

# The Impact of Exports on Innovation: Theory and Evidence\*

Philippe Aghion      Antonin Bergeaud

Matthieu Lequien      Marc Melitz

March 2017

## Abstract

A simple model of trade and innovation with heterogeneous firms predicts that a positive export shock should raise innovation more for more productive firms. Two channels coexist: the innovation effort increases for all firms because the accompanying rents increase with a firm's market size (*market size* effect); the innovation effort decreases because competition toughens. This *competition* effect dissipates with higher firm productivity. It is therefore most salient for the least productive firms and can potentially overturn the direct *market size* effect.

We test this prediction with patent, customs and production data covering all French firms. To disentangle the direction of causality between innovation and export performance, we construct various firm-level export demand measures. These variables capture the extent to which demand fluctuations in a firm's foreign markets should influence its exports and through them weigh on its innovation decisions; but they remain exogenous to those firm-level decisions. We show that patenting robustly increases more with demand for initially more productive firms. This effect is reversed for the least productive firms as the negative competition effect dominates.

---

\*Addresses - Aghion: Collège de France, LSE and Banque de France; Bergeaud: Banque de France and Paris School of Economics - EHESS; Lequien: Banque de France; Melitz: Harvard University and NBER. The authors would like to thank, without implicating Ufuk Akcigit, Giammario Impullitti, Dan Treffer, John Van Reenen and participants at the 2017 AEA meeting for useful discussions and advises. This work is supported by a public grant overseen by the French National Research Agency (ANR) as part of the "Investissements d'Avenir" program (reference: ANR-10-EQPX-17 - Centre d'accès securise aux donnees CASD)

# 1 Introduction

Does firms' access to export affect innovation? Modern trade and growth theories (Grossman and Helpman, 1991b; Acemoglu, 2009; Aghion and Howitt, 2009) suggest it should, if only because improved access to export markets should increase the *size of markets* that can be appropriated by successful innovators. Moreover, we know that trade induces knowledge spillovers. The *knowledge spillover* effect of trade underlies the work of Coe and Helpman (1995) among others.<sup>1</sup> However, on the empirical side, the question of how trade and more specifically *exports* should affect firms' innovation performance, has received little attention until recently (see our literature review below).

In this paper we use exhaustive firm-level data covering all French exporting firms to analyze how new export opportunities affects exporting firms' innovation performance. One of the most striking features that emerges from this data is a massive correlation between export and innovation performance across firms. This holds both at the extensive margin (exporters are substantially more likely to innovate, and innovators are more likely to export) as well as the intensive margin (large exporters tend to be big innovators and vice-versa). We describe these relationships in much more detail in Section 3. Does this correlation reflect a causal effect of export on innovation, or the effect of innovation on exports, or both? How does the innovation behavior of a firm react to its export markets' conditions? Is there a spillover effect of trade on innovation whereby, when a firm starts exporting into a country, the scope for subsequent innovations by other firms operating in that country increases? Our paper is a first attempt at understanding these firm-level patterns connecting innovation and trade using the matching between patenting, balance sheet, and customs exhaustive datasets.

In the first part of the paper we develop a simple model of trade and innovation with heterogeneous firms. The model builds on Mayer et al. (2014) but adds the innovation dimension to it. It features a continuous set of firms indexed by their heterogeneous production costs. Innovation allows firms to reduce their production costs by an amount that increases with the size of the innovation investment. Think of French firms which export to China. An increase in Chinese demand for French firms' products, will have two main effects on firms' innovation incentives. First, a direct market size effect: namely, the expanded market for

---

<sup>1</sup>Coe and Helpman (1995) construct, for each country, measures of domestic and foreign R&D capital stocks, where the latter are weighted averages of the domestic stocks of trade partners. They find that foreign R&D appears to have a beneficial effect on domestic productivity, and that the effect increases in strength with the degree of openness. Hence, not only are there important spillovers, but there is also some evidence that these are mediated by trade. However, one may argue that even if a correlation is observed between domestic productivity and foreign research, this may simply represent the outcome of common demand or input price shocks. Weighting the contribution of foreign research using data on bilateral trade flows, as in Coe and Helpman (1995), is likely to mitigate this problem but will not overcome it altogether.

exports will increase the size of innovation rents and thereby increase French firms' incentives to invest more in innovation. Second, a competition effect: namely, the expanded market for exports may attract new firms into the Chinese market and more generally it will raise competition between exporters into that market. This competition effect dissipates with higher firm productivity. It is therefore most salient for French firms with initially higher production costs (these firms will suffer more than -or at the expense of- more efficient exporting firms). Hence the prediction that a positive export shock should raise innovation more in more frontier firms; and that it may induce less innovation for those firms that are far from the frontier.

In the second part of the paper we take this prediction to the data. More specifically, we merge three exhaustive firm-level datasets - patenting data, production data, and customs data-, which cover the whole population of French firms to analyze how the access to export markets affects the stock and quality of patents by these firms.

The patent data are drawn from PATSTAT (Spring 2015 version) and contain information on all granted patents, including the country of residence of the applicant and a citation network between these patents. An algorithm matches a French firm's name with its unique administrative identifier, which allows us to link the innovation activities of a firm with all other firm data sources. The production datasets FICUS and FARE, from INSEE/DGFiP, contain balance sheet information for each firm registered in France from 1993 to 2012 (total and export sales, number of employees, sector, etc.). French customs trade data (1993-2014) cover nearly comprehensive export flows by firm and destination at a very detailed level of product disaggregation (over 10,000 product categories). We complement these firm-level data sets with bilateral trade data from BACI ([Gaulier and Zignago, 2010](#), updated to cover the period 1995-2013) at the product level (at a slightly higher level of aggregation than our French firm-level export data); and with country-level data (primarily GDP).

To disentangle the direction of causality between innovation and export performance, we construct a firm-level export demand variable following [Mayer et al. \(2016\)](#). This variable responds to aggregate conditions in a firm's export destinations but is exogenous to firm-level decisions (including the concurrent decisions for export-market participation and the forward looking innovation response). We show that: (i) Firms that are initially more productive (closer to their sector's technology "frontier") strongly respond to a positive export demand shock by patenting more; (ii) this effect dissipates for firms further from the "frontier" and is reversed for a subset of initially less productive firms. These results confirm the predictions of the model for both, a *market size* and a *competition effect* of the export shock.

Our analysis relates to several strands of literature. There is first the theoretical liter-

ature on trade, innovation and growth (e.g. see [Grossman and Helpman, 1991a,b<sup>2</sup>](#)). We contribute to this literature by uncovering a new -indirect- effect of market size on innovation working through competition, and by testing the market size effects of export expansion on innovation using exhaustive firm-level data. Second, our paper relates to recent papers on import competition, innovation and productivity growth (e.g. see [Bustos, 2011](#); [Bloom et al., 2016](#); [Iacovone et al., 2011](#); [Caldwell and Tabellini, 2015](#); [Autor et al., 2016](#)). These papers show that increased import competition induces firms to innovate more in order to escape competition as in [Aghion et al. \(2005\)](#). Instead we look at how the export side of trade affects innovation.

Most closely related to our analysis in this paper are [Clerides et al. \(1998\)](#), [Bernard and Jensen \(1999\)](#) and [Lileeva and Trefler \(2010\)](#), which look at the effects of exports on productivity. In particular, [Lileeva and Trefler \(2010\)](#) provide evidence of a causal effect on export on productivity and innovation by using the US tariff cut imposed in 1989 by the new Free Trade Agreement (FTA) between US and Canada, as an instrument for export expansion. Their main conclusion is that the FTA induced productivity gains by Canadian firms that saw their access to the US market improved by the FTA. Moreover, focusing on a small subsample of 521 firms for which they have survey information on innovative investment, the authors show that firms in that sample which experience higher productivity growth also invested more in technology adoption and product innovation. We add to their analysis in three main respects: first, by uncovering an indirect- competition-enhancing effect of increased export markets; second, by showing that this effect leads to the market size effect of a positive export shock, being stronger for more frontier firms; third, by using patenting data to measure firms' innovation performance and by merging these data with exhaustive administrative and customs data covering the whole population of French firms.

The remaining part of the paper is organized as follows. Section 2 develops our model of export and innovation, and generates the prediction that the market size effect of a positive export shock, is stronger for more frontier firms. Section 3 briefly presents the data and show some descriptive statistics on export and innovation. Section 4 describes our estimation methodology and present our empirical results and Section 5 concludes.

## 2 Theory

The model in this section is essentially a closed economy short-run version of the model in [Mayer et al. \(2014\)](#), augmented with innovation. We consider (all) French exporting firms

---

<sup>2</sup>See also [Akcigit et al., 2014](#).

that are selling in some export market destination  $D$ . We let  $L$  denote the number of consumers  $L$  in that destination – with income normalized to 1. Suppose that the representative consumer in country  $D$  has utility for good  $i$  which is quadratic<sup>3</sup> and equal to:

$$u(q_i) = \alpha q_i - \frac{\beta q_i^2}{2}.$$

where  $\alpha > 0$  and  $\beta > 0$ .

