is observed only if the firm is use the internet that will cause the problem of sample selection. For this reason, the purpose of this paper is to include the effect of sample correction and examine whether a relationship exists between penetration (Geographic concentration of industries) and internet intensity, and further to look at the factors determining the extent of the internet's influence. Following this Introduction, the literature on the influence of the factors related to internet intensity is reviewed in Section 2. In Section 3, we introduce the selection bias model and Heckman's two-step efficient estimation. A description of the sample and variables follows in Section 4.

This is followed by the empirical results in Section 5, and the Conclusion in Section 6.

2. Firm's Internet Intensity and Geographical Concentration

Forman et al. (2005a) proposed three related theories to the relationship between internet technology and urban penetration, namely, (1) global village theory, (2) urban density theory, and (3) industry composition theory. The Global village theory suggest that the new network technologies would help break down the barriers between individuals and groups. Internet technology can make up for the disadvantages faced by manufacturers due to their being located far away from the city's center of economic activity, and for this reason there is a substitutionary relationship exists between the adoption of internet technology and urban penetration.

The urban density theory suggests that as the density and scale of urbanization increase, the costs borne by manufacturers using internet technology will be reduced. In other words, if the manufacturer is located in the city center, a reduction in the cost of using internet technology will increase internet intensity, so that a complementary relationship exists between the adoption of internet technology and urban penetration.

The industry composition suggests that when the density and scale of urban areas increase, the benefits that manufacturers derive from using the internet will increase. Before network technology began to be widely used, 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. That is to say, the demand for the internet increased with the scale of urbanization. For this reason, a complementary relationship exists between the intensity of internet technology and urban penetration.

Forman et al. (2005a) use U.S. data to examine the relationship between internet intensity and urbanization and find that when the number of manufacturers in leading industries in urban areas increases, this will cause internet intensity in such regions to increase, indicating that the use of the internet will be enhanced as the scale of urbanization increases, that is, a complementary relationship exists between internet intensity and urban penetration. Later, Forman et al. (2005b) compare the influence of the location of enterprises and industrial penetration on internet intensity for the information intensity and the information-producing manufacturing industries and find that in the areas in which manufacturers are located, the larger the scale of industrial penetration, the more that the manufacturers use the internet. A similar result from U.S. businesses data from Kolko (1999) also indicated a complementary relationship between the internet intensity 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). Kauffman and Kumar (2007) test three hypotheses: (1) internet intensity reduce the market linkages; (2) whether the effects of internet intensity on market linkages will be the same for IT-related industry and information technology-related service industries; and (3) whether the effects of these market linkages in urban and non-urban areas will be the same. Their results indicate that internet intensity 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 effect of internet intensity in terms of the extent of its impact on IT-related manufacturing and information technology-related services is not significantly different.

Galliano and Roux (2008) used a French manufacturers' sample survey data for the year 2002 to examine the behavior of firms in the e-commerce industry in terms of their use of “Information and Communications Technology (ICT).” Their empirical research indicates that for those manufacturers located in the countryside, the extent to which they used the internet was lower than that for their counterparts in the urban areas. Moreover, for those industries for which there was a higher degree of penetration, the less that the manufacturers used the internet, which exist a substitutionary relationship between the extent of internet intensity and penetration.

Lal (1999) uses survey data for the year 1994 to investigate the factors affecting the manufacturers’ use of the internet for India manufacturing industry. Based on the extent to which the sampled firms used IT technology (IT), Lal 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, and referred to four categories of factors that affected internet intensity: (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 by manufacturers according to whether they (1) had not used B2B internet technology, (2) had used B2B internet technology, and (3) had extensively used B2B internet technology. They used factors which had been deemed in the past literature to have affected the manufacturer's adoption of new technologies, including the scale of the manufacturer's operations, the age of the plant, human capital and international competitive pressure, as well as variables that had not been considered in the earlier literature, such as electronic data interchange (EDI), which can be regarded as a precursor to B2B electronic commerce, and the bandwagon effect or herd behavior, and so on.

Bertschek and Fryges (2002) found that the scale of the 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; that the probability that manufacturers that had used EDI technology in the past would extensively use B2B technology in the future was extremely high; and that the more that other manufacturers within the same industry used internet technology, the greater the likelihood that they themselves would use new technologies.

Giunta and Trivieri (2007) looked into the factors determining the use of information technology (IT) by SMEs (Small and Medium-sized Enterprise) in Italy's manufacturing industry. Using sample survey data for 17,000 small and medium-sized firms covering the period from July 2001 to February 2002 and by focusing on the extent to which the manufacturers used information technology (IT), they categorized the manufacturers into those that: (1) did not use information technology, (2) had low use of information technology, (3) had medium use of information technology, and (4) had high use of information technology. They found that the factors that significantly affected the manufacturer's use of information technology 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, the more that enterprises within the same industry or geographical area used internet technology, the greater the contagion effect resulting from the internet technology, with there being a significant positive impact on the extent to which the enterprises used the internet. These empirical results lend support to the theories put forward by Mansfield (1963a, 1963b) and Saloner and Sheppard (1995).

As research literature above, many researchers focus on the problems associated with internet intensity related to urbanization, but with few studies looking into the relationship between industrial penetration and the extent to which firms use the internet. Therefore, this article will focus on the issue of internet use and industrial penetration.

3. Heckman Selection Model

Manufacturing firms may make decisions to use the internet and to purchase raw materials and components on line simultaneously, possibly leading to sample selection bias. Some enterprises that purchase online are a subset of manufacturing firms, forming a non-randomly selected sample from manufacturing firms, so that observations on the amount of internet purchases taken, and the corresponding firm specific characteristics, are available only for those who use the internet to purchase raw materials and components. Therefore, a manufacturing firm that uses the internet to purchase raw materials and components on line has a different preference structure from a non-user.

In order to draw conclusions about the larger population of all manufacturing firms in Taiwan, and not just the subpopulation of manufacturing firms from which the firm reports the internet purchase data are taken, the Heckman (1979) two-stage estimation procedure for a continuous decision variable can be used to incorporate the amount of internet purchases and the decision to join internet purchases (Lewis 1974; Heckman 1976, 1979; Greene, 2003). This method assumes the decisions to use the internet and purchase raw materials on line are made simultaneously (that is, the error terms of the two equations are correlated). It is assumed that zero observations represent the decision not to use the internet to purchase materials, so no individual firm is observed at the standard corner solution. Therefore, the demand curve for the internet purchaser is established only over manufacturing firms that have reports of internet purchases online. All non-users are assumed to not want to use the internet purchase mechanism, so firms that do not use the internet will not influence the demand curve for purchases online (Blaylock and Blissard, 1992).

