

Determinants of labour productivity in Mexico: an approach from the endogenous growth theory using artificial neural networks

Héctor Eduardo Díaz Rodríguez and Fidel Aroche Reyes

Abstract

The widespread adoption of information and communications technology (ICT) has increased research on the link between such technology and economic growth. The use of ICT has a significant impact on growth, but studies tend to ignore the determinants of this use and the differing conditions across the companies, industries and countries that adopt them. In this study, we analyse the determinants of the differentiated impact of ICT on labour productivity growth in Mexican companies, using microdata from the ICT survey of the National Institute of Statistics and Geography (INEGI) and based on factor analysis and artificial neural network techniques. The results indicate that the strength of companies' links with the external sector and their capacities (educational level of staff and organizational maturity) determine the impact of ICT on labour productivity in Mexico.

Keywords

Employment, labour, labour productivity, innovations, information technology, communication technology, computer networks, Mexico

JEL classification

C46, O47, O33

Authors

Héctor Eduardo Díaz Rodríguez is a professor and Director of Undergraduate Studies of the Professional Studies Division in the School of Economics of the National Autonomous University of Mexico (UNAM). Email: hectoreduardo12@comunidad.unam.mx.

Fidel Aroche Reyes is a professor in the Graduate Studies Division of the National Autonomous University of Mexico (UNAM). Email: aroche@unam.mx.

I. Introduction

The past 20 years have seen a boom in research on the impact of the adoption of technologies worldwide, in particular information and communications technology (ICT), as well as on their increasing importance in production processes. These studies began to multiply shortly after the emergence in the 1990s of endogenous growth theories (known as *AK* models), which hold that, in addition to technology, investment in human capital is an important driver of productivity growth. Paradoxically, however, most early analyses of the impact of ICT overlooked the importance of workforce education and training as a determinant of the benefits of technology (Brynjolfsson, 1993 and 1996; Brynjolfsson and Hitt, 1996; Bresnahan, 1997 and 1999; Lichtenberg, 1995; Jorgenson and Stiroh, 2000). In other words, product and technology innovation have played a major role in this regard.

Only recently have some studies recognized that the differences in the initial conditions various agents (companies, industries or countries) encounter are key to understanding the relationship between ICT and productivity; however, there remains a dearth of such studies in the existing literature and they focus primarily on developed economies (Schreyer, 2000; Daveri, 2000; Dutta and Bilbao-Osorio, 2012; Bilbao-Osorio, Dutta and Lanvin, 2013; ITU/UNESCO, 2013). Research on the impact of ICT on developing economies has been scarce at best, with little consideration given to the role of education in that nexus. This is evident from the few studies on the role of ICT and its impact on productivity in Mexico (Aravena, Cavada and Mulder, 2012; INEGI, 2013; UNCTAD, 2003). In view of the foregoing, this paper aims to add to the discussion on the role of workforce education and training as determining factors in the exploitation of technology in Mexico, drawing on some principles of endogenous growth theory. This is done through an analysis of microdata collected from the ICT survey (ENTIC) conducted by INEGI, using factor analysis and artificial neural networks methodologies (Haykin, 1999; Arrieta, Torres and Velásquez, 2009; Pitarque, Roy and Ruiz, 1998; Larrañaga, Inza and Moujahid, 2003).

The assumption is that while ICT is a prerequisite for labour productivity growth, it alone is insufficient for that growth to occur — organizations and workers must possess the requisite capacities that will enable them to incorporate technology into production and business processes and to fully exploit that technology. Without such capacities, ICT availability becomes moot.

The rest of the paper is organized as follows. Section II reviews the studies on the impact of ICT on productivity, with a view to describing some of the models derived from endogenous growth theory. Section III describes the factor analysis and artificial neural network methodologies used and section IV presents and analyses the results obtained from applying the method to the survey data. Section V concludes and presents possible public policy options.

II. Literature review

The well-known Solow model (1957) is one of the first approximations of the impact of technological change and factor growth on productivity, using a Cobb-Douglas production function that measures the efficiency with which factors are used in the production process. Thus, productivity growth is possible through two channels: as a result of an increase in the capital-per-worker ratio, or from the use of technology, which, in addition to widening that ratio, leads to more efficient use of the combination of factors. However, the model does not explain technological change or the reasons behind the improved exploitation thereof. The specific characteristics of the Solow model made it the model of choice for subsequent studies analysing the link between ICT and productivity.¹

¹ This is because the model can break growth down into both the underlying factors (capital, labour and more efficient use of combinations of the two) and the type of capital (ICT or non-ICT investment) that drive growth.

Because of the inability of the Solow model to explain the sources of technical change, in the early 1990s a new wave of studies emerged which are based on that model but incorporate the role of investment in education as a catalyst of technical progress (Romer, 1990 and 1994; Aghion and others, 1998).

Early research on the effect of computer use on productivity in the United States economy found no direct link between the two. This gave rise to what was known as the “productivity paradox”, prompting further research (based on longer and more reliable data series as well as more refined methods), with notable studies by Brynjolfsson (1993 and 1996), Brynjolfsson and Hitt (1996), Bresnahan (1997 and 1999), Lichtenberg (1995) and Jorgenson and Stiroh (2000). This collection of work pointed to a significant change in productivity resulting from the use of ICT. To differentiate the impact of the exponential use of ICT in organizations and from the adoption of other types of technology, this series of studies (based on the Solow model) identified a fundamental difference between investments in ICT and in other types of capital, which is that the former has a dual role. First, like other types of capital, ICT can be seen as a new production technology that improves labour productivity; second, it can also be considered a transformative element of production processes.

The impact of that second aspect, which triggers changes in business processes both within and between organizations, has been shown to be greater in studies by Bresnahan, 1997; Gurbaxani and Whang, 1991; Malone, Yates and Benjamin, 1989; Hollenstein, 2004; Capel and Bosch, 2004; Inklaar, O’Mahony and Timmer, 2005; Bayo-Moriones and Lera-López, 2007; Carrera Portugal, 2010; Tödtling, Grillitsch and Höglinger, 2012; and Peppard and Ward, 2016. In this sense, this body of research credits ICT with promoting changes in business processes and, with them, in multifactor productivity. Greenan, Mairesse and Topiol-Bensaid (2001) analysed the impact of ICT on some French companies and their findings were similar to those of Bresnahan (1997) and Lichtenberg (1995).