## 2.1 Consumer optimization

The representative consumer solves:

$$\max_{q_i \geq 0} \int_0^M u(q_i) di \quad \text{s.t.} \quad \int_0^M p_i q_i di = 1,$$

which yields by first order condition the *inverse* residual demand function (per consumer):

$$p(q_i) = \frac{u'(q_i)}{\lambda} = \frac{\alpha - \beta q_i}{\lambda}, \quad (1)$$

where  $\lambda = \int_0^M u'(q_i) q_i di > 0$  is the corresponding Lagrange multiplier, also equal to the marginal utility of income.

## 2.2 Firm optimization

Consider a firm with marginal cost  $c$  facing demand conditions  $\lambda$ . This firm chooses the optimal output per consumer  $q(c, \lambda)$  to maximize operating profits  $L [p(q)q - cq]$ . The corresponding first order condition yields

$$q(c, \lambda) = \frac{\alpha - c\lambda}{2\beta}.$$

This in turn leads us to the following expressions for equilibrium revenues and profits:

$$r(c, \lambda) = \frac{\alpha^2 - (c\lambda)^2}{4\beta\lambda},$$

---

<sup>3</sup>The analysis can be extended to a broader class of utility functions, in particular to those that satisfy Marshall's Second Law of Demand, whereby firms' inverse residual demand becomes more inelastic as consumption  $q$  increases.

and

$$\pi(c, \lambda) = \frac{(\alpha - c\lambda)^2}{4\beta\lambda}.$$

In particular we see that  $\pi(c, \lambda)$  is decreasing in  $c$  and  $\lambda$ . Below we shall interpret  $\lambda$  as a measure of the intensity of product market competition.

### 2.3 Innovation choice

A firm is characterized by its baseline cost  $\tilde{c}$ . The firm can reduce its actual marginal cost of production  $c$  below its baseline cost by investing in innovation. More formally, we assume that:

$$c = \tilde{c} - \varepsilon k,$$

where  $k$  is the firm's R&D investment and  $\varepsilon > 0$ . Without loss of generality, we assume that the cost of innovation is quadratic in  $k$ , equal to  $\frac{1}{2}c_I k^2$ .<sup>4</sup>

Thus a firm with baseline cost  $\tilde{c}$  will choose its optimal R&D investment  $k(\tilde{c}, \lambda)$  so as to maximize:

$$L\pi(\tilde{c} - \varepsilon k, \lambda) - \frac{1}{2}c_I k^2.$$

By virtue of the envelope theorem,  $k(\tilde{c}, \lambda)$  satisfies the first order condition:

$$L\varepsilon q(\tilde{c} - \varepsilon k, \lambda) = c_I k,$$

or<sup>5</sup>

$$\frac{\varepsilon L}{2\beta} = \frac{c_I k}{\alpha - (\tilde{c} - \varepsilon k)\lambda}. \quad (2)$$

This yields the equilibrium value of R&D investment  $k^*$ :

$$k^* = \frac{\alpha - \tilde{c}\lambda}{\frac{2\beta c_I}{\varepsilon L} - \varepsilon L}$$

As shown in Figure 1, we see that this equilibrium value of R&D investment is defined by the intersection between a horizontal line and an upward-sloping curve, which correspond respectively to the LHS and the RHS of (2).

Figure 2 shows the comparative statics with respect to  $L$ . Namely, an increase in  $L$  shifts

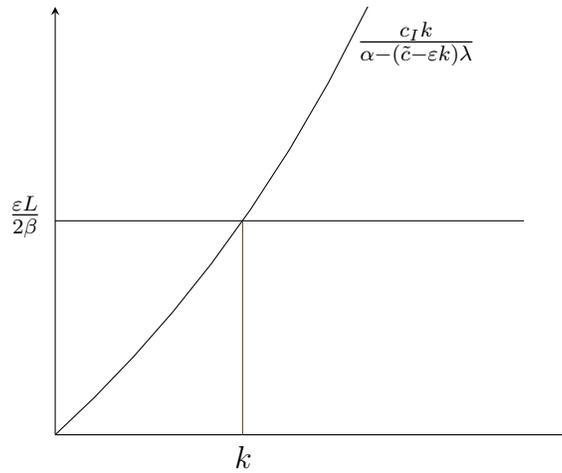
---

<sup>4</sup>Since we only consider a single sale destination  $D$  for our firms, we are implicitly assuming that the innovation is directed at the delivered cost to consumers in  $D$ . We should thus think of innovation as specific to the appeal/cost trade-off to consumers in  $D$ .

<sup>5</sup>We restrict attention to parameter values such that the equilibrium  $k$  is such that

$$\tilde{c} > \varepsilon k.$$

Figure 1: Equilibrium of R&D



the LHS curve upward, and therefore the equilibrium R&D investment  $k^*$  must increase for given  $\tilde{c}$  and  $\lambda$ . This corresponds to the *direct market size effect* of an increase in  $L$  on innovation by all firms.

Figure 2: Direct Market size Effect

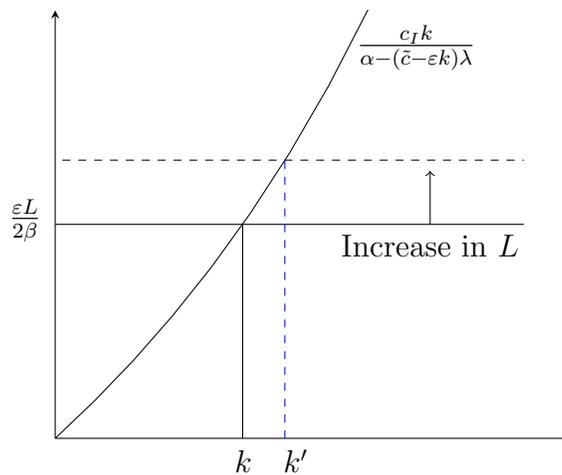
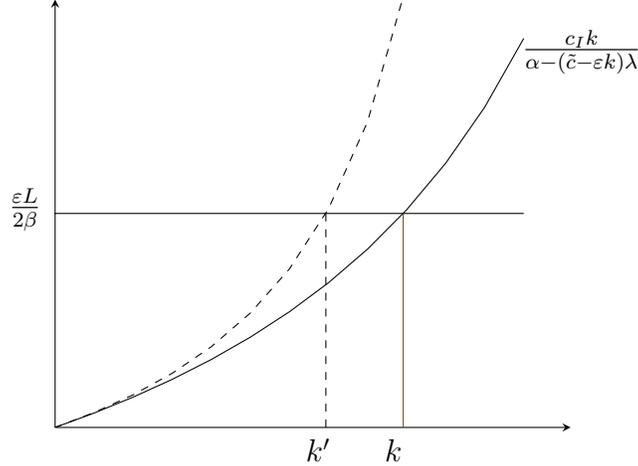


Figure 3 shows the comparative statics with respect to  $\lambda$ . Namely, an increase in  $\lambda$  shifts the RHS curve up upward, thereby reducing the equilibrium R&D investment  $k^*$ . However,

Figure 3: Competition Effect



the higher  $\tilde{c}$  the bigger the upward shift in the RHS curve for given  $k$  and therefore the bigger the reduction in the equilibrium R&D investment  $k^*$ . In other words, *an increase in competition will reduce innovation incentives more for less frontier firms.*

## 2.4 Free entry equilibrium and the indirect effect of market size

We shall focus attention on the free-entry equilibrium where the marginal firm is indifferent between paying a fixed ex post operating cost  $F$  and not entering the export market. Since operating profit is monotonic in a product's baseline cost  $\tilde{c}$ , the cutoff cost  $\hat{c}$  of the marginal firm will satisfy:

$$L\pi(c(\hat{c}, \lambda), \lambda) = F, \quad (3)$$

where the LHS corresponds to the aggregate ex post profit the firm makes by entering the export market. Thus, all French firms with baseline cost  $\tilde{c} \leq \hat{c}$  will find it profitable to export to country  $D$ .

We close the model by introducing an aggregate budget constraint. This budget constraint states that the aggregate spending by country  $D$ 's consumers on all exported products from the  $N$  exporting countries to  $D$ , must be equal to the aggregate revenue of country  $D$ 's consumers (which we normalized at 1). Then, given the distribution of baseline costs  $\Gamma(\tilde{c})$ , the consumers' budget constraint can be expressed as:

$$N \left[ \int_0^{\hat{c}} r(c(\tilde{c}, \lambda), \lambda) d\Gamma(\tilde{c}) \right] = 1. \quad (4)$$

Together, those two conditions (cutoff profit and budget constraint) jointly determine the toughness of competition  $\lambda$  in country  $D$  and the baseline cost cutoff  $\hat{c}$  as functions of market size  $L$ .

