Innovation and trade. Evidence from Italian manufacturing firms

Antonio Accetturo*, Matteo Bugamelli*, Andrea R. Lamorgese*
and Andrea Linarello†

first version: January 2012
this version: March 8, 2012

Abstract

Firms exposed to foreign demand have larger incentives to innovate, if market size matters for innovation. We test this hypothesis using Italian firm data. We measure innovation as the probability of applying for a patent to the European Patent Office, whereas most of the existing papers focus either on productivity or other self-reported measure of innovation. Using information on destination markets of province aggregate exports to construct an instrument for changes in exporting, we find that passing from the 25th to the 75th percentile of the export distribution increases the probability of applying for a patent by half a standard deviation. Our result is heterogenous across firms: export favors innovation to a larger extent for smaller and less productive firms.

JEL classification:

Keywords:

*Research Department, Bank of Italy

†UPF and Bank of Italy, Economic Research Unit, Turin, Italy. Corresponding author, email address: andrea.linarello@bancaditalia.it. We wish to thank . . . for their valuable advice, as well as seminar participants at . . . . The usual disclaimer applies. The views expressed herein are those of the author are not necessarily those of Bank of Italy.

1 Introduction

Innovation is one of the key determinant of aggregate productivity and economic growth. At aggregate level, several studies have shown how openness to trade is a key determinant for technological progress (ANY CITATION, CICCONE GUY?). At micro level, recent advances in international trade show a strong linkages between export and innovation activity. Indeed, exporters are generally larger, more productive and they spend more on R&D. Two possible explanations can rationalize this evidence. The first one concerns a *self-selection* mechanism. As Griliches (2000) points out, the effect of R&D investments on firm level productivity growth is huge. The productivity level also influences the exporting behavior of firms, as conceptualized by Melitz (2003), since only ex ante more productive firms self select into international markets. This implies that firms investing in R&D end up in being more competitive in the international markets. The second explanation refers to the *complementarity* between market size and technological change. As Rodrik (1988) and Yeaple (2005) point out, the incentives for a firm to invest in technology rises with the size of the final market they are going to serve. This implies that trade liberalization induce firms to innovate more. Moreover, trade flows facilitate the international knowledge spillovers (Coe and Helpman, 1995) and, therefore, may contribute to the adoption of new technologies (learning by exporting).

In this paper, we test the complementarity hypothesis. We use Italian firm-level data and the European Patent Office (EPO) records to uncover the causal link that goes from internation trade to incentives to innovate. We find that an increase in firms export has a effect on the probability a firm applies for a patent, which is our measure of innovation. Passing from the 25th to the 75th percentile of export distribution increases the probability of patenting by 15% (half standard deviation). Our contribution is twofold. First, we use a measure innovation output based on patent data, while most of the previous works use either productivity or measure of innovation input (i.e. R&D). Second, we use the growth rate of international demand to construct an instrument for firms export. Italy (as many advanced economies) did not experience any relevant trade liberalization in the recent past. In absence of a natural experiment it is difficult to identify the causal effect of trade on firms outcomes. Our instrument is build on the intuition that the growth rate of international demand has a direct impact on firms export but has no effect on firms outcome [**che vuoi dire con firms outcome?**].

Recently, several works have contributed to the understanding of the impact of exports on productivity. Lileeva and Trefler (2010), using Canada-U.S. Free Trade Agreement, shows that labor productivity of Canadian firms increases as a consequence to U.S. tariff cut. They also document substantial heterogeneity in the responses: ex-ante smaller and low productive firms experience the largest gains. Bustos (2011) shows how Argentinean firms respond to the MERCOSUR Free Trade Agreement by increasing both their export market participation and their spending in technology. This positive effect depends on firms size, indeed firms in the upper-middle range of firms-size distribution benefit more. Her work is among the few ones that focus on a specific channel through which trade might affect productivity, i.e. spending in technology. Our approach is close to hers, since our attention is on innovation that we measure using patent data. Empirical works which do not rely on an identification strategy based on trade liberalization also provide evidence on the impact of trade on productivity. De Loecker (2007) shows that productivity of Slovenian firms increases in response to export market participation. Aw et al. (2011) estimate a structural model and find that productivity evolves depending on export and R&D spending. Their results suggest that investment in R&D have a greater impact on productivity than export, and exports have little impact on the decision to invest in R&D and the following productivity dynamics. Our results, instead, suggests that export market participation has a positive impact on innovation output when measured with patent application.

