

**GROWING BY LEARNING:
FIRM-LEVEL EVIDENCE
ON THE SIZE-PRODUCTIVITY NEXUS**

2016

Enrique Moral-Benito

**Documentos de Trabajo
N.º 1613**

BANCO DE ESPAÑA
Eurosistema



**GROWING BY LEARNING: FIRM-LEVEL EVIDENCE
ON THE SIZE-PRODUCTIVITY NEXUS**

GROWING BY LEARNING: FIRM-LEVEL EVIDENCE ON THE SIZE-PRODUCTIVITY NEXUS

Enrique Moral-Benito ^(*)

BANCO DE ESPAÑA

Contact: enrique.moral@gmail.com. I thank Roberto Ramos for insightful suggestions. I also thank seminar participants at the Banco de España. The opinions and analyses are the responsibility of the author and, therefore, do not necessarily coincide with those of the Banco de España or the Eurosystem.

The Working Paper Series seeks to disseminate original research in economics and finance. All papers have been anonymously refereed. By publishing these papers, the Banco de España aims to contribute to economic analysis and, in particular, to knowledge of the Spanish economy and its international environment.

The opinions and analyses in the Working Paper Series are the responsibility of the authors and, therefore, do not necessarily coincide with those of the Banco de España or the Eurosystem.

The Banco de España disseminates its main reports and most of its publications via the Internet at the following website: <http://www.bde.es>.

Reproduction for educational and non-commercial purposes is permitted provided that the source is acknowledged.

© BANCO DE ESPAÑA, Madrid, 2016

ISSN: 1579-8666 (on line)

Abstract

It is a well-known empirical regularity that small firms are less productive than large firms. However, does size cause productivity or vice versa? Using matching methods, I find that productivity shocks are followed by significant increases in size defined by employment. In contrast, size shocks are not followed by productivity gains at the firm level. This finding casts doubt on the conventional wisdom that aggregate productivity in Spain is driven by a firm size distribution biased towards small firms in comparison with other developed countries. According to my findings, low firm-level productivity might play a crucial role in shaping the Spanish firm size distribution.

Keywords: firm-level data, productivity, size distribution.

JEL classification: L11, L25, D24.

Resumen

Las pequeñas empresas son menos productivas que las grandes empresas. Asimismo, las empresas españolas son más pequeñas y menos productivas que sus homólogas europeas. Sin embargo, ¿son menos productivas porque son más pequeñas o son más pequeñas porque son menos productivas? Este trabajo trata de dar respuesta a esta pregunta estimando la dirección de causalidad entre productividad y tamaño empresarial mediante técnicas de *matching* y explotando microdatos de empresas españolas. Los resultados indican que incrementos de productividad son seguidos por aumentos significativos en el tamaño empresarial medido por el número de empleados. Por el contrario, no se observan ganancias de productividad tras episodios de aumentos en el tamaño empresarial. Este hallazgo cuestiona el consenso establecido de que la baja productividad agregada en España se debe a un exceso de empresas pequeñas en comparación con otros países europeos. De acuerdo con mi interpretación de la evidencia empírica presentada en este trabajo, la baja productividad de las empresas españolas, tanto pequeñas como grandes, podría ser la principal causante del excesivo número de empresas pequeñas y, por tanto, de la baja productividad agregada.

Palabras clave: productividad empresarial, tamaño empresarial.

Códigos JEL: L11, L25, D24.

1 Introduction

In Spain and other Southern European countries there is a large number of small firms in comparison to other developed economies. On the other hand, it is well-known that large firms are more productive than small ones (see Section 3). As a result, it is typically argued that a firm size distribution biased towards small firms is at the root of the low productivity growth in the South of Europe. Understanding the obstacles to firm growth in those countries is thus at the center of the policy debate (IMF (2015)). For instance, Guner, Ventura, and Xu (2008) show that size-dependent policies, by preventing most productive firms to grow, can potentially explain a high fraction of the aggregate TFP differences across countries. However, Gourio and Roys (2014), Garicano, LeLarge, and Van-Reenen (2015) and García-Santana and Pijoan-Mas (2014) have found modest effects when evaluating real size-dependent policies in France and India.

This paper argues that low firm-level productivity may also generate a firm size distribution biased towards small firms. Canonical models of firm growth characterize industries as groups of heterogeneous-productivity firms under the assumption that they learn about their productivity as they operate in the market (see for instance Jovanovic (1982); Melitz (2003)). Firms maximize profits choosing the level of employment (and capital) given their productivity shocks, which are assumed to be exogenous. Low productivity firms are less likely to survive and thrive than their more efficient counterparts. This type of models rationalize the positive association between size and productivity by implicitly imposing that causality goes from productivity to employment. Under this hypothesis, the presence of a high number of small firms should be interpreted as the consequence rather than the cause of their low productivity.

On the empirical front, little is known about the direction of causality in the size-productivity nexus at the firm level. While there is a wide literature on the determinants of firm growth in terms of employment (e.g. Henrekson and Johansson (2010); Lopez-Garcia and Puente (2012)), studies analyzing the relationship between productivity and employment growth are scarce and provide contradicting results. Daunfeldt, Elert, and Johansson (2010) find that neither high-growth employment is associated to high-growth productivity nor the reverse. They conclude that a trade-off exists between employment and productivity. On the other hand, Du, Gong, and Temouri (2013) find that firms which exhibit higher productivity growth are more likely to become high-growth firms in terms of turnover and that high-growth firms are more likely to experience higher productivity growth in the future.

This paper provides evidence in favor of the hypothesis that firm productivity shocks generate firm growth but not the other way around. In particular, I consider two different treatment variables. First, I consider high-growth in terms of (employment) size as the treatment status and productivity

growth as the outcome variable. Second, I consider high-growth productivity as the treatment and employment growth as the outcome. In the baseline specification, high-growth episodes are defined as productivity or employment growth above 10% in a given year. Needless to say, both treatments are far from an experimental setting, which threatens the causal interpretation of the estimates. For instance, one can interpret high employment growth as a proxy for demand shocks, but a pure demand shock may also generate a large increase in productivity that will be ambiguously labeled as a TFP shock in this paper. Unfortunately, the lack of exogenous TFP shocks observable to the econometrician precludes the consideration of treatment assignments independent of firm observable and unobservable characteristics.

I attempt to control for reverse causality and self-selection by using matching techniques that are able to account for observable characteristics but ignore selection on unobservables (e.g. I cannot rule out the existence of demand shocks simultaneously affecting employment and TFP growth). Having these concerns in mind, I estimate the effects up to five years after each treatment. The estimated effects indicate that high-productivity growth is followed by statistically significant increases in size/employment growth; in contrast, employment growth is not followed by subsequent gains in productivity. These opposite patterns are reassuring. Selection on unobservables and reverse causality would imply, among other things, the presence of a third omitted factor simultaneously causing subsequent increases in size and TFP, e.g. a demand shock. Under this demand shock scenario, the positive association between size and TFP growth would show up regardless of which increase (either TFP or size) takes place in the first place. Therefore, I argue that my estimates provide suggestive evidence in favor of the hypothesis that TFP growth causes employment growth but not the other way around.

I label this pattern as growing by learning because those firms learning about their higher efficiency levels are those that happen to grow more; under this hypothesis, the process of learning and innovation can be interpreted as the mechanism that makes firms thrive as it is usually assumed in models of firm dynamics. According to the growing by learning hypothesis, Spanish firms are smaller because their productivity is lower. This finding raises an obvious question: Why do Spanish firms have worse abilities to convert inputs into output? Based on the survey by Syverson (2011), I highlight the potential role of managerial talent, quality of inputs, and R&D activities. Available cross-country evidence reveals that Spanish firms perform substantially worse than their EU4 counterparts in all the three factors, which may explain, at least partially, their dismal performance in terms of productivity.

The rest of the article is organized as follows. Section 2 describes the Spanish firm-level data. Section 3 documents the existence of a positive association between size and productivity at the firm level together with the fact that Spanish firms are less productive than their European counterparts.

Then, Section 4 presents the econometric methodology in detail, the main results regarding the growing by learning hypothesis as well as some robustness exercises. In Section 5, I discuss the implications of the growing by learning estimates for Spanish firms as well as the potential sources behind their low productivity in comparison to their EU4 counterparts. Some concluding remarks are provided in Section 6.

2 Data

I use a firm-level dataset which contains information of a representative sample of Spanish non-financial companies from 2000 to 2007. The sample contains an average number of 497,782 firms per year. This database is named Central Balance Sheet Data—or *Central de Balances Integrada (CBI)* in Spanish—and is provided by the Banco de España. This dataset covers not only manufacturing but also services, construction and trade sectors.

The database is comprised of two complementary datasets. The first one—*Central de Balances Anual (CBA)*—is based on a standardized voluntary survey handled to companies at the time of requesting compulsory accounting information. Each year, around 9,000 companies fill this survey. The information gathered is very detailed, but the sample size is low and big firms are over-represented. The second dataset—*Registros Mercantiles (CBB)*—contains the balance sheets of a much larger number of companies. It originates from the firms' legal obligation to deposit their balance sheets on the Mercantile Registry. Therefore, coverage is much wider.

The Bank of Spain Central Balance Sheet Office is in charge of collecting and cleaning these datasets. All of the variables contained in the latter database are also included in the former. For each firm, I observe its revenue, total wage bill, employment, book value of the capital stock (both physical and intangible), expenses in intermediate goods, and sector of activity at the 4-digit level (according to the NACE rev. 2 classification). Since most of the variables are recorded in nominal terms, I employ sector-specific deflators for capital and value added to compute real values with 2000 as the base year.¹ Using the information above, I also compute firm-specific measures of total factor productivity (see Appendix A for details). Finally, for some estimations I also consider the exporter status of each firm taken from the micro data used by the Banco de España to construct the official Spanish Balance of Payments. This dataset, which accounts for 97 per cent of aggregate Spanish exports in 2007, contains information on firms making transactions with foreign agents if they are worth more than €12,000 and are performed through a bank (€50,000 from 2008 on). Ramos and Moral-Benito (2013) provide more information on this dataset.

¹I take the capital deflators from Mas, Pérez, and Uriel (2013) and the value added deflator from Spanish National Accounts. Both sets of deflators are constructed at the 2-digit NACE classification.