In contrast, Lal (2001) found no link between ICT investments and productivity in the Indian textile industry. There is a general pattern of contrasting findings in studies on firms located in developed economies and those in developing economies, as shown by Kraemer and Dewan (2000) and Pohjola (2001), whose observations indicate that, at the firm level, ICT tends to bring about productivity changes in developed countries but not in developing countries. Kraemer and Dewan (2000) and Pohjola (2001) posit that these differences are attributable to labour costs in the two groups of countries. In developing countries, labour costs tend to be lower and the costs of access to capital high, making the substitution of labour by capital relatively more onerous. The opposite is true of developed countries.²

However, the plethora of studies based on the Solow model has failed to explain why, despite the exponential increase in ICT adoption worldwide, patterns of productivity and economic growth — which should follow therefrom — have diverged so markedly. More often than not, there has been a tendency towards stagnation, as exemplified by the Mexican economy. Consequently, more recent studies have begun to analyse factors associated with the context of economies and enterprises, as well as differences in their initial conditions, with a view to accounting for such disparate productivity outcomes despite similar levels of ICT uptake. These include papers by Aramendia-Muneta and Ollo-López (2013), Tarutè and Gatautis (2014) and a series of analyses conducted by international development and cooperation agencies (World Bank, 2011 and 2012; Dutta and Bilbao-Osorio, 2012; Bilbao-Osorio, Dutta and Lanvin, 2013; and OECD, 2012), which conclude that inflexible institutional factors which make it difficult to take full advantage of the potential benefits of ICT are the reason why technology has failed to create an impact on economies or the firms operating therein. The conclusion of these agencies has not changed from that drawn 30 years ago: general conditions for high and sustained growth must be created and these include openness to trade and capital flows, an appropriate institutional and legal

² In this study, we posit that the variations in how ICT affects countries and economic activities are owed mainly to differences in the level of education of the workforce, in the maturity of the companies in each group, and in the sectoral composition of the companies.

framework, flexibility in the production structure and efficient markets (World Bank, 2011 and 2012; Dutta and Bilbao-Osorio, 2012; Bilbao-Osorio, Dutta and Lanvin, 2013; OECD, 2012).

This paper does not establish a direct link between ICT availability and economic performance, which is seen as a change in factor productivity or aggregate economic growth; rather, that link is potential and is contingent on the economies, industries or firms involved possessing a set of characteristics and capacities to leverage the use of ICT infrastructure. How these determinants influence productivity depends on the situation, the economy in question, the possession of a set of skills concomitant with workers' educational level and organizational maturity and which is reflected in the ability to incorporate technology into production processes, and the very nature of the economic activities conducted.

Despite their focus on developed economies, the aforementioned Solow model and the studies reviewed in this section are incapable of explaining the differences in productivity associated with the adoption of ICT; in the case of developing economies, the paucity of research on this issue reinforces the need for a different perspective. In the specific case of the Mexican economy, the widespread adoption of ICT by companies has not led to increased productivity and it is worthwhile to consider possible explanations in that regard.

III. Methodology and characteristics of data

To understand why the adoption and expansion of ICT have not had a positive effect on productivity in Mexico, this paper analyses the database created from the ICT survey (ENTIC) published by INEGI in 2013. It contains information on 389 variables on ICT use for a sample of 6,468 companies, representative of 157,611 organizations that have ten employees or more and which were classified in 76 subsectors the previous year, based on the North American Industry Classification System (NAICS). With a view to applying the model presented below, 34 variables were selected to measure how organizations incorporate information and communications technology into their production and business processes: 13 variables were used directly and 21 variables were used to construct two indices, one on Internet use and the other on software use.³

The following criteria were used to group the variables in these indices: each of the available variables on Internet, software and cloud use was sorted based on the skills required; a rating was then applied to each variable,⁴ with a lower rating for uses classified as basic⁵ and a higher rating for advanced uses. Through this procedure, the 21 variables representing the various uses of the Internet and software were narrowed down to two: the Internet usage index and the software usage index. These indices are included as variables for factor analysis.

³ Use and exploitation is a concept that describes the ways in which organizations use available technology and benefit from its advantages. The concept of use is not linear, since there are different levels at which the use of technology is incorporated into production processes; the more these uses affect the development of processes and lead to organizational change, the greater the impact of the technology.

The aim of the proposed classification is not to assign a specific qualification for each use, but simply to rank them according to the level of skills required for a specific use.

⁴ The variables and the respective scores used to construct the Internet use index are as follows: general information search (4), information transfer (6), financial transactions (8), access to government portals (10), payment of non-government services (12), supplier care (14), recruitment and staffing (16), customer care (18), corporate advertising (20) and distance training (22). The variables and the respective scores used to construct the software use index are as follows: programs for invoicing (4), accounting (6), payroll administration (8), purchasing and payment management (10), human resources (12), sales (14), inventory (16), business administration (18), information analysis (20), design (22) and logistics (24).

⁵ A use is considered basic if it requires little skill; the more complex the use, the greater the skill required. For example, in the case of Internet use, the variable for general information search demands little capacity from human resources and organizations. In comparison, the distance training variable requires greater capacities, such as the development of a culture that promotes training, the necessary infrastructure, and technological know-how on the part of the workforce, among others.

1. Factor analysis

Factor analysis can be used to group correlated variables and find homogeneous groups, particularly in databases with a high number of series. In principle, these groups are independent of each other. Using factor analysis, it is possible to find the minimum number of dimensions capable of explaining the maximum amount of information contained in the data. Put another way, a factor is a qualitative dimension on a coordinate axis, which defines how entities (observations) differ, just as the size or taste of an object defines its qualitative dimensions (Domínguez and Brown, 2004).

