future internet

MDPI

Article

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

Chia-Lin Chang [1,*], Michael McAleer [2,3,4,5,6] and Yu-Chieh Wu [7]

1. Department of Applied Economics, Department of Finance, National Chung Hsing University, Taichung 40227, Taiwan
2. Department of Finance, Asia University, Taichung 41354, Taiwan; michael.mcaleer@gmail.com
3. Discipline of Business Analytics, University of Sydney Business School, Sydney, NSW 2006, Australia
4. Econometric Institute, Erasmus School of Economics, Erasmus University Rotterdam, 3062 PA Rotterdam, The Netherlands
5. Department of Economic Analysis and ICAE, Complutense University of Madrid, 28223 Madrid, Spain
6. Institute of Advanced Sciences, Yokohama National University, Yokohama 240-8501, Japan
7. Department of Applied Economics, National Chung Hsing University, Taichung 40227, Taiwan; jessica11037@gmail.com

* Correspondence: changchialin@email.nchu.edu.tw; Tel.: +886-4-22840350

† The authors are most grateful to three reviewers for very helpful comments and suggestions. For financial support, the first author wishes to thank the Ministry of Science and Technology (MOST), Taiwan, and the second author acknowledges the Australian Research Council and the Ministry of Science and Technology (MOST), Taiwan.

Received: 10 January 2018; Accepted: 13 March 2018; Published: 16 March 2018

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

1. Introduction

With the arrival of the internet era, internet intensity by business enterprises has continued to increase significantly in recent years. Furthermore, the proliferation of internet technology has 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, thus making it easier for businesses to communicate with each other over such long distances. Taiwan's overall industrial internet intensity (that is, the proportion of enterprises that uses the internet) has increased from 62% in 2002, to 79% in 2003, and to 94.3% in 2010. According to reports that were prepared by the Institute for Information Industry in 2008, 2009, and 2010 (see [1] Lin (2009) for 2008 and 2009, and [2] Lin (2010) for 2010), the growth of the internet has been increasing rapidly, especially 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 been able to decrease their costs of communicating and collecting information. Owing to the increased convenience that the internet has brought in enabling firms to communicate with each other, in reducing the costs of transportation, as well as an abundance of resources that has further sped up the exchange of information, the "distance" factor is clearly no longer as important as it was in the past for online purchases.

According to the 2009–2013 Global Competitiveness Report that was 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, with Taiwan being hailed as a model for the development of global innovation and industrial clusters. Despite its ranking falling to numbers 6 and 3, respectively, in the following two years, the state of its cluster development enabled Taiwan to receive a score of 5.5 (of a 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 ([3–6] Schwab and Sala-i-Martin, 2009, 2010, 2011, 2012).

In the existing literature, many scholars have focused on R&D and new technology ([7] Audretsch and Feldman, 1996; [8] Bertscheck and Fryges, 2002; [9] Chang and Oxley, 2009). Some scholars have examined the relationship between internet intensity and urbanization economics ([10–12] Forman et al. (2005a, b, c), [13] Goldfarb and Prince (2008), [14] Baptista and Swann (1998)), as well as a link between computers and productivity ([15] Atrostic and Nguyen, 2005). However, there has been little research undertaken on the relationship between internet intensity and industrial penetration. Moreover, when we consider that for the total expenditure on internet intensity, actual figures are observed only if the firm uses the internet, which will lead to the problem of sample selection bias.

For this reason, the purpose of this paper is to include the effect of sample selection bias correction, in order to examine whether a relationship exists between industrial penetration (geographic concentration of industries) and internet intensity, and to analyse the factors determining the extent of the internet's influence. The paper is the first to investigate the issue of Industrial Penetration and Internet Intensity, based on manufacturing industry data that cover 26 two-digit industries.

The remainder of the paper is as follows. The literature on the influence of the factors that are related to internet intensity is reviewed in Section 2. In Section 3, we introduce the sample 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 an analysis of the empirical results in Section 5, and some concluding remarks and a summary of the empirical findings in Section 6.

2. Firm's Internet Intensity and Geographical Concentration

[10] Forman et al. (2005a) proposed three related theories for the relationship between internet technology and urban penetration (urbanization), 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 to break down the barriers between individuals and groups. Internet technology can make up for the disadvantages that are faced by manufacturers due to their being located far from the city's center of economic activity. For this reason, there is a substitute relationship that exists between the adoption of internet technology and urban penetration.

Urban density theory suggests that, as the density and scale of urbanization increase, the costs that are 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.

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 used widely, manufacturers had already decided where to locate their activities, so that large numbers of manufacturers using the information-intensive technology industry tended to agglomerate in certain areas. 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. Therefore, 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.

[10] Forman et al. (2005a) used US data to examine the relationship between internet intensity and urbanization, and found that, when the number of manufacturers in leading industries in urban areas increases, this will cause internet intensity in such regions to increase. This indicates that the use of the internet will be enhanced as the scale of urbanization increases, so that a complementary relationship exists between internet intensity and urban penetration. [11] Forman et al. (2005b) compared the influence of the location of enterprises and industrial penetration on internet intensity for information intensity and information-producing manufacturing industries, and discovered that, in the areas in which manufacturers were located, the larger was the scale of industrial penetration, the greater would manufacturers use the internet. A similar result from US business data from [16] Kolko (1999) also indicated a complementary relationship between the internet intensity rate and the scale of urbanization.