Despite firms have the legal obligation to submit their statements, some observations are missing from my data because firms deposit their balance sheets late or on paper form, in which case they may not have been digitized. Panel A of Table 1 illustrates the size distribution of firms in my raw sample for the year 2001. The table also compares this distribution with that obtained from the Central Business Register available from the National Statistics Institute, which contains employment information for the universe of Spanish firms. There are two important aspects to highlight. First, the coverage of my raw sample is remarkably large in terms of both the number of firms (56% of the operating firms in Spain) and the level of employment (54% of total employment). Second, my sample provides an excellent representation of the firm size distribution in Spain. In particular, small firms (less than 10 employees) account for 83.90% of the total number of firms and 20.47% of the employment in my sample versus 83.07% and 20.23% in the population. At the other extreme, large firms (more than 200 employees) represent less than 0.5% of the total number of firms both in my sample and in the population, while they account for 33.47% of the employment in my sample and 32.13% in the population.

Table 1: Size distribution of firms in my sample and in the census.

Employees	Central Balance Sheet Dataset				Central Business Register			
	Firms		Labor		Firms		Labor	
	Total (#)	Share (%)	Total (#)	Share (%)	Total (#)	Share (%)	Total (#)	Share (%)
PANEL A: Raw Sample								
0-9	406,924	83.90	941,897	20.47	715,795	83.07	1,718,600	20.23
10-19	41,664	8.59	583,312	12.68	77,372	8.98	1,050,038	12.36
20-49	27,125	5.59	828,714	18.01	46,683	5.42	1,400,422	16.49
50-199	8,064	1.66	707,535	15.38	17,781	2.06	1,596,481	18.79
+200	1,245	0.26	1,540,260	33.47	4,082	0.47	2,728,958	32.13
All	485,022	100.00	4,601,718	100.00	861,713	100.00	8,494,499	100.00
PANEL B: Final Sample								
1-9	249,770	76.34	907,098	20.00	531,399	78.46	1,718,600	20.23
10-19	41,272	12.62	577,844	12.74	77,372	11.42	1,050,038	12.36
20-49	26,919	8.23	822,699	18.14	46,683	6.89	1,400,422	16.49
50-199	7,984	2.44	700,565	15.44	17,781	2.63	1,596,481	18.79
+200	1,219	0.37	1,528,178	33.69	4,082	0.60	2,728,958	32.13
All	327,164	100.00	4,536,384	100.00	677,317	100.00	8,494,499	100.00

Notes. Figures refer to the year 2001. Self-employed persons are not included.

From this original sample I drop observations with missing or non-positive values for the number of employees, value added, or capital stock. I also eliminate observations at the top and bottom 1% of these variables. I also exclude firms with 0 employees because these firms represent mostly firms with no production, being created merely for tax purposes.² I am left with around 350,000 firms

²Notice that these firms are not self-employed people, which are not covered in our dataset.

per year distributed across 518 4-digit industries. In Panel B of Table 1 I compare this sample to the population of firms in the Central Business Register, from which I have also deleted firms with 0 employees. The screening strategy has minor effects in the distributions of firms and employment, and the representativeness of my final sample remains noticeably good.³

3 The Size-Productivity Nexus

Small firms are less productive than larger ones. Table 2 confirms this fact for my sample of Spanish firms. Relative to the firms with 1-9 employees, firms with more than 250 employees are roughly 1.83 log points more productive (i.e. 208 per cent more productive) according to the figures in Column (1). This figure implies a TFP ratio of $e^{1.83} = 6.23$, which indicates that the average firm with more than 250 employees makes around six times as much output with the same measured inputs as the average firm with 1-9 employees. Also, firms with between 50-249 employees are roughly 105 per cent more productive than firms with between 1 and 9 employees, implying a TFP ratio of 2.5. Finally, firms with 20-49 and 10-19 employees are 67 and 41 per cent more productive than firms with 1-9 employees, which implies TFP ratios of 1.80 and 1.43, respectively. These differences cannot be explained by differences in the size distribution of firms across industries or years since they remain very similar when I regress log TFP on a set of size dummies accounting for 2-digit industry and year dummies in Column (2). Finally, Columns (3)-(6) report the same figures for each main sector of activity confirming the patterns discussed above.

Table 2: Average productivity by size class.

Size class	Overall (1)	Overall (2)	Manufacturing (3)	Construction (4)	Trade (5)	Services (6)
1 – 9	0.88	1.10	1.03	0.55	1.02	0.81
10 – 19	1.24	1.46	1.38	0.86	1.36	1.29
20 – 49	1.48	1.68	1.61	1.09	1.55	1.58
50 – 249	1.81	2.01	2.09	1.37	1.86	1.71
+250	2.71	2.84	3.02	2.51	2.68	2.34
Dummies	NO	YES	NO	NO	NO	NO

Notes. Figures refer to average log total factor productivity (TFP) measured at the firm level as described in Appendix A. Dummies refer to the inclusion of 2-digit industry and year dummies in a regression of log TFP on a set of size dummies. The sample covers the period 2000-2007.

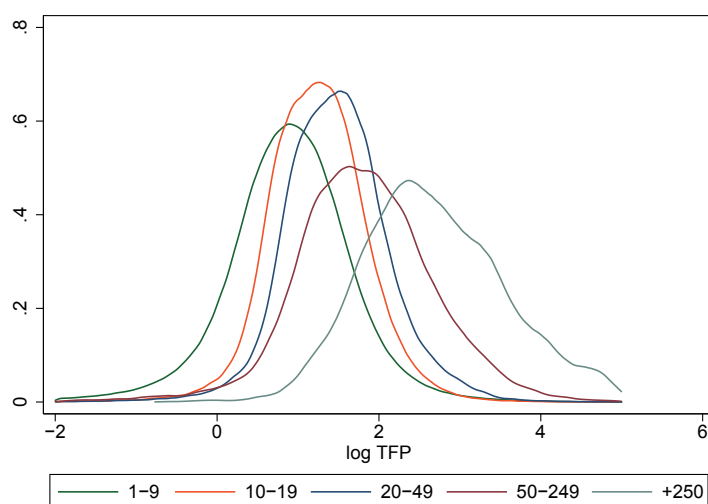
In addition to the descriptive statistics reported in Table 2, the size-productivity nexus is also illustrated by the correlation between log employees and log TFP, which is 0.33 ($t - stat = 12.59$). This correlation remains positive, large, and statistically significant when I account for a set of

³This is true for all years between 2000 and 2007 as illustrated in Figure B.1 in the Appendix.

industry and year dummies: 0.32 with $t - stat = 23.34$. I next compute the quintiles of TFP and size by industry and year confirming that firms in higher quintiles of TFP tend to be in higher quintiles of size. For instance, around 40% of the firms in the first quintile of TFP belong to the first size quintile while only 5% belong to the fifth size quintile. These figures imply that among the 20% most productive firms, 40% are also among the largest 20% while only 5% are among the smallest 20%.

Finally, Figure 1 illustrates that the whole distribution of productivity is shifted to the right for large firms, which implies that average figures reported above are not driven by a handful of outliers. Moreover, dispersion seems to be larger in the group of firms with more than 50 employees. This pattern is also documented by Farinas and Huergo (2015) for a sample of OECD countries using the OECD compendium of productivity indicators (see Figure 5.1 in Farinas and Huergo (2015)).

Figure 1: Distribution of productivity by size class in Spain



Notes. This figure plots the distribution of log TFP of Spanish firms over the period 2000-2007 by size class. The five size categories reported are: (1) from 1 to 9 employees; (2) from 10 to 19 employees; (3) from 20 to 49 employees; (4) from 50 to 249 employees; (5) more than 250 employees. The distribution is estimated by means of an epanechnikov kernel function using the optimal bandwidth that minimizes the mean integrated squared error.

3.1 Spanish Firms in the EU Context

In terms of cross-country data, the Spanish firm size distribution is biased towards small firms in comparison to other developed countries. For instance, firms with less than 9 employees account for 41% of total employment in Spain while this figure is 20% in Germany and 32% in France.⁴ On the

⁴See OECD Entrepreneurship at a Glance (2014) available at <http://goo.gl/501vw3>.

other hand, Spanish firms are also less productive than their European counterparts on average. For instance, average labour productivity of Spanish firms is €39,140 while this figure is €51,610 for Germany and €53,820 for France.⁵ The combination of these two results provides further evidence in favour of the size-productivity nexus based on cross-country variation of micro-level data.

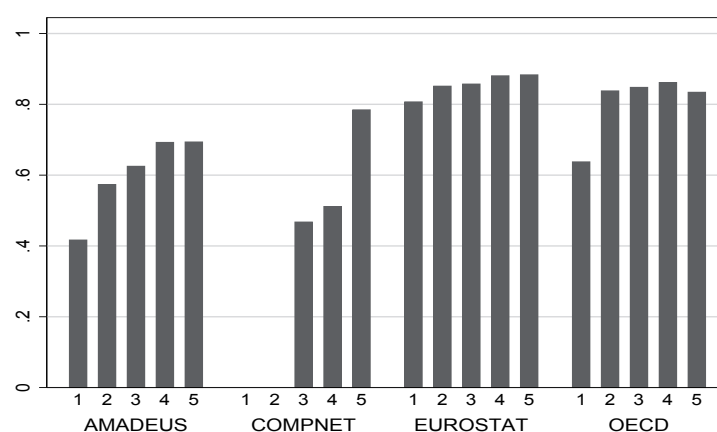
In light of the positive association between size and productivity, it is typically argued that the large number of small firms in comparison with other developed economies is at the root of the low productivity growth in Spain (e.g. IMF (2015)). However, the evidence in Figure 2 casts doubt on this conclusion since Spanish firms are less productive than their European counterparts for all size classes. This Figure shows four different measures of productivity by firm size in Spain relative to the other EU4 countries (Germany, France, and Italy).⁶ First, I compute firm-specific TFP for the four countries as described in Appendix A but using the Amadeus database. In particular, I follow the approach in Gal (2013) to ensure comparability of such productivity measures across countries. Second, I use the same TFP measure taken from the ECB's CompNet database that includes a usable sample of firms with more than 20 employees (see Lopez-Garcia, di Mauro, and Others (2015)). Third, I consider labour productivity, measured as value added per employee, from EUROSTAT's Structural Business Statistics (SBS). Fourth, I use labour productivity from an alternative source, the Entrepreneurship at a Glance (2014) report by the OECD. In all the four cases, Figure 2 reports the ratio of average productivity in Spain to average productivity in the remaining EU4 countries. Moreover, I consider five size categories: (1) from 1 to 9 employees; (2) from 10 to 19 employees; (3) from 20 to 49 employees; (4) from 50 to 249 employees; (5) more than 250 employees. All the ratios refer to the total economy including manufacturing, trade, construction and services.

