



'R&D and export performance: exploring heterogeneity along the export intensity distribution'

L. Benfratello^{1,2} · A. Bottasso³  · C. Piccardo⁴

Received: 31 August 2021 / Revised: 23 January 2022 / Accepted: 24 January 2022
© The Author(s) 2022

Abstract

This study analyses the relationship between firm-level innovative effort as measured by R&D expenditures and export intensity. We apply quantile regression techniques to a sample of Italian firms to verify whether R&D expenditures' effect varies along the conditional distribution of export intensity, after controlling for censoring and endogeneity issues. Empirical findings suggest that the effect of R&D expenditures on export intensity is positive and that firms taking most advantage from R&D activity are in the right tail of the export intensity distribution (from the 70th quantile onwards), that is, those exporting 50% of their sales or more. Overall, the results prove robust to several specification checks and suggest not only that firms' innovative efforts help explaining heterogeneity in export intensity performance, but also that its positive effect differs across the export to sales ratio distribution. This implies that innovation policy measures might be more effective for firms characterised by a relatively high export intensive margin.

Keywords Exports · R&D · Quantile regression

JEL Classification F14 · O32 · D22 · C31 · C36

A previous version of this paper has been presented at the following conferences: Italian Industrial Economics (Naples), Italian Economic Society (Bologna), Italian Econometric Association (Salerno), European Trade Study Group (Munich) and European Association for Research Industrial Economics (Maastricht). We thank those audiences as well as Pierluigi Murro and Enrico Rettore for helpful suggestions. Usual disclaimers apply.

✉ A. Bottasso
bottasso@economia.unige.it

¹ Polytechnic of Torino, Torino, Italy

² CSEF, University of Naples, Naples, Italy

³ Department of Economics, University of Genova, Via Vivaldi 5, 16125 Genova, Italy

⁴ University of Verona, Verona, Italy

1 Introduction

During the last two decades, vast amounts of empirical studies have analysed the relationship between firm characteristics and export activity. The different approaches in the empirical literature reflect the evolution of theoretical models that have gradually improved the explanation of the complex interactions between firm heterogeneity and participation in international markets.

One of the issues that has attracted increasing attention is the relationship between firms' innovative capacity and export activity. As suggested by Melitz and Redding (2014), comprehension of such relationships not only helps explain some empirical regularities observed in trade data (see, among others, Love & Roper, 2015) but also plays a major role in theoretical models linking growth to trade (e.g. Aghion et al. 2018). Based on the literature, the analysis of the export-innovation relationship must consider that productivity-enhancing investments, such as R&D expenditures, are endogenous decisions jointly made with trade participation. In these models, innovation investments can induce productivity improvements that allow firms to afford the costs associated with exporting activity and enable firms to achieve a greater ability to meet international market demand, thus making exporting more profitable. Moreover, participation into international markets can further stimulate innovative activity by inducing higher competitive pressure on firms and favouring technological transfer from destination markets.

This study aims to shed additional light on the effects of firms' innovative effort on their export intensity ('intensive margin' measured as the ratio of firm's export sales on total sales) by taking into account endogeneity issues and controlling for other firm characteristics that might favour the exporting activity. Particularly, we focus on firms' export intensity as a continuous measure of exporting activity, allowing us to apply a quantile analysis approach.

We believe that such an issue deserves priority given the role played by exporters in the economy. Indeed, empirical literature suggests that firms active in international markets perform better. Moreover, they strengthen over time, become more resilient to negative shocks, and are more capable of keeping up with external demand fluctuations (Bugamelli et al., 2018, among others). For this reason, investigating the factors favouring firms' exporting performance is crucial to better understanding the complexity of growth processes and countries' resilience capacity.

The main novelty of this study is that it investigates how the relationship between innovation and exports varies along the conditional distribution of export intensity. Since innovation is more valuable for exporters as it generates additional profits from export sales, a natural question is to investigate whether innovation rents are different along the export intensity distribution. Indeed, empirical evidence on this issue is scant. A similar research question was analysed by Wagner (2006), who applies quantile regression techniques to a panel of German plants, and by Orts and Martí (2018) on a European firms sample. However, these studies do not address the issue of endogeneity of firm innovative effort and exclude non-exporting firms from the analysis, thereby avoiding censoring issues. In our study, we address these methodological issues and apply the conditional quantile instrumental variable

(CQIV) estimator proposed by Chernozhukov et al. (2015) to a sample of Italian manufacturing firms contained in the Unicredit-Capitalia survey and observed in the 1992–2003 period. This approach has never been adopted in the literature. Regarding the choice of the instrument, we follow a widely used approach in applied industrial organisations and use the industry value of R&D expenditures in Spain as instruments for the R&D variable.

All estimated models confirm the positive impact of R&D expenditures on export intensity and, more importantly for the purposes of this study, suggest that such an effect displays significant heterogeneity along the export intensity conditional distribution. Particularly, our estimates show that the effects of R&D increase along the conditional distribution of export intensity, reach a maximum at the 70th quantile, and persist at higher quantiles. Such results imply that firms characterised by an export intensity of about 50% or more can take the highest advantages, in terms of further expansion of their sales in international markets, from investing in R&D activity.

The heterogeneous effect of R&D expenditures along the export intensity distribution might be related to an increase in firm competitiveness, spurred by innovations generated by R&D activity and the resulting possibility of entering new foreign markets. Furthermore, we might argue that firms with low export intensity do not devote large resources, particularly human resources, to export activity and/or are passive exporters, so they have lower returns from innovative effort.

This study contributes to the exporters' heterogeneity trade literature by providing new empirical evidence for the Italian case. Unlike most of the literature focusing on heterogeneity in firms' observed characteristics, our study highlights another dimension of heterogeneity, namely, the effect of some covariates. If some differentials are observed, we believe such information might be useful for designing policy interventions aimed at favouring exporting activity and productivity-enhancing policies at the micro level.

A possible drawback of this study is related to the available sample period, which is not so recent and does not reflect important events such as the great recession and recent world trade trends. However, the external validity of our analysis rests on the similarities of our sample (considering sectoral and size distributions and the presence of exporting and innovative firms) with the sample of Italian firms included in the more recent EFIGE dataset. Indeed, as discussed in Sect. 5, our results are consistent with those obtained by Orts and Marti (2018) on the EFIGE data.

The remainder of this paper is organised as follows. The next section synthesises the theoretical background and evolution of the empirical literature addressing the relationship between innovation and exports. Section 3 describes the data, and Sect. 4 illustrates our empirical strategy. We discuss the empirical results in Sect. 5 and we provide our concluding remarks in Sect. 6. An Appendix containing additional results concludes the paper.

2 Related literature

The relationship between innovation and exporting activity has traditionally been analysed within the frameworks of trade theory and growth theory. The recent trade literature (surveyed by Melitz & Redding, 2014) considers empirical regularities observed in trade data and proposes theoretical models based on firm heterogeneity in differentiated product markets, able to explain different mechanisms whereby trade liberalization affects the economy, including an endogenous increase of firm productivity. These models explain observed heterogeneity between exporters and non-exporters relative to performance indicators, different patterns of trade participation across firms and destination markets, and within-industry reallocations of resources following trade liberalization.

Within this literature, some models are particularly relevant for our study because they include the impact of the trading environment and trade participation on firms' decisions relative to productivity-enhancing investments, such as R&D expenditures (e.g. Aw et al., 2011; Bustos, 2011; Lileeva & Trefler, 2010). In such a framework, trade liberalisation can stimulate innovation and adoption of upgraded technology so that firm-level productivity becomes the outcome of several endogenous decisions jointly taken with trade participation. In these models, trade liberalisation has different impacts on firms; most productive firms will decide to both innovate and export, lower productive firms only export, firms with lower productivity neither innovate nor export, and the least productive firms exit. More recently, Bonfiglioli et al. (2018) developed and empirically validated on US firm-level data, a theoretical model that includes fixed and variable export costs; firm heterogeneity derives from technological choices; particularly, firms can select the size of their investment in projects of unknown quality, thus affecting the variance of the probability distribution from which firms' productivity is drawn. The model reveals that only the most productive firms export and that export opportunities raise the return to technological heterogeneity by reallocating profits in favour of the most productive firms, so that firms have higher incentives to invest in more ambitious innovation projects (with more dispersed outcomes).

The relationship between innovation and exports is also related to the endogenous growth literature linking trade and growth. A recent survey by Grossman and Helpman (2015) describes the main channels relating trade and growth, namely, knowledge spillovers, market dimension and competitive intensity, relative prices, and technology dissemination.

In such models, trade liberalisation generates an expansion of market size, thus increasing innovation incentives (market size effect). However, this effect can be counterbalanced by a competition effect as exporters are subject to a more intense competition that might reduce incentives to innovate by lowering expected innovation rents. Indeed, participation in international trade might stimulate innovation through different channels: enlargement of relevant markets that entail productivity improvements, requirement of catching the international technological frontier, exploitation of scale economies allowing firms to sustain fixed costs associated with R&D and innovative activities, and technological transfer from destination markets

(learning-by-exporting effect).¹ Particularly, Aghion et al. (2018) suggests that positive demand shock for exporters can induce higher innovation especially for firms closer to the frontier and that it may reduce innovation for firms far from the frontier. Authors test the predictions of their model on a sample of French firm-level data observed over the period 1994–2012.

Impullitti and Licandro (2018) develop a theoretical model with heterogeneous firms operating in an oligopolistic market and innovation-driven productivity growth. They consider three potential sources of gain from trade. First, through higher product market competition, the pro-competitive effect of trade leads to lower markups in the domestic market. Second, the selection effect of trade forces the least productive firms out of the market and prompts firms to enter the export market. Both the pro-competitive and selection effects reallocate resources to most productive firms, thereby raising the average aggregate level of productivity and lowering equilibrium prices. Moreover, both these effects stimulate firms to innovate, thereby increasing productivity growth rate. The model identifies innovation as a key driver of dynamic gains from trade. Using firm-level and aggregate data on the US economy, they calibrate and numerically solve the model and demonstrate how the interaction between selection and innovation-driven growth can generate larger welfare gains from trade.

From an empirical point of view, firms' innovative ability has been included among those factors that explain observed heterogeneity in firms' participation in international markets, others being productivity, size, and age. We refer to Bernard et al. (2012) for a survey of the related empirical evidence.²

The international evidence supports the 'self-selection hypothesis', that is, a positive causal effect of firms' innovative capacity on exporting activity, both on the extensive (i.e. exporting probability) and intensive (i.e. export intensity) margins. These findings are robust to the adoption of different innovation activity measures, as suggested in the literature. R&D expenditures or R&D employees, the share of workers with technical and scientific backgrounds, or the presence of joint R&D projects with external partners (Braymen et al., 2011) are usually employed to measure innovation inputs. Innovation output indicators include the number of product and/or process innovations and the number of patents (Cassiman & Golovko, 2011; Fernández-Mesa & Alegre, 2015; Monreal-Pérez et al., 2012). A positive impact of innovation on exporting margins has also been found on Italian data (e.g. Altomonte et al., 2013; Benfratello & Razzolini, 2008; D'Angelo, 2012; Dosi et al., 2015; Frazzoni et al., 2011).³ Recently, Brancati et al. (2018), using data on Italian

¹ Dhingra (2013) suggests a theoretical model where process and product innovations react differently to trade liberalization. The model predicts that a bilateral tariff reduction enlarges relevant markets available to new and continuing exporters, thus inducing an increase in process innovation. Particularly, more productive exporters benefit the most from the expansion of the market size and manage to overcome greater competition by expanding their product lines through higher product and process innovation. Non-exporting firms face import competition from new foreign brands and reduce their product lines.

² A recent strand of literature pioneered by Greenaway et al (2007) has investigated the effect of financial constraints on export activity. For applications to the Italian case see Secchi et al., (2016a, 2016b).

³ For an exhaustive survey on the empirical literature for Italy, see Bottasso and Piccardo (2013).

manufacturing and services sectors observed over the period 2007–2013 (MET dataset), analysed the determinants of firms' exporting activity. The authors found that firms' innovative efforts, as measured by product innovation and R&D expenditures, past trade experience in foreign markets, and productivity, are the main determinants of firms' decision to export. However, only past trade experience seems to have significant impact on export intensity. Slightly different results are provided by Mancusi et al. (2018), who empirically confirm the self-selection hypothesis for a sample of Italian firms observed over the period 2004–2009 (Unicredit-Capitalia and EFIGE datasets). Particularly, the authors demonstrate that product innovation is one of the main factors determining both extensive and intensive export margins.