In an alternative investigation on information technology-related manufacturing industry in the US (computer and peripheral parts manufacturing, semiconductors, and other components manufacturing) and information technology-related service industries (software publishing, computer systems design, and related services), [17] Kauffman and Kumar (2007) tested three hypotheses: (1) internet intensity reduces the market linkages of two manufacturing and two service IT industries; (2) the effects of internet intensity on the market linkages of two manufacturing and two service IT industries will be the same for the IT-related industry and information technology-related service industries; and, (3) the effects of market linkages of two manufacturing and two service IT industries in urban and non-urban areas would be the same. Their results indicated that internet intensity would lead to a reduction in the market linkages of two manufacturing and two service IT industries, and that the internet effect would be less pronounced in urban areas than in rural areas. However, the effect of internet intensity in terms of its impact on IT-related manufacturing and information technology-related services was not found to be significantly different from each other.

[18] Galliano and Roux (2008) used a French manufacturers' sample survey data for 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 indicated that, for those manufacturers that are located in rural areas, the degree 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 penetration, the less that manufacturers used the internet, there was a substitution relationship between internet intensity and penetration.

[19] Lal (1999) used survey data for 1994 to investigate the factors affecting the manufacturers' use of the internet for India's manufacturing industry. Based on how much the sampled firms used IT technology (IT), Lal grouped the manufacturers into the following categories: (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. [19] Lal (1999) referred to four categories of factors that affected internet intensity, namely: (1) the characteristics of entrepreneurs, which included the managers' qualifications, and their ability to understand R & D, and the degree of importance that they attached to product quality and market share; (2) international orientation (how much products were imported and exported); (3) human capital; and (4) the manufacturers' scale of operations. The empirical results showed that the education of managers and the scale of the manufacturers' operations and R & D had a significant and positive impact on the use of the internet.

Moreover, [19] 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.

[8] Bertschek and Fryges (2002) used sample survey data for German companies in both the services and manufacturing industry sectors for 2000, and examined the factors affecting how 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 (internet technology supports business-to-business e-commerce); (2) had used B2B internet technology; and (3) had used B2B internet technology extensively. They used factors that 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 the plant, human capital, and international competitive pressure, as well as factors that had not been considered in the extant 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, among others.

[8] Bertschek and Fryges (2002) highlighted the following empirical findings: 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 how 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 was the likelihood that the manufacturers would themselves use new technologies.

[20] Giunta and Trivieri (2007) examined the factors determining the use of information technology (IT) by SMEs (Small and Medium-sized Enterprises) 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 degree 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 that was provided by the manufacturers for their employees, how much they engaged in R&D, the amount of outsourcing that occurred, and the degree of cooperation with other manufacturers.

[21] 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 was the contagion effect resulting from the internet technology, with there being a significant positive impact on how much the enterprises used the internet. These empirical results provided support to the theories that were advanced by [22,23] Mansfield (1963a, b) and [24] Saloner and Sheppard (1995).

The literature shows that substantial research has focused on the problems that are associated with internet intensity related to urbanization. However, few if any studies have analysed the relationship between industrial penetration and the degree to which firms have used the internet. For this reason, this paper will focus on the important but as yet not thoroughly investigated issue of internet use and industrial penetration.

3. Heckman Sample Selection Model

Manufacturing firms may make decisions to use the internet and to purchase raw materials and components online 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. Therefore, 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 online has a different preference structure from that of a non-user.

In order to draw appropriate 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 [25] 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 have joint internet purchases ([26] Lewis 1974; [25,27] Heckman 1976, 1979; [28] Greene, 1981). This method assumes that the decisions to use the internet and purchase raw materials online are made simultaneously (that is, the error terms of the two equations are correlated).

It is assumed that a zero observation represents the decision not to use the internet to purchase materials, so no individual firm is observed at the standard corner solution, namely a special solution to an agent's maximization problem, in which the quantity of one of the arguments in the maximized function is zero. Therefore, the demand curve for the internet purchaser is established only for 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 ([29] Blaylock and Blissard, 1992).

In order to correct the problem of sample selection bias, this paper uses the Heckman selection model ([24] Lewis 1974; [23,25] Heckman 1976, 1979; [28] Greene, 2003), which assumes that there exists an underlying regression relationship, as given below:

Regression equation:

$$\begin{aligned} y_i = \boldsymbol{x}_i'\beta + u_{yi}, \quad i = 1,\ 2, \ldots, n \\ u_{yi} \sim \mathrm{N}\left(0, \sigma_y^2\right) \end{aligned} \tag{1}$$

where $\boldsymbol{x}_i'$ is a vector of explanatory variables for each i, and u_{yi} is the error term for observation i. 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 the dependent variable, y_i, is observed or unobserved (that is, selected or not selected). The selection equation is given as follows:

Selection equation:

$$\begin{aligned} z_i^* = \boldsymbol{\omega}_i'\gamma + u_{zi},\ i = 1,\ 2, \ldots, n \\ u_{zi} \sim \mathrm{N}(0,\ 1) \\ corr\left(u_{yi}, u_{zi}\right) = \rho \end{aligned} \tag{2}$$

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 ($z_i^* = 0$). Therefore, the selection mechanism and regression model is reformulated as follows:

Selection mechanism:

$$\begin{aligned} z_i = \boldsymbol{\omega}_i'\gamma + u_{zi} = 1,\ \text{if } z_i^* > 0 \\ z_i = \boldsymbol{\omega}_i'\gamma + u_{zi} = 0,\ \text{otherwise} \\ \mathrm{prob}(z_i = 1|\omega_i) = \Phi\left(\boldsymbol{\omega}_i'\gamma\right)\ \text{ and} \end{aligned} \tag{3}$$

where $\Phi(\cdot)$ is the standard normal cdf (cumulative distribution function).

