UNIVERSITAT DE BARCELONA
SCHOOL OF ECONOMICS

MSc Economics

Invention and sales of new products
Is the firm or the region key? *

Thor Donsby Noe † [a]

[a]Universitat de Barcelona, Spain

January 11, 2019
Project outline

Abstract

In this paper determinants of the sales of new products as a share of total sales is analyzed for Spanish firms for the period 2005-2016. Using the PITEC database the region of the firm is identified through the main location of its R&D personnel. Though regional controls sort out some of the heterogeneity firm-level R&D activity greatly surpasses potential regional spillovers from R&D spending in other firms. Especially continuous internal research drives the product innovation in the long-run but the more effective areas of research differ between manufacturing and services.

Keywords Research & Development • Product innovation • Research continuity • Spanish firms • Regional spillovers

*I am thankful to Enrique López-Bazo for his useful supervision and inspiring teaching of the course Regional & Urban Economics. I also acknowledge to him the identification strategy in which to assign firms to regions in the PITEC database. I would like to thank Signe Rix and Katrine Thornfeldt for their helpful inputs. All errors remain my own.

†*E-mail address:* thornoe@gmail.com (T. D. Noe).

1. Introduction

In a global competition a firm's ability to innovate and come up with new and successful products is crucial for the development of the firm as well as regional growth and jobs.

Innovations is not just the result of isolated research and development (R&D) processes within firms but can also be expected to be driven up by knowledge spillovers from a high share of R&D spending in other firms within the same region by the means of formal and informal interactions between firms and employees of different firms. Though I do not find significant spillovers from regional R&D expenses a few of the regional dummies correct some of the heteroscedasticity in the estimation. The internal R&D activity seem to be the more important, especially the continuity of in-house research.

I proceed by outlining the theoretical background in section 2, the empirical strategy and the data in section 3 and 4 respectively, before presenting the results in 6 and concluding in 6.

2. Background

2.1. Theory of firm-level innovations

As point of departure I use the same theoretical framework which Sternberg and Arndt (2001) use to analyze the probability of innovations in 10 different European regions. As a region level proxy for regional knowledge-spillovers I let the average share of R&D spending relative to total sales in other firms within the region affect firm-level innovations.

As analyzed by Harrison et al. (2014), though primarily positive, the effect of R&D spending and introduction of new products in other firms can be expected to capture effects in two directions: while there can be positive spill-overs through development, interaction and exchange of human capital and technology there is also the risk that sales in other firms might take over market shares. As an elaboration one could expect that

2.2. Results in other studies

Barrios Cobos et al. (2001) measure the effects of R&D spendings and regional spillovers on firms' export from Spain. Harrison et al. (2014) estimate the effects of innovations on employment growth in Spanish firms using the PITEC database. Vogel (2015) find effects of R&D spendings and evidence of local technology spillovers through her estimation of a long run model for convergence in Total Factor Productivity (TFP) between all European regions. Though they find that both region-level variables and firm-level variables are significant, Sternberg and Arndt (2001) found regional characteristics and regional between-firm spillovers to be of less importance than the characteristics of the individual firm itself.

3. Empirical strategy

The aim is to model the firm level inventions, measured as share of sales due to new products, as a function of within-firm R&D spendings as well as the average share R&D spendings in other firms within the region. That is, for firms within the same sector and in other sectors respectively. I use a panel regression controlling for time-constant firm-level random effects using GLS which provides efficient, though not consistent estimates due to the presence of endogeneity. As opposed to a Fixed Effects model this however allows for both the inclusion of time-invariant controls as well as using cluster robust standard errors.[1]

4. Data

The PITEC database is used which unfortunately is not exhaustive nor representative.[2] The Spanish region of the firm is identified using the firm's location based on information on the region in which they have more than 50% of their R&D department.[3] Other criteria are tried out as well in appendix B and C.

The baseline identification strategy for regions implies that the modified dataset will be limited to firms that have some internal R&D personnel employed for at least one of the years and do not have their internal R&D activities spread out to a degree where no region holds more than half of them.

The dependent variables includes products invented within the same year t as well as in the past two years. In the regression specification I include this variable as a lead one year ahead in period $t+1$. Thus, effectively the regressions is done for the period $t \in 2005, \ldots, 2015$ where all of the explanatory variables are either for year t or time-invariant.

