The Impact of Exports on Innovation: Theory and Evidence*

Philippe Aghion, Antonin Bergeaud,

Matthieu Lequien and Marc J. Melitz

December 22, 2017

Abstract

This paper investigates the effect of export shocks on innovation. On the one hand a positive shock increases market size and therefore innovation incentives for all firms. On the other hand it increases competition as more firms enter the export market. This in turn reduces profits and therefore innovation incentives particularly for firms with low initial productivity. Overall the positive impact of the export shock on innovation is magnified for high productivity firms, whereas it may negatively affect innovation in low productivity firms. We test this prediction with patent, customs and production data covering all French manufacturing firms. To address potential endogeneity issues, we construct firm-level export proxies which respond to aggregate conditions in a firm's export destinations but are exogenous to firm-level decisions. We show that patenting robustly increases more with export demand for initially more productive firms. This effect is reversed for the least productive firms as the negative competition effect dominates. JEL codes

*Addresses - Aghion: Collège de France, LSE and PSE; Bergeaud: Banque de France and PSE; Lequien: Banque de France and PSE; Melitz: Harvard University and NBER. The authors would like to thank, without implicating Daron Acemoglu, Ufuk Akcigit, Elhanan Helpman, Giammario Impullitti, Thierry Mayer, Torsten Persson, Stephen Redding, Thomas Sampson, Dan Trefler, John Van Reenen and participants at the 2017 AEA meeting, the NBER Summer Institute 2017, CIFAR and numerous seminars 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

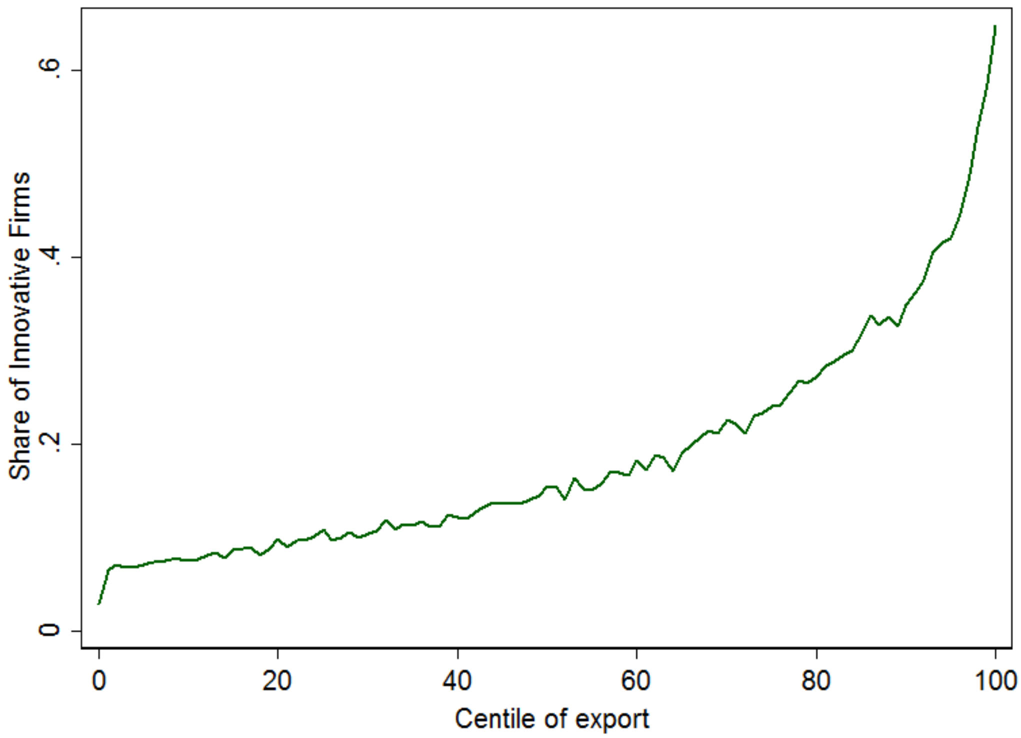
In this paper we combine modeling insights from the literature on firm heterogeneity and trade (see Melitz and Redding, 2014 for a survey) into the new growth theory (e.g., see Aghion and Howitt, 2009). The former literature has focused on the causality link from productivity to trade whereas the latter literature has focused on the reverse link from trade to productivity by investigating the various channels whereby trade liberalization affects innovation-led productivity growth across firms. Our model derives testable predictions on how firms' access to export markets affect innovation; and how this link will vary across exporters. We then take advantage of the availability of exhaustive firm-level data on productivity, trade, and patenting in France to test these predictions.

Figure 1 below motivates our analysis. The curve depicts, for each percentile in exports, the share of French innovative firms within that percentile. It clearly shows a positive relationship between exports and innovation.[1] One of the most striking features that emerges from our merged production-export-innovation dataset 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? 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 is one with monopolistic competition and a fixed cost of exporting, but adds the innovation dimension of new growth theory 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

[1] Obviously the relationship shown in Figure 1 is partly driven by a scale effect: large exporters are larger firms and larger firms are more likely to innovate. However, as we will see below, the distribution of export is more skewed than the distribution of sales or value added. In addition, the positive relationship still exists when centiles of exports intensity (exports divided by sales) are used.

Figure 1: The share of innovators jumps at the top of the export distribution



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

of French firms that export to China. An increase in Chinese demand for products produced by these firms will have two main effects on their innovation incentives. First, a direct market size effect: namely, the expanded market for exports will increase the size of innovation rents and thereby increase those firms' incentives to invest more in innovation. Second, a competition effect: namely, the expanded market for exports will attract new firms into the Chinese market as more firms find it profitable to export there; this in turn 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 our 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 2016 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 literature on trade, innovation and growth (e.g. see Grossman and Helpman, 1991a,b, Aghion and Howitt, 2009, chapter 13; and more recently Akcigit et al., 2014[2]). We contribute to this literature by uncovering

[2]Akcigit et al., 2014 develop and calibrate a new dynamic trade model where firms from different countries compete strategically for leadership in domestic import and export markets. Their model predicts that trade openness encourages innovation in advanced sectors and discourages it in backward sectors.

a new -indirect- effect of market size on innovation working through competition, and by testing the overall effect 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; Bombardini et al., 2017). These papers show that increased import competition induces firms to innovate more in order to escape competition as in Aghion et al. (2005).[3] 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.

[3] Interestingly, in this paper we use *firm-level* competition data, whereas Aghion et al. (2005) as well as previous papers by Nickell (1996) and Blundell et al. (1999) regress innovation and/or productivity growth on *sectoral* measures of product market competition.

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 that are selling in some export market destination D, and we let N denote the number of French firms that could potentially export. We let L denote the number of consumers L in that destination – with income normalized to 1, and we assume that these consumers spend a share of their income on French goods, which for simplicity we normalize at 1. Suppose that the representative consumer in country D has utility for good i which is quadratic and equal to:[4]

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

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

2.1 Consumer optimization

The representative consumer solves:

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

which yields the inverse residual demand function (per consumer):

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

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

Given the assumption of separable preferences, this marginal utility of income λ is the unique endogenous aggregate demand shifter. Higher λ shifts all residual demand curves downwards; we

[4]As we argue below, 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 increases.

thus interpret this as an increase in competition for a given exogenous level of demand.

2.2 Firm optimization

Consider a firm with marginal cost c facing demand conditions λ. This firm chooses the 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 to the following expressions for equilibrium revenues and operating profits:

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

and

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

In particular we see that all three of these per-consumer performance measures (output, revenue, operating profits) are decreasing in both firm level cost c and the endogenous competition measure λ. More productive firms (with lower cost c) are larger and earn higher profits than their less productive counterparts; and an increase in competition λ lowers production levels and profits for all firms.

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 investment in innovation and $\varepsilon > 0$. Without loss of generality, we assume that the cost of innovation is quadratic in k, equal to $c_I k + \frac{1}{2} c_{I2} k^2$.[5]

[5]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

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

$$L\pi(\tilde{c} - \varepsilon k, \lambda) - c_I k - \frac{1}{2} c_{I2} k^2.$$

From the envelope theorem, the optimal R&D investment $k^*(\tilde{c}, \lambda)$ satisfies the first order condition:[6]

$$\frac{\varepsilon L}{2\beta}(\alpha - (\tilde{c} - \varepsilon k^*)\lambda) = c_{I2} k^* + c_I. \tag{2}$$

Figure 2 describes the determination of the optimal innovation investment as the intersection between the marginal cost and marginal gain of innovation, respectively the right and left hand side of equation 2. As long as the marginal gain is above the marginal cost of investing in R&D, the firm wants to invest more, because the marginal profit made by investing one more unit of R&D, at R&D level k, exceeds its cost. The second order condition ensures that the slope of the marginal cost is strictly larger than the slope of the marginal gain, otherwise firms end up doing infinite R&D. When comparing a more productive firm (lower baseline cost, blue curve) and a less productive one (red curve), we see that they face the same marginal cost curve and have the same slope for the marginal gain, only the intercept of the marginal gain is different. Lower $\tilde{c}$ firms have a higher intercept, thus a higher marginal gain for a given level of R&D, and therefore invest more in R&D. Firms with costs too high don't innovate: the intercept of their marginal gain falls below c_I, so that even their first innovation unit would not be worth its cost.

trade-off to consumers in D.

[6]To guarantee an interior solution, we implicitly assume that :

$$\tilde{c} > \widetilde{c}_{\min}$$

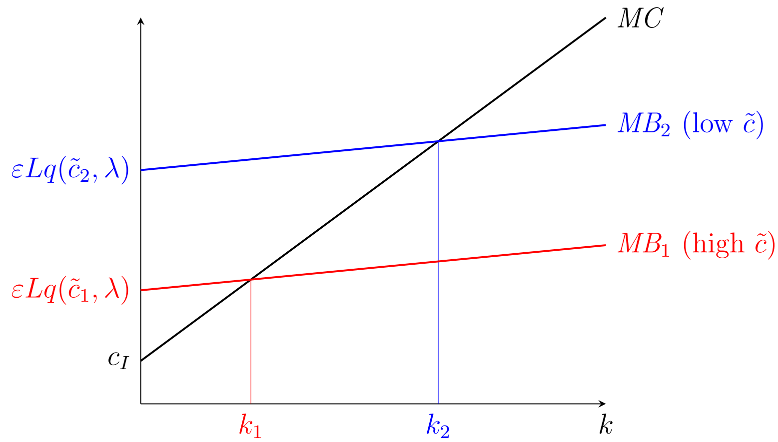
where $\widetilde{c}_{\min}$ satisfies:

$$\widetilde{c}_{\min} - \varepsilon k^*(\widetilde{c}_{\min}, \lambda) = 0$$

or equivalently:

$$\widetilde{c}_{\min} = \frac{1}{\lambda}\left(\alpha - \frac{2\beta c_I}{\varepsilon L}\right) > 0.$$

Figure 2: Optimal innovation is higher for more efficient firms



2.4 The impact of an increase in market size or competition on innovation

Marginal costs do not vary with L or λ, only the marginal gain curve is modified.

Figure 3 shows how innovation responds to an increase in market size. Both the intercept and the slope of the marginal gain curve increase. This leads to an unambiguous higher investment in innovation, for all firms; yet this innovation increase is stronger for more productive firms. More generally, we have:

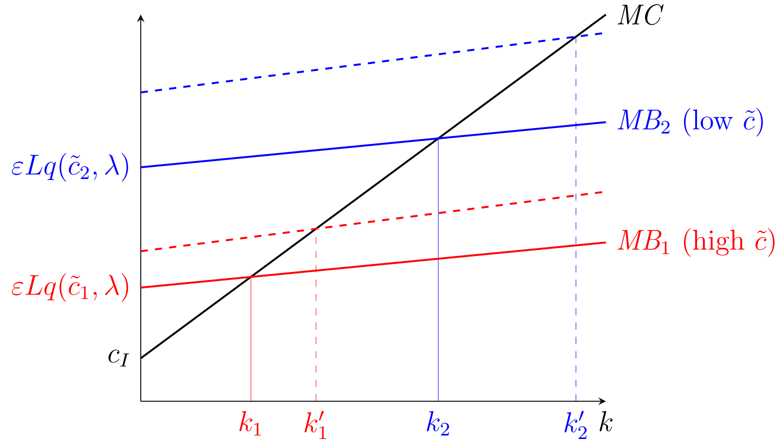
$$\frac{\partial^2 k}{\partial L \partial \tilde{c}} < 0.$$

An increase in the market size makes some firms begin R&D (those with the intercept of the marginal gain just below c_I).

Figure 4 shows how innovation responds to an increase in λ. The marginal gain slope increases but its intercept decreases; however the new dotted curve remains below the old plain one at least until it meets the marginal cost curve. Therefore tougher competition reduces investment in innovation for all firms. Furthermore an increase in λ decreases more the intercept when $\tilde{c}$ is bigger, so that a given competition increase reduces innovation more in less efficient firms:

$$\frac{\partial^2 k}{\partial \lambda \partial \tilde{c}} < 0.$$

Figure 3: Direct market size effect (increase in L)



Besides an increase in λ will make some firms stop R&D.