From the analysis of the applied literature that investigates firms' exporting probability determinants, some empirical regularities emerge: firms exhibiting higher productivity and a higher propensity to innovate are more likely to become exporters.

Overall findings on the heterogeneity of (potential) exporters with respect to non-exporters are robust to using different samples and econometric techniques. Most studies demonstrate that size has a nonlinear effect on export probability and that credit constraints do matter for internationalisation choices. Evidence on other control variables commonly included in empirical models, such as firm age, ownership structure, labour cost, workforce composition, and geographical location, is less conclusive and not sufficiently robust. Overall, empirical findings on Italian manufacturing firms are consistent with those obtained in international literature. Furthermore, the applied literature analysing firms' export intensity determinants has univocally confirmed that factors affecting the extensive margin are key in determining the intensive margin.

A related branch of the empirical literature investigates the impact of the exporting activity on innovative capacity, i.e. the so-called learning-by-exporting hypothesis predicting that firms can improve their performance by participating in international markets. Particularly, some studies have identified a positive effect of exporting activity on introducing product innovations (e.g. Aghion et al., 2018; Damijan et al., 2010; Van Beveren & Vandenbussche, 2010). Other studies show that firms increase their R&D activities and upgrade their technologies when they compete in international markets (e.g. Baum et al., 2016; Criscuolo et al., 2010; Wagner, 2012; Zhou & Song, 2016).

Turning to empirical evidence for Italy, Bratti and Felice (2012) test whether exporting activity improves firm innovativeness on the VIII and IX waves of the Unicredit-Capitalia database. They find that firm internationalisation seems to boost the probability of introducing product innovations after controlling for reverse causality issues. More recently, Fassio (2018) focused on analysing the learning-by-exporting mechanism in a sample of European firms (EFIGE). Stressing the role played by export destinations, the author finds that a higher level of technological knowledge in export destinations causes exporting firms to introduce product innovations (the

technological learning effect) while a foreign higher demand growth mainly induces exporters to develop process innovations (the foreign demand effect).⁴

Overall, the results stemming from the surveyed literature suggest that exporting firms with higher productivity levels are more likely to introduce product innovations and produce high-quality products. Such results have been obtained after accounting for the effect of a wide range of firm- and industry-level control variables such as size, workforce composition, ownership structure, age, affiliation to a group, geographical location, sector, and destination country characteristics.

In summary, the empirical literature has highlighted how firms' innovative capacity is closely related to firm performance in foreign markets. Indeed, innovation is more beneficial for exporters because it will generate additional profits from export sales. However, a dimension that remains almost unexplored is whether the relationship between innovative capacity and export activity is heterogeneous according to different levels of involvement in international activity.

To the best of our knowledge, only two studies have analysed this issue. Wagner (2006), using a sample of German exporting plants, estimates a model wherein export intensity depends on size, its square, the subsidiary branch plant status of the establishment, a dummy variable that identifies whether a firm operates in the crafts sector, the workforce composition, three dummies for different classes of R&D intensity, and a binary indicator for patent registration. The author finds that, for medium levels of R&D intensity (between 3.5% and 8.5%), the impact of R&D is significant and similar across different levels of export intensity distribution. In contrast, for low levels of R&D intensity (less than 3.5%), the impact is significantly positive only for the highest quantiles (75th and 90th) of export intensity. No significant effect of R&D on export intensity was found for high levels of R&D intensity (higher than 8.5%).

Similarly, Orts and Martí (2018) analyse the effect of different firm characteristics on export intensity and market scope (i.e. the number of markets to which the firm exports) on the EFIGE dataset with a quantile approach. The authors estimate a model wherein export performance is modelled as a function of firm productivity, size, capital intensity, R&D intensity, import intensity and a series of binary indicators for foreign affiliates and product and process innovations, respectively.⁵ The authors find that the positive effect of R&D intensity increases along the export intensity distribution, reaches its maximum around the 70th quantile, and is not statistically significant at the highest quantiles.

⁴ See Love and Roper (2015) for an exhaustive survey of the empirical literature on innovation, exporting activity, and growth for small and medium firms.

⁵ A few studies (e.g. Arnold and Hussinger (2010), Haller (2012), Powell and Wagner (2014)) have used quantile techniques to evaluate the impact of the exporting activity along firm productivity distribution. In particular, Powell and Wagner (2014) use a panel data quantile technique allowing to control for firm-level fixed effect and find that the exporter premium is positive over the whole productivity distribution, with the largest premium at the bottom of the distribution, for a panel of German manufacturing firms.

3 Data

The empirical analysis is based on a dataset of Italian manufacturing firms obtained by merging the VI, VII, VIII, and IX waves of the ‘Indagine sulle imprese manifatturiere italiane’ (Survey on Italian Manufacturing Firms) run every 3 years by the Unicredit-Capitalia Observatory of Medium and Small Firms. The sampling scheme was the same for all four waves. All firms above the 500-employee threshold are included whereas firms with a number of employees between 10 and 500 are included according to a stratified sampling strategy. Strata are then based on size class, geographical area, and industry (according to the Pavitt taxonomy) to represent the population of Italian manufacturing firms significantly.

Pavitt (1984) suggested a classification of industries based on innovation-related characteristics (e.g. product and process innovation, sources of knowledge, appropriability regimes), firm size, and competitive factors. Pavitt identifies four groups of sectors: (i) *the supplier-dominated* sector, which is the most traditional manufacturing industries relying on sources of innovation external to the firm; (ii) *the scale-intensive* sector, mainly characterised by large firms producing basic materials and consumer durables for which sources of innovation may be both internal and external to the firm with a medium level of appropriability; (iii) *the specialised suppliers* sector, composed of small, more specialised firms producing technology to be sold to other firms for which the level of appropriability is high; and (iv) *the science-based* sector with high-tech firms relying on R&D from both in-house sources and university research and developing new products or processes with a high degree of appropriability.

The four waves cover the periods 1992–1994, 1995–1997, 1998–2000, and 2001–2003, respectively, and provide qualitative and quantitative information concerning several firm-level characteristics such as ownership structure, workforce composition, internationalisation, and innovation activities. Some variables have annual frequency while others refer to the last year of each wave, and still others cover the 3-year time span.

This dataset (or portions of it) has been extensively used by researchers in the internationalisation and innovation fields. For instance, the merged VIII and IX waves have been used, along with the study by Bratti and Felice (2012), by Antonelli and Crespi (2013) to examine the allocation of public subsidies to R&D activities and by Antonietti and Cainelli (2011) to estimate a structural model of R&D, innovation, productivity, and export.

Survey data were integrated with balance sheet information derived from the AIDA repository, a database elaborated by Bureau Van Dijk. Although this dataset admittedly currently covers a somewhat distant period, it still represents an extremely valuable source of information, especially for analysing the relationships between exports and innovation. In fact, it contains several measures of both different innovation input variables (e.g. formal R&D expenditures, number of employees involved in R&D) and innovation outputs (i.e. process and product innovation), and a wide array of firm-level control variables, which are extremely difficult to find in a single dataset.

The combined dataset obtained by merging the four waves of the survey is composed of 12,679 firms; unfortunately, only a very small fraction was observed in all four waves (1.2%), and a major change in the sample occurred between the VII and VIII waves. In turn, most of the firms appear in only one wave (68.85%), and even when firms appear twice, they might appear in two non-contiguous waves. Furthermore, in the VIII wave, the question about the share of exported turnover, our dependent variable, has unfortunately been omitted from the questionnaire.

Because of this peculiar structure of the original dataset, we prefer using the subsample composed of those firms observed either in both the VI and VII waves or in both the VIII and IX waves. The use of firms present in two contiguous waves comes at the cost of not being able to control for time-invariant firm-specific individual effects; however, it allows us to use 3-year lags of the independent variables likely to exert their effect on exports after a sufficiently long period. Therefore, we apply cross-sectional techniques to a sample of 1,029 firms with current variables observed in 1997 or 2003 and lagged variables observed in 1994 or 2000 (or over the 1992–1994 or 1998–2003 periods).⁶

Table 1 presents the sectoral distribution of the original and final datasets across the manufacturing industries (2 digits Nace rev. 1 classification). Firms in the leather, electrical machinery, and medical and precision instruments industries (industry codes 19, 31, and 33) are underrepresented in the final dataset whereas the opposite occurs in the machinery industry (code 29). As for the extensive margin of export, that is, the share of exporting firms, the original and final samples differ only for the chemical and basic metal industries (codes 24 and 27), wherein exporting firms are overrepresented, and the electrical machinery industry (code 31), wherein they are underrepresented. These differences are mostly due to the unavailability of the instrument, constructed at the industry/size level (see later), for some industry/size class combinations and the increasing relationship between propensity to export and size. Indeed, the instrument is unavailable for the entire leather industry; for medium and large firms in the industry codes 22, 24, 27, 31, and 33; and for small firms in industry code 30. To avoid imputation of an instrument retrieved from a contiguous class size or industry, we excluded those observations.⁷ Notwithstanding these differences, the two datasets do not seem to differ dramatically, so the final sample reasonably represents the sectoral composition of the Italian manufacturing sector well.⁸

Table 2 shows the distribution of the sample concerning both the firms' Pavitt classification and export status. Such classification groups firms based on their

⁶ The 62 firms appearing in all four waves of the survey have been included only in the 1998–2003 period to avoid the need for standard errors clustering for these firms only.

⁷ As a robustness check, we estimate the model with the imputed instrument and results, available upon request to the authors, are very similar to those presented in the study.

⁸ The final sample does not significantly differ from the original one also in terms of geographical, foreign ownership and size distributions. For sake of space we do not report such analysis that is available upon request.

Table 1 Firms sectoral distribution

Industry (2 digits Nace Rev 1)	Original sample			Final sample				
	Number of firms	Percent	Number of exp	Share of exp	Number of firms	Percent	Number of exp	Share of exp
15—Food products and beverages	1131	8.95	734	64.90	91	8.84	53	58.24
16—Tobacco products	4	0.03	3	75.00	0	0	0	0
17—Textiles	1098	8.66	827	75.32	128	12.44	101	78.91
18—Wearing apparel	558	4.40	409	73.30	27	2.62	22	81.48
19—Leather	555	4.38	479	86.31	0	0	0	0
20—Wood and wood products	298	2.35	169	56.71	38	3.69	24	63.16
21—Pulp, paper and paper products	363	2.86	242	66.67	43	4.18	29	67.44
22—Publishing and printing	462	3.64	208	45.02	29	2.82	14	48.28
23—Coke and refined petroleum products	51	0.40	20	39.22	0	0	0	0
24—Chemicals and chemical products	622	4.91	494	79.42	28	2.72	13	46.43
25—Rubber and plastic products	705	5.56	558	79.15	86	8.36	76	88.37
26—Other non-metallic mineral products	781	6.16	425	54.42	74	7.19	38	51.35
27—Basic metals	494	3.90	355	71.86	19	1.85	10	52.63
28—Abricated metal products, except machinery and equipment	1525	12.03	931	61.05	118	11.47	85	72.03
29—Machinery and equipment n.e.c	1789	14.11	1552	86.75	205	19.92	192	93.66
30—Office machinery and computers	67	0.53	37	55.22	0	0	0	0
31—Electrical machinery and apparatus n.e.c	506	3.99	355	70.16	14	1.36	13	92.86
32—Radio, television and communication equipment and apparatus	256	2.02	191	74.61	21	2.04	18	85.71
33—Medical, precision and optical instruments	261	2.06	201	77.01	6	0.58	4	66.67
34—Motor vehicles, trailers and semi-trailers	260	2.05	197	75.77	34	3.30	25	73.53
35—Other transport equipment	134	1.06	104	77.61	10	0.97	8	80.00
36—Manufacture of furniture; manufacturing n.e.c	759	5.99	630	83.00	58	5.64	49	84.48

Table 1 (continued)

Industry (2 digits Nace Rev 1)	Original sample			Final sample		
	Number of firms	Percent	Share of exp	Number of firms	Percent	Share of exp
Total	12,679	100.0	71.94	1029	100.0	75.22

Table 2 Number of observations

Pavitt taxonomy	Number of firms	Share of exporting firms
Supply dominated	439	73.8
Scale intensive	270	62.6
Specialised supplier	280	90
Science based	40	80
Total	1029	75.2

technological competence and has often been adopted in many empirical studies on innovation, trade, and competitiveness at both the macro and micro levels.⁹

Supplier-dominated industries include textiles, footwear, food and beverages, paper and printing, and wood while the *scale-intensive* group includes basic metals, motor vehicles, trailers, and semi-trailers. *Specialised suppliers* include machinery and equipment; office, accounting, and computing machinery; and medical, precision, and optical instruments. Finally, *science-based* industries consist of chemicals, pharmaceuticals, and electronics.