2 Data

We use firm level data from Bank of Italy industrial survey (INVIND). Participation in the survey is on a voluntary base. The survey is submitted annually to a representative sample of italians manufacturing. Firms are contacted each year, when a firm exits the sample is because it decides not to participate. Exiting firms are replaced to preserve the representativeness of the survey. The attrition is large and does not occur randomly. Most of exiting firms have less than 50 employees. Of the almost 3000 firms that we observe in 2003, only 2000 were surveyed in the sample 2 years before, table 1 shows the distribution of the sample by year. Nonetheless the quality of the data is high because the survey is administrated and conducted directly by the regional branches of the Bank of Italy and ... [**which is the advantage? we need to write it**].[1] We use data starting from 2001 because the survey have been extended to firms with less than 50 but more than 20 employees and the sample size increase substantially. The methodological innovation is of particular relevance in the italian context since the size distribution of manufacturing is skewed towards small firms. Table 2 reports basics facts about the sample. The average log employment is 4.6 while the mean log sales per employee is 5.3 (sales are reported in thousand of euros). Around 77% of firms export, the share of export over total sales is 30%. The export participation and intensity is large compared to the figures reported for other countries and it reflects the minimum size

[1]Data are available upon request to external researcher by the system BIRD `http://www.bancaditalia.it/statistiche/indcamp/sondaggio/bird`

requirement for a firms to be included in the sample, i.e. 20 employees.

Table 1: Sample

	exit			
entry	2001	2002	2003	Sample
2001	330	324	1998	2697
2002		157	522	3001
2003			559	3124

Source INVIND. Authors' calculation. There are 45 firms present in 2001 and 2003. The last column report the actual sample size used in the empirical session.

Table 2: Summary Statistics

	Mean	SD	Min	Max
$\ln(empl)$	4.57	1.195	2.485	10.37
$\ln(sales/empl)$	5.315	.8055	-.8032	11
D_{exp}	.7783	.4154	0	1
$\ln(export)$	6.477	3.996	0	15.87
$Export_{sh}$	.2939	.2989	0	1
D_{pat}	.09794	.2972	0	1
N_{pat}	.2106	1.403	0	26.5

Source INVIND and PATSTAT. Authors' calculation.

Our measure of innovation use patent applications presented by firms and registered in PATSTAT database. PATSTAT is a commercial database compiled by the European Patent Office which contains several information like applicants' name and addresses and the priority date of the application. One common issue in using Patent data is the difficulty to match such data with other data sources. The main problem relies on the fact that patent applicants are recorded according to their name. Companies' names can vary, the same name can be registered in several ways over time or the same company can be registered in different way depending on the patent office to which it applies. Several routines have been developed by the NBER Patent Data Project to harmonize names and to allow the matching of patent applicants to other database. We use the matching between PATSTAT and Italian firms developed by Marin (2011). He follows the NBER routines to harmonize names and then he matches names recorded in PATSTAT to the harmonized names of the Italian firms in AIDA-*Bureau van Dijk* database. Finally we match INVIND firms to PATSTAT using fiscal record.

Some common drawbacks in the use of patent data in economics have been widely documented in the literature [Griliches-Hall] [**a questo punto abbiamo bisogno della citazione giusta, Andrea?**]. First, patent count do not reveal the economic value of patents. Within the same industry is not possible to distinguish between useless and one billion worth patents. Moreover the use of patent vary across industries. Second, which time horizon to observe is unclear because the timing of the innovation outcomes is uncertain: while we know that firms sustain fix costs to set up their research lab, we cannot infer from the information contained in this data set how long before the innovation they need to sunk such cost, we do not know either when the innovation takes place and how long it takes to the firm to file a patent request, which might depend on strategic considerations that have little to do with the production process of the innovation itself. We will tackle these problems by considering within industries differences in the empirical specification and by looking to a time window of four years, which correspond to the median average citation lag in the NBER patent database. The last two rows of table 2 report the average probability of a firm applying for at least one patent in a 4 year time period and the average number of patent applications presented by each firm.

To build our instrument we use aggregate data on exports from two different sources. We measure the growth rate of international demand using world trade flows from Baci-Cepii dataset. The Baci dataset build on UN comtrade but harmonize the data to reconcile flows reported by importing and exporting countries.[2] Another source of information is the aggregate export flows registered by the Italian Statistical Agency (ISTAT-COEWEB) at the sector-province-destination level of disaggregation. Provinces are administrative areas which correspond to NUTS3 areas. We observe 95 different areas. We use aggregate province export flows to build specific weights, the share of export for sector-destination countries, that we use to construct a sector-province specific growth rate of demand.