In order to correct the problem of selection bias, this paper use Heckman selection model (Lewis 1974; Heckman 1976, 1979; Greene, 2003)，which assumes that there exists an underlying regression relationship, as given below:

Regression equation：

$$y_i = \boldsymbol{x}_i'\beta + u_{yi} \quad , \quad i = 1,2,\dots,\mathrm{n} \qquad (1)$$

$$u_{yi} \sim \mathrm{N}\left(0, \sigma_y^2\right)$$

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 $(\boldsymbol{\omega}_i')$ are the variables thought to determine whether dependent variable y_i is observed or unobserved (selected or not selected). So the selectin equation can be given as:

Selection equation：

$$z_i^* = \boldsymbol{\omega}_i'\gamma + u_{zi}\ ,\ i = 1,2,\dots,n \qquad (2)$$

$$u_{zi} \sim N(0,1)$$

$$corr(u_{yi}, u_{zi}) = \rho$$

When $\rho \neq 0$, the Ordinary Least Square (OLS) estimation applied to the equation (1) yield an 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 and regression model as follows:

Selection mechanism:

$$z_i = \boldsymbol{\omega}_i'\gamma + u_{zi} = 1, \text{ if } z_i^* > 0$$

$$z_i = \boldsymbol{\omega}_i'\gamma + u_{zi} = 0, \text{ otherwise} \qquad (3)$$

$$\text{prob}(z_i = 1|\omega_i) = \Phi(\boldsymbol{\omega}_i'\gamma) \text{ and}$$

$$\text{prob}(z_i = 0|\omega_i) = 1 - \Phi(\boldsymbol{\omega}_i'\gamma)$$

where $\Phi(\cdot)$ is the standard normal. cdf[3]

Regression (or observation) equation:

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

$$corr(u_{yi}, u_{zi}) = \rho$$

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

---

[3] Cumulative distribution function

In the equation (3), the selection equation is estimated by maximum likelihood (for details, see Maddala, 1983) as an independent probit model to determine the decision to join using the available information. However, Heckman's (1979) two-step estimation procedure is usually used for both the selection mechanism and regression model estimations. The first step estimates the selection equation by maximum likelihood to obtain an estimate of γ in equation (3) and 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 least squares to obtain estimates of β and $\beta_\lambda = \rho\sigma_y$. Green (1981, 2003) provides the statistical proof for consistency of the estimators of the individual parameters ρ and σ_y^2 (see Greene, 1981, 2003).

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

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

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

where $\alpha_{\mathrm{z}} = -\boldsymbol{\omega}_i'\gamma/\sigma_{\mathrm{z}}$, $\lambda_i(\alpha_{\mathrm{z}}) = \emptyset(\alpha_{\mathrm{z}})/[1 - \Phi(\alpha_{\mathrm{z}})]$, and $\delta_i(\alpha_{\mathrm{z}}) = \lambda_i(\alpha_{\mathrm{z}})[\lambda_i(\alpha_{\mathrm{z}}) - \alpha_{\mathrm{z}}]$，$0 < \delta_i < 1$. $\lambda_i(\alpha_{\mathrm{z}})$ is called the inverse Mill's ratio，$\emptyset(\cdot)$ is the standard normal pdf，and $\Phi(\cdot)$is the standard normal cdf.

[4] The theorem of moments of the incidentally truncated bivariate normal distribution are given in Green (2003, pp.781).

The regression equation with observed data can be written as equation (6):

$$y_i|(z_i = 1) = \mathrm{E}[y_i|\mathrm{z}_i^* > 0] + \upsilon_i$$
$$= \boldsymbol{x}_i'\beta + \beta_\lambda\lambda_i(\alpha_\mathrm{z}) + \upsilon_i \quad (6)$$

where the disturbance υ_i is heteroscedastic.

Least squares regression of y_i on **x** and λ would give a consistent estimator, but if λ is omitted, then the speciation error of an omitted variable is committed (Green, 2003). The marginal effect of the regressors on y_i in equation (6) is given as equation (7):

$$\frac{\partial E[y_i|\mathrm{z}_i^*>0]}{\partial x_{\mathrm{ik}}} = \beta_\mathrm{k} - \gamma_\mathrm{k}\left(\frac{\rho\sigma_y}{\sigma_z}\right)\delta_i(\alpha_z) \quad (7)$$

where $\delta_i(\alpha_\mathrm{z}) = \lambda_i(\alpha_\mathrm{z})[\lambda_i(\alpha_\mathrm{z}) - \alpha_\mathrm{z}]$, $0 < \delta_i < 1$.

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 $\gamma_\mathrm{k}\left(\frac{\rho\sigma_\varepsilon}{\sigma_\mathrm{u}}\right)\delta_i(\alpha_\mathrm{u})$. Suppose ρ is positive and $\mathrm{E}[y_i]$ is greater when $\mathrm{z}_i^* > 0$ than otherwise. As $0 < \delta_i < 1$, for a particular independent variable, if it appears in the probability as $\mathrm{z}_i^* > 0$, then it will influence y_i through λ_i, and reduce the marginal effect (see Green 2003, p.783).

As shown above, the vector of inverse Mill's ratios (estimated expected error) can be generated from the parameter estimates. The level of intern purchase, *y*, is observed only when the selection equation equals 1 (that is, when a firm uses the internet) and is then regressed on the explanatory variables, *x*, and the vector of inverse Mill's ratios from the selection equation by ordinary least squares. Therefore, the second stage reruns the regression with the estimated expected error included as an extra explanatory variable, removing the part of the error term correlated with the explanatory variable, and thereby avoiding the bias. Sample selection bias has been corrected by the selection equation, which determines whether an observation is included in the nonrandom sample.

4. Data and Variables

In order to reflect the use of the internet by manufacturers from a geographical dimension, we use census data for Taiwan's manufacturing firms obtained from the Directorate-General of Budget, Accounting and Statistics (DGBAS) for 2006. Our sample comprises a total of 153,081 manufacturers that may be broken down into 26 items (at the 2-digit SIC level) and 212 items at the (at the 4-digit SIC level)[5]. The scope of coverage includes the island of Taiwan and the Penghu archipelago, there being a total of 358 urban and rural areas. The 26 industries associated with the 2-digit code and numbers of firms are given in the Table 1.

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 (*HHI* in short) and the top four-firms' concentration ratio (*CR*4). The concept of the degree of industrial concentration is further extended to the estimation of industrial penetration, in which case we use the Geographical Herfindahl-Hirschman index (*GHHI*) as a proxy variable for industrial penetration. The formulae for the degree of industrial concentration and the geographical concentration index may be simply explained as follows:

(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:

[5] SIC (Standard Industrial Classification)

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

where s_{ij}: the market share of firm *i* in industry j, and *n* is the number of firms in industry *j*, $i = 1,2,3 \ldots n$.

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 the *HHI* is that the manufacturer's market share serves as a weight, with smaller manufacturers being given smaller weights, and larger manufacturers being given larger weights. The lower that the *HHI* value is, the lower is the degree of concentration in the industry; the higher the value, the higher the degree of industrial concentration.

(2) Top Four-firms Concentration Ratio, (*CR*4 in short): *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:

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

where s_{ij}: the market share of firm *i* in industry j and $s_{ij} \geq s_{i'j}$ for all $i < i'$.

(3) Geographical Herfindahl-Hirschman index (*GHHI* in short): This is the Herfindahl 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:

$$GHHI_j = \sum_{k=1}^{M} v_{jk}^2 \text{，} 0 \leq GHHI_j \leq 1$$

where v_{jk}: the ratio of the number of firms in industry j in region k to the total number of firms in industry j , M is the number of regions in industry j , $k = 1,2,3 \ldots M$.