Regression (or observation) equation:

$$\begin{aligned}&y_i = \boldsymbol{x}_i'\beta + u_{yi},\ \text{observed only if } z_i = 1\\&corr(u_{yi}, u_{zi}) = \rho\\&(u_{zi,}, u_{yi}) \sim \text{bivariable normal } [0, 0, 1, \sigma_y^2, \rho].\end{aligned}$$

In Equation (3), the selection equation is estimated by maximum likelihood (for details, see [30] Maddala, 1983) as an independent probit model for determining the decision to join using the available information. However, [25] Heckman's (1979) two-step estimation procedure is typically 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 to compute $\hat{\lambda}_i = \varnothing(\boldsymbol{\omega}_i'\hat{\gamma})/\Phi(\boldsymbol{\omega}_i'\hat{\gamma})$ and $\hat{\delta}_i = \hat{\lambda}_i(\hat{\lambda}_i - \boldsymbol{\omega}_i'\hat{\gamma})$. The second step estimates the regression equation by ordinary least squares to obtain estimates of β and $\beta_\lambda = \rho\sigma_y$. [28,31] Green (1981, 2003) provides the statistical proof for consistency of the estimators of the individual parameters, ρ and σ_y^2.

The mean and variance of the incidentally truncated (or sample selection) bivariate normal distribution are given, respectively, as Equations (4) and (5) (the moments of the incidentally truncated bivariate normal distribution are given in [31] Green (2003b, p. 781)):

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

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

where $\alpha_z = -\boldsymbol{\omega}_i'\gamma/\sigma_z$, $\lambda_i(\alpha_z) = \varnothing(\alpha_z)/[1 - \Phi(\alpha_z)]$, $\delta_i(\alpha_z) = \lambda_i(\alpha_z)[\lambda_i(\alpha_z) - \alpha_z]$, $0 < \delta_i < 1$, $\lambda_i(\alpha_z)$ is called the inverse Mill's ratio, $\varnothing(\cdot)$ is the standard normal pdf, and $\Phi(\cdot)$ is the standard normal cumulative density function (cdf).

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

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

where the disturbance υ_i is heteroscedastic.

Ordinary least squares regression of y_i on $\boldsymbol{x}$ and λ would give a consistent estimator, but if λ is omitted, then the speciation error of an omitted variable is committed ([30] Green, 2003). The marginal effect of the regressors on y_i in Equation (6) is given as Equation (7):

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

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 $\gamma_{\mathrm{k}}\left(\frac{\rho\sigma_\varepsilon}{\sigma_{\mathrm{u}}}\right)\delta_i(\alpha_{\mathrm{u}})$. Suppose that ρ is positive and $\mathrm{E}[y_i]$ is greater when $z_i^* > 0$ than otherwise. As $0 < \delta_i < 1$, for a particular independent variable, if it appears in the probability as $z_i^* > 0$, then it will influence y_i through λ_i, and thereby reduce the marginal effect (see [31] Green 2003b, 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 internet 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 additional explanatory variable, removing the part of the error term correlated with the explanatory variable, and thereby avoiding the sample selection bias. In short, sample selection bias is corrected by the selection equation, which determines whether an observation is included in the nonrandom sample.

4. Data and Variables

In order to capture the use of the internet by manufacturers from a geographical dimension, we use unique census data for Taiwan's manufacturing firms obtained from the Directorate-General of Budget, Accounting and Statistics (DGBAS) for 2006. The sample comprises a total of 153,081 manufacturers that can be decomposed into 26 items (at the 2-digit SIC level) and 212 items at the (at the 4-digit SIC level) (SIC denotes Standard Industrial Classification). The scope of coverage includes Taiwan (Republic of China) and the Penghu archipelago, with 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 Table 1.

Table 1. Industry 2-digit codes and number of firms.

	Code	2-Digit Industry	Number of Firms
Traditional industries	08	Food	6165
	09	Beverages	644
	11	Textiles Mills	6439
	12	Wearing Apparel and Clothing Accessories	4084
	13	Leather, Fur and Related Products	1870
	14	Wood and Bamboo Products	2849
	15	Pulp, Paper and Paper Products	3605
	16	Printing and Reproduction of Recorded Media	9439
	23	Non-metallic Mineral Products	3677
	32	Furniture	2849
	33	Manufacturing Not Elsewhere Classified	5435
Technology-intensive industries	26	Electronic Parts and Components	6023
	27	Computers, Electronic and Optical Products	3717
	28	Electrical Equipment	6198
	29	Machinery and Equipment	18,545
	30	Motor Vehicles and Parts	3580
	31	Other Transport Equipment	2905
	34	Repair and Installation of Industrial Machinery and Equipment	3907
Basic industries	17	Petroleum and Coal Products	229
	18	Chemical Material	1549
	19	Chemical Products	2304
	20	Medical Goods	543
	21	Rubber Products	1756
	22	Plastic Products	11,012
	24	Basic Metal	4710
	25	Fabricated Metal Products	39,047
	Total	All manufacturing industries	153,081

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

(1) Herfindahl-Hirschman index (*HHI*): The degree of industry concentration is used to measure the extent of the competition that is faced by an industry. The HHI for industry j is calculated, as follows:

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

where s_{ij} denotes 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 *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 is the *HHI* value, the lower the degree of concentration in the industry; the higher the value, then the higher is the degree of industrial concentration.