In formal terms, factor analysis is a multivariate method that expresses p observable variables as a linear combination of hypothetical or latent variables (m), called factors. These variables or common factors are obtained and interpreted from the matrix of correlations between the variables:

$$R = \begin{pmatrix} 1 & r_{12} & \dots & r_{1p} \\ r_{21} & 1 & \dots & r_{2p} \\ \vdots & \vdots & \ddots & \vdots \\ r_{p1} & \vdots & \dots & 1 \end{pmatrix} \quad (1)$$

In the case of multifactor models, the p observable variables (variables from the ICT survey) depend on m latent variables or common factors, in addition to unique factor p , as denoted by the following linear model:

$$X_1 = a_{11}F_1 + \dots + a_{1m}F_m + d_1U_1 \quad (2)$$

$$X_2 = a_{21}F_1 + \dots + a_{2m}F_m + d_2U_2 \quad (3)$$

...

$$X_p = a_{p1}F_1 + \dots + a_{pm}F_m + d_pU_p \quad (4)$$

The multifactor analysis model is based on two fundamental assumptions:

(i) Common factors and unique factors are not correlated, i.e:

$$\text{Cor}(F_i, U_j) \quad i = 1, \dots, m \quad j = 1, \dots, p \quad (5)$$

(ii) The common factors and the unique factors have a mean of 0 and variance of 1:⁶

$$\mu = 0, \sigma = 1$$

In the factor analysis model, overall variance is explained by two groups of factors — common factors and unique factors— as illustrated below:

$$\text{Var}(X_i) = a_{i1}^2 + \dots + a_{im}^2 + d_i^2 \quad (6)$$

That is, a_{ij}^2 is the share of variation of X_i that is explained by the existence of a common factor (F_j), while d_i^2 is interpreted as the share that is explained solely by the unique factor U_i .

The share of the pooled variance of variable X that is explained by common factors is called the commonality variable and is derived from the following function:

$$h_i^2 = a_{i1}^2 + \dots + a_{im}^2 \quad (7)$$

⁶ However, since common factors are independent dimensions in the linear sense and both common and unique factors are conventional variables, there is a possibility that independent (common) vectors may not have a mean of 0 and a variance of 1 (Cuadras, 2007).

2. Artificial neural networks

A neural network is a massively parallel distributed processor [...] which has a natural propensity for storing experiential knowledge and making it available for use (Haykin, 1999). The author notes that it resembles the brain in two respects:

- (i) Knowledge is acquired by an artificial neural network, like its biological counterpart, through a learning process.
- (ii) Interneuron connection strengths, known as synaptic weights, are used to store the acquired knowledge.

The knowledge acquired in an artificial neural network is applied to estimate the behaviour of a given variable using variables with input information; between the input variables (input layer) and the output variables (output layer) is an intermediate area called the hidden layer, in which the information and the model used to analyse it are processed.

While a regression model can be described as a specific type of network, in the sense that regression coefficients can acquire and store knowledge, regression has a rigid model structure and an underlying structure to assumptions, which are both imposed from the outset. This is in contrast to an artificial neural network model, where the form of the relationships is determined during the learning process. If a linear relationship between dependent and independent variables is adequate, the results of the neural network should match those of the linear regression model as closely as possible. If a non-linear relationship is more appropriate, the neural network will automatically approximate the “correct” model structure.

Larrañaga, Inza and Moujahid (2003) hold that artificial neural networks can be described as a directed graph with the following properties:

- (i) Each node (neuron) i is associated with a state variable X_i .
- (ii) Each synaptic connection (i, j) between nodes (neurons) i and j is assigned a weight $w_{ij} \in \mathbb{R}$.
- (iii) Each node (neuron) i is associated with a threshold $\theta_i \in \mathbb{R}$.
- (iv) For each node i , a function $f_i(X_p, \dots, X_r, W_{ip}, \dots, W_{in}, \theta_i)$ is defined according to the connection weights, the threshold and the states of the nodes j that are connected to node i . The value of this function determines the new state of the node.

In an artificial neural network, the nodes are connected through contact between the neural functions in a process known as synapsis. The behaviour of an artificial neural network will ultimately be determined by the structure of the synaptic connections, known as the network architecture. This denotes the topology, structure or pattern formed by the synaptic connections of an artificial neural network. These networks can define both linear and non-linear parameters that are mapped from an input to an output, such as $Y=Y(X;W,A)$, which is a continuous function of both the inputs and the parameters (W) of the network architecture (A) (MacKay, 2003). Networks can be trained to perform regression, classification and prediction tasks.

The input layer of an artificial neural network is the layer composed of the set of neurons responsible for receiving data and information from the environment; in this layer of the structure, neurons do not have input synapses, only output synapses. Next are the hidden layers, which are not directly connected with the environment, but receive information and data from the input layer; it is here that the characteristics of the environment are estimated, modelled and represented. This is the only layer that is neither input nor output, and is formulated thus:

$$\text{Hidden layer: } a_j^1 = \sum_l W_{jl}^1 X_l + \theta_j^1; \quad h_j = f^1(a_j^1) \quad (8)$$

Lastly, the output layer comprises the sets of neurons that provide the network response. The neurons in this layer do not have output synapses.

$$\text{Output layer: } a_j^2 = \sum_l W_{jl}^2 X_l + \theta_j^2; \quad h_j = h^2(a_j^2) \quad (9)$$

One of the major differences between existing types of artificial neural networks is the number of layers used in the learning process (hidden layers). In general, using more layers makes it possible to model more complex processes. For the purposes of this study, we use a multilayer perceptron network.

(a) Multilayer perceptron

The multilayer perceptron is an artificial neural network that has one or more hidden layers, allowing for the detection of more complex patterns, which may or may not be separated by more than one hyperplane. The structure of this artificial neural network arguably makes it the most suitable for analysing the relationship discussed in this study as it can capture a wide range of functional relationships, even non-linear ones.⁷ A multilayer perceptron is a feedforward network, meaning that it reacts to changes in its environment to maintain a specific state in the system. It has three fundamental elements (Bishop, 1995):

- (i) Input layer (input variables)
- (ii) Hidden layers
- (iii) Output layer (output variables)

The hidden layers model the functional forms of the data obtained in the input layer through a process called neuron training. For this process, a subset of the data $D = \{X^{(n)}, t^{(n)}\}$ is used and the W of the input function is adjusted, minimizing the error function through the gradient descent method given by the backpropagation algorithm:

$$E_D(w) = \frac{1}{2} \sum n \sum i \left(t_i^{(n)} - y_i(x^{(n)}; w) \right)^2 \quad (10)$$

This minimization continually calculates the E_D gradient by means of the chain rule to find the derivatives through the following steps:

- (i) Synaptic weights and initial thresholds are established.
- (ii) One of the bases is processed to obtain a response from the network with respect to the r -th pattern.
- (iii) Errors are interpreted as error signals associated with the process and are calculated by:

$$\left(\sum_{k=1}^s \left(\sum_{j=1}^o W'_{kj} Y_j^r - \theta_k \right) W_{kj} \right) \frac{\partial f \left(\sum_{i=1}^n W_{ji} X_j^r - \theta_j \right)}{\partial \left(\sum_{i=1}^n W_{ji} X_j^r - \theta_j \right)} \quad (11)$$

- (iv) The total increase is calculated for all patterns of thresholds and weights $\Delta W'_{kj}$ and $\Delta W'_{ji}$.
- (v) The weights and thresholds are updated.