When $GHHI_j$ is close to 1, this means that the firms within the industry are more geographically concentrated; when $GHHI_{jk}$ is close to 0, this means that the firms within the industry are more geographically dispersed. The advantage of $GHHI_j$ is its simplicity of calculation. 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 the $GHHI_j$ is not part of a neighborhood messaging system, it is not possible to reveal the differences brought about by being either closer or further away, or to reflect the spatial correlation for different economic activities; all one can do is indicate that economic activities are unevenly distributed. (3) $GHHI_j$ 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 earlier literature in section 2, 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), calculated as 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 Keelung City, Hsinchu City, Taichung City, Chiayi City, Tainan City, Taipei City or Kaohsiung City. When area_city=0, this means that the manufacturers are located in Taipei County, Yilan County, Taoyuan County, Hsinchu County, Miaoli County, Taichung County, Changhua County, Nantou County, Yunlin County, Chiayi County, Tainan County, Kaohsiung County, Pingtung County, Taitung County, Hualien County, or Penghu County.

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 itself 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 that particular manufacturer. The computer2 variable is used to measure the contagion effect for the internet technology within a certain area. Table 2 shows variable definition and Table 3 represents the statistical description of explanatory variables.

As described in the section 3, we use Heckman two-stage estimation procedure to obtain the estimates of parameters of the sample selection model which is specified as equation (8):

$$y_i = \beta_0 + \beta_1 HHI_j + \beta_2 Export_i + \beta_3 GHHI_j + \beta_4 city_i + \beta_5 computer1_i + \beta_6 computer2_{jki} + \beta_7 size_i + \beta_\lambda \lambda_i + \varepsilon_{yi} \quad (8)$$

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_j is the Herfindahl-Hirschman index for the industry *j* that firm *i* belongs to, and, $export\_rate_i$ is export intensity for firm *i*, $GHHI_j$ is the Geographical Herfindahl-Hirschman index for the industry *j* in region *k* that firm *i* is located to, $city_i$ is dummy variable indicating that the firm's geographical location, when $\text{city}_i = 1$ if firm *i* is located in the city, $\text{city}_i = 0$, otherwise. $computer1_i$ is the cost on buying the computer equipment for firm *i*, and $computer2_i$ is the total cost on computer equipment within the same industry and same area, but deducting the expenditure on computer equipment of firm *i* itself. The variable "computer2_i" is to capture the contagion effect for the internet technology in the same area and industry. The variable "$size_i$" is to capture the firm's characteristics. The λ_i is obtained from the select equation which is given as equation (9):

$$z_i = \gamma_0 + \gamma_1 HHI_j + \gamma_2 export\_rate_i + \gamma_3 GHHI_j + \gamma_4 city_i + \gamma_5 size_i + \gamma_6 group_i + \varepsilon_{zi} \quad (9)$$

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

Table 4 shows the correlation coefficients for each variable. In addition to the correlation coefficient between $export_i$ and (HHI_j and $CR4_j$) 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 various variables. In the next section, we report the empirical results based on Heckman two-stage estimation.

5. Empirical Results

The Column 2 of Table 5 and Table 6 reports 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. The Table 5 reports the results with *HHI* as the proxy variable for the degree of industrial concentration, while the Table 6 reports the results with *CR4* as the proxy variable for the degree of industrial concentration instead. The Column 3 of both Table 5 and Table 6 gives the coefficient estimate for the select equation for equation (9), which is estimated by probit regression.

In order to enhance the efficiency in estimation, we also use bootstrapping methods to estimate the variances, both with and without bootstrapping standard deviation are reported in the Tables 5 and 6. The 2-digit industry dummies are included in the empirical model to control heterogeneity, for saving space, we do not report each of coefficient estimate of 2-digit industry in the tables.

Our empirical result shows that regardless of whether the bootstrapping method is used or not, a nonzero Mill's lambda (β_λ), reject the statistical hypothesis that β_λ equal zero at the 1% level of significance, indicating that sample selection bias should be taken into account into the model. In order to make the empirical results easier to read, we first present the results for whole manufacturing industry and then second present the results for individual 2-digit industries.

For the whole industry, we will firstly summarize the results of selection corrected equation of firm's internet use for the factors influencing the extent to which manufacturers use the internet and also marginal effect of explanatory variables, and then we summarize the results of selection equation for the factors determining manufacturers to use or not adopt the internet for their business.

The regression model with selection corrected for all industry:

The coefficient of HHI_j is positive (but insignificant) in the Column 2 of Table 5, while the coefficient of $CR4_j$ is positive and significant in the Column 2 of Table 6, respectively. These indicate that higher degree of industrial concentration increase firms' expenditure to internet use. The coefficient of $export_i$ is positive but insignificant in the Column 2 of Table 5 and Table 6 respectively, indicating that the export intensity has no statistical impact on the expenditure of firm to internet use.

The coefficient of $GHHI_j$ shows negative and significant in the Column 2 of Table 5 and Table 6, respectively, indicating that the lower the level of the industrial penetration, the greater the extent to which the manufacturers will use the internet. The coefficients of $city_i$ show a positive and significant effect in the Column 2 of Table 5 and Table 6, respectively.

The coefficients of computer1_i show a positive but insignificant effect in both the Column 2 of Table 5 and Table 6. These indicate that the manufacturers' expenditure on computer equipment has not statistical impact on the expenditure of firm to internet use. The coefficients of computer2_i show a positive but insignificant effect with bootstrapping standard deviation in both the Column 2 of Table 5 and Table 6. These indicate that the manufacturers' expenditure on computer equipment within the same industry and region has no statistical impact on the expenditure of firm to internet use.

We further calculate the marginal effect of equation (8) (also eq. (7)) and report the marginal effect in the Table 7. The Column 2 of Table 7 gives the industrial marginal effects with HHI_j as the proxy variable for the degree of industrial concentration, while the Coulun 3 of Table 7 gives the industrial marginal effects with $CR4_j$ as the proxy variable for the degree of industrial concentration respectively.

For the HHI_j variable, the marginal effect is -0.0902 for the Column 2 and -0.007 for the Column 3 in Table 7. For example, the figure -0.0902 means when the degree of industrial concentration rate increase by 1, the extent to which manufacturers use the internet reduce by 0.0902%, indicating that the lower the degree of industrial concentration, the greater the extent to which manufacturers use the internet. Not surprisingly, there are differences between the marginal effect of HHI_j and $CR4_j$ on the extent to which manufacturers use the internet, as we had described in the section 4 that HHI_j takes into account all firms in an industry, use manufacturer's market share as a weight, with smaller firm being given smaller weights and bigger firm being given bigger weights, while $CR4_j$ is only consider the weighted average of the market shares of the top four-firms in an industry. However, our findings of industrial concentration agree with those of Galliano and Roux (2008) and Galliano et al. (2011) who used French manufacturing industry data.

For the $export_i$ variable, the marginal effect is (0.2708, 0.2963) for the Column 2 and the Column 3 in Table 7. For example, the figure 0.2708 means when the export intensity is increased by 1, the extent to which the manufacturers use the internet will increase by 0.2708%.

For the $GHHI_j$ variable, the marginal effect is (-0.0245, -0.0133) for the Column 2 and the Column 3 in Table 7. For example, when the industrial penetration is reduced by 1, the extent to which the manufacturers use the internet will increase by 0.0245%. That is to say, there exists a substitutionary relationship between the extent to which the manufacturers use the internet and the level of industrial penetration, a result that accords with the results 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. The result confirms that the popularity of the internet is such that the distance factor is no longer so important, that is, the internet has overcome the problem of distance between manufacturers.

It is worth noting that for the dummy variable $city_i$, the marginal effect is (-0.0051, -0.0062) for the Column 2 and the Column 3 in Table 7. For example, manufacturers who are located in the city areas will use the internet to -0.0051% lesser than those who are located in the county areas. In other words, manufacturers who are located in county areas will use the internet to a greater extent than those who are located in the city areas. These results also confirm empirical finding by Forman et al. (2005) and Kolko (1999), in that a complementary relationship exists between internet intensity and urbanization.

We now continue to present the Column 3 of Table 5, and Table 6 that show probit estimations, as given by equation (9), which estimate the factors of whether manufacturers will use or not use the internet for their business.

Our empirical results show no matter *HHI* or *CR4* was used as the proxy variable for the degree of industrial concentration, the coefficients of HHI_j and $CR4_j$ are negative and significant at 1% level of significance in the column 3 of Table 5 and Table 6. These indicate that the more competition that the manufacturers face, in order to increase their ability to compete with other manufacturers, the more that they will be inclined to use the internet for business.

Export intensity is also an important factor for affecting the manufacturers' use of the internet. The coefficients of $export_i$ is positive and significant at 1% level of significance in the column 3 of Table 5 and Table 6. This is not surprised that the more that manufacturers rely on exports, the greater their export intensity, and the more that they need to use the internet to communicate with overseas customers.

The coefficient of the geographical location, $city_i$ in the column 3 of Table 5 and Table 6 show a negative and significant effect on manufacturers use or not use the internet for their business. This result suggests that manufacturers who are located in the county areas will be likely to use the internet for business than those who are located in the city areas. However, this result is contrast with the empirical results of coefficient of $city_i$ in the column 2 of Table 5 and Table 6, which suggest the manufacturers who are located in the city will expend more money on the internet use than firms in the county.

The coefficient of manufacturer's scale of operations, $size_i$ shows a positive and significant probability of manufacturers to use the internet for their business. It is not surprised that the bigger firm will be likely to use the internet for business. Also a positive and significant coefficient of $group_i$, which suggests that manufacturers with independent operations will be likely to use the internet for business than those who do have subsidiary (branch). It is not surprised that as Taiwan largely consists of manufacturers with independent operations, the likelihood of such manufacturers using the internet is relatively high.

While the impact of the degree of industrial penetration on the manufacturers' use of the internet is not significant in the column 3 of Table 5 and Table 6, the effect on the extent to which manufacturers use the internet is significant and negative in the column 2 of Table 5 and Table 6, indicating that the extent of the industrial penetration does not affect whether or not the manufacturers will use the internet, but it will affect the extent to which manufacturers who already use the internet.

The regression model with selection corrected for two digit industries:

In this section we only report the Heckman two-stage estimation with *HHI* as the proxy variable for the degree of industrial concentration and marginal effect for two digit industries in the Table 8 and Table 9, respectively. A nonzero Mill's lambda (β_λ), rejects the statistical hypothesis that β_λ equal zero 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, (32) Furniture. However, because the industries being different, the empirical results for the individual industries based on the two-digit level classifications also vary. For individual 2-digit industries, we will firstly summarize the results of selection corrected equation for the extent to which manufacturers use the internet, then the results of selection equation for the factors of whether or not manufacturers use the internet and finally summarize the marginal effect.

The effect of the degree of industrial penetration ($GHHI_j$) in terms of the extent to which manufacturers use the internet vary across 2-digit industries. In the case of traditional industries such as (08) Food, (12) Wearing Apparel and Clothing Accessories, (13) Leather, Fur and Related Products, (32) Furniture and also technology-intensive industries such as (28) Electrical Equipment, (30) Motor Vehicles and Parts, (31) Other Transport Equipment, and also basic industries such as (24) Basic Metal, show the lower the level of the industrial penetration, the greater the extent to which the 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 degree of the industrial penetration, the greater the extent to which the manufacturers will use the internet.

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

The variable $export_i$ show 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, and 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 negative effect on the extent to which manufacturers use the internet for traditional industries.

The effect of the geographic location, $city_i$ show manufacturers who are located in county areas will use the internet to a greater extent than those who are 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 shows manufacturers who are located in city areas will use the internet to a greater extent than those who are located in the county areas.

The variable of manufacturers' expenditure on computer equipment, computer1_i, has no statistical impact on the expenditure of firm to internet use for most of the 2-digit industries, except for traditional industries such as (16) Printing and Reproduction of Recorded Media, and 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.

Similar, computer2_i that use to capture the contagion effect for the internet technology in the same area show no statistical impact on the expenditure of firm to internet use for most of the 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.

In the following paragraph, we will present the probit estimation, as given by equation (9), which estimates the factors of whether or not manufacturers adopt or not adopt the internet for their business across 2-digit industries and the coefficient estimates also be shown in the Table 8.

The effect of the degree of industrial penetration ($GHHI_j$) in terms of whether or not manufacturers will use the internet shows different across 2- digit industries. As for traditional industries such as the (8) Food, (11) Textiles Mills, (13) Leather, Fur and Related Products, (14) Wood and Bamboo Products, and also technology-intensive industries such as (29) Machinery and Equipment, (31) Other Transport Equipment, and also basic industries such as (25) Fabricated Metal Products, when the degree of industrial penetration is high, manufacturers will be more inclined to use the internet, while traditional industries such as (15) Pulp, Paper and Paper Products, (16) Printing and Reproduction of Recorded Media, (32) Furniture, (33) Manufacturing Not Elsewhere Classified, and also technology-intensive industries such as (26) Electronic Parts and Components, (30) Motor Vehicles and Parts, and also basic industries such as (22) Plastic Products, when the degree of industrial penetration is high, manufacturers will be less inclined to use the internet. However, industrial penetration will not affect whether or not manufacturers use the internet for most of basic industries such as (18) Chemical Material, (19) Chemical Products, (20) Medical Goods, (21) Rubber Products, (24) Basic Metal, and traditional industries such as the (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 degree of industrial concentration (HHI_j) in terms of whether or not manufacturers will use the internet shows different across 2- digit industries. In terms of traditional industries such as (11) Textiles Mills, (15) Pulp, Paper and Paper Products, (23) Non-metallic Mineral Products, (32) Furniture, and technology-intensive industries such as (29) Machinery and Equipment, and basic industries such as (22) Plastic Products, when the degree of the industrial concentration increase, manufacturers will be more inclined to use the internet. On the contrary, in the case of 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 the industrial concentration decrease, manufacturers will be more likely to use the internet.

The effect of $export_i$ is important for affecting the manufacturers' decision to use the internet for many of 2-digit industries. In the case of traditional industries such as (14) Wood and Bamboo Products, (15) Pulp, Paper and Paper Products, (16) Printing and Reproduction of Recorded Media, and 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 increase, manufacturers will be more likely to use the internet. On the contrary, in the case of basic industries such as (18) Chemical Material, (19) Chemical Products, (21) Rubber Products, when the degree of export intensity increase, manufacturers will be less likely to use the internet.