One lesson emerges from Figure 2. Spanish firms are less productive than their European counterparts at all size categories and for all the four measures of productivity (i.e. the ratio is lower than 1 in all cases). For instance, labour productivity of Spanish firms with 1-9 employees (group 1) is 19% lower than the corresponding EU4 average according to EUROSTAT and 36% lower according to the OECD. In the case of large firms with more than 250 employees (group 5), Spanish firms' labour productivity is 12% and 17% lower according to EUROSTAT and the OECD. Turning to total factor productivity, the productivity of Spanish firms in group 3 (20-49 employees) is 38% and 53% lower than their EU4 counterparts according to AMADEUS and COMPNET, respectively. Large firms in group 5 (more than 250 employees) are 31% and 22% less productive in Spain. Finally, it is worth mentioning that these figures remain qualitatively unchanged when looking at sector-specific ratios (see Figure B.2). To sum up, the productivity of Spanish firms is lower for all size categories.

⁵See EUROSTAT Structural Business Statistics (SBS) available at <http://goo.gl/FDFj9G>.

⁶The lack of firm-level data precludes the consideration of a larger set of countries for the four measures analyzed here.

Figure 2: Spain-to-EU4 productivity ratio by firm size



Notes. Each bar plots the ratio of average productivity in Spain to average productivity in the remaining EU4 countries (Germany, France, and Italy). The five size categories reported are: (1) from 1 to 9 employees; (2) from 10 to 19 employees; (3) from 20 to 49 employees; (4) from 50 to 249 employees; (5) more than 250 employees. AMADEUS and COMPNET refer to total factor productivity ratios, while EUROSTAT and OECD refer to labour productivity ratios (see the main text for more details). AMADEUS and COMPNET figures refer to the period 2004-2012. EUROSTAT covers the period 2002-2013, and OECD refers to the year 2011. All the figures refer to the total economy including manufacturing, trade, construction and services.

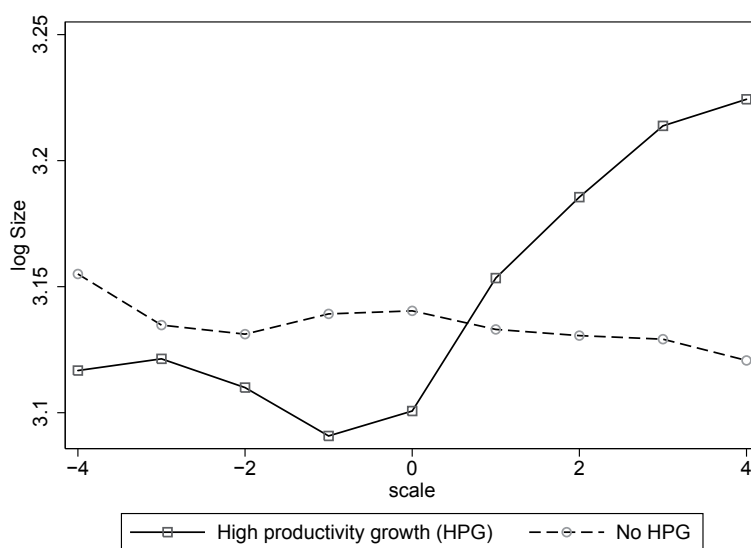
However, the productivity gap is smaller for large firms.

4 Growing by learning vs. Learning by growing

In order to graphically inspect whether firms grow once they enjoy a positive productivity shock, Figure 3 plots on the horizontal axis a time scale which is zero for the period where firms are hit by the productivity shock. These firms are labeled as High Productivity Growth (HPG). To be more concrete, HPG firms are those with TFP growth above 10% in 2003, which roughly corresponds to the 85th percentile in this year. The remaining firms are labeled as “No HPG” in the same year. On the vertical axis I plot average log size for these two different groups before and after 2003. We can see that after four years firms hit by a TFP shock are 17% larger. In contrast, firms in the control group are slightly smaller over the same four years. Moreover, Figure 3 illustrates that HPG firms are smaller than non-HPG firms before they receive the shock but become larger afterward.

I label the pattern depicted in Figure 3 as growing by learning because those firms learning how to improve their production strategies are those that happen to grow more; under this hypothesis, the process of learning and innovation can be interpreted as the mechanism that makes firms thrive.

Figure 3: TFP shock at $s = 0$ and subsequent size growth



Notes. HPG firms are those with TFP growth above 10% in 2003, labeled as $s = 0$. The remaining firms are labeled as “No HPG” in the same year. The vertical axis plots average log size for these two different groups before and after 2003.

Indeed, the learning by growing hypothesis is implicitly assumed in theoretical models of firm dynamics that characterize industries as groups of heterogeneous-productivity firms (e.g. Jovanovic (1982); Melitz (2003)). These models assume that firms learn about their productivity as they operate in the market. Low productivity firms are less likely to survive and thrive than their more efficient counterparts.

Learning by growing represents an alternative rationale to explain the size-productivity nexus. This would indicate that once the firms grow, they become more productive. The economic foundations for this alternative hypothesis are difficult to motivate. The standard profit maximization problem establishes that firms chose their input mix given their productivity shocks. It is questionable to assume that firms chose their productivities given their level of inputs, labor and capital. Indeed, Figure B.3 in the Appendix provides little evidence in favor of this hypothesis. Analogous to Figure 3, Figure B.3 plots the evolution of log TFP before and after an increase in size (number of employees) in 2003. High size growth (HSG) firms do not appear to experience productivity gains after the increase in size.

In any event, the remaining of the paper aims to formally test these two hypotheses using the sample of Spanish firms described above.

4.1 Econometric Approach

I test for the growing by learning (GBL) and learning by growing (LBG) hypotheses by creating control groups using matching techniques based on average treatment models as discussed in Heckman, Ichimura, and Todd (1997). The intuition of this technique is to estimate the effect of productivity (size) growth on size (productivity) growth by matching high-growth-productivity (-size) firms with non-high-growth-productivity (-size) firms. The method constructs a counterfactual which allows analyzing how size (productivity) of a firm would have evolved if it had not received a productivity (size) shock. This counterfactual is based on matching the high-growth-productivity (size) firm with a control group of similar characteristics that do not receive the productivity (size) shock.

I consider two parallel empirical specifications throughout this section: (i) one in which the treatment variable is a productivity shock and the outcome of interest is size growth (growing by learning — GBL); (ii) another one in which the treatment status is defined by a size/employment shock and the outcome variable is productivity growth (learning by growing — LBG).

In order to formally introduce the approach, I first rescale the time periods in such a way that a firm receives the productivity (size) shock at $s = 0$.⁷ Let y_{is} be the outcome at time s —either productivity growth or size growth— of firm i at period s . The variable HGF_i takes the value one if a firm i receives a productivity (or size) shock at $s = 0$. In the baseline estimation, a firm is labeled as a high growth firm if its productivity (or size) growth is above the 75th percentile, which roughly corresponds to an annual growth rate above 10% in both cases. Also, I restrict the sample to firms with more than 10 employees in order to reduce the over-representation of small firms in the population of high-growth firms.⁸

The causal effect can be verified by looking at the average effect of interest:

$$E(y_{is}^1 - y_{is}^0 | HGF_i = 1) = E(y_{is}^1 | HGF_i = 1) - E(y_{is}^0 | HGF_i = 1) \quad (1)$$

where the superscript denotes whether firm i received a shock at $s = 0$ or not. In particular, y_{is}^1 is the outcome if firm i receives a shock while y_{is}^0 denotes the outcome of firm i if it does not receive the shock. Therefore, the crucial problem in this analysis is that the second term in equation (1) is not observable. This is the effect that high growth firms would have experienced, on average, had they not received the shock. In order to identify this term I assume that all differences between high growth firms and the appropriate control group can be captured by a vector of observables

⁷I follow the strategy proposed by DeLoecker (2007) when estimating the productivity gains of exporting.

⁸A widely used definition of high-growth firms establishes that average employment growth must exceed 20% per year over a three-year period and the firm must have 10 or more employees at the start of the period. I do not consider this type of multi-period definition because the balancing hypothesis is not satisfied when using this criterion as explained below.

including the pre-shock size and productivity of a firm. For that purpose, I apply the propensity score matching method (see Rosenbaum and Rubin (1983))⁹ which provides me with an appropriate control group that is as close as possible to the high growth firm in terms of its predicted probability to receive a high growth/productivity shock.

To be more concrete, I model the probability of becoming a high growth firm (the propensity score) as follows:

$$Pr(HGF_{i,0} = 1) = \Phi[h(y_{i,-1}, X_{i,-1})] \quad (2)$$

where $\Phi(\cdot)$ refers to the normal cumulative distribution function. The subscript 1 denotes that the probability of being a high growth firm is regressed on variables prior to $s = 0$. I take a full polynomial in the elements of $h(\cdot)$ as to free up the functional form and improve the resulting matching (Woolridge (2002)). The lagged outcome variable $y_{i,-1}$ is included as a pre-treatment control together with other relevant firm-level covariates, namely, total factor productivity, size, export status, wages, and age. Finally, I estimate the propensity score by year and 2-digit NACE industry to control for common aggregated demand and supply shocks. Let the predicted high growth firm probability be denoted by p_i .¹⁰

The matching is based on the method of the nearest neighbor, which selects a non high growth firm j which has a propensity score p_j closest to that of the corresponding high growth firm (Becker and Ichino (2002) provide more details on this matching procedure). More concretely, I match within each year and 2-digit NACE sector and therefore create control groups within narrowly defined sectors. This matching strategy results in a group of matched HGF and non HGF firms that allows me to evaluate the causal impact of (i) productivity shocks on size growth, and, (ii) size shocks on productivity growth.

Having established a matched sample of treated (HGF) and control (non HGF) firms, I use a difference-in-differences (DID) methodology to estimate the effects of interest. In particular, every HGF firm i is matched with a group of N_i^c control firms denoted as $C(i)$. The weight of each control firm in the group of controls for the treated firm i is given by $w_{ij} = \frac{1}{N_i^c}$ if $j \in C(i)$ and zero otherwise. Moreover, I define y^1 and y^c as the estimated outcomes of the treated and the controls, respectively. Armed with these ingredients, I consider the following estimator in the spirit of DeLoecker (2007):

$$\beta_h^s = \frac{1}{N_s} \sum_i \left(y_{is}^1 - \sum_{j \in C(i)} w_{ij} y_{js}^c \right) \quad (3)$$

⁹This method boils down to estimating a probit model with a dependent variable equal to 1 if a firm is a high growth firm and zero elsewhere on lagged observables.