(2) Top Four-firms Concentration Ratio (hereafter *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:

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

where s_{ij} denotes the market share of firm i in industry j, and $s_{ij} \geq s_{i\prime j}$ for all $i < i\prime$.

(3) Geographical Herfindahl-Hirschman index (hereafter *GHHI*): This is the Herfindahl index (*HHI*) for industrial market concentration combined with a geographical concept that reflects how firms are dispersed within a particular area. The formula for calculating the index is as follows:

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

where v_{jk} denotes the ratio of the number of firms in industry j in region k to the total number of firms in industry j, and 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; and, 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, but 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 $GHHI_j$ is not a part of a neighborhood messaging system, then it is not possible to reveal the differences that are brought about by being either closer to or further from a neighborhood, or to reflect the spatial correlation for different economic activities; so it is only possible to indicate that economic activities are unevenly distributed. (3) $GHHI_j$ can only reveal the spatial concentration for a single industry, without taking into account the spatial distribution characteristics for all industries as a whole.

In accordance with the literature that was discussed in Section 2, we select those factors influencing the manufacturers' use of the internet, including industrial characteristics (concentration), manufacturers' characteristics (scale of operations, manufacturers' organization, manufacturers' export intensity), geographical concentration of the 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 that are 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 degree 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 definitions of cities and counties in Taiwan have changed within the past two years, but the empirical findings are not otherwise affected.

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 (*computer*1) refers to the manufacturer having itself incurred expenses, in addition to capital expenditure on investment in computer equipment. Computer expenditure 2 (*computer*2) 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 *computer*2 variable is used to measure the contagion effect for internet technology within a certain area. Table 2 gives the variable definitions, and Table 3 provides the descriptive statistics for the explanatory variables.

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 industry j
$CR4_j$	Top Four firms Concentration Index for industry j
$export_i$	Export share for firm i = export value/total sales
$GHHI_j$	Geographic Herfindahl-Hirschman lndex for industry j
$size_i$	Firm size: total number of employees for firm i
$computer1_i$	Total expenditure on computer equipment for firm i: unit NT$1000
$computer2_i$	Total expenditure on computer equipment within the same industry and same area, excluding the expenditure of firm i: unit NT$1000
$city_i$	$city_i$ =1, if firm i is urban; $city_i = 0$, if firm i is rural
$group_i$	$group_i = 1$, if firm i has no subsidiary (branch); $group_i = 0$, otherwise

Table 3. Descriptive Statistics.

Variables (Unit)	Mean	Std Dev.	Min	Max
y_i (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

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

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

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 term. HHI_j is the Herfindahl-Hirschman index for industry j to which firm i belongs, 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 a dummy variable, indicating 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 of buying the computer equipment for firm i, and $computer2_i$ is the total cost of computer equipment within the same industry and same area, after deducting the expenditure on computer equipment of firm i itself. The variable "computer2$_i$" captures the contagion effect for internet technology in the same area and industry. The variable "$size_i$" captures the firm's characteristics. The λ_i is obtained from the sample selection 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 a binary variable, $z_i = 1$ if firm i reports use of the internet, $z_i = 0$, otherwise, and ε_{zi} is a random error term. The explanatory variables to determine whether the dependent variable z_i is observed or unobserved include industry characteristics (HHI_j), export intensity ($export\_rate_i$), geographical concentration of the industry ($GHHI_{jki}$), geographical location ($city_i$), the firm's characteristics ($size_i$), and the 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, thus reflecting the low degree of correlation between the variables. In the next section, we report the empirical results based on the Heckman two-stage estimation.

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

5. Results and Discussion

Column 2 in Tables 5 and 6 reports the Heckman two-stage estimation for Equation (8), which estimates the factors affecting the degree to which manufacturers use the internet after correcting for sample selection 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 in Tables 5 and 6 gives the coefficient estimates for the sample selection equation for Equation (9), which is estimated using a probit analysis.

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

The empirical results show that, regardless of whether the bootstrapping method is used, a non-zero Mill's lambda (β_λ) rejects the statistical null hypothesis that β_λ equals zero at the 1% level of significance, indicating that sample selection bias should be taken into account in the model. In order to make the empirical results more readily accessible, we present the results for the whole manufacturing industry, and then the results for the individual 2-digit industries.

For the whole industry, we present the results of the sample selection-corrected equation of the firm's internet use for the factors influencing the degree to which manufacturers use the internet, the marginal effects of the explanatory variables, and summarize the results of the sample selection-corrected equation for the factors that determine whether manufacturers adopt the internet for their business purposes.