3 Empirical design

We asses the impact of export flows on innovation activity. To do that we estimate the following equation:

$$D\{Patent\}_{ijp}^{\{t,t+4\}} = \alpha + \alpha_j + \beta X_{ij} + \gamma Z_{ij} + \varepsilon_{ijp}, \qquad (1)$$

for $t = \{2001, 2002, 2003\}$, where $D\{Patent\}_{ijp}^{\{t,t+4\}}$ is a dummy variable taking value equal to 1 if the firm files the application for a patent in the following four years (between t and $t+4$), α_j is a sector fixed effect, X_{ij} is a variable which take values equal to $\ln(export)_{ij}$ if $export_{ij} > 0$ and 0 otherwise, Z_{ij} is a set of firm's control.

[2]For further details see `http://www.cepii.fr/anglaisgraph/bdd/baci.htm`

$D\{Patent\}_{ijp}^{\{t,t+4\}}$ is our proxy of innovation. Since patenting features some lumpiness, due to both the uncertainty of research and development and strategic considerations,[3] it makes little sense to proxy innovation with the probability of filing a patent application in a single year. We choose a four-years period for similarity with the median and average citation lag in NBER patent data, that is the time elapsing between the moment in which a certain patent is granted and the first time that patent is referred to in the application for a new patent. We assume that that anyone citing a previous patent starts its innovation process building on a previous patent as soon as it is granted and achieves the innovation when filing its own application, so that the average citation lag proxies the average duration of the innovation process.

Since both exporting and innovation/patenting are activities with huge start-up costs and a steep learning curve, it makes a big difference whether the firm is already an exporter or it has already filed a few patents for innovations it achieved. For this reason our firm's controls Z_{ij} include the stock of patents granted to the firm at $t-4$, the average annual growth rate of firm's exports between 1995 and $t-4$. Since regional heterogeneity is a specific feature of Italian data, in robustness checks we also include a dummy which takes value equal to 1 if the firm is located in either Central or Southern Italy and zero otherwise.

As it will became clear, our data are an unbalanced panel in which in each year we ask the firm for its current and its previous export flow. We then avoid to estimate a single panel, and we rather estimate our model separately in three different years, namely 2001, 2002 and 2003.

3.1 Causality

OLS estimation of equation (1) is potentially plagued by an endogeneity bias. More productive firms might simultaneously export more and have a larger probability of patenting. To obviate this problem, we recur to instrumental variable (henceforth IV) estimation. We use the changes of international demand as an instrument for exports. As in Lileeva and Trefler (2010) and Angrist and Imbens (1995), the actual instrument is a set of mutually orthogonal dummies built according to the within industry quartiles of the distribution of the changes in the international demand.

More in detail, we build the instrument in two steps. We first compute international demand for the sector x province cell: $\hat{X}_{jpt} = \sum_c X_{cjpt_{1995}} M_{cjt}/M_{cjt_{1995}}$, with $t = \{2001, 2002, 2003\}$. We then compute the within-industry distribution of the changes in international demand and build a set of mutually orthogonal dummies according to the quartile to

[3] A firm may decide not apply for a patent for an innovation it discovered in order to avoid the disclosure of specific knowledge required when filing a patent application (Reinganum, 1983, 1984, 1986).

which each firm belongs. That is, let's call $q^n_{\Delta \ln \hat{X}_{jt}}$ with $n = 25, 50, 75, 100$ the upper bound of respectively the 1st, the 2nd, the 3rd and the 4th quartile of the within sector distribution of $\hat{X}_{jpt}$, then build the following set of dummies:

$$D^1_{\Delta \ln \hat{X}_{pjt-1}} = \begin{cases} 1 & \text{if } \Delta \ln \hat{X}_{pjt-1} \leq q^{25}_{\Delta \ln \hat{X}_{jt-1}} \\ 0 & \text{otherwise} \end{cases} \quad (2)$$

$$D^2_{\Delta \ln \hat{X}_{pjt-1}} = \begin{cases} 1 & \text{if } q^{25}_{\Delta \ln \hat{X}_{jt-1}} < \Delta \ln \hat{X}_{pjt-1} \leq q^{50}_{\Delta \ln \hat{X}_{jt-1}} \\ 0 & \text{otherwise} \end{cases} \quad (3)$$

$$D^3_{\Delta \ln \hat{X}_{pjt-1}} = \begin{cases} 1 & \text{if } q^{50}_{\Delta \ln \hat{X}_{jt-1}} < \Delta \ln \hat{X}_{pjt-1} \leq q^{75}_{\Delta \ln \hat{X}_{jt-1}} \\ 0 & \text{otherwise} \end{cases} \quad (4)$$

$$D^4_{\Delta \ln \hat{X}_{pjt-1}} = \begin{cases} 1 & \text{if } \Delta \ln \hat{X}_{pjt-1} > q^{75}_{\Delta \ln \hat{X}_{jt-1}} \\ 0 & \text{otherwise.} \end{cases} \quad (5)$$