## 2.5 The direct and indirect effects of increased market size

Consider the overall effect of an increase in market size  $L$ . From equation (3), either  $\hat{c}$  or  $\lambda$  or both variables must increase. Now suppose for a moment that  $\hat{c}$  increases but  $\lambda$  decreases. Then the LHS of (4) should increase since: (i) the integration interval  $[0, \hat{c}]$  expands; (ii) for any  $\tilde{c}$  in that support,  $c(\tilde{c}, \lambda) = C$  must go down as  $k(\tilde{c}, \lambda)$  would increase; therefore

$$r(c(\tilde{c}, \lambda), \lambda) = \frac{\alpha^2 - (C\lambda)^2}{4\beta\lambda},$$

would necessarily increase. But this implies that if (4) held initially, it would cease to hold after the increase in  $L$ . This reasoning implies that an increase in  $L$  must necessarily lead to an increase in  $\lambda$ .

The intuition for this induced indirect -competition driven- effect of increasing export market size, can be summarized as follows: an increase in export market size  $L$  leads to an increase in the mass of products  $N\hat{c}$  exported to country  $D$  (free entry condition 3); but then each individual exporting firm to country  $D$  will face a more elastic curve as it faces more competition from other exporting firms to  $D$ , which in turn corresponds to an increase in  $\lambda$ .<sup>6</sup>

The indirect effect of market size will counteract the positive direct effect of market size on innovation, but less so for more frontier firms, i.e. for firms with lower initial  $\tilde{c}$ . This reasoning extends to more general utility functions,<sup>7</sup> and yields the prediction:

*Prediction:* An increase in market size will increase innovation more in more frontier firms.

---

<sup>6</sup>Another interpretation is that, as the number of products exported to country  $D$  increases, the representative consumer from country  $D$  gets a higher increase in utility for every extra dollar of income as she can increase consumption across a higher variety of products. This in turn implies that the marginal utility of income  $\lambda$  must increase.

<sup>7</sup>The marginal benefit of innovation  $\partial\pi/\partial k$  increases proportionately more for low cost firms. This is a direct consequence of the impact of higher competition  $\lambda$  on the relative output of firms: For any 2 firms with marginal costs  $c_1 < c_2$  the output ratio  $q(c_1, \lambda)/q(c_2, \lambda)$  increases with competition  $\lambda$  (proof in Mayer et al., 2016). It is straightforward to extend this result to the comparison of 2 firms with different baseline costs  $\tilde{c}_1 < \tilde{c}_2$ .

### **3 Exporters and Innovators: data and descriptive statistics**

In this section, we briefly present our datasets and show some descriptive evidence on the link between firms' innovation and exports. Further details about data construction can be found in Appendix A.

#### **3.1 Data sources**

We build a database covering all French firms and linking export, fiscal and innovation data from 1993 to 2012. Our database draws from three sources: (i) French customs, which reports yearly export flows at a very disaggregated HS8 product level (representing over 10,000 manufacturing products) by destination; (ii) Insee-DGFIP administrative fiscal datasets (**FICUS** and **FARE**), which provide extensive production and financial information for all firms operating in France; (iii) the 2015 **PATSTAT** patent dataset from the European Patent Office, which contains detailed information on all patent applications (including its citations of prior patents) from every patent office in the world by year 2015. In our analysis we will focus on all patent applications (whether ultimately granted or not) as well as the patents' citations by future patent applications (see Appendix A for details).

Although each French firm has a unique identifying number (Siren) across all French databases, patent offices do not identify firms applying for patents using this number but instead using the firm's name. We use the rigorous matching algorithm developed in [Lequien et al. \(in progress\)](#) to link each patent application (and its citations) back to the French firms' Siren numbers for all firms with more than 10 employees.

Finally, we use two additional datasets to construct measures of demand shocks across export destinations for French firms. CEPII's **BACI** database reports bilateral trade flows at the HS6 product level (covering more than 5,000 manufacturing products). IMF's **World Economic Outlook** provides country information such as GDP or GDP per capita.

#### **Sample restrictions**

Although our main firm-level administrative data source is comprehensive, with more than 47.1 million observations spanning over 7.3 million different firms from 1995 to 2012, we restrict our data sample for several reasons. The first is due to the matching with patent data mentioned above, which is most complete for firms above 10 employees. We therefore impose this size restriction, which drops a large number of firms but a relatively small

share of aggregate French production: 17.1% of employment, 15.6% of sales, and 13.6% of exports (predominantly within EU exports). Second, we restrict our attention to private firms (legal category 5 in the Insee classification). We thus drop state-owned firms, self-employed businesses, non-profit organizations, as we focus on profit-maximizing firms. This further reduces our sample from 1.7 million to 835,000 firms. Yet, the bulk of aggregate employment (74.2%), sales (77.7%), and exports (77.2%) remain in our dataset after imposing these restrictions. These remaining firms are matched with an average of 27,640 patents per year in PATSTAT. Lastly, since our detailed customs trade data only covers goods trade (and not services), we will further restrict our sample to the manufacturing sector for most of our analysis.<sup>8</sup> This reduces our working sample to 105,000 firms. Nevertheless the bulk of French aggregate exports and innovation are still concentrated in manufacturing: only 20.6% of aggregate exports and 33% of patents are recorded outside of the manufacturing sector.

### 3.2 Sector breakdown and skewness

Table 1 shows the breakdown of those firms across sectors, along with their average employment, exports, and patents (per firm) for 2007.<sup>9</sup> As has been widely reported in the empirical literature on micro-level trade patterns, many firms are only occasional exporters: they export in some years, but not in others. This pattern is even more pronounced for innovation: even firms with substantial ongoing R&D operations do not typically file patent applications year in and year out. We therefore use the broadest possible cross-year definition to classify firms as exporters and innovators. We label a firm as an *exporter* if it has exported at least once between 1993-2012; and as an *innovator* if it has filed at least one patent application between 1995-2012.<sup>10</sup> Thus, our reported export participation rates in Table 1 are higher than in other studies. However, even with this broadest classification, innovators represent only a small minority of manufacturing firms. For comparison, Table 1 also reports the share of exporters and innovators based on the more standard definition of current year (2007 for this table) exporting or patenting activity – shown in parentheses.

Even within the minority set of innovators, patenting activity is extremely skewed. This is clearly visible in Figure 4, which plots the Lorenz curve for manufacturing patents in

---

<sup>8</sup>Although the customs data also covers the wholesale sector, we also exclude those firms as they do not produce the goods that they export.

<sup>9</sup>Throughout, we define sectors at the 2-digit level of the NAF rev2 classification. We also eliminate the tobacco sector (# 12) as it only contains two firms.

<sup>10</sup>The initial year for both ranges do not coincide so as to best reflect our subsequent empirical analyses. We will use prior years of export data to construct exogenous export share weights (see section 4.1 for more details).

Table 1: EXPORTS AND INNOVATION IN THE MANUFACTURING SECTOR

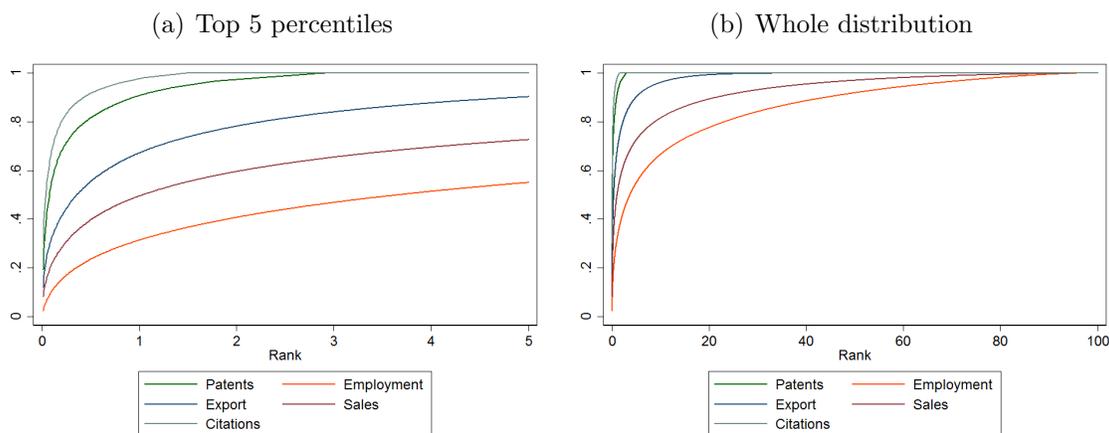
Sector	Description	Firms	Emp	Export	% Exporter	Patents	% Innov.
10	Food products	8,814	43	1,847	41 (26)	0.025	3 (0)
11	Beverages	1,463	47	5,974	80 (59)	0.012	2 (0)
13	Textiles	1,802	37	2,335	86 (63)	0.106	12 (2)
14	Wearing apparel	1,558	39	2,577	80 (59)	0.065	5 (1)
15	Leather	492	56	2,566	85 (59)	0.090	9 (2)
16	Wood	2,432	29	790	63 (36)	0.017	5 (1)
17	Paper	2,950	44	2,056	79 (44)	0.086	7 (1)
18	Printing	842	24	167	52 (20)	0.006	3 (0)
19	Coke	171	225	75,957	92 (69)	3.883	22 (9)
20	Chemicals	1,229	116	17,607	94 (79)	2.347	21 (7)
21	Basic pharmaceutical	357	288	42,065	96 (82)	3.282	35 (13)
22	Rubber and plastic	2,745	80	3,820	86 (64)	0.398	21 (6)
23	Other non-metallic	2,158	63	2,320	65 (38)	0.363	11 (2)
24	Basic metals	1,648	80	12,487	65 (44)	0.176	11 (3)
25	Fabricated metal	8,392	36	1,125	67 (40)	0.091	9 (2)
26	Computer and electronic	3,511	85	7,602	72 (54)	0.949	23 (8)
27	Electrical equipment	447	106	8,812	91 (70)	2.095	26 (8)
28	Machinery and equipment	4,668	80	8,252	79 (58)	0.584	23 (7)
29	Motor vehicles	791	61	2,549	79 (47)	0.200	15 (3)
30	Other transport equipment	558	215	54,910	83 (56)	2.555	18 (7)
31	Furniture	1,146	34	598	67 (36)	0.031	7 (1)
32	Other manufacturing	1,017	41	2,472	82 (58)	0.392	12 (3)
33	Repair and installation of machinery and equipment	3,430	28	302	54 (23)	0.029	6 (1)
	Aggregate manufacturing	52,621	56	4,634	68 (44)	0.335	11 (3)

**Notes:** This table presents the number of firms, average employment, average export (in thousands of Euros), average number of patents, and the shares of exporters and innovators (cross-year definitions). The shares in parentheses are calculated based on current year export participation or patent filing. Data are for 2007.