Table 2 shows that most firms operate in *supply-dominated* industries (about 42%); firms operating in *science-based* industries represent a small fraction of the sample (only about 4%) whereas less than 30% of firms operate in each of the other two sectors. Overall, 75% of firms are engaged in exporting activity. However, the Pavitt sector stratification reveals that such percentage is higher for firms operating in *specialised suppliers* and *science-based* industries (about 90% and 80% of firms export, respectively).

Table 3 shows some descriptive statistics for the whole sample for subsamples of exporting and non-exporting firms and for different ranges of the distribution of export intensity. Export intensity (*expint*), defined as the ratio of firms' export sales to total sales, is on average about 40% for exporters and displays a high level of dispersion around the mean. For almost half of the exporters (about 41%), sales derived from exporting activity are below 30% of total sales. In contrast, the remaining firms are uniformly distributed across higher values of export intensity. Therefore, the export intensity distribution appears to be skewed to the right and, hence, not to be normally distributed. This shape of the export intensity distribution is confirmed by both graphical methods and statistical tests.¹⁰

Firms' innovative efforts are measured by the expenditures on R&D activity (*R&D*). Unconditional to export status, around 44% of firms in our sample were involved in formal R&D activities. However, striking differences emerge between exporting and non-exporting firms. More than half of the exporting firms (51%)

⁹ We use the Pavitt classification industry dummies as our econometric estimates suffer from convergence problems when the large set of two-digit industry dummies (Nace rev. 1) is included.

¹⁰ The skewness (0.63) and kurtosis (2.1) values confirm that the export intensity distribution is not normal, and it is skewed to the right. Statistical tests (Skewness and Kurtosis, Shapiro–Wilk, and Shapiro–Francia tests for normality) formally support such evidence.

Table 3 Descriptive statistics for different ranges of export intensity

	Export firms	Non-export firms	Whole sample	0q-30q <i>exp</i> ≤ 3%	30q-50q 3% < <i>exp</i> ≤ 20%	50q-70q 20% < <i>exp</i> ≤ 50%	70q-100q <i>Exp</i> > 50%
<i>Expint</i>	40.09 (28.24)	0.00 (0.00)	30.16 (29.99)	0.37 (0.84)	12.20 (5.75)	37.34 (8.58)	72.34 (13.66)
<i>Size</i>	0.162 (0.399)	0.040 (0.058)	0.131 (0.351)	0.054 (0.209)	0.112 (0.481)	0.171 (0.248)	0.203 (0.420)
<i>R&D</i>	306.38 (1568.77)	17.46 (75.60)	234.78 (1366.59)	29.27 (193.57)	98.34 (606.76)	345.63 (2191.68)	483.455 (1591.98)
<i>Supply dominated</i>	0.41 (0.49)	0.45 (0.49)	0.42 (0.49)	0.44 (0.49)	0.54 (0.49)	0.38 (0.48)	0.35 (0.47)
<i>Scale intensive</i>	0.21 (0.441)	0.39 (0.49)	0.26 (0.44)	0.39 (0.49)	0.24 (0.42)	0.24 (0.43)	0.13 (0.33)
<i>Specialised suppl</i>	0.32 (0.46)	0.12 (0.32)	0.27 (0.44)	0.12 (0.33)	0.19 (0.39)	0.32 (0.46)	0.46 (0.49)
<i>Science based</i>	0.04 (0.19)	0.03 (0.17)	0.03 (0.19)	0.03 (0.17)	0.014 (0.1)	0.05 (0.22)	0.05 (0.22)
<i>North</i>	0.79 (0.40)	0.61 (0.48)	0.75 (0.43)	0.63 (0.48)	0.72 (0.44)	0.81 (0.38)	0.85 (0.35)
<i>Centre</i>	0.15 (0.36)	0.28 (0.45)	0.18 (0.39)	0.26 (0.44)	0.19 (0.39)	0.14 (0.35)	0.12 (0.33)
<i>South</i>	0.04 (0.20)	0.09 (0.29)	0.05 (0.23)	0.09 (0.29)	0.08 (0.27)	0.03 (0.18)	0.02 (0.13)

Standard deviations in parenthesis. *Expint* is the ratio of firms export sales on total sales (%), *Size* is the number of employees (in thousands), *R&D* is the expenditures in R&D activity (in thousands of Euro). All others variables are binary indicators

performed R&D activities whereas only about 18% of non-exporting firms did. Furthermore, average expenditures in R&D activity is much higher for exporters than for non-exporters and significantly increases along export intensity distribution.¹¹

The average size, as measured by the number of employees (*size*), is approximately 130, as expected, and size distribution is very asymmetric with about 55% of firms classified as small (i.e. less than 50 employees) and a lower fraction (about 13%) defined as large (i.e. more than 250 employees).

On average, exporting firms are significantly larger than non-exporting firms (162 vs. 40 employees), and size is positively correlated with export intensity, so larger firms are those exhibiting higher shares of sales derived from exporting activity. Area dummies (*North*, *Centre*, *South*) show that most firms (75%) are located in Northern Italy (Emilia-Romagna, Friuli-Venezia Giulia, Liguria, Lombardia, Piemonte, Trentino-Alto Adige, Valle D'Aosta, and Veneto), and 18% are located in the Centre (Abruzzo, Marche, Toscana, Lazio, and Umbria). When considering exporting firms, firms in the North appear to be more export-oriented compared to firms in the rest of the country, both on the intensive and extensive margins. The sectoral and size distribution as well as the firm-level characteristics in our final sample are consistent with the evidence of the Italian section of the more recent EFIGE survey covering 2007–2009 (see Altomonte et al., 2012; Altomonte & Aquilante, 2012).¹² Particularly, the share of exporters is around 72%, and 54% of firms have workers directly involved in R&D activities. Furthermore, in the EFIGE dataset, the share of exporting firms increases with size. A minor difference emerges in the intensive margin of exports, which is slightly lower for the EFIGE dataset. The overall similarity of our dataset with that of EFIGE is reassuring in terms of the representativeness of our dataset regarding R&D and export activity.

Pavitt industry dummies suggest that firms operating in *specialised suppliers* and *science-based* industries show higher export propensity and intensity. Particularly, the share of firms belonging to *specialised suppliers* and *science-based (supply dominated and scale intensive)* industries increases (decreases) along the export intensity distribution.

4 Empirical strategy

This study aims to investigate whether innovative activities and other export intensity determinants differently affect export intensity at various points in its distribution by applying quantile regression techniques. As previously mentioned, such an approach has been adopted only by Wagner (2006) and Orts and Martí (2018).

¹¹ R&D expenses are deflated with a weighted average of the consumer price index (0.8) and the aggregate business investment price index (0.2). As for the R&D expenditures over total sales ratio, exporting firms show an R&D intensity of about 0.82% while non-exporters invest a much lower percentage (about 0.28%). Moreover, the R&D intensity increases along the export intensity distribution, ranging from 0.3% for the class 0–3% of export intensity to 1.1% for the class of export intensity higher than 50%.

¹² The EFIGE (EU-EFIGE/Bruegel-Unicredit) dataset was collected in 2010 within the 'European Firms in a Global Economy: internal policies for external competitiveness' project.

However, the standard conditional quantile regression techniques adopted in these studies do not account for either the censored nature of our dependent variable or the possible endogeneity of the regressors. Therefore, we adopted a more comprehensive estimation approach.

To assess the relationship between firms' innovative effort and export intensity (firms' export sales on total sales, in percentage), along with the effects of other firm characteristics on their export performance, we consider the following empirical model:

$$\text{Expint}_i = \beta_0 + \beta_1 \text{Innov}_i + \beta_2 X_i + u_i \quad (1)$$

where subscript i denotes firms.

Firms' innovative effort (*Innov*) is measured by logarithmic transformation of firms' R&D expenditures, which captures the existence of a system of incentives toward intentional innovative activities and can be considered a proxy for the allocation of resources to research and other information-generating activities in response to perceived profit opportunities (Grossman & Helpman, 1991).

Although the Unicredit-Capitalia survey reports information on the introduction of process and product innovation, we do not use these variables as export activity is often associated with organisational and technical changes that might be erroneously considered product or process innovation. In contrast, these variables represent product differentiation activity or simple process reorganisation and consist of introducing minor (technical or aesthetic) changes. Indeed, the Oslo Manual (2018), which provides guidelines for collecting and using innovation data, specifies that "An innovation is a new or improved product or process (or combination thereof) that differs significantly from the unit's previous products or processes and that has been made available to potential users (product) or brought into use by the unit (process)".

X_i is a vector of covariates including firm size and its square, a set of Pavitt dummies to include the omission of sector specific time invariant characteristics that might bias our parameter estimates, a year dummy accounting for common macroeconomic factors, and three area dummy variables indicating whether firms operate in the North, Centre, and South of Italy to control for time-invariant locational effects. Indeed, the main destination markets of Italian exports are located in Europe (e.g. Bugamelli et al., 2018); therefore, being located in an area closer to destination markets might favour exporting performance. Moreover, economic activities are not evenly distributed across different geographical areas. For example, industrial districts are more concentrated in the North and Centre of Italy; in turn, the exploitation of agglomeration economies might differ across areas and affect firms' export performance. A related issue is the heterogeneity in infrastructure endowments, quality of institutions, and economic growth that characterises the Italian economy (e.g. Banca d'Italia 2019).

We do not include productivity among covariates since R&D may be considered as one of its major drivers.¹³ Size is expected to exert a positive effect on export intensity. Given the existence of relevant sunk costs necessary for entering foreign markets, larger firms might be able to utilise economies of scale in production and might show a higher capacity to take risks and obtain credit at lower costs (Wagner, 1995). Such predictions stem from several theoretical models suggested by different trade theory models (Melitz & Redding, 2014).

As the relationship between size and export intensity might be nonlinear, we also include the square of the size variable. Wakelin (1998) suggests that an inverted U-shaped relationship may be associated with the existence of major firms with monopolistic power, which may show less motivation to export.

Both size and R&D might be endogenous given the possible existence of reverse causality and simultaneity issues, as discussed by recent theoretical models of trade (e.g. Aghion et al., 2018; Impullitti & Licandro, 2018; Melitz & Ottaviano, 2008; Melitz & Redding, 2014). In principle, identifying valid instruments for both variables would be appropriate. However, given that the CQIV estimator does not allow us to apply IV to more than one endogenous variable at a time, we choose to find a valid instrument for the R&D variable. Moreover, to weaken the possible inverse relationship with export intensity, we include 3-year lagged values of R&D (the last year of the previous waves). Indeed, it is reasonable to assume that R&D effort takes time to exert its effect on firms' export performance.¹⁴

With regard to the choice of the instrument, we choose to follow a widely used approach in applied Industrial Organization—one that uses the value the endogenous variable in a market different from the one under analysis. For a discussion on this approach, see Berry and Haile (2016). To this end, we retrieve the 1998–2000 values of industry R&D expenditures in Spain from the ‘third Community Innovation Survey’ (CIS III), which has been conducted over a sample of European (mostly EU) countries. Among countries participating in this survey, we select Spain as the country whose industrial structure is most similar to that of Italy. We match Italian and Spanish firms according to both industry and size. The rationale is that, on one hand, R&D activity is obviously strongly related to industry characteristics (technology, product differentiation, etc.) while, on the other hand, it is related to firm size. Thus, by adopting this approach, we are confident in increasing the correlation between the instrument and endogenous variable.

Particularly, the CIS III survey contains information on the industry wherein the firm operates (Nace rev.1 at the 2 digits level) and a discrete measure of size built

¹³ Indeed, as productivity is an outcome of R&D, the insertion of productivity leads to a “bad control” issue. Nonetheless, following a referee suggestion we re-estimated the model including (Total Factor) Productivity as regressor, both in substitution and alongside with R&D. Results, available upon request to the authors, show that TFP exerts a positive effect on export intensity for lowest quantiles but negative for highest ones. When included together with R&D, TFP somehow reduces the effect of R&D and it exerts a negative impact on export intensity along all the quantiles. These results have to be taken with caution given that we could not instrument TFP.

¹⁴ The approach of including lagged values for input or output innovation variables has frequently been adopted in prior literature. See, for example, Brancati et al. (2018) and the literature cited by the authors.

as a three-class variable (small, medium, large), which is defined according to the number of employees. The instrument for an Italian firm operating in a given industry and size is thus the mean value of the log R&D expenditures for Spanish firms of that size and operating in the same industry observed in 1998–2000.