Table 5. Selection corrected internet intensity (with Herfindahl-Hirschman index (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]	-
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	153,081	
# of censored observation	31,924	
Wald Chi2 (df)	543.38 (32)	

Notes: Bootstrapping standard errors are in 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 top four-firms' concentration ratio (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] ***

Table 6. *Cont.*

Variables	Intensity of Internet Use (y_i)	Select (z_i)
$GHHI_j$	−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	153,081	
# of censored observation	31,924	
Wald Chi2 (df)	561.99 (32)	

Notes: As in Table 4.

5.1. Regression Model with Sample Selection Correction for All Industries

The coefficient of HHI_j is positive, though insignificant, in column 2 of Table 5, while the coefficient of $CR4_j$ is positive and significant in column 2 of Table 6. These findings indicate that a higher degree of industrial concentration increases a firms' expenditure on internet use. The coefficient of $export_i$ is positive but insignificant in column 2 of Tables 5 and 6, indicating that export intensity has no statistical impact on the expenditure of firm on internet use.

The coefficient of $GHHI_j$ is negative and significant in column 2 of Tables 5 and 6, indicating that the lower the level of industrial penetration, the greater the degree to which manufacturers will use the internet. The coefficient of $city_i$, accordingly, has a positive and significant effect in column 2 of Tables 5 and 6.

The coefficient of $computer1_i$ shows a positive but insignificant effect in column 2 of Tables 5 and 6. These findings indicate that the manufacturers' expenditure on computer equipment has no statistical impact on the expenditure of firms on internet use. The coefficients of $computer2_i$ show a positive but insignificant effect, with bootstrapping standard errors, in column 2 of Tables 5 and 6. These empirical findings 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.

We calculate the marginal effects of Equations (7) and (8), and report the marginal effects in Table 7. Column 2 gives the industrial marginal effects with HHI_j as the proxy variable for the degree of industrial concentration, while Column 3 gives the industrial marginal effects with $CR4_j$ as the proxy variable for the degree of industrial concentration.

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

For the HHI_j variable, the marginal effects are −0.0902 for column 2 and −0.007 for column 3, in Table 7. For example, the figure −0.0902 means that, when the degree of the industrial concentration rate increases by 1, the degree to which manufacturers use the internet is reduced by 0.0902%. This result indicates that the lower the degree of industrial concentration, the greater the degree to which manufacturers will use the internet.

Not surprisingly, there are differences between the marginal effects of HHI_j and $CR4_j$ on the degree to which manufacturers use the internet, as described in Section 4: HHI_j takes into account all the firms in an industry, use the manufacturer's market share as weights, with smaller firms being given lesser weights and larger firms being given greater weights, while $CR4_j$ only considers the weighted average of the market shares of the top four firms in an industry. However, the empirical findings of industrial concentration agree with those of [18] Galliano and Roux (2008) and [21] Galliano et al. (2011), who used French manufacturing industry data.

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

For the $GHHI_j$ variable, the marginal effects are (−0.0245, −0.0133) for columns 2 and 3, respectively, in Table 7. For example, when industrial penetration is reduced by 1, the degree to which manufacturers use the internet will increase by 0.0245%. Therefore, a substitute relationship exists between the degree to which manufacturers use the internet and the level of industrial penetration, an empirical result that accords with the results obtained in [17] Kauffman and Kumar (2007), who used US information technology-related manufacturing and service industry data, and [18] 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 especially 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 effects are (−0.0051, −0.0062) for columns 2 and 3, respectively, in Table 7. For example, manufacturers that are located in urban areas will use the internet less by −0.0051% compared with those that are located in rural areas. In other words, manufacturers that are located in rural areas will use the internet to a greater extent than those that are located in urban areas. These results confirm the empirical finding in Forman et al. (2005) and Kolko (1999) that a complementary relationship exists between internet intensity and the degree of urbanization.

We now present the results in column 3 of Tables 5 and 6 that show the probit estimates, as given by Equation (9), which estimate the factors that determine whether manufacturers will use the internet for their business purposes.

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_j and $CR4_j$ are negative and significant at the 1% level of significance in column 3 of Tables 5 and 6. These results indicate that the greater is the competition that manufacturers face, in order to increase their ability to compete with other manufacturers, the more likely they will be inclined to use the internet for their business purposes.

Export intensity is also an important factor for affecting the manufacturers' use of the internet. The coefficients of $export_i$ are positive and significant at the 1% level of significance in column 3 of Tables 5 and 6. This is not surprising as, the greater will manufacturers rely on exports, the greater will be their export intensity, and the greater will be the use of the internet to communicate with their overseas customers.

The coefficient of geographical location of $city_i$ in column 3 of Tables 5 and 6 shows a negative and significant effect on manufacturers' use of the internet for their business purposes. This empirical result suggests that manufacturers that are located in rural areas will be more likely to use the internet for business than those located in urban areas. However, this finding is in contrast with the empirical results of the coefficient of $city_i$ in column 2 of Tables 5 and 6, which suggest that manufacturers that are located in urban areas will spend more money on internet use than firms in rural areas.

