



# Productivity in the service sector in Colombia: the misallocation and innovation channels

**Autorías:** Sanguinetti, Pablo; Feroce, Andrés

**Fecha de publicación:** Diciembre 2024

## ¿Cómo citar este trabajo?

Sanguinetti, P., Feroce, A. (2024). *Productivity in the service sector in Colombia: the misallocation and innovation channels*. [Working Paper. Universidad Torcuato Di Tella]. Repositorio Digital Universidad Torcuato Di Tella.

<https://repositorio.utdt.edu/handle/20.500.13098/13576>

El presente documento se encuentra alojado en el Repositorio Digital de la **Universidad Torcuato Di Tella**, para su preservación, archivo y difusión.

**Dirección:** <https://repositorio.utdt.edu>

# Productivity in the service sector in Colombia: the misallocation and innovation channels

Pablo Sanguinetti and Andres Feroce

Universidad Torcuato Di Tella

This draft, December 2024

## 1. Motivation and research questions

Latin America and the Caribbean have underperformed in economic growth compared to other developing regions like Southeast Asia or Eastern Europe. This has implied a low rate of income convergence with developed economies. For example, the average income per capita was around 0.20 of that of the USA in 1960 and went up to only 0.26 at the end of the second decade of the XXI century. Most growth accounting exercises (see Jones, 2019; Alvarez et al., 2018) show that weakened productivity dynamics is a critical immediate cause of this lack of convergence. In other words, the LAC countries have not improved their efficiency when using their resources at the firm level, sectors, and the aggregate economy.

In the last 15 years, there has been a flourishing literature looking at the forces behind this sluggish productivity growth in developing economies and LAC, in particular. One channel that has been emphasized is the problem of misallocation. The difference in productivity across countries is partly associated with substantial productivity disparities across firms, even within narrowly defined sectors. These disparities are, in turn, related to the existence of distortionary policies (taxes, regulations, subsidies, etc.) and/or market failures (financial frictions) that affect high-productivity enterprises or entrepreneurs, which limits their participation in the market and their expansion while, at the same time, promoting the survival of small, low-productivity firms that absorb a large portion of the resources of the economies (Banerjee and Dufflo, 2005; Restuccia & Rogerson, 2008; Hsiew & Klenow, 2009).

Various methodologies has been proposed to measure the extend of misallocation in developing economies (Hsiew and Klenow, 2009; Bartelsman et al., 2013). Most empirical analyses have focused on (formal) manufacturing activities. For example, estimations for Latin America show that these discrepancies in productivity across firms, which are implicitly associated with policy distortions/market failures, could have a significant impact on aggregate productivity, with a fall between 50 and 60% in TFP in the manufacturing sector in many countries (Buzzo et al. 2012).

There are important reasons suggesting that this misallocation problem could also significantly affect service activities, one being their much higher level of informality. The lack of comprehensive surveys is one factor that has precluded the rigorous analysis for the various service activities. The first objective of this paper is to close this knowledge gap by evaluating this issue in the case of Colombia. To measure the extent of misallocation, we will apply (see below for more details) the approach that computes the covariance between firm size (or market share) and productivity. When using this method in formal manufacturing activities, various papers (see Olley and Peck, 2016; Bartelsman et al., 2013; Alvarez et al., 2018) find relatively small effects of the misallocation factor concluding that the significant reason behind the low productivity in many developing economies is the low average productivity at the firm

level. This, in turn, is related to the innovation effort the firm performs during its life cycle, from the start-up stage to more mature phases. We will also evaluate to what extent the innovation channel play a role in explaining the evolution of productivity for the formal service sector in Colombia. These more dynamic factor have proven relevant for manufacturing activities in some developing economies like Mexico and India explaining the different in firm size between these economies and those of developed countries (Hsiew and Klenow, 2014; Bento and Restuccia, 2017).

The case of Colombia offers an excellent opportunity to look at this issue given the availability of the innovation survey for the service sector, EDITS (Encuesta de Innovación de Servicios)<sup>1</sup>. The EDITS has a panel form and detailed information about whether service firms in Colombia have participated in the various government programs to support innovation efforts by private firms (i.e., Innpulsa, Colciencia, etc.). We can then evaluate whether these programs have impacted innovation inputs (say, total innovation expenditures, labor allocated to these activities) or outputs (number of innovations implemented, sales, total employment, and labor productivity).

The research questions the study will try to address are the following. To what extent is the productivity of the service sector in Colombia affected by a misallocation problem across firms or subsectors? Alternatively, how much the low productivity of the service sector is due to low productivity at the firm level? If this is the case, could we link this low productivity of the firm to a lack of research and development and innovation (R+D+i) efforts? How have R+D+i expenditures evolved in Colombia's services sector in recent years? What market failures justify government support for R+D+i in the service sector? What are the most common policies that Colombia has adopted, and what is the evidence of their impact on different innovation inputs and outcomes?

We are not the first to examine misallocation and innovation channels as determinants of productivity. As we will show below, many papers have already looked at these questions globally and in Latin America. Still, as indicated above, most of these papers have focused on the manufacturing sector<sup>2</sup>. Thus, it could be interesting to find out whether the analysis of the service in Colombia brings in new evidence and policy lessons that can be useful for governments in the region.

The rest of the paper is organized as follows. The next section describes a theoretical framework that helps to identify the misallocation and innovation channels as immediate determinants of the productivity of the economy and sectors and explains the Olley and Peck (1996) methodology

<sup>1</sup>There is also the Annual Service Survey, EAS (Encuesta Annual de Servicios). But, as we argued below, it has a lower coverage, so we will not use it at this stage of the project.

<sup>2</sup> An interesting exception is the report published by the IADB (see Navarro y Olivari, 2016), which analyzes the impact of misallocation on productivity in the service sector in various countries of the region using the World Bank Enterprise Survey. They also focus on the innovation channel, looking at the innovation surveys for Chile and Uruguay. For the case of Colombia Albis et al., 2017 analyze the innovation channel on firm productivity for the service and manufacturing sectors in Colombia using the innovation survey. In particular they perform impact evaluation analyses of innovation support programs. They only look at expenditures and productivity, while in this paper we look at some other input and output innovation variables like labor allocated to innovation activities, number of innovations reported by firms, sales and total employment. They also fail to include year-sector interaction terms which, as we will see, are quite relevant for the regression results.

for estimating these two components. Section 3 depicts the data sources and some descriptive statistics regarding the service sector in Colombia. Section 4 shows the results for the Olley and Peck (1996) decomposition of productivity. Section 5 concentrates on the issue of innovation as a source of productivity dynamics at the firm level and presents some descriptive data on innovation inputs and outputs for various services activities and subsectors. Section 6 describes different policy initiatives to promote innovation in Colombia and whether these policies affect innovation inputs and outputs variables. Section 7 concludes with policy recommendations and future areas of research.

## 2. Conceptual framework: misallocation and innovation as determinants of productivity.

As detailed by Alvarez et al. (2018), within each aggregate sector, like manufacturing or services (or even the economy as a whole), multiple and highly heterogeneous production units coexist like subsectors and firms. The large heterogeneity in productivity among these units implies that aggregate productivity depends on the average unit's productivity and the extent to which the available capital and human talent are allocated to high-productivity units. The former captures the "internal component" of productivity or, as it is usually called, "the innovation channel" within a given productive unit. In contrast, the latter captures the "external component," or the "allocative efficiency channel" across those productive units.

Of course, the innovation component is important: it captures the productivity level of an average unit. However, the higher the heterogeneity in productivity, the less representative the average unit is. Here, the allocation component of productivity comes into play, which captures a straightforward idea. If there are units with different productivity levels, aggregate productivity will be greater if more resources are allocated to those with high productivity.

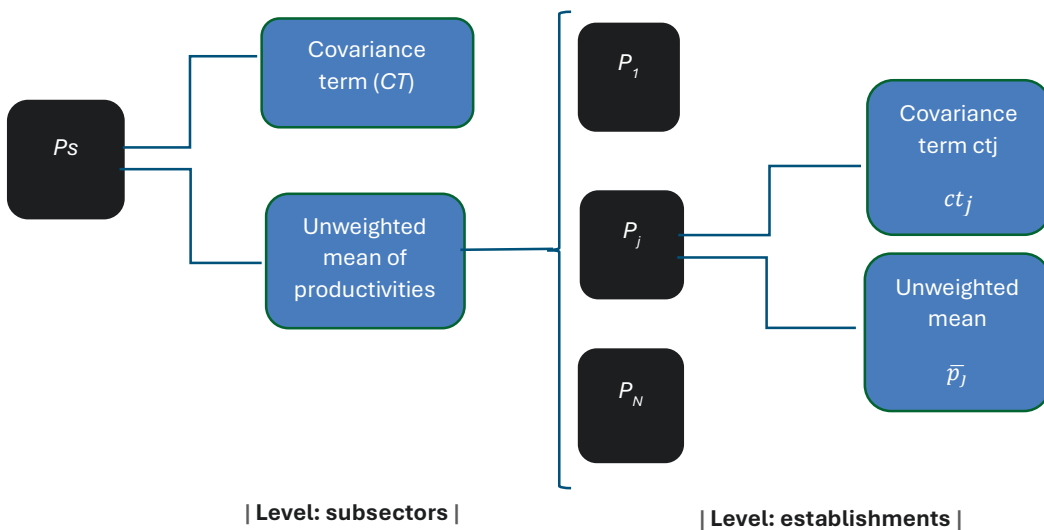
We will apply this simple decomposition framework of productivity between these two components to the service sector in Colombia to judge to what extent misallocation and innovation factors can explain the level and dynamic of productivity in this sector. For this purpose, we will apply Olley and Peck (1996) methodology. Consider a set of economics units and let  $p_i$ ,  $s_i$  represent, respectively, a productivity index and an input share (weight) of the unit  $i$ . The aggregate productivity,  $P$ , of a sector, subsector, or the economy as a whole can be expressed as:

$$P = \sum_i p_i s_i = \bar{p} + \sum_i (p_i - \bar{p})(s_i - \bar{s}) \quad (1)$$

where  $\bar{p}$  is the unweighted mean of productivity indexes  $p_i$  across all units of production while  $\bar{s}$  is the same for resource (or market) share  $s_i$ . Thus, we see that the productivity  $P$  of a sector can be decomposed into two parts: the average productivity of the production units belonging to that sector and covariance terms that describe the correlation between relative productivity and the size or market share of the unit of production. The first component reflects innovation efforts that determine the productivity of individual firms (the average). At the same time, the second describes the extent to which the allocation of resources across

individual units helps to increase aggregate sector-level productivity. Thus, a high positive value for this term implies a better allocation of resources.

We can apply equation (1) to explain the service sector’s overall productivity by using it at the subsector and establishment level. Figure 1 could help us understand the two steps of this process. First, taking the subsectors as units of analysis, the productivity can be decomposed by the covariance terms  $CT$  associating productivity and the size of each subsector plus the unweighted average of productivity across all subsectors  $P_j$ . In a second step, taking the establishments as a unit of analysis, the productivity of each subsector  $P_j$  can be divided between the covariance terms corresponding to each subsector  $ct_j$  -that associate productivity of individual firms with their size- plus the (unweighted) average productivity across firms within each subsector,  $\bar{p}_j$ .



Thus, we arrive at the following expression,

$$P_s = CT + \frac{1}{N} \sum_{j=1}^N ct_j + \frac{1}{N} \sum_{j=1}^N \bar{p}_j \quad (2)$$

We will use equation (2) to run the “anatomy of productivity” analysis for the service sector in Colombia.

*The innovation channel*

Of course, this decomposition only describes the “anatomy” of an economy or sector's productivity and not the ultimate factors behind it. A key issue behind the internal component is firms' innovation efforts. As indicated in the previous section, we will perform a detailed

analysis of this channel by looking at data describing the innovation activities of service firms in Colombia and linking these activities with firm-level productivity.

Now, innovation initiatives may be low because of various reasons. On the one hand, what we call government failures are associated with distortions caused by ill-designed public policies, implementation failures, corruption, and other factors that generate uncertainty and affect the return on these innovation investments, lowering the incentives of enterprises to conduct these initiatives<sup>3</sup>. One example could be very high tariffs for specific intermediate inputs critical for producing a new product that could be developed by investing in engineering and design. Thus, innovation efforts are affected by transversal public policies associated with tax regimes, trade openness and intensity of competition, labor market regulations, the macroeconomic regime, and public infrastructure investments, among others, that also affect productivity at the firm level through channels that goes beyond innovation (Alvarez et al. 2018).

The second group of reasons is associated with market failures that could weaken private sector motivations for making these investments. These reasons usually distinguish between externalities, or knowledge spillover, and financial market imperfections (Bloom et al., 2019).

Regarding externalities, new knowledge developed by a firm that could end up in a new product or a production process could be totally or partially a nonrival and nonexcludable good, which implies that other firms may be able to copy them without paying its development costs. These spillovers may create a wedge between private and social returns and a disincentive against private investment in knowledge production. This “public good” nature of knowledge applies more strongly to generic or scientific knowledge than to technological knowledge, which is more applicable and specific to the firm. As will be seen below, some entry barriers or regulations (say patent laws) could help avoid free riding, but many times, they do it partially. In addition, these interventions also restrict the diffusion of new ideas, which could affect economy-wide productivity. So, there is a trade-off that needs to be evaluated. We will see to what extent public grants could help to solve this public good nature affecting some innovation investments.

Another externality-type market failure is when innovation occurs through collaboration among various firms along a production chain. A firm that is launching a new product may need specific inputs that could only be produced by a few suppliers, which may invest in developing these products if enough potential clients demand them. This could also happen with specific critical infrastructure relevant to the production chain (i.e., a quality certification lab). Cluster-type policies may help to deal with this coordination issue.

Concerning financial market imperfections, innovation projects are subject to higher uncertainty about their outcomes and extended maturity periods than ordinary investments. This makes projections of potential cash flows the project generates challenging to predict. In addition, innovation investments typically include a large proportion of intangible assets with minimal use as collateral. All of this creates a wedge between the rate of return required by an innovator investing his or her own funds and that needed for an external credit institution. This would imply underinvestment in innovation from the point of view of the social optimum due to a lack of access to financing or the high cost of capital. This is why it is usually seen that funding for

---

<sup>3</sup> As we mention before these distortions also affects the allocation channel affecting aggregate productivity.

innovative projects is implemented through equity or venture capital instruments and, to a lesser extent, through the traditional banking system. We come back to this issue when discussing innovation policy below.

We will look at the various instruments the Colombian government designed to deal with these market imperfections affecting private sector investment in innovation and how they apply to the service sector. We will extend the analysis by discussing government capabilities (for example, lack of information) and other institutional restrictions (i.e., regulations on public procurement for technologically advanced products) that may limit what the government can do. Finally, we will empirically assess how these mechanisms have effectively promoted innovation inputs and outputs and how they impact firms' productivity.