Our instrument is the subset $\{D^2_{\Delta \ln \hat{X}_{pjt}}, D^3_{\Delta \ln \hat{X}_{pjt}}, D^4_{\Delta \ln \hat{X}_{pjt}}\}$. Being international demand ($\hat{X}_{jt-1}$) reasonably exogenous with respect to the probability to apply for patent of a single Italian firm in the following four years, we can assume that a set of dummies built on such variable to be a reasonable instrument for the lagged log-level of exports of that firm.

4 Results

Our baseline estimation of equation 1 includes only lagged stock of patents and lagged average export dynamics as controls. Table 3 shows that the probability of innovating is correlated with export: firms which at time $t-1$ have been exposed to larger foreign demand are more likely to innovate in the following four years. Results are very consistent across the three different years.

Table 3: Baseline OLS

	2001	2002	2003
X_{jpt-1}	0.0159***	0.0154***	0.0186***
	[0.0013]	[0.0013]	[0.0013]
R^2	0.123	0.120	0.141
Observations	2697	3001	3124

Notes: Robust standard errors in brackets. Each regression includes 2-digit sector fix effects. Significance: * 10%, ** 5%, *** 1%.

As said, estimates in table 3 are potentially plagued by an endogeneity bias. In table 4 we recur to IV estimation using the instrument described in section 3.1.

The coefficients of the first stage have the expected signs and the first stage F-statistics greater than 10, safely above the standard levels of the weak instruments literature (Bound et al., 1995). The estimates show that larger export increase the probability to file a patent and that such effect is causal, that is robust to reverse causality and omitted variable bias. For what concerns the magnitude of the estimated effect, we find that passing from the 25th to the 75th percentile of export distribution increases the probability of patenting by 15%, that is half a standard deviation. [The IV estimates also show a downward bias in the OLS estimates.] Results are very consistent across the three different years.

Table 4: Baseline IV

	2001	2002	2003
X_{jpt-1}	0.0333***	0.0319***	0.0266***
	[0.0106]	[0.0072]	[0.0085]
R^2	0.076	0.078	0.132
F-first stage	15.67	29.44	23.87
Observations	2693	2997	3121
first stage			
$D^2_{\Delta \ln X_{pjt-1}}$	0.9949***	0.9403***	1.0458***
	[0.2040]	[0.1961]	[0.1938]
$D^3_{\Delta \ln X_{pjt-1}}$	0.6103***	0.6773***	0.9425***
	[0.2079]	[0.1960]	[0.1911]
$D^4_{\Delta \ln X_{pjt-1}}$	-0.3392	-0.8220***	-0.3071
	[0.2255]	[0.207]	[0.2011]

Notes: Robust standard errors in brackets. Each regression includes 2-digit sector fix effects. Significance: * 10%, ** 5%, *** 1%.

4.1 Robustness

In check robustness of our estimates in four different exercises. The first row in table 5 also includes the geographical dummy to take into account regional heterogeneity. Results holds by and large unchanged. The estimated effect of export on patenting is not therefore generated by the fact that norther regions record the bulk of both exporting (75% of total Italian exports) and patenting (**[xx%]** of Italian patents) activities.

Italian exports are not only geographically concentrated in the North-

Table 5: IV estimates: Robustness checks

	2001	2002	2003
Controls			
X_{jpt-1}	0.0355**	0.0312***	0.0233*
	[0.0175]	[0.0116]	[0.0126]
Excluding superstars			
X_{jpt-1}	0.0370**	0.0324***	0.0202*
	[0.0169]	[0.0111]	[0.0119]
Alternative treatment: Export share			
$\{\text{export share}\}_{jpt-1}$	0.5677*	0.4786**	0.3569**
	[0.3045]	[0.1874]	[0.1820]
Alternative treatment: Export dummy			
$D\{X_{jpt-1} > 0\}$	0.2324*	0.2892***	0.1943*
	[0.1326]	[0.1103]	[0.1100]

Notes: Robust standard errors in brackets. Each regression includes 2-digit sector fix effects. Significance: * 10%, ** 5%, *** 1%.

ern regions, but as documented in most countries (Eaton et al., 2004) **[...aggiungere altre citazioni]** they are also mainly due to a subset of Italian manufacturing firms. These superstar exporters are large firms that might patent larger amount of innovations independently of the complementarity hypothesis we are testing in this paper. For this reason, in the second row of table 5 we exclude superstar innovators from this robustness check and we check that our estimates hold virtually unchanged.[4]

The last two robustness checks are concerned with changes in the treatment, that is in replacing our endogenous variable $\ln(exports)$ with the export share, that is **[...]**, or a binary variable taking value equal to 1 when exports are positive. The third and the fourth row in table 5 show that, although the magnitude is expectedly different, the qualitative result holds true.