⁷ The estimation of the two-layer artificial neural network model shows a better fit than the single-layer model when applied to data from the ICT survey. Only the best fitting results used in the model are reported here. Unused results are not reported.

- (vi) The total error is recalculated at $t+1$. If that error is not “satisfactory”, that is, if it is not the absolute minimum obtained through gradient descent, it is again interpreted as a process error signal and the third step is repeated until the minimum error is found.

In addition, an activation function, determined using the corresponding estimation algorithm, is calculated for each hidden layer and output layer. The activation function relates the weighted sum of units in a layer to the values of the units in the correct layer and differs for the hidden layers and the output layer. The following activation functions can be obtained in hidden layers:

- (i) Hyperbolic tangent, which takes arguments from real values and maps them to a range of -1 to 1 using the function:

$$\gamma(c) = \tanh(c) = \frac{e^c - e^{-c}}{e^c + e^{-c}} \quad (12)$$

- (ii) Sigmoid, which takes arguments from real values and maps them to a range of 0 to 1 using the function:

$$\gamma(c) = \frac{1}{(1 + e^{-c})} \quad (13)$$

For output variables, the activation function relates the weighted sum of units in a layer to the values of the units in the correct layer. The following activation functions can be obtained for the output layer:

- (i) Identity (or linear), which takes arguments from real values and returns them unchanged through the function:

$$\gamma(c) = c \quad (14)$$

- (ii) Softmax, which takes a vector of arguments from real values and transforms it into a vector that has elements between the values 0 and 1, with the sum equal to 1. This function cannot be used when there are scaled output variables; it is only available when the dependent (input) variables are categorical variables. It is formulated thus:

$$\gamma(c_k) = \exp(c_k) / \sum_j \exp(c_j) \quad (15)$$

The process is considered complete once the algorithm has minimized errors and the activation functions of the hidden and output layers have been found. The best way to compare the results obtained is to analyse a set or batch of data separately. As this set is not part of the training data, it is a good indicator of the network’s ability to estimate or recognize patterns.

IV. Analysis of the results

The first step in the factor analysis is to obtain the commonality coefficients of the analysed variables. If these coefficients are high (close to 1), this gives an approximation that a significant component of the variables can be explained by the existence of common factors. Otherwise (if coefficients are close to 0), it becomes evident that a large proportion of the variance can be explained by unique factors. The factors are obtained using the principal components method.

The commonality coefficients in the extraction column represent the proportion of the variance that can be explained by the factor solution model obtained. The commonalities explain most of the pooled variance of the variables (see table 1).

Table 1
Results of factor analysis

Component	Initial eigenvalues	Percentage of variance of each component	Rotation sums of squared loading	Commonality Variable	
	Total		Cumulative percentage	Commonalities (extraction)	
1	3.684	24.562	24.562	Landlines	0.526
2	2.074	13.824	38.386	Mobile phones	0.452
3	1.259	8.395	46.781	Social networks	0.625
4	1.002	6.681	53.462	Computers	0.803
5	0.913	6.090	59.552	Software	0.631
6	0.892	5.948	65.500	Internet	0.807
7	0.847	5.648	71.148	Cloud computing	0.321
8	0.824	5.496	76.644	Innovation	0.491
9	0.733	4.886	81.530	Salaries of postgraduate degree-holders	0.512
10	0.646	4.310	85.840	Salaries of bachelor's degree-holders	0.671
11	0.587	3.914	89.754	Salaries of technicians	0.612
12	0.542	3.612	93.366	ICT training	0.671
13	0.465	3.100	96.465	Systems department	0.461
14	0.371	2.470	98.936	Internet use index	0.651
15	0.160	1.064	100.000	Software use index	0.984

Source: Prepared by the authors.

Note: The extraction method applied is the principal components analysis method, using the Statistical Package for the Social Sciences (SPSS 23) software.

Table 1 shows the set of eigenvalues of the variance and covariance matrix, as well as the percentage of the total variance that each eigenvalue represents. The eigenvalues obtained express the share of variance that is explained by each of the factors. The extraction method is capable of yielding both factors and eigenvalues above 1 for the matrix analysed (four, in this case). Thus, the four components extracted explain 53.5% of the total variance. The next step is to calculate the component matrix, which shows the number of variables that saturate each factor. To obtain a number of “simple structure” variables (Thurstone, 1947), which are variables that saturate a single factor, as well as factors containing a limited number of explicitly and exclusively saturated variables, the component matrix obtained through the initial solution can be rotated.⁸ For the purposes of our analysis, the Varimax rotation method was used. The results of the rotated component matrix are shown in table 2.

These results indicate that the set of variables analysed can be grouped into three main factors: technology, capacity and innovation. Thus, six variables make up a factor termed “ICT infrastructure”, which explains 25% of the pooled variance, while a second factor comprising a set of human resource capacities explains 14% of the that variance. Here, salaries by level of education are seen as an approximation of capacities since, in theory, they are commensurate with work experience, education, training and skills, among other things. The use of new technologies (social networking and cloud computing), as well as innovation in products or processes, make up a third factor, termed “innovation”, which explains 8% of the pooled variance. Lastly, the software use variable explains 7% of the pooled variance.

⁸ Various methods of rotation exist, based on the premise that there is no universal solution for determining the matrix of weights. Multiplication by an orthogonal matrix of order $k \times k$ can be used such that the new model verifies the properties and weights of the original matrix but makes it more easily interpretable in terms of determining the variables that saturate factors.