2007, along with the Lorenz curves for exports, sales, and employment. Figure 4 confirms the previously reported finding that firm-level exports are significantly more skewed than sales and employment (e.g. see Mayer et al., 2014 and Bernard et al., 2016): The top 1% of exporters account for 67% of aggregate exports in 2007, whereas the top 1% of firms based on total size account for 50% of sales (ranked by sales) and 31% of employment (ranked by employment). But Figure 4 also shows that patenting is even significantly more skewed than exporting: the top 1% of patenting firms account for 91% of patents in 2007 and 98% of citations. Indeed fewer than 2.9% of manufacturing firms have patented in 2007. This fraction is significantly smaller than our previously reported 11% share of innovators in Table 1 measured across our full sample years. Similarly, only 44% of manufacturing firms report any exporting activity for 2007 compared to a 68% share when exporting is measured across our full sample years.

These univariate statistics for patenting and exporting do not capture the massive overlap between these two activities across firms – which we investigate in more detail below.

Figure 4: Lorenz curves - patents are more concentrated than exports, sales and employment



**Notes:** Lorenz curves plot cumulative distribution function for patents, citations, employment, export and sales. Data are for manufacturing firms in 2007.

### 3.3 The innovation-export nexus

Looking across our sample years (1994-2012), Table 2 reports different size-related performance measures (averages per firm) based on their exporter and innovator classification. This table confirms the well-documented size differential in favor of exporters. However, several new salient features regarding innovators pop-out from this table: 1) Innovating firms are massively concentrated among exporters: only 6% of innovators do not report any exporting. 2) Those non-exporting innovators do not look very different than their non-innovating coun-

terparts amongst non-exporters. All the various measures of firm size (employment, sales, value-added) are within ten percent of each other.<sup>11</sup> 3) On the other hand, the size differences between innovators and non-innovators amongst exporters are massive: innovators employ on average 4.6 times more workers and produce 7-8 times more output and value-added than non-innovating exporters. They export almost 10 times more than non-innovators and reach more than triple the number of export destinations. These size differentials are many multiples larger than those separating exporters and non-exporters. In the aggregate, this small subset of innovators accounts for over half of French manufacturing exports.

Table 2: EXPORTERS AND INNOVATORS ARE BIGGER

	Non-exporters		Exporters		Total
	Non-innovator	Innovator	Non-innovator	Innovator	
Firms	45,825	397	52,061	6,878	105,161
Employment	18	21	51	234	59
Sales	2,800	2,497	11,478	71,024	14,310
Value Added	719	899	2,722	16,279	3,359
Age	14	15	20	22	18
Exports	0	0	2,440	23,155	3,622
Countries	0	0	5	18	5

**Notes:** This table presents basic descriptive statistics across four categories of manufacturing firms whether they innovate, export, both or none. Employment is given in full-time equivalent on average over the year and exports, sales and value added are in thousand of euros. Countries is the number of destination countries for exports. Employment, Sales, Value Added, Age, Exports and Countries are taken as an average over the whole period 1995-2012.

We now examine these performance differentials in favor of exporters and innovators in greater detail. We first focus on the exporter premia in the top panel of Table 3. These premia are generated by regressing the performance measure of interest (listed in the rows) on our exporter indicator – with each cell representing a separate regression. Column (1) includes no other controls; Column (2) adds a 2-digit sector fixed effect (see Table 1); and Column (3) also controls for firm employment, in addition to the sector fixed effect. Since we are using a broad cross-year definition for exporter status, we expect these premia to be lower than measures based on current-year exporter status.<sup>12</sup> This is the case for the premia

<sup>11</sup>This is not the case outside of the manufacturing sector. In those other sectors, non-exporting innovators are substantially bigger than their non-exporting and non-innovating counterparts. We conjecture that this is driven by the fact that exporting no longer serves the same performance screening function outside of manufacturing.

<sup>12</sup>Since firms who drop in and out of export markets tend to be substantially smaller than year in year

in column (1) compared to similar number reported by [Bernard et al. \(2016\)](#) for U.S. firms in 2007. Yet, once we control for sectors in column (2), the reported premia are much more similar. In particular, we verify that even within sectors, exporters are substantially larger than non-exporters. And we also find that large differences in productivity and wages in favor of exporters persist after controlling for firm employment (within sectors).

Table 3: EXPORT AND INNOVATION PREMIA

<b>Exporter</b>	(1)	(2)	(3)	Obs.	Firms
log Employment	0.853	0.816	-	937,050	91,583
log Sales	1.614	1.595	0.509	979,413	104,368
log Wage	0.132	0.111	0.123	935,489	91,525
log Value Added Per Worker	0.216	0.194	0.196	537,050	91,563
<b>Innovative Exporters</b>	(1)	(2)	(3)	Obs.	Firms
log Employment	1.046	1.015	-	644,390	58,095
log Sales	1.293	1.278	0.217	655,033	58,803
log Wage	0.126	0.101	0.116	643,402	58,079
log Value Added Per Worker	0.207	0.184	0.190	634,056	57,708
log Export Sales (Current period exporters)	2.043	1.970	0.859	433,456	56,509
Number of destination countries	13	12	7	661,751	58,924

**Notes:** This table presents results from an OLS regression of firm characteristics (rows) on a dummy variable for exporting (upper table) or patenting (lower table) from 1994 to 2012. Column 1 uses no additional covariate, column 2 adds a 2 digit sector fixed effect, column 3 adds a control for the log of employment to column 2. All firm characteristic variables are taken in logs. All results are significant at the 1 percent level. Upper table use all manufacturing firms whereas lower table focuses on exporting manufacturing firms.

In the bottom panel, we focus on the subset of exporters from the top panel, and report the *additional* premia in favor of innovators within this subset. As with the top panel, those premia are calculated by running separate regressions on our innovator indicator. Even within this subset of bigger and better performing firms, innovators stand out: they are substantially bigger, more productive, and pay higher wages. They also export substantially more (and to more destinations) than non-innovative exporters. All these differences persist within sectors and controlling for firm employment.

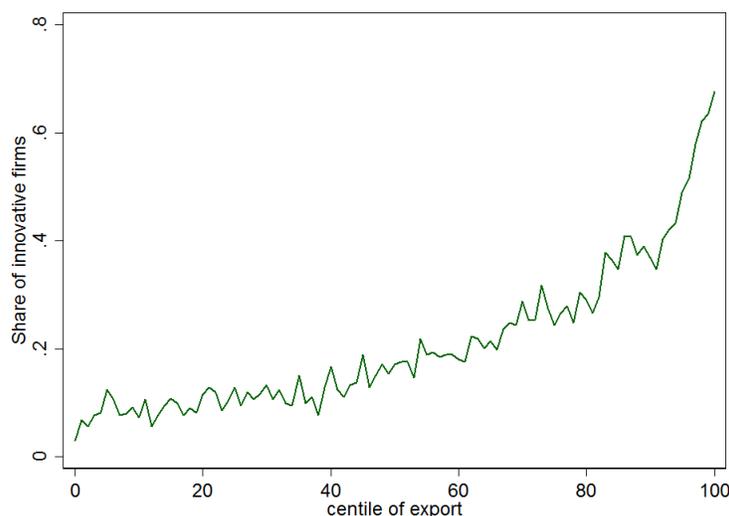
Even these large premia do not characterize the concentration of innovative and exporting activities within an even more restricted subset of exporters and innovators. To capture this concentration at the “intensive” margin of firm-exports, [Figure 5](#) plots the share of innovating

---

out exporters.

firms for each centile of the firm export distribution. We see that the innovative firms are highly concentrated within the top percentiles of the export distribution. At the 80th percentile of the export distribution, 30% of the firms have some patenting experience. And the increase in the share of innovative firms with the percentile of the export distribution is convex. Above the 95th percentile of the export distribution, a majority of firms are innovators; in the top percentile, 68% of the firms are innovators. Those firms in the top export percentile account for 41% of the aggregate share of French patents.

Figure 5: The share of innovators jumps at the top of the export distribution



**Notes:** Centile of exports are computed each year from 1995 to 2012 separately and then pooled together. For each centile, we calculate the share of innovators. Each centile contains the same number of firm, except for centile 0 that contains all the firm with no export.

## 4 Empirical framework and results

### 4.1 Identification strategy: firm-level export shocks

We have just documented the strong correlation between exporting and innovation in the cross-section of French manufacturing firms. This correlation does not shed light on the direction of causation: from innovation to exports (a brilliant innovation leads to growth in export demand and entry into new export markets, or from exports to innovation (as our theoretical model explains). In addition, other firm-level changes could generate concurrent changes in both innovation and exports (for example, a new management team). Thus, to identify the causal relationship from exports to innovation, we need to identify a source of variation in firm exports that is exogenous to changes within the firm (and in particular to

innovation activity at the firm). We follow Mayer et al. (2016) in building such a measure of exogenous export demand (to the firm).

Consider a French exporter  $f$  who exports a product  $s$  to destination  $j$  at an initial date  $t_0$ . Let  $X_{fjst_0}$  denote this export flow, where  $t_0$  is the firm's first observed export year in our sample.<sup>13</sup> The underlying idea is that subsequent changes in destination  $j$ 's imports of product  $s$  from the world (excluding France) will be a good proxy for the change in export demand faced by this firm. Let  $M_{jst}$  denote this trade flow into  $j$  at time  $t > t_0$ . By excluding French exports to this destination, we seek to exclude sources of variation that originate in France and may be correlated with changes for the firm.<sup>14</sup> We construct our exogenous export demand shock for firm  $f$  at time  $t$  by averaging the world export flows  $M_{jst}$  across destinations  $j$  and products  $s$  using firm  $f$ 's market shares at the initial date  $t_0$ :

$$D_{ft}^{M_s} = \frac{X_{ft_0}^*}{S_{ft_0}^*} \sum_{j,s} \frac{X_{fjst_0}}{X_{ft_0}} \log M_{jst}. \quad (5)$$

The weights  $(X_{ft_0}^*/S_{ft_0}^*)(X_{fjst_0}/X_{ft_0})$  represent firm  $f$ 's initial share of sales of product  $s$  to destination  $j$ .  $X_{ft_0} = \sum_{j,s} X_{fjst_0}$  represents the firm's total exports. In order to incorporate firm  $f$ 's sales to its domestic French market, we scale the firm's export share  $X_{fjst_0}/X_{ft_0}$  by its export intensity  $X_{ft_0}^*/S_{ft_0}^*$ , where  $X_{ft_0}^*$  and  $S_{ft_0}^*$  denote the firm's total exports and total sales taken from the production data. We use asterisks to denote that these data come from a different source than the customs data used to calculate the export shares.<sup>15</sup>

We note that the time variation in our demand shock  $D_{ft}^{M_s}$  only stems from the world export flow  $M_{jst}$  and not the firm-level weights, which are fixed in the initial export period  $t_0$ . We expect that a firm's innovation response at time  $t > t_0$  will induce changes to its pattern of exports at time  $t$  and beyond, including both intensive margin responses (changes in exports for a previously exported product  $s$  to a destination  $j$ ) and extensive margin responses (changes in the set of products  $s$  sold across destinations  $j$ ). By fixing the firm-level weights in the initial period  $t_0$  (including the extensive margin set of products and destinations), we exclude this subsequent endogenous variation from our demand shock. In order to ensure that the time variation in the world export flow  $M_{jst}$  is not influenced by time  $t > t_0$  decisions made by dominant French exporters in market  $(j, s)$  (including, most

---

<sup>13</sup>We consider this firm to be an exporter only if we observed positive exports in our customs data (so we can calculate destination market shares) as well as in our production/administrative data (so we can calculate export intensity).

<sup>14</sup>Later, we will also exclude firms with dominant market shares in any destinations.

<sup>15</sup>Total exports reported by customs and in the production data do not always exactly match, though they are very highly correlated.

importantly, their innovation decisions), we will investigate excluding those firms from our analysis (see appendix 4.4.2).

We will also experiment with alternate measures of this demand shock using more aggregated data (across products). First, we aggregate both the world and the firms' export shares at the 3-digit ISIC level:

$$D_{ft}^{M_I} = \frac{X_{ft_0}^*}{S_{ft_0}^*} \sum_{j,I} \frac{X_{fjIt_0}}{X_{ft_0}} \log M_{jIt},$$

where  $M_{jIt} = \sum_{s \in I} M_{jst}$  measures aggregate imports (excluding France) in destination  $j$  for industry  $I$ , and  $X_{fjIt_0} = \sum_{s \in I} X_{fjst_0}$  is the associated firm-level exports for that industry-destination pair in the initial year  $t_0$ . This measure will no longer reflect the cross-firm variation at the detailed product level. However, it captures some potential spillovers across related products in the construction of the demand shock (an increase in export demand for closely related products may induce a firm to direct innovation towards these related products). Lastly, at the most aggregate level for a destination  $j$ , we will measure changes in GDP (at current exchange rates), and weight those with the firms' export shares in those destinations:

$$D_{ft}^G = \frac{X_{ft_0}^*}{S_{ft_0}^*} \sum_j \frac{X_{fjt_0}}{X_{ft_0}} \log GDP_{jt},$$

where  $X_{fjt_0} = \sum_s X_{fjst_0}$  represents the firm's initial exports to destination  $j$ . One benefit of this aggregation is that it allows us to incorporate the domestic response to demand in destination  $j$ .

The construction of these export demand shocks generates some outliers for a few firms who export just a few products to small destinations (such as yachts to Seychelles and Maldives). We therefore trim from our sample the firms with extreme changes in the export demand variable. We regress the firm demand shock on a firm fixed effect and trim observations with a residual that is above/below the 97.5th/2.5th percentile of that distribution. Thus, firms with the largest variations in their export demand shock (relative to their firm mean) are eliminated.<sup>16</sup>

---

<sup>16</sup>The incidence of these outliers decreases as we aggregate the trade flows from products to industries, and then switch from industries to destination GDP. We have experimented with different threshold cutoffs in the 1-5% range. Our qualitative results are robust to these changes.

## 4.2 Estimation

Our baseline regression seeks to capture the impact of the exogenous demand shock on a firm’s innovation response. We expect this innovation response to be sluggish and incorporate the accumulated effects of the trade shocks over time. We therefore start with an estimation in levels (log-levels for the trade shock) focusing on the within-firm response by incorporating firm fixed-effects. We also add sector-time dummies to capture any sector-level changes over time.

For now, we restrict our analysis to the subset of innovating firms. We will separately investigate the entry margin into the set of innovating firms (a first patent after 1994).<sup>17</sup> We measure their innovation output  $Y_{ft}$  at time  $t$  using their cumulated patents since 1994. We will also use a measure of this stock adjusted for citation counts (whenever the citation is made in our sample years). Given our estimation strategy with firm fixed-effects, we seek to capture changes in this stock over time (driven by the introduction of new patents).

In order to capture the indirect competition effect of an export demand shock (which varies with a firm’s productivity level), we add an interaction between the demand shock and firm productivity. Just as we did with the firm-level export shares, we only use our initial year  $t_0$  to generate a productivity measure that does not subsequently vary over time  $t > t_0$ . We assign a 0-9 productivity index  $d_f$  to all firms based on their labor productivity (value-added per worker) decile in year  $t_0$  within their 2-digit sector.<sup>18</sup>

$$Y_{ft} = \alpha D_{ft} + \beta D_{ft} * d_f + \chi_{s,t} + \chi_f + \varepsilon_{ft}, \quad (6)$$

where the  $\chi_{s,t}$  and  $\chi_f$  capture the sector-time and firm fixed effects.

## 4.3 Baseline results

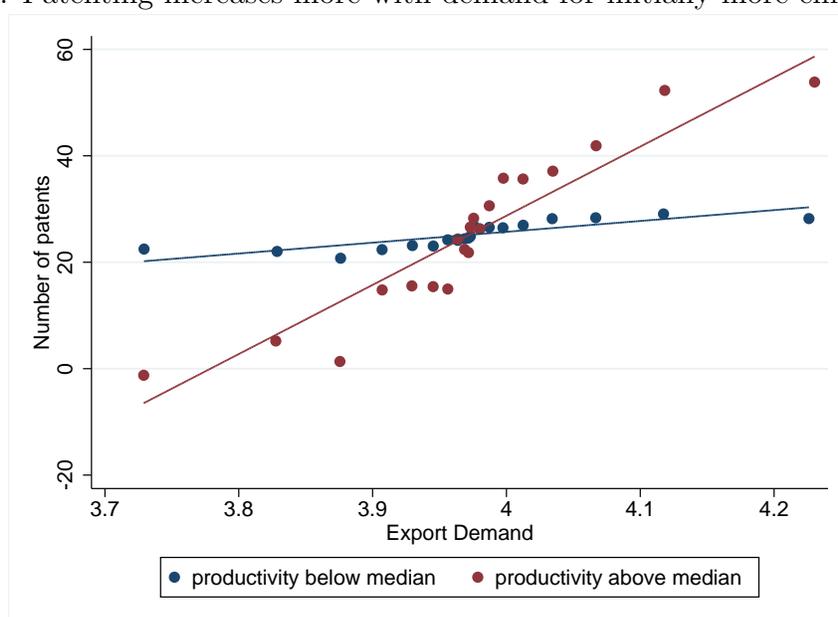
Our model in Section 2 predicts that an increase in market size should have both a positive market size effect and a counteracting competition effect which is most pronounced (and potentially dominant) for the least productive firms (section 2.5). Figure 6 shows that a firm’s patenting responds much more strongly to export demand for manufacturing firms that are initially more productive (a productivity decile  $d_f$  above the median).