Dependent variable at time $t+1$

- *Sales of new products in year t+1:* Share of total sales constituted by products, invented within the year or the past two years $(t+1, t, t-1)$.

Region-level variable of R&D expenses at time t

- *log R&D expenses at region-level:* Regional-level aggregate of R&D expenses as the percentage share of total sales.
  - The aggregate region-level share is calculated for each firm after deducting the R&D expenses and total sales for the firm.

[1] The STATA do-files can be accessed from github.com/thornoe/ub/tree/master/Regional_Urban/paper/stata_code

[2] Description of the methodology, the full questionnaires etc. in PITEC is available by the Spanish Foundation of Science & Technology at icono.fecyt.es/pitec

[3] This novel identification strategy for regions is an idea by Enrique López-Bazo.

  - Calculated each year for all firms in the PITEC database despite sector or missing information in other years.
  - If firm-level share is above the threshold C for one year, the firm's R&D expenses as well as total sales is left out of the sum for that year only.

Firm-level variables of R&D expenses at time t

- *log R&D expenses*: Firm-level R&D expenses as the percentage share of total sales. An intermediate variable created only for calculating the other variables and truncating the sample.
  - Constructed from dividing two simulated variables in PITEC which are imputed from the remaining Survey information. The extreme observations are as high as several million percent, especially within the service sector. While start-ups can run with higher expenses than sales for some years the more extreme values are more probably due to either poor imputation or measurement errors in the Survey data.
  - To contain a balanced panel, firms are dropped from the regression sample if having a value above the threshold point C for one or more year.
  - In the baseline specification C is set to 400 to maintain a high sample size.
- *log internal R&D expenses:* Internal R&D expenses as the percentage share of total sales.
- *log external R&D expenses:* External R&D expenses as the percentage share of total sales.

Firm-level variables of continuity of R&D expenses, time-invariant

- *Continuity of basic research; Continuity of applied research; Continuity of technological development*

  0 *Never:* expenses on the R&D activity in 0 out of 12 years.

  1 *Occasional:* expenses on the R&D activity in 1-6 out of 12 years.

  2 *Frequent:* expenses on the R&D activity in 6-11 out of 12 years.

  3 *Permanent:* expenses on the R&D activity in all 12 years.
- *Continuity of external R&D only*: Same criteria but only counting years with none of the internal R&D activities above.

Control dummies

- *Region dummies:* Dummies for each of the 17 autonomous communities of Spain.
  - In the baseline regression the region is assigned if more than 50% of the R&D personnel of the firm is located within that region. If for one year no region can be identified for the firm, the region is imputed from potential former or secondarily latter years such that the firm is not left out from the panel.

  - Modifications to this identification strategies are used for estimation in appendix B and C.

- *Year dummies:* For all years. Controlling for general time trends.
- *Industry dummies:* Based on the Statistical Classification of Economic Activities (CNAE93).
  - The manufacturing sector and service sector are each split into six industries.
  - Other sectors are dropped.

Descriptive statistics for the main variables are shown for the period 2005-2015 in table 1, split by whether firm belongs to the manufacturing or the service sector as well as jointly. The last row show a weak positive correlation between the sales of new products in year $t+1$ and log R&D expenses at region-level.

Table 1: Descriptive statistics

	0 Manufacturing	1 Services	Total
Sales of new products in year t+1	16.81 (28.54)	13.58 (26.56)	15.70 (27.92)
log R&D expenses at region-level	0.698 (0.321)	0.659 (0.349)	0.685 (0.332)
log internal R&D expenses	0.925 (0.871)	1.615 (1.594)	1.162 (1.217)
log external R&D expenses	0.222 (0.480)	0.376 (0.813)	0.275 (0.620)
Cont. of basic research	0.426 (0.588)	0.425 (0.606)	0.426 (0.594)
Cont. of applied research	1.577 (0.903)	1.381 (0.924)	1.509 (0.915)
Cont. of tehcnological development	1.729 (0.862)	1.674 (0.865)	1.710 (0.864)
Cont. of external R&D only	0.154 (0.386)	0.134 (0.355)	0.147 (0.376)
Number of firms	1,949	1,026	2,975
Total observations	21,439	11,286	32,725
Corr. w. log R&D expenses at region-level	0.0216	0.0466	0.0336

Mean coefficients; sd in parentheses.