¹⁰The balancing hypothesis states that for a given propensity score p_i , exposure to treatment is random and thus treated and control units should be on average observationally identical. In order to corroborate that this assumption is satisfied in my application, I perform the algorithm described in Becker and Ichino (2002).

where h denotes the hypothesis being tested, either growing by learning effect ($h = GBL$) or learning by growing effect ($h = LBG$). In the case $h = GBL$, the matching is performed considering as HGF firms those with high productivity growth while the matching is based on high size growth firms for the case $h = LBG$. In both cases, the matching is performed at the time a firm receives the shock and $s = 1, \dots, S$ denotes the time periods after the shock ($s = 0$). Therefore, β_h^s estimates the effect at every period s after the shock. The outcome variable of interest, y_{is} , can be either productivity growth or size growth.

In addition to the impact effect estimated by β_h^s , I also consider the cumulative effect gathered over a period S after the shock:

$$\beta_h^S = \frac{1}{N_S} \sum_i \left(\sum_{s=0}^S y_{is}^1 - \sum_{s=0}^S \sum_{j \in C(i)} w_{ij} y_{js}^c \right) \quad (4)$$

In words, β_h^S in equation (4) measures the productivity/size gains that HGF firms gather over S periods whereas β_h^s in equation (3) estimates the gain at each period s .

4.2 Baseline Results

The estimated growing by learning (GBL) and learning by growing (LBG) effects are reported in Table 3. Rows (a) and (b) refer to the growing by learning hypothesis. In row (a) I present the impact of a productivity shock on employment growth at every period s , whereas row (b) shows the cumulative employment growth effect, reflecting the size growth premium gathered after S periods. Rows (c) and (d) refer to the learning by growing hypothesis. Row (c) presents the estimated impact effects of an employment shock on productivity growth at every period s after the shock, whereas row (D) reports the corresponding cumulative effects.

Rows (a) and (b) illustrate that a productivity shock has a positive and significant treatment effect on size growth, which provides evidence in favour of the growing by learning hypothesis. High productivity-growth firms grow on average 1.7% more the year after they increase their productivity ($s = 1$). This figure is reduced to 1.2% and 1.0% in the second and third years, respectively. The effect on size growth is no longer significantly estimated after 4 years of receiving the productivity shock.

With respect to the GBL cumulative effects, row (b) shows that after 5 years ($s = 5$), employment growth of high productivity-growth firms is on average 8.4% higher than non-HPGF firms. According to these results, a productivity shock at $s = 0$ causes cumulative size growth to be significantly higher even after five years ($s = 5$). Note that the estimated cumulative effects β_h^S are not equal to the sum of the pure time effects β_h^s due to the unbalanced data, i.e., $\sum_s \beta_h^s \neq \beta_h^S$ since N varies with s .

Table 3: Estimated GBL and LBG effects.

s	1	2	3	4	5
(a) Outcome: employment growth					
β_{GBL}^s	0.017*** (0.004)	0.012*** (0.004)	0.010** (0.005)	0.004 (0.007)	0.003 (0.006)
(b) Outcome: cumulative employment growth					
β_{GBL}^S	0.017*** (0.004)	0.037*** (0.005)	0.045*** (0.006)	0.075*** (0.007)	0.084*** (0.009)
(c) Outcome: productivity growth					
β_{LBG}^s	0.017* (0.009)	-0.005 (0.005)	-0.014 (0.020)	-0.008 (0.009)	-0.010 (0.014)
(d) Outcome: cumulative productivity growth					
β_{LBG}^S	0.017* (0.009)	0.012 (0.012)	0.006 (0.017)	0.012 (0.008)	0.007 (0.031)

Notes. $s = 1, \dots, 5$ denotes the time periods after the shock that takes place as $s = 0$. The number of treated firms is 14,135 and the number of controls is 5,502 in rows (a) and (b). In rows (c) and (d) the number of treated firms is 12,533 and the number of controls is 5,697. I denote significance at 10%, 5% and 1% with *, ** and ***, respectively. Standard errors are denoted in parentheses.

The economic magnitude of the estimated GBL effects is also relevant. Consider two identical firms, A and B, with 60 employees. In a given year ($s = 0$), firm A receives a productivity shock while firm B does not. According to the estimates in Table 3, the next year ($s = 1$) firm A will hire an additional worker while firm B will not. After 5 years ($s = 5$), firm A will hire five additional workers but firm B will not.

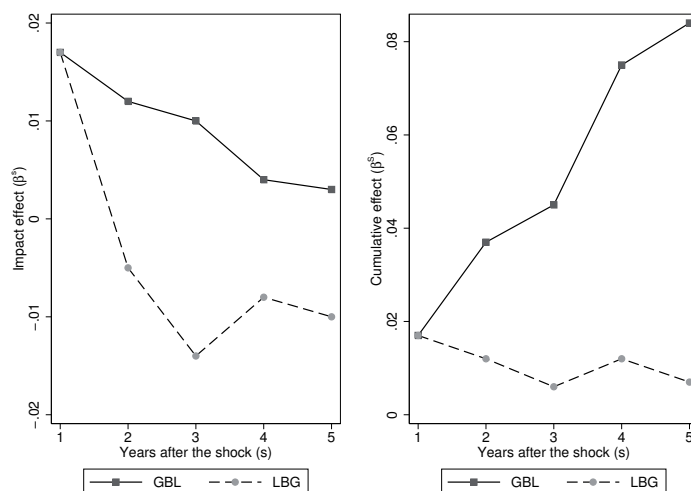
Turning to the LBG estimated effects in rows (c) and (d) of Table 3, they provide little support for the learning by growing hypothesis. To be more concrete, row (c) shows the impact effects of an employment shock on productivity growth at every period. In the first year ($s = 1$) the estimated LBG effect is very similar to the GBL effect in row (a) but it is less precisely estimated being significant only at the 10% level. However, the impact LBG effect vanishes after one year becoming even negative albeit not statistically significant.

The cumulative LBG effects reported in row (d) confirm that productivity growth does not increase after an employment shock. While there is a positive effect during the first year, it is only significant at the 10% level.¹¹ In any event, the effect of employment growth on cumulative productivity growth is not statistically significant after the first year, which confirms that size growth cause no effect on productivity growth.

¹¹Note that cumulative and impact effects coincide in the first year ($s = 1$).

Finally, Figure 4 illustrates the main findings. The estimated impact and cumulative GBL — growing by learning— effects are positive during 5 years after the shock. In particular, the impact GBL effect is decreasing while the GBL cumulative effect increases over time. In contrast, the learning by growing (LBG) effects are even negative on impact and roughly zero in cumulative terms.

Figure 4: Estimated GBL and LBG effects



Notes. The solid line plot the estimated growing by learning (GBL) effects. The dashed line plots the learning by growing (LBG) effects. The left panel shows the cumulative effect (β_h^S) on either employment (GBL) or productivity (LBG) growth while the right panel shows impact effects (β_{LBG}^I).

I interpret these opposing findings as conclusive evidence that productivity growth does cause firm growth, whereas firm growth does not cause productivity growth. The identification assumptions required for a causal interpretation of the GBL effects alone might be controversial. In particular, I need to assume that control and treated firms are identical prior to the productivity shock. However, I can only make them comparable in terms of the observable characteristics included in the propensity score while they might well differ in other non-observable dimensions. Analogously, the same concern applies to the causal interpretation of the LBG effects. The fact that the effect is positive and significant only in one direction (from productivity to size growth) indicates that the causal interpretation of my GBL estimates is not sharply at odds with the data.

4.3 Robustness Analysis

In this section I perform several robustness exercises. In particular, these robustness checks are related to the definition of high growth firms, the matching methodology and the inclusion of financial variables. First, in the baseline a firm is labeled as a high growth firm if its productivity (or size)

growth is above an annual growth rate above 10%, which roughly corresponds to the 75th percentile. Second, in the baseline the matching is based on the method of the nearest neighbor, which selects the non high growth firms with the closest propensity score. I consider below different definitions of high growth firms as well as different matching techniques to investigate the robustness of the baseline results. Third, I re-estimate the baseline specification and consider the asymptotic standard errors derived in Abadie and Imbens (2016), which take into account the estimation error in the propensity score. Fourth, I include the firms' cash-flow among the control variables in the matching to ensure that differences in firm's liquidity do not bias the baseline estimates.

4.3.1 Alternative HGF Definitions

The criterion for defining a high-growth firm (in terms of both size and productivity) determines the treatment and control groups. Therefore, it represents an assumption that might have non-negligible effects on the estimates. In my baseline results HGF firms are those with productivity (or size) annual growth above 10% that result in a control and treatment groups satisfying the balancing hypothesis.¹² Standard definitions of high-growth firms (HGF) in terms of employment are based on multi-period criteria more stringent than the one in my baseline estimates. For instance, the OECD labels as HGF a firm with average employment growth above 20% per year over a three-year period; also, the firm must have 10 or more employees. I cannot consider this type of definition because the resulting groups of treated and control firms do not satisfy the the balancing hypothesis, which is crucial for ensuring the validity of the matching estimates. Therefore, for the robustness exercises I consider two alternative definitions with alternative thresholds for the annual growth rate (in terms of both productivity —GBL— and employment —LBG—). In particular, I consider thresholds of 8% and 12% which roughly correspond to the 70th and 80th percentiles, respectively.

Table 4 shows the estimated GBL and LBG effects in cumulative terms under the labels HGF 8% and HGF 12%. As in the baseline results, GBL effects are highly significant and present similar magnitudes. If I consider a threshold above 12% (such as 20%, which corresponds to the 90th percentile), the balancing hypothesis is not satisfied. Thresholds below 8% (e.g. 4%) correspond to lower percentiles (around 55th) so that the identification of a size/productivity shock is less clear-cut and the HGF labeling becomes more controversial.¹³

The overlapping between HGF in size and productivity represents another concern related to the definition of high growth firms. In particular, 18% of high growth firms in terms of productivity are also high growth firms in terms of size. Analogously, 16% of HGF in size are also HGF in productivity.

¹²In particular, this assumption requires that for a given propensity score p_i , exposure to treatment is random and thus treated and control units should be on average observationally identical (see Becker and Ichino (2002)).

¹³In any event, the main results remain robust when using the 4% threshold and are available upon request.

Therefore, the productivity (size) shock considered in the baseline specification might include a size- (productivity-) shock component. In other words, around on sixth of the firms experiencing a productivity (size) shock also enjoy a size (productivity) shock, which might contaminate the baseline estimates of the GBL and LBG effects. Table 4 reports GBL and LBG effects estimated from a sample that does not include firms that are HGF in terms of both TFP and size.¹⁴ The GBL estimates are even larger in magnitude while the LBG remain statistically insignificant after the second period, which corroborates the robustness of the baseline estimates to different definitions of high growth firms.

