The European Journal of Development Research
https://doi.org/10.1057/s41287-019-0197-0

ORIGINAL ARTICLE

ICT Use, Investments in R&D and Workers' Training, Firms' Productivity and Markups: The Case of Ecuadorian Manufacturing

Jorge Antonio Rodríguez-Moreno[1,2] · **María Engracia Rochina-Barrachina**[3]

Abstract
We use data from the last Ecuador Economic Census, covering the universe of manufacturing firms, to study the relationship between firms' R&D and workers' training investments and ICT use and firms' productivity and markups. These knowledge-related investments may affect productivity. Moreover, investments in both knowledge and productivity can affect the ability of firms to set prices above marginal costs. Whether R&D and workers' training investments and ICT are important for productivity and the capacity to set higher markups in developing countries are interesting development policy questions. We find that good business practices, including access to internal capital markets or to external finance, encourage R&D and workers' training investments, and ICT use. These investments affect positively firms' productivity and markups. Their influence on markups operates in general through efficiency and prices. Finally, there is evidence about demand conditions to boost knowledge-related investments and markups.

Keywords R&D · ICT · Training · Productivity · Markups · Economic development · Ecuadorian manufacturing

✉ Jorge Antonio Rodríguez-Moreno
jarodrig@usm.edu.ec; joranro@alumni.uv.es

María Engracia Rochina-Barrachina
erochina@uv.es

1 Department of Economics, Universidad Santa María, Campus Guayaquil, Av. Carlos Julio Arosemena, Km 4.5, Guayaquil, Ecuador

2 Department of Applied Economics II, Faculty of Economics, University of Valencia, Av. Tarongers, s/n, 46022 Valencia, Spain

3 Department of Applied Economics II, Faculty of Economics and ERICES, University of Valencia, Av. Tarongers, s/n, 46022 Valencia, Spain

Published online: 13 February 2019

Introduction

To understand the relevance of the topics in this paper for development and growth, it is quite illustrative to follow Baldwin's view about world economic globalization stages (Baldwin 2016). The old stage of globalization, what he calls the Great Divergence, was characterized by the combination of low trade costs and high communication costs. Under these conditions, world globalization was embodied in trade of goods and manufacturing was concentrated in the 'North'. Policies that made sense at that time were sector-level policies. In this scenario, many developing countries were focused on protecting their infant industries through tariffs. Next came a new globalization stage named by Baldwin, the Great Convergence. This period, which we are probably still in, is characterized by both low trade costs and low communication costs. The arrival of information and communication technologies (ICT) was crucial for the movement of ideas in the world and, specially, for the movement of productive know-how from the 'North' to the 'South'. This was behind the proliferation of multinationals and of the global value chains (GVC) revolution. However, initially, this transfer of knowledge only affected a few developing countries. Only later did it spread to more countries through the import demand of materials. In this new type of globalization, policies that go beyond sectors and that focus on stages of production, specific tasks and occupations become more relevant. This new way of acquiring comparative advantages in specific parts of the value chain opens the door to the strategic positioning of developing countries. They face a trade-off between protectionism and liberalization to be integrated in GVCs. As Baldwin highlights, in this new world, developing economies do not only compete with developed ones but also with developing countries. The world is not yet flat and, roughly speaking, high-tech/high-wage countries coexist with low-tech/low-wage and high-tech/(still) low-wage countries.

If developing countries want to be integrated in this knowledge economy world and profit from steps in the value chain of higher value added, or even being themselves the headquarters (HQ) of the GVC, they should think about knowledge creation as a multidimensional investment that includes, among others, investments in R&D to facilitate innovation, knowledge transfer and absorptive capacity, investments in workers' training for the acquisition of relevant skills, and investments in ICT for the improvement of relevant knowledge and participation in global networks. These three investments even appear as more relevant when we consider the challenge for developing countries in closing the digital divide (with developed countries) and facing the Industry 4.0 era. Things such as digital devices, artificial intelligence, robotics, the Internet of Things, Big Data Analytics, machine learning, 3-D printing, etc., characterize this fourth industrial revolution (or Industry 4.0).

In this paper, we contribute to the general theme of how relevant knowledge is for development and growth, and we investigate a developing country, Ecuador, that will serve to provide empirical evidence on how knowledge creation activities at the firm level (and in the manufacturing sector) contribute to the firm's performance both in terms of productivity and markups. It is evident that higher firm productivity should be good for society. However, the contribution to society of higher markups could be a controversial issue. For this reason, in the "Reflection on Results for Markups" section, we include a reflection on this that rests on an important distinction: an increase in markups generated by the existence of barriers to competition is not the same as one promoted by knowledge investments that both increase efficiency and quality (or diversification) of goods, which likely allow greater appropriation of value added from GVCs and better access to more sophisticated export markets.

The Ecuadorian economy is still heavily dependent on the international prices of oil and agricultural products. Given this situation, the Ecuadorian government is committed to facilitating a transition from an economy based on the primary sector to a more knowledge-oriented one. A developing country should not only aspire to participate in GVCs through its participation as a supplier of materials.

Our analysis is based on large-scale firm-level data that covers the entire population of manufacturing firms. The source is the Economic Census of Ecuador (INEC 2010). Since this census is meant to have a 10-year periodicity, this last census is the one we use in this study. We will exploit information about firms' R&D and workers' training investments and use of ICT, as well as about general firm characteristics. The scarce utilization of this database for an empirical analysis which is not merely descriptive makes this work novel and pioneering for Ecuador. According to the information in the census, 88.32% of manufacturing firms are not involved in any of the three considered drivers of knowledge. This highlights that Ecuadorian manufacturing firms have not yet obtained all the benefits derived from the widespread performance of these activities.

From the empirical perspective, our particular objectives are manifold. First, we are interested in explaining the joint likelihood of firms carrying out R&D, workers' training and use of ICT. The joint likelihood of these activities will require the estimation of a trivariate probit model taking into account the potential interrelationships between them. Second, we are also interested in explaining firms' determinants of R&D and workers' training investment intensities. Unfortunately, the database does not have information about ICT expenditures. Selection issues are taken into account by estimating bivariate Heckman sample selection models. Third, we introduce estimates from previous stages in a Crépon–Duguet–Mairesse (CDM 1998) framework to study the linkages between knowledge-related investments and firms' productivity. We do this by employing alternative measures of productivity such as labor productivity and total factor productivity (TFP) estimates from *Cobb–Douglas* and *Translog* production functions. Finally, we check whether knowledge-related investments not only affect firms' productivity but also have an influence on the firms' capacity to set prices above marginal costs and, hence, markups. Estimated firms' markups follow from the production function estimation. The implementation of a CDM approach allows the incorporation of control function corrections (see Rivers and Vuong 1988; Wooldridge 2010) for testing and handling the possibility of endogeneity of drivers of knowledge in the productivity and markup equations.

The novelties in this paper are as follows. First, we use a broad definition of firms' knowledge investment that includes investments in R&D and workers' training as well as ICT usage. This is expected to contribute to the minimization of omitted variables bias when trying to understand the consequences for firms' performance of adoption and intensity of such investments. Second, we use different functional forms to analyze the relationship between knowledge-related investments and productivity. Third, we go one step further in the CDM framework by incorporating a second firm's performance measure, besides the typical one of productivity, that is, firms' markups. The extension to markups is one of the main contributions of the paper. Therefore, we are not only going to answer the question about why firms invest in different types of knowledge and how much they invest and, later, their effects on firms' TFP but also break new ground in investigating the role of knowledge-related investments and TFP on firms' markup formation. Fourth, in this final stage of our estimation procedure in the paper, we can distinguish, by conditioning to TFP in the markup regression, whether the effect of the knowledge investment variables on markups operates through efficiency, that is marginal costs proxy by TFP, and/or through the higher capacity of firms to set prices above marginal costs, as more knowledge-oriented firms are more likely to produce higher quality products. Finally, literature integrating all these elements in a unified framework is scarce. Hence, to find out whether these types of activities have a relevant role for developing countries is of considerable interest, not only for managers but also for policy makers, since whether this type of investments are important sources of productivity and capacity to fix higher markups in developing countries are interesting development policy questions. Furthermore, this is the first study of this type for Ecuador.

The main results in the paper can be summarized as follows. First, the professionalization and good business practice variables such as belonging to a business group, having access to finance, performing activities of market research, accountancy, and having environmental concerns, explain higher propensities and intensities of R&D and workers' training investments, and ICT use. Second, the three knowledge-related investments positively affect firms' TFP and markups. Third, except for ICT, part of their effect on markups operates through influencing prices and not only efficiency. Fourth, we detect some demand-driven knowledge-related investments and markups. Fifth, we also detect some evidence about learning and product quality requirements from international markets encouraging this type of investments. Finally, we obtain results that may be indicative of financial constraints affecting such investments, softened for firms belonging to a business group or with access to external finance.

This clearly demonstrates the important role for public policy in encouraging the spread of these activities among firms in order to obtain sound effects on firms' performance measures, such as productivity and markups. There is also room for government intervention in alleviating Ecuadorian manufacturing firms' financial constraints affecting these investments.

The paper is organised in the following sections. The "Literature Review" section provides an overview of related literature. The "Data" section introduces the dataset. The "Estimation Methodology and Results" section is devoted to methodological concerns and procedures at each stage of estimation and presents obtained results. The "Concluding Remarks" section concludes and introduces further discussion on policy issues.

Literature Review

One can think of two main channels by which the performance of R&D activities can translate into productivity improvements: first, firms investing in R&D may both increase its productive efficiency and obtain better products, increasing demand and reducing production costs. Second, firms investing in R&D likely face more favorable growth perspectives, which contribute to a better exploitation of economies of scale in production, with the associated costs reduction. However, in empirical studies, a positive and statistically significant relationship between firms' productivity and R&D performance is not always found. In particular, the survey by Hall (2011) highlights that results can be different depending on R&D investments being relatively more oriented towards product or process innovation (her paper does not consider R&D but the innovation outputs product and process innovation). The results of the survey are mostly negative for process innovation and positive for product innovation (suggesting that productivity is enhanced by new and better-quality products, and not always by process innovation). However, as Hall (2011) explains, the negative result for process innovation is primarily due to the fact that since firms' individual prices are commonly absent in most of the datasets, deflation of revenue by industry deflators yields real revenue rather than a totally physical output measure. Hence, if the typical firm operates in an inelastic part of its demand curve, so that real revenue productivity falls when it becomes more efficient (as firm's prices decrease while sales remain constant), it can be explained theoretically why some empirical studies find a negative relationship between process innovation and productivity. This would suggest they are not able to measure the real quantity effect of process innovation but the real revenue effect.

In the line of research focused on the relationship between productivity, innovation and R&D investment, we find the seminal work by Crépon et al. (1998) that, using cross-sectional data for French manufacturing firms, develops the so-called CDM model, which explains the above-mentioned relationships through sequential estimation steps. The idea behind the CDM is that R&D expenses generate knowledge for firms, and this knowledge can be measured by innovation outcomes (inter alia, patents, new processes, new products) that can generate a subsequent positive effect in the firms' productivity. Following this approach, Crespi and Zuniga (2012) use the CDM model and find in six Latin American countries (Argentina, Chile, Colombia, Costa Rica, Panama, and Uruguay) that firms' improvements in new technologies enhance their productivity. However, Benavente (2006) finds no impact of innovation (R&D) on productivity for Chilean firms.

However, there is not a unique path to apply the CDM model approach. It depends on the researchers' available information in a particular database and the nature of this information, as well as on researchers' interests. Thus, there are papers applying the CDM sequence going directly from innovation inputs to productivity (in the case of an absence of innovation outputs in the databases).

Since the available information in the Economic Census of Ecuador includes the entire population of manufacturing firms but has no information on innovation outcomes (patents, new processes, or new products), our application in this paper of the CDM approach relies on the direct use of inputs that likely contribute to knowledge creation. An example of these are R&D expenses. In this way, we treat R&D, workers' training and ICT as inputs in a knowledge production function that can ultimately generate innovation outputs but also contribute to the firm's capacity to absorb, assimilate and manage technical change (Cohen and Levinthal 1990). In developing countries, it is not only relevant to increase the capacity to innovate but also the capacity to assimilate newly acquired technology from abroad and new information. Building technological capabilities depends on investments such as R&D and on-the-job training, which may be considered investments in knowledge or in knowledge-producing (or -acquiring) activities (Aw et al. 2007). Corrado et al. (2005) include software, R&D and firms' on-the-job training as knowledge investments.

There is no doubt about human capital being an important driver of economic growth. Hence, the expected role for workers' training on firms' productivity has its roots on the maintenance and improvement of the human capital of workers. Better human capital may act on productivity through several channels: the firm can take better decisions, learning-by-doing is expected to be higher when workers have superior human capital, and innovation or adaptation to new technologies can be stimulated by the quality and training of workers. For Belgian firms, Konings and Vanormelingen (2015, p. 485) find, for instance, "that an increase in the share of trained workers by 10 percentage points is associated with 1.7% to 3.2% higher productivity." For German establishments, Zwick (2006) shows that increasing training intensity has a positive and significant effect on productivity. Colombo and Stanca (2014) obtain a similar result for Italian firms.

The arguments for ICT contributing to productivity are diverse. First, it enables firms to be better connected, to better manage information and to have access to external knowledge. Second, it may affect productivity through a more efficient production organization or the supply of new and/or better products and services. In this respect, Añón-Higón (2012) for UK SMEs finds that "ICT operate primarily as efficiency-enhancing technologies, although specific market-oriented applications (that is, website development) exhibit the potential to create competitive advantage through product innovation." In this sense, ICT should not only be understood as a general-purpose technology (Bresnahan and Trajtenberg 1995). Since the question about ICT use in the Economic Census of Ecuador is posed in a broad and general way ("Does the firm use the internet in its different processes?"), it is reasonable to think that with this type of variable we are jointly capturing both aspects in the use of ICT.