In the next subsection we endogeneize the competition variable λ by linking it to aggregate market size L and the resulting equilibrium mass of competing firms under free entry.

2.5 Endogenous determination of λ

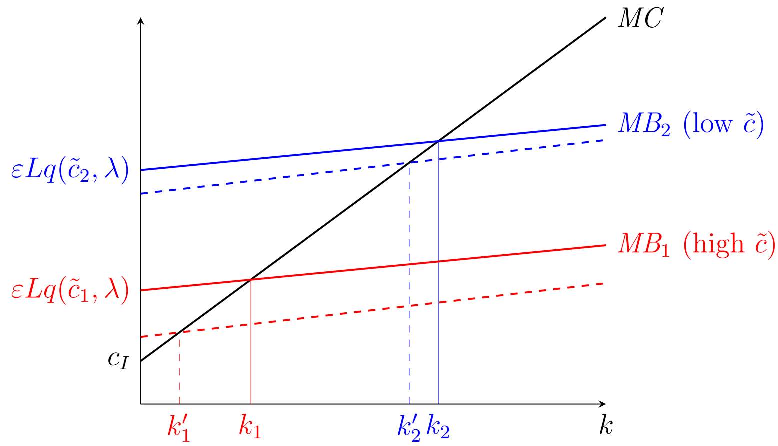
We shall focus attention on the free-entry equilibrium where the marginal firm is indifferent between paying a fixed ex post operating cost F and not entering the export market. Since operating profit is monotonic in a product's baseline cost $\tilde{c}$, the cutoff baseline cost $\hat{c}$ of the marginal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

$$L\pi(\hat{c} - \varepsilon k(\hat{c}, \lambda), \lambda) = F, \tag{3}$$

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

In fact the short-run equilibrium value of the baseline cost cutoff $\hat{c}$ and the equilibrium value of the competition variable λ will be jointly determined by (3) and by 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

Figure 4: Competition effect (increase in λ)



D's consumers spent of French products, i.e. must equal 1. More formally, letting $\Gamma(\tilde{c})$ denote the distribution of baseline costs across French firms, the budget constraint can be expressed as:

$$N\left[\int_{\widetilde{c}_{\min}}^{\widehat{c}} r(\hat{c}-\varepsilon k(\hat{c},\lambda),\lambda)d\Gamma(\tilde{c})\right]=1. \tag{4}$$

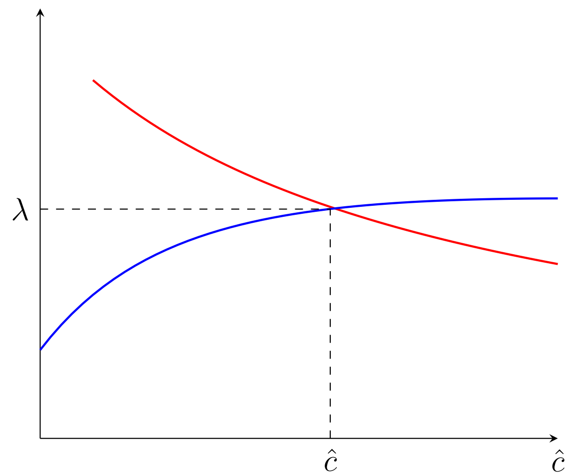
Together, the above two equations (cutoff profit and budget constraint) jointly determine the toughness of competition λ in country D and the baseline cost cutoff $\hat{c}$ as functions of market size L, and of the number of potential French exporters N to country D.

In Figure 5 we represent the above two equations in the $(\widehat{c},\lambda)$ space. The downward-sloping curve (1) represents equation (3) whereas the upward-sloping curve (2) represents equation (4). An increase in L will shift the red curve (1) upward, whereas an increase in N will shift the blue curve (2) upward. Either shift will result in an increase the equilibrium level of competition λ.

2.6 The direct and indirect effects of increased market size

The free-entry and budget balance conditions determine that the equilibrium λ must increase with L as we have just seen in Figure 5 above. The intuition for this induced *competition effect* of increasing export market size, can be explained as follows: an increase in export market size L leads to an increase in the mass of French firms $N\widehat{c}$ in the corresponding industry exporting to

Figure 5: Determination of λ



country D (free entry condition 3); but then each individual French firm in the industry exporting to country D will face a more elastic curve as it faces more competition from other exporting French firms in the same industry to country D, which in turn corresponds to an increase in λ.[7]

Overall, an increase in export market size L induces both:

1. a *direct market size effect* which fosters innovation, more so for more productive firms;

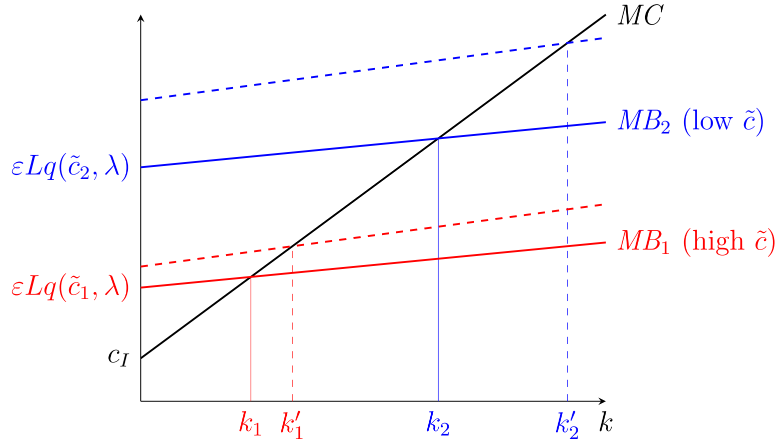
2. an *induced competition effect* which discourages innovation, but less so for more productive firms, i.e. firms with lower $\tilde{c}$.

Figure 6 depicts the overall response of innovation to an increase in market size L. We see that this response is more positive for the low cost firm (in blue) than for the high cost firm (in red), but for both firms the overall innovation response to an increase in market size L is positive.

But somehow this is an incomplete description of a true long-run equilibrium with free-entry: namely, it takes the number N of French firms exporting to country D as given when in fact that number should end up increasing when, as it is the case above, the equilibrium profits of existing firms are all strictly positive. But we know that an increase in N will result in a further increase

[7] 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 λ must increase.

Figure 6: Overall response of innovation to an increase in market size L



in the competition measure λ. Factoring in this extra increase in λ, we may end up with an overall response of innovation to an increase in market size which is positive for low cost firms but negative for high cost firms, as shown in Figure 7 below.

To better understand what underlies the possibility of a negative overall effect of effect of market size on innovation for low productivity (i.e. high cost) firms, in Figure 8(a) we depict the inverse demand curves in logs. We see that an increase in market size L has the direct effect of shifting the demand curve upward, but the induced increase in λ (taking into account the extra effect on competition from the increase in the number of exporters N in a free entry equilibrium) will shift the demand curve to such an extent that at the end the new demand curve will lie above the initial demand curve for high q but below the initial demand curve for low q.[8]

Figure 8(b) depicts the inverse demand curves for the case of CES utility functions: the corresponding inverse demand curves do not satisfy Marshall's Second Law of Demand. In that case, the overall effect of an increase in L amounts to a parallel upward shift in the demand curve, so that it will be positive and the same for all firms no matter their initial productivity levels.

Figures 9(a) and 9(b) depict the log of the equilibrium profit as a function of the log of the firm's productivity, measured by $\varphi = 1/c$. Figure 9(a) corresponds to the quadratic utility case,

[8] If N is fixed, the induced competition effect from increasing market size L will shift the blue demand curve downward, but not enough for it to cross the red demand curve.

Figure 7: Overall response of innovation to an increase in market size L

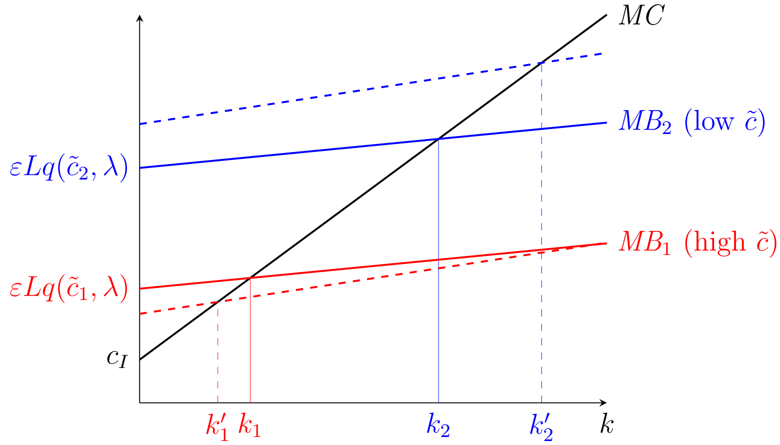


Figure 8: Inverse demand curves

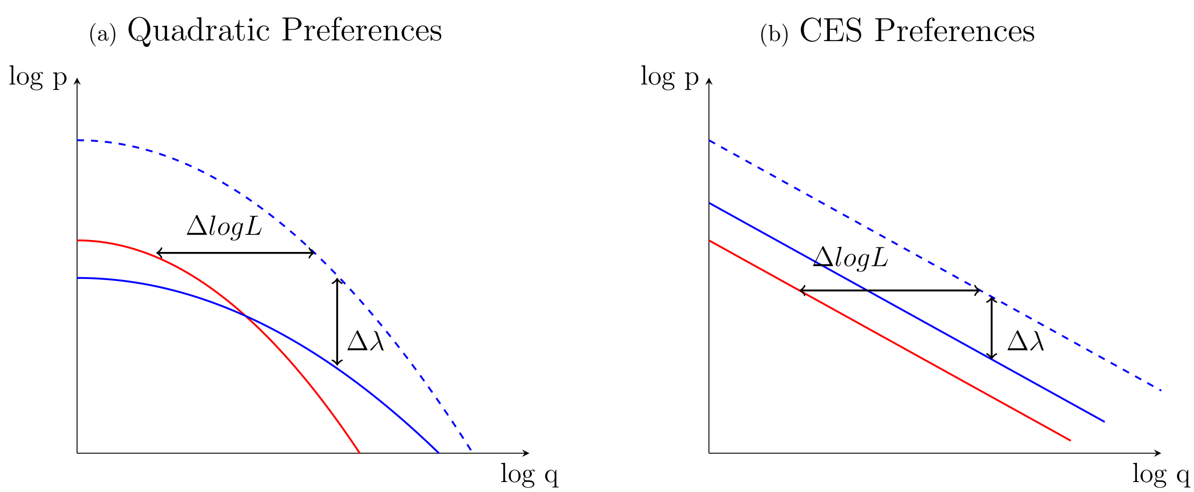
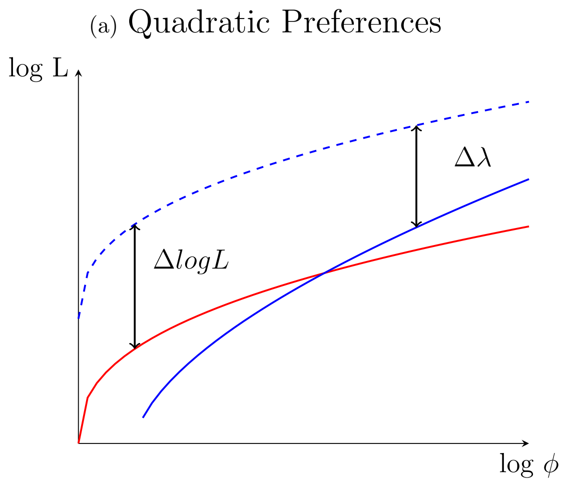


Figure 9: Operating Profit



whereas Figure 9(b) corresponds to the CES case. We see that the overall effect of an increase in L, i.e. factoring in the additional increase in competition λ induced by free entry and the resulting increase in N,is to increase equilibrium profits for high productivity firms (i.e. for firms with low unit cost c) and to reduce equilibrium profits for low productivity firms (i.e. for firms with high unit cost c).[9] This in turn implies that following an increase in market demand innovation will be encouraged for high productivity firms but discouraged for low productivity firms.

In contrast, looking at Figure 9(b) we see that in the CES case an increase in market demand will enhance profits and therefore innovation incentives for firms at all productivity levels. Our empirical analysis will confirm the prediction in Figure 9(a) as we shall see in the next sections.

[9]If N remains fixed, the induced competition effect from increasing market size L will shift the blue curve downward, but not enough for it to cross the red curve: then the overall effect of an increase in market size is to increase equilibrium profits - and therefore innovation incentives - for all firms (yet more for more productive firms, i.e more for firms with higher ϕ).

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, production and innovation data from 1994 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 Spring 2016 vintage of `PATSTAT` patent dataset from the European Patent Office, which contains detailed information on all patent applications from many patent offices in the world. In our analysis we will focus on all patent applications and on patents filed in some specific patent offices (see section 4.2 and 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. This name may sometime carry inconsistencies from one patent to another and/or can contain typos. Various algorithms have been developed to harmonized assignees' names (for example this is the case of the OECD Harmonized Assignee Name database, see Morrison et al., 2017 for a review) but none of those have been applied to French firms. One notable exception is the rigorous matching algorithm developed in Lequien et al. (in progress) to link each patent application with the corresponding French firms' Siren numbers, for all firms with more than 10 employees. This new method, based on supervised learning and described in Appendix A.4, provides significant performance improvements relative to previous methods used in the empirical patent literature: Its recall rate (share of all the true matching that are accurate) is 86.1% and its precision rate (share of the identified matches that are accurate) is 97.0%. This is the matching procedure we use for our empirical analysis in this paper.

Finally, we use CEPII's `BACI` database of bilateral trade flows at the HS6 product level (covering more than 5,000 manufacturing products, see Gaulier and Zignago, 2010) to construct measure of demand shocks across export destinations.

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 business corporations (legal category 5 in the Insee classification). We thus drop state-owned firms, self-employed businesses, and 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.[10] This reduces our working sample to 105,000 firms. Nevertheless the bulk of French aggregate exports and innovation are still concentrated in manufacturing as only 20.6% of aggregate exports and 33% of patents are recorded outside of this sector.

Our dataset does not allow us to properly take into account the case of multinational groups, an issue which often arises when dealing with national firm level data. Multinational groups tend to break the relationship between export shocks and patenting since these groups may locate their R&D activities in different countries from the location of production. In particular, the R&D activity for production based in France may be located elsewhere under a different entity of a multinational's group. In this case, we will not record the appropriate link between the export shocks for this producer and an induced innovation (patents). This measurement issue works against our obtained results of a positive response of patenting to export shocks that is increasing

[10] Although the customs data also covers the wholesale sector, we also exclude those firms as they do not produce the goods that they export.

with a firm's proximity to its industry frontier. Thus, we conjecture that our results would be strengthened if we had the needed information to exclude broken production/R&D links amongst the multinational groups in our sample.

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.[11] 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.[12] 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.

[11]Throughout, we define sectors at the 2-digit level of the European NACE rev2 classification. We also eliminate the tobacco sector (# 12) as it only contains two firms.

[12]The initial year for both ranges do not coincide in order to 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).

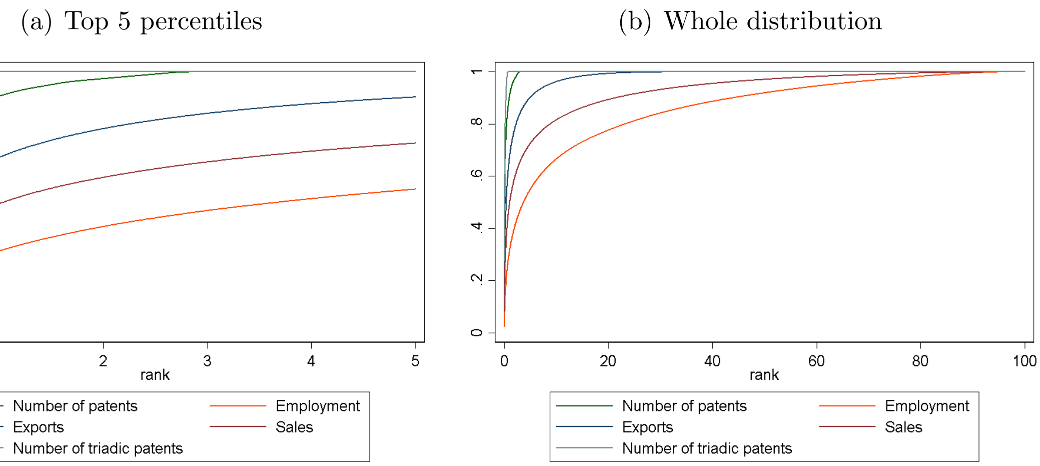
Table 1: Exports and innovation in the manufacturing sector

Sector	Description	Firms	Emp	Export	% Exporter	Patents	% Innov.
10	Food products	8,764	43	1,856	41 (26)	20	2 (0)
11	Beverages	1,452	47	6,018	79 (59)	11	2 (0)
13	Textiles	1,793	37	2,347	86 (63)	91	12 (2)
14	Wearing apparel	1,555	39	2,582	80 (59)	50	5 (1)
15	Leather	487	56	2,593	85 (59)	39	9 (2)
16	Wood	2,424	29	793	63 (36)	16	5 (1)
17	Paper	2,924	44	2,074	79 (44)	86	7 (1)
18	Printing	837	24	168	52 (20)	5	3 (0)
19	Coke	170	225	76,385	92 (69)	3,079	22 (9)
20	Chemicals	1,226	116	17,650	94 (79)	1,997	21 (6)
21	Basic pharmaceutical	356	288	42,183	96 (82)	2,816	35 (13)
22	Rubber and plastic	2,738	80	3,830	86 (64)	340	21 (6)
23	Other non-metallic	2,149	63	2,330	65 (38)	273	11 (2)
24	Basic metals	1,643	80	12,525	65 (44)	148	11 (3)
25	Fabricated metal	8,367	36	1,126	67 (40)	83	9 (2)
26	Computer and electronic	3,497	85	7,650	73 (54)	772	23 (8)
27	Electrical equipment	443	106	8,892	91 (71)	1,780	26 (8)
28	Machinery and equipment	4,656	80	8,274	79 (58)	560	23 (7)
29	Motor vehicles	789	61	2,555	78 (48)	174	15 (3)
30	Other transport equipment	555	215	55,207	83 (56)	2,305	18 (7)
31	Furniture	1,141	34	601	67 (36)	14	7 (1)
32	Other manufacturing	1,005	41	2,501	82 (58)	325	12 (3)
33	Repair of machinery	3,422	28	303	54 (23)	25	6 (1)
	Aggregate manufacturing	52,393	57	4,654	68 (44)	288	11 (3)

Notes: This table presents the number of firms, average employment, average export (in thousands of Euros), average number of patents (in thousands), 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.

Even within the minority set of innovators, patenting activity is extremely skewed. This is clearly visible in Figure 10, which plots the Lorenz curve for manufacturing firms in 2007, along

Figure 10: Lorenz curves - patents are more concentrated than exports, sales and employment



Notes: Lorentz curves plot cumulative distribution function for patents, triadic patents, employment, export and sales. Data are for manufacturing firms.

with the Lorenz curves for exports, sales, and employment. Figure 10 confirms the previously reported finding that firm-level exports are significantly more skewed than sales and employment (e.g. see Mayer and Ottaviano, 2008 and Bernard et al., 2016): 1% of firms 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 10 also shows that patenting is even significantly more skewed than exporting: 1% of all firms account for 91% of patents in 2007. And less than 1% of firms own all the triadic families - i.e. patent families which include patents filed in Asia, Europe and in the USA, see Section 4.2) - all applications are made by the top 1% firms. 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.

3.3 The innovation-export nexus

Looking across our sample years (1995-2012), 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 5% of innovators do not report any exporting; 2) non-exporting innovators do not look very different than non-exporting non-innovators, and the various measures of firm size (employment, sales, value-added) respectively for innovators and non-innovators among non-exporters remain close to each other;[13] 3) these same measures of firm size differ markedly between innovators and non-innovators among exporters: innovators employ on average 4.5 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 three times the number of export destinations. These size differentials are several times larger than those between exporters and non-exporters. In the aggregate, a small subset of innovators accounts for over half of French manufacturing exports.

[13]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.

Table 2: Exporters and innovators are bigger

	Non-exporters		Exporters		Total
	Non-innovator	Innovator	Non-innovator	Innovator	
Firms	45,707	385	51,221	6,770	104,083
Employment	17	21	52	235	59
Sales	2,173	2,530	11,671	69,906	14,075
Value Added	646	908	2,775	16,242	3,354
Age	14	15	20	22	18
Exports	0	0	2,440	23,155	3,622
Countries	0	0	5	18	5
Patents	0	0.2	0	2.6	0.3

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, Countries and Patents are taken as a yearly average over the whole period 1995-2012.

In order to compare exporters to non-exporters and innovators to non-innovators, within specific groups, we compute export and innovation premia. Consider first the exporter premia reported 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 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 since firms who drop in and out of export markets tend to be substantially smaller than year in year out exporters. This is the case for the premia 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 become much more similar. In particular, we find 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

Panel 1: Premium for being an exporter (among all manufacturing firms)					
	(1)	(2)	(3)	Obs.	Firms
log Employment	0.851	0.762	-	931,309	90,688
log Sales	1.613	1.474	0.417	972,956	103,404
log Wage	0.132	0.097	0.110	929,756	90,653
log Value Added Per Worker	0.217	0.171	0.176	918,062	90,055

Panel 2: Premium for being an innovator (among all exporting manufacturing firms)					
	(1)	(2)	(3)	Obs.	Firms
log Employment	1.038	0.993	-	639,938	57,267
log Sales	1.277	1.233	0.197	650,013	57,901
log Wage	0.15	0.095	0.110	638,955	57,253
log Value Added Per Worker	0.203	0.173	0.180	629,819	56,920
log Export Sales (Current period exporters)	2.043	1.970	0.859	433,456	56,509
Number of destination countries	13	12	7	656,609	57,991

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 3-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 have larger total wage bill. 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 fully reflect the concentration of innovative and exporting activities within the more restricted subset of firms that are both exporters and innovators. Figure 1 plots the share of innovating firms for each percentile 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 highly 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.

4 Empirical framework and results

4.1 Identification strategy: firm-level export demand shocks

We have just documented the strong correlation between exports and innovation in the cross-section of French manufacturing firms. However, this correlation does not shed light on the direction of causation: from innovation to exports (a major innovation leads to growth in export demand and entry into new export markets), or from exports to innovation (as we emphasize in our theoretical comparative statics). Moreover, 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 the innovation activity of the firm). We follow Mayer et al. (2016) in building such a measure of exogenous export demand.

To construct these export demand shocks, consider a French exporter f who exports a product s to destination j at an initial date t_0. Let M_{jst} denote the aggregate import flow in product s into country j from all countries except France at time $t > t_0$. M_{jst} reflects the size of the (s, j) export market at time t. We then sum M_{jst} across destinations j and products s weighted by the relative importance of market (s, j) in firm f's exports at the initial date t_0. 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. 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.[14] We then scale the weighted export demand variable by the firm's initial export intensity (at t_0) so that our demand shock scales proportionately with a firm's total production. (As a firm's export intensity goes to zero, so does the impact on any export shock on total production.)

More precisely, let X_{fjst_0} denote firm f's export flow to market (j, s) at time t_0. This is the firm's first observed export year in our sample.[15] The export demand shock for firm f at time t is then:

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

where the weight $(X_{ft_0}^*/S_{ft_0}^*)(X_{fjst_0}/X_{ft_0})$ represents firm f's initial share of sales of product s to destination j and $X_{ft_0} = \sum_{j,s} X_{fjst_0}$ represents the firm's total exports. The asterisks on firm f's initial export intensity $X_{ft_0}^*/S_{ft_0}^*$ indicate that the underlying data for total exports $X_{ft_0}^*$ and sales $S_{ft_0}^*$come from the production data (as opposed to customs data which we use to calculate the destination/product specific market shares).[16]

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.

We will also experiment with an alternate measure of this demand shock using more aggregated data (across products). We thus aggregate both the world and the firm's export shares at the 3-

[14] Another distinct potential source of endogeneity may arise in markets where a French firm has a dominant position. In this case, imports into those markets may respond to this firm's decisions (including innovation). We address this issue in Section 4.5.3.

[15] We consider this firm to be an exporter only if we observe positive exports in both customs data (so we can calculate destination market shares) as well as production data (so we can calculate export intensity).

[16] Total exports reported by customs and in the production data do not always exactly match, though they are highly correlated. One potential source of difference comes from small exports towards other European Union countries which are not reported in custom data (see Appendix A for more details).

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).

Constructing these export demand shocks generates outliers for a few firms that export a small set of products (often highly specialized) to small destinations (such as yachts to Seychelles and Maldives). We therefore trim our sample by removing firms with extreme changes in our constructed export demand. We regress this demand shock on a firm fixed effect and trim observations with a residual that is above/below the 97.5th/2.5th percentile. That is, observations with the largest variations in their export demand shock (relative to their firm mean) are eliminated from our sample.[17]

Before turning to the impact of the export demand shock on the firms' innovation response, we note that this demand shock has very significant explanatory power for a firm's total export response. (See Table B1, which uses a similar estimating strategy as the one we develop in the next section for innovation.)