---

<sup>17</sup>Our theoretical model highlights that the “entry” decision into innovation in response to an export demand shock will likely have a very different functional form: only the set of firms “close-enough” to the innovation threshold will start innovating in response to a positive demand shock.

<sup>18</sup>When a firm belongs to the manufacturing sector for a subset of our sample years, we only use those years in our estimation. For a firm not in our manufacturing sample at  $t_0$ , we compute its productivity decile within its previous sector at  $t_0$ .

Figure 6: Patenting increases more with demand for initially more efficient firms



**Notes:** Demand and the stock of patents increase together much more for manufacturing firms with productivity above their sectoral median at  $t_0$ . This binscatter absorbs firm fixed effects on both variables, which makes it very close to column (1) of table (4). In detail, the residual of the export demand  $D_{ft}^{Ms}$  on a firm fixed effect is broken down into 20 quantiles. Each dot represents the quantile-average of the residual of the stock of patent on a firm fixed effect (y-axis) against the quantile-average of the residual of the demand on a firm fixed effect (x-axis), for firms below (blue) or above (red) the productivity median at  $t_0$ .

This response of innovation to demand conditions differs markedly between the two groups of firms based on their initial productivity. Our regressions results strongly confirm this qualitative pattern, and show that it is robust to alternative measures of innovation, productivity, demand conditions, and the timing of our sample. The results from our baseline estimation of (6) using all three versions of our export demand shocks (different aggregation levels across products) and our two measures of innovation output (patents and citations) are presented in table 4. This table clearly shows that initially more productive firms respond to an export demand shock by innovating relatively more. For firms in the lowest productivity decile ( $d_f = 0$ ), we see that the effect of the export shock is strongly negative: a positive export shock induces those firms to introduce fewer patents relative to the sector average for that year (given the sector-time fixed effects). In all cases, the effect of the export shock is reversed and strongly positive for all productivity deciles above the median (see Table 7 further down for additional details on this impact broken down by decile).

Table 4: BASELINE REGRESSIONS

Dependent variable	Stock of patents			Stock of citations		
	$D_{ft}^{M_s}$ (1)	$D_{ft}^{M_I}$ (2)	$D_{ft}^G$ (3)	$D_{ft}^{M_s}$ (4)	$D_{ft}^{M_I}$ (5)	$D_{ft}^G$ (6)
Demand	-68.34*** (0.000)	-76.94*** (0.000)	-81.56*** (0.000)	-71.23*** (0.000)	-92.96*** (0.000)	-104.6*** (0.000)
Decile * Demand	19.91*** (0.000)	30.32*** (0.000)	41.64*** (0.000)	21.50*** (0.000)	32.66*** (0.000)	47.99*** (0.000)
Nb of observation	78,823	78,823	77,888	78,823	78,823	77,888

**Notes:** This table presents regression results of an OLS estimation of equation 6. Sample includes manufacturing firms with at least one patent in 1995-2012 for which we can compute the export demand shock (see section 4.1). Stock of patents and citations are computed from 1994. Coefficients and standard errors are obtained using an OLS estimator. p-values of the t statistics are presented in parentheses, computed with autocorrelation and heteroskedasticity robust standard errors. \*\*\*, \*\* and \* respectively indicate 0.001, 0.01 and 0.05 levels of significance.

Our results from column (1) using the product-level demand shock imply the following quantitative response in the number of patents: for a firm in the lowest productivity decile, the stock of patents grows 2.9 patents slower than the sector average ; each additional productivity decile increases the patent response by 0.9 patent; and thus for a firm in the top decile of productivity, the stock of patents increases by 4.7 patents faster than the sector average. Our results using the more aggregated demand shocks (columns (2-3)), or the citation-based measure of innovation (columns (4-6)) yield very similar results.

## 4.4 Additional results

Our main finding – that initially more productive firms respond to an export demand shock by innovating relatively more – is robust to many alternative specifications. In this subsection, we show the robustness of our main results to: (i) adding controls for firm size; (ii) excluding dominant firms in a destination market; (iii) allowing the interaction coefficient to vary across productivity deciles; and (iv) using regressions in long differences .

### 4.4.1 Controlling for size effect

We now control directly for the potential impact of other structural changes within firms. A firm in the midst of a growth spurt (not initially related to innovation) may respond to this increase in scale by innovating and exporting more (nevertheless, there is no strong reason why the firm-level growth spurt would be correlated with the export demand shock). Also, a potential confounding effect for our competition channel is that initially less productive firms may anticipate lower sales responses from a given demand shock – relative to initially more productive firms (although, by construction, our demand shock is independent of overall firm size). To assess the relevance of these channels, we directly control for firm size (at time  $t$ ) in our baseline regression. We select different empirical measures of size: employment, raw materials, net and gross capital stock, and sales. The corresponding regression results are reported in Table 5. They clearly show that a direct control for size does not affect our previously reported baseline coefficients (reported again in column (1)): the coefficients remain virtually unchanged.

### 4.4.2 Excluding markets where a firm is a leader

If a firm has a dominant market share in a market  $(j, s)$ , then the world exports  $M_{jst}$  may be correlated with the firm's exports  $X_{fst}$  (even though French exports are excluded from the construction of the world exports  $M_{jst}$ ). To investigate this further, we drop from our dataset the markets  $(j, s)$  (in all years) for a firm  $f$  whenever its export sales in market  $(j, s)$  are above 10% of world exports (including France) into this market for any given year. These instances represent 6.7% of the customs data observations and happen essentially in African countries. The results are reported in Table 6; and once again leave our baseline results virtually unchanged.

Table 5: CONTROL FOR FIRM SIZE

Dependent variable	Stock of patents					
	$D_{ft}^{Ms}$ (1)	$D_{ft}^{Ms}$ (2)	$D_{ft}^{Ms}$ (3)	$D_{ft}^{Ms}$ (4)	$D_{ft}^{Ms}$ (5)	$D_{ft}^{Ms}$ (6)
Demand	-68.34*** (0.000)	-70.62*** (0.000)	-67.79*** (0.000)	-68.60*** (0.000)	-69.39*** (0.000)	-68.42*** (0.000)
Decile * Demand	19.91*** (0.000)	20.81*** (0.000)	20.08*** (0.000)	20.08*** (0.000)	20.33*** (0.000)	20.03*** (0.000)
log of size		-0.970 (0.302)	0.386 (0.354)	-2.153 (0.010)	-2.055 (0.182)	1.579 (0.073)
Nb of observation	78,823	77,153	77,583	77,764	78,165	78,527

**Notes:** This table presents regression results of an OLS estimation of equation 6. Sample includes manufacturing firms with at least one patent from 1995-2012 for which we can compute the export demand shock (see section 4.1). Stock of patents is computed from 1994. Log of size control corresponds to different measure for firm size: column 2 uses raw material, column 3 uses net capital stock, column 4 uses gross capital stock, column 5 uses employment and column 6 uses sales. Coefficients and standard errors are obtained using an *OLS* estimator. p-values of the t statistics are presented in parentheses, computed with autocorrelation and heteroskedasticity robust standard errors. \*\*\*, \*\* and \* respectively indicate 0.001, 0.01 and 0.05 levels of significance.

Table 6: EXCLUDING FIRMS WITH A DOMINANT MARKET POSITION

Dependent variable	Stock of patents			Stock of citations		
	$D_{ft}^{Ms}$ (1)	$D_{ft}^{MI}$ (2)	$D_{ft}^G$ (3)	$D_{ft}^{Ms}$ (4)	$D_{ft}^{MI}$ (5)	$D_{ft}^G$ (6)
Demand	-70.07*** (0.000)	-80.86*** (0.000)	-75.93*** (0.000)	-75.69*** (0.000)	-101.2*** (0.000)	-99.10*** (0.000)
Decile * Demand	19.83*** (0.000)	27.38*** (0.000)	39.23*** (0.000)	21.30*** (0.000)	29.18*** (0.000)	44.04*** (0.000)
Nb of observation	78,713	78,713	77,799	78,713	78,713	77,713

**Notes:** This table presents regression results of an OLS estimation of equation 6. Sample includes manufacturing firms with at least one patent from 1995-2012 for which we can compute the export demand shock (see section 4.1). The Demand variable does not include country  $j$  and products  $s$  for a firm  $f$  with a market share above 10% for the pair  $(j, s)$ . Stock of patents and citations are computed from 1994. Coefficients and standard errors are obtained using an *OLS* estimator. p-values of the t statistics are presented in parentheses, computed with autocorrelation and heteroskedasticity robust standard errors. \*\*\*, \*\* and \* respectively indicate 0.001, 0.01 and 0.05 levels of significance.