The coefficient of $size_i$ shows a positive effect for affecting the manufacturers' decision to use the internet for most of 2-digit industries. Also the coefficient of $group_i$ shows a positive and significant effect on manufacturers' decision to use the internet for most of 2-digit industries.

In the following, we will present the total marginal effect of each of the explanatory variables on the extent to which the manufacturers use the internet for the individual 2-digit industries in Table 9. Of these 26 industries, seven 2-digit industries significantly reject null hypothesis that β_λ equal zero at 10% level of significance with bootstrapping standard deviation, namely, (08) Food, (09) Beverages, (22) Plastic

Products, (28) Electrical Equipment, (29) Machinery and Equipment, (30) Motor Vehicles and Parts (32) Furniture, indicating that these industries are affected by the problem of sample selection bias, thus making it necessary to correct this sample selection bias.

In the following paragraph, we will present the marginal effect as given by the equation (8) (also eq. (7)). In terms of industrial penetration ($GHHI_j$), among traditional industries, the largest value is 2.3761 for the (09) Beverages, while the smallest is -1.4581 for the (32) Furniture; for technology-intensive industries, the largest value is 5.5503 for the (27) Plastic Products, while the smallest is -12.6278 for the (30) Motor Vehicles and Parts; for basic industries the largest value is 21.886 for the (20) Medical Goods, while the smallest is -1.3668 for the (21) Rubber Products.

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

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

The marginal effect of geographic location ($city_i$), among traditional industries, the largest value is 0.0266 for the (08) Food, while the smallest is -0.0018 for the

(11)Textiles Mills; for technology-intensive industries, the largest value is 0.0527 for the (26) Electronic Parts and Components, while the smallest is -0.0249 for the (27) Plastic Products; for basic industries the largest value is 0.0578 for the (21) Rubber Products, while the smallest is -0.0216 for the (24) Basic Metal.

The marginal effect of manufacturer's scale of operations, ($size_i$), among traditional industries, the largest value is 0.0029 for the (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.

The marginal effect of manufacturers' expenditure on computer equipment, $\mathrm{computer1}_i$, among traditional industries, the largest value is 17.4643 for the (11) Textiles Mills, while the smallest is -0.0075 for the (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 the (30) Motor Vehicles and Parts; for basic industries the largest is 139.043 for (24) Basic Metal, while the smallest is -5.4236 for the (21) Rubber Products.

The marginal effect of the manufacturers' expenditure on computer equipment within the same industry and region ($\mathrm{computer2}_i$), 0.0045 for the (15) Pulp, Paper and Paper Products, 0.0025 for the (27) Plastic Products and, 0.0008 for the (24) Basic Metal, have the largest value for the traditional industries, for technology-intensive industries, and for the basic industries, respectively.

6. Conclusion

In this paper, we use 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 that the total expenditure on internet intensity, an actual figure is observed only if the firm uses the internet that will cause the problem of sample selection (selection bias). In order to correct the problem of selection bias, this paper use Heckman selection model and two-stage estimation procedure to obtain the estimates of parameters of the sample selection model.

In order to improve the effectiveness of our estimation, we further use bootstrapping approach to estimate the sample variance, our empirical results show that regardless of whether we use the bootstrapping approach, the Mill's lambda test statistic significantly reject null hypothesis that β_λ equal zero at the 1% level of significance for the aggregated full industry and 7 out of 26 industries significantly reject null hypothesis that β_λ equal zero at 10% level of significance, indicating the problem of the sample selection bias should be corrected. Our conclusions 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, the manufacturer's size of operations, and the independence of operations. As Taiwan largely consists of manufacturers with independent operations, it is not surprised that the likelihood of such manufacturers using the internet is relatively high and the manufacturers' independence of operations having the greatest impact. The second most influential factor is the manufacturers' export intensity, indicating that the more that manufacturers rely on exports, the greater their export intensity, and the more that they need to use the internet to communicate with overseas customers.

The third most influential factor is the degree of industrial concentration. The more competition that the manufacturers face, in order to increase their ability to compete with other manufacturers, the more that they will be inclined to use the internet. Our empirical results also show that manufacturers who are located in the county areas would be likely to use the internet for business than those who are located in the city areas, and the bigger firm would be likely to use the internet for business than smaller size firm. However, the impact of the degree of industrial penetration on the 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 penetration, geographical location, and the contagion effect. While the impact of the degree of industrial penetration 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, indicating that the extent of the industrial penetration does not affect whether or not the manufacturers will use the internet, but it will affect the extent to which manufacturers who already use the internet will use the internet. Our results seem to suggest there exists a substitutionary relationship between the penetration of localization and the extent to which manufacturers use the internet, indicating that internet technology has overcome the "distance" factor, so that the distance factor is no longer so important.

(3) The variable of industrial penetration show a negative marginal effect on the extent to which the manufacturers use the internet, indicating there exists a substitutionary relationship between the extent to which the manufacturers use the internet and the level of industrial penetration. Such results confirm the researches 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.

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

(5) The export intensity has the greatest marginal effect on the extent to which the manufacturers use the internet, indicating that international competition has relatively large influence on the extent of internet intensity. The second and third largest are the variables of manufacturers' expenditure on computer equipment and the contagion effect that have a positive marginal effect on the extent to which the manufacturers use the internet, thought the magnitudes for both marginal effects are quite small.

(6) Because of the industries being different, the empirical results for the individual industries based on the two-digit level classifications are quite varied. In terms of the variable of degree of industrial penetration, (09) Beverages and (32) Furniture are largest positive (2.376) and smallest negative (-1.458) marginal effect on the extent to which the manufacturers use the internet respectively for traditional industry; (27) Plastic Products and (30) Motor Vehicles and Parts are largest positive (5.550) and smallest negative (-12.628) marginal effect on the extent to which the manufacturers use the internet respectively for technology-intensive industry; (20) Medical Goods and (21) Rubber Products are largest positive (21.886) and smallest negative (-1.367) marginal effect on the extent to which the manufacturers use the internet respectively for basic industry.

(7) In terms of the marginal effect of localized penetration on the extent to which the manufacturers use the internet is also vary. The largest positive and smallest negative value for the traditional industries are 0.0266 for the (08) Food and -

0.0018 for the (11) Textiles Mills; the largest and smallest value for technology-intensive industries are 0.0527 for the (26) Electronic Parts and Components and -0.0249 for the (27) Plastic Products; the largest and smallest value for basic industries are 0.0578 for the (21) Rubber Products and -0.0216 for the (24) Basic Metal.