The coefficient of the manufacturer's scale of operations, $size_i$ shows a positive and significant probability to use the internet for their manufacturing business. It is not surprising that larger firms will be more likely to use the internet for business. Moreover, a 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 that have a subsidiary (branch). 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 penetration on the manufacturers' use of the internet is not significant in column 3 of Tables 5 and 6, the effect on the degree to which manufacturers use the internet is significant and negative in column 2 of Tables 5 and 6. This indicates that the degree of industrial penetration does not affect whether manufacturers will use the internet, but it will affect the degree to which manufacturers that already use the internet will continue to do so.

5.2. Regression Model with Sample Selection Correction for 2-Digit Industries

This section reports the Heckman two-stage estimates with *HHI* as the only proxy variable for the degree of industrial concentration, and marginal effects for two-digit industries, in Tables 8 and 9, respectively. A nonzero Mill's lambda (β_λ) rejects the statistical null hypothesis that β_λ equals 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, and (32) Furniture. However, as the industries are different, the empirical results for individual industries that are based on the two-digit level classifications also vary. For individual two-digit industries, we first discuss the results of the sample selection-corrected equation for how much manufacturers use the internet, then present the results of the sample selection-corrected equation for the factors that determine whether manufacturers use the internet, and finally summarize the marginal effects.

The effect of the degree of industrial penetration ($GHHI_j$) in terms of how much manufacturers use the internet vary across the 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, 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, all show that the lower is the level of industrial penetration, the greater is the degree 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 that the higher the degree of industrial penetration, the greater is the degree to which manufacturers will use the internet.

The effect of the degree of industrial concentration (HHI_j) in terms of how much manufacturers will use the internet also differs 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 that the higher is the degree of industrial concentration, the greater is the degree to which 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 that the lower the degree of industrial concentration, the greater the degree to which manufacturers will use the internet.

The variable $export_i$ show a positive and significant influence on how much 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 negative effect on the degree to which manufacturers use the internet for traditional industries.

The effect of geographic location, $city_i$ shows manufacturers that are located in rural areas will use the internet to a greater extent than those that are located in urban 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 urban areas will use the internet to a greater extent than those that are located in rural areas.

The variable of manufacturers' expenditure on computer equipment, $computer1_i$, has no statistical impact on the expenditure 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, computer2_i that is used to capture the contagion effect for internet technology in the same area shows no statistical impact on the expenditure of the firm on 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.

The following discussion presents the probit estimates, as given by Equation (9), which estimates whether or not manufacturers adopt the internet for their business across 2-digit industries. The coefficient estimates are given in Table 8.

The effect of the degree of industrial penetration ($GHHI_j$) in terms of whether or not manufacturers will use the internet shows differences across 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 penetration 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 penetration is high, manufacturers will be less inclined to use the internet.

However, industrial penetration will generally 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 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 manufacturers will use the internet is 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, 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, 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 industrial concentration decreases, 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 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, in the case of 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 $size_i$ shows a positive effect for affecting the manufacturers' decision to use the internet for most 2-digit industries. Moreover, the coefficient of $group_i$ has a positive and significant effect on manufacturers' decision to use the internet for most 2-digit industries.

The following discussion presents the total marginal effects of each of the explanatory variables on how 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 equal zero at the 10% level of significance, with bootstrapping standard errors, 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, thereby making it necessary to correct for sample selection bias.

In the following paragraphs, we present the marginal effects, as given by Equations (7) and (8). In terms of industrial penetration ($GHHI_j$), among 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.

The marginal effects in various industries can be summarized, as follows. Industrial concentration (HHI_j), among traditional industries, has the largest marginal effect at 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 the basic industries, the largest value is 2.3671 for (22) Plastic Products, while the smallest value is −0.2068 for (24) Basic Metal.

The marginal effect of export intensity ($export_i$), among traditional industries, has the largest at 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.

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

Among traditional industries, the marginal effect of manufacturer's scale of operations, ($size_i$) has the largest value at 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.

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

Notes: To save space, we do not present Industries (17) Petroleum and Coal Products industry and (34) Repair and Installation of Industrial Machinery and Equipment, in Tables 8 and 9. Some coefficients of explanatory variables were not estimated for the sample selection correction regression model.

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

The marginal effect of manufacturers' expenditure on computer equipment, computer1_i, among traditional industries, has the largest value at 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.

Finally, the marginal effect of the manufacturers' expenditure on computer equipment within the same industry and region ($computer2_i$), 0.0045 for (15) Pulp, Paper and Paper Products, 0.0025 for (27) Plastic Products, and 0.0008 for (24) Basic Metal, have the largest values for the traditional industries, for technology-intensive industries, and for basic industries, respectively.

6. Conclusions

The paper is the first to investigate the issue of Industrial Penetration and Internet Intensity, which is based on manufacturing industry data that cover 26 two-digit industries. This paper used Taiwanese manufacturing industry census data that was 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 considered the total expenditure on the internet for purchase and sales products, an actual figure was observed only if the firm used the internet, thereby causing the problem of sample selection bias. In order to correct the problem of sample selection bias, the paper used the Heckman sample selection model and two-stage estimation procedure to obtain consistent estimates of the parameters of the sample selection model.

In order to improve the efficiency in estimation, we used the bootstrapping approach to estimate the sample variance. The empirical results showed that, regardless of whether we used the bootstrapping approach, the Mill's lambda test statistic rejected the null hypothesis that β_λ equals zero at the 1% level of significance for the aggregated full industry, and 7 of 26 industries rejected the null hypothesis that β_λ equals zero at the 10% level of significance, indicating that the problem of sample selection bias needed to be corrected.