### 4.4.3 Heterogeneous effects across productivity deciles

Here, we slightly alter our econometric model and allow our key interaction coefficient  $\beta$  in equation (6) to vary with the productivity decile  $d_f$ . Our estimating equation then becomes:

$$Y_{ft} = \alpha D_{ft} + \sum_{d=1}^9 [\beta_d D_{ft} * d_f * \mathbb{1}_{d_f=d}] + \chi_{s,t} + \chi_f + \varepsilon_{ft},$$

where  $\mathbb{1}_{d_f=d}$  are indicator dummies for each productivity decile. This new specification allows us to relax the assumption that there is a constant slope shift across decile groups and to account for potential non linear effects of our export demand shock variable for different levels of productivity. Results are reported in Table 7. They suggest that the assumption of a constant effect across decile group is a good approximation, although the effect is the strongest for the 80-90th decile (the next to highest productivity decile). This table highlights that the effect of the demand shock is clearly negative for some of the lowest productivity deciles; and that this effect turns strongly positive for all deciles above the median (deciles 5 through 9).

### 4.4.4 Regression in long differences

We now explore an alternate estimation strategy in long differences. We decompose our full 1996-2011 sample into two periods  $p \in \{p_0, p_1\}$  of equal length. Our demand variable is now measured in log differences as:

$$\Delta D_f^{M_s} = \frac{X_{fp_0}^*}{S_{fp_0}^*} \sum_{j,s} \frac{X_{fjsp_0}}{X_{fp_0}} \log \frac{M_{jsp_1}}{M_{jsp_0}},$$

where all trade flows are aggregated over each period  $p_0$  and  $p_1$ . Similarly we measure innovation output  $\Delta Y_f$  as the introduction of new patents in period  $p_1$  (the change in the stock between  $p_1$  and  $p_0$ ). We also measure the citations associated with those new patents.

Our estimating equation then becomes (we drop the firm fixed-effects but keep a sector fixed effect, and add the firm's productivity decile as a control):

$$\Delta Y_f = \alpha \Delta D_f^{M_s} + \beta \Delta D_f^{M_s} * d_f^{M_s} + \gamma d_f + \chi_s + \varepsilon_f,$$

The results are reported in Table 8. The impact of the demand shock is no longer negative for the lowest productivity decile (it is positive but insignificant). However, we still clearly find a very strong increasing impact of the demand shock with the productivity decile.

Table 7: ALLOWING COEFFICIENT TO VARY ACROSS DECILE GROUP

Dependent variable	Stock of patents			Stock of citations		
	$D_{ft}^{Ms}$ (1)	$D_{ft}^{MI}$ (2)	$D_{ft}^G$ (3)	$D_{ft}^{Ms}$ (4)	$D_{ft}^{MI}$ (5)	$D_{ft}^G$ (6)
Demand	-37.63*** (0.000)	-30.56*** (0.000)	10.92 (0.336)	-41.60*** (0.000)	-37.14*** (0.000)	-1.201 (0.928)
Demand Decile 1	-3.709 (0.627)	-13.26 (0.066)	-54.21*** (0.000)	-5.602 (0.530)	-16.23* (0.050)	-60.27*** (0.000)
Demand Decile 2	-14.92* (0.045)	-7.159 (0.333)	-61.23*** (0.000)	-18.72* (0.026)	-11.29 (0.154)	-60.76*** (0.000)
Demand Decile 3	22.39** (0.009)	13.50 (0.073)	-35.14** (0.002)	22.96* (0.015)	6.968 (0.411)	-52.13*** (0.000)
Demand Decile 4	-11.42 (0.216)	4.747 (0.544)	-35.33** (0.002)	-14.45 (0.137)	2.020 (0.810)	-38.23*** (0.001)
Demand Decile 5	102.4*** (0.000)	98.72*** (0.000)	140.8** (0.003)	154.0*** (0.001)	138.9*** (0.001)	271.9** (0.006)
Demand Decile 6	64.08*** (0.000)	184.8*** (0.000)	232.4*** (0.000)	58.50*** (0.000)	94.78*** (0.000)	148.6*** (0.000)
Demand Decile 7	117.6*** (0.000)	167.4*** (0.000)	174.9*** (0.000)	148.4*** (0.000)	189.4*** (0.000)	187.9*** (0.000)
Demand Decile 8	166.3*** (0.000)	255.3*** (0.000)	339.5*** (0.000)	152.1*** (0.000)	298.8*** (0.000)	416.8*** (0.000)
Demand Decile 9	130.4*** (0.000)	176.7*** (0.000)	200.1*** (0.000)	151.1*** (0.000)	185.8*** (0.000)	248.0*** (0.000)
Nb of observation	78823	78823	77888	78823	78823	77888

**Notes:** This table presents regression results of an OLS estimation of equation 6 where the coefficient on interaction is allowed to vary across decile groups. Sample includes manufacturing firms with at least one patent from 1995-2012 for which we can compute the export demand shock (see section 4.1). Stock of patents and citations are computed from 1994. Coefficients and standard errors are obtained using an OLS estimator. p-values of the t statistics are presented in parentheses, computed with autocorrelation and heteroskedasticity robust standard errors. \*\*\*, \*\* and \* respectively indicate 0.001, 0.01 and 0.05 levels of significance.

Table 8: LONG DIFFERENCE REGRESSIONS

Dependent variable	Number of patents	Number of citations
Sample period	1996 – 2011	1996 – 2011
	(1)	(2)
Decile	0.303* (0.011)	0.306* (0.040)
Demand	0.797 (0.809)	0.939 (0.851)
Decile * demand	2.893** (0.009)	3.481* (0.014)
Nb of observation	5,072	5,072

Notes: \*\*\*, \*\* and \* respectively indicate 0.001, 0.01 and 0.05 levels of significance.

## 5 Conclusion

In this paper we analyzed the impact of export shocks on innovation for French firms. A model of trade and innovation with heterogeneous firms predicts that a positive export shock should raise innovation more for initially more productive firms. Demand shocks generate both market size and competition effects. A larger market size increases the incentives for innovation for all firms, whereas the increased competition generated by the larger market reduces the incentives for innovation most strongly for less productive firms. Our model highlights how the increased competition from higher demand can generate losses and hence lower incentives for innovation for a subset of less productive firms. We find very strong confirmation of both this market size and competition effect for French manufacturing innovators. Our empirical work merges three exhaustive firm-level datasets: customs, patent, and production data. We show that patenting responds very strongly to increases in export demand, but only for relatively more productive firm (closer to the technology frontier). This patenting response steadily increases for firms that are closer to the technology frontier (higher initial levels of productivity).

Our analysis can be extended in several directions. A first direction will be to use the same data to explore the effect of imports on innovation, using the same comprehensive databases. This would allow us to better understand why [Bloom et al. \(2016\)](#) and [Autor et al. \(2016\)](#) get opposite conclusions. A second direction would be to look at the impact of

exports on the citations to previous innovations, thereby shedding new light on the knowledge spillover effects of trade. These await future research.

## References

- ACEMOGLU, D. (2009): *Introduction to modern economic growth*, Princeton Univ. Press, OCLC: 612185973.
- AGHION, P., N. BLOOM, R. BLUNDELL, R. GRIFFITH, AND P. HOWITT (2005): “Competition and Innovation: an Inverted-U Relationship,” *The Quarterly Journal of Economics*, 120, 701–728.
- AGHION, P. AND P. HOWITT (2009): *The economics of growth*, MIT Press, OCLC: 610950320.
- AKCIGIT, U., S. T. ATES, AND G. IMPULLITTI (2014): “Technology Dynamics, Innovation Policy and the Gains from Globalization,” .
- AUTOR, D., D. DORN, G. HANSON, G. PISANO, AND P. SHU (2016): “Foreign Competition and Domestic Innovation: Evidence from U.S. Patents,” DOI: 10.3386/w22879.
- BERNARD, A. B. AND B. J. JENSEN (1999): “Exceptional exporter performance: cause, effect, or both?” *Journal of International Economics*, 47, 1–25.
- BERNARD, A. B., B. J. JENSEN, S. J. REDDING, AND P. K. SCHOTT (2016): “Global Firms,” NBER Working Paper w22727, National Bureau of Economic Research.
- BLOOM, N., M. DRACA, AND J. VAN REENEN (2016): “Trade Induced Technical Change? The Impact of Chinese Imports on Innovation, IT and Productivity,” *The Review of Economic Studies*, 83, 87–117.
- BUSTOS, P. (2011): “Trade Liberalization, Exports, and Technology Upgrading: Evidence on the Impact of MERCOSUR on Argentinian Firms,” *American Economic Review*, 101, 304–340.
- CALDWELL, S. AND M. TABELLINI (2015): “Import Competition and Innovation,” .
- CLERIDES, S. K., S. LACH, AND J. R. TYBOUT (1998): “Is Learning by Exporting Important? Micro-Dynamic Evidence from Colombia, Mexico, and Morocco,” *The Quarterly Journal of Economics*, 113, 903–947.
- COE, D. T. AND E. HELPMAN (1995): “International R&D spillovers,” *European Economic Review*, 39, 859–887.