The adoption of the CIS III survey is determined by the need to have the most similar time span to the one we consider. Unfortunately, the unavailability of the first two releases of the survey prevented us from using the years preceding our sample period. However, the relevance of the chosen instrument can be assessed by observing very high statistical significance of the instrument in first-stage regression (see below) whereas instrument exogeneity relies on the fact that a shock to exporting activity in Italy is very unlikely to have consequences for Spanish R&D expenditures. Moreover, we conduct IV estimates with an additional instrument, namely, the 1998–2000 values of industry R&D expenditures in Germany from the CIS III; thus, the over-identification of the model allows us to run a formal test of exogeneity.

Regarding the possible endogeneity of the size variable, given the impossibility of applying instrumental variable techniques to more than one endogenous variable at a time, we include size and its square in lagged form (3-year lag) to weaken the possible inverse relationship between size and export intensity given that lagged size might be less related to current export intensity values.

Another issue we need to control for is the censored nature of our dependent variable. Indeed, export intensity has a non-negligible mass at 0 due to the large fraction (23%) of non-exporting firms. To account for both endogeneity of R&D and censoring bias issues, we apply the CQIV estimator recently proposed by Chernozhukov et al. (2015). Note that this estimator allows for including only one continuous endogenous explanatory variable in the model. This also motivates the choice of our innovation activity variable. This estimator uses a control function approach by estimating the first stage for the endogenous regressor and the second stage, where the first-stage estimated residual is included as an additional regressor. Moreover, in the second stage, the probability of censoring is estimated for all observations, and standard quantile estimators for uncensored data are iteratively applied to the subset of observations for which the probability of censoring is sufficiently low (for technical details, see Chernozhukov et al., 2015, 2019).

For comparison purposes, we estimate different specifications of Eq. (1) with other quantile and mean estimation techniques that, in contrast to the CQIV estimator, only partially consider the issues of censoring, endogeneity, and heterogeneous effects along the distribution of the dependent variables raised by the nature of the current data.

Particularly, we also use three quantile estimators (the Censored Quantile [CQR], the Instrumental Variable Quantile [QIV] and the simple Quantile [QR]). These estimators address the heterogeneity issue but differ in the way they address the other two issues. The CQR approach treats the censored nature of the data, but it does not consider endogeneity whereas the QIV approach only controls for endogeneity, and standard quantile techniques do not consider both endogeneity and censoring issues. Coefficients obtained by quantile regressions can be interpreted similarly with the coefficients in an OLS regression (i.e. as the partial derivative with respect to a particular regressor). The difference is that the effect is on the conditional quantile

of the dependent variable and not on the conditional mean. Moreover, to verify whether the quantile approaches provide a more exhaustive picture of the relationship between firms' export intensity and their innovation potential, we estimate our models using both instrumental variables TOBIT and standard TOBIT models that have been widely used for analysing export intensity determinants. The first method allows us to include endogeneity issues while both approaches allow us to tackle the censored nature of the dependent variable. However, they rely on specific distributional assumptions about variables and are more sensitive to outliers, and the marginal effects are heterogeneous according to values assumed by covariates and not by values of the dependent variable, such as in quantile regressions. Finally, we also use standard IV and OLS techniques measuring only a constant effect of the regressors at the mean of the dependent variable.¹⁵

5 Empirical results

Tables 4 and 5 report the estimates of the basic specification of Eq. (1) obtained by applying the different econometric techniques mentioned above. Regarding the quantile approaches, since 23% of observed firms do not export, we present results for percentiles above the 30th; particularly, estimates are performed for each ten percentiles ranging from the 30th to the 90th percentiles.

The first-stage coefficients for the instrument ($R\&D_SP$) are always positive and highly statistically significant, thereby confirming the relevance of the instrument. More precisely, the first-stage F-statistics on excluded instruments are very high (above 25, being the square of the t-statistics for the null that the coefficient is equal to 0). This finding suggests our estimates do not suffer from a weak instrument problem and confirms that the instrument is significantly correlated with the endogenous variable. Furthermore, to determine instrument exogeneity, we estimate our basic specification after including an additional instrument, namely, 1998–2000 values of industry R&D expenditures in Germany recovered from the CIS III (see Table 7 in the Appendix). The results clearly show the instruments' validity; the Hansen J-statistics for over-identifying restrictions is equal to 0.028 with a p -value equal to 0.868.¹⁶

¹⁵ Some technical details must be explained. First, to make inference for the whole population, we estimate all regressions by weighing observations with sample weights that incorporate the stratified nature of the sample. Furthermore, to avoid missing values, the 0 values of the R&D expenditures are substituted by 1E(-10). Finally, all estimations are not performed with Stata 14 built-in routines, and quantile estimations are performed with a user-written routine (see Chernozhukov et al. (2019)). The Tobit IV estimator is the default full information maximum likelihood estimator.

¹⁶ The unconditional correlation between log R&D expenditure for Spanish and Italian firms is 0.81. Results are obtained using both instruments, which confirm the results obtained by using the single instrument reported below in the text, although marginal effects are somehow lower. We prefer showing, in the text, results with a single instrument only as the content of the two instruments partially overlaps (the F-statistics on excluded instruments is 12.94). Furthermore, Tables 4 and 5 report only the estimated coefficients of the excluded instrument from the first stage. Coefficients for the other controls in the first stage reveal that size positively and significantly affect firms' R&D spending and that firm in *specialised*

In addition to the estimated coefficients of R&D and control variables, we also report in Tables 4 and 5 the coefficient of the control function residual (labelled as *CF residual*), which is significant for all quantiles, thereby confirming the endogeneity concerns about the R&D variable.

Overall, R&D expenditures positively and significantly affect export intensity. However, the magnitude of the effect displays a significant variability associated with different estimation techniques. Figure 1 shows the pattern of the R&D coefficient estimated using the CQIV techniques for different percentiles along with the average partial effect of the Tobit IV estimates.

CQIV estimates suggest that an increase of 10% in R&D expenditures prompts an increase in export intensity, ranging between about 0.10 and 0.38 percentage points. Particularly, the impact of R&D increases along the export intensity distribution, reaches the highest value at the 70th percentile (where export intensity is about 50%), and decreases slightly at the highest percentiles (80th and 90th), but it is still higher than the lowest percentiles (30th and 40th). Therefore, CQIV regressions suggest that firms characterised by an export intensity of about 50% or higher benefit most from investing in R&D activity in terms of expansion in international markets. As expected, Tobit IV estimates lie in the ranges of the CQIV values and suggest an increase of 10% in R&D expenditures causes an increase in export intensity of about 0.14 percentage points. As shown in Fig. 1, most percentile coefficients lie outside the Tobit IV confidence interval, thereby supporting conditional quantile modelling with respect to a traditional conditional mean approach. A similar pattern of R&D coefficients is observed when applying the QIV estimator, although marginal effects are quantitatively lower, ranging between 0.2 and 3.1.

Consistent with previous international empirical literature, our findings suggest that investments in R&D activity have a positive effect on firms' export intensity by generating a higher probability of introducing product/process innovations and/or by increasing firm absorptive capacity (see the Sect. 2). Furthermore, by applying quantile regression techniques, we provide a clearer picture of the heterogeneity of such effects, which seems to depend on firms' export intensity distribution. Several not mutually exclusive considerations might explain our finding that R&D investments have a heterogeneous effect on export intensity distribution. The first, most intuitive explanation is that introducing new products and/or cost savings induced by innovative activities fosters firms' competitiveness, allowing them to enter new foreign markets and/or export new products, thus enlarging export baskets and the number of destination markets.¹⁷ Indeed, we do not have information on the evolution of export basket composition; however, our data suggest that higher export intensity is associated with a higher number of export destination markets. The

Footnote 16 (continued)

suppliers industries display an R&D advantage relative to other industries. All these results are available upon request.

¹⁷ Haddad et al. (2013) provide an interesting discussion on this issue from a more macroeconomic perspective. Particularly, by analysing a panel of developing and developed countries observed over the period 1976–2005, evidence supporting the role that exports diversification (both in terms of destination markets and export baskets) plays in reducing the vulnerability of countries to global shocks was found.

Table 4 Regression results

	CQIV						TOBITIV	
	30	40	50	60	70	80		90
$\ln(R\&D_{i-3})$	1.062*** (0.131)	1.871*** (0.170)	3.017*** (0.276)	3.040*** (0.292)	3.790*** (0.240)	2.547*** (0.762)	2.787*** (0.728)	1.366*** (0.202)
$size_{i-3}$	63.67*** (5.896)	29.25*** (5.460)	-19.86* (11.16)	-27.49** (11.69)	-51.20*** (12.18)	-48.82 (45.81)	-78.61** (37.54)	-7.702 (15.15)
$size^2_{i-3}$	-31.41*** (1.605)	-11.55*** (1.001)	-0.176 (2.849)	3.249 (2.120)	6.743*** (2.168)	12.19 (8.360)	19.54*** (6.744)	0.920 (2.800)
<i>North</i>	21.22*** (2.599)	30.47*** (2.454)	37.04*** (4.033)	17.09*** (5.542)	8.086** (3.839)	13.68 (15.38)	17.30** (7.167)	5.863 (3.566)
<i>Centre</i>	19.96*** (3.042)	23.84*** (3.252)	3.610 (4.451)	-10.32 (6.727)	0.248 (4.509)	5.525 (17.10)	-12.42 (8.921)	0.478 (5.933)
<i>Scale int</i>	-0.915 (1.424)	-1.498 (1.895)	-9.056*** (2.579)	-14.00*** (3.569)	-12.45*** (2.619)	-9.084 (9.023)	-13.60* (7.619)	-6.424 (4.054)
<i>Special.suppl</i>	10.92*** (1.790)	6.746*** (2.525)	-8.817** (4.468)	-9.153* (5.336)	-7.178 (4.682)	6.114 (18.56)	-2.156 (10.08)	-2.989 (4.611)
<i>Sc. Based</i>	-3.153* (1.893)	-5.709** (2.591)	-11.96** (4.713)	-18.31*** (5.076)	-21.20*** (4.278)	-2.668 (15.12)	-3.142 (12.99)	-7.969 (6.538)
<i>CF residual</i>	-0.954*** (0.118)	-1.663*** (0.162)	-2.632*** (0.253)	-2.606*** (0.311)	-3.328*** (0.249)	-2.224*** (0.843)	-2.344*** (0.811)	
<i>Const</i>	-2.111 (3.615)	8.867** (3.980)	33.72*** (7.252)	62.53*** (8.355)	88.70*** (6.148)	71.88*** (20.95)	103.4*** (16.38)	
<i>N obs</i>				1029				1029
<i>R&D_SP</i>				0.698*** (0.138)				

Table 4 (continued)

	CQR					TOBIT					
	30	40	50	60	70	80	90				
<i>ln(R&D_{t-3})</i>	0.0307** (0.0127)	0.169*** (0.0144)	0.674*** (0.0441)	0.717*** (0.114)	0.863*** (0.266)	0.662*** (0.208)	0.171 (0.142)	0.312** (0.153)			
<i>size_{t-3}</i>	106.9*** (1.000)	154.9*** (1.668)	93.03*** (6.603)	111.2*** (19.64)	159.1*** (38.64)	37.08 (28.83)	16.98 (12.25)	48.47*** (12.90)			
<i>size²_{t-3}</i>	-40.12*** (0.307)	-111.6*** (0.605)	-27.32*** (2.013)	-32.49*** (6.062)	-43.57*** (11.42)	-8.682 (8.286)	-1.153 (2.310)	-9.372*** (2.814)			
<i>North</i>	21.10*** (0.652)	29.77*** (0.706)	21.12*** (2.819)	24.50*** (6.670)	21.08** (10.42)	19.71*** (5.787)	20.42*** (5.362)	11.98*** (4.342)			
<i>Centre</i>	20.68*** (0.725)	28.62*** (0.839)	5.204* (3.061)	5.156 (7.584)	9.193 (11.33)	19.69** (9.229)	-4.235 (6.122)	7.117 (5.695)			
<i>Scale int</i>	-0.00883 (0.394)	-0.546 (0.386)	-6.932*** (1.235)	-13.90*** (3.637)	-9.694 (6.411)	-9.017* (5.158)	-6.471 (8.738)	-5.875 (4.389)			
<i>Special.suppl</i>	20.69*** (0.329)	26.97*** (0.450)	12.53*** (1.815)	19.65*** (5.232)	30.62*** (10.18)	30.99*** (7.532)	25.16*** (3.750)	13.85*** (4.419)			
<i>Science based</i>	7.683*** (0.377)	14.24*** (0.517)	5.530** (2.355)	4.342 (5.826)	5.629 (10.50)	8.723 (7.232)	20.03*** (6.749)	5.777 (5.249)			
<i>Const</i>	-22.78*** (0.705)	-27.54*** (0.812)	1.587 (2.894)	6.335 (6.852)	12.48 (10.66)	26.48*** (7.182)	53.56*** (7.225)				
<i>N obs</i>	1029					1029					