### **3. Data sources and descriptive statistics regarding the service sector in Colombia.**

This section will describe some basic features of the EDITS surveys, which will be used to analyze how the allocation and innovation channels affect productivity, characterize the innovation effort made by firms in the service sector, and perform an impact evaluation analysis of the government promotion policies<sup>4</sup>.

#### *The main features of the EDITS survey*

This is a biannual innovation survey that DANE applied to the service sector. It aims to describe different innovation activities performed by firms. The sampling comes from the census of large companies in the service and commerce sector, the Annual Services Survey (EAS), and the Annual Trade Survey (EAC). The data can be disaggregated by subsectors as defined by the CIUU rev 4 classification comprising 19 activities. Table 1 describes the sector classification, the number of firms in each sector participating in the survey, and the criteria used to include them in the database.

---

<sup>4</sup> We chose to work with the EDITS instead of the EAS survey because of its greater coverage. In particular, the former database includes wholesale and retailing activities, a major subsector within the service sector. As a consequence, the EDITS contains a larger sample of firms. See Annex X for the main features of the EAS survey.

Table 1. EDITS's sample features.

Description of the sector	# of firms	Criteria for inclusion (pesos of 2018; rate of exchange 1USD= 2900 pesos)
Electricity, gas, stema and air conditions supply	90	20 workers or more
Water collection, treatment and distributioun	154	20 workers or more
Wastewater Treatment and Waste Disposal	149	20 workers or more
Vehicule Trade, Maintenance and Repair	313	100 workers or more or income above 15000 millions
Wholesale trade, except vehicules	1780	100 workers or more or income above 15000 million
Retail trade, except trade in vehicules	1932	50 workers or more or income above 5.000 million
Automotive Public Ground Transportation	1302	20 workers or more
Air transport	48	20 workers or more
Mail and Courier Services	64	40 workers or more or income above 3.000 million
Accomodation and Food Services	712	40 workers or more or income above 3.000 million
Editing Activities	109	10 workers or more and income above 500 million
Film, video sound recording and music edition	36	40 workers or more or income above 2.000 million
Programming, broadcasting and /or news agency activities	37	40 workers or more and income above 3.000 million
Telecomunications	166	40 workers or more or income above 3.000 million
Develoment of software, computing systems and data processin	334	75 workers or more or income above 3.000 million
Banking and financial services	24	Census
Research and Development Centers	42	Census*
University Educacion	265	20 workers or more or income above 1.000 million
Human health activities	1255	40 workers or more or income above 3.000 million
Total number of firms	8812	

\*Recognized by COLCIENCIAS. Source: DANE

Wholesale Commerce and Retail Trade are the activities with the most significant number of enterprises (1932 and 1780, respectively). On the other hand, Film, Video, Sound Recording, and Music, as well as Programming, Broadcasting, and News Agency activities (37), have the smallest number of firms (36 and 37, respectively). As shown in the Table, for most activities, firms must have a minimum of 20 or 40 workers to be included in the survey. Only Editing has a relatively lower limit of 10 or more workers. When the income criteria are used for selection, most subsectors have a threshold of 1 million dollars of annual income. These collection data requirements would imply, as in the case of the EAS Survey, that many formal small firms (say, between 10 and 20 workers or income less than 1 million dollars) will not be included in the sample. For 2019, the total number of firms in the database was 8812.

The survey tracks information about 726 variables. It reports primary data, such as total sales, employment, wages, and exports. Most importantly, it also records variables associated with innovation initiatives like the number of innovations carried out by companies, by type of innovation; the total amount invested by companies in scientific, technological, and innovation activities; financing of scientific, technological, and innovation expenditures; the number of people employed by the company who participated in the implementation of scientific, technological and innovation activities; patenting; etc.

#### *Some basic descriptive statistics for 2019*

To get a flavor of the characteristics of the firms participating in the survey, Table 2 describes some basic statistics for the entire sample of establishments in 2019. The average mean sales is around 69 thousand million pesos of 2018 (around 23 current million dollars). The largest firm in that year had sales of almost 200 times that value (around 5 billion dollars). We have some very low sales observations (the minimum is about 530 thousand pesos, 182 dollars), but they are outliers, as almost 97% of the firms have sales above 100000 dollars (86% above 1 million dollars)<sup>5</sup>. The average firm in 2019 had 219 workers, though the largest reaches 11.218

<sup>5</sup> In the regression análisis we will drop this outliers.

laborers. The mean labor productivity is around 412 hundred million pesos (141 thousand dollars).

Table 2 describes some basic statistics of the EDITS survey corresponding to 2019.

Variables	Obs	mean	Std Dev	Min	Max
Total sales*	9,182	68800000	410000000	530	14400000000
Numer of workers	9,182	220	599	1	11208
Labor productivity*	9,182	411441,5	1475241	100	105000000

\*In thousands of pesos of 2018. Source: own elaboration and DANE

Table 3 describes the distribution of firms and employment shares by firm size. We classify firms into four size categories: micro, small, medium, and large. Given the criteria used to select firms that participate in the survey, it is unsurprising that establishments with less than 10 employees have a minimal portion in the sample (5,2% of the total number of firms and 0,16% share in total employment). We expect a representative sample from the entire service sector (including formal and informal firms) to have a much larger share of microenterprises. Small firms (11-50 employees) have a 1/3 participation in the total number of firms but still a very low share in employment (4,35%). Instead, medium and especially large enterprises have a much larger participation in employment. Both groups account for 52% of the firms and 75% of total employment (out of 2,02 million employees in 2019).

Table 3. Number of firms, employment share by size categories. EDITS 2019

Size	Micro 1-10	Small 11-50	Medium 51-250	Large 251 +	Total
#Firms	483	3047	3985	1688	9203
%	5,2%	33,1%	43,3%	18,3%	100%
#Employees	3177	88090	459676	1476267	2027210
%	0,16%	4,35%	22,68%	72,82%	100

Source: own elaboration and DANE

Table 4 presents additional descriptive statistics regarding employment and labor productivity, categorized by 1-digit ISIC rev 4 sectors. The service sector in Colombia is diverse, as evidenced by the data. For instance, the wholesale and retail trade sector has the most significant number of firms across service subsectors (4187) despite having the lowest average size of establishment in terms of employment (132 workers). On the other hand, the financial and insurance activities have the smallest number of firms in the sample but the most significant average number of employees per firm (4178). The average labor productivity of firms varies significantly, from 30000 dollars (of 2018) in Education, a labor-intensive sector, to 606000 dollars in Electricity and gas, a capital-intensive activity.

Table 4. Descriptive statistics for employees and labor productivity by firms and subsectors. EDITS 2019

sector	description	variables	min	max	average	firms
D	Electricity, gas, steam and air conditioning supply	employees	17,0	2221,0	400,7	90
		Labor productivity	34789,5	4800767,5	1758115,4	
E	Water supply; sewerage, waste management and remediation activities	employees	1,0	800,0	181,7	323
		Labor productivity	841,2	1281603,0	152732,2	
G	Wholesale and retail trade; repair of motor vehicles and motorcycles	employees	1,0	11208,0	132,5	4187
		Labor productivity	366,7	104813810,5	669398,7	
H	Transportation and storage	employees	1,0	3235,0	200,0	1468
		Labor productivity	100,1	8950104,5	111337,3	
I	Accommodation and food service activities	employees	1,0	2696,0	180,0	760
		Labor productivity	7128,1	1624603,8	97245,6	
J	Information and communication	employees	1,0	5388,0	233,2	732
		Labor productivity	571,4	43361112,0	446651,9	
K	Financial and insurance activities	employees	327,0	9336,0	4178,5	26
		Labor productivity	129051,3	2179323,7	916641,3	
M	Professional, scientific and technical activities	employees	7,0	1751,0	167,4	41
		Labor productivity	16304,2	452962,5	103291,8	
P	Education	employees	5,0	5247,0	815,5	260
		Labor productivity	2608,0	885507,5	86909,3	
Q	Human health and social work activities	employees	1,0	2429,0	341,1	1316
		Labor productivity	144,0	3413514,2	134934,0	

Source: own elaboration and DANE

#### 4. Results for the OP decomposition in the service sector of Colombia.

Table 5 provides a detailed breakdown of the results of applying the Olley-Peck methodology at the subsector level ( 1-digit ISIC rev 4 sectors) for the service sector of Colombia using the EDITS 2019. The largest subsector in terms of sales and employment is Wholesale and Retail activities (columns (1) and (2)). As indicated above, this sector accounts for 27% of total employment (column (3)), followed by health and social work, which accounts for 22%. On the other end, Professional, Scientific, and Technical Services is the smallest activity measured by total sales and employment share (0.3%).

Column (4) shows the value of weighted labor productivity by sector (total sales over total employment). It confirms that capital-intensive activities like Electricity and Gas services have the most significant value of labor productivity, while labor-intensive activities like Accommodation and Food and Education have the lowest. As described in equation (1), columns (5) and (6) decompose the value of weighted productivity in terms of the unweighted labor productivity and the covariance terms between labor productivity and the labor share of each firm within each subsector. The average unweighted labor productivity across all subsectors is 0,395 thousand million pesos of december 2018 (around 123000 USD current dollars). This is higher than the average weighted productivity across subsectors (0,377 thousand million pesos or 117813 USD current dollars). This is unsurprising given that the average covariance term is negative (-0,018 thousand million pesos or 5625 current dollars). Thus, on average, resources across firms within the service's activities are strongly misallocated, meaning that low-productive firms are larger (demand more labor) compared to high-productivity units. As a consequence, compared to a situation where the average covariance term is zero, the mean value of the subsector's productivity is reduced by 4,2% (see last row column 8). When looking at this covariance term across subsectors, we observe a large dispersion: Wholesale and Retail services show the lowest value, -0,21 (lowering productivity by 37% compared to the 0 value covariance in this activity ), while Financial and Insurance Services have the highest positive indicators (0,07) which implies that resource allocation raises productivity in this subsector.

Of course, this counterfactual of assuming a random allocation across firms within subsectors (0 covariance) is rather conservative. One would expect this cov term to be positive, as the evidence shows for the manufacturing sector in Colombia (Alvarez et al. 2018). For example, if the average  $ct_j$  is 0.05 instead of -0,018, the average increase in productivity across all subsectors would be around 21739USD or 22%  $((123000 + 21793)/117813)$

Table 5. Productivity decomposition using the O-P (1996) methodology. By subsector. 2019

sector	sales* (1)	employees (2)	employment share (3)	weighed labor productivity $P_j^*$ (1)/(2) (4)	Unweighted labor productivity $\bar{p}_j^*$ (5)	Covariance term $ct_j^*$ (6)	$\bar{p}_j/P_j$ (7)	$ct_j/P_j$ (8)
Electricity, gas, steam and air conditioning supply	54152,93	38544	0,0190133	1,405	1,416	-0,011	1,008	-0,008
Water supply; sewerage and waste management	8860,92	56384	0,0278136	0,157	0,146	0,011	0,929	0,071
Wholesale and retail trade; repair of motor vehicles and motorcycles	260619,92	554158	0,2733599	0,470	0,645	-0,175	1,371	-0,371
Transportation and storage	27537,20	230127	0,1135191	0,120	0,117	0,002	0,980	0,020
Accommodation and food service activities	12427,48	138529	0,0683348	0,090	0,094	-0,004	1,044	-0,044
Information and communication	49749,68	234005	0,115432	0,213	0,338	-0,125	1,588	-0,588
Financial and insurance activities	103512,76	108642	0,0535919	0,953	0,883	0,070	0,927	0,073
Professional, scientific and technical activities	763,73	6865	0,0033864	0,111	0,100	0,012	0,894	0,106
Education	20905,15	210616	0,1038945	0,099	0,084	0,016	0,843	0,157
Human health and social work activities	70563,19	449340	0,2216544	0,157	0,131	0,026	0,831	0,169
total/average	609092,95	2027210	1	0,377	0,395	-0,018	1,042	-0,042

Values in columns (1), (4), (5), and (6) are expressed in thousands of millions of pesos of 2018; \* sales divided by employment; \*\* simple average of values in columns. Source: own elaboration and DANE.

Table 6 presents the productivity decomposition analysis for the aggregate service sector, applying equation (2) of the conceptual framework. Here, we incorporate a measure of what we may call the “misallocation” of resources across subsectors described by the CT term, which represents the covariance of labor productivity and labor shares along subsector service activities. The analysis for the aggregate service sector shows that misallocation across subsectors is greater than misallocation within subsectors (-0,077 vs -0,018, respectively).

Thus, the aggregate service sector's overall weighted productivity is now 0,300 thousand million pesos of december 2018 (around 93750 current dollars). We arrive at this result by deducting from the unweighted mean of firms' productivity across all subsectors (0,395 thousand million or 123000 dollars) the average covariance terms across firms within subsectors (-0,018 thousand million or 5625 dollars) and the already mentioned covariance terms across subsectors (-0,077 thousand million or 24066 dollars). Thus, now shutting down (making it zero) the two covariances terms, the overall productivity of the service sector of Colombia could rise by around 32% (29690USD), as indicated by Column (7).

Table 6. Productivity decomposition using the O-P (1996) methodology for the aggregate service sector. 2019.

Aggregate labor productivity service sector, Ps (1)	Covarianza across subsector, CT (2)	Average covariance within subsectors, $ct_j$ bar (3)	Average unweighted productivity within subsectors, $p_j$ bar (4)	CT/Ps (2)/(1) (5)	$ct\_bar/Ps$ (3)/(1) (6)	$p_j\ bar/Ps$ (4)/(1) (7)
0,300	-0,077	-0,018	0,395	-0,26	-0,06	1,32

Source: own elaboration and DANE

We should be cautious about how we look at this result. Interpreting the covariance term across subsectors as reflecting the efficiency level of resource allocation could be misleading. Demand shocks could explain, in part, the expansion of specific subsectors within the service

activities despite their relatively low level of efficiency in production. This is, for example, the case for tourism or personal services. Other studies that use this methodology for manufacturing (see, for example, Alvarez et al. 2018) also find a larger “misallocation” term across subsectors than that estimated within sectors. They argue that these other demand factors could play a more significant role when evaluating the allocation of resources across activities that aggregate multiple products or services, as we have done previously when using 1-digit service categories. Thus, it is easier to discuss misallocation when looking at resource distribution across firms within (narrowly defined) subsectors.

#### *OP Analysis with the panel datasample*

We have extended the OP analysis using the panel data sample covering 2010-2019 (see details below in section 6.2). This allows, first to check whether the findings for 2019 are representative of a more extended period. We may also see whether there is any observable trend in the value of the different components of the OP decomposition. Appendix B presents the details of the various tables. Overall, the following conclusions can be drawn in the case of within-sector analysis: (i) Looking at the average for the 2010-2019 years, the Wholesale and Retail sector is the one with the most significant level of misallocation of resources across firms, while the Financial and Insurance category is the most efficient activity (strongly negative and positive ctj indicators, respectively). This result resembles what we find for 2019; (ii) for the average ctj across sectors, there is no clear trend in the covariance term through the considered period. This is, in part, not surprising given that the number of participating firms increased over time, making year-to-year comparisons difficult; (iii) for the entire period, the average value of the covariance term across sectors is slightly negative, which implies that the allocation of resources did not help to increase productivity in the different activities. It reduced it (by one-third of a percentage point). We already saw that for 2019, this factor reduced overall productivity by around 4,2% points.

Table 7 below shows the OP decomposition for the 2010-2019 period, including the term associated with the covariance across subsectors, CT, in addition to the average across sectors of the covariance term, mean ctj, and the mean of the non-weighted productivity across firms within sectors, mean pj\_bar<sup>67</sup>. We see that the negative value for the CT term found for 2019

---

<sup>6</sup> We will explain in detail the characteristics of the panel database in section 6.2. Still, it is helpful to anticipate that there is no information on firms' sales for the years 2008-2009, which is why the OP analysis starts from 2010 onwards.

<sup>7</sup> As done by other authors (see Bartelsman et al., 2013; Alvarez et al., 2018), we have run the OP decomposition using log productivity instead of its level (see appendix B). This has the advantage that the estimated values, for example, the covariance ctj, can be interpreted in terms of elasticities. Nevertheless, log calculations tend to moderate the effect of significant disparities in productivity across firms and sectors, so even though qualitative calculations in logs show the same patterns as in levels, they differ in quantitative terms. For example, the Wholesale and Retail sector in 2019, as we found before, has the lowest value for the covariance term ctj, -0,13, which means that if the allocation of resources across firms were to be done randomly, productivity in this sector could increase by 13% (we found an increase of 37% in levels). The mean value of ctj across sectors for that year is 0,14, which implies that, on average, resource allocation across firms within the sector has increased productivity by around 14% (we found a decrease of around 4% when calculating in levels). The average of the ctj estimated in log for the period 2010-2019 is 0,19, which, again, suggests that the allocation of resources across formal firms is collaborating with increasing productivity in the service sector of Colombia. Nevertheless, it is helpful to compare this number for calculated for manufacture for the same country which has been estimated at 0.42 for years 2004-2007 (see Alvarez et al 2018), and 0,51 for the USA (Bartelsman et al., 2013). Thus, we may conclude that misallocating resources across firms is still a

(see Table 6) is observed in all years, and actually, it has a clear increasing trend. We have already mentioned that we should be careful about interpreting this result. In part, it could be driven by demand factors pushing the expansion of certain service activities (say tourism, education) with low productivity levels (see Table 6). We already mentioned that, in comparison, there is no visible trend in the average covariance terms within sectors (ctj). On the other hand, the simple average of productivity across firms and sectors (mean pj bar) shows a positive trend, implying that its value almost doubled between 2010 and 2019. The increase in this indicator of firm productivity allowed the overall weighted productivity of the entire service sector in Colombia (column 4) to increase by around 50% during the whole period despite a negative reallocation across subsectors and also a small or even negative impact (for some years) coming from the level of efficiency of resource allocation across firms within sectors. Though potentially interesting, as we indicated before, we should be cautious about this interpretation of results because, as we will show below, the sample of firms increased significantly during this period (from 5000 thousand in 2010-2011 to 9000 in 2008-2019).

Table 7. Productivity decomposition using the O-P (1996) methodology for the aggregate service sector. 2010-2019. In thousand million pesos of December 2018.

year	CT (1)	ctj_mean (2)	pj_bar mean (3)	P_s (4)
2010	-0,022	0,039	0,196	0,213
2011	-0,023	0,036	0,180	0,194
2012	-0,036	-0,042	0,236	0,158
2013	-0,035	-0,046	0,228	0,147
2014	-0,043	-0,014	0,347	0,290
2015	-0,056	-0,021	0,364	0,287
2016	-0,070	0,011	0,375	0,316
2017	-0,060	0,023	0,341	0,303
2018	-0,076	-0,010	0,377	0,291
2019	-0,077	-0,018	0,395	0,300

Source: Own elaboration and DANE

Notwithstanding the key role that misallocation could play in determining aggregate productivity within sectors or for the whole economy, the previous analysis showed that a key determinant of aggregate productivity in the service sector is productivity at the firm level. Thus, the innovation channel that explains the evolutions of productivity throughout the life cycle of firms plays a key role. In the estimation presented above, assuming that allocation across subsectors is left unchanged, improving the allocation of resources across firms within subsectors could bring significant gains, say, between 4% and 22%, but still, we are well behind the level of productivity reached by developed economies (Triplet and Bosworth, 2004). Thus, firms and governments must strive to improve productivity at the production unit level, and innovation plays a key role in this regard. We will discuss this in the next section.

---

critical issue in explaining low productivity in Colombia's service sector. Finally, the CT component, which estimates the "efficiency" of the allocation across service subsectors, as found for the calculation in levels, has a decreasing trends throughout the period with a value of -0,02 during 2017-2019. This is comparable with that estimated for manufacturing (around 0,02) and even the US (-0,03).

## 5. The innovation channel and firm-level productivity.

The productivity decomposition exercise of the previous section shows that productivity at the firm level is a key driver of sector aggregate productivity. The evidence for LAC suggests that productivity growth along the firm's life cycle is low compared to developed economies (Hsieh and Klenow, 2014). Thus, a focus must also be placed on innovation activities taken by the firm. In this section, we will use the EDITS surveys to characterize firms' innovation efforts in the service sector in Colombia. We will also assess to what extent the various government policies to support these actions by private firms have positively impacted innovation inputs like expenditures and labor allocation, and innovation outputs, like the launching of new goods and services, and especially the impact on firms' performance indicators like sales, employment, and labor productivity. As we have done with the productivity decomposition analysis, we will first characterize innovation activities done by firms for the year 2019 and then describe the behavior of some of these variables across the considered period.

### 5.1 Innovation Indicators for the year 2019.

A first issue that must be clarified is that, as it is common in the specialized literature and policy practice, we will adopt a broader approach to innovation performed by firms that go beyond R&D and include the acquisition of machinery and equipment, information and telecommunications technologies (e.g., software and hardware development), innovation in marketing methods, technology transfer (e.g., purchasing licenses), technical assistance and consulting, engineering and industrial design, and education and training.

Taking this broad definition, table 8 shows some basic statistics for the year 2019. The first result is that, from the total sample of 9304 firms, only 2366 (25,4%) displayed a positive level for innovation expenditures during that year. For this subsample of firms, the mean value of outlays for this purpose was 2312026 thousand pesos in 2018 (equivalent to 797257 current dollars), representing 2,4% of total sales. There is an important variation across this mean value, with some firms reporting very low values of innovation effort (say, 200000 pesos, less than 100 dollars) while at the top of the distribution, expenditures reached more than 85 million USD). This is also reflected in the minimum and maximum indicators of innovation expenditures over total sales.

Table 8. Descriptive Statistics. Innovation expenditures. Whole Sample, 2019

Variable	Total observations	Obs with Innexp>0	Mean	Std Dev	Min	Max
Innovation expenditures (Innexp)*	9,304	2366	2312046	13300000	200	246000000
Innexp/total sales			0,024	0,072	0,00000356	1,48

\*in thousand pesos of 2018. Source: own elaboration and DANE.

Table 9 shows the number of firms, total innovation expenditures over total sales, and the percentage of firms with positive innovation expenditures by subsector. As before we group service activities at the ISIC classification's main division (letter), which implies a 10-category disaggregation.

As expected, the Professional, Scientific, and Technical subsector is one of the sectors with the highest share of innovative firms as 68.18% of them reported positive innovation expenditures in 2019, which represented, on average, 30,5% of these firm's sales. This is unsurprising because its main activity is developing and providing innovation and technical services. Two other relatively innovative sectors are Education and Information and Telecommunication, with relatively high shares of innovative firms (65,8% and 37%, respectively) that spend on average 4,5% and 3,8% of their sales in these activities. For the rest of the categories, innovation efforts are very low in the retail sector, where only 17,8% of the firms report positive expenditures in these activities, and for these firms, these expenses are only 0,5% of their sales.

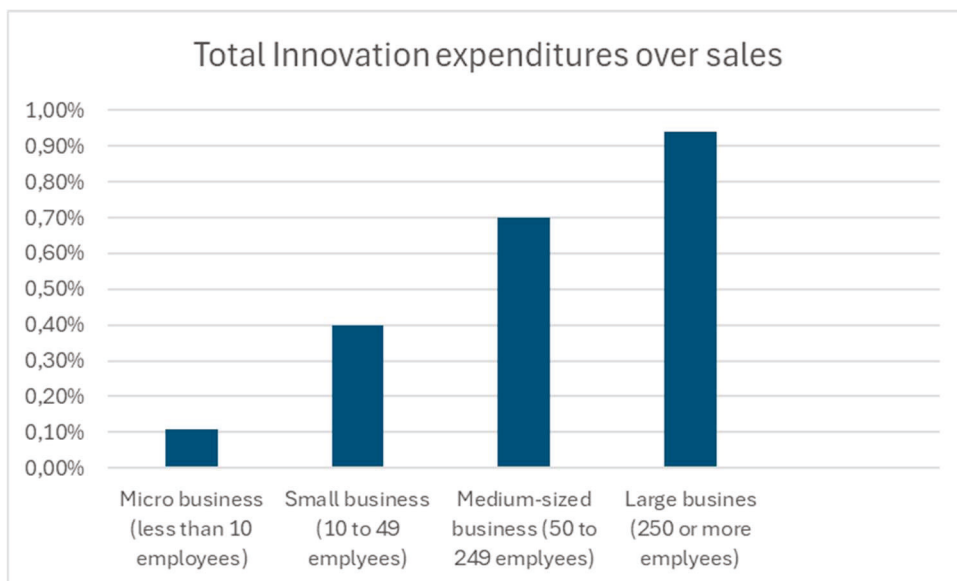
Table 9. Innovation expenditures by service subsectors. EDITS Survey. 2019

IC	Service Sector	Number of firms per sector	Expenditure in innovation / total sales	% of firms with positive innovation expenditures
D	Electricity, gas, steam and air conditioning supply	123	2,26%	48,7%
E	Water supply; sewerage, waste management and remediation activities	291	4,20%	36,20%
G	Wholesale and retail trade; repair of motor vehicles and motorcycles	4181	0,50%	17,81%
H	Transportation and storage	1582	2,82%	23,90%
I	Accommodation and food service activities	761	1,30%	25,80%
J	Information and communication	625	3,80%	36,80%
M	Professional, scientific and technical activities	44	30,50%	68,18%
Q	Human health and social work activities	1316	1,84%	32,20%
K	Financial and insurance activities	26	0,96%	80,77%
P	Education	258	4,50%	65,80%
Total/average		9207	5,27%	43,05%

Source: own elaboration and DANE

Beyond the distinction of innovation expenditures by service subsector, looking at the relationship between the firm's size and innovation outlays could be interesting. There is the presumption that because of market failures affecting the financing of these activities, small firms with fewer cash flows and/or lower levels of assets to put as guarantees for loans would spend less on R&D and other innovation initiatives. The data confirmed this hypothesis when we classify firms' size in micro, small, medium, and large, as shown in Figure 2.

Figure 2



Source: own elaboration and DANE

As we indicated before, there are various channels through which firms can improve the quality of their services or production processes. Table 10 shows the different innovation activities taken up by firms in each subsector, which go beyond R&D done within the firm or by external sources, including purchasing machinery and equipment, IT and software, marketing, licensing certain technologies, consulting support, industrial design, and managerial/training activities. Not surprisingly, R&D is most relevant in two subsectors: Professional, Scientific, and Technical Services and Higher Education. For the rest of the service's activities, Machinery and Equipment and IT are the most relevant channels for innovation. As we expected, marketing and training are important mechanisms that spur innovation in the case of Retailing.

Table 10. Innovation channels by subsectors (% of participation). 2019

ISIC	Sector	internal R&D	External R&D	Purchase of machinery and equipment	Information technologies, software and data analysis	Marketing	Licences	Consulting	Industrial Design	Training
D	Electricity, gas, steam and air conditioning supply	2,71%	18,30%	54,47%	14,65%	0,18%	0,74%	2,06%	3,65%	0,16%
E	Water supply, sewerage and waste management	70,13%	4,07%	11,98%	3,16%	0,11%	0,15%	1,16%	0,16%	0,17%
G	Wholesale and retail trade; repair of motor vehicles	4,57%	1,54%	32,79%	21,49%	8,63%	0,09%	6,99%	0,18%	12,31%
H	Transportation and storage	0,88%	0,23%	22,05%	23,01%	0,77%	0,11%	39,24%	0,32%	0,32%
I	Accommodation and food service activities	25,74%	0,06%	34,00%	15,05%	2,87%	0,02%	4,40%	1,53%	0,70%
J	Information and communication	10,29%	1,12%	25,04%	60,06%	0,89%	0,44%	1,33%	0,23%	0,14%
M	Professional, scientific and technical activities	65,17%	28,77%	2,58%	0,86%	0,00%	0,07%	1,20%	0,18%	1,05%
Q	Human health and social work activities	8,09%	1,19%	42,97%	19,29%	0,92%	1,40%	4,24%	1,27%	0,80%
K	Financial and insurance activities	20,13%	8,32%	24,14%	37,85%	4,44%	0,01%	4,13%	0,34%	0,38%
P	Education	46,72%	17,75%	5,31%	16,28%	1,55%	0,79%	2,01%	1,39%	2,00%
	Average	25,44%	8,14%	25,53%	21,17%	2,04%	0,38%	6,68%	0,92%	1,80%

Source: own elaboration and DANE

A final issue regarding innovation decisions by firms that we will examine is the sources of financing of these activities. Table 11 describes this information, distinguishing own resources, funds received from other firms within the same corporate group, public resources in the forms of grants, tax incentives, or subsidized loans, private bank loans, financing coming from clients or supplier firms, private capital resources like venture capital, and donations.

The evidence shows that own resources are the primary source of financing in most subsectors. This is not surprising given the market failures affecting access to finance due to

the impossibility of using intangible assets as collateral, asymmetric information about project details, and more significant uncertainty about potential outcomes. This is why public funds are vital in supporting R&D and innovation. We see this as the second most important source of financing, although there is a substantial disparity across subsectors<sup>8</sup>. Public support for R&D and innovation is relevant in the Professional, Scientific, and Technical subsector, as well as in Higher Education and Human Health (with a share of 77%, 23%, and 11%, respectively). On the other hand, government aid is a minor funding source for the rest of the activities. This is an expected result as these three subsectors generate more basic R&D and general knowledge, which are more subject to spillover effects, impairing incentives for private firms to use their own resources.

It is interesting to see that private capital in the service sector in Colombia is not a source of innovation funding, even though venture capital has been a critical mechanism for financing innovation and entrepreneurship in other countries.

Table 11. Financing sources of innovation by subsectors. 2019.

ISIC	Sector	Own resources	Funds from firms from same group	Public resources	Private banks	Clients or Suppliers firms	Private capital funds	Donations
D	Electricity, gas, steam and air conditioning supply	91,87%	0,00%	4,48%	0,00%	0,00%	0,00%	0,00%
E	Water supply; sewerage, waste management and remediation activities	87,07%	0,18%	4,10%	0,00%	0,00%	0,00%	0,00%
G	Wholesale and retail trade; repair of motor vehicles and motorcycles	93,89%	1,01%	0,18%	1,12%	0,15%	0,00%	0,54%
H	Transportation and storage	74,60%	0,01%	2,74%	0,00%	0,00%	0,00%	0,00%
I	Accommodation and food service activities	95,23%	0,00%	0,08%	0,00%	0,00%	0,00%	0,00%
J	Information and communication	97,99%	0,12%	0,93%	0,10%	0,10%	0,00%	0,00%
M	Professional, scientific and technical activities	18,74%	0,36%	77,12%	0,00%	0,30%	0,00%	2,61%
Q	Human health and social work activities	75,56%	3,09%	11,21%	0,42%	0,00%	0,00%	0,08%
K	Financial and insurance activities	100,00%	0,00%	0,00%	0,00%	0,00%	0,00%	0,01%
P	Education	67,78%	0,06%	22,92%	0,02%	1,56%	0,00%	2,25%
	Averages	80,27%	0,48%	12,38%	0,17%	0,21%	0,00%	0,55%

Source: own elaboration and DANE

Having analyzed innovation input indicators like expenditures, we now look at some innovation outcomes. One key objective of these activities is to increase productivity. In Table 12, we compare labor productivity in each subsector between firms that innovate (have positive innovation expenditures) and those that do not. If innovation expenditures successfully generate new services and/or improve production processes, we expect labor productivity to be more significant for firms that allocate resources to these activities. The evidence shows that this is only sometimes the case. Innovation efforts positively correlate with productivity for some activities, like Water Supply and Waste Management Professionals, Scientific and Technical Services, Human Health, and Education. One reason that may explain this is that these subsectors have the highest innovation expenditures (see Table 8). In any case, these are straightforward nonconditional comparisons. We know that many other factors affect productivity beyond the potential positive impact of innovation when successful in generating new products and production processes.

<sup>8</sup> In 2019, only 131 firms received some government support. Section 6 provides a more detailed description of these programs.

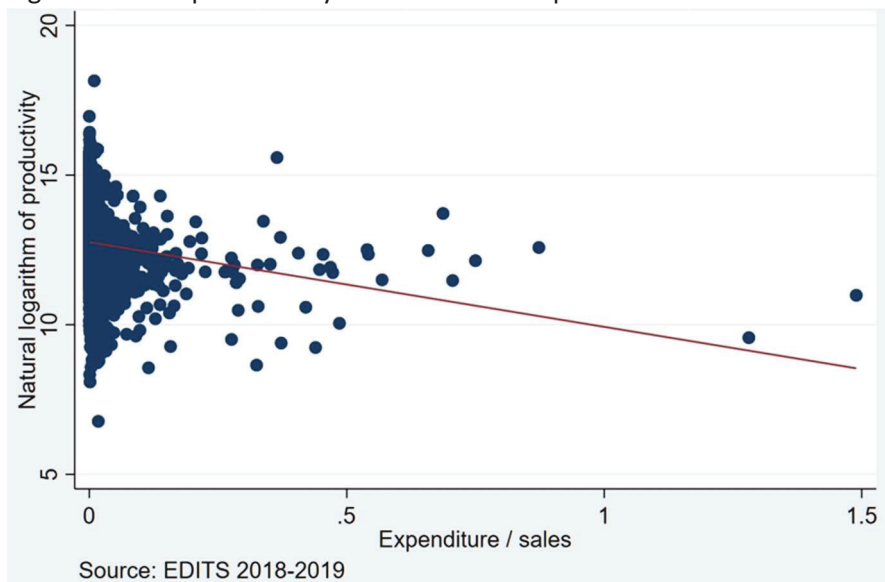
Table 12. Innovation expenditures and labor productivity by subsector. 2019.

ISIC	Sector	Average productivity of firm with positive innovation expenditures* (1)	Average productivity of firm with zero innovation expenditures* (2)	(1)/(2)
D	Electricity, gas, steam and air conditioning supply	\$ 3.191.163	\$ 3.872.829	0,82
E	Water supply; sewerage, waste management and remediation activities	\$ 441.693	\$ 224.911	1,96
G	Wholesale and retail trade; repair of motor vehicles and motorcycles	\$ 1.184.018	\$ 1.333.687	0,89
H	Transportation and storage	\$ 207.480	\$ 217.447	0,95
I	Accommodation and food service activities	\$ 186.891	\$ 196.780	0,95
J	Information and communication	\$ 699.312	\$ 854.291	0,82
M	Professional, scientific and technical activities	\$ 230.286	\$ 138.022	1,67
Q	Human health and social work activities	\$ 301.037	\$ 254.514	1,18
K	Financial and insurance activities	\$ 1.767.690	\$ 2.011.121	0,88
P	Education	\$ 189.964	\$ 140.162	1,36
Average		\$ 839.953	\$ 924.376	0,91

Source: own elaboration and DANE

We can also evaluate the relationship between innovation effort and labor productivity using the whole sample of firms with positive innovation expenditure levels (2364 firms out of 9300) without distinguishing which sector each firm belongs to. This analysis is presented in Figure 5. Somewhat consistent with the results we describe in Table 8, we find a negative correlation (in an OLS regression, the coefficient of innovation expenditures w.r.t log productivity is -2,8 and significant at 1%).

Figure. 3 Labor productivity and innovation expenditures. 2019



We should be careful how we interpret this result. This is an unconditional correlation, which, of course, should not be understood as a causal relationship. In the following subsection, taking advantage of the panel structure of the EDITS survey, which allows the introduction of fixed effects at the firm level, we will analyze the impact of public policies to support innovation and the causal link between these policies and productivity.

## 5.2 Descriptive analysis of the panel dataset

We will now turn to the description of the panel dataset covering the 2008-2019 period we will use in the regression analysis. Figure 4 shows the number of firms participating in each survey edition. We observe an increase in the scope of participating firms, going from around 3500 firms in 2008-2009 to more than 9000 enterprises in 2018-2019.

Figure 4



Figure 5 shows the number of firms with positive innovation expenditures each year and their share in the total surveyed enterprises. At the beginning of the period, and when the sample of firms was relatively small, there was a relatively high share of innovative firms (40%), but this changed after 2012-2013 when the proportion of firms with positive innovation outlay converged to values around 20-25%.

Figure 5

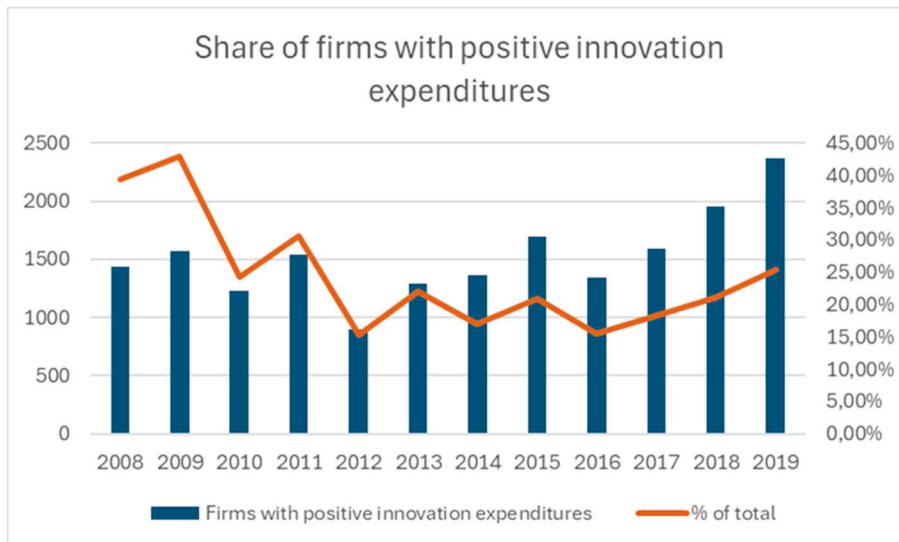
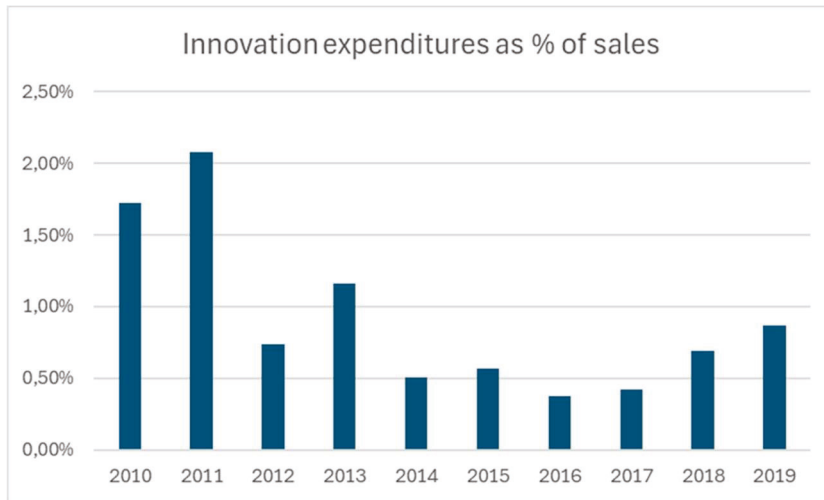


Figure 6 describes the evolution of expenditures as a share of total sales for the whole sample of firms in each year. We show the value for the complete sample, which includes

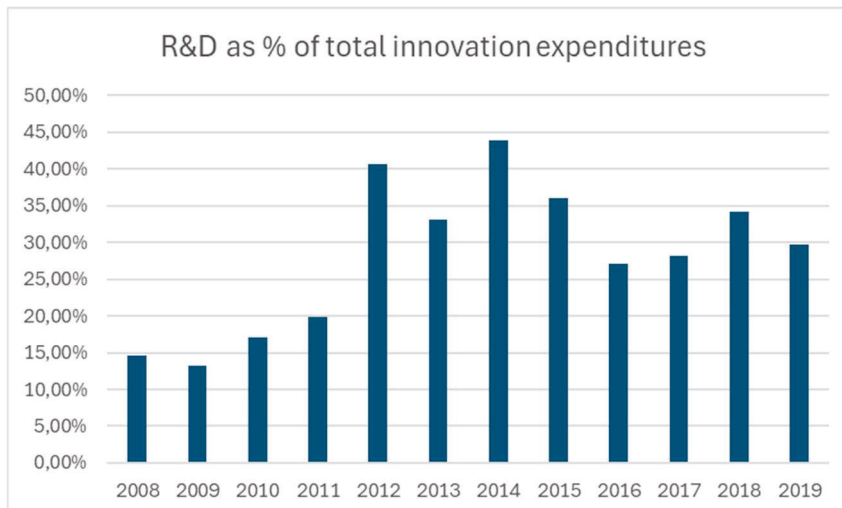
noninnovative firms. As the number of firms in the sample increases, the average value of expenditures over total sales decreases, approaching values between 0,5 and 1% in 2018 and 2019.

Figure 6



As mentioned, we are using a broad measure of innovation that goes beyond traditional research and development activities. This measure encompasses capital and intermediate good purchases associated with improving product quality or production process, engineering and design, adoption of new hardware and software systems, technology licensing, new management practices, and education and training. As shown in Figure 7, these other innovation channels have maintained a key participation in total innovation outlays (between 60 and 70% since 2012-2013).

Figure 7



A final issue regarding inputs that firms allocate to innovation activities is related to the number of employees dedicated to these activities. Table 13 shows that, as the number of firms surveyed increased, the total employment went up from 700 thousand workers in 2008 to around 2 million in 2019. On the other hand, the number of workers performing innovation activities increased but in a lower proportion (from almost 40 to 54 thousand), reducing the share of human capital allocated to R+D+I (from 5,5% to 2,67%). This result is explained by the

fact that in later years, the survey extended its sample to firms in sectors with lower innovation intensity, like retail.

Table 13. Employment in R+D+i activities

Year	2008	2009	2010	2011	2012	2013	2014	2015	2016	2017	2018	2019
Total employment	695875	723636	945050	998262	1383788	1450131	1655641	1713112	1866000	1887871	1985042	2027249
Workers in R+D+I	39644	45758	35244	39606	33066	44546	50433	62624	34484	41034	48060	54080
Share	5,70%	6,32%	3,73%	3,97%	2,39%	3,07%	3,05%	3,66%	1,85%	2,17%	2,42%	2,67%

As we did for innovation input variables, we can also look at the temporal pattern of innovation outcomes. For example, one variable of interest is the number of innovations that firms report. These innovations include new or improved goods or services, production processes, marketing, or management methods. We added all these categories into one index, defined as the sum of all these items. Table 14 presents the results computed for biannual periods, corresponding to each survey edition, starting in 2010 (no data for 2008-2009)<sup>9</sup>. The first thing that should be mentioned is that very few firms introduce innovations. In most biannual periods, innovative enterprises (now defined as those that introduce new or improved goods, services, or production processes) are below 1% of the sample. On the other hand, the average number of innovations seems large for those that innovate: between 20 (2014-2015) and a maximum of almost 60 in 2012-2013.

Table 14. Goods, services, and production innovations

Mean innovations/Year	2010-2011	2012-2013	2014-2015	2016-2017	2018-2019
mean # of innovations	52	59,4	20,46	33,9	32,5
Number of firms that report innovations	72	9	24	20	42
% of total firms in the sample	1%	0,15%	0,30%	0,23%	0,45%

An innovation output that is critical to evaluate is whether these efforts at innovation produce changes in productivity at the firm level. We have looked at labor productivity indicators for 2019 by sector (see Table 12). Here, taking advantage of the panel dataset, we can look at the evolution of this variable across time (since 2010) and make a distinction, as we did in Table 12, between firms that report positive innovation expenditures each year and those that do not<sup>10</sup>. Figure 8 presents this exercise evaluating the mean value of labor productivity (total sales over total employment by firms) at 2019 prices.