Table 2
Factors obtained through factor analysis

Factor	Percentage of variance explained	Variables	Rotated component matrix ^a			
			1	2	3	4
1. ICT infrastructure	24.5	Landline availability	0.724	0.008	-0.034	0.017
		Mobile phone availability	0.377	0.076	0.308	0.095
		Computer availability	0.896	0.010	0.035	0.009
		Software availability	0.469	0.069	0.032	-0.077
		Internet availability	0.893	0.019	0.095	0.013
		Internet use index	0.638	0.134	0.475	0.018
2. Human resource capacity	13.8	Salaries of postgraduate degree-holders	-0.004	0.698	0.158	-0.016
		Salaries of bachelor's degree-holders	0.073	0.802	0.149	0.006
		Salaries of technicians	0.047	0.780	-0.041	-0.002
		Systems department	0.243	0.460	0.436	0.003
		ICT training	0.037	0.573	0.205	-0.004
3. Innovation	8.4	Social networking	0.062	-0.063	0.562	0.034
		Cloud computing	0.028	0.033	0.564	-0.017
		Innovation	0.060	0.170	0.675	-0.053
4. Software use	6.7	Software use index	-0.019	-0.009	-0.015	0.991

Source: Prepared by the authors.

Note: The extraction method applied is the principal components analysis method, Rotation method: Varimax with Kaiser normalization. SPSS 23 software was used.

^a Rotation converged in five iterations.

After identifying the factors into which the linear transformations of variables collapse, we can establish whether these are determinants of labour productivity. According to endogenous growth theory, besides factors related to workforce training, there are determinants of productivity related to capital intensity per worker and the capacity of organizations to absorb knowledge from competition (learning-by-doing). We use investment per employee as a proxy for capital stock per worker; we also use years of market experience and the exports-to-jobs ratio as operating variables in the competitive environment.

The next step of the analysis is to test, using artificial neural network methodology, whether the theoretical variables included are capable of explaining productivity behaviour, and whether that behaviour is associated with the capacities of organizations or, rather, the technology they use. We model a multilayer perception artificial neural network with two hidden layers. Thus, labour productivity is a function of the variables of the input layer (obtained through factor analysis): 1. ICT infrastructure; 2. Human resources capacity; 3. Innovation; 4. Software use; 5. Investment per employee; 6. Exports per employee; and 7. Firm experience, in years. The output layer denotes labour productivity.

1. Summary of processing

From a total of 6,210 observations, 1,622 were excluded from the analysis because they lacked one of the variables considered to be explanatory variables of productivity. The artificial neural network dataset was partitioned as follows: 78% of observations were used to train the network, 14% for validation and 8% used as the holdout set. The training process is performed in the hidden layer and is used to approximate the correct functional relationship. However, more data are sometimes required for neuron training, in which case the training dataset can pull data from the validation and holdout sets. The validation dataset is used to verify that the learning process has been performed appropriately with regard to error function minimization.

The partition of the sample allocated to the holdout set is excluded from the training and validation processes and is used to check how close productivity data estimated by the artificial neural network model are to real values to rule out the possibility of bias in the estimation. As these data were not used for training, they are a good indicator of the predictive strength of the model in terms of the proximity between estimated and real values. Table 3 shows the distribution of cases following network processing.

Table 3
Artificial neural network processing data

Summary of processing			
		N	Percentage
Sample	Training	3 556	78
	Validation	665	14
	Holdout	367	8
Valid		4 588	100
Excluded		1 622	
Total		6 210	

Source: Prepared by the authors.

Note: SPSS 23 software was used.

2. Input layer

The input layer comprises the independent variables which, theoretically, are capable of representing the behaviour of the dependent or output layer variable. The form that this representation takes is determined through a learning process within the network (hidden layers). The seven variables mentioned above were used in the artificial neural network selected for this study.

3. Hidden layers

It is in the hidden layers of an artificial neural network that the type of relationship that exists between the input variables and the dependent variable is estimated. This is done through a learning process known as “neuron training”. A significant proportion of the available data is needed for this training, which allows for the approximation of the correct functional relationship. Thus, 78% of the dataset was required to ensure that the learning process was carried out appropriately.⁹ Specifically, the architecture of the selected artificial neural network comprises two hidden layers (maximum number of hidden layers) with a hyperbolic tangent activation function, which means that the input values are converted to values within a range of -1 and 1, starting with function 12. There are 20 units in the first hidden layer and 15 units in the second (see table 4).

⁹ A learning process is deemed “adequate” when it is determined that the error minimization process (using the gradient descent method, explained in the methodology section) is capable of finding the minimum error function. When the minimum error is found, the learning process in the hidden layers stops.

Table 4
Artificial neural network processing data, hidden layers

Network information, hidden layers		
Hidden layers	Number of hidden layers	2
	Number of units in hidden layer 1	20
	Number of units in hidden layer 2	15
	Activation function	Hyperbolic tangent
Training characteristics		
	Initial learning rate	0.4
	Lower boundary of learning rate	0.001
	Learning rate reduction, in epochs	10
	Momentum	0.9
	Interval centre	0
	Interval offset	± 0.5

Source: Prepared by the authors.

Note: SPSS 23 software was used.

4. Output layer

The output layer denotes labour productivity and, like the input layer, requires an activation function which makes it possible to compare the results obtained against what are considered to be the explanatory factors. For the artificial neural network modelled here, the output layer is an indicator of two things:

- (i) Whether labour productivity is linked to input variables (ICT infrastructure and use, human resource capacities, innovation, software use, investment per employee, exports per employee and firm experience); this is determined through the error function.
- (ii) The significance and order of significance of each independent variable over the dependent variable.

This is shown in table 5.

Table 5
Artificial neural network processing data, output layer

Training characteristics		
Output layer	Independent variables 1	Labour productivity
	Number of units	1
	Method of scale change for scale-dependent variables	Normalized corrected
	Activation function	Hyperbolic tangent
	Error function	Sum of squares

Source: Prepared by the authors.

Note: SPSS 23 software was used.

The table also shows information regarding the synaptic connections of the input layer, the hidden layers and the output layer. Each connection can denote an excitatory or inhibitory weight: in other words, this determines the signal that the independent variable fires in the dependent variable, and the strength of that impact. The results of the estimation by the network model can be analysed in terms of the model's predictive capacity. In this regard, the quality of the network model is ultimately validated — using the holdout sample — by comparing the values estimated by the model against real values to determine their proximity.

As can be seen in table 6, the errors in each process (thousands of pesos) are reasonably low, especially in the holdout sample. Based on the holdout dataset, this indicates that, on average, there is a difference of just over 1,900 pesos (equivalent to a 1.2% error) between the estimated and observed productivity data. Figure 1 shows how close the labour productivity values estimated by the artificial neural network model are to real labour productivity values, based on the holdout sample.