Standard errors in parentheses; * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. CQIV censored quantile instrumental variable; CQR censored quantile regression; CF residual control function residual from first stage. Average partial effects of the regressors are reported for TOBITIV and TOBIT. Omitted categories are firms located in the South operating in the supply dominated sector. A dummy for the year 1994 is included but not shown in all regressions. Estimates performed with Stata 14. Software

Table 5 Regression results

	QIV					IV		
	30	40	50	60	70		80	90
<i>ln(R&D_{t-3})</i>	0.226*** (0.0239)	0.619*** (0.0202)	1.457*** (0.0959)	2.634*** (0.318)	3.145*** (0.101)	2.404*** (0.673)	1.963*** (0.0940)	1.860*** (0.444)
<i>size_{t-3}</i>	77.13*** (0.973)	77.45*** (0.799)	51.76*** (3.472)	-4.069 (10.96)	-32.35*** (5.179)	-43.20 (40.27)	-52.17*** (4.853)	-3.801 (25.09)
<i>size²_{t-3}</i>	-21.59*** (0.180)	-22.59*** (0.149)	-12.71*** (0.643)	-0.0492 (1.974)	3.816*** (0.942)	11.12 (7.362)	13.40*** (0.873)	-0.447 (4.516)
<i>North</i>	0.198 (0.317)	-0.462 (0.284)	-1.532 (1.375)	-4.530* (2.575)	1.310 (1.121)	10.44 (9.061)	21.30*** (0.946)	6.583 (4.287)
<i>Centre</i>	-0.486 (0.395)	-1.407*** (0.338)	-4.729*** (1.610)	-5.130 (4.200)	-10.13*** (1.717)	3.981 (13.26)	-0.810 (1.109)	-0.956 (7.237)
<i>Scale intensive</i>	-0.666*** (0.242)	-1.001*** (0.223)	-1.852* (0.997)	-7.588*** (2.759)	-11.78*** (1.073)	-10.85 (7.645)	-11.27*** (0.989)	-6.722 (4.927)
<i>Special.suppl</i>	8.004*** (0.377)	19.46*** (0.350)	19.48*** (1.615)	-1.469 (5.358)	-2.080 (2.088)	5.483 (16.29)	5.174*** (1.330)	-0.317 (7.436)
<i>Science based</i>	4.806*** (0.418)	9.521*** (0.378)	1.743 (1.746)	-8.603 (5.427)	-19.55*** (1.873)	-1.871 (13.35)	3.501** (1.783)	-8.716 (10.24)
<i>CF residual</i>	-0.208*** (0.024)	-0.374*** (0.020)	-0.988*** (0.092)	-2.257*** (0.299)	-2.728*** (0.108)	-2.043*** (0.749)	-1.636*** (0.105)	50.38*** (10.71)
<i>Const</i>	3.606*** (0.637)	13.95*** (0.529)	35.72*** (2.459)	69.56*** (8.006)	86.57*** (2.599)	74.91*** (16.78)	83.17*** (2.141)	1029 (10.71)
<i>N obs</i>				1029				1029
<i>R&D_SP</i>				0.698*** (0.138)				

Table 5 (continued)

	QR					OLS					
	30	40	50	60	70	80	90				
$\ln(R\&D_{i-3})$	0.0575*** (0.0107)	0.210*** (0.00597)	0.593*** (0.0203)	0.576*** (0.125)	0.659*** (0.162)	0.671*** (0.186)	0.364*** (0.0837)	0.369*** (0.168)			
$size_{i-3}$	87.48*** (0.808)	101.1*** (0.491)	107.5*** (1.731)	114.0*** (11.48)	97.36*** (16.51)	36.49* (20.04)	12.77 (8.079)	57.43*** (18.14)			
$size^2_{i-3}$	-24.18*** (0.153)	-28.36*** (0.0944)	-27.91*** (0.338)	-27.93*** (2.275)	-18.88*** (3.312)	-7.283** (3.698)	0.0586 (1.551)	-11.39*** (3.791)			
<i>North</i>	0.980*** (0.297)	1.199*** (0.191)	0.573 (0.744)	3.000 (2.910)	9.274* (4.743)	16.80*** (5.504)	26.68*** (3.746)	11.26*** (3.376)			
<i>Centre</i>	0.461 (0.361)	0.0992 (0.223)	-0.107 (0.911)	9.548** (4.536)	3.385 (6.262)	17.06** (8.036)	5.089 (3.839)	5.946 (4.545)			
<i>Scale intensive</i>	-0.432* (0.247)	-0.699*** (0.146)	-0.680 (0.493)	-7.363** (3.084)	-11.40*** (3.902)	-11.89** (4.693)	-10.00** (4.216)	-4.374 (4.289)			
<i>Special.suppl</i>	14.39*** (0.291)	23.10*** (0.177)	29.00*** (0.640)	22.11*** (4.127)	27.57*** (5.753)	30.75*** (7.337)	22.52*** (2.417)	17.87*** (6.153)			
<i>Science based</i>	5.867*** (0.329)	10.44*** (0.211)	11.19*** (0.720)	10.76** (4.662)	6.748 (6.066)	8.682 (6.874)	23.58*** (6.008)	7.794 (6.640)			
<i>Const</i>	-0.572 (0.372)	3.315*** (0.221)	13.84*** (0.805)	17.76*** (3.968)	24.99*** (5.662)	29.35*** (6.430)	43.92*** (4.742)	13.97*** (5.095)			
<i>N obs</i>											1029

Standard errors in parentheses; * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. QR quantile regression; QIV quantile instrumental variable; CF residual control function residual from first stage. Omitted categories are firms located in the South operating in the supply dominated sector. A dummy for the year 1994 is included but not shown in all regressions. Estimates performed with Stata 14 Software

Unicredit-Capitalia database allows us to identify eight different destination markets: UE (15), other European countries, Africa, North America, South America, Asia, China, Australia, and Oceania. Firms included in our sample, on average, export to 3.2 foreign markets. Particularly, the average number of destination markets increases from 1.45 for firms exporting less than 3% of their turnover (30th percentile) to 4.22 for those with export intensity higher than 76% (90th percentile). Hence, the number of markets served increases along with the distribution of export intensity. A higher number of destination markets might have a multiplicative effect on innovation rents and might help diversify business risks by providing cross-subsidization of revenues. A similar argument has been suggested, among others, by Halpern and Muraközy (2012), who analyse the innovation and firm performance relation on Hungarian data. The authors find that innovation has a positive effect on all margins of trade and innovative firms export to more countries, more products, and more in one product–destination combination, on average.

Another possible explanation might be related to the different ways export activities are organised. Instead of directly (or actively) exporting, firms may engage in exports indirectly (or passively) via intermediaries. If firms with different export methods are also characterised by different export intensities, the heterogeneous impact of R&D might be demand-induced (or reduced) by intermediaries. Finally, firms exporting a non-negligible fraction of their sales tend to devote specific resources, such as export departments and export managers, to export activities. In turn, the return on innovative activities might be higher in terms of an increase in export shares; larger human resources are devoted to foreign markets.

The relationship between firms' innovative effort and exporting activity detected by applying the CQIV estimator is confirmed when Eq. (1) is estimated with the CQR and QR approaches. In both cases, the pattern of estimated coefficients for R&D appears to be similar to the CQIV results as they reach their maximum value around the 70th percentile, and marginal effects are found to be lower with respect to CQIV and QIV values since they range between 0.03 and 0.8. Likewise, the estimates obtained with the Tobit and OLS methods are smaller than those computed using the instrumental variables approach (Tobit IV and IV). Therefore, the results obtained when both endogeneity issues and the censored nature of the dependent variable are neglected provide the lower bound of the parameter values.

To further support our results, properly testing the statistical significance of the difference among parameter estimates obtained at different quantiles is important. Unfortunately, the CQIV estimator provided by Chernozhukov et al. (2015) does not provide a variance–covariance matrix that would allow us to conduct specific tests involving more than one quantile. To overcome this limitation, we can compare parameter estimates at different quantiles by inspecting confidence intervals in Fig. 1. If they do not overlap, we can infer that coefficients differ along export intensity distribution. Particularly, confidence intervals confirm the increasing pattern of the R&D coefficients until the 70th percentile and seem to suggest that coefficients estimated at the highest percentiles (from 70th onwards) do not significantly differ from each other.

Furthermore, for each pair of quantiles, we performed *t*-tests on the R&D coefficients from CQIV estimates. Particularly, we computed the following statistics:

$$t = \frac{\beta_{1i} - \beta_{1j}}{\sqrt{se_{\beta_{1i}}^2 + se_{\beta_{1j}}^2}} \tag{2}$$

where β_{1i} and β_{1j} are the CQIV estimated coefficients of R&D at quantiles i and j (Table 5), and $se_{\beta_{1i}}$ and $se_{\beta_{1j}}$ are the standard errors. The null hypothesis is that the coefficients are equal between pairwise quantiles. In Table 6, we report t -statistics, and we highlight in bold those values allowing us to reject the null hypothesis (at the 5% significance level).

Although these tests are not rigorous, their results are consistent with the results of the graphical inspection of Fig. 1. Particularly, the effect of R&D estimated at the 70th percentile is significantly higher than those estimated at other percentiles, both lower and higher. However, in the latter case, the null hypothesis is rejected only at the 20% significance level.

We cannot directly compare our results with those of Wagner (2006) on German data because the way R&D is included as a regressor is different. However, they are strikingly similar to those provided by Orts and Martí (2018) on EFIGE data. Despite this, we believe our results on the heterogeneous impact of R&D on export intensity along its conditional distribution can provide novel evidence on this issue because they are obtained by properly addressing the endogeneity of the R&D variable and the censored nature of the dependent variable.

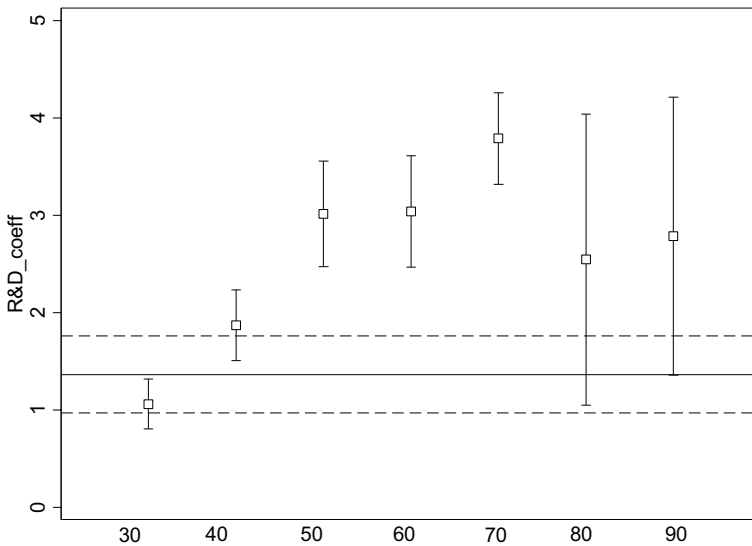


Fig. 1 Censored quantile instrumental variable (CQIV) and Tobit IV estimates. Figure shows estimated R&D coefficients (squares) and their confidence interval (vertical lines) obtained by applying censored quantile instrumental variable estimator (CQIV) at 30th, 40th, 50th, 60th, 70th, 80th and 90th percentiles of the export intensity distribution. The horizontal continuous line indicates the R&D average partial effects estimated by applying instrumental variables Tobit (Tobit IV) model and the horizontal dashed lines display the corresponding 95% confidence interval bounds

Table 6 *t*-test on pairwise quantiles

Quantiles	40	50	60	70	80	90
30	-3.778	-6.41	-6.181	-9.982	-1.922	-2.331
40		-5.348	-3.832	-5.997	-2.471	-1.185
50			-0.057	-2.116	0.581	0.296
60				-1.985	0.605	0.323
70					1.557	1.309
80						-0.228

The null hypothesis is that the R&D coefficients are equal between pairwise quantiles. A negative statistics means that the R&D coefficient at the column quantile is higher than the coefficient at the row quantile. Bold values indicate rejection of the null hypothesis at 5% level of significance

On the results concerning different control variables included in the basic specification of Eq. 1, the effect of firm size (evaluated at *t*-3) is not stable along the export intensity distribution and exhibits a significant variability associated with different estimation techniques. Coefficients from the CQIV regressions follow a bell-shaped pattern for lower quantiles while a U-shaped relationship between size and export intensity is observed in the right tail of the distribution.¹⁸ Given that firm size in our sample increases along the export intensity distribution, these findings are broadly consistent with those obtained by Sterlacchini (2001), who finds an inverted U-shaped relationship for small firms, a U-shaped relationship for large firms, and no impact of size on export intensity for medium-sized firms.¹⁹ Size coefficients show a similar pattern when the QIV estimator is applied, while Tobit IV and IV estimates do not show any impact of size on export intensity. A bell-shaped relationship between firm size and exporting activity that persists along most of the export intensity distribution is shown by the CQR and QR estimates and by Tobit and OLS regressions. As discussed above for the effect of R&D, the adoption of the most suitable estimation technique helps disentangle more nuanced evidence on the relationship between size and export intensity.