- DE RASSENFOSSE, G., H. DERNIS, D. GUELLEC, L. PICCI, AND B. VAN POTTELSBERGHE DE LA POTTERIE (2013): “The worldwide count of priority patents: A new indicator of inventive activity,” *Research Policy*, 42, 720–737.
- GAULIER, G. AND S. ZIGNAGO (2010): “BACI: International Trade Database at the Product-Level. The 1994-2007 Version,” Working Papers 2010-23, CEPII research center.
- GROSSMAN, G. M. AND E. HELPMAN (1991a): *Innovation and growth in the global economy*, MIT Press.
- (1991b): “Quality Ladders in the Theory of Growth,” *The Review of Economic Studies*, 58, 43.
- HALL, B., A. JAFFE, AND M. TRAJTENBERG (2005): “Market Value and Patent Citations,” *RAND Journal of Economics*, 36, 16–38.
- IACOVONE, L., W. KELLER, AND F. RAUCH (2011): “Innovation Responses to Import Competition,” .
- INSEE (2016): “NAF REV. 2 et CPF REV. 2 : Guide d’utilisation,” .
- LEQUIEN, M., L. PY, AND P. TRICHLER ( in progress): “Linking patents to firms: insights with French firms,” Working paper, Banque de France.
- LILEEVA, A. AND D. TREFLER (2010): “Improved Access to Foreign Markets Raises Plant-Level Productivity for Some Plants \*,” *Quarterly Journal of Economics*, 125, 1051–1099.
- MAYER, T., M. J. MELITZ, AND G. I. P. OTTAVIANO (2014): “Market Size, Competition, and the Product Mix of Exporters,” *American Economic Review*, 104, 495–536.
- (2016): “Product Mix and Firm Productivity Responses to Trade Competition,” NBER Working Papers 22433, National Bureau of Economic Research, Inc.

## A Data description

### A.1 Patent data

Our first database is PATSTAT Spring 2015 which contains detailed information about patent applications from every patent office in the world. Each patent can be exactly dated to the day of application, which is sometimes referred to as the “filling date”. Moreover, we can retrieve all the future patents making citations to the patents up to 2014.

**Counting patent applications** Each French firm is associated with a number of patent applications by that firm each year (how this match is explained in [Lequien et al., in progress](#)). If the firm shares a patent with another firm, then we only allocate a corresponding share of this patent to the firm. We restrict attention to patents that have been granted by 2015. This raises the well-documented issue of truncation bias [Hall et al. \(2005\)](#). Indeed as we come closer to the end of the sample, we observe a smaller fraction of all patents since many of them are not yet granted <sup>19</sup> In addition, there is a legal obligation to wait 18 months before publication in PATSTAT. With our version of Spring 2015 this implies that we can assume the data to be reasonably complete up to 2012. To avoid this issue, an alternative solution could be to use the year of granting instead of the year of application. However, the former is less relevant than the latter as it is affected by administrative concerns and also by potential lobbying activities that have little to do with the innovation itself. In order to be as close to the time of the innovation as possible, we follow the literature and consider the filling date. We count every patent owned by a French firm, regardless of the patent office that granted the patent rights. Here we need to be aware of the differences in regulations across intellectual property offices. Some patent offices, especially those of Japan and Korea, offer less breadth to a patent, which implies that more patents are needed to protect a given invention than in other patent offices (see [de Rassenfosse et al., 2013](#)). Since we only consider French firms, this would become an issue only if some French firms patent a lot in countries like Japan or Korea, in which case the number of patents by such firms would be artificially large. To check that this problem does not drive our results, we build three different measures of patent counts as a robustness test: (i) we count only priority patents which correspond to the earliest patents which relate to the same invention; (ii) we count the number of patent families: patents are indeed classified into the same patent family when they protect the same invention, even if these patents get granted by different patent

---

<sup>19</sup>The time between patent application and patent granting is a little more than 2 years on average but the distribution of this lag is very skewed with few patent applications still waiting for patent granting many years after the application.

offices; (iii) we count the number of triadic families, that is, families containing patents from the largest three intellectual property offices: Japan (JPO), the US (USPTO) and Europe (EPO). We also look at how our results are affected when Japan is being replaced by China, given the recent rise in Chinese patents.

**Citations** We use PATSTAT information on citations received by patents owned by French firms. Citations are often used to address the problem that all patents are not of equal quality and that simply counting the number of patent applications provides a noisy measure of the true innovation performance of a firm. However, the truncation bias issue is even worse with citations than with patent count. Patents from say 2010 have less time to be cited than patents from 1980 regardless of their respective qualities. Comparing different cohorts of patents can thus lead to misinterpreting what is reflected by the total number of citations received by a firm. One solution to this problem, is to consider the number of citations received within a certain time window after the application date (usually 3 or 5 years) and to set the last year of the sample accordingly (so for 3 year window, we should stop in 2009-2010).

## A.2 Firm-level accounting data

Our second data source provides us with accounting data for French firms from the DGFIP-INSEE, this data source is called FICUS and FARE. The corresponding data are drawn from compulsory reporting of firms and income statements to fiscal authorities in France. Since every firm need to report every year to the tax authorities, the coverage of the data is all French firms from 1994 to 2012 with no limiting threshold in terms of firm size or sales. This dataset provides us with information on the turnover, employment, value-added, the four-digit sector the firm belongs to,...This corresponds to around 47 million observations and the number of observations per year increases from 1.9m to 3.9m over the period we consider.

The manufacturing sector is defined as category C of the first level of the NAF (*Nomenclature d'Activits Franaise*), the first two digits of which are common to both NACE (Statistical Classification of Economic Activities in the European Community) and ISIC (International Standard Industrial Classification of All Economic Activities). Insee provides each firm with a detailed principal activity code (APE) with a top-down approach: it identifies the 1-digit section with the largest value added. Among this section, it identifies the 2-digit division with the largest value-added share, and so on until the most detailed 5-digit APE code (INSEE (2016)). It is therefore possible that another 5-digit code shows a larger value-added

share than the APEidentified, but one can be sure that the manufacturing firms identified produce a larger value-added in the manufacturing section than in any other 1-digit section, which is precisely what we rely on to select the sample of most of our regressions. The 2-digit NAF sector, which we rely intensively on for our fixed effects, then represents the most important activity among the main section of the firm. Employment each year is measured on average within the year and may therefore be a non-integer number. The age of the firm has been retrieved from the reported date of creation.

A unique 9-digit identifier called *Siren number* is associated to each firm, this number is given to the firm until it disappears and cannot be assigned to another firm in the future. When a firm merges with another firm, or is acquired by another firm, or makes significant changes in its organization, this number may change over time. Hence, new entrant *Sirens* in our database do not necessary correspond to new firms.

### A.3 Trade data

**Customs data for French firms** Detailed data on French exports by product and country of destination for each French firm are provided by the French Customs. These are the same data as in Mayer et al. (2014) but extended to the whole 1994-2012 period. Every firm must report its exports by destination country and by very detailed product (at a level finer than HS6). However administrative simplifications for intra-EU trade have been implemented since the Single Market, so that when a firm annually exports inside the EU less than a given threshold, these intra-EU flows are not reported and therefore not in our dataset. The threshold stood at 250 000 francs in 1993, and has been periodically reevaluated (650 000 francs in 2001, 100 000 euros in 2002, 150 000 euros in 2006, 460 000 euros in 2011). Furthermore flows outside the EU both lower than 1 000 euros in value and 1 000 kg in weight are also excluded until 2009, but this exclusion was deleted in 2010.

### A.4 Matching

Our paper is the first to merge those three very large -patent, administrative, and custom-datasets covering exporting French firms. Merging administrative firm-level data from FICUS/FARE and Customs data is fairly straightforward<sup>20</sup> as a firm can be identified by its *Siren* identifier in both datasets. Thus the main challenge is to match either of these two datasets with PATSTAT. Indeed, PATSTAT only reports the name of the patent owner. Not

---

<sup>20</sup>Although one must keep track of the different definitions of firms across these two datasets. Matthieu to expand here.

only can this name be slightly different from the name reported in the other two databases, but it may also change over time, for example because of spelling mistakes. We thus relied on the work of [Lequien et al. \(in progress\)](#) who developed a matching algorithm to map patents with the corresponding French firms. The match is not perfect for reasons that are detailed in their paper, but most of the patents associated to private firms have been successfully merged.

## **A.5 Other data**

two datasets are used to construct the trade demand variables. CEPII's database BACI provides bilateral trade flows in value and quantity for each pair of countries from 1995 to 2015 at the HS6 product level, which covers more than 5,000 products. IMF's *World Economic Outlook* provides country information such as GDP or GDP per capita in ppp.