Table 4: Estimated GBL and LBG effects — Robustness (I).

s	1	2	3	4	5
Panel A: Effect of productivity on cumulative size growth (β_{GBL}^S)					
HGF 8%	0.022*** (0.004)	0.037*** (0.005)	0.040*** (0.006)	0.057*** (0.007)	0.076*** (0.009)
HGF 12%	0.025*** (0.004)	0.042*** (0.004)	0.054*** (0.005)	0.073*** (0.006)	0.067*** (0.008)
HGF only TFP	0.046*** (0.007)	0.057*** (0.005)	0.056*** (0.007)	0.066*** (0.008)	0.089*** (0.009)
Panel B: Effect of size on cumulative productivity growth (β_{LBG}^S)					
HGF 8%	0.012** (0.005)	0.001 (0.006)	-0.015* (0.008)	-0.015 (0.009)	-0.011 (0.010)
HGF 12%	0.012 (0.012)	0.022 (0.016)	0.004 (0.024)	0.021* (0.011)	0.003 (0.039)
HGF only size	0.015* (0.008)	0.015* (0.009)	0.004 (0.011)	0.006 (0.012)	0.013 (0.014)

Notes. $s = 1, \dots, 5$ denotes the time periods after the shock that takes place as $s = 0$. HGF 8% refers to high growth firms defined as those with TFP (or size) growth above 8% in a given year, which corresponds to the 70th percentile in the sample. HGF 12% identifies firms with TFP (or size) growth above 12%, the 80th percentile. HGF only TFP (size) refer to high growth firms in terms of productivity (size) but not in terms of size (productivity) — see the main text for more details. I denote significance at 10%, 5% and 1% with *, ** and ***, respectively. Standard errors are denoted in parentheses.

4.3.2 Alternative Matching Techniques

The particular technique considered for matching control and treatment firms might also represent a decision with significant effects on the results. In the baseline estimates, I consider the method

¹⁴The estimated effects of the productivity and size shock are positive and strongly significant on employment growth, and negative and marginally significant on productivity growth.

of the nearest neighbor as explained in section 4.1. Two alternative methods are explored in the robustness exercises, kernel and stratification matching. None of the different techniques is ex-ante superior to the others given the trade-off between quality and quantity of the matches, but their joint consideration allows me to assess the robustness of the estimates. Dehejia and Wahba (2002) provide more details on the different matching methods based on the propensity score.

First, I consider kernel matching in which all treated firms are matched with a weighted average of all controls using weights that are inversely proportional to the distance between the propensity scores of treated and controls. In the nearest neighbor method considered in the baseline, for some treated units the nearest neighbor may have a very different propensity score, and, nevertheless, it would contribute to the estimation of the treatment effect independently of this difference. This caveat is addressed by the weighting scheme considered in kernel matching. Second, I consider the stratification method that consists of dividing the propensity score in intervals such that within each interval, treated and control units have on average the same propensity score. Then, within each interval the difference between the average outcomes of the treated and the controls is computed. The overall effect is obtained as an average of the effects for each block with weights given by the distribution of treated units across blocks. Note that with this method, some treated units may be discarded in case no control is available in their block.

Table 5 reports the cumulative GBL and LBG effects estimated when using Kernel and Stratification matching techniques. In the case of GBL cumulative effects in Panel A, the estimates are highly significant and very similar to those in the baseline results based on nearest neighbor matching (see row (b) of Table 3). Turning to the LBG results in Panel B, the estimated effects are more precisely estimated than those in the baseline results for the first two years after the size shock ($s = 1$ and $s = 2$). However, the effects completely vanish after the third period ($s = 3$) as in row (d) of Table 3 based on nearest neighbor matching. All in all, I conclude that the matching technique considered in the baseline results has no decisive impact on the estimated effects.

4.3.3 Other Robustness Checks

In the baseline results, inference is based on standard errors that ignore the estimation error in the propensity score. Abad e and Imbens (2016) take into account the fact that the propensity scores are random variables and are estimated from the data (instead of being constants), and they derive the adjustment to the large sample variance of the estimated treatment effects. Table 6 reports the cumulative GBL and LBG effects and their corresponding standard errors estimated when considering the approach in Abad e and Imbens (2016). In the case of GBL results in Panel A, the estimates are highly significant and similar to those in the baseline results not accounting for the randomness of the propensity scores (see row (b) of Table 3). Regarding the LBG results in Panel B, standard

Table 5: Estimated GBL and LBG effects — Robustness (II).

s	1	2	3	4	5
Panel A: Effect of productivity on cumulative size growth (β_{GBL}^S)					
Kernel matching	0.020*** (0.002)	0.040*** (0.004)	0.057*** (0.009)	0.073*** (0.010)	0.081*** (0.010)
Stratification matching	0.015*** (0.003)	0.035*** (0.005)	0.042*** (0.007)	0.070*** (0.008)	0.079*** (0.010)
Panel B: Effect of size on cumulative productivity growth (β_{LBG}^S)					
Kernel matching	0.018*** (0.004)	0.019** (0.008)	0.001 (0.007)	0.006 (0.009)	0.002 (0.010)
Stratification matching	0.021*** (0.003)	0.022** (0.009)	0.003 (0.008)	0.008 (0.009)	0.009 (0.011)

Notes. $s = 1, \dots, 5$ denotes the time periods after the shock that takes place as $s = 0$. Kernel and stratification matching refer to alternative techniques for constructing the control group (see the main text for more details). I denote significance at 10%, 5% and 1% with *, ** and ***, respectively. Standard errors are denoted in parentheses.

errors are now smaller with respect to the coefficient estimates for the first two years after the size shock ($s = 1$ and $s = 2$). However, the statistical significance of the effects completely vanishes after the third period ($s = 3$) as in row (d) of Table 3. I thus conclude that considering the asymptotic distribution in Abadie and Imbens (2016) has no impact on the conclusions from the baseline results.

The baseline estimates control for selection on observables by including the following firm covariates: lagged high growth status, total factor productivity, size, export status, wages, and age. Financial variables are thus not included in the set of observable characteristics that might correlate with the high-growth indicators. However, access to finance may well be correlated with firms' growth performance in terms of both productivity and employment. If this is the case, a positive effect from high-growth productivity to employment growth could be driven by differences in the access to finance. For instance, high-growth episodes may improve firms' ability to raise funding, and this improvement would generate employment growth rather than the productivity shock per se. In order to control for this possibility, we follow Fazzari, Hubbard, and Petersen (1988) and include firms' cash-flow among the set of firm covariates when performing the matching in the baseline specification.¹⁵ The estimated effects are reported in Table 6 under the label "Access to finance". The findings in the baseline specification remain robust to the inclusion of financial variables in the matching procedure. I can thus conclude that the estimated effect from productivity to employment growth is not driven by differences in access to finance between control and treated firms.

¹⁵Cash-flow is defined as the sum of net income, depreciation, and extraordinary income. Positive cash-flow indicates that a firm's liquid assets are increasing, enabling it to settle debts, reinvest in its business, return money to shareholders, pay expenses and provide a buffer against future financial challenges.

Table 6: Estimated GBL and LBG effects — Robustness (III).

s	1	2	3	4	5
Panel A: Effect of productivity on cumulative size growth (β_{GBL}^S)					
Abad�e and Imbens (2016)	0.019*** (0.004)	0.031*** (0.005)	0.039*** (0.007)	0.073*** (0.009)	0.082*** (0.011)
Access to finance	0.016*** (0.004)	0.033*** (0.005)	0.046*** (0.006)	0.055*** (0.007)	0.072*** (0.009)
Panel B: Effect of size on cumulative productivity growth (β_{LBG}^S)					
Abad�e and Imbens (2016)	0.019*** (0.005)	0.024*** (0.007)	0.003 (0.009)	0.005 (0.012)	0.002 (0.013)
Access to finance	0.023*** (0.005)	0.018*** (0.006)	0.002 (0.023)	-0.005 (0.032)	0.004 (0.010)

Notes. $s = 1, \dots, 5$ denotes the time periods after the shock that takes place as $s = 0$. “Abad e and Imbens (2016)” refers to the use of the asymptotic standard errors derived in Abad e and Imbens (2016), which take into account the estimation error in the propensity score. “Access to finance” refers to the inclusion of firms’ cash-flow among the control variables in the matching to ensure that differences in firm’s financial restrictions do not bias the baseline estimates. I denote significance at 10%, 5% and 1% with *, ** and ***, respectively. Standard errors are denoted in parentheses.

5 Discussion

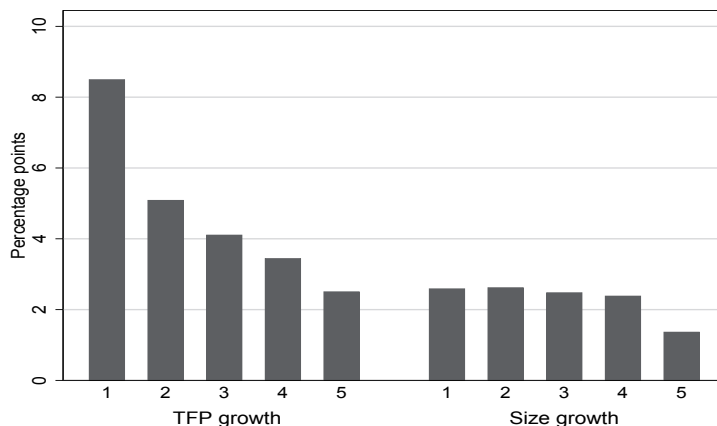
Three empirical facts have been discussed above: (i) large firms are more productive than small firms; (ii) the direction of causality appears to go from productivity to size as it is usually assumed in canonical models of firm dynamics; (iii) there are too many small and low-productivity firms in Spain in comparison with the rest of EU4 countries. A natural interpretation of these findings is that Spanish firms are smaller because they are less productive. According to the GBL estimates in section 4, Spanish firms should experience, on average, fewer and/or smaller TFP shocks than their European counterparts. This lack of productivity increases would result in lower employment growth rates, and, as a consequence, it would bias the Spanish firm size distribution towards small firms.

Figure 5 provides evidence in favor of this pattern by illustrating that average productivity growth at the firm level is substantially lower in Spain than in the remaining EU4 countries. In particular, TFP growth for firms with 1-9 employees is 8.5 pp. lower in Spain while this difference is 2.4 pp. for large firms with more than 250 employees.¹⁶ The fact that the TFP growth gap is substantially

¹⁶Differences in TFP growth between EU4 and Spanish firms are also positive and large for all size categories according to the COMPNET database.

larger for small firms is compatible with the larger gap of TFP levels illustrated in Figure 2. Turning to employment (size) growth, the differences are also positive for all size categories, which means that Spanish firms grow less than their European counterparts. To be more concrete, the annual growth rate of EU4 firms is around 2.5 pp. higher than that of Spanish firms. According to my interpretation, Spanish firms are smaller and less productive mainly because they experience much smaller productivity shocks that preclude them to thrive.