Examining the coefficients of geographical location dummies, heterogeneous results emerge according to the different quantiles analysed. Estimates reveal that being located in Northern Italy positively affects export intensity in most of the estimated models. This result is consistent with the evidence provided by previous literature for Italy showing that northern firms exhibit a higher export intensity relative to firms located in other areas (Brancati et al., 2018; Bugamelli et al., 2000; D'Angelo, 2012; Sterlacchini, 2001).

¹⁸ These results are supported by F-tests on joint significance of the size parameters.

¹⁹ Some authors identify an inverted U-shape relationship between the two variables while other authors do not investigate the nonlinearity of such relationship (e.g. Basile 2001; Castellani 2002; D'Angelo 2012; Nassimbeni 2001). Our results differ from those suggested by Wagner (2006), who finds a linear relationship between size and export intensity only at the 25th percentile of the export intensity distribution.

Furthermore, estimates suggest that firms whose export intensity is lower than 10% and located in the South of Italy seem to have a locational disadvantage relative to those located in the Centre, while such a differential disappears for firms exporting a higher percentage of their sales. This finding might suggest that southern firms may find it difficult to begin participating in export markets, for instance, because of geographical distance and, hence, higher transportation and informational costs. However, once they increase their participation, as measured by export intensity, they overcome these disadvantages. This pattern is not confirmed when we apply other estimation techniques that do not provide conclusive evidence.

As for sectoral dummies, coefficients turn to be at best weakly significant for highest quantiles of the export intensity distribution, once we control for formal R&D activity, size, and location. Hence, CQIV results suggest that the observed highest export intensity/propensity characterizing firms operating in *specialised suppliers* and *science-based* industries (as shown in Tables 2 and 3) is likely to be associated to their R&D activity rather than to sector specific characteristics. These findings are broadly confirmed by estimates obtained with Tobit IV, QIV, and IV, while the unconditional rankings of Tables 2 and 3 tend to emerge with estimation methods that do not consider the endogeneity of the innovation activity.

To check the robustness of our results, we follow different strategies. First, we replicate our estimates by using R&D intensity (i.e. R&D over total sales) instead of R&D expenditures, and all our findings are confirmed. Second, we augment our basic specification by including some firm characteristics that have been found to affect export intensity in previous empirical literature. We include firms' ownership structure to account for the fact that firms belonging to foreign owners might be more skilled to compete in international markets as they might have a larger international network. Alternatively, a binary variable indicating whether the firm has established, through either brownfield or greenfield, a foreign subsidiary in the previous 3-year period is also included. Indeed, FDI and export activity might be correlated since FDI might improve the knowledge of foreign markets, and that of the markets obtained by exporting might favour FDI investments. Most of the results discussed so far are broadly confirmed when we control for these covariates.²⁰

As an additional robustness check, we estimate Eq. (1) after splitting the sample according to different criteria. Estimates conducted stratifying the sample according to Pavitt classification suggest that our main results on the relationship between R&D and export intensity are confirmed regardless of the category considered.²¹

²⁰ Following a referee's suggestion we performed a further robustness check by considering a model where the (log) value of exports is chosen as an alternative dependent variable (other than export intensity) and estimating the model on three different sub-samples of exporting firms only stratified by firms export intensity (less than 20%, between 20 and 53% and above 53%). Results show that the elasticity of exports with respect to R&D is increasing along firm export intensity distribution thereby indirectly confirming our main result.

²¹ In particular, to maintain a reasonable size, we aggregated the four Pavitt sectors in two subsamples: the first one aggregates *scale intensive* and *science-based* industries, which mostly rely upon R&D activity as internal source of innovation, and the second one aggregating *supply dominated* and *specialised supplier* industries, for which internal R&D expenditure is not the main source of innovation.

This notwithstanding, as expected, the effect of R&D on export intensity is stronger for companies operating in industry that base their innovative activity more on internal R&D expenditures (i.e. *scale intensive* and *science-based* industries). Moreover, a stronger heterogeneity on the relationship between R&D expenditures and export intensity emerges when we split the sample according to firms' geographical location. Indeed, our main results seem to be mostly driven by firms operating in the Centre and South of Italy, where the impact of R&D increases along the export intensity distribution and reaches the highest value between the 70th and 80th percentiles. As for firms located in Northern Italy, instead, a quite stable impact of the R&D on export intensity emerges. This impact is also very strong, being larger (for all quantiles but the first) to the maximum effect estimated for the rest of the country. Finally, the overall findings are also confirmed when we estimate our model on subsamples based on firms' size (small vs. medium large firms as defined in the Data Section).²² Indeed, estimates conducted on separate sub-samples have to be taken with caution given the small size of the different subsamples.

6 Concluding remarks

In this paper, we analyse firms' export intensity determinants with a particular focus on the role of innovative effort. To better understand the relationship between firms' innovative effort, measured by R&D expenditures and export intensity, we study whether such a relationship varies along the export intensity distribution by applying quantile regression techniques on a cross-sectional sample of Italian manufacturing firms. We also deal with the censored nature of the export intensity variable and with the endogeneity of the innovation proxy by applying the CQIV estimator recently proposed by Chernozhukov et al. (2015).

The empirical results obtained by applying CQIV and other estimation techniques confirm that R&D expenditures positively affect export intensity. Particularly, this effect increases along the export intensity distribution, reaches the highest value at the 70th percentile (where export intensity is about 50%), and persists at the highest percentiles (80th and 90th). Therefore, our results suggest that firms characterised by an export intensity of about 50% or more can obtain the highest advantages from investing in R&D activity in terms of expansion in international markets. Point estimates suggest that an increase of 10% in R&D expenditures prompts an increase in export intensity, which ranges between about 0.10 and 0.38 percentage points according to different export intensity levels.

²² We also estimate a model often adopted by the empirical literature on export, namely, the fractional logit model, that considers the special nature of the dependent variable (a fraction between zero and one) better compared to the Tobit estimator. The results strongly confirm those obtained with the Tobit estimator, and, due to the unavailability of an immediate way to incorporate endogenous regressors, we present the Tobit results in the text. All estimates related to robustness analysis are reported in the Appendix.

We have discussed different possible explanations for these results, grounded on competition effects, the way exporting activities are organised, the number of destination markets, and the amount of resources devoted to the exports. As for external validity, assessing whether our results hold for other countries and/or in recent years is difficult. Particularly, our sample period covers only the beginning of the first decade of this century, thereby neglecting the huge increase in world trade and the great recession. In turn, a natural extension of this study is to replicate our analyses with data covering the years following the great recession. However, the similarity of our results with those obtained by Orts and Martí (2018) using more recent data covering several countries (e.g. the EFIGE dataset) is partially reassuring in this respect. Indeed, our sample shows very similar characteristics relative to the Italian data included in the EFIGE dataset. Particularly, the sectoral and size distribution, percentage of firms investing in R&D, and percentage of exporting firms do not significantly differ. The overall similarity of our dataset with the more recent EFIGE database can provide some intuition on the external validity of our results.

Our findings show the importance of considering the heterogeneous effects of covariates along export intensity distribution, not only to better understand the relationship between export intensity and firms' innovative effort but also to provide useful insights for the design of policy instruments aimed at favouring export activity. Particularly, R&D incentives should consider their impact on export performance (see Altomonte et al., 2013), and our results suggest that such policy measures are likely to have a heterogeneous effect, being stronger for firms characterised by an already high level of export intensity. Overall, the results point towards the importance of supporting the growth of firms' export-intensive margin to better exploit the positive impact of innovation incentives. Results of the heterogeneity analysis, to be taken with caution given the small size of different subsamples, suggest that the overall effect is homogenous across firm size but seems to vary in magnitude according to Pavitt sectoral classification and geographical location. Indeed, the relationship between R&D and export performance seems to be stronger for firms that mostly base their innovative activity on internal R&D expenditures (i.e. *scale intensive* and *science-based* industries) and for firms operating in the North of Italy. Hence, policies which uniformly affect incentives to perform R&D for all firms will tend to exacerbate existing differences in exporting performance across sectors and regions thereby suggesting the need to adopt other kinds of interventions to boost export performance of firms in South-Central Italy and operating in supply dominated and in specialised suppliers industries. Further research using wider samples might be useful to clarify this issue.

Appendix

See. Tables [7](#), [8](#), [9](#), [10](#), [11](#), [12](#), [13](#), [14](#), [15](#), [16](#) and [17](#).

Table 7 Estimates using R&D expenditures in Germany as additional instrument

	CQIV					TOBITIV				
	30	40	50	60	70	80	90			
$\ln(R\&D_{i-3})$	0.685*** (0.112)	1.394*** (0.101)	3.315*** (0.370)	3.022*** (0.384)	4.001*** (0.265)	2.552*** (0.769)	2.797*** (0.942)			1.386*** (0.213)
$size_{i-3}$	89.15*** (5.430)	58.349*** (3.306)	-30.59** (14.72)	-17.89 (10.973)	-63.99*** (16.47)	-48.07 (45.42)	-77.11 (48.30)			-9.097 (15.85)
$size^2_{i-3}$	-48.26*** (1.489)	-18.649*** (0.605)	2.184 (3.729)	1.878 (1.938)	13.05*** (4.601)	11.93 (8.293)	18.98** (8.679)			1.178 (2.942)
<i>North</i>	20.40*** (2.261)	31.737*** (1.581)	36.71*** (5.285)	16.082** (4.868)	8.967** (3.669)	13.86 (13.98)	17.88* (9.275)			5.756 (3.590)
<i>Centre</i>	19.15*** (2.729)	25.35*** (2.128)	3.188 (5.738)	-10.612* (6.121)	0.314 (4.282)	5.885 (15.87)	-11.72 (11.75)			0.367 (5.859)
<i>Scale int</i>	-1.535 (1.133)	-2.200** (1.037)	-8.765*** (3.330)	-18.72*** (3.503)	-12.62*** (2.498)	-8.353 (8.849)	-13.59 (9.816)			-6.380 (4.044)
<i>Special.suppl</i>	13.26*** (1.615)	14.248*** (1.520)	-11.35* (5.820)	-8.997 (6.788)	-10.65** (4.449)	7.191 (18.00)	-3.496 (13.13)			-3.404 (4.838)
<i>Science based</i>	-0.0386 (1.685)	-0.014 (1.548)	-16.84*** (6.131)	-21.305*** (6.378)	-23.30*** (4.222)	-2.550 (14.81)	-3.715 (16.88)			-8.290 (6.644)
<i>CF residual</i>	-0.652*** (0.104)	-1.226*** (0.099)	-2.913*** (0.340)	-2.470*** (0.367)	-3.557*** (0.265)	-2.226*** (0.835)	-2.383** (1.052)			
<i>Const</i>	-7.466** (3.207)	-2.913 (2.530)	39.22*** (9.394)	63.796*** (9.712)	93.92*** (6.524)	71.14*** (20.57)	103.3*** (21.22)			
<i>N. obs</i>				1029						1029
<i>R&D_SP</i>				0.730*** (0.165)						0.747*** (0.156)
<i>R&D_GE</i>				-0.063 (0.149)						-0.113 (0.130)

Table 7 (continued)