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

Table 1

Industry 2-digit codes and number of firms

	code	2-digit industry	Number 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 variable	
y_i	the extent to which the firm *i* use the internet = (online purchase amount + online sales amount) / total sales
z_i	z_i=1, if firm *i* use an internet equipment for business information z_i=0, otherwise
Independent variable	
HHI_j	Herfindahl-Hirschman Index for the industry *j* that firm *i* belongs to.
$CR4_j$	Top Four firms Concentration Index for the industry *j* that firm *i* belongs to.
$export_i$	Export share for firm *i*= export value / total sales
$GHHI_j$	Geographic Herfindahl-Hirschman lndex for the industry j in the region that firm *i* is located to
$size_i$	Firm size Total number of employees for the firm *i*
$computer1_i$	Total expenditure on the computer equipment for firm *i* unit: NT$1000
$computer2_i$	Total expenditures on computer equipment within the same industry and same area, exclude the expenditure of firm *i* itself unit: NT$1000
$city_i$	$city_i$ =1, if firm *i* locate at the city $city_i$ =0, if firm *i* locate at the county
$group_i$	$group_i = 1$, if firm *i* has no subsidiary (branch) $group_i = 0$, otherwise

Table 3

Statistical descriptions

Variables (unit)	Mean	Std Dev.	Min	Max
y_i (100%)	1.9998	43.2231	0	7153.077
z_i	0.6069	0.4884	0	1
HHI_j	0.0322	0.0656	0.0020	1
$CR4_j$	0.2053	0.1683	0.0407	1
$export_i$	0.0709	0.1669	0	1
$GHHI_j$	0.0031	0.0239	0	0.4752
$size_i$	16.7994	113.8733	0	17,040
$computer1_i$ (NT$1000)	0.0029	0.2871	0	99.2
$computer2_i$ (NT$1000)	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

	HHI_j	$CR4_j$	$GHHI_j$	$export_i$	$city_i$	$computer1_i$	$computer2_i$	$size_i$
HHI_j	1							
$CR4_j$	0.8518	1						
$GHHI_j$	-0.0078	0.0011	1					
$export_i$	0.1558	0.1780	0.0413	1				
$city_i$	0.0261	0.0290	-0.0428	0.0093	1			
$computer1_i$	0.0028	0.0066	-0.0008	-0.0032	-0.0002	1		
$computer2_i$	0.0077	0.0155	0.0140	-0.0149	0.0010	0.0401	1	
$size_i$	0.0803	0.0863	-0.0000	0.1729	0.0072	0.0010	-0.0062	1

Table 5

Selection corrected internet Intensity (with HHI) for all industries

Variables	Intensity of internet use (y_i)	Select (z_i)
HHI_j	0.148 (3.660) [2.732]	-1.369 (0.065)*** [0.067]***
$export_i$	1.086 (1.284) [1.336]	3.807 (0.207)*** [0.057]***
$GHHI_j$	-2.774 (1.057)*** [5.237]	0.051 (0.237) [0.201]
$city_i$	0.852 (0.523)* [0.378]**	-0.201 (0.013)*** [0.010]***
$computer1_i$	0.239 (51.880) [0.432]	-
$\mathrm{computer2_i}$	0.069 (0.119) [0.019]***	-
$size_i$	0.002 [0.002]	0.003 (0.001)*** [0.0002]***
$group_i$	-	58.543 (16.397)*** [0.005]***
constant	2.643 (0.755)*** [0.882]***	-57.606 (16.400)***
Mills lambda (λ)	-7.229 (2.595)*** [2.193]***	
# of observations	153081	
# of censored observation	31924	
Wald Chi2(df)	543.38(32)	

Note: Bootstrapping standard errors are in the parentheses and standard errors without bootstrapping appear in square brackets. The asterisks ***, ** and * denote significance at the 1%, 5% and 10% levels, respectively. 2-digit industry dummies are included in the empirical equation to control heterogeneity, but not report in the table for saving space.

Table 6

Selection corrected internet Intensity (with CR4) for all industries

Variables	Intensity of internet use (y_i)	Select (z_i)
$CR4_j$	4.137 (1.160)*** [1.244]***	-0.645 (0.028)*** [0.025]***
$export_i$	0.532 (1.143) [1.342]	3.813 (0.214)*** [0.057]***
$GHHI_j$	-1.861 (1.064)* [5.246]	0.071 (0.203) [0.202]
$city_i$	0.904 (0.344)*** [0.377]**	-0.201 (0.011)*** [0.010]***
$computer1_i$	0.240 (55.104) [0.432]	-
computer2_i	0.069 (0.142) [0.019]***	-
$size_i$	0.001 [0.002]	0.004 (0.001)*** [0.0002]***
$group_i$	-	61.607 (22.335)*** [0.007]***
constant	1.876 (0.763)** [0.894]**	-60.585 (22.243)***
Mills lambda (λ)	-8.067 (2.444)*** [2.172]***	
# of observations	153081	
# of censored observation	31924	
Wald Chi2(df)	561.99(32)	

Note: same as Table 4

Table 7

Marginal effect of internet intensity

unit:%

Variables	Internet intensity (1)	Internet intensity (2)
$GHHI_j$	-0.0243	-0.0133
HHI_j	-0.0897	
$CR4_j$		-0.0069
$export_i$	0.2643	0.2908
$city_i$	-0.0049	-0.0060
$size_i$	0.0002	0.0003
$computer1_i$	0.0024	0.0024
$computer2_i$	0.0007	0.0007

Table 8. Selection corrected internet intensity (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
$GHHI_j$	-16.06	22.91	-39.48	206.89	-9.10	10.74	-0.31	0.23	-11.52	35.24	-38.27	98.98	-46.92	-193.30
	(3.53)***	(5.37)***	(16.07)**	(237.25)	(10.16)	(2.38)***	(0.16)*	(0.19)	(3.68)***	(19.24)*	(67.17)	(61.42)	(69.32)	(29.47)***
HHI_j	10.06	-8.14	-0.38	-3.84	-2.95	3.81	1.70	-1.80	17.76	-16.81	13.47	-10.51	-1.46	3.66
	(3.81)***	(0.74)***	(0.89)	(60.43)	(3.53)	(0.98)***	(1.70)	(0.61)***	(7.47)**	(6.49)***	(22.95)	(7.58)	(1.68)	(1.34)***
$export_i$	0.84	18.50	0.80	4.26	4.22	21.50	1.10	12.96	-0.27	17.69	3.93	676.48	-0.14	916.90
	(1.59)	(545.92)	(0.39)**	(237.99)	(5.22)	(86.23)	(1.02)	(366.33)	(0.17)	(989.31)	(3.09)	(353.91)*	(0.59)	(243.48)***
$city_i$	-0.73	1.15	-0.21	218.03	-0.09	-0. 21	0.37	0.46	0.24	-0.12	0.62	-0.20	-0.15	-0.55
	(0.29)**	(0.22)***	(0.06)***	(139.58)	(0.74)	(0.06)***	(0.30)	(0.06)***	(0.27)	(0.15)	(0.36)*	(0.11)*	(0.21)	(0.06)***
$size_i$	-0.0003	0.05	0.001	0.22	-0.01	0.00005	0.003	0.003	-0.001	0.01	0.11	0.01	0.03	0.01
	(0.003)	(0.01)***	(0.002)	(0.25)	(0.01)	(0.002)	(0.002)	(0.002)	(0.002)	(0.02)	(0.07)	(0.005)	(0.03)	(0.01)
$computer1_i$	24.42		-0.07		1751.71		86.13		-0.73		46.31		26.20	
	(18.66)		(10.86)		(1497.42)		(84.32)		(4.86)		(114.46)		(145.49)	
$computer2_i$	0.01		-0.22		-5.02		0.07		-1.94		0.90		0.60	
	(0.45)		(0.31)		(4.49)		(0.13)		(0.58)***		(1.31)		(1.47)	
$group_i$		91.68		313.54		7.09		12.05		25.35		14.07		16.86
		(30.83)***		(193.88)		(1.38)***		(2.84)***		(41.1)		(5.28)***		(7.63)**
$constant$	0.74	-90.28	0.20	-311.96	0.05	-6.78	-0.41	-11.47	0.13	-24.39	-0.63	-12.82	-0.31	-15.63
	(0.16)***	(30.86)***	(0.07)***	(194.32)	(1.45)	(1.38)***	(0.36)	(2.82)***	(0.10)	(41.13)	(0.98)	(5.30)**	(0.31)	(7.68)**
# of observations	6165		644		6439		4084		1870		2849		3605	
# of censored	1081		106		1783		936		306		329		595	
Mills Lambda	-3.01		-1.28		-2.10		0.93		0.14		0.12		1.04	
	(1.15)***		(0.67)*		(2.61)		(0.73)		(0.49)		(1.85)		(1.08)	
Wald Chi2(ddl)	31.13(7)		27.53(7)		3.48(7)		15.84(7)		17.80(7)		19.65(7)		5.97(7)	