The first thing that comes as a relative surprise is that labor productivity is, on average, lower for those enterprises that report positive innovation expenditures. This result was partially advanced when looking at labor productivity for the year 2019 by sector (see Table 12). Here, we show that this pattern is repeated for all years for the simple mean across firms (not distinguishing across industries). For both groups of firms, mean labor productivity follows a similar time pattern, being relatively constant during the first years of the considered period, then increasing sharply between 2013 and 2015 (more significantly for zero innovation expenditures enterprises), and then declining and stabilizing at higher values than the first years in 2016-2019<sup>11</sup>. Of course, as we indicated before, these are just descriptive indicators

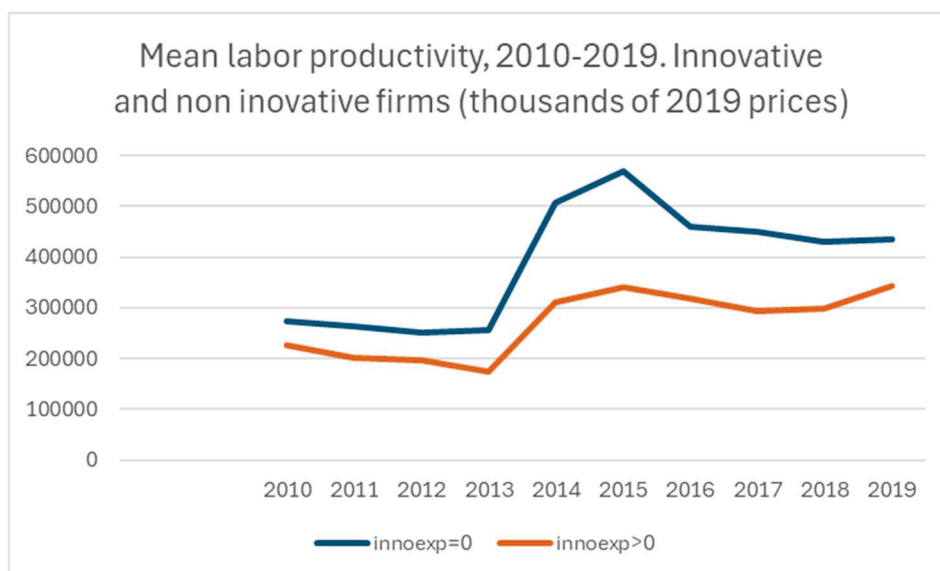
<sup>9</sup> The survey collects information asking firms for the total number of innovations introduced in these two-year periods.

<sup>10</sup> The data on labor productivity can be calculated only since 2010, given that there is no information on total sales by firms for 2008 and 2009.

<sup>11</sup> We should be cautious about the conclusions drawn from the trends shown in this Figure. To maintain confidentiality, the data available through the DANE web page have imposed certain upper limits for some variables, such as sales applied for some very large firms within specific sectors. This could bias

that in no way suggest any causation between innovation and productivity. We turn to this causal analysis in the following section.

Figure 8



Source: own elaboration and DANE.

## 6. Impact estimation of innovation public support programs

In this section, we want to use the panel data set to estimate the impact of public support policies on innovation indicators by private firms in Colombia's service sector. First, it will be helpful to briefly describe the various institutions and programs that the government has established to that end.

### *Public programs to support private investment in R+D+i*

DANE classifies these government programs into two categories: grants that cofinance innovation activities and concessional credit lines.

Various programs run by different public sector institutions have established co-financing grants to promote innovation by private firms. One such institution is **MinCIT** (for Ministerio de Comercio, Industria y Turismo). Within this Ministry, there are various programs like INNpulsa that launch calls mainly to support technical assistance, though, on some occasions, they also provide seed capital for start-ups. Another program within the MinCIT is “Colombia Productiva,” which also supports technical assistance but focuses on fostering productivity within firms and helping them reach external markets by obtaining international certifications. Finally, MinCIT has a program coordinated with Bancoldex, Colombia's leading public development bank, which offers grants for cofinancing entrepreneurial training and consulting services.

---

downward the variation for the labor productivity variable, which may more significantly affect innovative firms. We plan to correct this bias by accessing the uncensored data set.

Another critical institution that supports innovation, again taking a broad definition of this concept, is SENA (Servicio Nacional de Aprendizaje). It offers free professional training programs and a program called “Fondo Emprender,” which offers seed capital to entrepreneurs trained by the institution. It has also established the program “SENAINnova,” which cofinances inputs, technological services, and equipment purchases for projects to improve product quality or production processes.

A third crucial public institution in Colombia's innovation ecosystem is **MinCiencias** (for Ministerio de Ciencia Tecnología e Innovación, former ColCiencias). This institution runs a program that grants tax subsidies (through competitive calls) to firms that usually perform I+D+i activities and asks for public support for specific innovation projects.

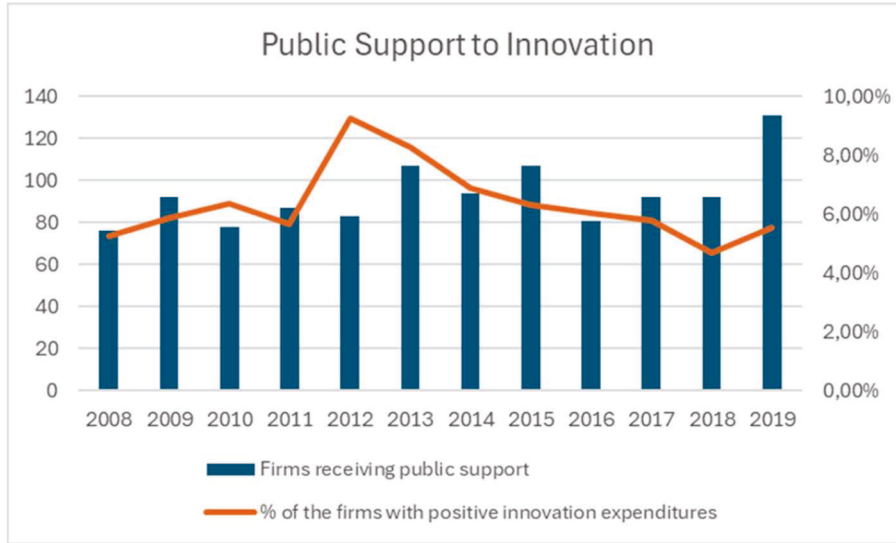
Finally, there is the **MinTIC** (Ministerio de la Información y las Comunicaciones), which provides funds (via the SETIC program) to finance I+D+i projects presented by firms or research groups considered strategic, usually belonging to the TIC sector.

Regarding credit lines, in Colombia, the majority of concessional loans offered by the public sector for financing I+D+i activities are provided by **Bancoldex**. They are managed directly by the Bank and through a network of “strategic partners” like private commercial banks, other national and subnational corporations, and cooperatives. These are credit lines (up to a given amount) with interest rates and maturity periods more favorable than those offered by private institutions. For example, in 2023, while the market interest rate was above 20%, Bancoldex offered credit lines charging rates between 10 and 18%. Other products offered by Bancoldex that could be relevant for financing innovation activities are leasing, leaseback, factoring, and guarantees. Regional and local governments, such as Departments and Municipalities, and those from the mineral and oil royalties system, also manage credit lines.

The relative importance of these various public support programs for R+D+i changes yearly. For example, in 2019, out of around 23 billion pesos in total support for services activities (about 8 million dollars), co-financing programs represented almost 85% of the funding.

Notwithstanding the effort made by the public institution to support I+D+I activities in Colombia, Figure 9 shows that they reach a small number of firms in the service sector. In 2008, less than 80 enterprises participated in any of the abovementioned programs, which went up to 130 in 2019. This represents a tiny share of the firms that participated in the EDITS survey (between 1 and 1,2%). Also, as shown in Figure 9, these participating firms are about 6% of those that report positive innovation expenditures at the end of the period.

Figure 9



Source: own elaboration and DANE

### Estimation

Since the subsidies and concessional loans described above are not granted randomly, beneficiaries may differ from non-beneficiaries due to selection bias. Beneficiaries are more likely to be innovative (invest in R&D and other innovation activities) and more productive than non-beneficiaries. Therefore, beneficiaries would show different outcomes than non-beneficiaries even without program support. A major advantage of using longitudinal firm-level datasets is that they allow to account for constant unobservable factors that may affect both the outcome of interest and participation in the program. In particular, we will use panel fixed effects to run a difference-in-difference (DID) estimation methodology. We will run this estimation for the whole sample and use more restricted versions to ensure that the control groups (those firms that do not participate in the programs) are similar to those that are treated. In one version of these exercises, we apply propensity score matching (PSM) to identify a common support for treated and not-treated firms. The equation that we will estimate has the usual specification of DID format,

$$Y_{i,t} = \tau D_{i,t} + \delta_t + \mu_i + \delta_t * \mu_j + e_{i,t} \quad (1)$$

$Y_{i,t}$  is the variable of interest (innovation inputs and output indicators),  $D_{i,t}$  is the treatment variable that indicates if a firm receives public support,  $\delta_t$  is a time dummy,  $\mu_i$  is a firm-fixed effect,  $\delta_t * \mu_j$  is an interaction term between a sectoral and the year dummy, and  $e_{i,t}$  is an error term. Firm fixed effects  $\mu_i$  fully absorb any permanent heterogeneity at the firm level, and  $\delta_t$  represents yearly shocks that affect all firms. Regarding the interaction term,  $\delta_t * \mu_j$ , it

controls for time-specific shocks that affect the productivity of all firms in subsector  $j$ <sup>12</sup>. As indicated  $e_{i,t}$  is the usual error term assumed to be uncorrelated with  $D_{i,t}$ . In this case, the identifying assumption is independence of treatment status and potential outcomes, conditional on time-invariant unobservable and observable factors and time-varying observable confounders<sup>13</sup>.

We will use alternative specifications for the treatment variable  $D_{i,t}$ . First, we will define it as the total monetary amount the firm receives from the various public support programs. In this estimation,  $D_{i,t}$  helps to test whether public support crowds in or out private funds allocated to R+D+i. The literature has widely analyzed this hypothesis (i.e., Sanguinetti, 2005; Benavente et al., 2007; Lopez et al., 2010; Crespi et al., 2011; Rocha, 2015; Aboal y Garcia, 2016; Chavez, 2020).

Alternatively, we can define  $D_{i,t}$  as equal to one starting when the firms first participate in any public support program and zero otherwise. This alternative definition will be used especially when evaluating public support's impact on input or output indicators that could take some time to change, like labor allocated to innovation activities or firms' productivity<sup>14</sup>.

Before running the regression, it could be helpful to describe the panel structure of the data (see A.3 in the appendix). We will work with an unbalanced panel of 12603 firms that appeared in at least two years between 2008 and 2019. This last case amounts to 19.2% of the firms (2426); on the other extreme, 17.4% (2198) were observed during the entire period. The time pattern varies considerably, though a trend can be identified where more observations are placed in the second half of the considered period, coinciding with the fact that the survey sample was expanded in those years. As we will see below, the number of observations for the complete sample is 81116, which comprises, as indicated, 12603 firms that, on average, appear in 6,4 years.

Table 15 describes the descriptive statistics for a few variables of the treated firms (they receive public support in at least one year) and control firms (never participated in any public support program). We see that treated firms have a much higher level of innovation expenditures (1.16 million dollars against 81.6 thousand dollars)<sup>15</sup>, are much larger in terms of the number of employees (817 against 183 workers), report substantially higher levels of total sales, and have a more qualified labor force as measured by the share of workers that have a university degree.

---

<sup>12</sup> As indicated before (see footnote 2), the inclusion of these interaction terms is quite relevant for estimation outcomes. The paper by Albis et al., 2017 does not control for these outcomes which could bias their results.

<sup>13</sup> Another key identifying assumption is that outcome trends between treated and control groups would be the same without the R&D support programs. Although testing for this during the treatment period is impossible, in a revised version of the paper, we will explore whether both groups exhibit similar trends in the variable of interest during the pre-treatment period. Following Castillo et al. (2019) and Cerulli and Ventura (2019), we will include in Eqs. (1) lags and leads of the treatment variable.

<sup>14</sup> This definition of the treatment variable is used, for example, in the paper by Crespi et al., (2020) when looking at the impact of the FONTEC/FONDEC programs on productivity in Chile.

<sup>15</sup> We use the exchange rate of 2900 pesos per dollar to express these amounts in current dollars.

Tabla 15. Mean values. Basic characteristics of treated and non-treated firms

Treated firms	Obs	Mean
Innovation expenditures*	3.836	3385054
Total sales*	3.362	173000000
Employees	3.386	817
University degree workers	3.382	214
Non treated firms	Obs	Mean
Innovation expenditures*	77.280	236747
Total sales*	70.430	45400000
Employees	77.280	183
University degree workers	77.280	38

\*In thousand pesos of 2018. Source: own calculations and DANE

These findings suggest that even when we include fixed effects that control for constant unobservable and observable characteristics that may affect selection into the program and firm's performance variables, we should be careful about selecting the sample when running the regressions, trying to include within the control groups firms that are as close as possible to those that are treated.

Table 16 shows the first set of regressions that tries to evaluate the impact of public support on total innovation expenditures by firms. The treatment indicator is the amount a given firm  $i$  receives in year  $t$ . As indicated, this analysis aims to evaluate to what extent public funds crowd in (the estimated coefficient should be greater than one) or crowd out (coefficient less than 1) own private funds allocated to innovation by the firm. As explanatory variables, we add the size of the firms represented by total employment and a set of year-sector interaction dummies.

We run the regressions using four alternative samples to address the mentioned issue of differences in observable characteristics of treated and control firms: (i) the baseline estimation that includes the whole sample, (ii) innovative firms (those with positive expenditures in at least one year), (iii) we use Propensity Score Matching (PSM) to construct a common support for treated and control firms so that the estimated probability of receiving public funds of the latter equals the range of that of the former<sup>1617</sup>; (iv) a much-reduced set of firms that included those that obtained funds from public support programs in at least one year during the considered period.

We see that across all samples, the estimated coefficient for the public support variable is very similar, around 2.06 (raises to 2,23 in the PSM estimation), and statistically significant at 1%. The estimated coefficients suggest that one peso of public support produces more than two

<sup>16</sup> To apply the PSM method, we used data from 2008, dropping firms treated that year. We estimate a probit model using the total number of employees, total sales, total innovation expenditures, number of university-educated workers, and sector dummies as controls. The dependent variable is a dummy indicating whether the firms received public support in any year between 2009 and 2019. See details of the Sata-do file and the first stage regression in Appendix C.