Figure 5: EU4-Spain difference in annual TFP/size growth



Notes. Each bar plots the difference measured in percentage points between annual productivity (and size defined by employment) growth in Spain and average productivity (and size) growth in the remaining EU4 countries (Germany, France, and Italy). The five size categories reported are: (1) from 1 to 9 employees; (2) from 10 to 19 employees; (3) from 20 to 49 employees; (4) from 50 to 249 employees; (5) more than 250 employees. The source of TFP growth is AMADEUS over the period 2004-2012 (see Section 3 for more details) while size growth figures are taken from EUROSTAT over the period 2002-2013. All the figures refer to the total economy including manufacturing, trade, construction and services.

This interpretation raises an obvious and crucial question. Why do Spanish firms are less productive than their European counterparts? Syverson (2011) provides an in-depth overview of the literature on the determinants of productivity at the firm level. Managerial talent, quality of inputs, and R&D activities are identified as internal drivers of measured productivity differences across producers. Cross-country evidence reveals that Spanish firms perform worse than their EU4 counterparts in all the three factors, which may explain, at least partially, their dismal performance in terms of productivity.

Managerial talent represents an obvious lever that can impact the productivity of a firm. Indeed, management can be interpreted as an unmeasured input in the production function. Managers coordinate the application of labor, capital, and intermediate inputs, which determines the ability of

the firm to convert inputs into output. Data limitations have precluded the measuring of managerial talent across firms and countries until the World Management Survey (WMS) was released a decade ago. Bloom and Reenen (2007) and their team collected plant-level management practices data across multiple sectors and countries. The measured practices revolve around day-to-day and close-up operations on eighteen specific management practices in operations, monitoring, targets, and incentives. According to the latest release of the survey, in a scale from 1 to 5 the average score of Spanish firms was 2.75. This figure was 3.18, 3.00, and 2.95 in the cases of Germany, France, and Italy, respectively (see Bloom, Lemos, Sadun, Scur, and Reenen (2014)). To gauge the magnitude of these differences, which are economically important, note that the highest score, 3.29, corresponds to the US, while the lowest, 2.03, corresponds to Mozambique. I argue that poor managerial practices in Spanish firms may partially explain their low productivity. In fact, Bloom and Reenen (2007) show that the lower average management practice scores in some developing countries are driven by a large left tail of very poorly management practices, which goes in line with the larger gap in productivity of small Spanish firms with respect to EU4 firms.

The productive effects of innovation, information technology, and research and development (R&D) have been the subject of intense study. For instance, Doraszelski and Jaumandreu (2013) find that R&D explains a substantial amount of productivity growth using a panel of Spanish firms, and they also develop a model in which firm productivity growth is the result of R&D expenditures with uncertain outcomes.¹⁷ While R&D expenditures is only one of the many components of firms' innovative efforts, it is easy to measure. Indeed, cross-country statistics indicate that gross domestic expenditure on R&D is substantially lower in Spain than in other European countries. Over the years 2003-2014, the average R&D expenditure as a share of GDP was 1.22% in Spain, and 1.17%, 2.14%, and 2.63% in Italy, France, and Germany, respectively. On the other hand, the share of innovative firms is 33.6% in Spain according to EUROSTAT. However, this figure is 56.1%, 53.4%, and 66.9% in Italy, France, and Germany, respectively.¹⁸ It seems reasonable to think that the low innovative content of Spanish firms may also explain part of their productivity gap with respect to their European counterparts.

Standard input measures such as the number of employees or employee-hours do not capture differences in input quality. However, there exists a broad literature based on matched employer-employee datasets that documents the importance of labor quality as a driver of productivity. For instance, Ilmakunnas, Maliranta, and Vainiomäki (2004) show that workers' productivity is increasing in workers' education. Along these lines, there are reasons to think that Spanish workers are, on

¹⁷Rochina-Barrachina, Manez, and Sanchis-Llopis (2010) also find that the implementation of process innovations produces an extra productivity growth using a panel of Spanish firms.

¹⁸The proportion of innovative firms in Spain is also lower by size categories.

average, less productive than their EU4 counterparts. In particular, data from The Programme for the International Assessment of Adult Competencies —PIAAC— corroborates this concern. The PIAAC is a worldwide study of cognitive and workplace skills aiming to assess the skills of literacy, numeracy and problem solving in technology-rich environments among the working-age population (for more details see OECD (2013)). The average score is 249 in Spain, while it is 249, 258, and 271 in Italy, France, and Germany, respectively. These figures indicate that workplace skills of employees in Spanish firms are worse on average, which may well result in lower productivity at the firm level.

6 Concluding Remarks

Large firms are more productive than small firms. The levers that firm managers can use to increase size and productivity are the same, namely, managerial talent, employee quality, and/or innovation. Therefore, establishing the direction of causality between size and productivity represents a challenge. This paper investigates this issue by exploiting a bidirectional identification strategy together with standard matching methods applied to a representative sample of Spanish firms. My results indicate that productivity growth may cause employment growth. In contrast, size growth defined by employment does not result in significant productivity gains.

This finding casts doubt on the view that low productivity growth in Spain can be explained by the presence of too many small firms. In particular, I argue that the low productivity of Spanish firms with respect to their European counterparts might be at the root of the Spanish size distribution excessively biased towards small firms rather than the other way around. As a consequence, the low productivity at the firm level could also explain the dismal evolution of aggregate productivity in Spain. According to my own interpretation of the literature and the international evidence, poor managerial talent, low quality of the labor input, and scarce R&D activities by Spanish firms may explain, at least partially, their dismal performance in terms of productivity.

References

- ABADÍE, A., AND G. IMBENS (2016): “Matching of the Estimated Propensity Score,” *Econometrica*, 84, 781–807.
- BECKER, S., AND A. ICHINO (2002): “Estimation of average treatment effects based on propensity scores,” *STATA Journal*, 2, 358377.
- BLOOM, N., R. LEMOS, R. SADUN, D. SCUR, AND J. V. REENEN (2014): “JEEA-FBBVA Lecture 2013: The New Empirical Economics of Management,” *Journal of the European Economic Association*, 12, 835876.
- BLOOM, N., AND J. V. REENEN (2007): “Measuring and Explaining Management Practices Across Firms and Countries,” *Quarterly Journal of Economics*, 122, 13511408.
- DAUNFELDT, S., N. ELERT, AND D. JOHANSSON (2010): “The Economic Contribution of High-Growth Firms: Do Definitions Matter?,” Stockholm: The Ratio Institute.
- DEHEJIA, R., AND S. WAHBA (2002): “Propensity Score-Matching Methods for Nonexperimental Causal Studies,” *The Review of Economics and Statistics*, 84, 151–161.
- DELOECKER, J. (2007): “Do Exports Generate Higher Productivity? Evidence from Slovenia,” *Journal of International Economics*, 73, 69–98.
- DORASZELSKI, U., AND J. JAUMANDREU (2013): “R&D and Productivity: Estimating Endogenous Productivity,” *Review of Economic Studies*, 80, 1338–1383.
- DU, J., Y. GONG, AND Y. TEMOURI (2013): “High Growth Firms and Productivity - Evidence from the United Kingdom,” Nesta Working Paper No. 13/04.
- FARINAS, J., AND E. HUERGO (2015): “Demografía Empresarial en España: Tendencias y Regularidades,” FEDEA No. 2015/24.
- FAZZARI, S., G. HUBBARD, AND B. PETERSEN (1988): “Financing Constraints and Corporate Investment,” *Brookings Papers on Economic Activity*, 1, 141–195.
- GAL, P. (2013): “Measuring Total Factor Productivity at the Firm Level using OECD-ORBIS,” OECD Economics Department Working Papers, No. 1049.
- GARCÍA-SANTANA, M., AND J. PIJOAN-MAS (2014): “The Reservation Laws in India and the Misallocation of Production Factors,” *Journal of Monetary Economics*, 66, 193–209.

- GARICANO, L., C. LELARGE, AND J. VAN-REENEN (2015): “Firm Size Distortions and the Productivity Distribution: Evidence from France,” FEDEA No. 2015/24.
- GOURIO, F., AND N. ROYS (2014): “Size dependent regulations, firm size distribution, and reallocation,” *Quantitative Economics*, 5, 377–416.
- GUNER, N., G. VENTURA, AND D. XU (2008): “Macroeconomic Implications of Size-Dependent Policies,” *Review of Economic Dynamics*, 11(4), 721–744.
- HECKMAN, J., H. ICHIMURA, AND P. TODD (1997): “Matching as an econometric evaluation estimator,” *The Review of Economic Studies*, 65, 261–294.
- HENREKSON, M., AND D. JOHANSSON (2010): “Gazelles as job creators: a survey and interpretation of the evidence,” *Small Business Economics*, 35, 227–244.
- ILMAKUNNAS, P., M. MALIRANTA, AND J. VAINIOMÄKI (2004): “The Roles of Employer and Employee Characteristics for Plant Productivity,” *Journal of Productivity Analysis*, 21, 249–276.
- IMF (2015): “Obstacles to Firm Growth in Spain,” IMF Country Report No. 15/233.
- JOVANOVIC, B. (1982): “Selection and the Evolution of Industry,” *Econometrica*, 50, 649–670.
- LEVINSOHN, J., AND A. PETRIN (2003): “Estimating Production Functions Using Inputs as Controls for Unobservables,” *The Review of Economic Studies*, 70, 317–342.
- LOPEZ-GARCIA, P., F. DI MAURO, AND OTHERS (2015): “Assessing European competitiveness: the new CompNet microbased database,” European Central Bank Working Papers, No. 1764.
- LOPEZ-GARCIA, P., AND S. PUENTE (2012): “What makes a high-growth firm? A dynamic probit analysis using Spanish firm-level data,” *Small Business Economics*, 39, 1029–1041.
- MAS, M., F. PÉREZ, AND E. URIEL (2013): “Inversión y stock de capital en España (1964-2011). Evolución y perspectivas del patrón de acumulación,” *Bilbao: Fundación BBVA*.
- MELITZ, M. (2003): “The Impact of Trade on Intra-Industry Reallocations and Aggregate Industry Productivity,” *Econometrica*, 71(1965-1725).
- OECD (2013): *OECD Skills Outlook 2013: First Results from the Survey of Adult Skills*. OECD Publishing.
- OLLEY, S., AND A. PAKES (1996): “The Dynamic of Productivity in the Telecommunications Equipment Industry,” *Econometrica*, 64, 1263–1297.