Among the empirical studies that find evidence of positive effects of ICT on productivity with firm-level data, we have van Leeuwen and van der Wiel (2003) for The Netherlands, Brynjolfsson and Hitt (2003) and Rincón et al. (2013) for the US, and Castiglione (2012) for Italy. Furthermore, from those who also consider R&D, Bloom et al. (2010) find strong effects of R&D on productivity but little evidence for ICT. However, studies of the impacts of ICT on productivity are scarce in developing countries. Two recent exceptions are Commander et al. (2011) for Brazil and India and Aboal and Tacsir (2018) for Uruguay.

It is relevant to control simultaneously for the three knowledge-related investments (R&D, workers' training and ICT), both in the productivity and markups equations, in order to avoid potential omitted variable bias. However, studies at the firm level jointly controlling for the three knowledge investments are scarce. Instead, we mainly find works focused on one or two of them. See, for instance, the works by Greenan et al. (2001), Polder et al. (2010) and Hall et al. (2013), for French, Dutch and Italian firms, respectively, which look at R&D and ICT; or the works by Black and Lynch (2001) and Bresnahan et al. (2002) for US firms that consider ICT and human capital. Some exceptions are Arvanitis and Loukis (2009) for Greece and Switzerland and Aboal and Tacsir (2018) for Uruguay, which consider investment in R&D, investment in training (or a measure of the firm's human capital), and investment in ICT (or ICT use). Unfortunately, the Economic Census of Ecuador does not have any variable capturing the stock of firm-specific human capital and, hence, in estimation we can only control for between-industry variation of this stock through the inclusion of industry fixed effects but not for within-industry variation. In any case, it seems reasonable to expect that firms that in a given industry invest in workers' training versus those that do not contribute to skills upgrading of their workers. In addition, as discussed by Crespi and Zuniga (2012) and Janz et al. (2004), the introduction of a measure of human capital that includes R&D personnel may overlap in some way with R&D expenditure.

Finally, none of the above-mentioned studies considers the effects of these knowledge-related activities on firms' markups.

Data

In this paper, we use the Economic Census of Ecuador 2010. This is the most recent Ecuadorian census since it is intended to have a 10-year periodicity. However, the previous one goes back to 1980. It includes firms' characteristics like, for instance, age, location, legal status, industrial sector, employment, sales and main clients, costs, revenues and fixed assets that, among others, will be used in the empirical analysis.

For this study, as we are focused on the manufacturing sector, the number of firms of this type in the census is 44,109. The census covers the entire population of manufacturing firms. After cleansing the data of firms with missing information for the relevant variables in this study, we end up with a sample of 42,292 firms, which certainly guarantees that our analysis is based on large-scale firm-level data from the Ecuadorian manufacturing sector.

Since the survey includes explicit and particular questions about R&D, workers' training and ICT adoption, we find that 412 firms (0.97%) perform R&D, 1828 (4.32%) invest in workers' training and 4173 (9.87%) use ICT. This clearly shows that these types of activities are not widespread among manufacturing firms in Ecuador (88.32% of firms are not involved in any of them). Moreover, concentrating on the group of firms that performs at least one of these activities, the highest percentage corresponds to firms performing only one of them (74.58%) and the lowest to firms performing all three (4.46%), while firms performing two activities represent 20.96%. If we consider non-excluding categories, the most common activity among performers is ICT (84.50%), followed by workers' training (37.01%) and R&D (8.34%).

The variables used in this paper are defined in "Appendix 1". In addition, their descriptive statistics can be found in "Appendix 2". In this appendix, we not only present information for the whole sample but also separate information for two subsamples: the subsample of performers of at least one of the considered activities (R&D, workers' training and ICT use) and the subsample of non-performers of any of them.

Simply from the comparison of means, we can, from a descriptive point of view, characterize performers versus non-performers. First, looking at the regional distribution, performers are relatively more located in Pichincha than non-performers (37.72% and 22.05%, respectively).[1] Second, as regards composition by industry, performers are relatively more concentrated than non-performers in sectors such as Wood, Paper and Printing, Chemicals and Petroleum products, Rubber and Plastics, Office Machinery and Electrical Equipment, Communication–Precision–Optical and Medical Equipment, and Transport Equipment.[2] Third, as regards legal forms, they are more concentrated than non-performers in the category of Private Company (25.57% vs. 0.74%, as most, 98.93%, of the non-performers are categorized as Natural Persons).[3] Non-performers have no firm in the category of Foreign Control Company, but the percentage is also quite low among performers (0.12%). Fourth, for variables that we consider in this paper as proxy for good business practices and professionalization, in all of them the composition of performers outperforms non-performers (59.07% vs. 13.86% firms belong to an Enterprise Network, 11.11% vs. 1.23% firms perform market research, 42.77% vs. 3.34% firms perform accounting, 36.65% vs. 23.11% firms have access to finance, and 10.91% vs. 0.48% firms declare environmental concerns). Fifth, although there are not many firms for which main customer is foreign, the percentage of them among performers is 3.30% and among non-performers 0.19%. Sixth, performers are on average both larger (about 3.37 times) and older (about 1.68 times) than non-performers. Finally, performers are more capital-intensive (about 3.36 times) and more materials-intensive (about 2.47 times) than non-performers, and have higher labor productivity (2.67 times).

[1] There are 24 provinces in the country (see "Appendix 1 or 2"), three of which (Pichincha, Guayas and Azuay) account for 53.01% of manufacturing firms.

[2] Industries are food, beverages and tobacco; textiles and wearing apparel; leather and footwear; wood, paper and printing; chemicals and petroleum products; rubber and plastics; non-metallic mineral products; metal products; office machinery and electrical equipment; communication, precision, optical and medical equipment; transport equipment; and furniture and n.e.c. This classification is based on the International Standard Industrial Classification at the two-digit level for manufacturing.

[3] Legal forms are Natural Persons, Non-profit Company, Private Company, Foreign Company, Public Company, Local Government Company, Cooperative, and Association. These correspond to one of the questions in the survey.

Therefore, in general, the sub-sample of performers could be defined as the most 'advanced' part of Ecuadorian manufacturing. Furthermore, the sub-sample of performers concentrates 58.97% of employment and 93.93% of turnover in the manufacturing industry.

Estimation Methodology and Results

The Firms' Decisions on Knowledge Creation Activities: R&D, Workers' Training and ICT Use

We estimate a standard trivariate probit model for the three discrete choices involved in the first stage of our analysis. Let $y^{*}_{1i,j}$ denote a latent variable underlying firm i's ($i = 1, \ldots, N$) propensity to invest in activity j (j = R&D, WT or ICT) given firm and structural characteristics x_{1i}. Formally:

$$y^{*}_{1i,j} = \beta'_{1,j} x_{1i} + \varepsilon_{1i,j}, \qquad (1)$$

where $\beta'_{1,j}$ captures the effects of explanatory variables on the propensity to perform knowledge-related activity j and $\varepsilon_{1i,j}$ denotes idiosyncratic errors that affect $y^{*}_{1i,j}$. The observed dependent variables, $y_{1i,j}$, corresponding to $y^{*}_{1i,j}$ are defined as:

$$y_{1i,j} = \mathbf{1}\left[y^{*}_{1i,j} > 0 \right], \qquad (2)$$

where 1[] denotes the indicator function taking the value one if the condition between squared brackets is satisfied, and zero otherwise. We assume that the three error terms $\varepsilon_{1i,j}$ involved follow a trivariate normal distribution. This specification allows correlations between the three choices to be non-zero. If these correlations are not considered, we would not only lose efficiency but also biased and inconsistent parameter estimates would be possible due to the relationships between the different types of knowledge-related activities (Greene 2003). Estimation of the three-equation system by pseudo-simulated maximum likelihood is performed with the Stata command *cmp* developed by Roodman (2011).

The explanatory variables considered are the same for the three choices, although their impact may differ. First, we include the group of variables intended to capture the degree of professionalism and modernization of the firm in terms of quality and diligence in management and business practices. Here are included the variables about whether the firm belongs to an enterprise network or business group, performs activities of market research, performs accounting, has access to external finance, and/or if the company carries out activities for environmental improvement.[4]

[4] Our justification for the inclusion of this last variable in the set of variables that tries to characterize firms from the point of view of good business practices is that, once controlling for industry fixed effects in estimation, the answers to this question in the survey respond, at least in part, to the firm's attitude about the environmental aspects of its activity.

Second, there are other explanatory variables included such as whether the main customer of the firm is foreign, whether the firm has a craft certification, whether the firm has its own local HQ, whether the surveyed firm is the mother company, and whether it has a male manager. We expect a positive sign for the indicator of the main customer of the firm being foreign, since competition in international markets is expected to put pressure on firms' innovation and knowledge-related activities. It also makes access to information and communication tools more crucial. Thus, for instance, Bratti and Felice (2012) show that there exists a positive relationship between firms' openness to trade and firms' innovation activities. In addition, one may argue that more knowledge-oriented firms are more likely to enter foreign markets. However, we do not consider this possibility here. The reason has to do with data limitations in the survey as regards the lack of a proper question about whether or not firms are exporters. The questionnaire does not ask this question, but rather whether the main client of the firm is foreign. Hence, answering *no* to this question does not guarantee it being a non-exporting firm. Some evidence of this is the small number of firms that would be considered exporters in the Ecuadorian manufacturing sector, 236 out of 42,292 firms (0.56%).

Additionally, we consider firm age (in logs and its square) and firm size (as measured by number of workers in logs and its square). Finally, all estimations in the paper also include three groups of dummy variables that characterize firms from the point of view of locations (24 provinces), industries (12) and legal forms (8).[5] For the sake of brevity, the results for all these dummy variables are not displayed in the tables of this paper. However, the results obtained are discussed.

The estimated mean marginal effects for the explanatory variables in the vector x_{1i} in models (1)–(2) corresponding to the three firm's choices (R&D, workers' training and ICT) are presented in Table 1. As regards the group of variables intended to capture the degree of professionalism and modernization of the firm in terms of quality and diligence in management and business practices, the results in Table 1 indicate that all of them are statistically significant and have a positive sign in explaining the propensities to invest in R&D, workers' training and ICT. For the R&D decision, the strongest effect is estimated for the realization of activities of market research, what may indicate that R&D performance is driven in a relevant manner by firms' demand concerns. For the workers' training decision, the strongest effects are found for firms carrying out activities for environmental improvement and, again, for market research activities. Finally, for ICT use, the strongest effects are found for firms performing accounting and for firms belonging to an enterprise network or business group, very likely due to the required software and hardware for accounting activities and the more vital necessity of those firms to be interrelated inside the business group. To illustrate the estimated magnitude of marginal effects for these variables, let us take as an example the estimated marginal effect for the enterprise network dummy in the ICT use equation, 0.0982 (see Table 1). The interpretation of this value means that, should the firm enter a business group, its average

[5] See the "Data" section above and Appendix 1 or 2.

estimated probability of ICT use in the data, which is 0.098 (see the heading of column 3 in Table 1), would instead be 0.1962 (19.62%).

Additionally, size has a positive and statistically significant effect in the three choices, but at a decreasing rate (excluding the case of workers' training); also, age has a positive (but at a decreasing rate) and statistically significant effect, but in this case only for workers' training and ICT use; to have a male manager is negatively related with workers' training and ICT use; to be the mother company is positively related with workers' training and mainly with ICT use; to have a craft certification seems to require investments in workers' training but is negatively related with the use of ICT; to have your own local HQ is either non-relevant or it is negatively related with ICT use; and, finally, to have a foreign main customer is not statistically significant to explain the R&D and workers' training decisions, but contributes to explain higher likelihood of ICT use.

According to location variables (where the reference category is Pichincha), we obtain that all statistically significant marginal effects are negative, indicating that Pichincha is in general outperforming other Ecuadorian provinces in terms of R&D, workers' training and ICT use. Among industries, with reference category Food, Beverages and Tobacco, the ones with a higher probability to invest in R&D are Chemicals and petroleum products, Office machinery and electrical equipment, and Communications, precision, optical and medical equipment, all of them classified by the OECD as of medium and high technology and, thus, more R&D oriented. However, Non-metallic mineral products presents a lower R&D investment probability. For the performance of workers' training, Chemicals and petroleum products, and Communication, precision, optical and medical equipment also repeat, but Rubber and plastics also appears on the scene. For ICT use, all industries have a higher probability than the reference category, highlighting especially the industries of Rubber and plastics, Wood, paper and printing, Chemicals and petroleum products, Office machinery and electrical equipment, and Communication, precision, optical and medical equipment.

With respect to legal forms, none is statistically significant for the R&D investment decision (the reference category corresponds to the largest group of firms belonging to natural persons). However, for the workers' training and ICT use decisions, to be a Private company increases the likelihood of firms performing both activities, although the legal form associated with the highest likelihood of investing in workers' training is Association and the one for ICT use is Cooperative.[6]

Correlation coefficients for the error terms associated to the three choices are included at the bottom of Table 1. They are positive and highly statistically significant ($\rho_{\mathrm{RD,WT}} = 0.435$; $\rho_{\mathrm{RD,ICT}} = 0.339$; $\rho_{\mathrm{WT,ICT}} = 0.332$), giving support to our joint maximum likelihood estimation versus estimation of three independent probit models. The correlation coefficient measures the correlation between the disturbances in

[6] For the R&D and ICT decisions it was not possible to estimate coefficients associated with the legal form Foreign company, the reason being a perfect prediction of zeros in the corresponding firm's choices when the Foreign company dummy also has a value of zero. Firms with this legal form account for only 0.01% of total manufacturing firms.