Table 8. Selection corrected internet intensity (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
$GHHI_j$	14.21 (13.76)	-40.82 (3.26)***	-176.22 (221.18)	12.04 (183.62)	86.35 (240.43)	9.18 (107.81)	2103.67 (1324.42)	4.75 (162.27)	138.09 (351.43)	17.62 (37.44)	321.85 (236.26)	-33.14 (12.28)***
HHI_j	-0.49 (2.15)	0.08 (1.62)	-3.63 (3.02)	0.51 (1.27)	8.30 (5.50)	1.13 (1.84)	54.14 (40.35)	-0.22 (10.75)	0.26 (6.91)	-0.97 (1.17)	72.07 (76.82)	25.45 (8.33)***
$export_i$	4.42 (3.23)	1155.05 (573.23)**	5.19 (2.14)**	-3.46 (0.30)***	3.13 (2.54)	-1.87 (0.31)***	9.68 (6.88)	1662.65 (797.32)**	4.25 (3.34)	-2.96 (0.20)***	-1.60 (1.10)	1.32 (0.64)**
$city_i$	-0.02 (0.05)	-0.13 (0.03)***	-0.52 (0.44)	0.07 (0.32)	-0.10 (0.41)	0.12 (0.16)	0.80 (1.63)	-0.24 (0.28)	2.11 (2.01)	-0.23 (0.21)	0.77 (0.44)*	-0.31 (0.05)***
$size_i$	0.02 (0.02)	0.01 (0.003)***	-0.003 (0.01)	0.06 (0.01)***	-0.004 (0.01)	0.04 (0.02)**	-0.03 (0.02)	0.02 (0.06)	-0.002 (0.01)	0.11 (0.02)***	-0.01 (0.004)***	0.03 (0.01)***
$computer1_i$	126.63 (32.73)***		-22.34 (116.68)		-76.94 (166.20)		-84.18 (620.14)		-530.75 (290.49)*		444.96 (185.26)**	
$computer2_i$	-0.02 (0.02)		-0.60 (2.21)		-0.11 (0.29)		0.05 (1.43)		-0.05 (1.08)		-0.11 (0.07)	
$group_i$		11.30 (1.60)***		218.90 (67.82)***		39.08 (15.73)**		18.50 (38.31)		304.01 (106.46)***		43.54 (8.21)***
constant	-0.32 (0.33)	-10.74 (1.60)***	1.48 (0.48)***	-216.90 (67.79)***	0.76 (0.49)	-37.20 (15.76)**	-0.16 (2.41)	-17.10 (38.74)	0.51 (0.91)	-302.68 (106.49)***	2.68 (0.78)***	-42.42 (8.24)***
# of observations	9439		1549		2304		543		1756		11012	
# of censored observation	2790		455		499		142		249		1487	
Mills Lambda	0.40 (0.56)		0.63 (6.84)		-0.72 (7.40)		-30.66 (19.23)		15.50 (9.90)		-10.60 (2.75)***	
Wald Chi2(ddl)	36.82(7)		10.16(7)		10.78(7)		11.61(7)		8.59(7)		24.98(7)	

Table 8. Selection corrected internet Intensity (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
$GHHI_j$	-2.12 (1.63)	0.35 (0.63)	-105.59 (60.97)*	-0.78 (4.26)	-55.07 (34.47)	15.17 (2.74)***	-79.22 (67.35)	-10.87 (2.87)***	543.29 (464.33)	3.88 (11.80)	-42.62 (6.84)***	2.74 (5.40)
HHI_j	0.58 (2.55)	1.19 (0.57)**	-20.61 (12.82)	-0.05 (0.25)	13.17 (6.60)**	-4.71 (0.32)***	26.67 (9.45)***	0.18 (0.33)	-13.53 (18.21)	-0.38 (0.28)	-3.89 (1.75)**	-0.23 (1.30)
$export_i$	3.83 (2.59)	8.69 (390.12)	-7.67 (3.46)**	5.72 (313.81)	3.57 (4.11)	68.80 (65.41)	6.63 (2.59)**	18.18 (1.45)***	8.71 (8.04)	7.59 (266.85)	0.51 (0.49)	11.06 (877.09)
$city_i$	-0.27 (0.38)	-0.25 (0.10)***	-0.85 (1.31)	-0.37 (0.07)***	0.31 (0.95)	-0.09 (0.03)***	5.29 (6.71)	0.09 (0.06)	-2.89 (3.29)	0.10 (0.10)	0.22 (0.18)	-0.05 (0.08)
$size_i$	0.01 (0.01)	0.03 (0.01)***	0.01 (0.01)	0.01 (0.01)	0.05 (0.02)**	0.003 (0.001)***	-0.002 (0.004)	0.0001 (0.0002)	-0.005 (0.01)	0.003 (0.002)**	-0.01 (0.002)***	0.01 (0.005)*
$computer1_i$	140.06 (185.01)		13895.31 (9161.63)		44.14 (8.61)***		272.09 (303.94)		7.24 (3224.05)		-1.10 (76.65)	
$computer2_i$	-0.32 (0.38)		0.08 (1.31)		0.05 (0.06)		-0.03 (0.02)		0.25 (5.34)		-0.04 (0.03)	
$group_i$		70.24 (24.95)***		12.51 (8.95)		10.52 (0.93)***		6.99 (2.64)***		18.74 (5.59)***		22.70 (11.12)**
$constant$	0.32 (0.59)	-69.12 (24.99)***	0.64 (1.63)	-11.41 (9.00)	1.56 (0.51)***	-9.93 (0.93)***	0.09 (0.56)	-6.59 (2.65)**	5.98 (6.47)	-17.80 (5.63)***	1.69 (0.32)***	-21.69 (11.13)*
# of observations	3677		4710		39047		6023		3717		6198	
# of censored	684		861		8496		1558		716		1065	
Mills Lambda	0.06 (2.64)		-5.90 (5.90)		-0.59 (1.35)		2.34 (4.30)		-9.37 (20.51)		-3.76 (1.16)***	
Wald Chi2(ddl)	9.66(7)		10.64(7)		56.58(7)		30.65(7)		5.68(7)		42.45(7)	

Table 8. Selection corrected internetiIntensity (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
$GHHI_j$	5.05 (5.32)	5.31 (1.57)***	-1270.42 (339.41)***	-58.65 (26.40)**	-97.88 (44.32)**	23.42 (9.05)***	-142.35 (55.20)***	-40.96 (15.36)***	-1.70 (23.26)	-19.63 (3.91)***
HHI_j	-27.89 (2.38)***	7.32 (1.26)***	0.66 (9.77)	0.90 (0.59)	-8.66 (11.39)	-0.67 (1.33)	-75.58 (37.44)**	15.18 (5.73)***	-21.36 (14.65)	0.83 (1.44)
$export_i$	5.80 (0.90)***	99.56 (291.32)	25.94 (5.32)***	2270.49 (582.85)***	0.73 (0.99)	6.68 (3.12)**	-3.08 (4.87)	524.98 (408.23)	0.83 (0.78)	47.78 (534.54)
$city_i$	0.12 (0.20)	-0.27 (0.04)***	1.32 (2.05)	-0.36 (0.08)***	1.08 (0.58)*	-0.35 (0.10)***	0.49 (0.91)	-0.09 (0.12)	0.47 (0.40)	0.02 (0.07)
$size_i$	-0.01 (0.002)***	0.01 (0.004)***	-0.02 (0.01)**	0.01 (0.005)*	0.01 (0.01)	0.01 (0.01)	0.01 (0.02)	0.01 (0.01)***	0.01 (0.01)	0.004 (0.004)
$computer1_i$	60.40 (44.06)		-461.96 (295.56)		589.45 (360.87)		37.02 (67.99)		-0.76 (136.12)	
$computer2_i$	-0.15 (0.06)**		0.02 (0.39)		-0.08 (0.08)		-0.45 (0.63)		0.88 (1.15)	
$group_i$		25.93 (6.44)***		27.24 (9.91)***		24.65 (15.21)		12.47 (2.30)***		9.96 (3.71)***
$constant$	1.61 (0.16)***	-25.33 (6.46)***	3.16 (0.99)***	-26.49 (9.94)***	0.96 (0.98)	-23.84 (15.26)	4.19 (1.97)**	-11.45 (2.32)***	1.61 (0.57)***	-9.06 (3.74)**
# of observations	18545		3580		2905		2849		5435	
# of censored	3076		686		521		367		780	
Mills Lambda	-0.88 (0.39)**		5.77 (2.53)**		-1.74 (2.25)		-14.24 (7.45)*		-0.68 (1.60)	
Wald Chi2(ddl)	156.24(7)		47.75(7)		30.84(7)		21.73(7)		30.94(7)	

For saving space, we did not present (17) Petroleum and Coal Products industry and (34) Repair and Installation of Industrial Machinery and Equipment in the both Table 8 and Table 9, some coefficients of explanatory variables were not able to get the estimate for the correction regression model, so in the both Table 8 and Table 9, we do not present (17) and (34) industry for saving some space.

Table 9

Marginal effect of the internet intensity (with HHI) for two digit industries

unit:%

	Marginal Effects											
	(8)	(9)	(11)	(12)	(13)	(14)	(15)	(16)	(19)	(20)	(21)	(22)
$GHHI_j$	0.5189	2.2452	-0.0516	-0.0037	-0.1271	-0.3827	-0.4692	0.1421	0.9273	21.0367	-1.3420	0.5087
HHI_j	-0.1408	-0.0529	-0.0155	0.0222	0.1833	0.1347	-0.0146	-0.0049	0.0908	0.5414	0.1526	2.8016
$export_i$	0.5573	0.0624	0.1211	-0.0261	-0.0087	0.0393	-0.0014	0.0442	0.0183	0.0968	0.5001	0.0921
$city_i$	0.0268	-	-0.0018	0.0025	0.0024	0.0062	-0.0015	-0.0002	-0.0002	0.0080	0.0573	-0.0185
$size_i$	0.0014	0.0028	0.0001	0	0	0.0011	0.0003	0.0002	-0.0002	-0.0003	-0.0177	0.0020
$computer1_i$	0.2442	-0.0007	17.5171	0.8613	-0.0073	0.4631	0.2620	1.2663	-0.7694	-0.8418	-5.3075	4.4496
$computer2_i$	0.0001	-0.0022	-0.0502	0.0007	-0.0194	0.0090	0.0060	-0.0002	-0.0011	0.0005	-0.0005	-0.0011
$group_i$	-	-	0.1031	-0.0941	-0.0318	0	0	0	0.2633	0	-	-

	Marginal Effects										
	(23)	(24)	(25)	(26)	(27)	(28)	(29)	(30)	(31)	(32)	(33)
$GHHI_j$	-0.0214	-1.0797	-0.5480	-0.7682	5.6157	-0.3664	0.0505	-12.7042	-0.7477	-1.4235	-0.0170
HHI_j	0.0051	-0.2077	0.1308	0.2663	-0.1531	-0.0439	-0.2789	0.0066	-0.0932	-0.7558	-0.2136
$export_i$	0.0328	0.0976	0.0480	0.0261	0.4451	0.2464	0.0580	0.2594	0.0732	-0.0308	0.0083
$city_i$	-0.0025	-0.0205	0.0030	0.0527	-0.0241	0.0011	0.0012	0.0132	0.0072	0.0049	0.0047
$size_i$	0.0001	0.0003	0.0005	0	0.0001	0.0001	-0.0001	-0.0002	0.0002	0.0001	0.0001
$computer1_i$	1.4006	138.9531	0.4414	2.7209	0.0724	-0.0110	0.6040	-4.6196	5.8945	0.3702	-0.0076
$computer2_i$	-0.0032	0.0008	0.0005	-0.0003	0.0025	-0.0004	-0.0015	0.0002	-0.0008	-0.0045	0.0088
$group_i$	-	0.6456	0.0451	-0.1015	1.4723	0.7677	0.1331	0	0.3966	0	0.0139

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), pp. 630-640.

Atrostic, B. K., and S. V. Nguyen (2005), "IT and Productivity in U.S. Manufacturing: Do Computer Networks Matter?", *Economic Inquiry* 45(3): 493-506.

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

Baptista, R., and P. Swann (1998), "Do firms in clusters innovate more?", *Research Policy*, 27(5), pp. 525–540.

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.

Blaylock J. R. and W. N. Blissard (1992), "Self-valuated Health Status and Smoking Behaviour", *Applied Economics*, 25, pp.429-435.

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), pp. 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), pp. 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 Commercial Internet? 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), pp. 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), pp. 67-86.

Giunta, A., and F. Trivieri (2007), "Understanding the Determinants of Information Technology Adoption", *Applied Economics*, 39(10), pp. 1325-1334.

Goldfar A. and J. Prince (2008), "Internet Adoption and Usage Patterns are Different: Implications for the Digital Divide", *Information Economics and Policy*, 20, pp. 2–15.

Greene, W. (2003), "Sample Selectin Bias As a Specification Error: Comment", *Econometrica*, 49(3), pp. 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), pp. 475–492.

Heckman, J. J. (1979), "Sample Selection Bias as a Specification Error", *Econometrica*, 47(1), pp. 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), pp. 667-680.

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

Maddala, 1983

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