5. Results

The baseline model for the longitudinal panel is estimated using Random Effects estimation for the subsamples of manufacturing and services respectively. The estimated coefficients for the main variables are shown in table 2 and for the regional dummies in table 3.

No spillover-effects are found from the region-level R&D expenses. Altering the identifications strategy does not provide significant estimates either as shown in appendix A, B, and C. The effect is however found to be positive for the service sector in table 12 and

13 but all of the estimates are insignificant as the sample size decreases largely by only including firms for which the region is identified uniquely for each year.

At the firm-level internal R&D expenses are highly significant. Taking the robustness checks in appendix A into account the effect of doubling the share of internal R&D expenses is about a 2 percentage point increase in the share of new products in total sales. Continuity of applied research is highly significant for manufacturing firms. Occasional basic research and frequent technological development has a significant effect within services. However there is a causality issue as the continuity variable is calculated for all of the time period.

For a few regions there are significant negative regional effects relative to Andalusia.

Table 2: Estimated sales of new products as share of total sales

	(1) manufact b/se	(2) services b/se
log R&D expenses at region-level	-1.200	-0.253
	(1.227)	(1.526)
log internal R&D expenses	1.597***	1.974***
	(0.375)	(0.316)
log external R&D expenses	0.376	0.366
	(0.544)	(0.416)
Continuity of basic research: Never	0.000	0.000
	(.)	(.)
Occasional	1.112	2.062**
	(0.743)	(0.938)
Frequent	0.598	0.071
	(1.947)	(2.183)
Permanent	-2.266	7.740
	(4.659)	(7.170)
Continuity of applied research: Never	0.000	0.000
	(.)	(.)
Occasional	3.409***	1.125
	(1.102)	(1.276)
Frequent	4.991***	1.035
	(1.204)	(1.441)
Permanent	4.023***	0.211
	(1.325)	(1.801)
Continuity of tech. development: Never	0.000	0.000
	(.)	(.)
Occasional	1.242	2.314
	(1.596)	(1.715)
Frequent	2.116	4.301**
	(1.581)	(1.825)
Permanent	0.910	2.513
	(1.669)	(2.011)
Continuity of external R&D only: Never	0.000	0.000
	(.)	(.)
Occasional	-0.812	-0.633
	(1.046)	(1.200)
Frequent	-2.075	0.854
	(2.538)	(6.525)
Region dummies	Yes	Yes
Industry dummies	Yes	Yes
Year dummies	Yes	Yes
Constant	10.654	13.332
Observations	25,223	13,233

Cluster robust standard errors are in parentheses. * p<0.10, ** p<0.05, *** p<0.01

Table 3: Estimated sales of new products as share of total sales

	(1) manufact b/se	(2) services b/se
Andalucía	0.000	0.000
	(.)	(.)
Aragón	0.262	-2.121
	(2.194)	(3.388)
Asturias (Principado de)	-8.556***	-2.283
	(2.365)	(3.057)
Balears (Illes)	-13.213***	0.342
	(3.589)	(6.616)
Canarias	-0.926	-3.913
	(3.030)	(2.898)
Cantabria	-4.591	-8.205***
	(2.962)	(2.771)
Castilla y León	-0.350	-3.855
	(2.372)	(2.489)
Castilla-La Mancha	-5.547**	-7.632***
	(2.462)	(2.599)
Cataluña	-3.076*	-2.072
	(1.626)	(1.775)
Comunidad Valenciana	-1.847	-1.622
	(1.742)	(2.033)
Extremadura	1.088	-5.141
	(5.554)	(3.551)
Galicia	-4.960**	-3.046
	(2.033)	(2.261)
Madrid (Comunidad de)	-3.366*	-2.249
	(1.904)	(1.717)
Murcia (Región de)	-0.936	-3.336
	(2.870)	(3.169)
Navarra (Com. Foral de)	-3.870*	-3.392
	(2.034)	(3.625)
País Vasco	-2.332	1.043
	(1.876)	(2.263)
Rioja (La)	-3.518	-5.040
	(2.637)	(5.264)
R&D variables	Yes	Yes
Continuity dummies	Yes	Yes
Industry dummies	Yes	Yes
Year dummies	Yes	Yes
Constant	10.654	13.332
Observations	25,223	13,233

Cluster robust standard errors are in parentheses. * p<0.10, ** p<0.05, *** p<0.01

6. Conclusion

The internal R&D expenses strongly dominate any regional effects. For a few regions there is a significant regional effect on shares of newly developed products in total sales is extraordinarily small. In the long term continuity of applied research is a strong determinant for having a high share of new products in the portfolio for manufacturing firms and continuity of basic research and technological development for services.