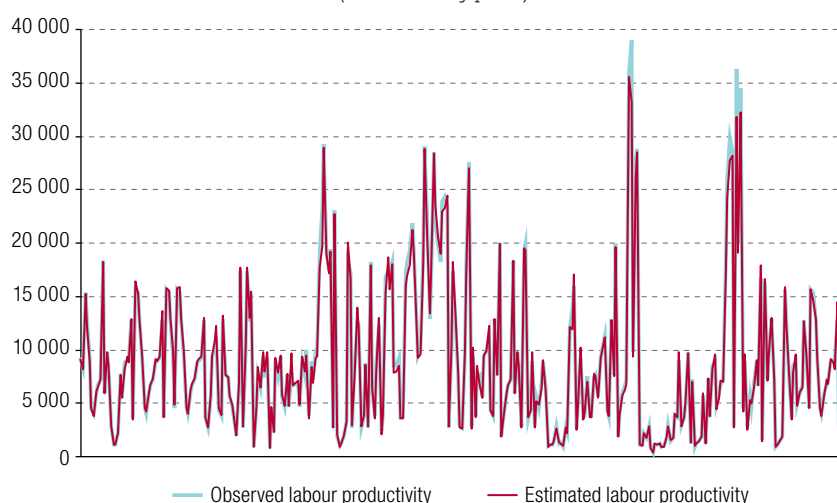
Table 6
Artificial neural network processing data (training, validation and holdout sets)

Model summary		
Training	Sum of squares error (<i>thousands of pesos</i>)	2.575
	Relative error (<i>thousands of pesos</i>)	1.006
	Stopping rule used	Maximum training time exceeded (15 minutes)
	Training time (<i>hours</i>)	27:48.7
Validation	Sum of squares error (<i>thousands of pesos</i>)	0.001
	Relative error (<i>thousands of pesos</i>)	1.862
Holdout	Relative error (<i>thousands of pesos</i>)	1.932

Source: Prepared by the authors.

Note: SPSS 23 software was used.

Figure 1
Observed labour productivity and labour productivity estimated by the artificial neural network model, holdout sample
(Thousands of pesos)



Source: Prepared by the authors, on the basis of data from the ICT survey (ENTIC).

Note: Estimation based on an artificial neural network with two hidden layers.

Figure 1 illustrates two interesting aspects: first, on average, for the entire holdout sample, the artificial neural network model is able to reasonably predict the behaviour of labour productivity; second, for companies with production per employee in excess of 10 million pesos, the model tends to underestimate the values. This could indicate that the factors that explain productivity in higher-income firms are more complex than the seven variables included as determinants of productivity in the model.

Unlike in econometric models, in artificial neural network models it is not possible to observe how coefficients are obtained within the hidden layers. However, they do provide information related to the significance of the impact of independent variables on the dependent variable. This is done through variable importance analysis, as seen in table 7.

Table 7
Variable importance analysis

Variable	Importance	Normalized importance (percentages)
Innovation	0.133	55.9
Software use	0.132	55.3
ICT availability and use	0.142	59.7
Human resources capacity	0.146	61.5
Investment per employee	0.128	53.7
Exports per employee	0.238	100.0
Experience	0.08	33.8

Source: Prepared by the authors.

Note: SPSS 23 software was used.

The importance of an independent variable is a measure of how much the network's model-predicted value changes for different values of the independent variable. Normalized importance is simply the importance values divided by the largest importance values and expressed as percentages (IBM, 2011).

The data in the table above show that for the simulated artificial neural network model, the key variable for understanding productivity is the volume of exports per employee. The educational qualifications of the workforce and the availability and use of ICT in organizations, respectively, rank second and third in order of importance as determinants of productivity. Innovation (be it in products or processes) is the fourth most important factor and also helps to explain productivity in Mexican companies with more than ten employees. However, it is the largest firms, with exports that account for more than 20% of total income, that are the most innovative. When this group of firms (which represents less than 4% of the total sample) is excluded from the analysis, exports and innovation fall to fourth and fifth position, respectively, while human resource capacities and the availability and use of ICTs move up to first and second position, respectively, in the order of importance, with investment per employee in third position.

V. Conclusions

The theoretical frame of reference of most of the research on the impact of ICT on productivity is neoclassical growth theory, in which technological change plays a central role in explaining productivity growth. However, as we have seen from the studies discussed in the present analysis, this theory is incapable of explaining why there are such discrepancies in productivity among companies with similar levels of ICT adoption.

The existence of differentiated initial conditions, in particular those involving the capacities of organizations, has been largely overlooked in the literature. Our research shows that including these differences in the analysis significantly improves the understanding of variation in productivity. This is particularly true in the case of Mexico, where the issue has not been examined extensively.

This paper identifies four important factors that explain productivity. The first is the availability of ICT in companies. Coupled with this is a set of capacities that determine how well technology is incorporated into production and business processes. We call this factor "organizational capacities", which include employees' education level, the training they receive and the existence of a specialized ICT department. The third factor involves the use and exploitation of technology, associated with the level of innovation (in products or processes), and the fourth is capital intensity per worker.

Organizations' exposure to competition and their market experience increase their capacity to absorb knowledge and raise productivity, as shown by the importance of exports and company experience as key production variables. The combination of these factors makes it possible to understand

differences in labour productivity. At firm level, the analysis by artificial neural networks shows that the combination of exports, human resources and ICT use seems to adequately explain a significant aspect of productivity behaviour in organizations with an income per worker of less than 10 million pesos. In firms with higher labour productivity, other factors appear to explain productivity behaviour.

This study has shown that it is not only technology, but also factors such as education and organizational maturity that make production processes more dynamic and lay the foundations for high productivity. This points to a need for complementarity in public policy in three areas:

- (i) Encouraging appropriate development of ICT infrastructure, particularly in economic sectors or activities considered priority, either because they occupy a strategic position in the value chain or because of their high impact in terms of production linkages and employment generation. This policy should be geared towards developing sufficient and suitable ICT infrastructure (particularly the Internet and software, which are indispensable for the operation of other types of services, such as cloud services) for key sectors.
- (ii) As this study has established, ensuring the availability of infrastructure is the first essential step towards improving labour productivity. However, this alone does not suffice. Government policies to enhance infrastructure must be accompanied by policies on capacity-building in the use of technological tools. One way of promoting and transmitting the importance of training to facilitate the effective use of technology could be the establishment of public-private partnerships. Such partnerships could, for example, allow for training costs to be covered in part by public funds with the remainder borne by the company, or for some form of tax exemption for firms that regularly provide technological training for their employees.
- (iii) In addition to developing specific policies on ICT training for current employees in companies, there is a need for a long-term policy on capacity-building. For this to be achieved, the best strategy appears to be that proposed in early studies on endogenous growth theory, namely, greater investment in education as a prerequisite for long-term economic growth. It would appear that this is already part of the debate in Mexican public policy, but it is far from being reflected in a formal policy aimed at increasing investment in education.

Bibliography

- Aghion, P. and others (1998), *Endogenous Growth Theory*, Cambridge, MIT Press.
- Aguilera, L., H. Cuevas-Vargas and M. González (2015), "The impact of information and communication technologies on the competitiveness: evidence of manufacturing SMEs in Aguascalientes, Mexico", *International Review of Management and Business Research*, vol. 4, No. 3, Teheran, Islamic Azad University.
- Aravena, C., C. Cavada and N. Mulder (2012), "Contribución al crecimiento económico de las tecnologías de la información y las comunicaciones y de la productividad en la Argentina, el Brasil, Chile y México", *Statistical Studies series*, No. 76 (LC/L.3439), Santiago, Economic Commission for Latin America and the Caribbean (ECLAC).
- Aramendia-Muneta, M. and A. Olló-López (2013), "ICT Impact on tourism industry", *International Journal of Management Cases*, vol. 15, No. 2, Bosnia and Herzegovina, CIRCLE Research Centre.
- Arrieta, J., J. Torres and H. Velásquez (2009), "Predicciones de modelos econométricos y redes neuronales: el caso de la acción de SURAMINV", *Semestre Económico*, vol.12, No. 25, Medellín, University of Medellín.
- Balboni, M., S. Rovira and S. Vergara (eds.) (2011), *ICT in Latin America: A Microdata Analysis* (LC/R.2172), Santiago, Economic Commission for Latin America and the Caribbean (ECLAC).
- Basu, S. and J. G. Fernald (2006), "Information and communications technology as a general purpose technology: evidence from U.S. industry data", *Working Paper*, No. 2006-29, San Francisco, Federal Reserve Bank of San Francisco.
- Baumol, W., J. Panzar and R. Willing (1982), *Contestable Market and the Theory of Industry Structure*, New York, Harcourt Brace Jovanovich.

- Bayo-Moriones, A. and F. Lera-López (2007), "A firm-level analysis of determinants of ICT adoption in Spain", *Technovation*, vol. 27, No. 6-7, Amsterdam, Elsevier.
- Bayo-Moriones, A., M. Billón and F. Lera-López (2013), "Perceived performance effects of ICT in manufacturing SMEs", *Industrial Management & Data Systems*, vol. 113, No. 1, Bingley, Emerald Publishing.
- Bilbao-Osorio, B., S. Dutta and B. Larvin (eds.) (2013), *The Global Information Technology Report 2013: Growth and Jobs in a Hyperconnected World*, Geneva, The Business School of the World (INSEAD)/World Economic Forum.
- Bishop, C. M. (1995), *Neural Networks for Pattern Recognition*, Oxford, Oxford University Press.
- Bresnahan, T. (1999), "Computerization and wage dispersion: an analytical reinterpretation", *The Economic Journal*, vol. 109, No. 456, Oxford, Royal Economic Society.
- _____(1997), "Testing and measurement in competition models", *Advances in Economics and Econometrics: Theory and Applications. Seventh World Congress*, vol. 3, D. Kreps y K. Wallis (eds.), Cambridge, Cambridge University Press.
- Brynjolfsson, E. (1996), "The contribution of information technology to consumer welfare", *Information Systems Research*, vol. 7, No. 3, Catonsville, Institute for Operations Research and the Management Sciences.
- _____(1993), "The productivity paradox of information technology: review and assessment", *Communications of ACM*, vol. 36, No. 12, New York, Association for Computing Machinery.
- Brynjolfsson, E. and L. M. Hitt (1996), "Paradox lost? Firm-level evidence on the returns to information systems spending", *Management Science*, vol. 42, No. 4, Catonsville, Institute for Operations Research and the Management Sciences.
- Capel, L. and J. Bosch (2004), "El districte industrial de les Tecnologies de la Informació i la Comunicació (TIC) a Barcelona. Comparació amb altres ciutats europees", *Scripta Nova*, vol. 8, No. 170, Barcelona, University of Barcelona.
- Carrera Portugal, A. (2010), "Las TIC como parámetro de competitividad urbana: un escenario para las economías emergentes", *Biblioteca Universitaria*, vol. 13, No. 2, Mexico City, National Autonomous University of Mexico (UNAM).
- Cuadras, C. M. (2007), *Nuevos métodos de análisis multivariante*, Barcelona, CMC Editions.
- Daveri, F. (2000), "Is growth an information technology story in Europe too?", *IGIER Working Paper*, No. 168, Milan, Bocconi University.
- Dedrick, J., V. Gurbaxani and K. Kraemer (2003), "Information technology and economic performance: a critical review of the empirical evidence", *ACM Computing Surveys*, vol. 35, No. 1, New York, Association for Computing Machinery.
- Domínguez, L. y F. Brown (2004), "Measuring technological capabilities in Mexican industry", *CEPAL Review*, No. 83 (LC/G.2231-P), Santiago, Economic Commission for Latin America and the Caribbean (ECLAC).
- Dutta, S. and B. Bilbao-Osorio (eds.) (2012), *The Global Information Technology Report 2012: Living in a Hyperconnected World*, Geneva, The Business School of the World (INSEAD)/World Economic Forum.
- Gatautis, R. (2008), "The impact of ICT on public and private sectors in Lithuania", *Engineering Economics*, vol. 59, No. 4, Kaunas, Kaunas University of Technology.
- Greenan, N., J. Mairesse and A. Topiol-Bensaid (2001), "Information technology and research and development impacts on productivity and skills: looking for correlations on French firm level data", *NBER Working Paper*, No. 8075, Cambridge, National Bureau of Economic Research (NBER).
- Gurbaxani, V. and S. Whang (1991), "The impact of information systems on organizations and markets", *Communications of the ACM*, vol. 34, No. 1, New York, Association for Computing Machinery.
- Haykin, S. (1999), *Neural Networks. A Comprehensive Foundation*, Singapore, Prentice-Hall.
- Hollenstein, H. (2004), "Determinants of the adoption of Information and Communication Technologies (ICT): an empirical analysis based on firm-level data for the Swiss business sector", *Structural Change and Economic Dynamics*, vol. 15, No. 3, Amsterdam, Elsevier.
- IBM (2011), *IBM SPSS Neural Networks 20* [online] https://www.spss.ch/upload/1316002997_SPSS%20Neural%20Networks%2020.pdf.
- INEGI (National Institute of Statistics and Geography) (2013), *Sistema de Cuentas Nacionales de México: productividad total de los factores 1990-2011*, Mexico City.
- Inklaar, R., M. O'Mahony and M. Timmer (2005), "ICT and Europe's productivity performance: industry-level growth account comparisons with the United States", *The Review of Income and Wealth*, vol. 51, No. 4, Hoboken, Wiley.
- ITU/UNESCO (International Telecommunication Union/United Nations Educational, Scientific and Cultural Organization) (2013), *The State of Broadband 2013: Universalizing Broadband*, Paris.

- Jorgenson, D. (2001), "Information technology and the U.S. economy", *The American Economic Review*, vol. 91, No. 1, Nashville, American Economic Association.
- Jorgenson, D. and K. Stiroh (2000), "Raising the speed limit: U.S. economic growth in the information age", *Brookings Papers on Economic Activity*, vol. 31, No. 1, Washington, D.C., The Brookings Institution.
- Jorgenson, D., M. Ho and K. Stiroh (2006), "Productivity growth in the new millennium and its industry origins", paper presented at the 2006 Intermediate Input-output Meeting, Sendai, 26-28 July.
- Katz, R. (2009), *El papel de las TIC en el desarrollo. Propuesta de América Latina a los retos económicos actuales*, Barcelona, Ariel/Fundación Telefónica.
- Kraemer, K. and S. Dewan (2000), "Information technology and productivity: evidence from country-level data", *Management Science*, vol. 46, No. 4, Catonsville, Institute for Operations Research and the Management Sciences.
- Lal, K. (2001), "The determinants of the adoption of information technology: a case study of the Indian garments industry", *Information Technology, Productivity, and Economic Growth: International Evidence and Implications for Economic Development*, M. Pohjola (ed.), Oxford, Oxford University Press.
- Larrañaga, P., I. Inza and A. Moujahid (2003), "Modelos probabilísticos para la inteligencia artificial y la minería de datos: selección de variables" [online] <http://leo.ugr.es/doctopgm/cartel.pdf>.
- Lichtenberg, F. R. (1995), "The output contributions of computer equipment and personnel: a firm level analysis", *Economics of Innovation and New Technologies*, vol. 3, No. 3-4, Abingdon, Taylor & Francis.
- MacKay, D. J. (2003), *Information Theory, Inference, and Learning Algorithms*, Cambridge, Cambridge University Press.
- Malone, T., J. Yates and R. Benjamin (1989), "The logic of electronic markets", *Harvard Business Review*, vol. 67, No. 3, Boulder, Harvard University Graduate School of Business Administration.
- OECD (Organization for Economic Cooperation and Development) (2012), *Estudio de la OCDE sobre políticas y regulación de telecomunicaciones en México*, Paris, OECD Publishing.
- Peppard, J. and J. Ward (2016), *The Strategic Management of Information Systems: Building a Digital Strategy*, Hoboken, Wiley.
- Pitarque, A., J. F. Roy and J. C. Ruiz (1998), "Redes neuronales vs modelos estadísticos: simulaciones sobre tareas de predicción y clasificación", *Psicológica*, vol. 19, Warsaw, De Gruyter.
- Pohjola, M. (2001), "Information technology and economic growth: a cross-country analysis", *Information Technology, Productivity, and Economic Growth: International Evidence and Implications for Economic Development*, Oxford, Oxford University Press.
- Rivera, E. (2004), "Teorías de la regulación en la perspectiva de las políticas públicas", *Gestión y Política Pública*, vol. 13, No. 2, Mexico City, Center for Research and Teaching in Economics (CIDE).
- Romer, P. M. (1994), "The origins of endogenous growth", *The Journal of Economic Perspectives*, vol. 8, No. 1, Nashville, American Economic Association.
- _____(1990), "Endogenous technological change", *The Journal of Political Economy*, vol. 98, No. 5, Chicago, The University of Chicago Press.
- Schreyer, P. (2000), "The contribution of information and communication technology to output growth: a study of the G7 countries", *OECD Working Paper*, No. 2000/02, Paris, OECD Publishing.
- Silva, E. and A. Teixeira (2011), "Does structure influence growth? A panel data econometric assessment of 'relatively less developed' countries, 1979-2003", *Industrial and Corporate Change*, vol. 20, No. 2, Oxford, Oxford University Press.
- Spiezia, V. (2011), "Are ICT users more innovative? An analysis of ICT-enabled innovation in OECD firms", *OECD Journal: Economic Studies*, vol. 2011, No. 1 [online] <http://dx.doi.org/10.1787/19952856>.
- Solow, R. M. (1957), "Technical change and the aggregate production function", *The Review of Economics and Statistics*, vol. 39, No. 3, Cambridge, MIT Press.
- Taruté, A. and R. Gatautis (2014), "ICT impact on SMEs performance", *Procedia-Social and Behavioral Sciences*, vol. 110, Amsterdam, Elsevier.
- Thurstone, L. L. (1947), *Multiple Factor Analysis*, Chicago, The University of Chicago Press.
- Tödting, F., M. Grillitsch and C. Höglinger (2012), "Knowledge sourcing and innovation in Austrian ICT Companies—How does geography matters", *Industry and Innovation*, vol. 19, No. 4, Abingdon, Taylor & Francis.
- UNCTAD (United Nations Conference on Trade and Development) (2003), *Information and Communication Technology Development Indices*, Geneva.

- Wimble, M. and H. Singh (2015), "A multilevel examination of information technology and firm performance: the interaction of industry and firm effects", *PACIS 2015 Proceedings*, Pacific Asia Conference on Information Systems [online] <https://aisel.aisnet.org/pacis2015/129>.
- World Bank (2012), *2012 Information and Communications for Development: Maximizing Mobile*, Washington, D.C.
- (2011), *Little Data Book on Information and Communication Technology*, Washington, D.C. [online] <http://documents.worldbank.org/curated/en/338701468335685584/pdf/635700REPLACEM0B0ISBN09780821388600.pdf>.