4.2 Innovation measures

We consider three main measures of innovation for each French firm f in any year t. The first measure counts all patent applications filed by the firm during year t. To better reflect the firm's individual contribution, we use fractional counts for patents shared with other firms (so that a patent filed with 2 other applicants counts as a third). The second measure selects higher quality innovations by counting triadic families of patents: when the same innovation (within a patent

[17]The incidence of these outliers decreases as we aggregate the trade flows from products to industries. We have experimented with different threshold cutoffs in the 1-5% range. Our qualitative results are robust to these changes (see Table B3 in Appendix B).

family) is filed at three different patent offices in Europe (the European Patent Office, EPO), the United States (U.S. Patent Office, USPTO) and at least one of the major Asian economies (Japan, China or Korea).[18] The rationale behind this measure is that the best ideas are more likely to be protected in the three main economic regions worldwide. Another feature of this triadic patent measure is that it is more immune to geographical and institutional biases, i.e. to the possibility that different patent offices would differ with regard to quality standards and/or the enforcement of Intellectual Property Rights (see Park, 2008). When aggregating triadic patent families by firm, we also use fractional counts to reflect a given firm's contribution to the patent family. The third measure counts (fractional) patent applications in Europe (at the EPO) – the main domestic market for French firms. This measure is thus less restrictive than the triadic measure (which requires filing in 3 regions including Europe); but it provides a homogeneous institutional framework for the assessment of intellectual property. Other innovation measures yield similar results (see section 4.5.1). Appendix A provides additional details on the construction of our patent measures.

4.3 Main estimation strategy

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 thus do not use year-to-year differences in the demand shock (first differences) and use that shock as constructed (in "levels") along with a firm fixed-effect in order to capture its within-firm variation (relative to the firm mean over time). We also add sector-time fixed-effects to remove any time variation that is common to the firm's sector. We restrict our analysis to the subset of innovating firms (i.e. firms with at least one patent since 1994).[19]

To capture the indirect competition effect of an export demand shock (which varies with a firm's initial productivity level), we add an interaction between the demand shock and the firm's initial 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 $t > t_0$.[20] We assign

[18]This definition is slightly broader than the definition of triadic patent families used elsewhere in the literature as we consider China and Korea in addition to Japan.

[19]Investigating the entry margin into - or the exit margin out of - the set of innovating firms is also an important topic, but we leave it for further research.

[20]Recall that t_0 is the first year since 1994 in which the firm reports positive exports. This year is equal to 1994

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.[21]

The left-hand panel of Figure 11 shows a bin-scatter of our main patent measure (all applications in year t) against the firm's export demand $D_{ft}^{M_s}$ for the same year – absorbing a firm fixed-effect. This clearly shows that there is a very strong correlation between changes in export demand and changes in patent flows (within firms over time); and that a linear relationship provides a very good functional form fit for that correlation. We thus use a linear OLS specification as our baseline regression equation to quantitatively assess this relationship:

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

where Y_{ft} is one of our innovation measures based on the flow of patent applications during year t by firm f and D_{ft} is one of our export demand proxies ($D_{ft}^{M_s}$ or $D_{ft}^{M_I}$) for that same year t. Those proxies are constructed so that they are exogenous to the firm's decision in year $t > t_0$.[22] The $\chi_{s,t}$ and χ_f capture the sector-time and firm fixed effects and ε_{ft} is an error term. In most of our regressions, we estimate coefficients and standard errors using a Newey-West estimator for the covariance matrix which allow for heteroskedasticity and autocorrelation in the error terms with a maximum lag of 5 years (see Newey and West, 1987 and Wooldridge, 2010).[23]

While the linear panel fixed effect model is our preferred specification, we also show results

for almost 50% of the firms and is always removed from the estimation.

[21] 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.

[22] Serial correlation in the innovation shocks could induce some correlation between a firm's export structure a time t_0 (which we use to construct our export demand shock for year t) and the subsequent innovation shock in year t. First, we note that a time-persistent effect in the innovation shock will be captured by the firm fixed-effect. To ensure that our results are not driven by transitory serial correlation, we also experiment with dropping additional years of data following t_0. Instead of starting our sample at $t_0 + 1$, we have tried starting at $t_0 + 2$ or $t_0 + 3$. We report those alternative specifications in Table B4, which are qualitatively very similar to our baseline results (even though are sample size is reduced in our critical time dimension).

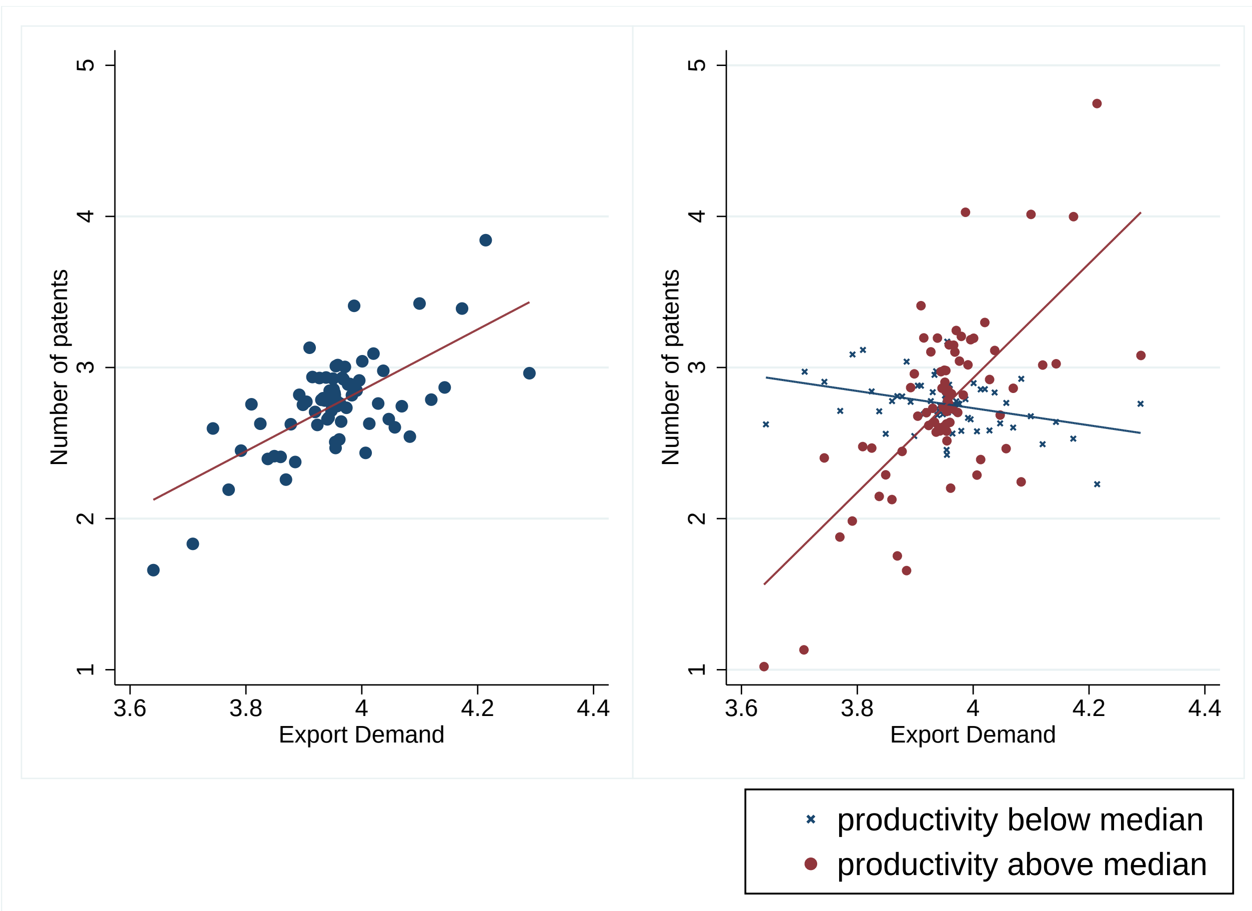
[23] We also show robustness of our baseline results using standard errors clustered by firm (see Table B5). However, given the relatively long time dimension of our sample, we find this latter specification too conservative and prefer the Newey-West estimator. As argued by Abadie et al. (2017), in such fixed-effect models, clustering standard error only matters if we expect heterogeneity in the treatment effect.

from a Poisson model in Table B6 using the following specification:

$$\ln\left(\mathbb{E}\left[\tilde{Y}_{ft}\right]\right) = a\ D_{ft} + b\ D_{ft} * d_f + \chi_{s,t} + \chi_f, \tag{7}$$

where $\tilde{Y}_{ft}$ is a measure of total patent that does not use fractional counts and is therefore an integer. We estimate a and b and corresponding standard errors using maximum likelihood.

Figure 11: Patenting increases more with demand for initially more efficient firms



Notes: Each dot represents the quantile-average of the residual of the number of patent on a firm fixed effect (y-axis) against the quantile-average of the residual of the demand variable $D_{ft}^{M_s}$ on a firm fixed effect (x-axis), for all firms (left-hand side panel) or for both firms below (blue) or above (red) the productivity median at t_0 (right-hand side panel).

Table 4: Baseline results

Dependent variable	All patents		Triadic patents		EPO patents	
Demand measure	$D_{ft}^{M_s}$ (1)	$D_{ft}^{M_I}$ (2)	$D_{ft}^{M_s}$ (3)	$D_{ft}^{M_I}$ (4)	$D_{ft}^{M_s}$ (5)	$D_{ft}^{M_I}$ (6)
Demand	-3.260***	-2.578**	-0.265***	-0.224***	-0.368***	-0.447***
	(1.014)	(1.056)	(0.0786)	(0.0820)	(0.123)	(0.143)
Decile × Demand	0.960***	0.909***	0.0859***	0.0862***	0.125***	0.114***
	(0.255)	(0.304)	(0.0195)	(0.024)	(0.029)	(0.039)
Cutoff decile	4	3	4	3	3	4
Nb of observation	77,901	77,918	77,901	77,918	77,901	77,918
R^2	0.897	0.888	0.759	0.747	0.849	0.836

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 and for which we can compute the export demand shock (see section 4.1). Cutoff decile corresponds to the first value of d_f fo which the overall effect becomes positive. Coefficients and standard errors are obtained using a panel fixed-effect estimator. Autocorrelation and heteroskedasticity robust standard errors using the Newey-West variance estimator with a bandwidth set to 5 years. ***, ** and * respectively indicate 0.01, 0.05 and 0.1 levels of significance.

4.4 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 2.6). The right-hand side panel in Figure 11 summarizes these main findings from our baseline regression: namely, 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) although the overall effect is positive.

The results from our baseline regression (6) 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. This pattern holds for all three patent and two demand shock measures. In all those cases, the effect of the export shock is reversed and strongly positive for all firms in productivity deciles that are large enough (usually the effect becomes positive for $d_f \geq 4$, which almost correspond to the median).

Our coefficients from column 1 using the product-level demand shock imply the following quantitative response in the number of patents to the average demand shock: for a firm in the lowest productivity decile, the number of patents (relative to the firm average) is 3.3 patents lower than the sector average; and each additional productivity decile increases this patent response by 0.96 patent. Thus the response of a firm in the top productivity decile amounts to 8.7 more patents than the sector average (still relative to the firm average). Our coefficients when using the industry demand shocks (column 2), yield very similar results.

4.5 Robustness analysis

Our main finding – that initially more productive (proximity to frontier) firms respond to an export demand shock by innovating relatively more – is robust to various alternative specifications. In this subsection, we show the robustness of our main results to: (i) considering other patent indicators; (ii) controlling for firm specific characteristics; (iii) excluding dominant firms in a destination market; (iv) considering alternative measures and heterogeneous effects for a firm's proximity to frontier; and (v) controlling for pre-trends with sector-decile specific time trends.

4.5.1 Other patent indicators

There are many alternative ways of aggregating patent counts, which yield different measurement of a firm's innovation output. In Table 5 we consider 6 alternative measures of Y_{ft} and run our baseline model (6) with these new dependent variables. (1) We first restrict our triadic measure to a dyadic one by only considering the number of patent families at the USPTO and at the EPO (dropping Asia); (2) We then expand the definition by counting all patent families (instead of individual patents as in our baseline); (3) We return to a count of individual patents but restrict this to EPO patent applications; (4) We count only the number of priority patent applications; (5) We drop the construction of fractional patent counts ("raw" number of patent); and (6) We only count the number of patents applications that will ultimately be granted in any patent office.[24] The results for all these alternative patent measures are similar to those in our baseline Table 4: the export demand shock has a positive effect on the corresponding measure of innovation in frontier firms but has a negative effect on innovation in lagging firms.

[24]See Appendix A for additional details on the construction of all these indicators.