References

Barrios Cobos, Salvador, Holger Görg, Eric Strobl, et al. (2001). *Explaining firms export behaviour: the role of RD and spillovers*. Tech. rep. Research paper/Leverhulme Centre for Research on Globalisation and Economic ...

Harrison, Rupert et al. (2014). "Does innovation stimulate employment? A firm-level analysis using comparable micro-data from four European countries." In: *International Journal of Industrial Organization* 35, pp. 29–43.

Sternberg, Rolf and Olaf Arndt (2001). "The firm or the region: what determines the innovation behavior of European firms?" In: *Economic Geography* 77.4, pp. 364–382.

Vogel, Johanna (2015). "The two faces of R & D and human capital: Evidence from W estern E uropean regions." In: *Papers in Regional Science* 94.3, pp. 525–551.

A. Estimates with different truncation points

Same specification as for the baseline model in table 2 but truncated at different thresholds $C \in 50, 100, 200, 400, \infty$ of the percentage of firm level R&D expenses relative to total sales.

Table 4: Estimated sales of new products as share of total sales, $C = 50$

	(1) manufact b/se	(2) services b/se
log R&D expenses at region-level	-2.377**	-0.899
	(1.119)	(1.634)
log internal R&D expenses	1.874***	1.720***
	(0.399)	(0.439)
log external R&D expenses	0.951	0.737
	(0.681)	(0.776)
Dummies for continuity, region, industry & year	Yes	Yes
Constant	Yes	Yes
Observations	24,134	9,471

Cluster robust standard errors are in parentheses. * p<0.10, ** p<0.05, *** p<0.01

Table 5: Estimated sales of new products as share of total sales, $C = 100$

	(1) manufact b/se	(2) services b/se
log R&D expenses at region-level	-1.229	-0.007
	(1.098)	(1.500)
log internal R&D expenses	1.780***	2.038***
	(0.385)	(0.375)
log external R&D expenses	0.604	0.447
	(0.617)	(0.544)
Dummies for continuity, region, industry & year	Yes	Yes
Constant	Yes	Yes
Observations	24,827	11,011

Cluster robust standard errors are in parentheses. * p<0.10, ** p<0.05, *** p<0.01

Table 6: Estimated sales of new products as share of total sales, $C = 200$

	(1) manufact b/se	(2) services b/se
log R&D expenses at region-level	-1.449	-0.573
	(1.181)	(1.588)
log internal R&D expenses	1.599***	2.096***
	(0.376)	(0.329)
log external R&D expenses	0.555	0.348
	(0.568)	(0.472)
Dummies for continuity, region, industry & year	Yes	Yes
Constant	Yes	Yes
Observations	25,091	12,540

Cluster robust standard errors are in parentheses. * p<0.10, ** p<0.05, *** p<0.01

A. Estimates with different truncation points

Table 7: Estimated sales of new products as share of total sales, $C = 400$

	(1) manufact b/se	(2) services b/se
log R&D expenses at region-level	-1.200	-0.253
	(1.227)	(1.526)
log internal R&D expenses	1.597***	1.974***
	(0.375)	(0.316)
log external R&D expenses	0.376	0.366
	(0.544)	(0.416)
Dummies for continuity, region, industry & year	Yes	Yes
Constant	Yes	Yes
Observations	25,223	13,233

Cluster robust standard errors are in parentheses. * p<0.10, ** p<0.05, *** p<0.01

Table 8: Estimated sales of new products as share of total sales, no C

	(1) manufact b/se	(2) services b/se
log R&D expenses at region-level	-0.941	-0.388
	(1.211)	(1.473)
log internal R&D expenses	1.376***	1.440***
	(0.363)	(0.290)
log external R&D expenses	0.306	0.167
	(0.506)	(0.345)
Dummies for continuity, region, industry & year	Yes	Yes
Constant	Yes	Yes
Observations	25,388	13,970

Cluster robust standard errors are in parentheses. * p<0.10, ** p<0.05, *** p<0.01

B. Higher criteria for the share of R&D personnel

Same model specification as for the baseline model in table 2 but raising the criteria for regional identification to regions having a share of the firm's R&D personnel $\geq S \in 75, 90, 100$ percent.