<sup>17</sup> Other papers that have combined DID/FE regressions with PSM to evaluate the impact of public support programs for innovation are Crespi et al, 2011; Belmar, 2014; Aboal and Garcia, 2016; and Crespi et al., 2020.

pesos of private innovation expenditures. This implies a solid *crowding in* response from firms that participated in the public programs. There are no significant differences in the estimated coefficients when introducing the sector-year dummies. The public fund variable explains between 6% and 24% of the variation in innovation expenditures across firms, where the higher explanatory power corresponds to the regressions using the more restricted sample of firms.

Table 16. Regressions in levels. Total Innovación expenditures

Dependent variable: Innexp	Whole sample		Innexp>0		PSM		Psupport>0	
Di,t= amount of funds received	2.06	2.05	2.06	2.04	2.23	2.23	2.06	2.04
z	9.77***	9.65***	9.77***	9.649.64***	6.76***	6.71***	9.76***	9.77***
Sector-Year Dummies	NO	YES	NO	YES	NO	YES	NO	YES
R2								
within	0.0446	0.0664	0.0446	0.0701	0.0388	0.0701	0.1308	0.2122
between	0.3657	0.3016	0.3677	0.3003	0.1294	0.0094	0.5965	0.5664
Overall	0.1069	0.0668	0.1061	0.0652	0.0722	0.0044	0.2402	0.2334
Number of obs	80,396	80,396	46,468	46,468	32,612	32,612	3,824	3,824
number of groups	12,484	12,484	5,862	5,862	3,341	3,341	421	421
Obs per group								
Min	2	2	2	2	2	2	2	2
Max	12	12	12	12	12	12	12	12
Average	6,4	6,4	7,9	7,9	9,5	9,5	9,1	9,1

Source: Own calculations and DANE.

Table 17 confirms the strong response of innovation expenditures to changes in public support when running the model in logs. As in the previous regressions, the estimated elasticity for public funds varies very little across samples. It implies that a 1% increase in public support raises private firms' innovation expenditures by around 0.47%. As we restrict the sample to more similar firms, the overall R2 rose, and this is in part driven by the increase in the Within indicator, showing that identification comes in a more significant proportion by variation across time in firms' participation in public support programs as a consequence of the increase in the average number of time observations reported for each firm. As in the previous analysis, these results do not change much when we add time-sector dummies.

Table 17. Regressions in logs. Total Innovación expenditures

Dependent variable: log (Innexp)	Whole sample		Innexp>0		PSM		Psupport>0	
Di,t= log amount	0.49	0.47	0.49	0.47	0.48	0.46	0.49	0.47
z	24.56***	24.18***	24.56***	23.87***	18.65***	17.91***	24.53***	23.89***
Sector-Year Dummies	NO	YES	NO	YES	NO	YES	NO	YES
R2								
within	0.0208	0.0685	0.0208	0.0888	0.0238	0.0938	0.2327	0.2982
between	0.1263	0.0887	0.1293	0.1439	0.1688	0.0426	0.3162	0.2560
Overall	0.0745	0.0918	0.0658	0.1116	0.0780	0.0890	0.2492	0.2671
Number of obs	80,396	80,396	46,468	46,468	32,612	32,612	3,824	3,824
number of groups	12,484	12,484	5,862	5,862	3,435	3,341	421	421
Obs per group								
Min	2	2	2	2	2	2	2	2
Max	12	12	12	12	12	12	12	12
Average	6,4	6,4	7,9	7,9	9,5	9,5	9,1	9,1

Source: own calculations and DANE.

So far, we have analyzed regressions for the input variable defined as total innovation expenditures. It could also be interesting to look at another input indicator: the allocation of workers to innovation activities. This is another variable that the innovation survey reports. We have already shown (see Table 14) that the number of employees in these activities grew for

the total sample of firms between 2008 and 2019. Table 18 shows the regression result where the treatment variable, D<sub>it</sub>, now equals one starting in the year the firms first participated in any public support program and zero otherwise. We adopted this specification because it could take time for the firm to adjust the number of workers it allocates to innovation tasks, mainly if this implies hiring new personnel, given that there are contract and search costs that the firm should pay.

Table 18 presents the regression results when running the model in logs. As we can see, participation in the support programs also expands the number of employees attached to innovation activities. It increases it by 69% in the case of the model using the whole sample of firms. It goes up 75% when the sector-year dummies are included. This value for the estimated coefficient is equivalent to an average increase of around 13 employees in the labor force performing innovation activities. These results are similar when we run the estimations using the other alternative samples, except when we restrict observations to firms that received support at least one year during the considered period and introduce sector-time dummies. In this specification, the increase in labor allocated to innovation activities more than doubled (the estimated coefficient is 1.10).

Table 18. Regressions for (log of) Innovation workers

Dependent variable: log (Innworkers)	Whole sample		Innexp>0		PSM		Psupport>0	
Participation Dummy (D <sub>it</sub> )	0.69	0.75	0.69	0.77	0.56	0.69	0.69	1.10***
z	7.44***	8.27***	7.43***	8.41***	4.70***	5.78***	7.43***	10.31***
Sector-Year Dummies	NO	YES	NO	YES	NO	YES	NO	YES
R <sup>2</sup>								
within	0.0057	0.0332	0.0057	0.0432	0.0038	0.0402	0.0365	0.1360
between	0.2012	0.0956	0.1814	0.1001	0.2287	0.0668	0.0808	0.0008
Overall	0.1214	0.0753	0.1015	0.0769	0.1167	0.0650	0.0497	0.0191
Number of obs	80,396	80,396	46,468	46,468	32,612	32,612	3,824	3,824
number of groups	12,484	12,484	5,862	5,862	3,341	3,341	421	421
Obs per group								
Min	2	2	2	2	2	2	2	2
Max	12	12	12	12	12	12	12	12
Average	6.4	6.4	7,9	7,9	9.8	9.8	9.1	9.1

Source: Own calculations and DANE.

We can now analyze public support's impact on innovation outcomes. We will first consider an indicator of new or improved goods or services and/or the implementation of new production, management, and marketing methods. As explained before, this indicator is defined as the sum of all these innovation outcomes that firms report in the survey.

Table 19 shows the result of the regressions for the indicator associated with the number of innovations reported by firms expressed in log terms<sup>18</sup>. When looking at the whole sample, participation in public support programs increases the number of innovations by around 10% when the sector-year dummy is included (18% without checking for the sector-year effects). When using other samples, the results change slightly, except for the much more limited database that considers only firms that received public funds where the estimated coefficient is 0.05 and not statistically significant.

The estimated elasticities are relatively small values if we compare them with the previously estimated impact on some input variables like innovation expenditure or labor dedicated to

<sup>18</sup> As this variable is calculated as the total innovations done by firms for biannual periods corresponding to each edition of the innovation survey, we have allocated half of the reported number of innovations to each year.

these tasks. We can partially explain these relatively weaker impacts because the number of innovations firms report is recorded for biannual periods, so there is less variability in the explanatory variable (see footnote 16). In any case, we also saw (Table 14) that compared to innovation expenditure efforts, the proportion of firms that report actual innovation results is a tiny part of the sample (less than 0.5 in all years).

Table 19. Regressions for (log of) the number of innovations by firms

Dependent variable: log (loginnovations)	Whole sample		Innexp>0		PSM		Psupport>0	
Participation Dummy (Dit)	0.18	0.10	0.18	.10	0.19	0.12	0.18	0.05
z	4.07***	2.74***	4.07***	2.76***	3.40 ***	2.39**	4.06 ***	1.16
Sector-Year Dummies	NO	YES	NO	YES	NO	YES	NO	YES
R2								
within	0.0042	0.0461	0.0042	0.0496	0.0050	0.0640	0.0114	0.1411
between	0.0946	0.0059	0.0972	0.0081	0.1248	0.0129	0.0345	0.0887
Overall	0.0551	0.0034	0.0546	0.0035	0.0569	0.0084	0.0176	0.0029
Number of obs	80,396	80,396	46,468	46,468	32,612	32,612	3,824	3,824
number of groups	12,484	12,484	5,862	5,862	3,341	3,341	421	421
Obs per group								
Min	2	2	2	2	2	2	2	2
Max	12	12	12	12	12	12	12	12
Average	6.4	6.4	7,9	7,9	9.8	9.8	9.1	9.1

Source: Own calculations and DANE.

We now look at different indicators of firms performance. The ultimate objective of innovation activities within firms is to expand sales, employment, and productivity. Table 20 shows the results for the (log of) sales variable. For most estimations, the participation dummy positively and significantly affects the total firm's sales. Nevertheless, the estimated impact size depends on whether sector-year interactive effects are included. When they do, firms participating in innovation support programs increase sales by around 82% compared to those that do not. However, when controlling for temporary shocks affecting different activities, we see that sales in participating firms increase much less, around 22%, and in the case of the PSM estimation, this goes down to 18% and is very marginally significant. These results call attention to including these potential cofounders in the impact evaluation of government programs to support firms as sectoral changes due to shocks in demand and supply (technology) could be relevant factors affecting firms' sales within specific activities.

Table 20. Regressions for (log of) sales

Dependent variable: log (sales)	Whole sample		Innexp>0		PSM		Psupport>0	
Participation Dummy (Dit)	0.82***	0.24***	0.82***	0.22***	0.85***	0.18	0.82***	0.23***
z	11.01	3.09	11.01	2.81	7.48	1.53	11.00	2,74
Sector-Year Dummies	NO	YES	NO	YES	NO	YES	NO	YES
R2								
within	0.0063	0.1739	0.0099	0.2154	0.0070	0.1742	0.0881	0.2495
between	0.0071	0.0319	0.0045	0.0205	0.0045	0.0316	0.0011	0.0224
Overall	0.0083	0.0234	0.0066	0.0290	0.0048	0.0250	0.0062	0.0473
Number of obs	73,262	73,262	41,516	41,516	25,936	25,936	3,354	3,354
number of groups	12,392	12,392	5,831	5,831	3,199	3,199	419	419
Obs per group								
Min	2	2	2	2	2	2	2	2
Max	10	10	10	10	10	10	10	10
Average	5,9	5,9	7,1	7,1	8.1	8,1	8	8

Source: Own calculations and DANE.

Another key performance variable that we want to examine is total employment at the firm level. A key development feature of Latin America is the spread of informal labor, so it would be interesting to see whether innovation activities taken by firms fostered by government

subsidies could help expand formal employment. Table 21 shows the results. Similarly to the previous sales results, program participation positively and significantly affects the total firms' employment. Still, the estimates vary greatly depending on whether we control for temporal sector-level shocks. When these controls are included, the estimated elasticity goes from 43% to 20%. The change in the estimated elasticities is of lower absolute magnitude than those observed for sales. This may be because labor for various reasons (among them, regulation) is a more sticky variable.

Table 21. Regressions for (log of) employment

Dependent variable: log (employment)	Whole sample		Innexp>0		PSM		Psupport>0	
Participation Dummy (Dit)	0.43***	0.19***	0.43***	0.16***	0.47***	0.18***	0.43***	0.20***
z	9.25	4.01	9.25	3.44	7.40	2.75	9.24	3.62
Sector-Year Dummies	NO	YES	NO	YES	NO	YES	NO	YES
R2								
within	0.0048	0.0628	0.0073	0.0926	0.0067	0.0876	0.0643	0.1342
between	0.0311	0.0018	0.0277	0.0108	0.0222	0.1083	0.0005	0.0013
Overall	0.0301	0.0028	0.0262	0.0157	0.0210	0.0492	0.0037	0.0041
Number of obs	80,480	80,480	46,473	46,473	32,629	32,629	3,824	3,824
number of groups	12,559	12,559	5,865	5,865	3,349	3,349	421	421
Obs per group								
Min	2	2	2	2	2	2	2	2
Max	12	12	12	12	12	12	12	12
Average	6,4	6,4	7,9	7,9	9,7	9,7	9,1	9,1

Source: Own calculations and DANE.

Finally, we may want to examine labor productivity. Innovation and technology activities within the firm affect the productivity of all factors used in production, called Total Factor Productivity, TFP. Recent models (see, for example, Hsiew and Klenow, 2014) that assume firm heterogeneity due to TFP's differences, as well as monopolistic competition, predict that, in equilibrium, (log of) labor demand and (log of) total production/sales are proportional to (log of) TFP. The proportionality factor depends on the elasticity of substitution,  $\sigma$ , across products within a given industry ( $\sigma-1$  for labor demand and  $\sigma$  for the case of total production). We already found evidence consistent with these theoretical predictions in the case of sales and employment. In this theoretical framework, (log of) labor productivity also depends positively on changes in TFP (in this case, with a coefficient of one).

Table 22 shows the econometric estimation for labor productivity. As expected, the results reflect in part the results we already observed for sales and employment. First, we should recall that the total number of observations is reduced because, as indicated previously, there is no sales data for 2008 and 2009. Second, results differ markedly depending on whether we control for sector-year interaction dummies as was the case for sales (see Table 20). When these variables are omitted, the coefficient for the program participation variable suggests that public support raises productivity by around 45%, and this result does not change across different samples. This is because as predicted by theory, the estimated elasticity of changes in TFP on sales (in turn produced by innovation spurred by public support) is much higher than that on employment. Nevertheless, this result is found when sector-year interactions are not controlled for. When they are taken into account, we find that an essential part of the effect of program participation on sales is due to time-varying sectoral shocks affecting sales, so the estimated coefficients on labor productivity became statistically insignificant.

Table 22. Regressions for (log of) labor productivity

Dependent variable: log (laborprod)	Whole sample		Innexe>0		PSM		Psupport>0	
Participation Dummy (Dit)	0.45	0.03	0.45	0.03	0.48	-0.15	0.45	0.003
z	8.59***	0.57	8.58***	0.63	7.81***	-0.22	8.57***	0.06
Dummy sector-year	NO	YES	NO	YES	NO	YES	NO	YES
R2								
within	0.0026	0.1475	0.0044	0.1789	0.0030	0.1495	0.0456	0.2322
between	0.0035	0.0686	0.0063	0.0208	0.0025	0.0001	0.0006	0.0334
Overall	0.0029	0.0552	0.0032	0.0399	0.0018	0.0195	0.0054	0.0902
Number of obs	73,192	73,192	41,514	41,514	25,931	25,931	3,354	3,354
number of groups	12,322	12,322	5,829	5,829	3,194	3,194	419	419
Obs per group								
Min	2	2	2	2	2	2	2	2
Max	10	10	10	10	10	10	10	10
Average	5,9	5,9	7,1	7,1	8,1	8,1	8	8

Source: Own calculations and DANE.