Table 1 Firms' choices: R&D, training and ICT (multivariate probit)

	Variables	(1) R&D y = Pred. P(R&D = 1) = 0.010 dy/dx (Aver. Marg. Eff.)	(2) Training y = Pred. P(training = 1) = 0.043 dy/dx (Aver. Marg. Eff.)	(3) ICT y = Pred. P(ICT = 1) = 0.098 dy/dx (Aver. Marg. Eff.)
Professionalization	Enterprise network	0.00742***	0.0443***	0.0982***
		(0.00129)	(0.00310)	(0.00437)
	Market research	0.0356***	0.0753***	0.0750***
		(0.00450)	(0.00845)	(0.0101)
	Accountancy	0.0119***	0.0379***	0.123***
		(0.00223)	(0.00467)	(0.00749)
	Access to finance	0.00440***	0.0156***	0.0165***
		(0.000961)	(0.00203)	(0.00263)
	Environment	0.0205***	0.0951***	0.0710***
		(0.00376)	(0.0106)	(0.0117)
Other regressors	Main customer foreign	0.00330	0.00825	0.0494**
		(0.00282)	(0.00821)	(0.0217)
	Craft certification	− 0.000617	0.00374*	− 0.00809***
		(0.00100)	(0.00204)	(0.00261)
	Own local HQ	− 0.000476	− 0.00246	− 0.0105***
		(0.000858)	(0.00171)	(0.00228)
	Mother company	− 0.000400	0.00974**	0.0200***
		(0.00130)	(0.00386)	(0.00610)
	Male manager	0.000911	− 0.00366*	− 0.0111***
		(0.000996)	(0.00215)	(0.00293)
	Log workers	0.00408***	0.0138***	0.0669***
		(0.000896)	(0.00196)	(0.00292)
	(Log workers)^2	− 0.000350**	4.37e−05	− 0.00637***
		(0.000147)	(0.000408)	(0.000784)
	Log age	0.00127	0.0110***	0.0135***
		(0.00124)	(0.00271)	(0.00360)
	(Log age)^2	− 0.000397	− 0.00307***	− 0.00330***
		(0.000320)	(0.000716)	(0.000935)
	Constant	− 3.146***	− 2.432***	− 2.554***
		(0.098)	(0.057)	(0.050)
Observations		42,292	42,292	42,292
Log pseudo-likelihood			− 14,207.293	
Correlation coefficients			ρ_{12} = 0.435, p val. = 0.000; ρ_{13} = 0.339, p val. = 0.000; ρ_{23} = 0.332, p val. = 0.000	

Estimations control for the firm's location (province), industry, and legal form fixed effects. Robust standard errors in parentheses.

dy/dx for dummy variables is the discrete change from the 0 to the 1 category.

***$p < 0.01$, **$p < 0.05$, *$p < 0.1$

the equations, the omitted factors. That is, ρ measures the correlation between the outcomes after the influence of the included factors such as explanatory variables is accounted for. Thus, for instance, the value 0.435 measures the correlation between the firm's choices of investing in R&D and in workers' training that remains after all explanatory variables effects on this correlation have been taken into account. Interrelations between error terms is not uncommon in empirical studies. The first reason is that the same actor (firm) according to a given strategy takes decisions and, thus, they may be actually related. The second source of interdependence in the errors of different choice equations is directly related to some possible model misspecification. If, for instance, the true model for the probabilities of different choices contained omitted variables, the error terms would be correlated by construction. These two reasons make the use of estimation methods allowing for this type of correlations generally advisable.

Firms' Investments: R&D and workers' Training

Let $y^{*}_{2i,j}$ (j = R&D or WT) denote the firm's latent R&D effort or workers' training intensity, which are defined as the log of the annual expenditure per employee in R&D or workers' training. These two latent knowledge investment intensities are formally modeled as:

$$y^{*}_{2i,j} = \beta^{'}_{2,j} x_{2i} + \varepsilon_{2i,j}, \tag{3}$$

where $\beta^{'}_{2,j}$ captures the effects of explanatory variables on the potential knowledge investment intensity j and $\varepsilon_{2i,j}$ denotes idiosyncratic errors that affect $y^{*}_{2i,j}$. The observed counterparts to $y^{*}_{2i,j}$ are defined as:

$$y_{2i,j} = \mathbf{1}\left[y^{*}_{1i,j} > 0\right] y^{*}_{2i,j} \equiv y_{1i,j} \cdot y^{*}_{2i,j}, \tag{4}$$

where 1[] again denotes the indicator function taking the value one if the condition between squared brackets is satisfied, and zero otherwise. This notation reflects that R&D or workers' training intensities of firm i are observed to be positive only if firm i performs R&D or workers' training activities, respectively ($y_{1i,j}=1$, see Eqs. (1) and (2) above). In estimation, we both allow for correlation of firm idiosyncratic error terms of each knowledge investment intensity equation and its corresponding associated dichotomous decision (i.e., correlation of $\varepsilon_{2i,j}$ in Eq. (3) with $\varepsilon_{1i,j}$ in Eq. (1)) and for correlation between idiosyncratic error terms in the two knowledge investment intensity equations (i.e., correlation between $\varepsilon_{2i,\text{R\&D}}$ and $\varepsilon_{2i,WT}$). For computational and convergence purposes, we follow a two-step estimation procedure. In the first step, we use the estimated coefficients from the trivariate probit model $\beta^{'}_{1,j}$ in Eq. (1) (for j = R&D or WT) to construct two Heckman's lambda terms (also called inverse Mill's ratios) that are used in the second step as additional regressors for sample selection bias correction in the estimation of two equations, one for the log of R&D intensity and another for the log of workers' training intensity. These omitted variables, $\lambda(\beta'_{1,j}x_{1i})$, are calculated by the ratio $\phi(\beta'_{1,j}x_{1i})/\Phi(\beta'_{1,j}x_{1i})$, where ϕ and Φ are, respectively, the density and the cumulative distribution function of a normal distribution.[7] In the second step, the two knowledge investment intensity equations corrected for sample selection bias are jointly estimated with the Stata command *cmp* (Roodman 2011) by pseudo-simulated maximum likelihood. As this procedure, besides corrections for sample selection, also allows for a non-zero correlation between the two knowledge investment intensities error terms, in this paper we call it a bivariate Heckman. This two-step Heckman procedure allows for consistent estimation of parameters in the R&D and workers' training intensity equations that can be extrapolated to population in spite of being estimated with the sub-population of R&D or workers' training performing firms, respectively (Heckman 1979). This is a suitable method with our data since there are many manufacturing firms in Ecuador not performing R&D and/or not performing workers' training activities. The implemented method will allow testing for the presence of sample selection in each one of the two intensity equations and for the interrelation between firms' expenditures in R&D and workers' training.

The explanatory variables in x_{2i} are the same than in x_{1i} with the exception of the variable log workers squared, which is not included in the vector x_{2i} (the top panel in "Appendix 3" reports the correlation coefficients among the main variables in x_{1i} and, consequently, in x_{2i}). This exclusion restriction will contribute to identification in the knowledge investment intensity equations. Notice that, although we use as dependent variables in these equations the log of knowledge expenditures per worker, the variable size (log workers) is included in these equations to allow for knowledge expenditures not necessarily being proportional to size.

Estimation results for the knowledge investment intensity equations are presented in Table 2. At the end of the table, we have the estimated coefficients associated with the Heckman's lambda terms for R&D and workers' training sample selection

[7] The two Heckman's lambda terms, Lambda R&D and Lambda training, are calculated as the expected value of the error term in the corresponding equation of interest (log of R&D or log of workers' training intensity equations) conditional to the explanatory variables x_{1i} in Eq. (1) and x_{2i} in Eq. (3), and conditional on the observability of positive values for the R&D or workers' training investments, respectively. As in our case x_{2i} is a subset of x_{1i}, for regressors we only need to condition on the vector of explanatory variables x_{1i}. The Heckman's method assumes that the errors in the two equations involved for sample selection correction (in our case $\varepsilon_{2i,j}$ in Eq. (3) and $\varepsilon_{1i,j}$ in Eq. (1)) have a bivariate normal distribution with 0 means and standard deviations $\sigma_{\varepsilon_{2,j}}$ and 1, respectively, being $\rho_{21,j} = (\sigma_{21,j}/\sigma_{\varepsilon_{2,j}}\sigma_{\varepsilon_{1,j}}) = \left((\sigma_{21,j})/\sigma_{\varepsilon_{2,j}}\right)$ the correlation coefficient between error terms. Under these assumptions, the calculous of the following conditional mean gives rise to the lambda term: $E\left(\varepsilon_{2i,j}\middle|x_{1i}, y_{1i,j} = 1\right) = E\left(\varepsilon_{2i,j}\middle|x_{1i}, y^*_{1i,j} > 0\right) = E\left(\varepsilon_{2i,j}\middle|x_{1i}, \beta'_{1,j}x_{1i} + \varepsilon_{1i,j} > 0\right)$ $= \sigma_{\varepsilon_{2,j}}E\left(\frac{\varepsilon_{2i,j}}{\sigma_{\varepsilon_{2,j}}}\middle|x_{1i}, \varepsilon_{1i,j} > -\beta'_{1,j}x_{1i}\right) = \rho_{21,j}\sigma_{\varepsilon_{2,j}}E\left(\varepsilon_{1i,j}\middle|x_{1i}, \varepsilon_{1i,j} > -\beta'_{1,j}x_{1i}\right) = \sigma_{21,j}\lambda\left(-\beta'_{1,j}x_{1i}\right)$, where $\lambda\left(-\beta'_{1,j}x_{1i}\right)$ is the lambda term or the inverse Mill's ratio and $\sigma_{21,j}$ is the covariance between $\varepsilon_{2i,j}$ and $\varepsilon_{1i,j}$. According to the properties of truncated normal distributions, and since the lambda term is no more than the mean of a standard normal random variable truncated from below at $-\beta'_{1,j}x_{1i}$, it can be calculated as the ratio of the density function ϕ over the cumulative distribution function Φ of a standard normal distribution evaluated at $\beta'_{1,j}x_{1i}$, that is $\lambda\left(-\beta'_{1,j}x_{1i}\right) = \left\{\phi\left(-\beta'_{1,j}x_{1i}\right)\big/\left[1 - \Phi\left(-\beta'_{1,j}x_{1i}\right)\right]\right\} = \phi\left(\beta'_{1,j}x_{1i}\right)\big/\Phi\left(\beta'_{1,j}x_{1i}\right)$ $= \lambda\left(\beta'_{1,j}x_{1i}\right)$. Notice that on its final expression we have used the symmetry property of the normal distribution.

corrections. They are positive and statistically significant, indicating that it was relevant to use sample selection methods for estimation instead of simple ordinary least squares and that unobserved factors that increase firms' likelihoods to invest in R&D and workers' training activities are positively correlated with their R&D and workers' training investment intensities. Therefore, endogenous selection of firms into these activities was an issue to take care of. Furthermore, at the bottom of Table 2 we also present information about the correlation coefficient between the error terms in the two intensity equations. The null hypothesis of $\rho_{\mathrm{R\&D,WT}}=0$ is rejected ($\rho_{\mathrm{R\&D,WT}}=0.535$, with p value$=0.000$). Hence, there is a positive and statistically significant correlation between firms' unobservables in the R&D and workers' training intensity equations. This finding supports the use of our bivariate Heckman estimation procedure.

In columns 1 and 2 of Table 2, we present the results for the R&D and workers' training intensity equations. It is interesting to notice that all the variables in the group of professionalization and good business practices explain, with positive and statistically significant coefficients, higher intensity in both R&D and workers' training investments. Beyond indicating firms' good business practices, variables such as belonging to a business group or declaring access to finance might be indicative of lack of financial constraints to carry out these investments. Similarly, the performance of market research activities might indicate deficiencies in demand that are responded to with firms' knowledge creation investments (especially with R&D).

Additionally, for the R&D intensity equation, we find that for older firms there is a negative effect of age and, more interestingly, a positive effect of the main customer being foreign. For the workers' training intensity equation, we obtain that larger firms have a lower intensity, while the opposite happens for firms with a male manager, mother companies, and firms for which main customer is foreign. It seems again that competition in international markets exerts pressure on firms' knowledge creation efforts.

In the R&D intensity equation, the location variables that are positive and statistically significant with respect to the reference category (Pichincha) are Carchi, Napo and Zamora. For the workers' training intensity equation, they are Imbabura, Napo, Zamora, Galápagos and Orellana. Those that are negative and statistically significant for R&D are El Oro, Guayas, Loja, Manabí and Orellana. For the workers' training equation, they are Azuay, Chimborazo, Loja, Los Ríos and Pastaza. As regards industries, the ones justifying higher R&D intensities are Chemicals and petroleum products, Rubber and plastics, Office machinery and electrical equipment, and Communication, precision, optical and medical equipment. For the case of workers' training intensities, they are Chemicals and petroleum products, and Communication, precision, optical and medical equipment (and those with the lowest intensities, Textiles and wearing apparel and Furniture). With respect to legal forms, most of them are positive and significantly associated with both expenditures as regards the reference category (individual firms). This applies to Private and Public companies, Cooperatives and Associations for R&D intensity, and to Private, Foreign companies and Cooperatives for workers' training intensity. For R&D intensity, the highest coefficient is found for Public companies and for workers' training intensity for Foreign companies.

Table 2 Firms' investments: R&D and training (bivariate *Heckman*)

	Variables	(1) R&D intensity	(2) Training intensity
Professionalization	Enterprise network	1.327***	0.547***
		(0.372)	(0.157)
	Market research	2.783***	0.691***
		(0.695)	(0.167)
	Accountancy	2.198***	0.719***
		(0.517)	(0.149)
	Access to finance	0.569***	0.247***
		(0.216)	(0.080)
	Environment	1.695***	0.564***
		(0.501)	(0.177)
Other regressors	Main custom foreign	0.524*	0.344*
		(0.315)	(0.180)
	Craft certification	0.036	− 0.068
		(0.217)	(0.078)
	Own local HQ	− 0.072	0.001
		(0.161)	(0.068)
	Mother company	− 0.165	0.507***
		(0.244)	(0.100)
	Male manager	0.312	0.151*
		(0.214)	(0.083)
	Log workers	− 0.062	− 0.396***
		(0.114)	(0.062)
	Log age	0.501	− 0.024
		(0.321)	(0.133)
	$(\text{Log age})^2$	− 0.150**	0.013
		(0.075)	(0.033)
	Lambda R&D	3.123***	
		(0.902)	
	Lambda training		0.812***
			(0.295)
	Constant	− 5.745*	2.592***
		(3.022)	(0.816)
Observations		412	1828
Log pseudo-likelihood		− 3833.363	
Correlation coefficient		$\rho_{R\&D,WT} = 0.535$ p val. = 0.000	

Estimations control for the firm's location (province), industry, and legal form fixed effects. Robust standard errors in parentheses.

***$p < 0.01$, **$p < 0.05$, *$p < 0.1$

Firms' Productivity, R&D/Training Investments and ICT Choices

There are methods for estimation of firms' productivity that require panel data availability. These are, for instance, the semi-parametric methods developed by Olley and Pakes (1996), Levinsohn and Petrin (2003), or Wooldridge (2009). However, the cross-sectional nature of our data prevents us from using these methods.[8] Instead, for robustness of results, we use in this paper three productivity measures as dependent variables. The first one is the standard one of labor productivity, calculated as the log of sales per worker. The second and the third ones rely on estimation of Cobb–Douglas and Translog production functions and proxy TFP by their estimated residuals, where firms' output is measured by log sales and firms' inputs by log number of workers, log capital (stock at book values of tangible fixed assets) and log materials. The Cobb–Douglas and the Translog production functions are estimated for each one of the 12 considered industries. In the Cobb–Douglas production function, the average elasticity for materials (β_{m}) is 0.59, for labor (β_{l}) 0.39 and for capital (β_k) 0.12. In the Translog production function, the average elasticity for materials (β_{m}) is 0.58, for labor (β_{l}) 0.36 and for capital (β_k) 0.11. The estimated industry specific input elasticities are shown in "Appendix 4".

We perform a set of regressions to uncover the role of our main variables of interest (R&D and workers' training investments and ICT use) on TFP. In these regressions, we also control for firms' geographical location, industries, legal forms and other regressors.

The results are in Table 3. Column 1 corresponds to the log linear productivity regression relating labor productivity to labor and its squared, age and its squared, capital per worker, materials per worker, performance of market research studies, the presence of a foreign main customer, belonging to a business network, having environmental concerns, the existence of own local HQ, having a male manager, and knowledge creation activities.[9] This regression is the only one in Table 3 that includes capital and materials intensities. The reason is that, since the dependent variable in this case is labor productivity, controlling for capital and materials intensities facilitates results from estimation to be interpreted as TFP effects (Crépon et al. 1998). In columns 2 and 3 of Table 3, we present results for the residual TFP from Cobb–Douglas and ranslog production functions, respectively.[10] Knowledge creation activities are captured by their latent counterparts (i.e. the potential R&D and workers' training intensities and the propensity to use ICT).[11] Thus, we can write:

[8] In a recent survey on TFP estimation, Van Beveren (2012) performs an empirical evaluation of TFP estimation methods as regards yielding different conclusions when conducting policy or impact evaluations (e.g., trade liberalization, deregulation, etc.). He shows that comparing OLS estimates with more sophisticated methods available for panel data, high correlations between different estimated TFP measures emerge (higher than 0.8 or 0.95 depending on the methods) and, more importantly for us, similar conclusions are obtained when evaluating the effect of some policy change with different TFP measures.

[9] The mean of log labor productivity in our sample is 8.858.

[10] The means of the Cobb–Douglas and Translog TFP (in logs) in the sample are 3.231 and 5.141, respectively.

[11] In addition to the seminal CDM model (Crépon et al. 1998), our paper is in line with Griffith et al. (2006), Crespi and Zuniga (2012) and Aboal and Tacsir (2018), who also estimate a CDM model not only with performing (innovative) firms, but with all firms.

$$y_{3i} = \beta_3' x_{3i} + \gamma_{3j} y_{i,j}^* + \varepsilon_{3i}, \tag{5}$$

where γ_{3j} is a vector of three elements associated with the potential knowledge investment intensities $y_{2i,j}^*$ in Eq. (3), with $j=1$ or 2 being referred to R&D or workers' training, respectively, and with the potential ICT propensity, $P(y_{1i,j}=1)=P(y_{1i,j}^*>0)$ in Eqs. (1)–(2), with $j=$ICT. The coefficients γ_{3j} capture the effects of knowledge creation activities on productivity, β_3' captures the effect of all the other explanatory variables and controls (x_{3i}) in the regression for productivity, and ε_{3i} denotes idiosyncratic errors. The explanatory variables in x_{3i} are a subset of the variables in x_{2i} and in x_{1i} (notice that x_{2i} only differs from x_{1i} in that it does not include the variable log workers squared). To work in Eq. (5) with two estimated intensities (for R&D and workers' training) and one estimated probability (for ICT use) allows taking endogeneity concerns of knowledge creation variables in the productivity equation into account. For instance, the most productive firms could raise more internal and external funds for these investments, implying reverse causality from productivity to drivers of knowledge and, therefore, simultaneity bias. Furthermore, knowledge creation variables could be affected by measurement errors that especially affect their expenditures. The excluded variables in x_{3i} with respect to x_{2i} (or x_{1i}) allow for the endogenization of R&D and workers' training intensities and ICT use in the productivity regressions.[12] Although, in this paper, instead of using predicted regressors for drivers of knowledge to correct for endogeneity, we use the equivalent method of substituting predicted regressors by their observed value and the estimated residual calculated as the difference between their observed value and their predicted value (control function approach; see Rivers and Vuong 1988; Wooldridge 2010). In this way, the included estimated residuals not only clean coefficients from observed values of endogeneity bias, but also deliver coefficient estimates for the residuals, which statistical significance provides tests of endogeneity for the knowledge

[12] The variables acting as instruments are whether the firm is the mother company, declares to have access to finance, carries company accounting, and has a craft certification. Since the cross-sectional nature of our dataset limits our choice of instruments, they have been empirically selected to guarantee validity. That is, with our selection of exclusion restrictions, we accomplish simultaneously two objectives: (1) that the instruments used are significantly correlated with the R&D, workers' training and ICT variables; and, (2) that they are not correlated with the error term in the productivity regressions. We have checked 1 by performing Wald tests of joint non-significance of instruments in the first stage estimates for the variables to be endogeneized. Accordingly, with the first stage estimates for ICT that are in column 3 of Table 1, we obtain a $\chi^2_{(4)}=602.91$ (p value$=0.0000$); with the first stage estimates for R&D and workers' training intensities that are in columns 1 and 2, respectively, of Table 2, we obtain $\chi^2_{(4)}=18.52$ (p value$=0.0010$) and $\chi^2_{(4)}=49.10$ (p value$=0.0000$). Hence, we reject the null of non-significance. The results for the verification of 2 will be presented below when commenting the estimates from the productivity regressions. All papers in the CDM framework assume exclusion restrictions in order to identify and estimate the model. Typically, they estimate a more parsimonious productivity equation as regards regressions in previous stages (such as the estimation of innovation inputs effort and/or innovation output equations). For example, Arvanitis and Loukis (2009) besides the variables being instrumented include physical capital, R&D and controls (for size and sector). Griffith et al. (2006) and Crespi and Zuniga (2012), besides the predicted innovation variables only include in the productivity regressions physical capital investment per worker and the usual controls for firm size and industry dummies. Aboal and Tacsir (2018) further include the ratio of professionals and technicians in the workforce.

Table 3 Firms' productivity, R&D/Training investments and ICT choices

	Variables	(1) Log labor productivity	(2) Cobb–Douglas total factor productivity	(3) Translog total factor productivity
	Log R&D intensity	0.108***	0.097***	0.080***
		(0.024)	(0.024)	(0.024)
	Log training intensity	0.259***	0.224***	0.183***
		(0.033)	(0.033)	(0.028)
	ICT use	0.125***	0.098***	0.106***
		(0.016)	(0.015)	(0.015)
Other regressors	Market research	− 0.074***	− 0.066***	− 0.053**
		(0.027)	(0.025)	(0.024)
	Main customer foreign	0.010	0.009	0.001
		(0.056)	(0.057)	(0.052)
	Enterprise network	− 0.009	− 0.013	0.000
		(0.010)	(0.011)	(0.010)
	Environment	0.061**	0.059*	0.043
		(0.030)	(0.032)	(0.028)
	Own local HQ	− 0.230***	− 0.231***	− 0.223***
		(0.007)	(0.007)	(0.007)
	Male manager	− 0.051***	− 0.051***	− 0.033***
		(0.011)	(0.010)	(0.010)
	Log workers	0.225***	0.090***	0.130***
		(0.025)	(0.024)	(0.022)
	(Log workers)^2	− 0.003	− 0.004	− 0.015***
		(0.003)	(0.003)	(0.003)
	Log age	0.104***	0.093***	0.099***
		(0.012)	(0.013)	(0.013)
	(Log age)^2	− 0.020***	− 0.017***	− 0.019***
		(0.004)	(0.004)	(0.004)
	Log capital per worker	0.113***		
		(0.003)		
	Log mater. per worker	0.557***		
		(0.004)		
	Constant	1.956***	− 1.409***	− 1.219***
		(0.171)	(0.176)	(0.153)
	Resid. Log R&D inten.[a]	− 0.108***	− 0.097***	− 0.084***
		(0.023)	(0.023)	(0.023)
	Resid. Log training. inten.[a]	− 0.249***	− 0.216***	− 0.179***
		(0.033)	(0.033)	(0.028)
	Resid. ICT use[a]	− 0.010	0.017	− 0.013
		(0.067)	(0.071)	(0.067)
Observations		41,665	41,665	41,665
R^2		0.637	0.067	0.055

Table 3 (continued)

Estimations control for the firm's location (province), industry, and legal form fixed effects. Robust bootstrapped standard errors in parenthesis (500 replications). Since the dependent variables are in logs, coefficient estimates for explanatory variables that are in logs are to be interpreted as elasticities. Those for dummy variables have the interpretation of semielasticities.

[a]Residual terms from the previously estimated R&D intensity, training intensity and ICT use equations, respectively. These terms correct for endogeneity of R&D and training intensities and ICT use in the productivity equations.

***$p<0.01$, **$p<0.05$, *$p<0.1$

creation variables. This is a sensible way to instrument drivers of knowledge in the productivity equation. The same method has been applied in Arvanitis and Loukis (2009) when investigating the effects of ICT, human capital and organizational practices on labor productivity in Greek and Swiss firms with cross-sectional data.

Our main interest from the productivity regressions is the estimation of elasticities for the R&D and workers' training intensities and the corresponding semi-elasticity for ICT use. These magnitudes appear at the top of Table 3, and the coefficient estimates for their associated residual terms at the bottom of that table.[13] The results in Table 3 indicate that the residuals for R&D and workers' training are negative and statistically significant, indicating not only the convenience of correcting for endogeneity as regards these knowledge creation variables in the productivity equations but also the very likely presence of measurement errors in these variables, since the negative sign of coefficients for residuals is a signal of attenuation bias. Differently, the residual for ICT is non-significant.[14,15] The estimated elasticity for R&D intensity indicates that an increase of 1% increases productivity in a range from 0.08% (using the Translog TFP) to 0.108% (using labor productivity). For workers' training intensity, an increase of 1% increases productivity in a range from 0.183% (in the Translog case) to 0.259% (with labor productivity). The estimated semi-elasticity for ICT use indicates that using ICT increases productivity in a range from 9.8% (with the Cobb–Douglas) to 12.5% (with labor productivity), being for the Translog 10.6%. Estimated values for the Cobb–Douglas TFP regression are quite close to the Translog ones.

13 Estimated residuals for the two knowledge investment intensity variables come from the bivariate Heckman in "Firms' Investments: R&D and workers' training" section. The estimated residual for the ICT dichotomous decision comes from the difference between $y_{1i,ICT} - P\left(y^*_{1i,ICT} > 0\right)$ obtained from results in "The Firms' Decisions on Knowledge Creation Activities: R&D, workers' training and ICT Use" section. Similar results were obtained when alternatively using a generalized residual for ICT.

14 Hence, the estimates reported in Table 3 already eliminate the instrumentation of the ICT variable.

15 We have checked that the instruments are not correlated with the error term in the productivity regressions by performing Sargan—Hansen tests of overidentifying restrictions. Notice that we can perform such tests since we have four instruments to instrument R&D and workers' training intensities. For the labor productivity regression in column 1 of Table 3, we get a $\chi^2_{(2)}=1.98$ (p value $=0.3722$); for the *Cobb–Douglas* TFP regression in column 2 of Table 3, we get a $\chi^2_{(2)}=2.52$ (p value $=0.2838$); and, for the *Translog* TFP regression in column 3 of Table 3, we get a $\chi^2_{(2)}=1.52$ (p value $=0.4681$). Hence, we do not reject the nulls of no correlation.

Among the group of other regressors in the productivity regressions, there are two variables included to control for demand factors and competition. These variables are whether the firm invests in market research and whether the firm's main customer is foreign. The second variable is never statistically significant (although with positive sign), giving support to our thought about this variable not isolating exporters from non-exporters (notice that one firm can be an exporting firm without its main customer being foreign). The first variable has a negative sign and is statistically significant. Obtained results for this variable are in favour of the market research dummy to be an indicator for the firm's demand conditions. In particular, it can proxy for bad demand conditions that require market research. Since we are working with 'revenue' productivity and not with 'physical' productivity, a downturn in demand puts pressure on firms' prices to go down and, therefore, 'revenue' productivity decreases. The inclusion of this variable is relevant to clean the effects of other variables from demand conditions, which can affect 'revenue' productivity through output prices instead of through efficiency. Firms' individual prices, and not only industry prices, would be required for 'physical' productivity, and they are commonly absent in most of the datasets.[16]

For the variables age and size, we obtain that firms' age explains higher productivity but at a decreasing rate, while for size the decreasing rate only appears in the Translog case. In addition, having a male manager, or its own local, does not seem to work in favor of productivity. As regards geographical location of firms, most of the provinces show a lower productivity than the reference category, Pichincha. Exceptions are El Oro, Los Ríos, Galápagos, Sucumbíos, Orellana and Santo Domingo, with similar productivity to Pichincha. Industries with clearly lower productivity than the reference category, Food, beverages and tobacco, are Wood, paper and printing, Rubber and plastics, and Metal products. Others are only negative and statistically significant in two out of three of the productivity regressions. This is the case for Non-metallic mineral products, Furniture, and Office machinery and electrical equipment. The only sector with higher productivity than the reference category seems to be Chemicals and petroleum products (and this happens only for two of the three productivity regressions). With respect to legal forms, the one with higher associated productivity is Private Company (in two of the three productivity regressions). The ones with lower are Non-profit Company, Cooperatives and, sometimes, Associations. The reference category corresponds to Natural person companies.

Firms' Markups, R&D/Training Investments and ICT Choices

We estimate firm specific markups (defined as the ratio of the price over marginal cost) following De Loecker and Warzyinski (2012) as:

$$\mu_{is} = e_{is}^{X} \big/ sh_{i}^{X}, \tag{6}$$

[16] If we had a proper export dummy, it could also contain relevant demand side information when firm prices are set differently in domestic than in export markets (Aw et al. 2011).

where μ_{is} is the markup of firm i in industry s, e^X_{is} is the output elasticity of variable input X (obtained for each one of the 12 considered industries) and sh^X_i is the firm's revenue share of variable input X. The revenue share of variable input X is defined as the total cost of that input over firm's total sales.

This methodology stems from Hall (1988), which was the first work that used data on production to estimate industry markups. The main advantage of the method proposed by De Loecker and Warzyinski (2012) is that it allows deriving an expression for calculating firm-specific markups under two mild assumptions. The first is cost-minimizing producers and the second is the existence of at least one variable input of production. The flexibility of this method, compared to structural models that simultaneously combine production or cost functions and markups equations, comes, among others, from the fact that there is no need to use a particular functional form for the demand function to estimate the markups. In relation to this, Corchón and Moreno (2010) show that standard forms for the demand system, such as the linear or the constant elasticity of substitution demand systems, do not provide a good explanation for the markups of Spanish manufacturing companies during the period 1990–2005. In addition, there is no need to make assumptions about the mode of competition. However, the value of the markup itself does depend on the specific nature of competition among firms. The intuition behind is that, under perfect competition, prices are equal to marginal costs and, hence, input choices of cost minimiser firms will make the revenue share to be equal to the output elasticity of the input. In this case, the value of the markup will be 1. However, under imperfect competition, the revenue shares are lower than the output elasticities, which implies markups above 1 (prices above marginal costs).

Since only one variable input is required, in our case the one chosen has been materials. We have a double reason for this choice. On the one hand, as the alternative variable input is labor, this production factor is more likely affected by adjustment costs. On the other hand, the dataset has many more missing values on the information of firms' wages than on the information about materials costs. Opting for the variable input labor will imply discarding a big proportion of the sample size used for estimation of total factor productivity measures. In particular, selecting materials as the freely adjustable input we can estimate the markups equation with 41,647 observations. Should we have selected labor the number of observations would have been only 16,763.

The denominator in (6), $sh^{\mathrm{Materials}}_i$, can be directly computed with firms' ratios of materials costs over sales using the data available in the dataset. However, the output elasticity of the input materials has to be estimated from a production function. In the previous section ("Firms' Productivity, R&D/Training Investments and ICT Choices"), we have estimated both a Cobb–Douglas and a Translog production function. However, since the Cobb–Douglas restricts output elasticities of inputs to be constant for all firms in a given industry, we rely on the Translog estimates, which allow between firms' variation in markups both coming from the numerator and the denominator in (6).[17]

[17] De Loecker and Warzyinski (2012) compare in their empirical work estimation results obtained with a translog and with a Cobb–Douglas production technology, although their main empirical specification relies on a translog. In their online appendix it is shown that estimated percentage differences between exporters and non-exporters in terms of markups are very similar under both production technologies, although somewhat lower with a *translog*.

Expressed in natural logarithms, the previously estimated Translog production function in section "Firms' Productivity, R&D/Training Investments and ICT Choices" above was as follows:

$$\begin{aligned} \text{Sales}_{is} = \alpha_s + \beta_{ls} l_i + \beta_{ks} k_i + \beta_{ms} m_i + \beta_{ll,s} l_i^2 + \beta_{kk,s} k_i^2 \\ \beta_{mm,s} m_i^2 + \beta_{lk,s} l_i k_i + \beta_{lm,s} l_i m_i + \beta_{km,s} k_i m_i + tfp_{is}, \end{aligned} \tag{7}$$

From where the output elasticity of materials is computed as:

$$e_{is}^{\text{Materials}} = \beta_{ms} + 2\beta_{mm,s} m_i + \beta_{lm,s} l_i + \beta_{km,s} k_i \tag{8}$$

Notice that the corresponding elasticity from a Cobb–Douglas would be only $e_{is}^{\text{Materials}} = \beta_{ms}$.

In Table 4, we present information about our markups estimates. We find that the median firm in our sample charges around 43% markup over marginal cost. This median markup estimate is comparable to, though slightly larger than, the estimates in De Loecker and Warzyinski (2012) for Slovenian manufacturing firms using also a gross output production function relying on materials to compute markups. Their 1.22 estimated median markup implies that a median Slovenian manufacturing firm charges around 22% markup over marginal cost. Figures in Table 4 show some variation in markups across industries. We find the largest median markups for Chemicals and petroleum products and Textiles and wearing apparel (62–63%), while the smaller one is for Transport Equipment (26%). Overall, the median markups estimates in all industries fall in the reasonable range of 1–2. However, at the bottom of Table 4, we provide the calculated overall standard deviation of firms' markups (1.56), the between industries' markups standard deviation (0.22), and the within-firms in an industry standard deviation (1.55), which uncover that most of overall variation in firms' markups is not due to firms belonging to different industries but to markup heterogeneity within firms in an industry.

Again, similarly to the TFP regressions, in the markups regressions knowledge creation activities depend on their latent counterparts (i.e. the potential R&D and workers' training intensities and the propensity to use ICT). Thus, we can write

$$\log \mu_{is} = \beta_4' x_{4i} + \gamma_{4j} y_{i,j}^* + \varepsilon_{4i}, \tag{9}$$

where γ_{4j} is a vector of three elements associated with the potential knowledge investment intensities $y_{2i,j}^*$ in (3), with $j = 1$ or 2 being referred to R&D or workers' training, respectively, and with the potential ICT propensity, $P(y_{1i,j} = 1) = P(y_{1i,j}^* > 0)$ in (1)–(2), with $j = \text{ICT}$. The coefficients γ_{4j} capture the effects of knowledge creation activities on markups, $\beta_4{}'$ captures the effect of all the other explanatory variables and controls (x_{4i}) in the regression for markups, and ε_{4i} denotes idiosyncratic errors. To work in (9) with two estimated intensities (for R&D and workers' training) and one estimated probability (for ICT use) allows taking endogeneity concerns of knowledge creation variables in the markups equation into account. Markups can proxy for market power and, therefore, may influence firms' knowledge investment decisions. Additionally, drivers of knowledge may suffer from measurement errors. As we follow in this paper the control function approach, estimation of (9) requires

Table 4 Markup and materials output elasticity estimates (from translog production function), and materials revenue shares (from balance sheet data)

Industry	(1) Markup[a] Mean/median (SD)	(2) Materials output elasticity Mean/median (SD)	(3) Materials revenue share Mean/median (SD)
Food, beverages and tobacco	1.80/1.38 (1.38)	0.62/0.63 (0.11)	0.46/0.46 (0.20)
Textiles and wearing apparel	2.27/1.63 (1.89)	0.51/0.51 (0.09)	0.35/0.31 (0.21)
Leather and footwear	1.84/1.42 (1.29)	0.62/0.62 (0.13)	0.44/0.45 (0.20)
Wood, paper and printing	1.98/1.47 (1.55)	0.57/0.57 (0.11)	0.40/0.40 (0.20)
Chemicals and petroleum products	2.31/1.62 (2.04)	0.61/0.62 (0.11)	0.39/0.37 (0.20)
Rubber and plastics	1.99/1.33 (1.85)	0.54/0.54 (0.15)	0.42/0.43 (0.22)
Non-metallic mineral products	1.88/1.43 (1.47)	0.62/0.63 (0.10)	0.44/0.44 (0.20)
Metal products	1.69/1.34 (1.24)	0.58/0.59 (0.15)	0.44/0.45 (0.20)
Office machin. and electrical equip.	1.97/1.48 (1.59)	0.56/0.56 (0.17)	0.39/0.39 (0.20)
Communi./prec./optic./medic. equip.	1.92/1.56 (1.49)	0.52/0.53 (0.20)	0.34/0.32 (0.18)
Transport equipment	1.50/1.26 (1.20)	0.51/0.52 (0.19)	0.42/0.42 (0.21)
Furniture and n.e.c.	1.92/1.46 (1.55)	0.61/0.61 (0.09)	0.43/0.42 (0.20)
Total	1.94/1.43 (SD overall: 1.56)	0.58/0.58 (SD overall: 0.12)	0.42/0.41 (SD overall: 0.21)

Table 4 (continued)

Industry	(1) Markup[a] Mean/median (SD)	(2) Materials output elasticity Mean/median (SD)	(3) Materials revenue share Mean/median (SD)
	(SD betw. industries: 0.22)	(SD betw. industries: 0.04)	(SD betw. industries: 0.03)
	(SD within firms in an industry: 1.55)	(SD within firms in an industry: 0.11)	(SD within firms in an industry: 0.20)

[a]Although the markup is defined as the ratio of the materials output elasticity over the materials revenue share, notice that figures in column (1) do not coincide with the ratio of figures in column (2) over those in column (3). The reason is that the mean of a variable that is a ratio of other two variables does not mathematically coincide with the value of the ratio calculated by dividing the means of the involved variables. The median markup (instead of the mean) is less severely affected by this distinction and, hence, independently of calculating the ratio of (2) over (3) with mean or median values for the involved variables, the obtained results would be closer to the calculated firms' median markups presented in column (1).

substituting predicted regressors by their observed values and the estimated residual terms from the previously estimated R&D intensity, workers' training intensity and ICT use equations. The inclusion of these predicted residuals corrects for endogeneity of knowledge creation variables in the markups equation and allows testing for it. The excluded variables in x_{4i} with respect to x_{2i} (or x_{1i}) allow for the endogenization of R&D and workers' training intensities and ICT use in the markups regressions.[18]

In column 1 of Table 5, we present results for the estimation of our baseline markups equation in (9). Since the markups are in logs, coefficient estimates for explanatory variables that are in logs are to be interpreted as elasticities. Those for dummy variables have the interpretation of semi-elasticities. Column 2 of Table 5 augments the baseline equation in (9) by including also among regressors the Translog TFP. The estimated residual calculated as the difference between the Translog TFP and its prediction is also included.[19] At the top of that table are the coefficient estimates for our knowledge creation variables of interest in this paper (and that of the TFP in column 2). The coefficient estimates for their associated residual terms are presented at the bottom of the table. The coefficients for the residuals of R&D and workers' training intensities are negative and statistically significant.[20,21] This confirms both the convenience of correcting for endogeneity of knowledge creation variables in the markups equations and the likely presence of measurement errors in some of them, since the negative sign of coefficients for residuals is a signal of attenuation bias. Differently, the coefficient of the residual for TFP is positive and statistically significant (see column 2), indicating that the TFP regressor in the markups equation may suffer more from reverse causality bias than from attenuation bias coming from measurement errors. It is then also important our correction for endogeneity of this regressor.

According to the estimated elasticities from column 1 of Table 5, we find that a 1% increase in R&D intensity increases markups 0.094%. For workers' training intensity, the increase in markups is 0.255%. The estimated semi-elasticity for ICT use indicates that using ICT justifies an increase in markups of 12.5%. Coefficients

[18] The variables acting as instruments are the same ones than in the productivity regressions. See footnote 12.

[19] In column 2 of Table 5 we exclude a variable in x_{4i} with respect to x_{3i}, log workers squared, which contributes additionally to the endogenization of TFP in this markups specification. This variable is significantly correlated with TFP (see its statistical significance in column 3 of Table 3).

[20] In our baseline specification (column 1), the residual for ICT is non-significant. Instead, in column 2 is positive and significant. Hence, the estimates reported in Table 5 for column 1 already eliminate the instrumentation of the ICT variable.

[21] We have checked that the instruments are not correlated with the error term in the markups regressions by performing Sargan–Hansen tests of overidentifying restrictions. In the regression in column 1 of Table 5, we have four instruments to instrument R&D and workers' training intensities, and we get a $\chi^2_{(2)} = 1.99$ (p value $= 0.3703$). In the regression in column 2 of Table 5, we have the same four instruments plus the extra-instrument log workers squared to instrument R&D and workers' training intensities, ICT use, and TFP. In this case, we obtain a $\chi^2_{(1)} = 0.67$ (p value $= 0.4123$). Hence, we do not reject the nulls of no correlation.

for R&D and workers' training intensities are roughly halved when we control for productivity in the regression (column 2). This supports the idea that firms investing in R&D and workers' training charge higher markups because of two reasons: one being that they are also more productive and the other one probably related to the generation of other firms' advantages such as, for instance, higher quality products, allowing firms to charge higher prices. In addition, in column 2, the ICT variable loses its statistical significance, indicating that the effect of this variable on markups acts only through efficiency. Considering TFP as a proxy for marginal costs (as in De Loecker and Warzyinski 2012), elasticities and semi-elasticities in column 2 of Table 5 would be net of the effect of variables on markups acting through the channel of productivity and, hence, they would pick up the effect on markups of these variables through the firm's capacity to fix prices above marginal costs. Therefore, the elasticities for R&D and workers' training intensities are reduced in column 2 to values 0.057% and 0.146%, respectively. Comparing with the magnitudes from column 1, we see that roughly half of the effects of these variables on markups act through decreasing marginal costs, that is, increasing firm's efficiency. The other half accounts for the effect on higher selling prices. Finally, the TFP elasticity in the markups regression is 0.507%.

Following De Loecker and Warzyinski (2012), in the group of other regressors we have included the three inputs in the production function. They recommend their inclusion in the markups regressions in order to eliminate a potential bias that may emerge in firms' investments coefficients when inputs are correlated with unobserved firm's output price variation. Furthermore, and similarly to the TFP regressions, we also include the variables squared labor, age and its square, whether the firm belongs to a business network, has environmental concerns, a male manager, has its own local, performs market research, and its main customer is foreign. The last two variables, market research and main customer foreign, are included to capture demand shocks and market power affecting markups that could bias drivers of knowledge coefficients in the markups equations (since not only markups depend on competition but also competition affects knowledge creation investments).[22] The coefficient for market research is negative and statistically significant, likely related to a lack of firm's demand that calls for market studies. However, the coefficient for main customer being foreign is non-significant.

We obtain that the variables age and size explain higher markups but at a decreasing rate. In addition, belonging to a business network, having a male manager, or its own local , does not seem to work in favor of markups (the opposite happens for firms with environmental concerns). Regarding geographical location variables, most of the provinces present a lower markup than the reference category (Pichincha). Exceptions are El Oro, Los Ríos, Galápagos, Sucumbíos, Orellana and Santo Domingo. All the industries, with the exception of Chemicals and petroleum

[22] The reason is that 'revenue' productivity may still potentially capture differences in firms' prices. In any case, De Loecker and Warzyinski (2012) show that using 'revenue' productivity affects only the level of the markup estimates, and not the correlation between markups and firm-level characteristics.

Table 5 Firms' markups (from translog), R&D/training investments and ICT choices

	Variables	(1) Log markup	(2) Log markup
	Log R&D intensity	0.094***	0.057***
		(0.024)	(0.007)
	Log training intensity	0.255***	0.146***
		(0.032)	(0.016)
	ICT use	0.125***	0.010
		(0.015)	(0.023)
	Translog TFP		0.507***
			(0.083)
Other regressors	Market research	− 0.072***	− 0.040***
		(0.024)	(0.008)
	Main customer foreign	− 0.002	− 0.002
		(0.055)	(0.020)
	Enterprise network	− 0.006	− 0.006**
		(0.011)	(0.003)
	Environment	0.060**	0.044***
		(0.030)	(0.010)
	Own local HQ	− 0.228***	− 0.115***
		(0.007)	(0.019)
	Male manager	− 0.053***	− 0.038***
		(0.010)	(0.004)
	Log workers	0.354***	0.260***
		(0.022)	(0.010)
	Log age	0.107***	0.049***
		(0.013)	(0.009)
	$(\text{Log age})^2$	− 0.021***	− 0.009***
		(0.004)	(0.002)
	Log capital	− 0.238***	− 0.216***
		(0.004)	(0.002)
	Log material	0.055***	0.060***
		(0.003)	(0.001)
	$(\text{Log workers})^2$	− 0.008**	
		(0.003)	
	Constant	0.336**	0.801***
		(0.168)	(0.097)
	Resid. Log R&D inten.[a]	− 0.097***	− 0.056***
		(0.023)	(0.007)
	Resid. Log training inten.[a]	− 0.245***	− 0.139***
		(0.032)	(0.016)
	Resid. ICT use[a]	0.025	0.050**
		(0.071)	(0.022)
	Resid. Translog TFP[b]		0.489***
			(0.083)

Table 5 (continued)

Variables	(1) Log markup	(2) Log markup
Observations	41,647	41,647
R^2	0.175	0.968

Estimations control for the firm's location (province), industry, and legal form fixed effects. Robust bootstrapped standard errors in parentheses (500 replications). Since the dependent variables are in logs, coefficient estimates for explanatory variables that are in logs are to be interpreted as elasticities. Those for dummy variables have the interpretation of semi-elasticities.

[a]Residual terms from the previously estimated R&D intensity, Training intensity and ICT use equations, respectively. These terms correct for endogeneity of R&D and Training intensities and ICT use in the markup equation.

[b]Residual term from the previously estimated TFP regression (using the *Translog* estimates) and taking into account innovation variables and other controls. This term corrects for endogeneity of TFP in the markup equation.

***$p<0.01$, **$p<0.05$, *$p<0.1$

products, display lower markups than the reference category (Food, beverages and tobacco). There is not a clear pattern of markups as regards legal forms. It seems that the lowest markup is for Non-profit companies and the highest for Private and Local government companies.

Excluding these final groups of dummy variables, in the bottom panel in "Appendix 3" we report the correlation coefficients bwtween the variables in x_{4i} and x_{3i}.

Reflection on Results for Markups

In this section, we discuss the previous section's results from the point of view of convenience for society. Why should it be good that firms investing in drivers of knowledge enjoy higher markups? A textbook perfectly competitive market is characterized by, among other things, the following: 1) No exit or entry barriers; 2) totally homogeneous products, that is products are perfect substitutes for each other (i.e., the qualities and characteristics of products do not vary between suppliers); and, 3) suppliers and consumers are 'price-takers', namely no individual action has any effect on market price. This is rationalized by assuming that each producer and each consumer are 'small' (as regards quantities) relative to the whole market. All firms have a relatively small market share. However, in the real world, there are, for instance, some barriers to entry/exit, some product differentiation and/or some marginal costs asymmetries (see the latest generation of models with heterogeneous producers under a market structure of monopolistic competition, e.g., Melitz 2003). In fact, in this work, we have already presented some evidence of within-industry firms' heterogeneity, since most of the variation in markups between firms' observations comes from within-industry variation (and not so much from between-industry variation).

This paper acknowledges that (similarly to De Loecker and Warzyinski 2012 for exporters vs. non-exporters) the performance of R&D, workers' training or the use of ICT can create or modify two-dimensional sources of within-industry firms' heterogeneity: productivity (costs) and quality of products (prices). Hence, once we eliminate the effect of within-industry productivity dispersion by controlling for productivity in the markups regression, we can shed some light on the role of other factors on prices. This differs from the model of Bernard et al. (2003) in which productivity is the only source of markups variation among firms and, hence, firms with equal productivity charge equal markups.

It is relevant to keep in mind that, according to our markups regressions, we find that firms in a given industry investing in R&D, workers' training or that use ICT, on average, have higher markups than those that do not. If the performance of these activities increases firm-level productivity, as we have also found, the effect of them on markups is not necessarily something explaining an increase in the capacity of firms to set prices above marginal costs. For this to be true, we need to control for productivity in estimation. If not, a firm can enjoy higher markups simply because its marginal cost has fallen and it has no pressure to reduce prices when it coexists in the industry with firms with higher marginal costs. Performers of R&D and workers' training activities or users of ICT in the industry can maintain (because of their cost advantage) higher markups without losing competitiveness and without being forced to reduce prices. In this direction, Melitz and Ottaviano' (2008) model shows that when a firm is relatively more productive it can charge a higher markup and enjoy higher profits.

Nevertheless, the performance of these activities can have a direct effect on markups, in the sense of not working through their effect on productivity. This effect operates through the firm's capacity to set prices above marginal cost as an active strategy. Notice, however, that, in general, the source of this firm's capacity is not innocuous for society. It is not the same when it has its origin in a rigid and strict regulation that creates entry barriers to the industry than when it is explained by some firm's investments and strategies allowing them the production of higher quality products, with higher value added, and/or the access to richer (higher-income) markets. The second source may explain a different pricing behavior of firms performing R&D and workers' training activities or using ICT as regards the ones in the same Ecuadorian industry not performing these activities. In summary, an increase in markups due to the appearance of better products in a product-differentiated market is not the same as an increase in markups of a homogenous product. In the second case, the government should care about the existence of a possible anti-competitive behavior of the industry. However, in the first case, the government should contribute to the diffusion of new knowledge and business practices among firms in the industry (for instance, alleviating sunk costs such as the ones associated with R&D investments).

Under the scenario in which markups increase simply because marginal costs decrease, the government may be required to develop two types of measures. The first one is also about diffusion and access to new knowledge and technologies. According to the second one, the government may consider facilitation of exit of less productive firms from the industry or the reallocation of market shares from less efficient to more efficient producers, increasing the global efficiency of the industry and contributing to aggregate productivity growth.

Although in the short run, firms with new knowledge and technology can have positive economic profits, in the long run, if followers either switch to the new technology or exit the industry, the industry can come back to normal profits. This is somehow a dynamic process of creative destruction that would make higher quality products available to consumers and allow efficient producers to win market share, with the subsequent gain in terms of exploitation of economies of scale in production.

Concluding Remarks

As regards the questions raised in this paper, we can summarize the results as follows. First, the variables included to signal professionalization and good business practices in Ecuadorian manufacturing firms, such as belonging to an enterprise network, having access to finance, performing activities of market research, accountancy, and having environmental concerns, explain both higher propensities to invest in R&D, workers' training , and ICT use, and also higher R&D and workers' training intensities (for ICT, the database does not contain information on spending).

Second, the three considered knowledge creation activities have a positive and relevant effect on firms' TFP, with an estimated elasticity for R&D intensity around 0.08% (in response to a 1% increase in R&D intensity), for workers' training intensity 0.183% (in response to a 1% increase in training intensity), and a semi-elasticity for ICT use that implies that performing this activity justifies around 10.6% higher TFP. Third, they are also relevant to explain higher firms' markups, since the statistically significant estimated elasticity for R&D intensity in the markups regression is around 0.094%, for workers' training intensity 0.255%, and the semi-elasticity for ICT use justifies around 12.5% higher markups.

Additionally, the estimated markups regression that includes among regressors the variable TFP allows discerning whether knowledge creation activities influence markups by affecting efficiency and/or by affecting firms' capacity to set prices. In particular, we find that around half of the effect of R&D and workers' training intensities on markups acts through increasing firms' efficiency, and the other half through higher selling prices. These drivers of knowledge probably generate higher quality products and allow access to better markets. Differently, the ICT effect on markups acts only through efficiency.

Worth mentioning are results for the variables market research and the main customer being foreign. The realization of market research activities by the firm is associated both with a higher propensity to perform R&D and workers' training investments and to a higher intensity in these investments. This may indicate some demand problems requiring innovations. This is reinforced by the fact that market research is associated with lower markups in the markups regressions (demand conditions pressuring prices to go down). The positive association between the main customer being foreign and R&D and workers' training intensities may suggest that learning and competition from international markets encourage firms' innovation efforts.

As regards firms' geographical location, Pichincha is in general outperforming other Ecuadorian provinces in terms of R&D, workers' training, ICT use, TFP, and markups. Furthermore, the Private company is the legal form associated with higher TFP and with higher markups (local government companies are also associated with higher markups). Among industries, Chemicals and petroleum products is the industry that stands out in all analysed dimensions.

It is worth mentioning that ours and in general CDM models intend to deal with problems of selectivity bias and endogeneity (Crespi and Zuniga 2012). The problem of selectivity emerges because not many firms report positive investments in R&D, workers' training, or use of ICT. The problem of endogeneity may appear with cross-sectional data if there are, for instance, time-invariant unobservables affecting productivity and markups that may also affect firms' R&D and workers' training investments or the use of ICT. For this reason, besides other regressors, in the productivity and markups regressions we not only control for industry, province and legal form fixed effects but also use the methodology developed by Rivers and Vuong (1988) to correct and test for possible endogeneity problems that may affect our main interest regressors (R&D and workers' training investments, the use of ICT, and TFP). Hence, we have done our best to control for endogeneity of our central variables in the analysis.

From a policy point of view, the fact that firms belonging to a business group or declaring to have access to finance have both higher likelihoods to perform knowledge creation activities and a higher intensity in their investments, may point to the relevance of easing Ecuadorian manufacturing firms' financial constraints in order to promote these activities. In a developing country such as Ecuador, the distribution of firms' sizes is very much skewed to small sizes (only around 5% of manufacturing firms had more than 10 workers). The small size of Ecuadorian firms is likely a limiting factor for the widespread performance of these activities. Large firms are more able to exploit economies of scale, can be more effective in protecting these investments from imitation, and can better face the uncertainty about their final returns. In addition, the results for the variable Main customer being foreign, although an imperfect proxy for firms' export status, possibly point out to firms that face a larger market size. This variable increases the propensity of ICT use and the intensity in R&D and workers' training investments. A similar market size interpretation could be applied to the variable firm size, which increases the propensity to perform the three activities.

Spreading these activities among more manufacturing firms will promote both firms' productivity and markups and will contribute to counteract the threat of deceleration and slowdown in the Ecuadorian economy, making it stronger against drops in oil prices and U.S. dollar appreciations. This requires comprehensive policies encouraging not only R&D investments but also education and training on-the-job to increase skills, and extending ICT use, which facilitates profiting from global networks of knowledge and global value chains.

The need for these policies is even more pressing today, given the arrival of a possible fourth industrial revolution, embodied in what has been called Industry 4.0. Following UNIDO (2016, p. 6), "the technologies today include artificial intelligence, robotics, the Internet of Things, autonomous vehicles, 3-D printing, nanotechnology, biotechnology, materials science, energy storage, and quantum computing". The arrival of all these technologies into manufacturing will move this sector from physical to digital, with the resulting change of skills that this requires. In this environment, ICT is a "key prerequisite and will help to engage in automation, in Big Data Analytics, in connecting global value chains" (UNIDO 2016, p. 7). However, for the new industrial revolution, it is not enough the widespread access and use of ICT but it also requires turning the focus of the economy, firms and society into innovation and an increase of the capabilities of workers and managers to understand the full potential of new technologies for manufacturing improvements ("Skills are needed to bridge the gap between engineering and computer science, machine learning, and artificial intelligence"; UNIDO 2016, p. 13). The third industrial revolution was already based on digital technology, personal computing and the development of the internet. Developing countries have been mainly working and investing in the physical infrastructure required by those technologies. However, there is still a digital divide with developed countries probably due to insufficient complementary assets in developing countries (such as access to finance, digital and technological skills, size of firms or markets that allow the scalability of investments, etc.). The triplet formed by the public sector, the business sector and universities (the educational system in general) must work closely to prevent the new industrial revolution from exacerbating the digital divide with the most advanced economies.

In this respect, the current National Development Plan for Ecuador 2017–2021 (Senplades 2017) includes as one of the aims of economic policy "To promote research, training, development and transfer of technology, innovation and entrepreneurship, and protection of intellectual property, to encourage the change of the productive matrix through the linkage between the public sector, the productive sector and universities" (policy 5.6, p. 83). This materializes with the establishment of the following goals for the year 2021: (1) increase the ICT Development Index (IDI) from 4.6 (position 101 in the 2016 rank ITU 2017) to 5.6, which is a relevant indicator for the information society that can be used to measure the digital divide across countries (for comparison, the average IDI for Europe in 2016 was 7.35); (2) increase the percentage of R&D expenditure on gross domestic product from 0.44% to 0.48% (according to the Global Innovation Index, WIPO 2017, Ecuador is in position 100 among 128 countries in 2016); (3) increase the number of national patent applications from 78 to 153; and (4) increase the number of scientific publications. In addition, the recent industrial policy plan for Ecuador 2016–2025 (Ministerio de Industrias y Productividad 2016) highlights the necessity of increasing diversification and sophistication of the productive matrix through investments in systemic competitiveness. As an example of this type of investments, the document mentions that the country has increased since 2006 the speed of internet by 40 times and the optic fiber network from 3500 km to around 60,000 km. It has also increased human capital through education and vocational training. In particular, in higher education and training, Ecuador climbed, during the same period, 38 positions in the ranking of the Global Competitiveness Index (World Economic Forum 2015). However, the percentage of workers in the country that receive training is still low, 12.9% in 2012 (vs. 14.2% for Latin America and the Caribbean or 56.3% for OECD). Both institutions (SENPLADES and Ministerio de Industrias y Productividad) recognize how powerful connectivity, innovation and education and training are to improve productivity in the digital age and in a knowledge economy.

From the perspective of micro-, small- and medium-sized enterprises, it is also relevant in the country a further development of knowledge services to support production and systemic competitiveness. These types of services face barriers for growth due to their difficulties in getting credit since they offer intangibles as collateral. The objective of industrial policy in this field is the creation of a virtual tool that facilitates contact between user firms and firms that provide knowledge-intensive services (Ministerio de Industrias y Productividad 2016). Additionally, the Ecuadorian public sector intends to introduce co-financing in the contracting of software or training, and the promotion of crowd funding for micro-, small- and medium-sized enterprises.

It is the time not only of sectoral policies but also of the most transversal ones that help firms in developing countries to compete on a more level field with both the group of countries with high technology/high salary and with high technology/(still) low salary. At this level of competition, knowledge creation activities are crucial, as they contribute to innovation, to the assimilation capacity of available technology and to a possible connection with global networks.

Acknowledgements J.A. Rodríguez-Moreno gratefully acknowledges financial support from the Ecuadorian Secretary of Higher Education, Science, Technology and Innovation, under its Grants Program Open Call 2011; and research funds for the period 2014–2019 from the Universidad Santa María Guayaquil, Ecuador. M.E. Rochina-Barrachina also acknowledges financial support from the Spanish Ministerio de Economía, Industria y Competitividad and the Spanish Agencia Estatal de Investigación (reference number ECO2017-86793-R, co-financed with FEDER funds, European Union).

Compliance with Ethical Standards

Conflict of interest On behalf of all authors, the corresponding author states that there is no conflict of interest.

Appendix 1: Variables Description

Variable	Description
R&D	Dummy variable taking value 1 if the firm does R&D activities and 0 otherwise
Workers' training	Dummy variable taking value 1 if the firm performs training programs for the employees and 0 otherwise
ICT	Dummy variable taking value 1 if the firm uses Information and Communication Technologies, and 0 otherwise.
R&D intensity	Expenditure in R&D per employee
Training intensity	Expenditure in training programs per employee

Variable	Description
Provinces	Dummy variables taking value 1 if the firm is located in a particular province and 0 otherwise. Provinces are: Azuay, Bolívar, Cañar, Carchi, Cotopaxi, Chimborazo, El Oro, Esmeraldas, Guayas, Imbabura, Loja, Los Ríos, Manabí, Morona Santiago, Napo, Pastaza, Pichincha, Tungurahua, Zamora, Galápagos, Sucumbíos, Orellana, Santo Domingo, y Santa Elena
Industries	Dummy variables taking value 1 if the firm belongs to a particular industry and 0 otherwise. Industries are Food, beberages and tobacco; textiles and wearing apparel; leather and footwear; wood, paper and printing; chemicals and petroleum products; rubber and plastics; non-metallic mineral products; metal products; office mach. and elect. equipment; communi./prec./optic./med. equipment; transport equipment; furniture and n.e.c.
Natural person	Dummy variable taking value 1 if the business has a natural person recognition by the National Tax Service, and 0 otherwise
Non-profit company	Dummy variable taking value 1 if the firm is a non-government, non-lucrative organization, and 0 otherwise
Private company	Dummy variable taking value 1 if the firm is a private company and 0 otherwise
Foreign company	Dummy variable taking value 1 if the firm has foreign control and 0 otherwise
Public company	Dummy variable taking value 1 if the firm is under the central government control and 0 otherwise
Local Gov. company	Dummy variable taking value 1 if the firm is under the local government control and 0 otherwise
Cooperative	Dummy variable taking value 1 if the firm is a cooperative and 0 otherwise
Association	Dummy variable taking value 1 if the firm is considered an association and 0 otherwise
Enterprise network	Dummy variable taking value 1 if the firm is a member of an enterprise network or business group, and 0 otherwise
Market research	Dummy variable taking value 1 if the firm does market research and 0 otherwise
Accountancy	Dummy variable taking value 1 if the firm has accounting control and 0 otherwise
Access to finance	Dummy variable taking value 1 if the firm has access to external finance and 0 otherwise
Environment	Dummy variable taking value 1 if the firm does some activity to improve the environment and has environmental concerns, and 0 otherwise
Main customer foreign	Dummy variable taking value 1 if the firm has a foreign main customer and 0 otherwise
Craft certification	Dummy variable taking value 1 if the firm has a Craft Certification and 0 otherwise. It is giving by the government to 'natural persons' (self-employed) who demonstrate long experience (in years) with handmade or technician work (non-professional). The 'natural persons' enjoy a tax benefit with this type of certification
Own local HQ	Dummy variable taking value 1 if the local of the firm is own by the firm, and 0 otherwise
Mother company	Dummy variable taking value 1 if the firm is a mother company and 0 otherwise
Male manager	Dummy variable taking value 1 if the firm manager is a male and 0 otherwise
Log workers	Number of employees of the firm. This variable is in log form
(Log workers)^2	Number of log employees squared
Log age	Number of years since the firm was born. This variable is in log form
(Log age)^2	Log age squared
Log labor productivity	Sales per employee in log form

Variable	Description
Log capital	Stock of tangible fixed assets at book values. This variable is in log form
Log capital/worker	Stock of tangible fixed assets at book values *per* worker. This variable is in log form
Log material	Amount of materials. This variable is in log form
Log materials/worker	Amount of materials *per* worker. This variable is in log form

Appendix 2: Descriptive Statistics

Variable	Obs.	Mean	SD	Mean performers (4938 obs.)	Mean non-performers (37,354 obs.)
R&D	42,292	0.0097	0.0982	0.0834	–
Workers' training	42,292	0.0432	0.2033	0.3701	–
ICT	42,292	0.0987	0.2982	0.8450	–
Log R&D intensity	412	5.1555	1.7565	5.1555	–
Log training intensity	1828	4.3227	1.4649	4.3227	–
Azuay	42,292	0.1026	0.3034	0.1160	0.1008
Bolívar	42,292	0.0080	0.0891	0.0026	0.0087
Cañar	42,292	0.0194	0.1379	0.0101	0.0206
Carchi	42,292	0.0078	0.0883	0.0022	0.0086
Cotopaxi	42,292	0.0294	0.1690	0.0141	0.0314
Chimborazo	42,292	0.0398	0.1955	0.0360	0.0403
El Oro	42,292	0.0377	0.1905	0.0257	0.0393
Esmeraldas	42,292	0.0166	0.1280	0.0068	0.0179
Guayas	42,292	0.1887	0.3913	0.1769	0.1902
Imbabura	42,292	0.0393	0.1944	0.0409	0.0391
Loja	42,292	0.0382	0.1918	0.0269	0.0397
Los Ríos	42,292	0.0292	0.1685	0.0145	0.0312
Manabí	42,292	0.0561	0.2301	0.0348	0.0589
Morona Santiago	42,292	0.0096	0.0977	0.0054	0.0101
Napo	42,292	0.0046	0.0680	0.0028	0.0048
Pastaza	42,292	0.0066	0.0815	0.0038	0.0070
Pichincha	42,292	0.2388	0.4264	0.3772	0.2205
Tungurahua	42,292	0.0605	0.2384	0.0641	0.0600
Zamora	42,292	0.0069	0.0830	0.0016	0.0076
Galápagos	42,292	0.0020	0.0447	0.0022	0.0019
Sucumbíos	42,292	0.0078	0.0879	0.0050	0.0081
Orellana	42,292	0.0050	0.0707	0.0040	0.0051
Santo Domingo	42,292	0.0297	0.1698	0.0214	0.0308
Santa Elena	42,292	0.0139	0.1170	0.0038	0.0152
Food, beverages and tobacco	42,292	0.2211	0.4150	0.1462	0.2310

Variable	Obs.	Mean	SD	Mean performers (4938 obs.)	Mean non-performers (37,354 obs.)
Textiles and wearing apparel	42,292	0.2188	0.4134	0.2069	0.2204
Leather and footwear	42,292	0.0277	0.1642	0.0330	0.0270
Wood, paper and printing	42,292	0.0998	0.2997	0.1761	0.0897
Chemicals and petroleum products	42,292	0.0091	0.0949	0.0449	0.0043
Rubber and plastics	42,292	0.0111	0.1050	0.0469	0.0064
Non-metallic mineral products	42,292	0.0578	0.2334	0.0415	0.0600
Metal products	42,292	0.1655	0.3716	0.1188	0.1717
Office mach. and elect. equipment	42,292	0.0129	0.1130	0.0332	0.0102
Communi./prec./optic./medic. equip.	42,292	0.0053	0.0730	0.0107	0.0046
Transport equipment	42,292	0.0101	0.1002	0.0164	0.0093
Furniture and n.e.c.	42,292	0.1602	0.3668	0.1249	0.1649
Natural persons (self-employed)	42,292	0.9591	0.1980	0.7300	0.9893
Non-profit company	42,292	0.0007	0.0270	0.0028	0.0004
Private company	42,292	0.0364	0.1874	0.2557	0.0074
Foreign company	42,292	0.0001	0.0119	0.0012	0.0000
Public company	42,292	0.0008	0.0299	0.0014	0.0008
Local gov. company	42,292	0.0005	0.0233	0.0004	0.0005
Cooperative	42,292	0.0002	0.0153	0.0016	0.00005
Association	42,292	0.0018	0.0426	0.0064	0.0012
Enterprise network	42,292	0.1914	0.3934	0.5907	0.1386
Market research	42,292	0.0238	0.1526	0.1111	0.0123
Accountancy	42,292	0.0794	0.2704	0.4277	0.0334
Access to finance	42,292	0.2469	0.4312	0.3665	0.2311
Environment	42,292	0.0170	0.1295	0.1091	0.0048
Main customer foreign	42,292	0.0056	0.0744	0.0330	0.0019
Craft certification	42,292	0.2949	0.4560	0.3566	0.2868
Own local HQ	42,292	0.4768	0.4994	0.5028	0.4733
Mother company	42,292	0.0342	0.1819	0.1318	0.0213
Male manager	42,292	0.7422	0.4373	0.7624	0.7396
Log workers	42,292	0.7496	0.8394	1.8234	0.6077
$(\text{Log workers})^2$	42,292	1.2666	3.2160	5.2440	0.7408
Log age	42,292	1.7227	1.0917	2.1795	1.6624
$(\text{Log age})^2$	42,292	4.1600	3.9232	5.7758	3.9464
Log labor productivity	41,665	8.8584	1.0648	9.7154	8.7315
Log capital	41,647	8.1121	1.7503	10.2519	7.8233
Log capital/worker	41,665	7.3610	1.4038	8.4293	7.2161
Log materials	41,647	8.5995	1.6346	10.4658	8.3469
Log materials/worker	41,665	7.8484	1.2371	8.6417	7.7382

Appendix 3: Correlation Coefficients

	Top panel: among the variables in x_{1i} and x_{2i}						
	Enterprise network	Market research	Accountancy	Access to finance	Environment	Main custom. foreign	Craft certification
Enterprise network	1.0000						
Market research	0.1177***	1.0000					
Accountancy	0.2661***	0.1775***	1.0000				
Access to finance	0.0947***	0.0599***	0.0602***	1.0000			
Environment	0.1595***	0.1720***	0.2434***	0.0566***	1.0000		
Main custom. foreign	0.0870***	0.0673***	0.1517***	0.0278***	0.1322***	1.0000	
Craft certification	0.3655***	− 0.0023	− 0.0340***	0.0455***	− 0.0140***	− 0.0060	1.0000
Own local	0.0230***	0.0059	0.0526***	0.0083*	0.0295***	0.0346***	0.0337***
Mother company	0.1565***	0.1026***	0.2383***	0.0600***	0.1367***	0.0836***	0.0352***
Male manager	0.0183***	0.0131***	0.0316***	− 0.0197***	0.0288***	0.0049	0.0203***
Log workers	0.3162***	0.1713***	0.5351***	0.1179***	0.2708***	0.1990***	0.0271***
(Log workers)^2	0.3015***	0.1896***	0.5506***	0.0903***	0.3187***	0.2843***	− 0.0280***
Log age	0.2246***	0.0366***	0.1682***	0.0159***	0.0825***	0.0633***	0.2103***
(Log age)^2	0.2192***	0.0426***	0.1740***	− 0.0025	0.0907***	0.0733***	0.2004***

	Top panel: among the variables in x_{1i} and x_{2i}						
	Own local	Mother company	Male manager	Log workers	(Log workers)^2	Log age	(Log age)^2
Enterprise network							
Market research							
Accountancy							
Access to finance							
Environment							
Main custom. foreign							
Craft certification							
Own local	1.0000						
Mother company	0.0047	1.0000					
Male manager	0.0253***	0.0099**	1.0000				
Log workers	0.0730***	0.2785***	0.0622***	1.0000			
(Log workers)^2	0.0849***	0.3124***	0.0579***	0.8588***	1.0000		
Log age	0.2478***	0.1051***	0.0920***	0.2142***	0.2081***	1.0000	
(Log age)^2	0.2502***	0.1091***	0.0932***	0.2160***	0.2295***	0.9491***	1.0000