Table 5: Alternative ways of counting patents

Dependent variable	Dyadic	Families	EPO*	Prior	Nb Appln	Granted
Demand Measure	$D_{ft}^{M_s}$	$D_{ft}^{M_s}$	$D_{ft}^{M_s}$	$D_{ft}^{M_s}$	$D_{ft}^{M_s}$	$D_{ft}^{M_s}$
	(1)	(2)	(3)	(4)	(5)	(6)
Demand	-0.233**	-0.936***	-0.883***	-1.364***	-4.732***	-1.197*
	(0.098)	(0.330)	(0.222)	(0.502)	(1.253)	(0.629)
Decile × Demand	0.072***	0.271***	0.245***	0.408***	1.375***	0.363**
	(0.024)	(0.087)	(0.050)	(0.133)	(0.315)	(0.154)
Nb of observation	77,901	77,901	77,901	77,901	77,901	77,901
R^2	0.819	0.802	0.848	0.830	0.881	0.900

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 and for which we can compute the export demand shock (see section 4.1). See Appendix A for a complete definition of the different indicators used in this Table. Coefficients and standard errors are obtained using a panel fixed-effect estimator. Autocorrelation and heteroskedasticity robust standard errors using the Newey-West variance estimator with a bandwidth set to 5 years. ***, ** and * respectively indicate 0.01, 0.05 and 0.1 levels of significance.

4.5.2 Direct control for firm size

A firm experiencing an increase (decrease) in market size which is not initially related to innovation, may still respond to it by innovating and exporting more (less). We do not control for firm size directly in our baseline as our theoretical model suggests that changes in size that are driven by the export shocks should be incorporated into our measure of the impact of exports on innovation. However, in order to eliminate a direct impact of firm scale on our estimates, we now include such a control for firm size (at time t). We select different empirical measures of size from the production data: employment, raw materials, net and gross capital stock, and sales. The corresponding regression results are reported in Table 6. 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. Similar results are obtained when using our two other main measures of innovation (triadic and EPO families) and are available upon request.

4.5.3 Excluding markets where a firm is a leader

When a French 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_{fjst} (even though French exports are excluded

Table 6: Control for firm size

Dependent variable	Number of patents					
Demand measure	$D_{ft}^{M_s}$ (1)	$D_{ft}^{M_s}$ (2)	$D_{ft}^{M_s}$ (3)	$D_{ft}^{M_s}$ (4)	$D_{ft}^{M_s}$ (5)	$D_{ft}^{M_s}$ (6)
Demand	-3.260***	-3.434***	-3.285***	-3.218***	-2.799***	-3.280***
	(1.014)	(1.039)	(1.024)	(1.017)	(1.004)	(1.012)
Decile × Demand	0.960***	0.970***	0.954***	0.960***	0.874***	0.976***
	(0.255)	(0.263)	(0.257)	(0.256)	(0.258)	(0.255)
Size		0.696***	1.257***	2.007***	2.007***	1.227***
		(0.114)	(0.096)	(0.201)	(0.350)	(0.187)
Nb of observation	77,901	76,236	76,678	76,860	77,240	77,605
R^2	0.897	0.900	0.898	0.898	0.898	0.898

Notes: This table presents regression results of an OLS estimation of equation 6 where we add a control for firm size. Column 1 uses no control, column 2 controls for the log of raw material inputs, column 3 (resp. 4) controls for the log of net (resp. gross) capital stock, column 5 controls for the log of employment and column 6 controls for the log of sales (we obtain similar results when we control jointly by any subset of these covariates). Sample includes manufacturing firms with at least one patent in 1995-2012 and for which we can compute the export demand shock (see section 4.1). Coefficients and standard errors are obtained using a panel fixed-effect estimator. Autocorrelation and heteroskedasticity robust standard errors using the Newey-West variance estimator with a bandwidth set to 5 years. ***, ** and * respectively indicate 0.01, 0.05 and 0.1 levels of significance.

from the construction of the world exports M_{jst}) as those other Foreign exporters may respond to actions take by this French firm. 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 predominantly reflect firms exporting to African destinations. The results are reported in Table B7 in the Online Appendix; and once again leave our baseline results virtually unchanged.

4.5.4 Other measures of proximity to frontier

So far, our model uses initial productivity deciles to measure a firm's proximity to its sector's technology frontier. We now consider alternative measures for this proximity. Table7 shows that our baseline results (1) are robust to (2) measuring productivity deciles using sales instead of value added per worker; as well as (3)-(6) measuring proximity to frontier using a binary threshold for initial productivity set at 50%, 75%, 90% and 95%. Those results highlight how the impact of the export shock is magnified for firms very close to the frontier (at the very top of the distribution of initial productivity).[25]

In light of these results, which suggest that the positive effect of export demand on innovation may be concentrated among the most productive firms, we allow our key interaction coefficient β in equation (6) to vary with the productivity decile d_f. Our estimating equation then becomes:

$$Y_{ft} = \sum_{d=0}^{9} \left[\beta_d \; D_{ft} * \mathbb{1}_{d_f=d}\right] + \chi_{s,t} + \chi_f + \varepsilon_{ft}, \tag{8}$$

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. Coefficients and corresponding confident interval are graphically reported in Figure 12. They show that the assumption of a constant effect across decile group is a good approximation – with the possible exception of the top-decile where the effect is magnified relative to the linear trend (this confirms the results from Table 7). This figure also highlights that the effect of the demand shock

[25]Similar results are obtained when using Triadic and EPO families. Tables are available upon request to the authors.

Table 7: Alternative definition of frontier

Dependent variable	Number of patents					
Demand measure	$D_{ft}^{M_s}$ (1)	$D_{ft}^{M_s}$ (2)	$D_{ft}^{M_s}$ (3)	$D_{ft}^{M_s}$ (4)	$D_{ft}^{M_s}$ (5)	$D_{ft}^{M_s}$ (6)
Demand	-3.260***	-1.316	-1.042**	-0.0904	0.520	0.578
	(1.014)	(0.890)	(0.501)	(0.680)	(0.699)	(0.671)
Interaction	0.960***	0.578**	4.438***	5.456***	8.375**	16.63**
	(0.255)	(0.235)	(1.117)	(1.835)	(3.570)	(7.276)
Nb of observation	77,901	77,901	77,901	77,901	77,901	77,901
R^2	0.905	0.905	0.905	0.905	0.905	0.906

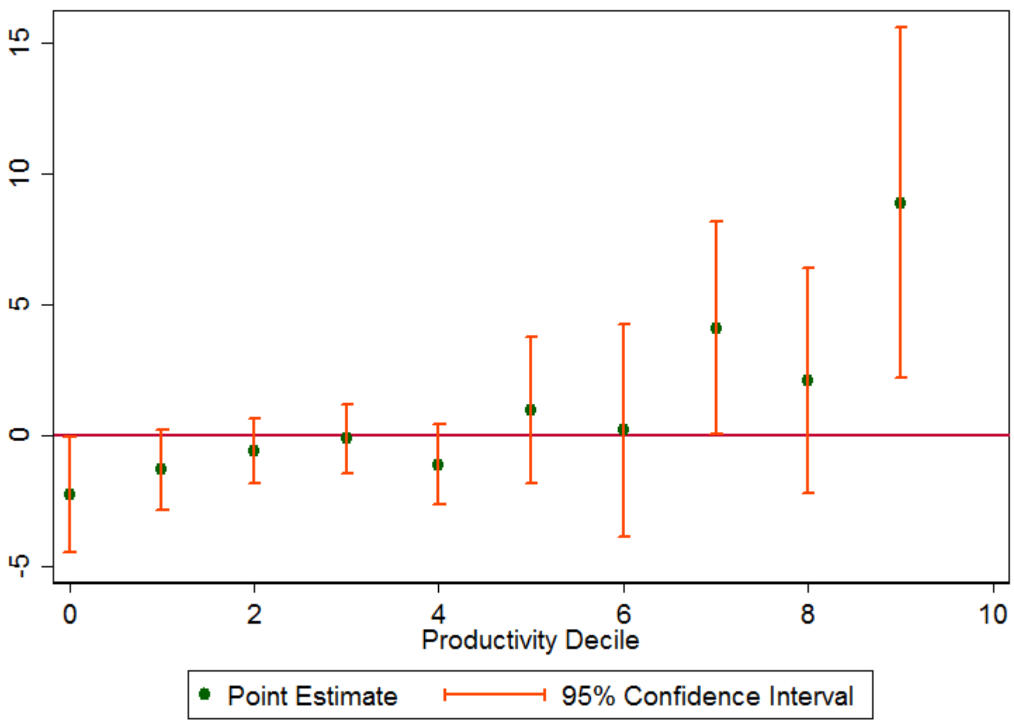
Notes: This table presents regression results of an OLS estimation based on equation 6. Sample includes manufacturing firms with at least one patent from 1994 and to which we can define the variable Demand (see section 4.1). Column (1) is our baseline model, column (2) defines productivity using sales instead of value added, columns (3) to (6) no longer construct decile groups but use a dummy variable for being above the 50^{th}, 75^{th}, 90^{th} and 95^{th} percentile of the initial productivity distribution. This dummy is interacted with the demand variable. Coefficients and standard errors are obtained using an OLS estimator. Autocorrelation and heteroskedasticity robust standard errors using the Newey-West variance estimator with a bandwidth set to 5 years. ***, ** and * respectively indicate 0.01, 0.05 and 0.1 levels of significance.

is clearly negative for some of the lowest productivity deciles; and that this effect turns positive for all deciles above the median (deciles 5 through 9). The confidence intervals remain relatively wide since those decile indicators induce substantial collinearity.

4.5.5 Controlling for sector-decile specific time trends

To deal with the possibility that firms in different productivity deciles and different sectors may evolve differently over time and in particular may follow different innovation trends independently from the export demand shock, we add dummies for all year-productivity decile bins. Those results are reported in the off-numbered columns (1), (3), (5) in Table 8. In addition, we also consider dummies for the triple interaction of all year-sector-productivity decile bins. Those results are reported in the even-numbered columns (2), (4), (6). This sectoral classification is the same that we use to construct the productivity deciles – which are now orthogonal to any sector-level changes over time. Table 8 highlights that these controls do not change the message from our baseline results (though the negative impact for low productivity firms is no longer significant for the case of EPO patent families).

Figure 12: Response to Demand by decile of productivity



Notes: Regression coefficients are estimated using a panel fixed effect estimator and corresponding 95% confident intervals are constructed with Newey-West estimated standard errors of equation (8).

Table 8: Decile group specific evolution

Dependent variable	All patents		Triadic patents		EPO patents	
Demand measure	$D_{ft}^{M_s}$ (1)	$D_{ft}^{M_s}$ (2)	$D_{ft}^{M_s}$ (3)	$D_{ft}^{M_s}$ (4)	$D_{ft}^{M_s}$ (5)	$D_{ft}^{M_s}$ (6)
Demand	-2.876***	-2.135**	-0.179**	-0.135**	-0.168	-0.0755
	(0.975)	(0.868)	(0.0716)	(0.0638)	(0.121)	(0.120)
Decile × Demand	0.900***	0.727***	0.0689***	0.0515***	0.0833***	0.0614**
	(0.267)	(0.251)	(0.0200)	(0.0188)	(0.0316)	(0.0309)
Nb of observations	77,901	77,744	77,901	77,744	77,901	77,744
R^2	0.897	0.899	0.759	0.763	0.849	0.849

Notes: This table presents regression results of an OLS estimation of equation 6, to which a productivity decile × year fixed effect is added (columns 1, 3 and 5) or replacing the sector × time fixed effect with a sector × productivity decile × time fixed effect (columns 2, 4 and 6). Sample includes manufacturing firms with at least one patent in 1995-2012 and for which we can compute the export demand shock (see section 4.1). Coefficients and standard errors are obtained using a panel fixed-effect estimator. Autocorrelation and heteroskedasticity robust standard errors using the Newey-West variance estimator with a bandwidth set to 5 years. ***, ** and * respectively indicate 0.01, 0.05 and 0.1 levels of significance.

4.6 Extensions

In this last section, we extend our empirical analysis in different directions. First, we categorize export destinations based on a separate measure of competition in those destinations and show that the competition effect is most salient in high competition export destinations. We then show how we obtain similar results with an alternate specification based on long time-differences (splitting our sample years into two intervals). This provides an alternate way of capturing the slow-moving changes in the variables of interest (changes in export demand and the associated innovation response). And third, we contrast our main results using measures of innovation output (patents) with results obtained with measures of innovation inputs (R&D inputs) – which are available for a subsample of firms in our sample.