Of course, this finding does not necessarily imply that, in practice, innovation activities induced by public policies, which affect TFP at the firm level, do not impact labor productivity. These results tell us that we cannot capture these effects using the data and econometric methods we have at hand. This estimate may be particularly affected by the problem of the lack of enough observations to render our estimated coefficients statistically significant. We already showed that the estimated impact of program participation on total sales, even when controlling for sector-year interaction terms, was somewhat larger than that for employment (23% against 20%), so this simple calculation would suggest a positive impact on the ratio between these two variables. Still, the problem may be that we don't have enough statistical power to identify it.

Another problem may be measurement error in both variables, which could be potentiated when we combine them in another indicator. In this regard, as we indicated before, to maintain confidentiality, the data available through the DANE web page have imposed certain upper thresholds for total employees and sales applied for some very large firms within specific sectors. This could bias downward the variation for the labor productivity variable, affecting the results of the estimations<sup>19</sup>.

On the other hand, it is natural that the link connecting innovation efforts to firm-level performance variables could be weakened at some point. It is not automatically that innovation efforts generate new or improved products or production processes, and this, in turn, positively affects firms' performance indicators like productivity. Most importantly, we should remember that we are comparing these effects between firms that have received public funds and those that do not participate in these programs. We have already shown that the coverage of these public initiatives is low among formal service sector firms in Colombia. Thus, so far, the above results suggest that the effects of public programs on innovation and productivity are somewhat debilitated at the end of the innovation effort- firm performance's chain. Maybe positive and statistically significant results are obtained for a larger period than the one we are considering for the average firm in the sample (depending on the dataset, this period varies between 6 to 8 years in the case of the productivity regressions). If confirmed, these results call for evaluating existing public support programs focusing on the impact on innovation outcomes. There may be other necessary complementary measures (training,

<sup>19</sup> We plan to solve this potential problem by running the regressions with the uncensored data.

mentoring, etc.) to ensure that public resources to finance expenditure in innovation foster firms' performance, especially productivity.

## **7. Concluding remarks**

Latin America's growth performance in the last 60 years has been disappointing compared to other developing economies in Asia and Eastern Europe. A key factor explaining this fact is sluggish productivity growth. One channel that has been emphasized is the problem of misallocation. The difference in productivity across countries is partly associated with substantial productivity disparities across firms, even within narrowly defined sectors. A second channel is related to low productivity growth at the firm level. This is partly related to low innovation efforts that keep firms relatively small compared to developed economies.

This paper has looked at these two channels of productivity growth in Colombia's service sector. Using the innovation survey applied to this sector, we have identified that improving the allocation of resources among firms within service activities could imply gains in overall sector productivity. We found these potential gains even when the survey covers only formal firms of relatively large size (more than 40 employees). So, gains could be even more significant if we included the informal firms participating significantly in various service activities. The extension of the analysis incorporating a large sample of firms, including those that are informal, could be the subject of future research. Also, even considering the actual set of formal firms, an analysis within subsectors to identify the regulations/taxes or market failures behind the misallocation of resources could be very relevant from a public policy perspective.

The innovation channel, meaning the increase in productivity at the firm level, is also relevant in explaining the overall lag in productivity affecting the service sector in Colombia. We showed that even when considering a sample of service firms that are formal and relatively large, only a small proportion (around 25%) allocated funds to innovation activities. Public programs to support these activities have had a positive impact on firms' innovation expenditures and labor allocation to R+D+i, but the overall effect on the service sector would be minor given the small number of firms that participate in these public initiatives (close to 1%).

For participating firms, public support has positively affected the number of innovations they report (new products, services, or improved production processes). This, in turn, has been associated with an expansion of sales and employment. Still, labor productivity does not show a statistically significant increase in participating firms after controlling for temporal shocks at the sector level. This may be, in part, due to data restrictions associated with the number of observations or the fact that there is a certain level of censoring in some key variables for very large firms belonging to specific sectors. Alternatively, innovation outcomes are not as strong as in affecting labor productivity (though they expand sales and employment) significantly compared to firms that have not received public support. If confirmed, these results suggest a future research agenda where a more detailed analysis of existing public support programs is launched on the design and implementation of these initiatives, looking at the potential causes of its weakened results on innovation outcomes. As indicated above, there may be other necessary complementary measures (training, mentoring, etc.) to ensure that public resources to finance expenditure in innovation foster firms' performance, especially productivity.

## References

- Aboal, D., & Garda, P. (2016). Does public financial support stimulate innovation and productivity? An impact evaluation. *CEPAL Review*, 2015(115), 41–61. <https://doi.org/10.18356/3fdbef72-en>
- Aboal, D., Rojas, G., Servin, B., & Queraltó, P. (2019). How Effective are Innovation Support Programs to Stimulate Innovation? Evidence from Paraguay. *Dotec-Colombia.org*. <http://www.dotec-colombia.org/index.php/series/518>.
- Albis, N., Garcia, J., Sánchez, E., Bayona-Rodríguez, H (2017). Evaluación de Políticas de apoyo a la innovación sobre la productividad de la empresas manufactureras y de servicios en Colombia. CAF - Documento de trabajo N° 2017/29.
- Álvarez, F., Eslava, M., Toledo, M., Alves, G., Daude, C., & Allub, L., and Sanguinetti, P (2018). Institutions for productivity: towards a better business environment. *Economic and Development Report (RED)*. Caracas: CAF.
- Arbeláez, M. and Parra-Torrado, M. (2011): Innovation, Research and Development Investment and Productivity in Colombian Firms. *IDB Working Paper Series*, No. IDB-WP-251
- Banerjee, A. and E. Duflo (2005). Growth theory through the lens of development economics. *Handbook of Economic Growth*, Volume 1A .
- Bartelsman, E., Haltiwanger, J. y Scarpetta, S. (2013). Cross-country differences in productivity: The role of allocation and selection. *American Economic Review*, 103(1), 305-334.
- Becker, B. (2015). Public R&D Policies and Private R&D Investment: A Survey of the Empirical Evidence. Mimeo. Economics and Strategy Group, Aston Business School, Aston University
- Benavente, J., Crespi, G., and Maffioli, A. (2007): Public Support to Firm-Level Innovation: An Evaluation of the FONTEC Program. OVE Working Paper No. 05/07. Office of Evaluation and Oversight, Inter-American Development Bank, Washington, DC.
- Belmar, J. M. (2014). Innovación y fomento : una evaluación no experimental de programas CORFO. *Uchile.Cl*. <https://doi.org/http://repositorio.uchile.cl/handle/2250/116389>
- Bento, P. and Restuccia, D. (2017). Misallocation, establishment size and productivity. *American Economic Journal: Macroeconomics*.
- Bloom, N., Van Reenen, J. and Williams, H. (2019): A Toolkit of Policies to Promote Innovation. *CEP Discussion Paper No 1634*, July.
- Bronzini, R. and Lachini, E. (2014): Are Incentives for R&D Effective? Evidence from a Regression Discontinuity Approach. *American Economic Journal: Economic Policy* 2014, 6(4): 100–134
- Busso, M., Madrigal, L. y Pages, C. (2013). Productivity and resource misallocation in Latin America. *The B.E. Journal of Macroeconomics*, 13(1), 903-932
- Calderón-Madrid, A. (2011). A Micro-Econometric Analysis of the Impact of Mexico's

R&D Tax Credit Program on Private R&D Expenditure. Paper presented at the international conference, Mind the Gap: From Evidence to Policy Impact, June 15–17, Cuernavaca, Mexico.

Cattaneo, M. D., Idrobo, N., & Titiunik, R. (2019a). A Practical Introduction to Regression Discontinuity Designs. <https://doi.org/10.1017/9781108684606>

Cattaneo, M. D., Titiunik, R., & Vazquez-Bare, G. (2019b). Power calculations for regression-discontinuity designs. *The Stata Journal: Promoting Communications on Statistics and Stata*, 19(1), 210–245. <https://doi.org/10.1177/1536867x19830919>

CEPAL (2022). *Innovación para el desarrollo: La clave para una recuperación transformadora en América Latina y el Caribe*. Santiago de Chile.

Chavez, E. (2020): The Effects of Public R&D Subsidies on Private R&D Activities in Mexico. <https://shs.hal.science/halshs-02355106v3>

Chudnovsky, D., López, A., Rossi, M. and Ubfal, D. (2006): Evaluating a Program of Public Funding of Private Innovation Activities: An Econometric Study of FONTAR in Argentina. OVE Working Paper No. 16/06. Office of Evaluation and Oversight, Inter-American Development Bank, Washington, DC.

Crespi, G., Solís, G. and Tacsir, E. (2011): Evaluación del impacto de corto plazo de SENACYT en la innovación de las empresas panameñas. IDB Technical Note No. 263. Inter-American Development Bank, Washington, DC.

Crespi, G., Figal-Garone, L., Maffioli, A. and Melendez, M. (2015): Long-Term Productivity Effects of Public Support to Innovation in Colombia. *Emerging Markets Finance & Trade*, 51:48–64.

Crespi, G., Fernandez-Arias, E. and Stein, E. (2016): Rethinking Productive Development. IADB, (DIA). Washington, DC.

Crespi, G., Figal-Garone, L., Maffioli, A., and Stein, E. (2020): Public support to R&D, productivity, and spillover effects: Firm-level evidence from Chile. *World Development*.

Cusolito, A. and Maloney, W. (2018). *Productivity Revisited. Shifting Paradigms in Analysis and Policy*. International Bank for Reconstruction and Development. The World Bank, 1818 H Street NW, Washington, DC.

David, P., Hall, B., and Toole, A. (2000): Is Public R&D a Complement or Substitute for Private R&D? A Review of the Econometric Evidence. *Research Policy*.

Fernández-Sastre, J., & Martín-Mayoral, F. (2015). The effects of developing-countries' innovation support programs: evidence from Ecuador. *Innovation*, 17(4), 466–484. <https://doi.org/10.1080/14479338.2016.1157447>

Hsieh, C. y Klenow, P. (2009). Misallocation and Manufacturing TFP in China and India. *Quarterly Journal of Economics*, 124(4), 1403-1448.

Hsieh, C. y Klenow, P. (2014). The life cycle of plants in India and Mexico. NBER Working Papers No 1833.

- Jones (2015). The Facts of Economic Growth. NBER Working Paper 21142.
- Kannebley, Jr., S., and Porto, G. (2012): Incentivos fiscais à pesquisa, desenvolvimento e inovação no Brasil: uma avaliação das políticas recentes. IDB Discussion Paper No. 236. Inter-American Development Bank, Washington, DC.
- Lemaire, M., Doebbel, H., Pertuzé, J., & Trinidad Alvarez, M. (2017). Resultados de voucher de innovación y contratos tecnológicos. In Laboratorio de Estudios de I+D+i Empresarial. Centro de Innovación UC: Anacleto Angelini.
- López, A., Reynoso, A. and Rossi, M. (2010): Impact Evaluation of a Program of Public Funding of Private Innovation Activities: An Econometric Study of FONTAR in Argentina. OVE Working Paper No. 03/10. Office of Evaluation and Oversight, Inter-American Development Bank, Washington, DC.
- Navarro, J.C. y Olivari, J. Ed. (2016). La Política de Innovación en América Latina y El Caribe. Nuevos Caminos. BID. Washington, DC.
- OECD (2018). The Oslo Manual. Paris.
- OECD (2020): Innovation for Development Impact Lessons from the OECD Development Assistance Committee. OECD Publishing, Paris.
- OIT-UNESCO. EL ESTADO DE LA CIENCIA (2023). Principales Indicadores de Ciencia y Tecnología Iberoamericanos / Interamericanos.
- Olley, S. and Pakes, A. (1996) The Dynamics of Productivity in the Telecommunications Equipment Industry. *Econometrica*, 64, 1263-1298.
- Petrin, T. (2018). A literature review on the impact and effectiveness of government support for R&D and innovation. [http://www.isigrowth.eu/wp-content/uploads/2018/02/working\\_paper\\_2018\\_05.pdf](http://www.isigrowth.eu/wp-content/uploads/2018/02/working_paper_2018_05.pdf)
- Pless, J. (2022): Are “Complementary Policies” Substitutes? Evidence from R&D Subsidies in the UK. Mimeo. MIT Sloan School of Management.
- Restuccia, D. y Rogerson, R. (2008). Policy distortions and aggregate productivity with heterogeneous establishments. *Review of Economic Dynamics*, 11(4), 707-720.
- Rocha, F. (2015): Does governmental support for innovation have a positive effect on R&D investments? Evidence from Brazil. *Rev. Bras. Inov., Campinas (SP)*, 14, n. esp., p. 37-60
- Sanguinetti, P. (2005): Innovation and R&D Expenditures in Argentina: Evidence from a firm-level survey. Mimeo. Department of Economics, Universidad Torcuato Di Tella.
- Triplett and Boswort, B. (2004): Productivity in the U.S. Services Sector: New Sources of Economic Growth. Brookings Institution Press
- VERDE. (2016). EVALUACIÓN DEL PROGRAMA START-UP CHILE DE CORFO INFORME FINAL. <http://ctie.economia.cl/wp-content/uploads/2017/07/Evaluaci%C3%B3n-Start-Up-Chile-2016.pdf>

What Works Centre for Local Economic Growth. (2015). Innovation: grants, loans and subsidies. <https://whatworksgrowth.org/policy-reviews/innovation/>

Zuñiga et al. (2014): Assessing the Effect of Public Subsidies on Firm R&D Investment: A Survey. Working Paper 12-15. Departamento de Economía. Economic Series Universidad Carlos III de Madrid.

## Appendix A.

### Main features of the EAS survey

This survey is performed annually by the DANE (official statistics department of the Colombian government) to examine firms' activity in the services sector. It included enterprises that were formally established. The sample comes from a list of firms with commercial registries identified in the Economic Census 1990. This list has been updated with information from various public and private entities (i.e., Societal Registry, Enterprise Associations, etc.). Firms are grouped according to the la CIIU Rev. 4, 2-digit classification covering 20 subsectors. Table A.1 details the subsector classification, the number of firms by each activity (for the 2019 edition), and the inclusion criteria set by DANE, which vary by subsector. A key feature from the description of the various subsectors is that the wholesale and retailing activities are not included within the sample (see Table A.2).