Table 9: Estimated sales of new products as share of total sales, $S \geq 75$

	(1) manufact b/se	(2) services b/se
log R&D expenses at region-level	-0.944	-0.605
	(1.157)	(1.375)
log internal R&D expenses	1.706***	2.107***
	(0.409)	(0.340)
log external R&D expenses	0.079	0.241
	(0.601)	(0.446)
Dummies for continuity, region, industry & year	Yes	Yes
Constant	Yes	Yes
Observations	21,857	11,506

Cluster robust standard errors are in parentheses. * p<0.10, ** p<0.05, *** p<0.01

Table 10: Estimated sales of new products as share of total sales, $S \geq 90$

	(1) manufact b/se	(2) services b/se
log R&D expenses at region-level	-1.296	-1.529
	(1.280)	(1.543)
log internal R&D expenses	2.005***	2.100***
	(0.439)	(0.358)
log external R&D expenses	0.150	0.008
	(0.665)	(0.470)
Dummies for continuity, region, industry & year	Yes	Yes
Constant	Yes	Yes
Observations	19,338	10,186

Cluster robust standard errors are in parentheses. * p<0.10, ** p<0.05, *** p<0.01

Table 11: Estimated sales of new products as share of total sales, $S = 100$

	(1) manufact b/se	(2) services b/se
log R&D expenses at region-level	-1.449	-1.870
	(1.284)	(1.543)
log internal R&D expenses	1.982***	2.089***
	(0.448)	(0.378)
log external R&D expenses	0.211	-0.048
	(0.698)	(0.499)
Dummies for continuity, region, industry & year	Yes	Yes
Constant	Yes	Yes
Observations	18,139	9,284

Cluster robust standard errors are in parentheses. * p<0.10, ** p<0.05, *** p<0.01

C. No imputation when region is missing

Same model specification as for the baseline model in table 2 but without imputing region from other years when it is missing. Run for different criteria for regional identification where regions have a share of the firm's R&D personnel $\geq S \in 50, 75, 100$ percent.

Table 12: Estimated sales of new products as share of total sales, $S \geq 50$

	(1) manufact b/se	(2) services b/se
log R&D expenses at region-level	-2.026	1.407
	(1.914)	(2.099)
log internal R&D expenses	0.854	1.211
	(0.840)	(1.162)
log external R&D expenses	-0.200	-0.255
	(1.085)	(0.965)
Dummies for continuity, region, industry & year	Yes	Yes
Constant	Yes	Yes
Observations	5,951	2,277

Cluster robust standard errors are in parentheses. * p<0.10, ** p<0.05, *** p<0.01

Table 13: Estimated sales of new products as share of total sales, $S \geq 75$

	(1) manufact b/se	(2) services b/se
log R&D expenses at region-level	-0.728	0.658
	(2.084)	(2.089)
log internal R&D expenses	0.956	2.920
	(1.274)	(1.835)
log external R&D expenses	1.156	1.155
	(1.870)	(1.830)
Dummies for continuity, region, industry & year	Yes	Yes
Constant	Yes	Yes
Observations	3,047	979

Cluster robust standard errors are in parentheses. * p<0.10, ** p<0.05, *** p<0.01

Table 14: Estimated sales of new products as share of total sales, $S = 100$

	(1) manufact b/se	(2) services b/se
log R&D expenses at region-level	4.504	-2.590
	(4.091)	(7.236)
log internal R&D expenses	3.939	6.517*
	(2.422)	(3.597)
log external R&D expenses	-0.381	-1.874
	(3.538)	(2.782)
Dummies for continuity, region, industry & year	Yes	Yes
Constant	Yes	Yes
Observations	1,155	352

Cluster robust standard errors are in parentheses. * p<0.10, ** p<0.05, *** p<0.01