- RAMOS, R., AND E. MORAL-BENITO (2013): “Agglomeration Matters for Trade,” Banco de España Working Paper 1316.
- ROCHINA-BARRACHINA, M., J. MANEZ, AND J. SANCHIS-LLOPIS (2010): “Process innovations and firm productivity growth,” *Small Business Economics*, 34, 147–166.
- ROSENBAUM, P., AND D. RUBIN (1983): “The central role of the propensity score in observational studies for causal effects,” *Biometrika*, 70, 41–55.
- SYVERSON, C. (2011): “What Determines Productivity?,” *Journal of Economic Literature*, 49, 326–365.
- WOOLDRIDGE, J. (2009): “On estimating firm-level production functions using proxy variables to control for unobservables,” *Economics Letters*, 104, 112–114.
- WOOLDRIDGE, J. (2002): “Econometric Analysis of Cross Sections and Panel Data,” The MIT Press, Cambridge.

A Measuring firm-level TFP

In order to obtain firm-level total factor productivity (TFP), I resort to relatively standard estimation techniques, namely, the Wooldridge (2009) GMM approach to implement the Olley and Pakes (1996) and Levinsohn and Petrin (2003) identification strategy.

The object of interest is the TFP of firm i in year t , labeled as a_{it} in the following equation derived from a firm-specific Cobb-Douglas production function:

$$y_{it} = \alpha_L l_{it} + \alpha_K k_{it} + \alpha_M m_{it} + a_{it} \quad (5)$$

where y refers to logged output, and l , k , and m are logged labour, capital, and materials, respectively.

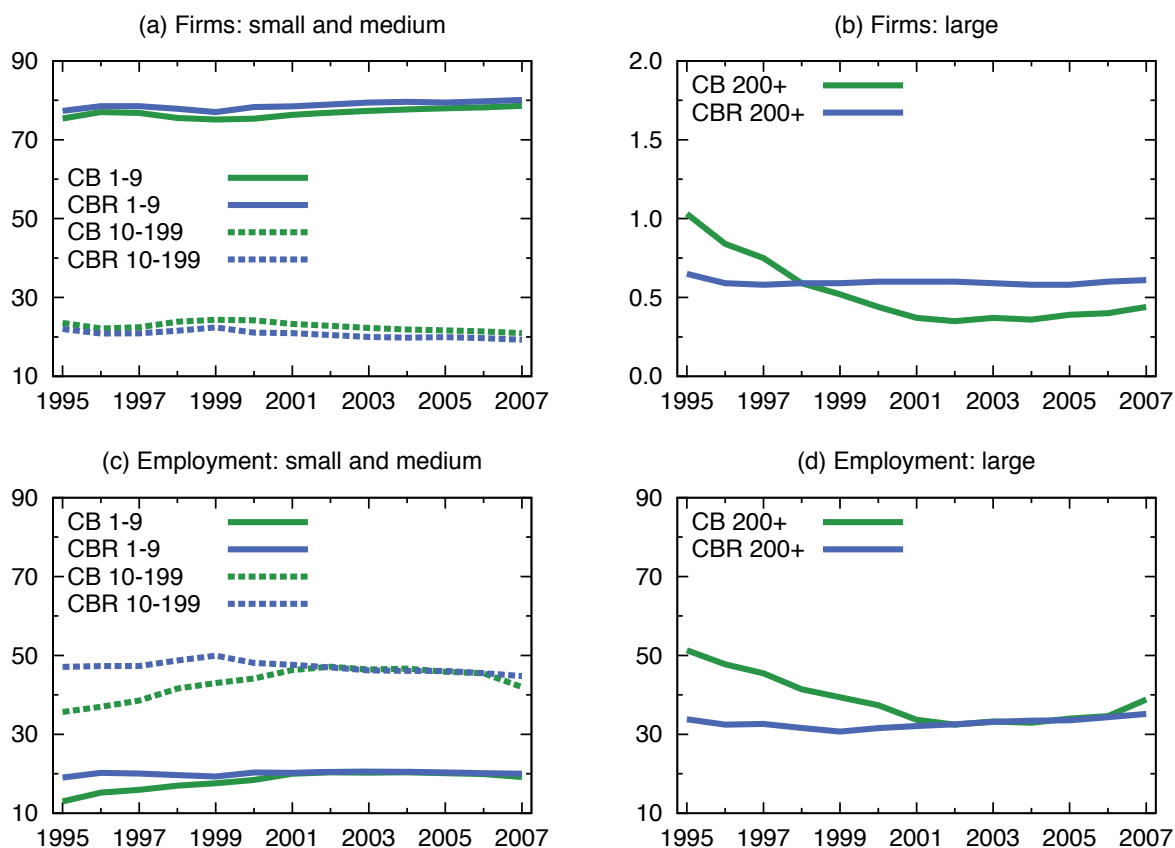
The estimation of equation (5) is performed on a 2-digit industry level. However, in order to obtain consistent estimates with sufficient degrees of freedom, a cutoff of a minimum of 25 observations per sector and year is introduced. Sectors that do not meet the minimum cutoff are flagged and their TFP estimates are replaced by an estimated value obtained on the corresponding macro-sector level. A full set of year dummies is included to control for sector-specific trends.

To estimate the parameters in the production function, I assume that a_{it} is the sum of two firm-specific and unobserved components, namely, a component which is known to the firm (ω_{it}), and a component unknown to the firm and with no impact on firm's decisions (v_{it}). The endogeneity problem that renders OLS estimates biased when estimating equation (5) arises from the correlation of ω_{it} with the input choices. One of the solutions provided for solving this problem is introduced by Olley and Pakes (1996) who proposed a structural approach to the problem, by using observed input choices to instrument for unobserved productivity. In particular, their approach relied upon the assumption that investment, i_{it} , installed in period t only becomes productive at $t + 1$, so that $i_{it} = i(\omega_{it}; k_{it})$ can be inverted to yield $\omega_{it} = \omega(i_{it}; k_{it})$ under the assumption of increasing monotonicity of i_{it} in ω_{it} .

Levinsohn and Petrin (2003) argued that the strict monotonicity of the investment function was broken given the many zeros reported by firms. If all observations with zero or negative investment had to be dropped, there would be an important efficiency loss in the estimation. Therefore, Levinsohn and Petrin (2003) advocate to invert the materials' demand function $m_{it} = m(\omega_{it}; k_{it})$ to obtain $\omega_{it} = \omega(m_{it}; k_{it})$, also under monotonicity plus some additional assumptions. Finally, Wooldridge (2009) introduces an appealing GMM approach that implements this identification strategy reducing the computational burden as well as producing more efficient estimates.

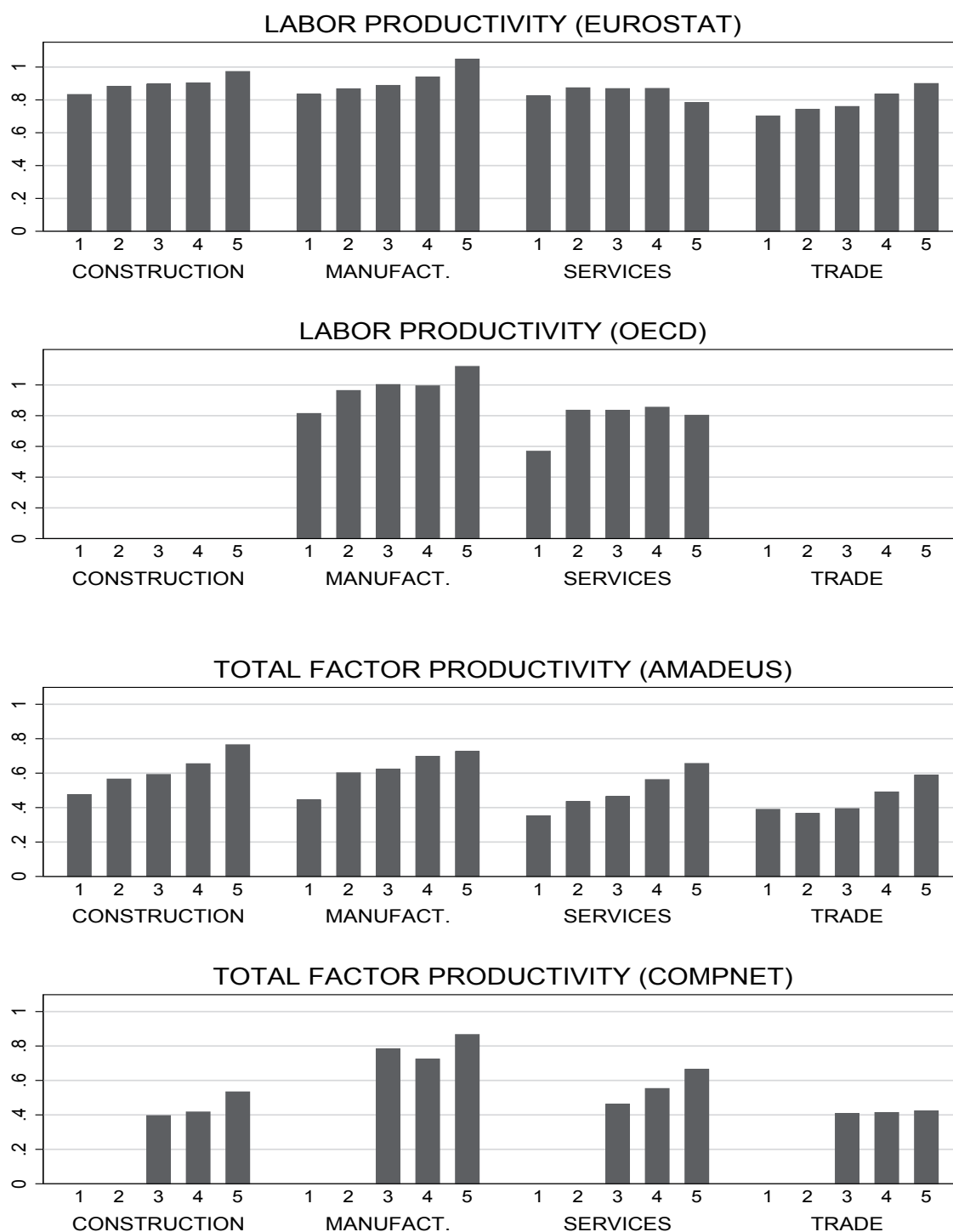
B Additional Figures

Figure B.1: Size distribution of firms, 1995-2007



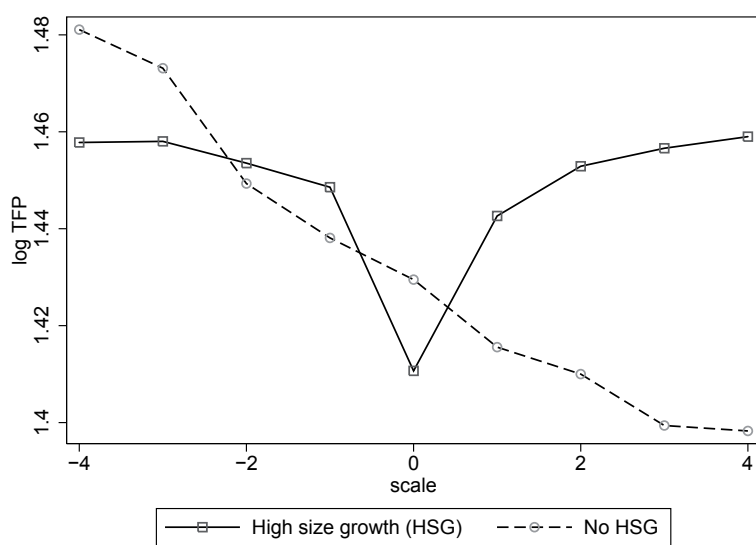
Notes. Panel (a) plots the percentage number of firms with 1-9 workers and 10-199 workers both for our sample from the *Central de Balances* (CB) and for the census from the Central Business Register (CBR). Panel (b) does the same for firms with 200+ employees. Panels (c) and (d) report the employment shares in the same size categories.