4.6.1 Direct competition effect

We now highlight how the skewed response of innovation to the export shock is driven by the induced competition effect (a demand-side explanation) – as opposed to supply-side effects (such as skewness in the costs or returns to R&D). Towards this end, we use an index of market competition from Djankov et al. (2002) to separate all French export market destinations into high- and low- competition categories. These data on competition levels across countries are now regularly updated and reported in the World Bank's "Doing Business" database. There are several different measures for competition; we use the index reflecting the ease of opening up a business in a country. This generates a time-invariant index by destination on a 0-100 scale.[26]

We then separate destinations into high (H, above median) and low (L, below median) competition according to this index and construct two separate export demand shock measures for those two categories:

$$D_{ft,H}^{M_s} = \frac{X_{ft_0}^*}{S_{ft_0}^*} \sum_{j,s} \mathbb{1}_{C_j > \hat{C}_t} \frac{X_{fjst_0}}{X_{ft_0}} \log M_{jst},$$

and

$$D_{ft,L}^{M_s} = \frac{X_{ft_0}^*}{S_{ft_0}^*} \sum_{j,s} \mathbb{1}_{C_j \leq \hat{C},t} \frac{X_{fjst_0}}{X_{ft_0}} \log M_{jst},$$

where C_j denotes the country-specific competition index (ease of doing business) and $\hat{C}_t$ is the median of this value in year t. Hence $\mathbb{1}_{C_j \leq \hat{C}_t}$ is equal to 1 if country j is less competitive than the median country. These two new demand shocks sum up to our baseline measure $D_{ft}^{M_s}$ and capture separate export demand proxies for destinations with high/low competition.

We then estimate the following model:

$$Y_{ft} = \alpha_H \; D_{ft,H}^{M_s} + \beta_H \; D_{ft,H}^{M_s} * d_f + \alpha_L \; D_{ft,L}^{M_s} + \beta_L \; D_{ft,L}^{M_s} * d_f + \chi_{s,t} + \chi_f + \varepsilon_{ft}. \tag{9}$$

Our theory predicts that we should observe the skewness impact of the demand shocks more (or entirely) for the high-competition destinations. The results reported in Table 9 strongly confirm this prediction. This table considers two separate ways of measuring the threshold value $\hat{C}$. In the odd-numbered columns (1), (3), (5), $\hat{C}$ is the yearly median value once the measure of competition has been aggregated by product (i.e. on a sample containing one observation per product/firm). In the even-numbered columns (2), (4), (6), we use the threshold value when we keep only one observation per country. Both threshold measures for high/low competition confirm that the skewness effect is predominantly driven by the impact of export demand in high competition destinations.

[26] See Appendix A for more details about these data and for an explanation on why we can only construct a time-invariant measure of competition by country.

Table 9: More direct estimation of the competition effect

Dependent variable	All patents		Triadic patents		EPO patents	
Demand Measure	$D^{M_s}_{ft,H}$,	$D^{M_s}_{ft,L}$	$D^{M_s}_{ft,H}$,	$D^{M_s}_{ft,L}$	$D^{M_s}_{ft,H}$,	$D^{M_s}_{ft,L}$
	(1)	(2)	(3)	(4)	(5)	(6)
Low Competition	-0.030	-0.113	-0.026	-0.041	-0.147	-0.140
	(0.457)	(0.465)	(0.031)	(0.032)	(0.092)	(0.092)
High Competition	-5.686***	-5.275***	-0.373**	-0.387**	-0.691***	-0.902***
	(1.980)	(1.769)	(0.153)	(0.151)	(0.222)	(0.221)
Interact. Low	0.086	0.008	0.012	0.013**	0.024	0.001
	(0.144)	(0.086)	(0.009)	(0.006)	(0.029)	(0.014)
Interact. High	2.134***	1.338**	0.182***	0.103**	0.306***	0.179***
	(0.652)	(0.522)	(0.052)	(0.043)	(0.084)	(0.059)
Nb of observation	76,821	76,821	76,821	76,821	76,821	76,821
R^2	0.892	0.896	0.836	0.756	0.861	0.843

Notes: This table presents regression results of an OLS estimation of equation (9). Demand (low comp.) corresponds to $D^{M_s}_{ft,B}$ as defined in section 4.6.1 and Demand (high comp.) to $D^{M_s}_{ft,A}$. Interaction low comp. (resp high comp.) is defined as the interaction between the productivity decile of the firms and $D^{M_s}_{ft,B}$ (resp $D^{M_s}_{ft,A}$). Columns 1, 3 and 5 define the competition median at the firm $\times$ product level each year to compute demand shocks while columns 2, 4 and 6 compute the median at the country level (see section 4.6.1). Sample includes manufacturing firms with at least one patent in 1995-2012 and for which we can compute the export demand shock (see section 4.1). Coefficients and standard errors are obtained using a panel fixed-effect estimator. Autocorrelation and heteroskedasticity robust standard errors using the Newey-West variance estimator with a bandwidth set to 5 years. ***, ** and * respectively indicate 0.01, 0.05 and 0.1 levels of significance.

4.6.2 Regression in long differences

In this section, we explore an alternate estimation strategy based on long differences over time. We decompose our full 1995-2012 sample into two periods $p \in \{p_0, p_1\}$ of equal length. Our demand variable is then measured in log differences, at the product (6 digit HS) or industry (3 digit ISIC) level, as:

$$\Delta D_f^{M_s} = \frac{X^*_{fp_0}}{S^*_{fp_0}} \sum_{j,s} \frac{X_{fjsp_0}}{X_{fp_0}} \log \frac{M_{jsp_1}}{M_{jsp_0}},$$
$$\Delta D_f^{M_I} = \frac{X^*_{fp_0}}{S^*_{fp_0}} \sum_{j,I} \frac{X_{fjIp_0}}{X_{fp_0}} \log \frac{M_{jIp_1}}{M_{jIp_0}},$$

where all trade flows are aggregated over each period p_0 and p_1. Similarly, we measure innovation output ΔY_f as the difference in patent introduction between both periods (same measures as for our baseline analysis).

The firm fixed-effect is differenced-out but we keep the sector fixed effect; and we add the firm's productivity decile as an additional (pre-trend) control. Our estimating equation then becomes:

$$\Delta Y_f = \alpha\ \Delta D_f^{M_s} + \beta\ \Delta D_f^{M_s} * d_f^{M_s} + \gamma\ d_f + \chi_s + \varepsilon_f, \tag{10}$$

The results are reported in Table 10. The impact of the demand shock for the lowest productivity decile is still negative though barely significant. However, the interaction between the demand shock and the firm's initial productivity decile is positive and much more strongly significant. These results confirm our previous baseline findings.

4.6.3 Measures of innovation inputs: R&D investment

Up to now, we have used measures for the output of innovation based on patents. Another commonly used measure for innovation is based on R&D investment (the inputs for the innovation process). As with any input-based measure, the latter generates biases against firms that use those inputs more efficiently to generate innovation. A separate issue is that this measurement of R&D inputs is only available based on survey responses covering a subsample of firms. The sample is exhaustive for the largest innovators, but the sampling frequency decreases steeply with firm size. Thus, smaller firms are not consistently surveyed over time, thwarting the construction of

Table 10: Long Difference regressions

Dependent variable	Nb patents	Triadic families	EPO families
Demand Shock	$\Delta D_f^{M_s}$	$\Delta D_f^{M_s}$	$\Delta D_f^{M_s}$
	(1)	(2)	(3)
Demand	-5.382**	-0.314*	-0.514*
	(2.431)	(0.191)	(0.302)
Decile × demand	1.260**	0.108**	0.140**
	(0.562)	(0.0461)	(0.0612)
Decile	-0.0345	-0.00493	0.00832
	(0.0554)	(0.00356)	(0.00661)
Nb of observation	4,707	4,707	4,707
R^2	0.0197	0.0171	0.0138

Notes: This table presents regression results of an OLS estimation. Sample includes one observation per manufacturing firm with at least one patent in 1995-2012 and for which we can compute the export demand shock (see section 4.1). 1995-2012 is broken down in 2 periods (1995-2003 and 2004-2012), over which trade flows and firm characteristics are aggregated. ***, ** and * respectively indicate 0.01, 0.05 and 0.1 levels of significance.

time-varying measures of innovation.

Table 11 reports the correlations between our three main patent measures and a firm's total R&D budget and the associated number of R&D researchers – whenever this survey data is available. The left-hand side panel reports the between-firm correlations based on firm averages across years. Although the correlations across innovation inputs (R&D) and outputs (patents) in the bottom-left rectangle are weaker than the correlations within a set of inputs or outputs, the between firm correlations are nevertheless substantial and highly significant. However, those correlations between inputs and outputs drop precipitously when focusing on within-firm variations – whereas the correlations within the set of either inputs or outputs remain strong. Those correlations are reported in the middle and right-hand side panels in Table 11. (The middle panel reports within-firm correlations across all years after absorbing a firm fixed-effect; while the right-hand side panel reports the within-firm correlation between periods p_0 and p_1 as defined in our previous long-difference regressions.) Those very low correlations could be driven by the fact that R&D investments within a firm occur at discrete time intervals and slowly translate into increased patents – along with unmeasured changes in the efficiency/utilization of those R&D inputs.

Table 11: Correlations between R&D and patent measures of innovation

Between

	Nb patents	Triadic families	EPO families	R&D budget	Nb researchers
Nb patents	1.00				
Triadic families	0.90	1.00			
EPO families	0.93	0.85	1.00		
R&D budget	0.45	0.39	0.46	1.00	
Nb researchers	0.39	0.33	0.42	0.89	1.00

Within

	Nb patents	Triadic families	EPO families	R&D budget	Nb researchers
Nb patents	1.00				
Triadic families	0.51	1.00			
EPO families	0.68	0.36	1.00		
R&D budget	0.09	0.01	0.07	1.00	
Nb researchers	0.08	0.04	-0.02	0.59	1.00

Long Difference

	Δ Nb patents	Δ triadic fam	Δ EPO fam	Δ R&D budget	Δ Nb researcher
Δ Nb patents	1.00				
Δ triadic fam	0.78	1.00			
Δ EPO fam	0.69	0.34	1.00		
Δ R&D budget	0.11	0.05	0.03	1.00	
Δ Nb researchers	0.07	0.13	-0.08	0.54	1.00

Notes: This table presents the correlations between innovation measures from `PATSTAT` and from the R&D survey. It is based on the sample of manufacturing firms with at least one patent in 1995-2012, for which we can compute the export demand shock (see section 4.1). Each correlation is made on the largest sample possible. Between correlations are the average over all years of the correlations between 2 variables in a given year. Within correlation are the correlations after taking out the firm fixed effect. The long difference correlations are the correlations between the period 1 innovations.

In Table 12, we report the regression results for both our baseline specification as well as the long-difference one using both R&D input measures. In order to separate out the impact of the reduction in sample size associated with the availability of R&D data, we report results using our main patent innovation output variable with the same subsample of firm-years. In the left-hand columns reporting the level regressions, we see that the reduction in sample size does not affect our main results for the skewness impact of the export demand shock on the patent response. The coefficients using the R&D inputs have the same signs, but are not significant. We conjecture that this is due to the fact that the patent measure better captures the within-firm changes in innovation intensity at a yearly frequency. This is confirmed by the results for long-differences, where the coefficients for the R&D inputs are now substantially stronger and significant for the case of the number of researchers. As was the case in our full sample, the significance of the patent response is reduced when moving to the long-difference specification. In this case with a much smaller sample, those coefficients for the patent response are no longer significant.

Table 12: R&D SURVEY

Dependent variable	Regression in levels			Regression in long differences		
Innovation measure	R&D budget (1)	Nb researchers (2)	Nb patents (3)	Δ R&D budget (4)	Δ Nb researchers (5)	Δ Nb patents (6)
Demand	-1441.1	-17.67*	-8.859***	-4228.2	-34.01**	-6.506
	(3015.6)	(9.531)	(3.340)	(4279.9)	(15.03)	(4.242)
Decile $\times$ Demand	1072.7	4.181*	2.219***	1541.8	8.554**	1.560*
	(720.5)	(2.321)	(0.714)	(1048.4)	(3.438)	(0.823)
Decile				-97.65	0.657	0.0835
				(173.3)	(0.568)	(0.102)
Nb of observations	20,030	20,662	21,480	1,746	1,713	1,827
R^2	0.918	0.844	0.896	0.0112	0.0288	0.0301

Notes: This table presents regression results of an OLS estimation of equation 6 (columns 1-3) and 10 (columns 4-6). Sample includes manufacturing firms with at least one patent in 1995-2012 and for which we can compute the export demand shock (see section 4.1); the sample is further reduced to pairs (firm, year) in the R&D survey (columns 1-3) or to firms surveyed at least once in p_0 and p_1 in the R&D survey (columns 4-6). For the regressions in levels, coefficients and standard errors are obtained using a panel fixed-effect estimator. Autocorrelation and heteroskedasticity robust standard errors using the Newey-West variance estimator with a bandwidth set to 5 years. ***, ** and * respectively indicate 0.01, 0.05 and 0.1 levels of significance.