The main points arising from the paper, which are summarized in Table 10, 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 surprising that the likelihood of such manufacturers using the internet is relatively high, with the manufacturers' independence of operations having the greatest impact. The second most influential factor is the manufacturers' export intensity, indicating that the more 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 greater is the competition that manufacturers face, in order to increase their ability to compete with other manufacturers, the more likely that they be inclined to use the internet. The empirical results also show that manufacturers that are located in rural areas would be likely to use the internet for business than those that are located in urban areas, and larger firms would be more likely to use the internet for business than smaller firms. However, the impact of the degree of industrial penetration on the manufacturers' use of the internet is not significant.

(2) The degree to which manufacturers use 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. This suggests that the extent of industrial penetration does not affect whether or not manufacturers will use the internet, but it will affect the extent to which manufacturers that already use the internet will continue to use the internet. The results suggest

there exists a substitute 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 it is no longer especially important.

(3) The variable of industrial penetration has a negative marginal effect on the extent to which the manufacturers use the internet, indicating that there exists a substitute relationship between the extent to which the manufacturers use the internet and the level of industrial penetration. Such results confirm the research by [17] Kauffman and Kumar (2007), who used US information technology-related manufacturing and service industry data, and [18] Galliano and Roux (2008), who used French manufacturing data.

(4) The more competitive is the industry, manufacturers will increasingly need to use the internet to communicate and trade with other entities, and to increase their competitiveness. The empirical findings agree with those of [18] Galliano and Roux (2008) and [21] Galliano et al. (2011), who used French manufacturing industry data.

(5) The 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 internet intensity. The second and third largest effects are the manufacturers' expenditure on computer equipment and the contagion effect, both of which have a positive marginal effect on the degree to which manufacturers use the internet, thought that the magnitudes for both marginal effects are quite small.

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

(7) The marginal effect of localized penetration on how the manufacturers use the internet also varies widely. The largest positive and smallest negative values for the traditional industries are, respectively, 0.0266 for (08) Food and -0.0018 for (11) Textiles Mills; the largest and smallest values for technology-intensive industries are, respectively, 0.0527 for (26) Electronic Parts and Components and -0.0249 for (27) Plastic Products; the largest and smallest values for basic industries are, respectively, 0.0578 for (21) Rubber Products and -0.0216 for (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. The 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.

Further to what has been investigated in the paper, internet technology has rapidly replaced long-distance non-electronic communications, and has also reduced the costs of relaying information over long distances. The empirical results that have been presented above have demonstrated clearly the relationship between industrial penetration and internet intensity, specifically that the lower the level of industrial penetration, the greater the degree to which manufacturers will use the internet. For future research, we intend to evaluate the economic performance of firms for which their location is far removed from others, while using internet technology to advance their business operations.

Table 10. Summary of Main Points

Overall results	1	Manufacturers decide to use the internet for five primary reasons.	Five primary reasons: (i) degree of industrial concentration; (ii) export intensity; (iii) geographical location; (iv) manufacturer's size of operations; (v) independence of operations.
	2	The extent to which manufacturers use the internet is influenced by three factors.	The three factors are: (i) degree of industrial penetration; (ii) geographical location; (iii) contagion effect.
	3	Industrial penetration has a negative marginal effect on the extent to which manufacturers use the internet.	Such empirical results support the research by [17] Kauffman and Kumar (2007), who used US information on technology-related manufacturing and service industry data, and [18] Galliano and Roux (2008), who used French manufacturing data.
	4	The more competitive is the industry, manufacturers will need to increase their competitiveness and use the internet more to communicate and trade with other entities.	The empirical findings agree with those of [18] Galliano and Roux (2008) and [21] 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.	This empirical finding indicates that international competition has a relatively large influence on the extent of internet intensity.
Decomposed industry to 2-digit industry	6	The empirical results for the individual industries based on the two-digit level classifications can vary substantially.	The same outcome holds for the marginal effect of localized penetration on the variable extent to which manufacturers use the internet.
	7	Industries with a higher degree of export intensity, and with a greater reliance on exports, will have a higher degree of internet intensity among manufacturers that use the internet.	This empirical finding seems to be a novel result.