Table A.1. EAS 2019: Sectors, number of firms, and criteria for inclusion in the survey

Sector description	# of firms	Criteria for inclusion (pesos of 2018; rate of exchange 1USD= 2900 pesos)
Storage and transport activities	460	More than 40 employees or income higher than 3000 million
Mail and Courier Services	80	More than 40 employees or income higher than 3000 million
Accommodation	730	More than 40 employees or income higher than 3000 million
Food and Beverage Activities	483	More than 75 employees or income higher than 3000 million
Editing	115	More than 10 employees or income higher than 1500 million
Film, video and television program production	42	More than 40 employees or income higher than 2000 million
Programming, broadcasting and /or news agency activities	44	More than 40 employees or income higher than 3000 million
Telecommunications	199	More than 40 employees or income higher than 3000 million
Computer system development, computer consulting and related activities	412	More than 75 employees or income higher than 3000 million
Real Estate, Rental and Leasing	342	More than 75 employees or income higher than 3000 million
Professional, scientist and technical activities	675	More than 75 employees or income higher than 3000 million
Advertising	167	More than 75 employees or income higher than 3000 million
Travel agencies, Tour Operators, Booking and Related Services	84	More than 75 employees or income higher than 3000 million
Employment Activities, Security and Private Investigations, Building Services	880	More than 75 employees or income higher than 3000 million
Administrative and office support activities and other business support, except call centers	140	More than 75 employees or income higher than 3000 million
Call Center	78	More than 75 employees or income higher than 3000 million
Private Higher Education	161	More than 20 employees or income higher than 1000 million
Human health activities	1.139	More than 40 employees or income higher than 3000 million
Gambling, sports, recreation and leisure activities	177	More than 40 employees or income higher than 2000 million
Other service activities	196	More than 40 employees or income higher than 2000 million
Total	6.604	

Source: DANE

The activities with the most enterprises in the sample are Human Health Services (1139) and Employment Activities, Security, and Private Investigation (880). On the other hand, Call Centers (78) and Film, Video, and Television (42) have the smallest number of firms participating in the survey. The criteria DANE sets for firms to be included in the study are set in terms of minimum employment or income level. Regarding employment, most subsectors have a minimum requirement of 40 or 75 employees. Only Editing has a relatively lower limit of 10 or more workers. When the minimum income threshold applies, we see that at the 2018 prices and exchange rates, this implies annual sales equal to or above 1 million dollars in most

activities. These data features entail that many formal small firms (between 10 and 40 workers or with income less than 1 million) will not be included in the sample.

Another interesting characteristic of the survey is that since 2019, there has also been a regional disaggregation in 26 departments. The data is available since 2006 and up to 2021. The most relevant variables covered by the survey are total sales and production, value-added, intermediate consumption, employment, and wages, among other indicators. For 2019, the number of firms incorporated in the sample was 6605. A key characteristic is that firms are followed through time, so the data has a panel format.

Table A. 2 presents a more detailed description of the activities included in each subsector. We can corroborate that wholesale and retailing services are not included in the survey. As indicated in the main text, this has been the main reason for working with the EDITS survey.

Table A.2

Table A.1. Detailed description of subsectors	
Administrative and office support activities and other business support activities.	It includes administrative and office support activities; call center activities; Business support services activities n.e.c., such as collections and credit rating bureaus, packaging activities. (Division 82, ISIC, Revision 4 a.c.) For EAS 2019, information from class 8220 Call Center activities is disaggregated from N4 in subsection N5.
Editing activities	It includes book publishing activities, periodicals, and other publishing activities. Not included are the edition of computer programs (Division 58, except Class 5820 ISIC Revision 4 a.c.)
Employment Activities, security and Private Investigation, Building Services	It includes the activities of employment agencies and the provision of human resources, private security activities, and services to buildings, such as facilities support and cleaning. (Divisions 78, 80, 81, except Class 8130 ISIC Revision 4 a.c.)
Television Programming and Broadcasting Activities	It includes programming and transmission activities in the sound broadcasting service; television programming and broadcasting activities and news agency activities. (Division 60 and Class 6391 ISIC Revision 4 a.c.)
Telecommunication	It includes wired, wireless, satellite and other telecommunications activities. (Division 61, ISIC Revision 4 a.c.)
Human Health	It includes the activities of hospitals and inpatient clinics; outpatient medical and dental practice activities; other human health-related care activities, such as diagnostic support, therapeutic support, etc. (Division 86 ISIC Revision 4 a.c.)
Advertising	It includes the provision of a full range of advertising services (through in-house resources or by subcontracting), including consultancy services, creative services, production of advertising material and use of the media. (Class 7310 ISIC Revision 4 a.c.)
Motion picture production	It includes activities of production, post-production, distribution and exhibition of motion pictures, videos and production of television programs, commercials and commercials; (Division 59, except Class 5920 ISIC Revision 4 a.c.)
Mail & Courier Services	It includes activities of national apostals; and messaging activities. (Division 53, ISIC Revision 4 a.c.)
Food & Beverage Service Activities	Includes restaurant, cafeteria, and mobile food service activities; catering activities for events and other food services; Sale of alcoholic beverages for consumption within the establishment. (Division 56, ISIC Revision 4 a.c.)
Real estate and rental activities	It includes real estate activities carried out with owned or leased property and real estate activities carried out in exchange for remuneration or under contract. Rental and leasing of motor vehicles; renting and leasing of personal items and household goods; Rentals and leases of other types of machinery and equipment and tangible goods n.e.c. without operators. (Divisions 68 and 77 ISIC Revision 4 BC)
Professional, scientific and technical activities	It includes legal and accounting activities; business administration activities, management consulting activities; architectural and engineering activities, technical testing and analysis; scientific research and development; market research and conducting surveys; and other professional, scientific and technical activities. (Divisions 69, 70, 71, 72, Class 7320 and Division 74 ISIC Revision 4 BC)
Travel agencies	It includes the activities of travel agencies and tour operators; and other booking services and related activities. (Division 79, ISIC Revision 4 a.c.)
Almacenamiento y actividades complementarias al transporte	It includes storage and warehousing activities; activities of stations, roads and complementary services for land transport, port activities and complementary services for water transport; airport activities, air navigation services and other activities related to air transport; cargo handling, and other activities complementary to transportation. (Division 52, ISIC, Revision 4 a.c.)
Accommodation	Includes accommodation in hotels, apart-hotels, resorts, rural accommodation, other types of visitor accommodation, camping area and recreational park activities, hourly services, and other types of accommodation n.e.c. (Division 55 ISIC Revision 4 a.c.)
Development of computer systems and data processing	It includes computer systems development (planning, analysis, design, programming, testing), computer consulting, and related activities, such as computer facility management and other information technology activities and computer service activities. Data processing, hosting and related activities such as web portals. (Divisions 62 and 63 ISIC Revision 4 BC)
Private Higher Education	It includes the activities of vocational technical education, technological education, education of university institutions or technological schools, education of universities. (Group 854 ISIC Revision 4 a.c.)
Gambling, sports, recreation and leisure activities	It includes gambling and betting activities; Sports, recreation and leisure activities (management of sports facilities, sports clubs, amusement parks and others). (Divisions 92 and 93, ISIC Revision 4 BC)
Other services activities	It includes creative, artistic, and entertainment activities; maintenance and repair of computers and peripheral equipment; and other personal service activities, such as laundries, hairdressers, funeral services and other personal service activities n.e.c. (Division 90, Class 9511 and Group 960 ISIC Rev. 4 a.c.)

Appendix B. Further calculations of the OP productivity decomposition using the panel data sample.

Tabla B1. Covariance across firms within sectors. 2010-2019. In thousand pesos dec 2018.

sector	2010	2011	2012	2013	2014	2015	2016	2017	2018	2019	Average
1	368071.1	357754.3	87744.71	57151.11	100348.9	85499.15	45924.39	191464.7	45987.57	-11459.2	132848.7
2	-4169.28	-3492.82	-194077	-180720	57229.2	63498.34	38546.88	41097.53	-19266.3	11226.47	-19012.7
3	-142132	-130450	-110049	-121054	-287450	-361010	-208819	-202195	-177414	-174540	-191511
4	-4948.75	-24036.1	-27178.8	-41457.8	55884.91	50645.92	150434.1	127484.9	2195.263	2358.304	29138.2
5	40227.53	36357.15	-24370.6	-24446.8	-10997.7	-10664	-11200.9	-9092.95	-1362.02	-3925.78	-1947.61
6	88569.91	76546.6	-16759.7	-21949.1	-107785	-108431	-67308	-54430.3	-128203	-124953	-46470.3
7	22336.88	22158.84	27363.33	19049.61	46142.61	52293.75	119288.6	89003.18	141365.4	69703.59	60870.57
8	-23545.8	-14266	-7875.39	-5641.59	1988.441	10162.18	9745.071	2057.677	8104.514	11739.42	-753.147
9	9492.033	9699.194	-86793.4	-76834.6	11765.94	8779.276	3338.376	4699.342	10956.24	15591.76	-8930.59
10	38261.92	33105.08	-67430.2	-63347.7	-4923.37	-2991.36	30614.4	37336.9	21671.41	26477.3	4877.441
Average	39216.31	36337.6	-41942.6	-45925	-13779.6	-21221.8	11056.42	22742.63	-9596.53	-17778.1	-4089.07

Table B2. Covariance across firms within sectors. 2010-2019. In logs

sector	2010	2011	2012	2013	2014	2015	2016	2017	2018	2019	Average
1	1.693744	1.677212	0.543147	0.439513	0.427474	0.42553	0.610673	0.606874	0.189812	0.131	0.674498
2	0.320408	0.328429	0.144285	0.189545	0.402021	0.477148	0.293285	0.280076	0.072831	0.133	0.264103
3	-0.12237	-0.14963	-0.23677	-0.26967	-0.28066	-0.23674	-0.20596	-0.19088	-0.1694	-0.13163	-0.19937
4	0.486505	0.467582	-0.075	-0.05584	0.073789	0.112982	0.138174	0.11129	0.132843	0.173897	0.156622
5	0.558484	0.497046	0.01898	-0.0252	0.005719	-0.00359	0.024082	0.038151	0.001419	-0.0008	0.111429
6	0.59497	0.556061	0.191949	0.182647	-0.07023	-0.05605	0.017308	0.038255	0.159579	0.15941	0.17739
7	0.43423	0.46374	0.176421	0.144631	0.162455	0.194436	0.292774	0.26753	0.166584	0.122211	0.242501
8	0.30637	0.354077	-0.26324	-0.19395	0.104769	0.246779	0.337459	0.313785	0.204142	0.255584	0.166577
9	0.295908	0.31974	0.136752	0.143773	0.228811	0.197642	0.125848	0.147385	0.271905	0.327316	0.219508
10	0.425047	0.438856	-0.00679	-0.02293	0.039628	0.028865	0.182214	0.19535	0.268893	0.293074	0.184221
Average	0.499329	0.495311	0.062974	0.053251	0.109378	0.1387	0.181585	0.180782	0.129861	0.146306	0.199748

Table B.3. OP decomposition. 2010-2019. In logs

sector	year	CT	ctj mean	pj mean	Ps
1	2010	0.160743	0.499329	10.92286	11.58293
2	2011	0.159064	0.495311	10.85795	11.51233
3	2012	0.065213	0.062974	11.11215	11.24034
4	2013	0.055534	0.053251	11.0721	11.18088
5	2014	0.050899	0.109378	11.6952	11.85548
6	2015	0.022411	0.1387	11.68364	11.84475
7	2016	-0.01216	0.181585	11.7185	11.88793
8	2017	-0.00735	0.180782	11.69808	11.87151
9	2018	-0.03078	0.129861	11.86641	11.96548
10	2019	-0.02688	0.146306	11.87523	11.99465

## Appendix C. Panel dataset information

Table A4. Panel structure of the dataset

id:	10001, 10002, ..., 85113939	n =	12603
year:	2008, 2009, ..., 2019	T =	12
	Delta(year) = 1 year		
	Span(year) = 12 periods		
	(id*year uniquely identifies each observation)		

Distribution of T_i:	min	5%	25%	50%	75%	95%	max
	2	2	4	6	10	12	12

Freq.	Percent	Cum.	Pattern
2198	17.44	17.44	111111111111
1713	13.59	31.03	.....111111
1698	13.47	44.51	.....11
1684	13.36	57.87	...11111111
931	7.39	65.25	.....1111
741	5.88	71.13	..1111111111
635	5.04	76.17	1111.....
447	3.55	79.72	..11.....
381	3.02	82.74	.....1111..
295	2.34	85.08	.....11....
224	1.78	86.86	...111111..
175	1.39	88.25	1111111111..
1481	11.75	100.00	(other patterns)
12603	100.00		XXXXXXXXXXXX

## Appendix D. PSM estimation

### PSM Do File

\*Generate probit dummy

```
sort id year
gen dummy_condition = (fondos_total > 0 & year > 2008)
bysort id: egen dummy_fondos = max(dummy_condition)
drop dummy_condition
```

\*drop firm that received public funds in 2008

```
sort id year
gen fondos_2008 = (fondos_total > 0 & year == 2008)
gen fondos_other_years = (fondos_total > 0 & year != 2008)
bysort id: egen max_fondos_2008 = max(fondos_2008)
bysort id: egen max_fondos_other_years = max(fondos_other_years)
gen only_in_2008 = (max_fondos_2008 == 1 & max_fondos_other_years == 0)
drop if only_in_2008 == 1
drop only_in_2008 fondos_other_years fondos_2008 max_fondos_2008
max_fondos_other_years
```

\*drop firms that are not in 2008

```
sort id year
egen firm_in_2008 = total(year == 2008), by(id)
keep if firm_in_2008 == 1
drop firm_in_2008
```

\*run probit model for year 2008

```
probit dummy_fondos personal_total personal_universitario dummy_d dummy_e dummy_g
dummy_h dummy_i dummy_j dummy_k dummy_m dummy_p dummy_q invertido_total if
year == 2008
```

\*getting the predicted values

```
predict predicted_probit
```

```
bysort id: replace predicted_probit = predicted_probit[_n-1] if year != 2008 & _n > 1
```

\*keeping the firms that are only in the range of the predicted values of dummy =1

```
summarize predicted_probit if dummy_fondos == 1 & year == 2008, detail
```

```
scalar min_pred_probit = r(min)
```

```
scalar max_pred_probit = r(max)
```

```
keep if dummy_fondos == 1 | (dummy_fondos == 0 & predicted_probit >= min_pred_probit &
predicted_probit <= max_pred_probit)
```

Table D1. First Stage PSM regression

<i>First Stage PSM (Probit) Regression. Depend variable = 1 if firm i received funds 2009-2019</i>	
University personnel	0,0003873***
t	3,02
innovation expenditures (in current pesos)	0,00000000898***
t	2,94
<i>Sectors' dummies</i>	
Electricity, gas, steam and air conditioning supply	0,195281
t	1,42
Water supply; sewerage and waste management	0,1698601
t	1,03
Wholesale and retail trade	-0,9115466***
t	-6,9
Transportation and storage	-0,8093282***
t	-5,04
Accommodation and food service activities	-1,047697***
t	-5,42
Information and communication	-0,3716678***
t	-2,85
Financial and insurance activities	-0,6728336
t	-1,36
Professional, scientific and technical activities	2,787943***
t	9,28
Education	0,8469883***
t	5,87
Human health and social work activities	Omitted
Pseudo R2	0.3032
Number of Observations	3592