5 Conclusion

In this paper we analyzed the impact of export shocks on innovation for French firms. On one hand those shocks increase market size and therefore innovation incentives for all firms. On the other hand they increase competition as more firms enter the export market. This in turn reduces profits and therefore innovation incentives particularly for firms with low initial productivity. Overall an export demand shock has a more positive effect on innovation in high productivity firms, whereas it may negatively affect innovation in low productivity firms. We tested this prediction with patent, customs and production data covering all French firms. To address potential endogeneity issues, we constructed firm-level variables which respond to aggregate conditions in a firm's export destinations but are exogenous to firm-level decisions. We showed 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. Moreover, we showed that the positive interaction between a firm's initial productivity and the export demand shock is primarily driven by those export destinations where product market competition is highest. This further confirms the fact that export demand shocks involve both, a market size and a competition effect for French manufacturing innovators.

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

Abadie, Alberto, Susan Athey, Guido W. Imbens, and Jeffrey Wooldridge, "When Should You Adjust Standard Errors for Clustering?," Working Paper 24003, National Bureau of Economic Research November 2017.

Aghion, Philippe and Peter Howitt, *The economics of growth*, MIT Press, 2009. OCLC: 610950320.

_ , Nicholas Bloom, Richard Blundell, Rachel Griffith, and Peter Howitt, "Competition and Innovation: an Inverted-U Relationship," *The Quarterly Journal of Economics*, 2005, *120* (2), 701–728.

Akcigit, Ufuk, Sina T. Ates, and Giammario Impullitti, "Technology Dynamics, Innovation Policy and the Gains from Globalization," 2014. Unpublished manuscript.

Autor, David, David Dorn, Gordon Hanson, Gary Pisano, and Pian Shu, "Foreign Competition and Domestic Innovation: Evidence from U.S. Patents," Technical Report 22879, National Bureau of Economic Research 2016. DOI: 10.3386/w22879.

Bernard, Andrew B. and Bradford J. Jensen, "Exceptional exporter performance: cause, effect, or both?," *Journal of International Economics*, 1999, *47* (1), 1–25.

_ , _ , Stephen J. Redding, and Peter K. Schott, "Global Firms," NBER Working Paper 22727, National Bureau of Economic Research 2016.

Bloom, Nicholas, Mirko Draca, and John Van Reenen, "Trade Induced Technical Change? The Impact of Chinese Imports on Innovation, IT and Productivity," *The Review of Economic Studies*, 2016, *83* (1), 87–117.

Blundell, Richard, Rachel Griffith, and John van Reenen, "Market Share, Market Value and Innovation in a Panel of British Manufacturing Firms," *Review of Economic Studies*, 1999, *66* (3), 529–554.

Bombardini, Matilde, Bingjing Li, and Ruoying Wang, "Import Competition and Innovation: Evidence from China," Technical Report 2017.

Bustos, Paula, "Trade Liberalization, Exports, and Technology Upgrading: Evidence on the Impact of MERCOSUR on Argentinian Firms," *American Economic Review*, 2011, *101* (1), 304–340.

Caldwell, Sydney and Marco Tabellini, "Import Competition and Innovation," 2015. mimeo MIT.

Clerides, Sofronis K., Saul Lach, and James R. Tybout, "Is Learning by Exporting Important? Micro-Dynamic Evidence from Colombia, Mexico, and Morocco," *The Quarterly Journal of Economics*, 1998, *113* (3), 903–947.

de Rassenfosse, Gaetan, Helene Dernis, Dominique Guellec, Lucio Picci, and Bruno van Pottelsberghe de la Potterie, "The worldwide count of priority patents: A new indicator of inventive activity," *Research Policy*, 2013, *42* (3), 720–737.

Djankov, Simeon, Rafael La Porta, Florencio Lopez de Silanes, and Andrei Shleifer, "The Regulation of Entry," *The Quarterly Journal of Economics*, 2002, *117* (1), 1–37.

Gaulier, Guillaume and Soledad Zignago, "BACI: International Trade Database at the Product-Level. The 1994-2007 Version," Working Papers 2010-23, CEPII research center October 2010.

Grossman, Gene M. and Elhanan Helpman, *Innovation and growth in the global economy*, MIT Press, 1991.

_ and _ , "Quality Ladders in the Theory of Growth," *The Review of Economic Studies*, 1991, *58* (1), 43.

Hall, Bronwyn, Adam Jaffe, and Manuel Trajtenberg, "Market Value and Patent Citations," *RAND Journal of Economics*, 2005, *36* (1), 16–38.

Iacovone, Leonardo, Wolfgang Keller, and Ferdinand Rauch, "Innovation Responses to Import Competition," 2011. Unpublished manuscript.

INSEE, "NAF REV. 2 et CPF REV. 2 : Guide d'utilisation," 2016.

Lequien, Matthieu, Loriane Py, and Paul Trichelair, "Linking patents to firms: insights with French firms," Working Paper, Banque de France in progress.

Lileeva, Alla and Daniel Trefler, "Improved Access to Foreign Markets Raises Plant-Level Productivity for Some Plants," *Quarterly Journal of Economics*, 2010, *125* (3), 1051–1099.

Mayer, Thierry and Gianmarco I. P. Ottaviano, "The Happy Few: The Internationalisation of European Firms," *Intereconomics - Review of European Economic Policy*, 2008, *43* (3), 135–148.

_ , Marc J. Melitz, and Gianmarco I. P. Ottaviano, "Market Size, Competition, and the Product Mix of Exporters," *American Economic Review*, 2014, *104* (2), 495–536.

_ , _ , and _ , "Product Mix and Firm Productivity Responses to Trade Competition," NBER Working Papers 22433, National Bureau of Economic Research, Inc July 2016.

Melitz, Marc J. and Stephen J. Redding, "Heterogeneous Firms and Trade," in Elhanan Helpman, Kenneth Rogoff, and Gita Gopinath, eds., *Handbook of International Economics*, Vol. 4 of Handbook of International Economics, Elsevier, 2014, pp. 1–54.

Morrison, Greg, Massimo Riccaboni, and Fabio Pammolli, "Disambiguation of patent inventors and assignees using high-resolution geolocation data," *Scientific Data*, 2017, *4.*

Newey, Whitney K and Kenneth D West, "A Simple, Positive Semi-definite, Heteroskedasticity and Autocorrelation Consistent Covariance Matrix," *Econometrica*, May 1987, *55* (3), 703–708.

Nickell, Stephen J, "Competition and Corporate Performance," *Journal of Political Economy*, August 1996, *104* (4), 724–746.

Park, Walter G, "International patent protection: 1960–2005," *Research policy*, 2008, *37* (4), 761–766.

Wooldridge, Jeffrey, *Econometric Analysis of Cross Section and Panel Data* Econometric Analysis of Cross Section and Panel Data, MIT Press, 2010.

Online Appendix for
“The impact of Exports on Innovation: Theory and Evidence”

A Data description

A.1 Patent data

Our first database is `PATSTAT Spring 2016` 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”.

Counting patent applications Each French firm is associated with a number of patent applications by that firm each year (see section A.4). If the firm shares a patent with another firm, then we only allocate a corresponding share of this patent to the firm. 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 [1] In addition, there is a legal obligation to wait 18 months before publication in `PATSTAT`. With our version of Spring 2016 this implies that we can assume the data to be reasonably complete up to 2012. The sector-time fixed effects also deal with the truncation bias in our regressions. 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 filing date. We consider 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

[1] 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.

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 different measures of patent counts as detailed below.

Different counts of patents The various indicators from `PATSTAT` used in the regressions are described in detail below. All these indicators, based on different ways of counting or selecting patents, have pros and cons and shed a different light on our analysis. As stated by de Rassenfosse et al. (2013), it is virtually impossible to define a measure of innovation based on patents that is immune to the various biases that are associated with such data.

Following the innovation literature, we always only select patents of invention (the bulk of patents), thus dropping utility model and design patents.

- `Number of patents`: Each year, we sum over the patents filed by a firm f. When a patent has other applicants than f, we only count the share that f represents among all the co-applicants (one third if f has 2 other applicants). This variable thus is a fractional count, as most of the variables shown in the regressions.

- `Triadic families`: when the same invention is filed in different patent offices, in practice the firm typically files for a different patent at each office, each referring to the first it has filed (called priority patent): these patents relate to each other, they belong to the same (DOCDB) family.[2]. Triadic families refer to such families with at least one patent filed at the EPO, one patent filed at the USPTO, and one patent filed at either the Japanese, the Chinese or the Korean Patent Office. We want to select innovations filed in the 3 main economic regions worldwide (Europe, USA and Asia). We depart slightly from the literature regarding the treatment of Asia: we do not want to consider Japan as the only relevant country, but instead add the two main other innovating countries, China and Korea. Finally the family is weighted with how much f contributes to it: $\frac{\sum_{k \text{ patents} \in \text{ family}} \frac{\mathbb{1}_{\text{f is applicant of } k}}{\text{nb applicants}_k}}{\text{nb patents in the family}}$. The date of the family corresponds to the earliest filing year of the patents in this family.

[2]The `PATSTAT` data catalog states that "a large DOCDB family might indicate that the applicant seeks a wide geographical protection for the invention", and that "if two applications claim exactly the same prior applications as priorities (these can be e. g. Paris Convention priorities or technical relation priorities [. . .], then they are defined by the EPO as belonging to the same DOCDB simple family."

- `EPO families`: The construction is very similar to that of the triadic families, except that the family will be taken into account if there is at least one patent in it filed at the EPO.

- `Dyadic families`: The construction is very similar to that of the triadic families, except that the family will be taken into account if there is at least one patent in it filed at the EPO, and another filed at the USPTO.

- `Families`: The construction is very similar to that of the triadic families, except that we take into account all the families containing a patent applied for by f.

- `EPO*`: We use the fractional count of the patents filed by firm f at the EPO.

- `Raw number of patents`: we use the (non-fractional) count of the number of patents filed by firm f.

- `Only Granted`: We use the fractional count of the granted patents filed by firm f.

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 needs 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 APE identified, 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.

Country-product bilateral trade flows CEPII's database `BACI`, based on the UN database `COMTRADE`, 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.

A.4 Matching

Our paper is the first to merge those three very large - patent, administrative, and customs - datasets covering exporting French firms. Merging administrative firm-level data from FICUS/FARE and Customs data is fairly straightforward[3] 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 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 advanced methodology, described below, is a leap forward compared with other methods proposed by the literature.

Lequien et al. (in progress) proceed in three main steps to merge `PATSTAT` and `SIRENE`:

1. For each *Siren* number from `SIRENE`, find a small subset of applicant firms in Patstat with phonetic similarities:
   - perform cleaning, splitting and phonetic encoding on firms' name in both databases. Too common words are deleted (THE, AND, CO, FRANCAISE ...).
   - sort each name by least frequent encoding in `SIRENE`. The more often a word appears in the database, the less information it can convey to identify firms.
   - for each `SIRENE` firm, the first (ie least frequent) cleaned word of the firm's name is compared with every `PATSTAT` name. All the `PATSTAT` names containing this word form a first subset of possible matches. Then the second word of the firm's name is compared with every name in this subset, reducing it further. This procedure stops before arriving at a null subset, and yields a set of likely `PATSTAT` matches for each `SIRENE` name. Very often this set is null because the majority of firms do not patent. On average, this subset contains 10 applicants, reducing a lot the computationally intensive comparisons.
2. Computation of parameters on these possible matches

[3] Although one must keep track of the different definitions of firms across these two datasets.

- Comparison of the names (raw names, and cleaned names), using Levenshtein distances and an inclusion parameter (all the words in one name are included in the name from the other database)
- zip code comparison (*code postal*)
- date comparisons (a firm cannot have patented before its creation)

3. Matching with supervised learning

   - Sample from INPI (*Institut National de la Propriété Intellectuelle*) with 15,000 true matches between *Siren* number and `PATSTAT` *person id* (and in total 170,000 pairs, with the corresponding known mismatches).
   - This sample is randomly split into a learning sample and a verification sample (this procedure is repeated 10 times, and the recall and precision measures are averaged over them, so that the choice of the sample does not alter the results). This allows to choose the relevant variables and estimate the parameters.
   - apply this model on all the possible matches identified in the previous step.
   - in 90% of cases, unique matching. In the remaining 10% of cases, filter further with a decision tree (is the date of creation of the firm lower than the first filing of the applicant?, which couple has the minimum Levenshtein distance between raw names, between cleaned names, is one of the names included in the other?, which firm has the maximum number of employees?)

   The recall rate (share of all the true matchings that are accurate) is at 86.1% and the precision rate (share of the identified matches that are accurate) is at 97.0%.

A.5 Other data

We also use additional databases at the country level for our analysis. First we use the October 2015 vintage of the IMF's `World Economic Outlook` which provides country information such as GDP and population with a coverage as wide as possible. Second, to measure the level of competition for each country, we use the "Doing Business" project, based on the work of Djankov et al. (2002) and updated by the World Bank. Among all the available information, we consider the "ease of starting a business" which is the variable with the largest spatial coverage. This is a rating of all country for 0 to 100 that measures the constraints when one want to open a new company in the country. Because most countries are not surveyed each year, we choose to take a time invariant average value of this measure as our competition indicator.

A.6 R&D survey variables and sample

A.6.1 The survey

The annual survey on R&D expenses in firms exists since 1963. It describes the private sector R&D in terms of financial means (spending and financing) and mobilized workers. It covers firms established in France and doing R&D, and gathers information on previous year R&D activities. Usually surveys on firms are sampled with the *rpertoire Sirene*, but this database has no information on R&D activities to select firms that one would like to cover. firms with R&D activities represent 1/200 among active firms in *Sirene*.

The Ministry of Higher Education and Research therefore selects firm according to the following procedure:

1. **The historical repertoire**: All units having had a R&D activity are considered. This repertoire is updated with the newest information from the previous survey: takeover, absorption . . . Firms answering they do not do R&D the year of survey but they might do some the following year are kept in this sample.

2. **External sources**: Administrative files and surveys allow to detect new firms possibly doing R&D: firms receiving the *Crdit Impt Recherche*, having the *young innovating firm* status, receiving help from firms incubators, firms reporting R&D activities in other surveys (Community Innovation Survey . . .)

3. **Updating with Sirene cessations**: Firms known as having shut down in *Sirene* are eliminated.

4. **Stratification**:

- Firms with internal spending of R&D over 400 000 are exhaustively interrogated (and above 2 million , they fill a bigger questionnaire).
- New firms (CIR, JEI . . .) are exhaustively interrogated as well (but only since 2001, see below).
- The rest stays only two years in a row in the panel: firms interrogated in N-1 and N-2 are excluded, those interrogated in N-1 are kept, and some newly selected firms are drawn.

Main changes in the survey methodology

- 1992: reform leading to broadly the survey as it exists today. Most variables exist since 1992 or 1993.
- 2000: Increase of the threshold separating the simplified from the general questionnaire, from 5 million francs to 10 million francs. Some variables therefore are missing for firms that filed the simplified questionnaire in 2000.
- 2001: New firms are all interrogated in the first year. Until 2000, only 1 in 2 new firm was interrogated, the rest was kept for the following year.
- Change in units: in 1998, the answer is in francs and not in thousands francs anymore, because many errors were seen. In 2004, the answer is in thousands euros and not in euros anymore. After 2008, the answer is again asked in euros.

Some firms provide a "group" answer. Indeed for larger firms, R&D activity is more often organized at the group level than at the legal unit level. A variable lists the legal units concerned, but only after 2009.

A.6.2 The variables

- Total R&D Budget : total spending of a firm on R&D activities. It is the sum of internal and external spending. One has to be careful, this variable can count twice contracts made between two firms of the same group, once in internal spending and a second time in external spending.

- *of which* Current spending : gross wages of R&D workers and general expenses (spending in capital excluded), such as small tools, raw materials, administrative costs ...
- *of which* Gross wages of R&D workers. It includes all fiscal and social contributions.
- R&D workers : Researchers and technicians (support). in full time equivalent (prorata of time spent in R&D activities, with a minimum of 10%).
- *among which* Nb of researchers : scientists and engineers working at creating knowedge, products, processes, methods or new systems. It includes PhDs paid by the firm or high-level staff responsible with animating the researchers' teams. In full time equivalent.

B Additional Tables

Table B1 presents regression results of an OLS estimation of equation 6 replacing the innovation left-hand-side with (the log of) total firm exports and dropping the export intensity from the demand shock computation. We thus use the following unadjusted export demand variables (at both the product M_s and industry M_I level):

$$\tilde{D}_{ft}^{M_s} = \sum_{j,s} \frac{X_{fjst_0}}{X_{ft_0}} \log M_{jst}$$

$$\tilde{D}_{ft}^{M_I} = \sum_{j,I} \frac{X_{fjIt_0}}{X_{ft_0}} \log M_{jIt}.$$

We see that those unadjusted export demand variables very strongly predict a firm's export response (first two columns). On the other hand, there is no evidence for a skewness effect for that export response – according to a firm's proximity to frontier. Our theoretical model predicts that the skewness effect evolves slowly over time as competition increases. The innovation response is forward looking and captures this anticipated effect, whereas the export measure does not.

Table B1: Impact of the Demand shock on firm's exports

Dependent variable	log(Exports)	log(Exports)	log(Exports)	log(Exports)
Demand Shock	$\tilde{D}_f^{M_s}$	$\tilde{D}_f^{M_I}$	$\tilde{D}_f^{M_s}$	$\tilde{D}_f^{M_I}$
	(1)	(2)	(3)	(4)
Demand	0.0419***	0.0592***	0.0334	0.0537**
	(0.0135)	(0.0164)	(0.0228)	(0.0260)
Decile × demand			0.00203	0.00135
			(0.00396)	(0.00453)
Nb of observation	72,380	72,416	72,380	72,416
R^2	0.855	0.856	0.855	0.856

Notes: Sample includes manufacturing firms with at least one patent in 1995-2012 and for which we can compute an the export demand shocks $\tilde{D}_{ft}^{M_s}$ and $\tilde{D}_{ft}^{M_I}$. Coefficients and standard errors are obtained using a panel fixed-effect estimator. Autocorrelation and heteroskedasticity robust standard errors using the Newey-West variance estimator with a bandwidth set to 5 years. ***, ** and * respectively indicate 0.01, 0.05 and 0.1 levels of significance.

Table B2 considers more aggregated (across industries and products) demand shock measure using the GDP of destination j at t instead of world imports (excluding France) for a particular industry or product. Namely:

$$D^G_{ft} = \frac{X^*_{ft_0}}{S^*_{ft_0}} \sum_j \frac{X_{jt_0}}{X_{ft_0}} \log GDP_{jt}.$$

Table B2: Other demand shock

Dependent variable	All patents	Triadic patents	EPO patents
Demand measure	$D^{M_s}_{ft}$	$D^{M_s}_{ft}$	$D^{M_s}_{ft}$
	(1)	(2)	(3)
Demand	-3.037**	-0.217*	-0.448**
	(1.466)	(0.119)	(0.195)
Decile × Demand	0.852**	0.103***	0.149***
	(0.394)	(0.0328)	(0.0544)
Nb of observations	77,002	77,002	77,002
R^2	0.891	0.753	0.838

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 and for which we can compute an export demand shock D^G_{ft}. Coefficients and standard errors are obtained using a panel fixed-effect estimator. Autocorrelation and heteroskedasticity robust standard errors using the Newey-West variance estimator with a bandwidth set to 5 years. ***, ** and * respectively indicate 0.01, 0.05 and 0.1 levels of significance.

Table B3: Different trimming

Dependent variable	All patents					
Demand Measure	$D_{ft}^{M_s}$	$D_{ft}^{M_s}$	$D_{ft}^{M_s}$	$D_{ft}^{M_s}$	$D_{ft}^{M_s}$	$D_{ft}^{M_s}$
Trimming	2.5% (baseline)	0%	1%	2%	3%	5%
	(1)	(2)	(3)	(4)	(5)	(6)
Demand	-3.260***	-0.751**	-1.994***	-3.052***	-3.491***	-3.474***
	(1.014)	(0.305)	(0.745)	(0.925)	(1.119)	(1.256)
Decile × Demand	0.960***	0.166**	0.510***	0.842***	1.065***	1.086***
	(0.255)	(0.0816)	(0.169)	(0.227)	(0.285)	(0.327)
Nb of observation	77,901	82,043	80,378	78,722	77,077	73,784
R^2	0.897	0.881	0.892	0.896	0.898	0.908

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 and for which we can compute the export demand shock (see section 4.1). Different trimming on extreme variations of the Demand variable are done. Coefficients and standard errors are obtained using a panel fixed-effect estimator. Autocorrelation and heteroskedasticity robust standard errors using the Newey-West variance estimator with a bandwidth set to 5 years. ***, ** and * respectively indicate 0.01, 0.05 and 0.1 levels of significance.significance.

Table B4: Removing first years

Dependent variable	All patents		Triadic patents		EPO patents	
Demand Measure	$D_{ft}^{M_s}$	$D_{ft}^{M_s}$	$D_{ft}^{M_s}$	$D_{ft}^{M_s}$	$D_{ft}^{M_s}$	$D_{ft}^{M_s}$
Years removed	$t < t_0 + 2$	$t < t_0 + 3$	$t < t_0 + 2$	$t < t_0 + 3$	$t < t_0 + 2$	$t < t_0 + 3$
	(1)	(2)	(3)	(4)	(5)	(6)
Demand	-2.703***	-2.378**	-0.233***	-0.203***	-0.339***	-0.321**
	(1.003)	(0.938)	(0.0788)	(0.0733)	(0.129)	(0.129)
Decile × Demand	0.811***	0.654***	0.0752***	0.0620***	0.112***	0.0896***
	(0.249)	(0.229)	(0.0194)	(0.0172)	(0.0298)	(0.0292)
Nb of observation	72,265	66,684	72,265	66,684	72,265	66,684
R^2	0.911	0.920	0.763	0.763	0.870	0.885

Notes: This table presents regression results of an OLS estimation of equation 6. First years following t_0 are excluded from the estimation. Sample includes manufacturing firms with at least one patent in 1995-2012 and for which we can compute the export demand shock (see section 4.1). Coefficients and standard errors are obtained using a panel fixed-effect estimator. Autocorrelation and heteroskedasticity robust standard errors using the Newey-West variance estimator with a bandwidth set to 5 years. ***, ** and * respectively indicate 0.01, 0.05 and 0.1 levels of significance.

Table B5: Baseline results - Clustered standard errors

Dependent variable	All patents		Triadic patents		EPO patents	
Demand measure	$D_{ft}^{M_s}$ (1)	$D_{ft}^{M_I}$ (2)	$D_{ft}^{M_s}$ (3)	$D_{ft}^{M_I}$ (4)	$D_{ft}^{M_s}$ (5)	$D_{ft}^{M_I}$ (6)
Demand	-3.260**	-2.578*	-0.265**	-0.224**	-0.368**	-0.447**
	(1.475)	(1.530)	(0.115)	(0.112)	(0.168)	(0.200)
Decile × Demand	0.960***	0.909**	0.0859***	0.0862***	0.125***	0.114**
	(0.372)	(0.444)	(0.0287)	(0.0327)	(0.0397)	(0.0554)
Nb of observation	77,901	77,918	77,901	77,918	77,901	77,918
R^2	0.897	0.888	0.759	0.747	0.849	0.836

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 and for which we can compute the export demand shock (see section 4.1). Coefficients and standard errors are obtained using a panel fixed-effect estimator. Heteroskedasticity robust standard errors clustered at the firm level ***, ** and * respectively indicate 0.01, 0.05 and 0.1 levels of significance.

Table B6: Poisson regressions

Dependent variable	All patents	Triadic patents	EPO patents
Demand measure	$D_{ft}^{M_s}$ (1)	$D_{ft}^{M_s}$ (2)	$D_{ft}^{M_s}$ (3)
Demand	-0.652***	-0.937***	-0.466*
	(0.208)	(0.360)	(0.246)
Decile × Demand	0.127***	0.164***	0.103***
	(0.0292)	(0.0481)	(0.0336)
Nb of observations	73,488	18,410	47,648

Notes: This table presents regression results of a Poisson estimation of equation 7. To obtain integer dependent variables, we do not use the fractional count (using these integer number with OLS has negligible effects) . Coefficients and standard errors are obtained using a maximum likelihood estimator. Autocorrelation and heteroskedasticity robust standard errors using the Newey-West variance estimator with a bandwidth set to 5 years. Sample includes manufacturing firms with at least one patent in 1995-2012 and for which we can compute the export demand shock (see section 4.1). ***, ** and * respectively indicate 0.01, 0.05 and 0.1 levels of significance.

Table B7: EXCLUDING LEADERS

Dependent variable	All patents		Triadic patents		EPO patents	
Demand measure	$D_{ft}^{M_s}$ (1)	$D_{ft}^{M_I}$ (2)	$D_{ft}^{M_s}$ (3)	$D_{ft}^{M_I}$ (4)	$D_{ft}^{M_s}$ (5)	$D_{ft}^{M_I}$ (6)
Demand	-2.820***	-2.730**	-0.251***	-0.219***	-0.357***	-0.482***
	(0.970)	(1.074)	(0.0838)	(0.0849)	(0.122)	(0.145)
interac	0.766***	1.246***	0.0790***	0.104***	0.122***	0.181***
	(0.239)	(0.318)	(0.0199)	(0.0244)	(0.0296)	(0.0405)
Nb of observation	77790	77806	77790	77806	77790	77806
R^2	0.897	0.887	0.751	0.748	0.846	0.835

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 and 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). Coefficients and standard errors are obtained using a panel fixed-effect estimator. Autocorrelation and heteroskedasticity robust standard errors using the Newey-West variance estimator with a bandwidth set to 5 years. ***, ** and * respectively indicate 0.01, 0.05 and 0.1 levels of significance.