	QIV					IV			
	30	40	50	60	70	80	90		
$\ln(R\&D_{i-3})$	0.194*** (0.0305)	0.536*** (0.0195)	1.398*** (0.0287)	2.577*** (0.0573)	3.050*** (0.306)	2.430*** (0.686)	2.269*** (0.132)	1.853*** (0.441)	
$size_{i-3}$	78.84*** (1.249)	81.56*** (0.785)	54.20*** (1.063)	-1.979 (2.010)	-29.08* (15.64)	-43.58 (40.90)	-66.81*** (6.806)	-2.797 (25.15)	
$size^2_{i-3}$	-22.01*** (0.231)	-23.60*** (0.146)	-13.40*** (0.197)	-0.325 (0.363)	3.354 (2.846)	11.12 (7.477)	16.38*** (1.224)	-0.498 (4.526)	
<i>North</i>	0.305 (0.398)	-0.274 (0.275)	-2.305*** (0.290)	-4.619*** (0.457)	1.450 (3.356)	10.42 (9.143)	20.54*** (1.332)	6.605 (4.288)	
<i>Centre</i>	-0.392 (0.496)	-1.083*** (0.325)	-5.182*** (0.392)	-5.001*** (0.741)	-9.092* (5.109)	4.263 (13.31)	-1.355 (1.582)	-0.924 (7.122)	
<i>Scale int</i>	-0.602** (0.304)	-0.729*** (0.217)	-1.864*** (0.296)	-8.168*** (0.483)	-11.45*** (3.176)	-10.59 (7.652)	-10.96*** (1.395)	-6.710 (4.921)	
<i>Special.suppl</i>	8.473*** (0.482)	20.60*** (0.342)	19.71*** (0.489)	-1.778* (0.953)	-1.279 (6.148)	5.563 (16.30)	2.645 (1.866)	-0.231 (7.481)	
<i>Science based</i>	5.155*** (0.529)	10.76*** (0.368)	2.564*** (0.525)	-9.549*** (0.962)	-18.35*** (5.573)	-1.899 (13.42)	0.0984 (2.508)	-8.639 (10.23)	
<i>CF residual</i>	-0.173*** (0.0304)	-0.284*** (0.0192)	-0.911*** (0.0274)	-2.205*** (0.0541)	-2.616*** (0.326)	-2.069*** (0.758)	-1.973*** (0.147)	50.21*** (10.61)	
<i>Const</i>	2.819*** (0.806)	11.87*** (0.509)	35.17*** (0.743)	69.21*** (1.439)	84.23*** (7.819)	75.14*** (17.09)	90.40*** (3.015)	1029 (1029)	

Table 7 (continued)

	QIV				IV		
	30	40	50	60	70	80	90
<i>R&D_SP</i>				0.730***			
				(0.165)			
<i>R&D_DE</i>				-0.063 (0.149)			

Standard errors in parentheses; * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. *CQIV* censored quantile instrumental variable; *CF residual* control function residual from first stage. Average partial effects of the regressors are reported for *TOBITIV*. Omitted categories are firms located in the South operating in the Supply Dominated sector. A dummy for the year 1994 included but now shown in all regressions. Estimates performed with Stata 14 Software

Table 8 Fractional logit estimates (average partial effects)

	<i>Expint_{i,t}</i>
<i>ln(R&D_{t,3})</i>	0.003** (0.0014)
<i>size_{t,3}</i>	0.558*** (0.186)
<i>size²_{t,3}</i>	-0.148** (0.059)
<i>North</i>	0.152*** (0.055)
<i>Centre</i>	0.104 (0.065)
<i>Scale int</i>	-0.049 (0.052)
<i>Special.suppl</i>	0.141*** (0.044)
<i>Science based</i>	0.067 (0.053)
<i>N. obs</i>	1029

Standard errors in parentheses; **p* < 0.10, ***p* < 0.05, ****p* < 0.01. Omitted categories are firms located in the South operating in the Supply Dominated sector. A dummy for the year 1994 included but now shown in all regressions. Estimates performed with Stata 14 Software

Table 9 CQIV estimates: subsample of firms operating in Scale intensive and Science based sectors

	30	40	50	60	70	80	90
<i>ln(R&D_{t,3})</i>	1.816*** (0.154)	3.369*** (0.449)	5.050*** (0.0172)	5.737*** (0.351)	7.231*** (1.376)	8.982*** (2.005)	6.930*** (1.661)
<i>size_{t,3}</i>	-2.519 (7.062)	-94.72*** (22.90)	-150.7*** (0.978)	-176.3*** (20.07)	-262.6*** (81.33)	-352.2*** (122.0)	-206.6** (103.4)
<i>size²_{t,3}</i>	-1.166 (1.845)	25.93*** (6.149)	38.84*** (0.253)	44.84*** (5.191)	68.77*** (20.96)	90.18*** (31.42)	53.22* (27.40)
<i>North</i>	-8.620*** (1.665)	-39.10*** (3.045)	-8.729*** (0.0977)	-9.994*** (1.656)	-4.595 (5.563)	0.812 (6.954)	6.052 (19.98)
<i>Centre</i>	-26.71*** (2.238)	-80.62*** (6.032)	-65.93*** (0.220)	-74.33*** (4.362)	-71.86*** (16.86)	-86.52*** (22.27)	-59.55** (24.44)
<i>Scale int</i>	6.773*** (1.518)	13.31*** (4.410)	21.28*** (0.162)	24.97*** (3.286)	42.23*** (13.35)	54.05*** (19.39)	35.46 (29.48)
<i>CF residual</i>	-1.459*** (0.144)	-3.016*** (0.435)	-4.636*** (0.0178)	-5.306*** (0.363)	-7.073*** (1.435)	-9.126*** (2.149)	-7.062*** (1.673)
<i>Const</i>	35.65*** (3.109)	96.17*** (7.807)	90.97*** (0.270)	103.6*** (5.225)	124.3*** (19.57)	156.1*** (26.65)	147.2*** (34.03)
<i>N. obs</i>				310			
<i>R&D_SP</i>				0.556*** (0.176)			

Standard errors in parentheses; **p* < 0.10, ***p* < 0.05, ****p* < 0.01

Table 10 CQIV estimates: subsample of firms operating in supply dominated and specialised supplier based sectors

	30	40	50	60	70	80	90
$\ln(R\&D_{t-3})$	1.272*** (0.0416)	2.597*** (0.000897)	2.744*** (0.145)	3.826*** (0.439)	3.982*** (0.0972)	2.627*** (0.477)	1.223 (0.830)
$size_{t-3}$	40.22*** (1.534)	10.37*** (0.0344)	10.92** (4.566)	-35.88*** (13.30)	-69.44*** (5.367)	-102.9 (64.71)	-29.81 (65.68)
$size_{t-3}^2$	-6.632*** (0.223)	-2.833*** (0.00506)	-3.191*** (0.654)	3.398* (1.898)	7.979*** (0.788)	82.37 (56.99)	25.80 (52.17)
<i>North</i>	17.91*** (0.249)	32.01*** (0.0120)	45.46*** (1.565)	31.97*** (5.677)	23.96*** (1.726)	20.27*** (6.080)	14.22 (11.22)
<i>Centre</i>	19.37*** (0.579)	34.10*** (0.0153)	15.35*** (1.804)	36.45*** (6.999)	19.20*** (2.018)	23.09*** (6.115)	6.784 (12.62)
<i>Suppl_domin</i>	-4.393*** (0.490)	2.762*** (0.0129)	-0.836 (1.804)	13.64** (6.452)	9.118*** (1.683)	-5.407 (7.003)	-23.90** (11.64)
<i>CF residual</i>	-1.227*** (0.0389)	-2.291*** (0.000858)	-1.887*** (0.128)	-3.388*** (0.393)	-3.451*** (0.0997)	-2.077*** (0.431)	-0.704 (0.833)
<i>Const</i>	9.918*** (0.455)	21.34*** (0.0141)	16.88*** (1.930)	47.65*** (6.525)	71.86*** (1.966)	76.00*** (9.449)	105.0*** (14.32)
<i>N. obs</i>				719			
<i>R&D_SP</i>				0.685***	(0.182)		

Standard errors in parentheses, * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table 11 CQIV results, R&D intensity (instead of R&D expenditure)

	30	40	50	60	70	80	90
$R\&Dint_{t-3}$	6.879*** (0.847)	21.799*** (0.6377)	52.36*** (0.498)	64.314*** (0.694)	67.64*** (19.97)	59.08*** (7.508)	38.06*** (18.77)
$size_{t-3}$	83.85*** (2.103)	80.232*** (1.489)	14.62*** (0.799)	7.923*** (0.906)	-22.55 (43.27)	-53.30*** (22.67)	-43.31 (46.43)
$size^2_{t-3}$	-24.58*** (0.617)	-23.047*** (0.411)	-6.221*** (0.218)	-1.698*** (0.165)	5.518 (12.20)	14.37*** (4.214)	11.97 (8.553)
<i>North</i>	-3.712*** (0.673)	28.048*** (0.632)	29.28*** (0.395)	25.977*** (0.589)	6.015 (12.42)	3.398 (6.124)	13.83 (10.84)
<i>Centre</i>	-5.106*** (1.041)	25.030*** (0.749)	-1.375*** (0.438)	2.846*** (0.691)	-11.07 (16.49)	-16.69* (9.643)	-15.03 (16.12)
<i>Scale int</i>	-0.872** (0.409)	-1.545*** (0.413)	-4.493*** (0.232)	-3.900*** (0.314)	-11.45 (8.631)	-8.343* (4.887)	-6.846 (11.29)
<i>Special.suppl</i>	17.98*** (0.825)	11.550*** (0.642)	-2.836*** (0.417)	-9.585*** (0.685)	-21.64 (21.38)	-12.62 (10.16)	2.295 (16.93)
<i>Science based</i>	12.01*** (0.810)	1.729*** (0.615)	-8.722*** (0.409)	-22.350*** (0.639)	-26.78 (18.34)	-12.71 (8.255)	-4.119 (19.35)
<i>CF residual</i>	-6.987*** (0.840)	-21.084*** (0.635)	-48.25*** (0.490)	-62.495*** (0.699)	-65.07*** (19.89)	-56.58*** (7.499)	-37.83*** (19.23)
<i>Const</i>	1.208** (0.606)	-33.29*** (0.657)	-38.40*** (0.438)	-30.614*** (0.630)	-0.155 (13.08)	16.36*** (6.186)	39.92*** (10.19)
<i>N. obs</i>				1029			
<i>R&D_SIP</i>				0.039*** (0.009)			

Standard errors in parentheses; * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. CQIV censored quantile instrumental variable; CF residual control function residual from first stage. Omitted categories are firms located in the South operating in the Supply Dominated sector. A dummy for the year 1994 included but now shown in all regressions. Estimates performed with Stata 14 Software

Table 12 CQIV, firms' ownership structure as additional control

	30	40	50	60	70	80	90
$\ln(R\&D_{t-3})$	0.646*** (0.0865)	1.838*** (0.101)	3.025*** (0.293)	3.038*** (0.365)	3.761*** (0.156)	2.640*** (0.816)	1.994*** (0.746)
$size_{t-3}$	123.9*** (4.116)	29.43*** (3.273)	-22.09* (12.51)	-28.36* (15.01)	-54.25*** (8.114)	-59.87 (49.55)	-48.41 (39.02)
$size^2_{t-3}$	-102.9*** (1.335)	-11.61*** (0.600)	0.319 (3.365)	3.412 (2.721)	7.236*** (1.446)	14.91* (9.036)	12.85* (7.009)
<i>North</i>	29.18*** (0.940)	29.47*** (1.588)	37.45*** (4.308)	17.41** (7.510)	9.267*** (2.521)	31.69* (16.84)	30.32*** (7.655)
<i>Centre</i>	26.75*** (1.633)	22.08*** (2.136)	4.321 (4.711)	-9.973 (8.892)	1.359 (2.954)	24.81 (18.84)	7.429 (9.407)
<i>Scale int</i>	-1.552* (0.930)	-1.527 (1.186)	-8.658*** (2.724)	-14.68*** (4.500)	-12.37*** (1.696)	-12.05 (9.815)	-11.10 (7.572)
<i>Special.suppl</i>	11.73*** (1.216)	6.128*** (1.494)	-8.292* (4.740)	-8.987 (6.727)	-7.232** (3.045)	6.960 (19.77)	10.18 (10.52)
<i>Science based</i>	0.979 (1.286)	-4.498*** (1.549)	-11.81** (5.014)	-18.98*** (6.720)	-20.56*** (2.793)	-3.754 (16.07)	2.649 (13.93)
<i>Foreign_{t-3}</i>	11.24*** (0.983)	3.577*** (1.281)	15.71*** (3.885)	14.05** (6.377)	29.06*** (2.369)	23.31** (11.64)	4.463 (7.861)
<i>CF residual</i>	-0.597*** (0.0828)	-1.639*** (0.0970)	-2.622*** (0.271)	-2.580*** (0.390)	-3.271*** (0.161)	-2.247** (0.900)	-1.628* (0.837)
<i>Const</i>	-18.66*** (2.040)	9.115*** (2.509)	32.86*** (7.629)	62.33*** (10.91)	87.08*** (3.977)	54.95** (23.45)	74.60*** (17.04)
<i>N. obs</i>				1021			
<i>R&D_SP</i>				0.702***	(0.139)		