4.2 Heteregenous Effects

Having verified that export activity has a causal effect on patenting, in this section we check whether this effect is heterogenous across firms. We look into two dimensions of heterogeneity, namely productivity and firm's size. To obviate to the likely endogeneity of both measures, we use sample splits

[4]Excluding superstars also meet another concern. If a firms is both a patent and an export superstar ... international demand might not be exogenous to its export performance ... **Giacomo's point on the exogeneity of our instrument**.

above and below the median of the distributions of sales per employee and employment which are predetermined with respect to the treatment, i.e. in 2000, whereas our oldest treatment is exports in 2001.

Table 6 show that the effect of export on patenting is totally driven by smaller and less productive firms. As expected from Lileeva and Trefler (2010) smaller and less productive firms are the one for which technological change is most likely to show complementarity with an expansion of market size induced by foreign demand.

Table 6: IV estimates: heterogeneity

	$\left[\frac{sales}{empl}\right]_{2000}$		$empl_{2000}$	
	<median	>median	< median	>median
X_{jpt-1}	0.0249*	0.0074	0.0399*	0.0122
	[0.0146]	[0.0553]	[0.0238]	[0.0280]
R^2	0.163	0.204	-0.081	0.221
F first step	7.62	1.04	2.60	4.77
Observations	1001	1011	997	1015

Robust standard errors in brackets. Each regression includes 2-digit sector fix effects. Significance: * 10%, ** 5%, *** 1%.

5 Concluding remarks

References

Angrist, J. D. and Imbens, G. W. (1995). Identification and estimation of local average treatment effects. NBER Technical Working Papers 0118, National Bureau of Economic Research, Inc.

Aw, B. Y., Roberts, M. J., and Xu, D. Y. (2011). R&d investment, exporting, and productivity dynamics. *American Economic Review*, 101(4):1312–44.

Bound, J., Jaeger, D., and Baker, R. (1995). Problems with instrumental variables estimation when the correlation between instruments and the endogenous explanatory variables is weak. *Journal of American Statistical Association*, 90(430):443–450.

Bustos, P. (2011). Trade liberalization, exports, and technology upgrading: Evidence on the impact of mercosur on argentinian firms. *American Economic Review*, 101:304–340.

Coe, D. T. and Helpman, E. (1995). International r&d spillovers. *European Economic Review*, 39(5):859–887.

De Loecker, J. (2007). Do exports generate higher productivity? evidence from slovenia. *Journal of International Economics*, 73(1):69–98.

Eaton, J., Kortum, S., and Kramarz, F. (2004). Dissecting trade: Firms, industries, and export destinations. *American Economic Review*, 94(2):150–154.

Griliches, Z. (2000). *R & D, education, and productivity*. Loeb classical library. Harvard University Press.

Lileeva, A. and Trefler, D. (2010). Improved access to foreign markets raises plant-level productivity ...for some plants. *The Quarterly Journal of Economics*, 125(3):1051–1099.

Marin, G. (2011). Matching of patstat applications to aida firms. discussion of the methodology and results. Working papers, IMT Lucca.

Melitz, M. (2003). The impact of trade on intra-industry reallocations and aggregate industry productivity. *Econometrica*, 71(6):1695–1725.

Reinganum, J. (1983). Uncertain innovation and the persistence of monopoly. *American Economic Review*, 73:741–48.

Reinganum, J. (1984). Uncertain innovation and the persistence of monopoly: Reply. *American Economic Review*, 74:243–246.

Reinganum, J. (1986). Innovation and industry evolution. *Quarterly Journal of Economics*, 100:81–100.

Rodrik, D. (1988). Closing the technology gap: Does trade liberalization really help? NBER Working Papers 2654, National Bureau of Economic Research, Inc.

Yeaple, S. (2005). A simple model of firm heterogeneity, international trade, and wages. *Journal of international Economics*, 65(1):1–20.

Appendix 1: