estructuraeconòmica
ECONOMÍA APLICADA II

Innovation and productivity: evidence for 4 Latin American countries manufacturing industry

by

M. Constanza Demmel
Juan A. Máñez
María E. Rochina-Barrachina
and
Juan A. Sanchis-Llopis

March 2013

Working Papers in Applied Economics

WPAE-2013-07

ECONOMÍA APLICADA II

VNIVERSITAT DE VALÈNCIA

Departamento de Estructura Económica (Economía Aplicada II), Facultad de Economía, Universitat de València
Avda dels Tarongers, s/n, 46022 – Valencia, Spain
Phone +34 96 382 83 49 / Fax +34 96 382 83 54
http://www.estructuraeconomica.es
e-mail: investigacion.estructura@uv.es

ISSN 2172-3036

INNOVATION AND PRODUCTIVITY: EVIDENCE FOR 4 LATIN AMERICAN COUNTRIES MANUFACTURING INDUSTRY

M. Constanza Demmel
Juan A. Máñez
María E. Rochina-Barrachina
Juan A. Sanchis-Llopis

Universitat de València and ERI-CES[*]

Abstract

The literature on firm level productivity in developed countries recognizes the important role played by firm innovation activities on the evolution of firm productivity. However, the literature on this topic for developing and emerging economies is scarcer and far from conclusive. The aim of this paper is to study the innovation-productivity link at the firm level for four Latin American countries (Argentina, Mexico, Colombia and Peru) for the manufacturing sector. The paper distinguishes between different innovations types such as process and product innovations. The data used have been drawn from the World Bank panel Enterprise Surveys, which provides data for these countries for the years 2006 and 2010. Our estimation strategy follows two-steps: first, we estimate TFP measures following De Loecker (2010) approach and Wooldridge (2009) estimation procedure (this allows us to compare results both considering an exogenous or endogenous Markov process for the dynamics of productivity); and, second, we use the estimated TFPs as dependent variables in several models with innovation activities as covariates, in order to disentangle the effects of those variables on the TFP. From our results we confirm that the most productive firms self-select into innovation activities under the endogenous Markov, being this result driven by product innovations, and only appearing for Argentina and Mexico. Further, we obtain that there are returns to innovation in terms of productivity for all innovation types, both under an exogenous or endogenous Markov process but, again, only for Argentina and Mexico.

Key words: process and product innovation, Total Factor Productivity (TFP), GMM, endogenous Markov

[*] We acknowledge financial support from Ministerio de Ciencia e Innovación (projects ECO2011-25033 and ECO2011-30323-C03-02) and Generalitat Valenciana (project PROMETEO/068).

JEL Classification: O12, O31, D24

1. Introduction.

The ability of a country and its enterprises to generate and assimilate technological change is a key factor for development. Schumpeter (1943) already identified a strong relationship between innovation and productivity and Griliches (1986) developed a model that embodied the impact of knowledge capital on productivity, providing formalization for empirical practitioners. The fact that innovating is crucial for firm' performance is well established in developed countries. There are numerous studies that show this positive correlation between the innovative activity and the level of development of a country. However, innovation activities are not limited to the developed world and it is rather a widely globalised phenomenon.[1]

Schumpeter (1943) established a clear difference between inventions (ideas about new processes or products), innovations (first successful commercial introduction of inventions in the market) and technological diffusion (transmission of the innovation). Innovation was related to innovators who were able to create new processes or products not available to the market. However, the concept of innovation nowadays can be understood under a broader perspective, in which the new process or product not only should be considered as a radical technological invention, but also a new way of producing goods or the ability of companies to adapt existing technologies to the environment in which they operate (incremental innovations). In this broader perspective, innovation may be as relevant for developing countries as for developed ones, Fagerberg *et al.* (2009). As put forward by Navarro *et al.* (2010), the acquisition, adaptation and creation of knowledge is a

[1] The relevance of the innovation process in firms doing business in developing countries is not always properly acknowledged, Chudnovsky *et al.* (2006).

major factor in economic development and is the common denominator in the successful development strategies followed by many countries (either rich economies or developing countries).[2] Innovation is a fundamental activity in today's world, not only for rich countries but also for less developed countries as it makes possible to implement more effective catching-up strategies by adapting knowledge devised in advanced economies. In the catching-up process, concepts like technological capability (Kim, 1980), absorptive capacity (Cohen and Levintal, 1990) and innovation system (Lundvall, 1992, Nelson, 1993, Edquist, 1997) turn out to be relevant, and many studies have focused on these aspects of development (see Fagerberg and Godinho, 2004, for a survey).

The working hypothesis in Navarro *et al.* (2010) is that the stagnation of productivity[3] in Latin America can be traced, in part, to an innovation deficit, which motivates researching the links between innovation and productivity in those economies. However, there are not many studies that analyse the relevance of undertaking innovation activities and its relation to productivity in countries of Latin America, being one of the main reasons the lack of data at the firm level and, above all, the lack of panel data. This paper attempts to fill this gap in the literature using a panel drawn form the Enterprise Surveys of the World Bank for four Latin American countries (Argentina, Colombia, Mexico and Peru) for the years 2006 and 2010.

As regards the relationship between innovation and productivity, there is a large

[2] Raffo *et al.* (2008) indicate that there is some evidence that developing countries, particularly Latin American countries, are now capable of transforming innovation on productivity. Although this evidence is quite heterogeneous due to the diverse indicators used.

[3] Daude (2010) points at TFP in Latin American countries as the reason why the region is not catching-up with developed countries, claiming that these economies should absorb and create more knowledge.

tradition within the Industrial Organization literature that studies the direction of causality between innovation activities and productivity. Recent findings suggest the existence of a double causal relationship. On the one hand, Mañez *et al*. (2005), Rochina-Barrachina *et al.* (2010) and Mañez *et al.* (2009a) find evidence in favour of self-selection of the most productive firms into innovation activities.[4] On the other hand, there is a large stream that analyses the impact of innovation activities on productivity using firm level data (see Griliches, 2000, for a survey). Although most of the studies focus on the effects of R&D investments on productivity, recent availability of data has made possible to analyse the impact of innovation output on productivity. Thus, Crépon and Duget (1998), Verspagen (1999), Gu and Tang (2004), Huergo and Jaumandreu (2004), Parisi *et al*. (2006), Rochina-Barrachina *et al*. (2010), and Máñez *et al*. (2009a), among others, considering direct measures of innovation output (such as patents, products or process innovations), find that innovations foster productivity growth.

Hall (2011) provides a survey of papers on productivity and innovation for different countries (most of them being developed countries), based the majority of them on data from the Community Innovation Survey (CIS). Given that most of these works use cross-sectional data, the estimation approach used is generally the well-known approach of Crepon-Duguet-Mairesse (CDM).[5] The empirical evidence collected in Hall (2011)

[4] Some support for the self-selection hypothesis into R&D activities (formal part of innovation activities) of the most productive firms can be found, among others, in Hall (2002), who uses a financial constraint argument, González and Jaumandreu (1998), González *et al*. (1999) and Máñez *et al.* (2005).

[5] Crepon-Duguet-Mairesse approach consists of a sequential relationship between R&D, innovation and productivity (measured as sales per employee, i.e., labour productivity).

concludes that product innovations have an economic impact on productivity, while the impact of process innovations is more ambiguous.

However, the existing empirical evidence for Latin American companies is not as conclusive. Crespi and Zuniga (2012) underline that some of the shortcomings in the results obtained are due to differences in the surveys and the sampling methodology that are not harmonized. Raffo *et al.* (2008) find a significant impact of product innovation on productivity in Brazil and Mexico, but not for Argentina (with the CDM approach). With the same method, Crespi and Zuniga (2012) found a strong positive relationship between innovation and productivity in five of the six Latin American countries studied (Argentina, Chile, Colombia, Costa Rica, Panama and Uruguay). Goedhuys (2007a), based on data from the PICS (Productivity and Investment Climate Survey) of the World Bank, however, fails to confirm, through the application of two stages least squares, any significant relationship between innovation and productivity in Brazil. Benavente (2006) also found a significant effect of innovation on productivity for Chilean companies using the CDM approach.

Navarro *et al.* (2010) analyze several works on innovation using the CDM approach, for countries of Latin America (Argentina, Chile, Colombia and Uruguay), and conclude that there is a clear positive impact of product innovation on productivity, except for Chile. On the other hand, process innovation does not seem to have a significant effect on productivity, with the exception of Uruguay. The authors explain that possibly this is due to the fact that in Latin American economies the learning process implicit in the adoption of a new technology needs some time to start having an effect.[6]

The purpose of this paper is to analyse in depth the effect of different firm innovating activities (either process or product or both) on total factor productivity (TFP). For this purpose we will carry out a four-stage strategy. First, we will estimate TFP using a procedure that introduces some novelties that will be describe later on. In a second step we will consider if ex-ante more productive firms are those engaged in product or process (or both) innovations (self-selection hypothesis). Third, we will study whether firms' process or product (or both) innovative strategies causes subsequent higher productivity (returns to innovation hypothesis). Finally, we will study if there is a contemporaneous correlation between firms' innovation strategies and productivity (contemporaneous correlation hypothesis).

To estimate TFP we follow De Loecker (2010) approach and Wooldridge (2009) estimation procedure. The TFP estimation procedure used in this study presents various novelties with respect to the typical control-function based estimation methods used to analyse the effects of, for instance, trade policy, i.e. Olley and Pakes, 1996 (OP, hereafter) and Levinshon and Petrin, 2003 (LP, hereafter). First, we allow for different demands of intermediate materials according to firm innovation strategies (process innovator, product

[6] The lack of significance of some types of innovation on productivity occurs not only for Latin America. For example, Lee and Kang (2007), with data for Korea, find a positive effect of process innovation on productivity, but not for product innovation. Goedhuys (2007b) and Goedhyus *et al.* (2008), using data from the PICS for several developing countries (Brazil, Ecuador, South Africa, Tanzania and Bangladesh), fail to confirm any significant impact of innovation on productivity.

innovator or both process and product innovator). Second, we move from an exogenous law of motion of productivity to an endogenous law of motion in which we allow past innovations (process, product or both) to affect productivity (following De Loecker, 2007, 2010).

From our results in this paper we confirm the self-selection hypothesis of the most productive firms into innovation under the endogenous Markov process, driven by product innovations, and only for Argentina and Mexico. Further, we obtain that there are returns to innovation in terms of productivity for all innovation types, both under an exogenous or endogenous Markov but, again, only for Argentina and Mexico. Finally, both the self-selection and the returns to innovation links with productivity working together generate, as expected, cotemporaneous correlation between all innovation activities and productivity for Argentina and Mexico.

The rest of the paper is organized as follows. Section 2 summarises related literature and poses our research hypothesis. Section 3 describes the data. Section 4 explains the main features pursued by our production function estimation method. In section 5 we discuss the results. And, finally, section 6 concludes.

2. Related Literature and hypotheses.

In this paper we aim to analyse the effect of different firm innovation strategies (either implementing process innovations, product innovations or both) on TFP. For this purpose we will organise our research by steps. First, we analyse whether more productive firms self-select into implementing innovation activities. Second, we check if there are returns from innovation for those firms that engage in innovation.

On theoretical and empirical grounds, it is expected the productivity level of product/process innovators not to be lower than that of non-innovators. This is so as the expenditures associated to innovation activities (both formal and informal) limit the access to innovation activities to the most productive firms. This is especially relevant for the case of R&D investments due to its sunk costs nature (Sutton 1991; Máñez *et al*. 2009b). Further, performing innovation activities is likely to be a previous condition to obtain innovations (the so called *self selection hypothesis*). Some support for the self-selection hypothesis into R&D activities (formal part of innovation activities) of the most productive firms can be found, among others, in Hall 1990 (who uses a financial constraint argument), González and Jaumandreu (1998), González *et al*. (1999) and Máñez *et al.* (2005). On the basis of these arguments, we formulate our first hypothesis:

Hypothesis 1: The most productive firm self-selection into implementing innovations.

According to this hypothesis, and from an applied perspective, we expect those firms that introduce innovations in period *t* to be already more productive in period *t*-1.

Furthermore, there are theoretical arguments that support the existence of returns to innovation associated to both process and product innovations. As regards process innovations, there are, at least, three strands in the literature supporting a positive relationship between the introduction of process innovations and firms' productivity growth. The first strand is based on the well-known *R&D capital stock model* of Griliches (1979) that analyses the relationship among R&D investments, achievement of innovations and productivity growth. Since this seminal work, other authors have incorporated more explicitly the role of process innovations on productivity growth. For instance, Klette and

Johansen (1998) incorporated the output elasticity of knowledge capital to point out the opportunity for process innovations, and Smolny (1998) assumed that process innovations reduce production costs by increasing the productivity of labour and/or capital. The second strand in the literature rendering theoretical support to the relationship between process innovations and productivity growth is the *active learning model* (Ericson and Pakes 1992, 1995; Pakes and Ericson 1998). According to this model, R&D investments, if successful, contribute to improve firm' productivity over time. If the successful output of firms R&D activities is the production of a process innovation and this is actually implemented, we expect an increase in the productivity growth of such a firm. Therefore, in the *active learning model* the relationship between R&D activities and productivity growth runs through the achievement and implementation of process innovations. Finally, the endogenous growth theory is the third strand of the literature stressing the importance of innovations that develop new processes for productivity growth (see, e.g., Romer 1990; Aghion and Howitt 1992).

As for product innovations, Bernard *et al.* (2007) relate firms' performance to their ability to create products. Product innovations can increase firms' sales and/or can be related to product quality and, therefore, to productivity. An early paper by Vernon (1966) develops a product lifecycle theory where product innovation should have an impact on firm productivity. A study by Foster et al. (2008) provides some evidence about firm's product innovation to positively affect firm productivity through positive firm-specific demand variations. This paper finds that firm-specific demand shocks seem to account for differences in measured firm productivity. Therefore, not only production efficiency (more related to process innovation) but also demand driven shocks (more related to product innovation) are expected to have effects on productivity.

In line with the above theoretical arguments, we formulate the following hypothesis related to firm innovation returns:

Hypothesis 2: Past innovations should have a positive effect on firm productivity.

From an empirical perspective this hypothesis implies testing whether firms that introduce innovations in period *t-1* are more productive in period *t*.

Finally, the joint consideration both of self-selection and the possibility of past innovations affecting positively productivity (i.e. the joint consideration of hypothesis 1 and 2) lead us to hypothesize that in a given period firms that introduce innovations should be more productive that firms that do not do it. Being this higher productivity the result of a combination of a process of self-selection and/or returns on productivity of past innovations. Therefore, our third hypothesis is as follows:

Hypothesis 3: For a given period t innovative firms are more productive than non-innovative ones.

3. The data.

The data we use in this research has been drawn from the World Bank *Enterprise Surveys* (ES)[7]. The ES collects data from key manufacturing and service sectors in every region of the world. The surveys use standardized survey instruments and a uniform sampling

[7] For accessing these data check http://www.enterprisesurveys.org

methodology to minimize measurement error and to yield data that are comparable across countries. Most importantly, the ES are designed to provide panel data sets, as panel data is one of the best ways to pinpoint how and which of the changes in the business environment affect firm level productivity.[8] This unique source of information can be used to measure TFP across a set of developing countries.

To produce internationally comparable data, apart from country specific questions derived from the objectives of the World Bank programmes for each country, the ES contains various data sections homogeneous across countries. These are the *Core* questionnaires, with questions that are asked in all countries for all industries where the survey is implemented, and the *Manufacturing Module* and *Services Module*, with questions that are asked to establishments in the manufacturing and services sectors, respectively. Firms provide qualitative and quantitative information that deals with the characteristics of the business in which they operate, such as labour, sales and supplies, etc.

The sampling procedure used in the ES aims to build samples with an appropriate size to provide information on all firms in an economy, not only manufacturing, and to have sample sizes large enough in the industries considered to carry robust statistical analysis. Therefore, the universe of industries has been stratified by region, size and industry. The number of manufacturing sectors in each country is related to the country GDP.[9]

We have a sample of 1896 observations from two years (2006 and 2010) for four Latin-American countries: Argentina, Colombia, Mexico and Peru. We have chosen these Latin-

[8] Although the ES provides establishment level data, we refer to establishment as firms.

[9] Detailed information on the methodology followed to build the ES datasets is available from the World Bank at in the http://www.enterprisesurveys.org.

American countries as they provide data from the ES for the same years[10]. In this study some industries are joined to achieve a sufficiently large number of observations (see Table 1). Following Saliola and Seket (2011) industries are grouped together in six major industry groups (Food, textiles, Garments, Chemicals, Non-metallic and basic metals, Fabricated metal and machinery) based on similarities in the type of activity and factor intensity. The group "Other Manufacturing" is a residual category that includes observations that are outside the six major industry groups. The concentration of firms in six major industry groups is the result of a sample design, used in most countries, where selected industries were targeted to facilitate industry-level analysis.

We exclude from the sample those firms with missing information in the relevant variables (such as capital, sales, labour or materials) we need to carry out our analysis (we require data both from the year 2006 and 2010). Therefore, after cleansing the data we end up with 463 observations distributed as follow: 157 for Argentina, 133 for Colombia, 88 for Mexico and 85 for Peru.

In order to obtain a TFP measure we estimate a production function in which we measure firm output using firm annual total sales, capital as the net value of machinery, vehicles and equipment (accounting for depreciation), labour as the total compensation of workers including wages, salaries, bonuses and social security, and intermediate materials as the cost of raw materials and intermediate materials. As all monetary values are at local

[10] Bolivia, Panama, Paraguay, Uruguay and Venezuela, are excluded from our analysis due to the lack of enough observations to carry out the TFP estimations.

currencies, we convert them to U.S. dollars and then deflate them by the GDP deflator in U.S. dollars (base year 2005).[11]

To analyse the effect of innovation activities on TFP we have also extracted the information on firm innovation activities using two questions. In particular, we use information about whether a firm has introduced a process and/or a product innovation. We obtain this information from the following questions in the questionnaire: *"Over the last three years, did this establishment introduce any new or significantly improved processes for producing or supplying products (goods or services)?"* and *"Over the last three years, did this establishment introduce any new or significantly improved products (goods or services)?*" From these questions we build a set of dummy variables: one for firms that implement process innovations, one for firms that implement product innovations and one for firms that implement both innovation strategies altogether.

Table 1 reports the proportion of firms implementing either process or product innovations only or both process and product innovations, by industry and country. In general, we can see that the majority of firms implement both process and product innovations (51.10%). We observe that only 22.01% do not implement any innovation activity. By industries, we observe that the *Non metallic and basic metal industry* is the one with the highest proportion of firms implementing only product innovations (20.51%), whereas the *Garments industry* is the one with the highest proportion of firms implementing process innovations (13.08%). Further, the *Chemical industry* seems to be the most innovating one as 82% of firms in that industry implement innovations (either

[11] Exchange rates and deflators are obtained from the *World Bank Development Indicators* database, at http://data.worldbank.org/indicator.

process or product or both). By countries, we observe that Peru is the country with the highest rate of innovating firms (85%) and Mexico the one with the lowest rates (58%).

It is important to notice that the highest rate of innovations observed are in line with the results obtained by other authors on developing countries. Fagerberg *et al.* (2010) present a summary of innovation surveys for several countries and find that the innovation activities in developing countries is quite frequent with high rates of innovation in many of them. Other studies with similar results are Crespi and Zuniga (2012), Dabla-Norris *et al.* (2010) and Waheed (2011). These results for developing countries can be explained as follows: innovation is defined broadly as the introduction of new production processes or product lines, so that it includes firms adopting existing technology (Dabla-Norris *et al.*, 2010). In this sense, innovation is the application of new knowledge in a firm that is not necessarily new for competitors, the market or the rest of the world. Using a Schumpeter perspective, this could be classified as imitation more than as a *de novo* innovation (Fagerberg *et al.*, 2009). Further, it is important to notice how innovation is perceived in Latin-American countries. As put forward by Crespi and Zuniga (2012), innovation is a wide concept and firms consider small changes in their products or services as innovations. Many times, innovation means the adoption of a technology developed in other firms. They also explain that technological innovation in these countries can be considered as incremental more than radical. Imitation and acquisition of external technology allows firms who are far from the technological frontier to reduce the technological gap through learning. This explains why the innovation rates for developing countries are so high as compared to the developed world. [12] Furthermore, the innovation rate is higher in manufacturing than in the rest of the economy, both in developed and developing countries.

Next, we identify some stylized facts about innovative and non-innovative firms using a simple regression analysis (see Table 2). The objective is to explore the relationship between firm innovation strategies (product innovation only, process innovation only or both) and some basic firm characteristics. In particular, we estimate an equation of the form:

$$\log\left(y_{it}\right)=\beta_0+\beta_1 proc_{it}+\beta_2 prod_{it}+\beta_3 both_{it}+\sum_{j=2}^{7}\gamma_j ind_j+\sum_{k=2}^{4}\gamma_k country_k+\gamma_t year_t+e_{it} \quad (1)$$

where the dependent variable y_{it} is alternatively output, capital and intermediate materials per worker, sales and size (as measured by the number of employees). The variables $proc_{it}$, $prod_{it},$ $both_{it}$ capture firm innovation strategies. Thus, $proc_{it}$ is equal to one if the firm i implements process innovations only in t (and zero otherwise), $prod_{it}$ is equal to one if the firm i implements product innovations only in t (and zero otherwise), and $both_{it}$ is equal to one if the firm i implements both product and process innovations in t (and zero otherwise). We also control for size (except for the size regression), industry, country and year dummies.

The differences (in %) between firms undertaking any innovation strategy and non innovators for each of the firm characteristics considered are computed from the corresponding estimated coefficient β_i as 100(exp(β_i)-1). In Table 2, we observe that firms that implement both process and product innovations or product innovations only are

[12] Crespi and Zuniga (2012) find that Costa Rica, Colombia and Argentina have innovation rates above 40%, whereas countries like France, Norway and Japan haver much lower innovation rates.

bigger (both in terms of employment and sales), more intermediate materials intensive, and have higher labour productivity than non-innovators. However, we do not find any significant difference between firms that implement process innovations only and non-innovators.

Finally, we check whether there exist significant differences in the five variables under analysis between process innovators only and firms that introduce both process and product innovations. Our results suggest that firms that introduce both process and product innovations are larger both in terms of sales and employees than product innovators only. However, we do not find any significant difference in labour productivity and intermediate materials per worker (whereas the difference $\beta_2 - \beta_3$ is significant at 1% level for sales and size it is not statistically significant for the rest of variables under consideration), see Table 3.

4. Methodology.

Our empirical strategy consists of two steps. In the first step, we estimate TFP using a procedure that introduces some novelties described later in this section. In the second step, with the aim of testing hypotheses 1, 2 and 3, we use this estimated TFP as dependent variable in a series of reduced form equations that include as covariates firms' innovation status variables and country dummies. In these regression analyses we pool TFP estimates for all countries and use OLS regressions with robust standard errors. In what follows we will organise the results according to our working hypotheses.

In this two-step analysis it is crucial to decide whether or not to include the innovation strategy variables as additional covariates in the TFP estimation. Therefore, we devote the following paragraphs to carry out a detailed reasoning of our choice.

4.1. Some features of the two-step method.

First, analogously to De Loecker (2007) with exports, to capture differences in market structure among firms with different innovation strategies (related to the mode of competition, demand conditions, and exit barriers), we allow for different demands of materials according to firm innovation strategies (*non innovators, process innovators only, product innovators only* and firms that introduce *both process and product innovations*). Thus, we do not include firm innovation strategy variables as other inputs in the production function but only in the demand for intermediate inputs.[13] There are at least three reasons that advice to do so: i) including a vector of innovation strategy dummies as inputs (process innovators only, product innovators only and, process and product innovators dummies) implies to assume, for example, that the impact of introducing a process innovation (without introducing product innovations simultaneously) in productivity is deterministic, i.e. the productivity of all process innovators will only increase by the estimate of the *process innovators only* dummy (the same applies to *product innovators only* or firms that introduce *both process and product innovations*); ii) including innovation strategy variables in the demand for intermediate inputs allows innovation strategies to have a different

[13] In an analogous work but in the field of trade, Van Biesebroeck (2007) and Kasahara and Rodrigue (2008) suggest that the export and import dummies should be treated also as state variables, in particular, as non-deterministic state variables. However, De Loecker (2007, 2010) and De Loecker and Warzyinski (2011), argue that these variables affect the characteristics of the markets where firms operate. Notwithstanding, from an empirical point of view both alternatives conduct to the same end.

impact for firms with different characteristics (capital, intermediate materials, etc.); and, iii) the Cobb-Douglas production function implies that if we include innovation strategy variables as additional inputs, a firm can substitute any input with introducing process/product innovations with a unit elasticity of substitution.

Second, we move from an exogenous law of motion of productivity to an endogenous law of motion in which we allow past innovation experience to affect productivity (following Doraszelski and Jaumandreu, 2010). An exogenous Markov process is only appropriate when productivity shocks are exogenous to the firm but not if future productivity is determined endogenously by firm choices, such as innovation decisions. Therefore, those methods that do not use an endogenous Markov process suffer from an internal inconsistency as do not accommodate endogenous productivity processes in which past innovations may affect productivity.

4.2. Productivity estimation.

We assume that firms produce a good using a Cobb-Douglas technology:

$$y_{it} = \beta_0 + \beta_l l_{it} + \beta_k k_{it} + \beta_m m_{it} + \gamma_j + \omega_{it} + \eta_{it} \tag{2}$$

where y_{it} is the natural log of production of firm i at time t, l_{it} is the natural log of labour, k_{it} is the natural log of capital, m_{it} is the natural log of intermediate materials, and γ_j are industry effects. As for the unobservables, ω_{it} is the productivity (not observed by the econometrician but observable or predictable by firms) and η_{it} is a standard *i.i.d.* error term that is neither observed nor predictable by the firm.

It is also assumed that capital evolves following a certain law of motion that is not directly related to current productivity shocks (i.e. it is a state variable), whereas labour and intermediate materials are inputs that can be adjusted whenever the firm faces a productivity shock (i.e. they are variable factors).[14]

Under these assumptions, OP show how to obtain consistent estimates of the production function coefficients using a semiparametric procedure; see also LP for a closely related estimation strategy. However, here we follow Wooldridge (2009), who argues that both OP and LP estimation methods can be reconsidered as consisting of two equations which can be jointly estimated by GMM: the first equation tackles the problem of endogeneity of the non-dynamic inputs (that is, the variable factors); and, the second equation deals with the issue of the law of motion of productivity. Next we consider each in detail.

Let us start considering first the problem of endogeneity of the non-dynamic inputs. Correlation between labour and intermediate inputs with productivity complicates the estimation of equation (2), because it makes the OLS estimator biased and the fixed-effects and instrumental variables methods generally unreliable (Ackerberg *et al.*, 2006). Both OP and LP methods use a control function approach to solve this problem, by using investment in capital and materials, respectively, to proxy for "unobserved" firm productivity.

---

[14] The law of motion for capital follows a deterministic dynamic process according to which. $k_{it} = (1-\delta)k_{it-1} + I_{it-1}$ Where I_{it-1} is investment in capital for firm *i* in period *t-1*. Thus, it is assumed that the capital the firm uses in period *t* was actually decided in period *t-1* (it takes a full production period for the capital to be ordered, received and installed by the firm before it becomes operative). Labour and materials (unlike capital) are chosen in period *t*, the period they actually get used (and, therefore, they can be a function of ω_{it}). These timing assumptions make them non-dynamic inputs, in the sense that (and again unlike capital) current choices for them have no impact on future choices.

In particular, the OP method assumes that the demand for investment in capital, $i_{it} = i\left(k_{it}, \omega_{it}\right)$, is a function of firms' capital and productivity. To circumvent the problem of firms with zero investment in capital, the LP method uses the demand for materials, $m_{it} = m\left(k_{it}, \omega_{it}\right)$, instead, as a proxy variable to recover "unobserved" firm productivity. Since we follow this last approach, we focus on the demand of materials hereafter.[15]

Therefore, when estimating productivity using these general versions of OP and LP in a sample where some firms do not introduce innovations and others innovate either introducing process innovations only, product innovations only or both, it is assumed that the demand of intermediate materials for the different types of firms according to their innovation strategy is identical. However, as explained in the previous section, heterogeneity in innovation strategies may influence the demand of intermediate inputs. Therefore, analogously to De Loecker (2007, 2010) when analysing the effects of exporting on firms' productivity, and Kasahara and Rodrigue (2008), when analysing the effects of importing on firms' productivity, we consider different demands of intermediate materials for non-innovators, process innovators only, product innovators only and both process and product innovators. Thus, we write the demand of materials as:

$$m_{it} = m_{IS}\left(k_{it}, \omega_{it}\right) \tag{3}$$

where we include the subscript *IS* (innovation strategy) to denote different demands of intermediate inputs for non innovators, process innovators only, product innovators only and firms that introduce both process and product innovations. Also, since the demand of

[15] Both the investment of capital demand function and the demand for intermediate materials are assumed to be strictly increasing in ω_{it} (in the case of the investment of capital this is assumed in the region in which i_{it}>0). That is, conditional on k_{it}, a firm with higher ω_{it} optimally invests more (or demands more materials).

intermediate materials is assumed to be monotonic in productivity, it can be inverted to generate the following inverse demand function for materials:

$$\omega_{it} = h_{IS}\left(k_{it}, m_{it}\right) \tag{4}$$

where h_{IS} is an unknown function of k_{it} and m_{it}. Then, substituting the above expression (4) into the production function (2) we get:

$$y_{it} = \beta_0 + \beta_l l_{it} + \beta_k k_{it} + \beta_m m_{it} + \gamma_j + h_{IS}\left(k_{it}, m_{it}\right) + \eta_{it} \tag{5}$$

Finally, by considering four different demand functions for intermediate materials (for non-innovators, process innovators only, product innovators only, and firms that introduce both process and product innovations), our first estimation equation results in:

$$\begin{aligned} y_{it} = {} & \beta_0 + \beta_l l_{it} + \gamma_j + \\ & 1(ni) H_{ni}\left(k_{it}, m_{it}\right) + 1(proc) H_{proc}\left(k_{it}, m_{it}, proc_{it}\right) + \\ & 1(prod) H_{prod}\left(k_{it}, m_{it}, prod_{it}\right) + 1(both) H_{both}\left(k_{it}, m_{it}, both_{it}\right) + \eta_{it} \end{aligned} \tag{6}$$

where 1(*ni*), 1(*proc*), 1(*prod*) and 1(*both*) are indicator functions that take value one for non innovators (*ni*), process innovators only (*proc*), product innovators only (*prod*) and firms that introduce both process and product innovations (*both*), respectively. Further, the unknown functions *H* in (6) are proxied by second-degree polynomials in their respective arguments.

With the specification in equation 6, the difference in the inverse demand function of firms with different innovation strategies arises not only from differences in the coefficients of k_{it} and m_{it} but also by the fact that each inverse demand function includes a dummy variable capturing the corresponding innovation strategy. This is not equivalent to introduce the set of innovation strategies dummies as additional inputs in the production function, as each one of these dummies is interacted with all the terms k_{it} and m_{it} in its corresponding polynomial. For example, introducing a process innovation only dummy as an input in the production function will cause at least two problems. First, an identification problem, as we will need another estimation step to identify the parameter associated to that variable. Second, introducing a process innovation only dummy in the Cobb-Douglas production function implies that a firm can substitute any input with introducing a process innovations at constant unit elasticity (see De Loecker, 2007, 2010, for similar arguments applied to export dummies).

Notice, however, that we cannot identify β_k *and* β_m from (6). This is achieved by the inclusion of a second estimation equation in the GMM-system that deals with the law of motion of productivity.

The standard OP/LP approaches consider that productivity evolves according to an exogenous Markov process:

$$\omega_{it} = E\left[\omega_{it}\right] + \xi_{it} = f\left(\omega_{it-1}\right) + \xi_{it} \tag{7}$$

where f is an unknown function that relates productivity in t with productivity in t-1 and ξ_{it} is an innovation term uncorrelated by definition with k_{it}. However, this assumption neglects the possibility of previous innovation strategies to affect productivity. Consequently, here we consider a more general (endogenous Markov) process in which previous innovations can influence the dynamics of productivity:

$$\begin{aligned}\omega_{it} &= E\left[\omega_{it} \mid \omega_{it-1}, proc_{it-1}, prod_{it-1}, both_{it-1}\right] + \xi_{it} \\ &= f\left(\omega_{it-1}, proc_{it-1}, prod_{it-1}, both_{it-1}\right) + \xi_{it}\end{aligned} \tag{8}$$

where $proc_{it\text{-}1}$, $prod_{it\text{-}1}$ and $both_{it\text{-}1}$ indicate whether the firm, in period t-1, chose to introduce process innovations only, product innovations only, or both process and product innovations, respectively. Obviously, the reference category is to be a non-innovator.

Let us now rewrite the production function (2) using (8) as:

$$y_{it} = \beta_0 + \beta_l l_{it} + \beta_k k_{it} + \beta_m m_{it} + \gamma_j + f\left(\omega_{it-1}, proc_{it-1}, prod_{it-1}, both_{it-1}\right) + \xi_{it} + \eta_{it} \tag{9}$$

Further, since $\omega_{it-1} = h_{is}\left(k_{it}, m_{it}\right)$, we can rewrite $f\left(\omega_{it-1}, proc_{it-1}, prod_{it-1}, both_{it-1}\right)$ as:

$$\begin{aligned}& f\left(\omega_{it-1}, proc_{it-1}, prod_{it-1}, both_{it-1}\right) \\ &= f\left[h_{IS}\left(k_{it-1}, m_{it-1}\right), proc_{it-1}, prod_{it-1}, both_{it-1}\right] = F_{IS}\left(k_{it-1}, m_{it-1}\right) \\ &= 1(ni)F_{ni}\left(k_{it-1}, m_{it-1}\right) + 1(proc)F_{proc}\left(k_{it-1}, m_{it-1}, proc_{it-1}\right) \\ &\quad + 1(prod)F_{prod}\left(k_{it-1}, m_{it-1} prod_{it-1}\right) + 1(both)F_{both}\left(k_{it-1}, m_{it-1}, both_{it-1}\right)\end{aligned} \tag{10}$$

with F being unknown functions to be proxied by second degree polynomials in their respective arguments.

Lastly, substituting (10) into (9), our second estimation equation is given by:

$$\begin{aligned}& y_{it} = \beta_0 + \beta_l l_{it} + \beta_k k_{it} + \beta_m m_{it} + \gamma_j \\ &+ 1(ni)F_{ni}\left(k_{it-1}, m_{it-1}\right) + 1(prod)F_{prod}\left(k_{it-1}, m_{it-1}, prod_{it-1}\right) \\ &+ 1(proc)F_{prc}\left(k_{it-1}, m_{it-1}, proc_{it-1}\right) + 1(both)F_{both}\left(k_{it-1}, m_{it-1}, both_{it-1}\right) + u_{it}\end{aligned} \tag{11}$$

where $u_{it}=\xi_{it}+\eta_{it}$ is a composed error term.

Wooldridge (2009) proposes to estimate jointly equations (6) and (11) by GMM using the appropriate instruments and moment conditions for each equation. This joint estimation strategy has several advantages: i) it increases efficiency relatively to the two step traditional procedures; ii) it makes unnecessary to do bootstrapping for the calculus of standard errors; and, iii) it solves the problem, pointed out by Ackerberg *et al.* (2007), of identification of the labour coefficient in the separate estimation of equation (6).

In the standard case, in which we assume a unique demand of materials for all firms regardless of their innovations strategy and an exogenous Markov process for the law of motion of productivity, the estimations equations (6) and (11) get simplified to (12) and (13), respectively.

$$y_{it} = \beta_0 + \beta_l l_{it} + \mu_j + H\left(k_{it}, m_{it}\right) + \eta_{it} \tag{12}$$

$$y_{it} = \beta_0 + \beta_l l_{it} + \beta_k k_{it} + \beta_m m_{it} + \mu_j + F\left(k_{it-1}, m_{it-1}\right) + u_{it} \tag{13}$$

Hereafter, we will call the model that considers a unique demand of materials and an exogenous Markov process Model 1. We will call Model 2 to the model that considers four different demands of materials according to the firm innovation strategy (non innovators, process innovators only, process innovators only and firms that implement both process and product innovations), and an endogenous Markov process for the law of motion of productivity. Comparison of the results for hypothesis 1, 2, 3 will allow us to check whether previous tests of the impact of innovation strategies based on the use of the standard OP and LP estimators could be flawed.

This procedure allows us to obtain (both for Models 1 and 2), *per* each one of the 4 countries considered, both the coefficient estimates of the production function (shown in Table 4) and firms' productivity estimates as

$$tfp_{it}^{s} = y_{it} - \beta_l l_{it} - \beta_k k_{it} - \beta_m m_{it} - \gamma_j Industry_j \qquad (14)$$

where tfp_{it}^{s} is the log of the estimated TFP for firm *i* at time *t,* for country *s*.

5. Results.

5.1. Testing hypothesis 1: self-selection.

To shed light on the empirical validity of the self selection hypothesis – namely, that the most productive firms in *t-1* self-select into implementing innovations in *t*– we are going to investigate whether innovative firms in *t* were already more productive in period *t-1* as compared to non-innovative firms in *t*. More formally, we estimate the following reduced form equation model (Specification 1):

$$tfp_{it-1} = \beta_0 + \beta_1 (D_{A,M} \cdot Innovation_t) + \beta_2 (D_{C,P} \cdot Innovation_t) + \beta_3 D_{C,P} + u_{it} \qquad (15)$$

where the dependent variable is TFP in period *t-1*, $Innovation_t$ is a dummy variable taking value one if the firm innovates in *t* (without distinguishing by the type of innovation), and $D_{A,M}$ and $D_{C,P}$ are two dummy variables controlling for data coming from Argentina and Mexico, and from Chile and Peru, respectively.[16] With this regression we aim to detect

[16] As we need for this regression an estimate of the TFP in period *t-1* (2006 in this study), we estimate it using the input estimates (labour, capital and intermediate materials) from the production function in year *t* (2010 in this study) to the inputs in year 2006 (production is also available for this year).

differences between innovators and non-innovators in *t* in terms of productivity one year before innovation, in *t-1*. Given that our estimation method has estimated separately the production function by country, we expect both the level of TFP and the effects of innovation on TFP to be country dependent. For this reason, the innovation variable has been interacted with two country group dummies, one for data corresponding to Argentina and Mexico and another one for data corresponding to Colombia and Peru. Also, a dummy variable controlling for data from Colombia and Peru *versus* data from Argentina and Mexico has been included in estimation without interaction. In exploratory work we have found that this is the more parsimonious way to group countries in order to distinguish differentiated effects between them, as regards the link between TFP and innovation. The results, reported in Table 5, for the model that considers an endogenous Markov process for the law of motion of productivity confirm the self-selection hypothesis for innovative firms in Argentina and Mexico but not for Colombia and Peru (for these countries the coefficients on innovation are non-significant). However, we do not find any evidence of self-selection when productivity is estimated under the hypothesis of an exogenous process for the law of motion of productivity.

To check whether not only innovating or not innovating but also the kind of innovation influences the results of the test of self-selection hypothesis, we estimate Specification 2:

$$\begin{aligned} tfp_{it-1} = \beta_0 &+ \beta_1(D_{A,M} \cdot OnlyProc_t) + \beta_2(D_{A,M} \cdot OnlyProd_t) \\ &+ \beta_3(D_{A,M} \cdot Both_t) + \beta_4(D_{C,P} \cdot OnlyProc_t) \\ &+ \beta_5(D_{C,P} \cdot OnlyProd_t) + \beta_6(D_{C,P} \cdot Both_t) + \beta_7 D_{C,P} + u_{it} \end{aligned} \tag{16}$$

where we distinguish between process innovators only, product innovators only and firms that introduce both process and product innovations. Results shown in Table 5 suggest no evidence of self-selection under the assumption of an exogenous process for the law of motion of productivity. The results of the estimation with an endogenous Markov process can be summarised as follows: i) they reveal the existence of a process of self selection by which the more productive Argentinian and Mexican firm in *t*-1 are the ones that introduce either product innovations or jointly process and product innovations in *t*; ii) we do not find evidence of this process of self-selection into the introduction of product innovation or the introduction of product and process innovations jointly for Colombian and Peruvian firms; finally, iii) we do not find evidence for any country of self-selection into the introduction of process innovations (there is no statistically significant difference between the productivity in *t-1* of the firms that introduce process innovations in *t* and those that do not introduce them). Thus, we do not find any evidence of self selection into the introduction of process innovations.

In light of our results it is interesting to check if there is any difference between productivity in *t-1* for firms that introduce just product innovations in *t* and those that introduce both process and product innovations. This implies to compare the coefficients between the variables $(D_{A,M} \cdot OnlyProd_t)$ and $(D_{A,M} \cdot Both_t)$. The result of this test does not allow rejection of the null of equality of the two coefficients[17] and, thus, leads to our Specification 3 (the preferred and more parsimonious for the self-selection hypothesis):

[17] We test he null H_0: $(D_{A,M} \cdot OnlyProd_t) - (D_{A,M} \cdot Both_t) = 0$, and obtain a value of the *t*-statistic equal to -0.63 with a *p*-value of 0.527.

$$tfp_{it-1} = \beta_0 + \beta_1 \cdot (D_{A,M} \cdot OnlyProc_t) + \beta_2 \cdot (D_{A,M} \cdot Prod_t) + \beta_3 \cdot (D_{C,P} \cdot Innovation_t) + \beta_4 \cdot D_{C,P} + u_{it} \tag{14}$$

where the variables $(D_{A,M} \cdot OnlyProd_t)$ and $(D_{A,M} \cdot Both_t)$ have been grouped into the variable $(D_{A,M} \cdot Prod_t)$ that takes value 1 for firms that introduce product innovations, regardless of whether or not they introduce process innovations simultaneously. Further, as none of the innovation strategy dummies in Specification 2 is significant for Colombia and Peru we have grouped their three innovation strategies dummies into a unique innovation dummy that does not distinguish by innovation types (such as in Specification 1).

The results of the estimation with an endogenous Markov process of specification 3 qualify the results of specification 1 for Argentina and Mexico. The estimates of Specification 1 suggest that the more productive firms in *t-1* innovate in *t*. However, results in Specification 3 suggest that this self-selection effect operates for product innovations and not for process innovations. For Colombia and Peru, as in Specification 1, we do not find any evidence of self-selection.

5.2. Testing hypothesis 2: Returns to innovation.

To test the second hypothesis – namely, that innovation fosters productivity – we are going to check whether more productive firms in *t* are the ones that introduced innovations in period *t-1*. Again, we consider a sequence of different reduced forms. Specification 1, which results are reported in Table 6, is as follows:

$$tfp_{it} = \beta_0 + \beta_1 (D_{A,M} \cdot Innovation_{t-1}) + \beta_2 (D_{C,P} \cdot Innovation_{t-1}) + \beta_3 D_{C,P} + u_{it} \tag{15}$$

where the dependent variable is the TFP in period *t* and $Innovation_{t-1}$ is a dummy variable taking value one if the firm implemented an innovation in *t-1* (without distinguishing by innovation strategies). As in the test of self-selection we group Mexico and Argentina on the one hand and Peru and Colombia on the other hand.

The results for this specification (see Table 6), both for the models with exogenous and endogenous Markov process, confirm the returns to innovation hypothesis for firms in Argentina and Mexico, but not for Colombia and Peru (for these countries the coefficients on innovation are non-significant).

As we are interested in disentangling whether the returns to innovation differ with firms' innovation strategies, we proceed to estimate Specification 2:

$$\begin{aligned} tfp_{it} = {} & \beta_0 + \beta_1(D_{A,M} \cdot OnlyProc_{t-1}) + \beta_2(D_{A,M} \cdot OnlyProd_{t-1}) \\ & + \beta_3(D_{A,M} \cdot Both_{t-1}) + \beta_4(D_{C,P} \cdot OnlyProc_{t-1}) + \beta_5(D_{C,P} \cdot OnlyProd_{t-1}) \\ & + \beta_6(D_{C,P} \cdot Both_{t-1}) + \beta_7 D_{C,P} + u_{it} \end{aligned} \tag{16}$$

where we distinguish between firms according to their innovation strategies. It is possible to observe in Table 6, that both under the assumption of an exogenous Markov (Model 1) and an endogenous Markov (Model 2) all innovation strategies render positive returns for the case of Argentina and Mexico (the estimates of $(D_{A,M} \cdot OnlyProc_{t-1})$, $(D_{A,M} \cdot OnlyProd_{t-1})$ and $(D_{A,M} \cdot Both_{t-1})$ are positive and significant). However, in the case of Peru and Colombia we do not find such evidence ($(D_{C,P} \cdot OnlyProc_{t-1})$, $(D_{C,P} \cdot OnlyProd_{t-1})$ and $(D_{C,P} \cdot Both_{t-1})$ are not significant).

Analogously to our procedure when testing for self-selection and to test whether the innovation returns depend on the kind of innovation strategy, we carry out pairwise comparisons between $(D_{A,M} \cdot OnlyProc_{t-1})$, $(D_{A,M} \cdot OnlyProd_{t-1})$ and $(D_{A,M} \cdot Both_{t-1})$. The results of these pairwise tests suggest that there are not significant differences between the innovation returns rendered by the different innovation strategies (the null of equality of the estimated coefficient is never rejected either in models 1 or 2)[18].

Therefore, we can summarise our results on the test of hypothesis 2 as follows: i) for Argentina and Mexico all innovation strategies render positive returns in terms of productivity without significant differences among them; and, ii) for Peru and Colombia we do not find any evidence of innovation returns.

5.3. Testing hypothesis 3: contemporaneous correlation.

The existence of self-selection and/or returns to innovation can lead to detect contemporaneous correlation between productivity and innovation. To test this possibility we check whether those firms that introduce innovations in period *t* are more productive in period *t*. Our Specification 1 for this hypothesis is as follows:

$$tfp_{it} = \beta_0 + \beta_1(D_{A,M} \cdot Innovation_t) + \beta_2(D_{C,P} \cdot Innovation_t) + \beta_3 D_{C,P} + u_{it} \tag{17}$$

where the dependent variable is the TFP in period *t* and we use as regressor a contemporaneous innovation dummy. The results for this Specification are in Table 7 and both for Models 1 and 2 confirm a contemporaneous correlation between innovation and

[18] We perform the following tests. First, H_0: $(D_{A,M} \cdot OnlyProc_{t-1}) - (D_{A,M} \cdot OnlyProd_{t-1}) = 0$; second, H_0: $(D_{A,M} \cdot OnlyProc_{t-1}) - (D_{A,M} \cdot Both_{t-1}) = 0$; and, third, H_0: $(D_{A,M} \cdot OnlyProd_{t-1}) - (D_{A,M} \cdot Both_{t-1}) = 0$. These tests yield for the GMM-exogenous (Model 1) as statistics: -0.06 and *p*-value=0.952; -0.01 and *p*-value=0.993 and 0.07 and *p*-value=0.941, respectively. For the GMM-endogenous (Model 2) the value of the statistics are: -0.15 and *p*-value=0.880; 0.20 and *p*-value=0.839 and 0.44 and *p*-value=0.660, respectively.

productivity for innovative firms in Argentina and Mexico but not for Colombia and Peru (once more, for these countries the coefficients on innovation are non-significant).

As we are interested in detecting possible different patterns of correlation depending on the type of innovation, we estimate Specification 2:

$$\begin{aligned} tfp_{it} &= \beta_0 + \beta_1(D_{A,M} \cdot OnlyProc_t) + \beta_2(D_{A,M} \cdot OnlyProd_t) \\ &+ \beta_3(D_{A,M} \cdot Both_t) + \beta_4(D_{C,P} \cdot OnlyProc_t) + \beta_5(D_{C,P} \cdot OnlyProd_t) \\ &+ \beta_6(D_{C,P} \cdot Both_t) + \beta_7 D_{C,P} + u_{it} \end{aligned} \tag{18}$$

where we have distinguished among process innovators only, product innovators only, and firms that introduce both types of innovations. Estimates of Specification 2 both for Models 1 and 2 confirm a positive correlation between all innovations strategies and productivity for Argentina and Mexico (the corresponding estimates are positive and significant) but not for and Colombia and Peru.

To analyse whether for Argentina and Mexico the correlation between TFP and innovation depends on the type of innovation, we perform pairwise tests between the estimates of $(D_{A,M} \cdot OnlyProc_t)$, $(D_{A,M} \cdot OnlyProd_t)$ and $(D_{A,M} \cdot Both_t)$.[19] We reject the null of equality of the estimated coefficients except in the comparison between $(D_{A,M} \cdot OnlyProd_t)$ *versus* $(D_{A,M} \cdot Both_t)$. Thus, this leads to our preferred and more

[19] H_0: $(D_{A,M} \cdot OnlyProc_t)$ - $(D_{A,M} \cdot OnlyProd_t)$ = 0, $(D_{A,M} \cdot OnlyProc_t)$ - $(D_{A,M} \cdot Both_t)$ = 0 and $(D_{A,M} \cdot OnlyProd_t)$ - $(D_{A,M} \cdot Both_t)$ = 0) give for the GMM-exogenous (Model 1) the following statistics -1.81 and p-value=0.071, -2.06 and *p*-value=0.040, -0.33 and *p*-value=0.744, respectively. For the GMM-endogenous the value of the statistics are -1.90 and *p*-value=0.058, -2.09 and *p*-value=0.037, -0.22 and *p*-value=0.825, respectively.

parsimonious specification for the contemporaneous correlation hypothesis, that is Specification 3:

$$\begin{aligned} tfp_{it} &= \beta_0 + \beta_1(D_{A,M} \cdot OnlyProc_t) + \beta_2(D_{A,M} \cdot Prod_t) \\ &\quad + \beta_3(D_{C,P} \cdot Innovation_t) + \beta_4 D_{C,P} + u_{it} \end{aligned} \tag{19}$$

where the variables $(D_{A,M} \cdot OnlyProd_t)$ and $(D_{A,M} \cdot Both_t)$ have been grouped in the variable $(D_{A,M} \cdot Prod_t)$ that accounts for product innovators independently of being simultaneously process innovators or not; and the variables $(D_{C,P} \cdot OnlyProc_t)$ $(D_{C,P} \cdot OnlyProd_t)$ and $(D_{C,P} \cdot Both_t)$ are grouped in the variable $(D_{C,P} \cdot Innovation_t)$ as none of them is significant in the estimation of Specification 2.

The results of Specification 2 suggest the existence of a positive correlation with productivity between introducing process innovation only, and product innovations (regardless of the simultaneous introduction of a process innovation) for Argentina and Mexico. For Colombia and Peru, we do not find any evidence of correlation between innovation and productivity. [20]

As the contemporaneous correlation hypothesis does not take into account the direction of causality between innovation and productivity, the results of this hypothesis are a mixture of the results we find for the other two hypotheses. First, on the one hand, the positive correlation between all types of innovation and productivity reflects the found positive effects of all innovation types on productivity according to the returns to innovation hypothesis. On the other hand, the fact that the coefficients for the innovation variables are larger for only product innovators and both, according to the

[20] These results hold both for Models 1 and 2.

contemporaneous correlation results than according to the returns to innovation results, is due to the fact that we also add the positive self-selection effect obtained according to the self-selection hypothesis results for these types of innovations but not for only process innovation.

6. Conclusions.

The literature on firm level productivity in developed countries recognizes the important role played by firm innovation activities on the evolution of firm productivity. However, the literature on this topic for developing and emerging economies is scarce and far from conclusive. One of the main reasons is the lack of data at firm level and, above all, the lack of panel data. Our paper in an attempt to fill this gap in the literature using a panel drawn from the Enterprise Surveys of the World Bank for four Latin American countries (Argentina, Colombia, Mexico and Peru), for the years 2006 and 2010.

For the countries involved in our analysis, we have studied the innovation-productivity link at the firm level and for the manufacturing sector, distinguishing between different innovations types, such as process and product innovations. Our estimation strategy has followed two-steps: first, we have estimated TFP measures following De Loecker (2010) approach and Wooldridge (2009) estimation procedure (this has allowed us to compare results both considering an exogenous or endogenous Markov process for the dynamics of productivity); and, second, we have use the estimated TFPs as dependent variables in several models with innovation activities as covariates, in order to disentangle the effects of those variables on the TFP. In this second stage we have tested three hypotheses. In particular, we have considered if ex-ante more productive firms are those engaged in product or process (or both) innovations (self-selection hypothesis). Second, we have studied whether firms' process or product (or both) innovative strategies cause subsequent higher productivity (returns to innovation hypothesis). Finally, and as a natural consequence of the other two hypotheses, we have checked whether the coexistence of the effects of the aforementioned hypotheses generates contemporaneous correlation between productivity and innovation.

From our results we confirm the self-selection hypothesis of the most productive firms into innovation under the endogenous Markov process, driven by product innovations, only for Argentina and Mexico. We do not confirm self-selection for process innovations, or for any kind of innovations in Colombia and Peru. Further, we obtain that there are returns to innovation in terms of productivity for all innovation types, both under an exogenous or an endogenous Markov but, again, only for Argentina and Mexico. Finally, both the self-selection and the returns to innovation links with productivity working together generate, as expected, cotemporaneous correlation between all innovation activities and productivity for Argentina and Mexico.

This paper has put forward the relevance of considering an endogenous Markov process for productivity, which may affect to the testing results in the literature about the innovation-productivity link. Second, our results point to the convenience of distinguishing among different types of innovations given that some directional relationships may work for only particular types of innovations. And, finally, the results open the question about why some countries do not report evidence of any relationship between innovation and productivity.

References

Ackerberg, D.A., Caves, K. and Frazer, G. (2006): "Structural Identification of Production Function", Working paper, Department of Economics, UCLA.

Aghion, P. and Howitt. P. (1992): A model of growth through creative destruction. Econometrica 60: 323-51.

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., Jensen, J. B., Redding, S., Schott, P. (2007): "Firms in international trade", *Journal of Economic Perspectives* 21(3), 105-130.

Chudnovsky, D., López, A. and Pupato, G. (2006): "Innovation and productivity in developing countries: A study of Argentine manufacturing firm's behavior (1992-2001)", *Research Policy* 35: 266-288.

Cohen, M. W. and Levinthal, D. A. (1989): "Innovation and learning: the two faces of R&D", *The Economic Journal* 99: 73-96.

Cohen, M. W. and Levinthal, D. A. (1990): Absorptive Capacity: A New Perspective on Learning and Innovation. *Administrative Science Quarterly*, 35, 128-152.

Crépon B, Duguet E, Mairesse J. (1998): Research, Innovation and Productivity: An Econometric Analysis at the Firm Level. Economics of Innovation and New Technology 7: 115-158.

Crespi, G. and Zuniga, P. (2012): "Innovation and Productivity: Evidence from Six Latin American Countries", *World Development* 40 (2): 273-290.

Daude, C. (2010): Innovation, productivity and economic development in Latin America and the Caribbean, OECD Development Centre working paper 288.

Dabla-Norris, E., Kersting, E. and Verdier, G. (2010): "Firm Productivity, Innovation, and Financial Development", Working Paper, IMF.

De Loecker, J. (2007): "Do exports generate higher productivity? Evidence Slovenia", *Juornal of International Economics* 73 (1): 69-98.

De Loecker, J. (2010): "A Note on Detecting Learning by Exporting", NBER Working Papers 16548, National Bureau of Economic Research, Inc.

De Loecker, J. and F. Warzyniski (2011): Markups and firm-level status, NBER Working Papers 15198, National Bureau of Economic Research, Inc.

Doraszelski, U. and Jaumandreu, J. (2010): "R&D and Productivity: Estimating Endogenous Productivity", Working Paper, Harvard University.

Edquist, C. (1997): *Systems of Innovation: Technologies, Institutions and Organizations*. London: Pinter.

Ericson, R., Pakes, A. (1992): An alternative theory of firm and industry dynamics. Cowles Foundation Discussion Papers 1041, Yale University.

Ericson, R., Pakes, A. (1995): Markov-perfect industry dynamics: a framework for empirical work. Review of Economic Studies 62: 53-82.

Fagerberg, J., Godinho, M. M., (2004): Innovation and Catching-up, in Fagerberg, J., Mowery, D., Nelson, R. (Eds.), The Oxford Handbook of Innovation. Oxford: Oxford University Press, 514-544.

Fagerberg, J., Srholec, M. and Verpagen, B. (2009): "Innovation and economic development", UNIT-MERIT Working Paper 2009-032, United Nations University - Maastricht Economic and social Research and training centre on Innovation and Technology.

Foster, L., Haltiwanger, J., Syverson, C. (2008): "Reallocation, firm turnover, and efficiency: selection on productivity or profitability?", *American Economic Review* 98(1), 394-425.

Goedhuys, M. (2007a): "The impact of innovation activities on productivity and firm growth: Evidence from Brazil", UNIT-MERIT Working Paper 2007-002, United Nations University - Maastricht Economic and social Research and training centre on Innovation and Technology.

Goedhuys, M. (2007b): "Learning, Product Innovation, and Firm Heterogeneity in Developing Countries: Evidence from Tanzania", *Industrial and Corporate Change* 16: 269-292.

Goedhuys, M., Janz, N. and Mohnen, P. (2008): "Knowledge-based productivity in low-tech industries: Evidence from firms in developing countries", UNIT-MERIT Working Paper 2008-007, United Nations University - Maastricht Economic and social Research and training centre on Innovation and Technology.

González, X., Jaumandreu, J. (1998): "Threshold effects in product R&D decisions: Theoretical framework and empirical analysis. Studies on the Spanish Economy", FEDEA.

González, X., Jaumandreu, J., Pazó, C. (1999): "Innovación, costes irrecuperables e incentivos a la I+D", Papeles de Economía Española 81: 155-166.

Griliches, Z. (1979): Issues in assessing the contribution of R&D to productivity growth. Bell Journal of Economics 10: 92-116.

Griliches, Z. (1986). Productivity, R&D, and the Basic Research at the Firm Level in the 1970's, *American Economic Review*, American Economic Association, 76, 141-54.

Griliches, Z. (2000): "R&D, Education and Productivity". Masachusetts: Harvard University Press.

Gu, W. and J. Tang, (2004): "The link between innovation and productivity in Canadian manufacturing industries", *Economics of Innovation and New Technology*, 13, 7,671-686.

Hall B. (2002): The financing of research and development. Oxford Review of Economic Policy 18(1): 35-51.

Hall, B. (2011): "Innovation and productivity", UNIT-MERIT Working Paper Series 2011-028, United Nations University - Maastricht Economic and social Research and training centre on Innovation and Technology.

Huergo, E. and Jaumandreu, J. (2004): "Firm's Age, Process Innovation and Productivity Growth", *International Journal of Industrial Organization* 22: 541-559.

Kasahara, H. and J. Rodrigue (2008): Does the use of imported intermediates increase productivity? Plant-level evidence, Journal of Development Economics 87, 106–118.

Kim L., (1980): Stages of development of industrial technology in a developing country: a model. *Research Policy*, 9, 254-277.

Klette, T. J., Johansen, F. (1998): Accumulation of R&D capital and dynamic firm performance: A not-so-fixed effect model. Annales d' Economie et de Statistique 49/50: 389-419.

Lee, K. and Kang, S-M. (2007): "Innovation Types and Productivity Growth: Evidence from Korean Manufacturing Firms", *Global Economic Review* 36: 343-359.

Levinsohn, J. and Petrin, A. (2003): "Estimating Production Functions Using Inputs to Control for Unobservables", *The Review of Economic Studies* 70 (2): 317-341.

Lundvall, B.Å. (ed.) (1992): "National System of Innovation: Towards a theory of innovation and interactive learning". Londres, Pinter.

Mañez Castillejo, J. A., Rincón Aznar, A., Rochina Barrachina, M. E. and Sanchis Llopis, J. A. (2005): "Productividad e I+D: Un Análisis No Paramétrico", *Revista de Economía Aplicada* 39 (13): 47-86.

Máñez J A, Rochina-Barrachina M E, Sanchis A, Sanchis J A (2009a) Self-selection into exports: productivity and/or innovation. Applied Economics Quarterly 55(3): 219-241.

Máñez, J. A., Rochina-Barrachina, M. E., Sanchis, A., Sanchis, J. A. (2009b): The role of sunk costs in the decision to invest n R&D. Journal of Industrial Economics 57(4): 717-735.

Navarro, J. C., Llisteri, J. J. and Zuñiga, P. (2010): "The importance of ideas: innovation and productivity in Latin America" en C. Pagés (Ed), *The age of productivity: transforming economies from the botton up.* Developement in the Americas Report. Washington, DC, United States: Inter-American Development Bank/Palgrave-McMillan.

Nelson, R. (1993): *National Innovation Systems: A Comparative Analysis*. New York: Oxford University Press.

Olley, S. and A. Pakes (1996), "The Dynamics of Productivity in the Telecommunications Equipment Industry", *Econometrica* 64 (6): 263-1298.

Pakes, A., Ericson, R. (1998): Empirical implications of alternative models of firm dynamics. Journal of Economic Theory 79: 1-45.

Parisi, M. L., Schiantarelli, F. & Sembenelli, A. (2006). Productivity, innovation and R&D: Micro evidence for Italy. *European Economic Review, 50,* 2037-2061

Raffo, J., Lhuillery, S. and Miotti, L. (2008): "Northern and Southern Innovativity: A Comparisson across European and Latin American Countries", *European Journal of Development Research* 20: 219-239.

Rochina-Barrachina, M.E., Mañez, J.A. and Sanchis-Llpois, J.A. (2010): "Process Innovations and Firm Productivity Growth", *Small Business Economics* 34: 147-166.

Romer, P. (1990): Endogenous technological change. Journal of Political Economy, 98(5): 71-102.

Saliola, F. and Seker, M. (2011): "Total Factor Productivity Across Developing World", Enterprise Note Nº 23, World Bank.

Schumpeter, J. (1943): "The Theory of Economic Development". Harvard University Press.

Smolny, W. (1998): Innovations, prices and employment: A theoretical model and an empirical application for West German manufacturing firms. The Journal of Industrial Economics 3: 359-381.

Sutton, J. (1991): Sunk Costs and Market Structure. Massachussets: MIT Press.

Van Biesebroeck, J. (2007): "Robustness of Productivity Estimates", *Journal of Industrial Economics* 55: 529-569.

Vernon, R. (1966): "International investment and international trade in the product cycle", *Quarterly Journal of Economics* LXXX, 190-207.

Verspagen, B. (1999). European Regional Clubs: Do They Exist and Where Are They Heading? On the Economic and Technological Differences between European Regions. (In J. Adams & F. Pigliaru (Eds.), *Economic Growth and Change: National and Regional Patterns of Convergence and Divergence.* Cheltenham: Edward Elgar.)

Table 1. Proportion of innovating firms by manufacturing industries and country.

Industry	Process innovators (%)	Product innovators (%)	Process and product innovators (%)	None (%)
Food	11.16	16.49	51.77	20.04
Textile	12.22	17.19	48.42	22.17
Garments	13.08	14.24	50.58	22.09
Chemical	7.47	19.16	54.22	19.16
Non metallic and basic metals	5.13	20.51	38.46	35.90
Fabricated metals and machinery	6.28	18.83	53.36	21.52
Other manufacturing	6.54	15.03	47.71	30.72
Country				
Argentina	8.48	20.32	52.78	18.24
Colombia	12.18	13.45	54.06	20.30
Mexico	9.70	12.71	35.45	42.14
Peru	10.77	17.44	57.18	14.62
Total	10.02	16.86	51.10	22.01

Table 2. Differences across innovating and non-innovating firms.

	Process innovators	Product innovators	Both process and product
Sales	18.543	80.016**	197.087***
Sales per worker	5.249	33.155**	41.014**
Capital (net value) per worker	23.938	-16.668	19.368
Materials per worker	-13.570	23.488**	25.737**
Size	11.318	28.072**	101.850***

Note: ***, ** mean significance at the 1% and 5% levels, respectively.

Table 3. Differences between innovating and non-innovating firms.

	Coefficient	s.e
Sales		
Product *vs.* process and product	-0.501***	(0.165)
Sales per worker		
Product *vs.* process and product	-0.057	(0.111)
Materials per worker		
Product *vs.* process and product	-0.018	(0.086)
Size		
Product *vs.* process and product	-0.455***	(0.100)

Notes:

1. Robust standard errors are in parenthesis.
2. Significance level: ***p<1%.

Table 4. Country production function estimates.

Method	N	Capital		Labour		Materials	
		β_k	s.e.	β_l	s.e.	β_m	s.e.
Argentina							
OLS	157	0.046	(0.042)	0.316***	(0.069)	0.601***	(0.069)
GMM-exogenous	157	0.066*	(0.039)	0.265***	(0.054)	0.530***	(0.054)
GMM-endogenous	157	0.041	(0.036)	0.238***	(0.051)	0.566***	(0.116)
Colombia							
OLS	133	0.086**	(0.033)	0.399***	(0.068)	0.506***	(0.058)
GMM-exogenous	133	0.021	(0.028)	0.406***	(0.049)	0.430***	(0.121)
GMM-endogenous	133	0.053*	(0.028)	0.407***	(0.043)	0.365***	(0.098)
Mexico							
OLS	88	0.128**	(0.049)	0.479**	(0.155)	0.418**	(0.145)
GMM-exogenous	88	0.208**	(0.071)	0.264***	(0.070)	0.492***	(0.154)
GMM-endogenous	88	0.107**	(0.051)	0.229***	(0.065)	0.653***	(0.089)
Peru							
OLS	85	0.143**	(0.063)	0.345***	(0.091)	0.484***	(0.145)
GMM-exogenous	85	0.050	(0.047)	0.367***	(0.060)	0.476***	(0.135)
GMM-endogenous	85	0.047	(0.044)	0.347***	(0.058)	0.418***	(0.110)

Notes:

1. Robust standard errors in parenthesis. Significance level: ***p<1%, **p<5% and * p<10%.
2. The production function estimates control for industry dummies.

Table 5. Hypothesis1: Self-selection.

	GMM-exogenous		GMM-endogenous	
TFP_{t-1}	β	s.e.	β	s.e.
Specification 1				
$D_{A,M}$*$Innovation_t$	0.060	(0.088)	0.360***	(0.124)
$D_{C,P}$*$Innovation_t$	0.126	(0.097)	0.066	(0.137)
Specification 2				
$D_{A,M}$*Only $Process_t$	-0.016	(0.164)	-0.025	(0.229)
$D_{A,M}$*Only $Product_t$	0.039	(0.105)	0.352**	(0.147)
$D_{A,M}$*$Both_t$	0.088	(0.096)	0.431***	(0.135)
$D_{C,P}$*Only Process $_t$	0.084	(0.143)	0.051	(0.200)
$D_{C,P}$*Only Product $_t$	0.170	(0.130)	0.066	(0.182)
$D_{C,P}$*Both $_t$	0.122	(0.103)	0.061	(0.145)
Specification 3 (preferred)				
$D_{A,M}$*Only Process $_t$			-0.024	(0.227)
$D_{A,M}$*Product $_t$			0.399***	(0.124)
$D_{C,P}$*Innovation $_t$			0.060	(0.136)

Notes:

1. Robust standard errors are in parenthesis.
2. Significance level: ***p<1%, **p<5% and * p<10%.
3. We also control for $D_{C,P}$ *versus* $D_{A,M}$ in estimation.
4. In Specification 2, under GMM-endogenous, we have performed a *t*-statistic for the difference of coefficients between the variables $D_{A,M}$**Only Product$_t$* and $D_{A,M}$**Both$_t$* (H_0: $D_{A,M}$**Only Product$_t$* - $D_{A,M}$**Both$_t$* = 0; value of the statistic -0.63 and *p*-value=0.527) not rejecting the null hypothesis and, thus, leading to Specification 3 (more parsimonious).

Table 6. Hypothesis 2: Returns to innovation.

	GMM-exogenous		GMM-endogenous	
TFP_t	β	s.e.	β	s.e.
Specification 1 (preferred)				
$D_{A,M}$*Innovation $_{t-1}$	0.661***	(0.103)	1.100***	(0.167)
$D_{C,P}$*Innovation $_{t-1}$	-0.086	(0.156)	0.115	(0.252)
Specification 2				
$D_{A,M}$*Only Process $_{t-1}$	0.663***	(0.188)	1.135***	(0.305)
$D_{A,M}$*Only Product $_{t-1}$	0.676***	(0.169)	1.189***	(0.275)
$D_{A,M}$*Both $_{t-1}$	0.664***	(0.108)	1.076***	(0.176)
$D_{C,P}$*Only Process $_{t-1}$	-0.120	(0.209)	0.032	(0.339)
$D_{C,P}$*Only Product $_{t-1}$	-0.268	(0.216)	-0.133	(0.350)
$D_{C,P}$*Both $_{t-1}$	-0.049	(0.159)	0.168	(0.258)

Notes:

1. Robust standard errors are in parenthesis.
2. Significance level: ****p*<1%, ***p*<5% and * *p*<10%.
3. We also control for $D_{C,P}$ *versus* $D_{A,M}$ in estimation.
4. In Specification 2 we have performed *t*-statistics for pairwise comparisons between the variables $D_{A,M}$*$Only\ Process_{t-1}$, $D_{A,M}$*$Only\ Product_{t-1}$ and $D_{A,M}$*$Both_{t-1}$ (H_0: $D_{A,M}$*$Only\ Process_{t-1}$ - $D_{A,M}$*$Only\ Product_{t-1}$ = 0, $D_{A,M}$*$Only\ Process_{t-1}$ - $D_{A,M}$*$Both_{t-1}$ = 0 and $D_{A,M}$*$Only\ Product_{t-1}$ - $D_{A,M}$*$Both_{t-1}$ = 0. For the GMM-exogenous the value of the statistics are: -0.06 and *p*-value=0.952; -0.01 and *p*-value=0.993; and 0.07 and *p*-value=0.941, respectively. For the GMM-endogenous the value of the statistics are: -0.15 and *p*-value=0.880; 0.20 and *p*-value=0.839; and 0.44 and *p*-value=0.660, respectively. Therefore, we cannot reject the null hypotheses, which indicates that Specification 1 is more parsimonious.

Table 7. Hypothesis 3: Contemporaneous correlation.

	GMM-exogenous		GMM-endogenous	
TFP_t	β	s.e.	β	s.e.
Specification 1				
$D_{A,M}$*Innovation $_t$	0.725***	(0.111)	1.295***	(0.179)
$D_{C,P}$*Innovation $_t$	-0.091	(0.123)	-0.061	(0.198)
Specification 2				
$D_{A,M}$*Only Process $_t$	0.352*	(0.206)	0.638**	(0.332)
$D_{A,M}$*Only Product $_t$	0.716***	(0.132)	1.253***	(0.213)
$D_{A,M}$*Both $_t$	0.753***	(0.121)	1.293***	(0.195)
$D_{C,P}$*Only Process $_t$	-0.158	(0.180)	-0.196	(0.290)
$D_{C,P}$*Only Product $_t$	-0.201	(0.164)	-0.187	(0.263)
$D_{C,P}$*Both $_t$	-0.029	(0.130)	0.022	(0.209)
Specification 3 (preferred)				
$D_{A,M}$*Only Process $_t$	0.351*	(0.205)	0.667**	(0.330)
$D_{A,M}$*Product $_t$	0.742***	(0.112)	1.318***	(0.181)
$D_{C,P}$*Innovation $_t$	-0.090	(0.123)	-0.060	(0.198)

Notes:

1. Robust standard errors in parenthesis.
2. Significance level****p*<1%, ***p*<5% and * *p*<10%.
3. We also control for $D_{C,P}$ *versus* $D_{A,M}$ in estimation.
4. In Specification 2 we have performed *t*-statistics for pairwise comparisons between the variables $D_{A,M}$*Only Process$_{t-1}$, $D_{A,M}$*Only Product$_{t-1}$ and $D_{A,M}$*Both$_{t-1}$ (Ho: $D_{A,M}$*Only Process$_{t-1}$ - $D_{A,M}$*Only Product$_{t-1}$ = 0, $D_{A,M}$*Only Process$_{t-1}$ - $D_{A,M}$*Both$_{t-1}$ = 0 and $D_{A,M}$*Only Product$_{t-1}$ - $D_{A,M}$*Both$_{t-1}$ = 0; for the GMM-exogenous the value of the statistics are -1.81 and p-value=0.071, -2.06 and p-value=0.040, -0.33 and p-value=0.744, respectively; for the GMM-endogenous the value of the statistics are -1.90 and p-value=0.058, -2.09 and p-value=0.037, -0.22 and p-value=0.825, respectively) not rejecting the null hypotheses only in the case of the comparison $D_{A,M}$*Only Product$_{t-1}$ *versus* $D_{A,M}$*Both$_{t-1}$ and, thus, leading to Specification 3 (more parsimonious).