	Bottom panel. Among the variables in x_{3i} and x_{4i}							
	Log R&D intensity	Log training intensity	ICT use	Log labor productivity	Translog TFP	Market research	Main custom. Foreign	Log workers
Log R&D intensity	1.0000							
Log Training intensity	0.3013***	1.0000						
ICT use	0.2210***	0.3391***	1.0000					

	Bottom panel. Among the variables in x_{3i} and x_{4i}							
	Log R&D intensity	Log training intensity	ICT use	Log labor productivity	Translog TFP	Market research	Main custom. Foreign	Log workers
Log labor productivity	0.1289***	0.2079***	0.3006***	1.0000				
Translog TFP	0.0158***	0.0304***	0.0660***	0.6056***	1.0000			
Market research	0.2274***	0.2217***	0.1878***	0.1067***	0.0173***	1.0000		
Main custom. foreign	0.0979***	0.1156***	0.1402***	0.0942***	0.0061	0.0673***	1.0000	
Log workers	0.1984***	0.2945***	0.4823***	0.2540***	– 0.0008	0.1713***	0.1990***	1.0000
(Log workers)^2	0.2396***	0.3304***	0.4720***	0.2746***	0.0002	0.1896***	0.2843***	0.8588***
Log age	0.0556***	0.0959***	0.1554***	0.0855***	0.0219***	0.0366***	0.0633***	0.2142***
(Log age)^2	0.0596***	0.0974***	0.1553***	0.0667***	0.0078	0.0426***	0.0733***	0.2160***
Log capital per worker	0.1163***	0.1873***	0.2744***	0.4392***	– 0.0036	0.0881***	0.0729***	0.1660***
Log mater. per worker	0.1035***	0.1736***	0.2407***	0.7613***	– 0.0012	0.0887***	0.0790***	0.2081***
Log capital	0.1885***	0.2918***	0.4512***	0.4754***	– 0.0033	0.1528***	0.1543***	0.6117***
Log material	0.1805***	0.2832***	0.4301***	0.7092***	– 0.0014	0.1555***	0.1623***	0.6704***

	Bottom panel. Among the variables in x_{3i} and x_{4i}						
	(Log workers)^2	Log age	(Log age)^2	Log capital per worker	Log mater. per worker	Log capital	Log material
Log R&D intensity							
Log Training intensity							
ICT use							

	Bottom panel. Among the variables in x_{3i} and x_{4i}						
	(Log workers)^2	Log age	(Log age)^2	Log capital per worker	Log mater. per worker	Log capital	Log material
Log labor productivity							
Translog TFP							
Market research							
Main custom. foreign							
Log workers							
(Log workers)^2	1.0000						
Log age	0.2081***	1.0000					
(Log age)^2	0.2295***	0.9491***	1.0000				
Log capital per worker	0.2057***	0.1147***	0.1026***	1.0000			
Log mater. per worker	0.2224***	0.0482***	0.0349***	0.3864***	1.0000		
Log capital	0.5758***	0.1946***	0.1859***	0.8817***	0.4095***	1.0000	
Log material	0.6089***	0.1461***	0.1372***	0.3780***	0.8653***	0.6249***	1.0000

***$p<0.01$; **$p<0.05$; *$p<0.10$

Appendix 4: Estimated Industry-Specific Input Elasticities from Translog and Cobb–Douglas Production Functions

Industry	Materials		Labor		Capital	
	Translog (β_m)	Cobb–Douglas (β_m)	Translog (β_l)	Cobb–Douglas (β_l)	Translog (β_k)	Cobb–Douglas (β_k)
Food, beverages and tobacco	0.62***	0.60***	0.32***	0.33***	0.14***	0.16***
Textiles and wearing apparel	0.51***	0.52***	0.35***	0.41***	0.12***	0.12***
Leather and footwear	0.62***	0.62***	0.36**	0.35***	0.10***	0.10***
Wood, paper and printing	0.57***	0.59***	0.36***	0.37***	0.15***	0.14***
Chemicals and petroleum product.	0.61***	0.66***	0.42**	0.31***	0.15***	0.18***
Rubber and plastics	0.54***	0.53***	0.49***	0.52***	0.14*	0.12***
Non-metallic mineral products	0.62***	0.64***	0.33***	0.36***	0.06*	0.07***
Metal products	0.58***	0.58***	0.41***	0.45***	0.09**	0.10***
Office machin. and electrical equip.	0.56***	0.62***	0.42***	0.32***	0.16**	0.18***
Communi./prec./optic./medic. equip.	0.52***	0.56***	0.41*	0.41***	0.07**	0.08**
Transport equipment	0.51***	0.54***	0.50***	0.49***	0.08***	0.09***
Furniture and n.e.c.	0.61***	0.61***	0.37***	0.38***	0.08***	0.08***
Total[a]	0.58***	0.59***	0.36***	0.39***	0.11***	0.12***

Cobb–Douglas estimates are common to all firms belonging to the same industry. However, Translog input elasticities are firm specific and, hence, the estimated values presented in the table correspond to the calculated average of estimated input elasticities for all firms belonging to the same industry.

[a]The Total estimates presented in the table are the ones obtained by estimation of a unique production function pooling all industries observations and controlling for industry fixed effects.

$^{***}p<0.01$, $^{**}p<0.05$, $^{*}p<0.1$

References

Aboal, D., and E. Tacsir. 2018. Innovation and productivity in services and manufacturing: The role of ICT. *Industrial and Corporate Change* 27 (2): 221–241.

Añón-Higón, D. 2012. The impact of ICT on innovative activites: Evidence for UK SMEs. *International Small Business Journal* 30: 684–699.

Arvanitis, S., and E.N. Loukis. 2009. Information and communication technologies, human capital, workplace organization and labor productivity: A comparative study based on firm-level data for Greece and Switzerland. *Information Economics and Policy* 21: 43–61.

Aw, B.Y., M.J. Roberts, and T. Winston. 2007. Export market participation, investments in R&D and worker training, and the evolution of firm productivity. *The World Economy* 30 (1): 83–104.

Aw, B.Y., M.J. Roberts, and D.Y. Xu. 2011. R&D investment, exporting and productivity dynamics. *American Economic Review* 101: 1313–1344.

Baldwin, R. 2016. *The great convergence: Information technology and the new globalization*. Cambridge, MA: Harvard University Press.

Benavente, J.M. 2006. The role of research and innovation in promoting productivity in Chile. *Economics of Innovation and New Technology* 55: 301–315.

Bernard, A.B., J. Eaton, J.B. Jensen, and S. Kortum. 2003. Plants and productivity in international trade. *American Economic Review* 93: 1268–1290.

Black, S.E., and L.M. Lynch. 2001. How to compete: The impact of workplace practices and information technology on productivity. *Review of Economics and Statistics* 83: 434–445.

Bloom, N., M. Darca, T. Kretschmer, R. Sadun, and J. Van Reenen. 2010. *The economic impact of ICT*. LSE: Center for Economic Performance.

Bratti, M., and G. Felice. 2012. Are exporters more likely to introduce product innovation? *The World Economy* 35: 1559–1598.

Bresnahan, T.F., and M. Trajtenberg. 1995. General purpose technologies 'Engines of growth'? *Journal of Econometrics* 65: 83–108.

Bresnahan, T.F., E. Brynjolfsson, and L.M. Hitt. 2002. Information technololgy, workplace organization, and the demand for skilled labor: Firm-level evidence. *The Quarterly Journal of Economics* 117: 339–376.

Brynjolfsson, E., and L.M. Hitt. 2003. Computing productivity: Firm-level evidence. *The Review of Economics and Statistics* 85 (4): 793–808.

Castiglione, C. 2012. Technical efficiency and ICT investment in Italian manufacturing firms. *Applied Economics* 44: 1749–1763.

Cohen, W.M., and D.A. Levinthal. 1990. Absorptive capacity: A new perspective on learning and innovation. *Administrative Science Quarterly* 3: 128–152.

Colombo, E., and L. Stanca. 2014. The impact of training on productivity: Evidence from a panel of Italian firms. *International Journal of Manpower* 35: 1140–1158.

Commander, S., R. Harrison, and N. Menezes-Filho. 2011. ICT and productivity in developing countries: New firm-level evidence from Brazil and India. *Review of Economics and Statistics* 93: 528–541.

Corchón, L.C., and L. Moreno. 2010. *Interpreting markups in Spanish manufacturing: The exponential model*. MPRA paper 26609. Munich: University Library of Munich.

Corrado, C., C. Hulten, and D. Sichel. 2005. Intangible capital and economic growth measuring capital and technology: An expanded framework. In *Measuring capital in the new economy, studies in income and wealth*, vol. 65, ed. C. Corrado, C. Hulten, and D. Sichel. Chicago: University of Chicago Press.

Crépon, B., E. Duguet, and J. Mairesse. 1998. Research and development, innovation and productivity: An econometric analysis at the firm level. *Economics of Innovation and New Technology* 7 (2): 115–158.

Crespi, G., and P. Zuniga. 2012. Innovation and productivity: Evidence from six Latin American countries. *World Development* 40 (2): 273–290.

De Loecker, J., and F. Warzyinski. 2012. Markups and firm-level export status. *American Economic Review* 102: 2437–2471.

Greenan, N., A. Topiol-Bensaid, and J. Mairesse. 2001. Information technology and research and development impacts on productivity and skills: Looking for correlations on French firm-level data. In *Information technology, productivity, and economic growth*, ed. M. Pohjola. Oxford: Oxford University Press.

Greene, W.H. 2003. *Econometric analysis*, 5th ed. New York: Prentice Hall.

Griffith, R., E. Huergo, J. Mairesse, and B. Peters. 2006. Innovation and productivity across four European countries. *Oxford Review of Economic Policy* 22 (4): 483–498.
Hall, R.E. 1988. The relation between price and marginal cost in U.S. industry. *Journal of Political Economy* 96 (5): 921–947.
Hall, B. 2011. Innovation and productivity. *Nordic Economic Policy Review* 2: 165–203.
Hall, B.H., F. Lotti, and J. Mairesse. 2013. Evidence on the impact of R&D and ICT investments on innovation and productivity in Italian firms. *Economics of Innovation and New Technology* 22: 300–328.
Heckman, J. 1979. Sample selection bias as a specification error. *Econometrica* 47: 153–161.
INEC. 2010. *Economic census of Ecuador*. Quito: INEC.
ITU. 2017. Measuring the information society, vol. 1. Geneva: International Telecommunication Union. https://www.itu.int/en/ITU-D/Statistics/Pages/publications/mis2017.aspx. Accessed 20 Aug 2018.
Janz, N., H. Loof, and B. Peters. 2004. Innovation and productivity in German and Swedish manufacturing firms: Is there a common story? *Problems and Perspectives in Management* 2: 184–204.
Konings, J., and S. Vanormelingen. 2015. The impact of training on productivity and wages: Firm-level evidence. *Review of Economics and Statistics* 97: 485–497.
Levinsohn, J., and A. Petrin. 2003. Estimating production functions using inputs to control for unobservables. *Review of Economic Studies* 70: 317–342.
Melitz, M. 2003. The impact of trade on intra-industry reallocations and aggregate industry productivity. *Econometrica* 71: 1695–1725.
Melitz, M., and G. Ottaviano. 2008. Market size, trade and productivity. *Review of Economic Studies* 75: 295–316.
Ministerio de Industrias y Productividad. 2016. *Política Industrial del Ecuador 2016-2025*. https://www.industrias.gob.ec/wp-content/uploads/2017/01/politicaIndustrialweb-16-dic-16-baja.pdf. Accessed 20 Aug 2018.
Olley, G.S., and A. Pakes. 1996. The dynamics of productivity in the telecommunications equipment industry. *Econometrica* 64: 1263–1297.
Polder, M., G. van Leeuwen, P. Mohnen, and W. Raymond. 2010. Product, process and organizational innovation: Drivers, complementarity and productivity effects. UNU-MERIT Working Paper Series 035.
Rincón, A., M. Vecchi, and F.Venturini. 2013. ICT as a general purpose technology: Spillovers, absorptive capacity and productivity performance. Discussion Paper 416. London: NIESR.
Rivers, D., and Q.H. Vuong. 1988. Limited information estimators and exogeneity tests for simultaneous probit models. *Journal of Econometrics* 39: 347–366.
Roodman, D. 2011. Fitting fully observed recursive mixed-process models with cmp. *Stata Journal* 11: 159–206.
SENPLADES. 2017. Plan Nacional de Desarrollo 2017-2021. Quito-Ecuador. http://www.planificacion.gob.ec/wp-content/uploads/downloads/2017/10/PNBV-26-OCT-FINAL_0K.compressed1.pdf. Accessed 20 Aug 2018.
UNIDO. 2016. Industry 4.0 opportunities and challenges of the new industrial revolution for developing countries and economies in transition (Panel discussion). https://www.unido.org/sites/default/files/2017-01/Unido_industry-4_NEW_0.pdf. Accessed 20 Aug 2018.
Van Beveren, I. 2012. Total factor productivity estimation: A practical review. *Journal of Economic Surveys* 26 (1): 98–128.
Van Leeuwen, G., and H. van der Wiel. 2003. *Do ICT spillovers matter? Evidence from Dutch firm-level data*, 26. The Hague: CPB Netherlands Bureau for Economic Policy Analysis.
WIPO. 2017. *The Global Innovation Index 2017: Innovation feeding the world*. Geneva: WIPO. http://www.wipo.int/publications/es/details.jsp?id=4193. Accessed 20 Aug 2018.
Wooldridge, J.M. 2009. On estimating firm-level production function using proxy variables to control for unobservables. *Economics Letters* 104: 112–114.
Wooldridge, J.M. 2010. *Econometric analysis of cross section and panel data*. Cambridge, MA: MIT Press.
World Economic Forum. 2015. The global competitiveness report 2015-2016. Geneva: World Economic Forum. https://www.weforum.org/reports/global-competitiveness-report-2015. Accessed 20 Aug 2018.
Zwick, T. 2006. The impact of training intensity on establishment productivity. *Industrial Relations: A Journal of Economy and Society* 45: 26–46.