Figure B.2: Spain-to-EU4 productivity ratio by firm size and sector



Notes. Each bar plots the ratio of average productivity in Spain to average productivity in the remaining EU4 countries (Germany, France, and Italy). The five size categories reported are: (1) from 1 to 9 employees; (2) from 10 to 19 employees; (3) from 20 to 49 employees; (4) from 50 to 249 employees; (5) more than 250 employees. Labor productivity figures cover the period 2002-2013 in the case of EUROSTAT and the year 2010 in the case of the OECD. Total factor productivity ratios from AMADEUS and COMPNET refer to the period 2004-2012.

Figure B.3: Size shock at $s = 0$ and subsequent TFP growth



Notes. HSG firms are those with employment growth above 10% in 2003, labeled as $s = 0$. The remaining firms are labeled as “No HSG” in the same year. The vertical axis plots average log TFP for these two different groups before and after 2003.

BANCO DE ESPAÑA PUBLICATIONS

WORKING PAPERS

- 1501 LAURA HOSPIDO and EVA MORENO-GALBIS: The Spanish productivity puzzle in the Great Recession.
- 1502 LAURA HOSPIDO, ERNESTO VILLANUEVA and GEMA ZAMARRO: *Finance for all*: the impact of financial literacy training in compulsory secondary education in Spain.
- 1503 MARIO IZQUIERDO, JUAN F. JIMENO and AITOR LACUESTA: Spain: from immigration to emigration?
- 1504 PAULINO FONT, MARIO IZQUIERDO and SERGIO PUENTE: Real wage responsiveness to unemployment in Spain: asymmetries along the business cycle.
- 1505 JUAN S. MORA-SANGUINETTI and NUNO GAROUPA: Litigation in Spain 2001-2010: Exploring the market for legal services.
- 1506 ANDRES ALMAZAN, ALFREDO MARTÍN-OLIVER and JESÚS SAURINA: Securitization and banks' capital structure.
- 1507 JUAN F. JIMENO, MARTA MARTÍNEZ-MATUTE and JUAN S. MORA-SANGUINETTI: Employment protection legislation and labor court activity in Spain.
- 1508 JOAN PAREDES, JAVIER J. PÉREZ and GABRIEL PEREZ-QUIRÓS: Fiscal targets. A guide to forecasters?
- 1509 MAXIMO CAMACHO and JAIME MARTINEZ-MARTIN: Monitoring the world business cycle.
- 1510 JAVIER MENCÍA and ENRIQUE SENTANA: Volatility-related exchange traded assets: an econometric investigation.
- 1511 PATRICIA GÓMEZ-GONZÁLEZ: Financial innovation in sovereign borrowing and public provision of liquidity.
- 1512 MIGUEL GARCÍA-POSADA and MARCOS MARCHETTI: The bank lending channel of unconventional monetary policy: the impact of the VLTROs on credit supply in Spain.
- 1513 JUAN DE LUCIO, RAÚL MÍNGUEZ, ASIER MINONDO and FRANCISCO REQUENA: Networks and the dynamics of firms' export portfolio.
- 1514 ALFREDO IBÁÑEZ: Default near-the-*default*-point: the value of and the distance to default.
- 1515 IVÁN KATARYNIUK and JAVIER VALLÉS: Fiscal consolidation after the Great Recession: the role of composition.
- 1516 PABLO HERNÁNDEZ DE COS and ENRIQUE MORAL-BENITO: On the predictability of narrative fiscal adjustments.
- 1517 GALO NUÑO and CARLOS THOMAS: Monetary policy and sovereign debt vulnerability.
- 1518 CRISTIANA BELU MANESCU and GALO NUÑO: Quantitative effects of the shale oil revolution.
- 1519 YAEL V. HOCHBERG, CARLOS J. SERRANO and ROSEMARIE H. ZIEDONIS: Patent collateral, investor commitment and the market for venture lending.
- 1520 TRINO-MANUEL NÍGUEZ, IVAN PAYA, DAVID PEEL and JAVIER PEROTE: Higher-order risk preferences, constant relative risk aversion and the optimal portfolio allocation.
- 1521 LILIANA ROJAS-SUÁREZ and JOSÉ MARÍA SERENA: Changes in funding patterns by Latin American banking systems: how large? how risky?
- 1522 JUAN F. JIMENO: Long-lasting consequences of the European crisis.
- 1523 MAXIMO CAMACHO, DANILO LEIVA-LEON and GABRIEL PEREZ-QUIROS: Country shocks, monetary policy expectations and ECB decisions. A dynamic non-linear approach.
- 1524 JOSÉ MARÍA SERENA GARRALDA and GARIMA VASISHTHA: What drives bank-intermediated trade finance? Evidence from cross-country analysis.
- 1525 GABRIELE FIORENTINI, ALESSANDRO GALESÌ and ENRIQUE SENTANA: Fast ML estimation of dynamic bifactor models: an application to European inflation.
- 1526 YUNUS AKSOY and HENRIQUE S. BASSO: Securitization and asset prices.
- 1527 MARÍA DOLORES GADEA, ANA GÓMEZ-LOSCOS and GABRIEL PEREZ-QUIROS: The Great Moderation in historical perspective. Is it that great?
- 1528 YUNUS AKSOY, HENRIQUE S. BASSO, RON P. SMITH and TOBIAS GRASL: Demographic structure and macroeconomic trends.
- 1529 JOSÉ MARÍA CASADO, CRISTINA FERNÁNDEZ and JUAN F. JIMENO: Worker flows in the European Union during the Great Recession.
- 1530 CRISTINA FERNÁNDEZ and PILAR GARCÍA PEREA: The impact of the euro on euro area GDP per capita.
- 1531 IRMA ALONSO ÁLVAREZ: Institutional drivers of capital flows.
- 1532 PAUL EHLING, MICHAEL GALLMEYER, CHRISTIAN HEYERDAHL-LARSEN and PHILIPP ILLEDITSCH: Disagreement about inflation and the yield curve.
- 1533 GALO NUÑO and BENJAMIN MOLL: Controlling a distribution of heterogeneous agents.

- 1534 TITO BOERI and JUAN F. JIMENO: The unbearable divergence of unemployment in Europe.
- 1535 OLYMPIA BOVER: Measuring expectations from household surveys: new results on subjective probabilities of future house prices.
- 1536 CRISTINA FERNÁNDEZ, AITOR LACUESTA, JOSÉ MANUEL MONTERO and ALBERTO URTASUN: Heterogeneity of markups at the firm level and changes during the great recession: the case of Spain.
- 1537 MIGUEL SARMIENTO and JORGE E. GALÁN: The influence of risk-taking on bank efficiency: evidence from Colombia.
- 1538 ISABEL ARGIMÓN, MICHEL DIETSCH and ÁNGEL ESTRADA: Prudential filters, portfolio composition and capital ratios in European banks.
- 1539 MARIA M. CAMPOS, DOMENICO DEPALO, EVANGELIA PAPAPETROU, JAVIER J. PÉREZ and ROBERTO RAMOS: Understanding the public sector pay gap.
- 1540 ÓSCAR ARCE, SAMUEL HURTADO and CARLOS THOMAS: Policy spillovers and synergies in a monetary union.
- 1601 CHRISTIAN CASTRO, ÁNGEL ESTRADA and JORGE MARTÍNEZ: The countercyclical capital buffer in Spain: an analysis of key guiding indicators.
- 1602 TRINO-MANUEL NÍGUEZ and JAVIER PEROTE: Multivariate moments expansion density: application of the dynamic equicorrelation model.
- 1603 ALBERTO FUERTES and JOSÉ MARÍA SERENA: How firms borrow in international bond markets: securities regulation and market segmentation.
- 1604 ENRIQUE ALBEROLA, IVÁN KATARYNIUK, ÁNGEL MELGUIZO and RENÉ OROZCO: Fiscal policy and the cycle in Latin America: the role of financing conditions and fiscal rules.
- 1605 ANA LAMO, ENRIQUE MORAL-BENITO and JAVIER J. PÉREZ: Does slack influence public and private labour market interactions?
- 1606 FRUCTUOSO BORRALLÓ, IGNACIO HERNANDO and JAVIER VALLÉS: The effects of US unconventional monetary policies in Latin America.
- 1607 VINCENZO MERELLA and DANIEL SANTABÁRBARA: Do the rich (really) consume higher-quality goods? Evidence from international trade data.
- 1608 CARMEN BROTO and MATÍAS LAMAS: Measuring market liquidity in US fixed income markets: a new synthetic indicator.
- 1609 MANUEL GARCÍA-SANTANA, ENRIQUE MORAL-BENITO, JOSEP PIJOAN-MAS and ROBERTO RAMOS: Growing like Spain: 1995-2007.
- 1610 MIGUEL GARCÍA-POSADA and RAQUEL VEGAS: Las reformas de la Ley Concursal durante la Gran Recesión.
- 1611 LUNA AZAHARA ROMO GONZÁLEZ: The drivers of European banks' US dollar debt issuance: opportunistic funding in times of crisis?
- 1612 CELESTINO GIRÓN, MARTA MORANO, ENRIQUE M. QUILIS, DANIEL SANTABÁRBARA and CARLOS TORREGROSA: Modelling interest payments for macroeconomic assessment.
- 1613 ENRIQUE MORAL-BENITO: Growing by learning: firm-level evidence on the size-productivity nexus.