Author Contributions: Chia-Lin Chang conceived the idea of the paper, and collected the data; Michael McAleer analysed the model specifications; Yu-Chieh Wu estimated the models; Chia-Lin Chang discussed the empirical results; Michael McAleer checked the empirical results and the contents of the paper.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Lin, Y.C. *2009 Industry Information Application Status and Demand Survey Report*; Department of Industrial Technology, Ministry of Economic Affairs, R.O.C.: Taipei, Taiwan, 2009. (In Chinese)
2. Lin, Y.C. *2010 Industry Information Application Status and Demand Survey Report*; Department of Industrial Technology, Ministry of Economic Affairs, R.O.C.: Taipei, Taiwan, 2010. (In Chinese)
3. Schwab, K.; Sala-i-Martin, X. *The Global Competitiveness Report, 2009–2010*; World Economic Forum: Geneva, Switzerland, 2009.
4. Schwab, K.; Sala-i-Martin, X. *The Global Competitiveness Report, 2010–2011*; World Economic Forum: Geneva, Switzerland, 2010.
5. Schwab, K.; Sala-i-Martin, X. *The Global Competitiveness Report, 2011–2012*; World Economic Forum: Geneva, Switzerland, 2011.
6. Schwab, K.; Sala-i-Martin, X. *The Global Competitiveness Report, 2012–2013*; World Economic Forum: Geneva, Switzerland, 2012.
7. Audretsch, D.B.; Feldman, M.P. R&D Spillovers and the Geography of Innovation and Production. *Am. Econ. Rev.* **1996**, *86*, 630–640.
8. Bertschek, I.; Fryges, H. *The Adoption of Business-to-Business E-Commerce Empirical Evidence for German Companies*; ZEV Discussion Paper; Centre for European Economic Research (ZEW): Mannheim, Germany, 2002; pp. 02–05.
9. Chang, C.-L.; Oxley, L. Industrial Agglomeration, Geographic Innovation and Total Factor Productivity: The Case of Taiwan. *Math. Comput. Simul.* **2009**, *79*, 2787–2796. [CrossRef]
10. Forman, C.; Goldfarb, A.; Greenstein, S. Geographic Location and the Diffusion of Internet Technology. *Electron. Commer. Res. Appl.* **2005**, *4*, 1–13. [CrossRef]
11. Forman, C.; Goldfarb, A.; Greenstein, S. How Do Industry Features Influence the Role of Locationon Internet Adoption? *J. Assoc. Inf. Syst.* **2005**, *6*, 383–408.
12. Forman, C.; Goldfarb, A.; Greenstein, S. How Did Location Affect the Adoption of the Commercial Internet? Global Village vs. Urban Density. *J. Urban Econ.* **2005**, *58*, 389–420. [CrossRef]
13. Goldfarb, A.; Prince, J. Internet Adoption and Usage Patterns are Different: Implications for the Digital Divide. *Inf. Econ. Policy* **2008**, *20*, 2–15. [CrossRef]
14. Baptista, R.; Swann, P. Do firms in clusters innovate more? *Res. Policy* **1998**, *27*, 525–540. [CrossRef]
15. Atrostic, B.K.; Nguyen, S.V. IT and Productivity in U.S. Manufacturing: Do Computer Networks Matter? *Econ. Inq.* **2005**, *45*, 493–506. [CrossRef]
16. Kolko, J. The Death of Cities? The Death of Distance Evidence from the Geography of Commercial Internet Usage. In *Mimeo*; Harvard University: Cambridge, MA, USA, 1999.
17. Kauffman, R.; Kumar, A. The New Economic Geography of IT Industries the Impacts of the Internet on Their Market Linkages and Agglomeration. In Proceedings of the 28th Annual International Conference on Information Systems (ICIS), Montreal, QC, Canada, 9–12 December 2007.
18. Galliano, D.; Roux, P. Organisational Motives and Spatial Effects in Internet Adoption and Intensity of Use: Evidence from French Industrial Firms. *Ann. Reg. Sci.* **2008**, *42*, 425–448. [CrossRef]
19. Lal, K. Determinants of the Adoption of Information Technology: A Case Study of Electrical and Electronic Goods Manufacturing Firms in India. *Res. Policy* **1999**, *28*, 667–680. [CrossRef]
20. Giunta, A.; Trivieri, F. Understanding the Determinants of Information Technology Adoption. *Appl. Econ.* **2007**, *39*, 1325–1334. [CrossRef]
21. Galliano, D.; Roux, P.; Soulie, N. ICT Intensity of Use and the Geography of Firms. *Environ. Plan. A* **2011**, *43*, 67–86. [CrossRef]
22. Mansfield, E. The Speed of Response of Firms to New Techniques. *Q. J. Econ.* **1963**, *77*, 290–311. [CrossRef]
23. Mansfield, E. Intrafirm Rates of Diffusion of an Innovation. *Rev. Econ. Stat.* **1963**, *45*, 348–359. [CrossRef]

24. Saloner, R.; Sheppard, J.M. Adoption of Technologies with Network Effects: An Empirical Examination of the Adoption of Automated Teller Machines. *Rand J. Econ.* **1995**, *26*, 479–501. [CrossRef]
25. Heckman, J.J. Sample Selection Bias as a Specification Error. *Econometrica* **1979**, *47*, 153–161. [CrossRef]
26. Lewis, H.G. Comments on Selectivity Biases in Wage Comparisons. *J. Polit. Econ.* **1974**, *82*, 1145–1155. [CrossRef]
27. Heckman, J.J. The Common Structure of Statistical Models of Truncation, Sample Selection and Limited Dependent Variables and a Simple Estimator for Such Models. *Ann. Econ. Soc. Meas.* **1976**, *5*, 475–492.
28. Greene, W. Sample Selection Bias as a Specification Error: Comment. *Econometrica* **1981**, *49*, 795–798. [CrossRef]
29. Blaylock, J.R.; Blissard, W.N. Self-valuated Health Status and Smoking Behaviour. *Appl. Econ.* **1992**, *25*, 429–435. [CrossRef]
30. Maddala, G.S. *Limited-Dependent and Qualitative Variables in Economics*; Cambridge University Press: New York, NY, USA, 1983.
31. Greene, W. *Econometric Analysis*; Macmillan: New York, NY, USA, 2003.