Standard errors in parentheses; * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. CQIV censored quantile instrumental variable; CF residual control function residual from first stage. Omitted categories are firms located in the South operating in the Supply Dominated sector. A dummy for the year 1994 included but now shown in all regressions. Estimates performed with Stata 14 Software. Foreign is a binary variable equal to one if the primary shareholder is foreign

Table 13 CQIV, FDI dummy as additional control

	30	40	50	60	70	80	90
$\ln(R\&D_{t-3})$	1.859*** (0.0256)	1.893*** (0.223)	3.080*** (0.312)	2.968*** (0.273)	3.781*** (0.250)	2.605*** (0.808)	2.563*** (1.039)
$size_{t-3}$	17.44*** (1.060)	22.32*** (6.949)	-23.37* (12.76)	-25.62** (11.09)	-51.76*** (12.74)	-59.31 (48.74)	-84.83 (53.98)
$size^2_{t-3}$	-7.864*** (0.283)	-9.462*** (1.273)	0.768 (3.277)	2.941 (2.011)	6.770*** (2.274)	14.68* (8.890)	20.93*** (9.704)
North	25.20*** (0.527)	29.73*** (3.365)	21.23*** (5.627)	16.66*** (5.173)	8.697** (3.979)	14.67 (15.93)	17.82* (10.22)
Centre	6.231*** (0.552)	22.52*** (4.446)	-11.91** (6.047)	-9.546 (6.280)	0.893 (4.685)	6.520 (17.63)	-11.35 (12.83)
Scale int	-0.442 (0.284)	-1.433 (2.597)	-8.480*** (2.896)	-14.57*** (3.331)	-12.49*** (2.718)	-8.847 (9.596)	-9.805 (10.78)
Special.suppl	5.447*** (0.319)	7.671** (3.194)	-8.963* (5.057)	-9.794** (4.979)	-7.339 (4.847)	6.180 (19.68)	0.127 (14.50)
Science based	-4.197*** (0.359)	-2.528 (3.312)	-12.27** (5.354)	-17.95*** (4.747)	-20.98*** (4.437)	-3.821 (16.03)	-0.693 (18.62)
FDI_{t-3}	3.227*** (0.319)	-0.314 (2.834)	8.863* (4.919)	15.39*** (5.693)	14.38*** (5.136)	24.38 (18.52)	46.42** (20.83)
CF residual	-1.354*** (0.0222)	-1.628*** (0.211)	-2.689*** (0.286)	-2.529*** (0.291)	-3.304*** (0.258)	-2.252** (0.893)	-2.196* (1.160)
Const	7.237*** (0.697)	7.708 (5.416)	50.15*** (8.891)	61.53*** (7.807)	88.11*** (6.382)	72.36*** (21.78)	98.17*** (23.44)
N. obs				1029			
$R\&D_SP$				0.699***	(0.138)		

Standard errors in parentheses; * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. CQIV censored quantile instrumental variable; CF residual control function residual from first stage. Omitted categories are firms located in the South operating in the Supply Dominated sector. A dummy for the year 1994 included but now shown in all regressions. Estimates performed with Stata 14 Software. FDI binary variable indicating whether the firm has established, through either brownfield or greenfield, a foreign subsidiary

Table 14 CQIV, subsample of firms operating in the North of Italy

	30	40	50	60	70	80	90
$\ln(R\&D_{t-3})$	2.496*** (0.0312)	3.663*** (0.134)	3.121*** (0.00443)	2.857*** (0.746)	3.961*** (0.291)	3.985*** (0.711)	3.081*** (1.002)
$size_{t-3}$	13.39*** (1.432)	-34.35*** (5.068)	-26.42*** (0.142)	-13.60 (25.03)	-79.66*** (20.66)	-135.3*** (53.62)	-145.5*** (61.50)
$size^2_{t-3}$	-28.92*** (0.430)	-13.52*** (1.461)	2.730*** (0.0230)	0.731 (4.239)	15.08** (6.515)	50.13*** (17.10)	49.75*** (9.627)
$Scale_{int}$	0.681* (0.356)	-1.855 (1.307)	0.437*** (0.0456)	-3.790 (6.612)	-6.823*** (3.035)	5.972 (6.873)	3.302 (11.12)
$Special_{suppl}$	3.313*** (0.454)	-12.24*** (1.699)	-4.172*** (0.0710)	-2.808 (13.64)	-8.075 (5.718)	0.427 (14.15)	-11.72 (13.94)
$Science_{based}$	-7.386*** (0.498)	-17.13*** (1.803)	-16.01*** (0.0716)	-17.48 (12.52)	-21.82*** (5.296)	-9.717 (12.47)	-10.27 (16.56)
$CF_{residual}$	-1.811*** (0.0286)	-2.916*** (0.115)	-2.455*** (0.00395)	-2.142*** (0.681)	-3.339*** (0.306)	-3.343*** (0.738)	-2.982*** (1.147)
$Const$	39.76*** (0.583)	70.41*** (2.353)	70.69*** (0.0903)	70.31*** (16.11)	101.3*** (6.098)	107.8*** (14.39)	122.1*** (18.90)
N_{obs}				776			
$R\&D_{SP}$				0.703***	(0.164)		

Standard errors in parentheses; * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. CQIV censored quantile instrumental variable; CF residual control function residual from first stage. Omitted categories are firms operating in the Supply Dominated sector. A dummy for the year 1994 included but now shown in all regressions. Estimates performed with Stata 14 Software

Table 15 CQIV, subsample of firms operating in the Centre and South of Italy

	30	40	50	60	70	80	90
$\ln(R\&D_{t-3})$	-0.112*** (0.0140)	1.803*** (0.650)	1.874*** (0.0160)	1.111 (0.789)	2.688 (2.212)	2.221*** (0.674)	1.661 (2.423)
$size_{t-3}$	39.52*** (0.400)	-29.14 (20.24)	24.21*** (0.586)	61.20* (31.61)	14.49 (107.7)	10.64 (22.75)	177.7*** (65.10)
$size^2_{t-3}$	-4.703*** (0.114)	13.32** (5.833)	0.499*** (0.164)	-9.018 (8.734)	3.686 (27.48)	4.086 (6.026)	-36.60** (17.68)
<i>Centre</i>	0.673*** (0.105)	-10.07** (3.992)	-9.287*** (0.123)	-3.633 (5.896)	11.95 (13.83)	19.02*** (4.630)	9.107 (16.39)
<i>Scale int</i>	1.106*** (0.127)	-9.948*** (3.483)	-18.16*** (0.0872)	-16.02*** (4.891)	-35.59** (17.96)	-42.61*** (4.642)	-43.07*** (15.57)
<i>Special.suppl</i>	1.732*** (0.125)	-5.707 (3.771)	-8.151*** (0.105)	-6.636 (5.969)	-0.292 (17.37)	3.245 (5.616)	1.501 (22.11)
<i>Science based</i>	0.493*** (0.172)	-13.78** (6.491)	-19.81*** (0.165)	-16.96* (9.144)	-28.58 (59.54)	-32.91** (13.71)	-160.7*** (36.03)
<i>CF residual</i>	0.147*** (0.0136)	-2.049*** (0.633)	-2.265*** (0.0161)	-1.437* (0.798)	-2.958 (2.274)	-2.583*** (0.658)	-1.546 (2.439)
<i>Const</i>	-3.304*** (0.326)	46.98*** (15.39)	52.85*** (0.404)	36.00* (19.71)	66.90 (51.52)	70.64*** (15.74)	82.82 (56.89)
<i>N. obs</i>				253			
<i>R&D_SP</i>				0.585**		(0.267)	

Standard errors in parentheses; * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. CQIV censored quantile instrumental variable; CF residual control function residual from first stage. Omitted categories are firms located in the South operating in the Supply Dominated sector. A dummy for the year 1994 included but now shown in all regressions. Estimates performed with Stata 14 Software

Table 16 CQIV, subsample of small firms (employees < 50)

	30	40	50	60	70	80	90
$\ln(R\&D_{t-3})$	0.364*** (0.00692)	0.197*** (0.0223)	0.728*** (0.262)	1.124** (0.484)	3.830*** (0.370)	3.776*** (0.521)	4.050*** (0.751)
$size_{t-3}$	9.464*** (2.393)	65.06*** (8.478)	301.2*** (127.6)	274.1 (212.8)	-642.0*** (175.3)	-759.1*** (177.2)	-2108.5*** (468.2)
$size_{t-3}^2$	142.3*** (8.472)	-51.76* (30.42)	-882.7* (449.4)	-785.7 (768.3)	2409.8*** (643.2)	2666.8*** (646.1)	31,941.8*** (5036.1)
<i>North</i>	13.54*** (0.112)	-0.914*** (0.258)	38.47*** (3.699)	30.78*** (6.613)	12.32*** (2.341)	18.58*** (3.551)	18.19*** (4.644)
<i>Centre</i>	7.923*** (0.114)	-1.463*** (0.322)	33.33*** (3.728)	36.10*** (6.887)	-3.303 (3.721)	1.148 (4.853)	-14.12** (6.403)
<i>Scale int</i>	-0.343*** (0.0605)	-0.309* (0.180)	-2.412 (1.932)	-7.696** (3.111)	-8.660*** (2.358)	-13.26*** (2.909)	-21.27*** (3.930)
<i>Special suppl</i>	8.169*** (0.0897)	14.84*** (0.315)	21.05*** (3.517)	21.47*** (6.372)	5.945 (4.852)	-1.405 (6.652)	-19.25** (9.512)
<i>Science based</i>	3.181*** (0.122)	5.316*** (0.361)	5.452 (4.481)	10.84 (7.723)	-32.25*** (5.565)	-39.67*** (7.972)	-25.96** (10.94)
<i>CF residual</i>	-0.211*** (0.00663)	-0.208*** (0.0220)	-0.315 (0.247)	-1.101** (0.460)	-3.362*** (0.389)	-3.491*** (0.541)	-3.575*** (0.797)
<i>Const</i>	-4.902*** (0.205)	4.663*** (0.709)	-23.26*** (8.051)	-2.755 (14.16)	102.9*** (10.88)	108.5*** (14.52)	158.5*** (21.07)
<i>N. obs</i>				559			
<i>R&D_SP</i>				0.71***	(0.202)		

Standard errors in parentheses; * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. CQIV censored quantile instrumental variable; CF residual control function residual from first stage. Omitted categories are firms located in the South operating in the Supply Dominated sector. A dummy for the year 1994 included but now shown in all regressions. Estimates performed with Stata 14 Software

Table 17 CQIV, subsample of medium and large firms (employees ≥ 50)

	30	40	50	60	70	80	90
$\ln(R\&D_{i-3})$	0.960*** (0.167)	4.654*** (0.0428)	8.747*** (0.00360)	9.306*** (0.125)	11.56*** (0.0336)	12.35*** (0.429)	6.522*** (1.686)
$size_{i-3}$	35.56*** (2.640)	-40.78*** (0.604)	-92.46*** (0.0560)	-100.3*** (1.927)	-138.4*** (0.426)	-149.6*** (5.420)	-88.09*** (20.97)
$size^2_{i-3}$	-26.30*** (0.579)	5.451*** (0.0964)	12.65*** (0.0117)	14.33*** (0.401)	25.34*** (0.0711)	28.56*** (0.907)	16.78*** (3.528)
<i>North</i>	24.05*** (1.365)	4.081*** (0.382)	-40.03*** (0.0331)	-32.88*** (1.110)	-92.89*** (0.294)	-91.93*** (3.804)	-47.61*** (12.97)
<i>Centre</i>	9.170*** (0.790)	0.959*** (0.278)	-19.79*** (0.0200)	-7.925*** (0.692)	-35.74*** (0.160)	-31.86*** (2.044)	-13.90*** (5.152)
<i>Scale int</i>	-4.316*** (0.470)	-16.77*** (0.136)	-18.22*** (0.00829)	-21.86*** (0.286)	-11.75*** (0.0627)	-14.12*** (0.824)	-1.170 (2.392)
<i>Special.suppl</i>	13.70*** (1.354)	-22.36*** (0.370)	-54.75*** (0.0260)	-58.78*** (0.901)	-74.53*** (0.238)	-80.40*** (3.127)	-31.43*** (13.39)
<i>Science based</i>	6.945*** (1.394)	-23.32*** (0.369)	-47.28*** (0.0298)	-46.47*** (0.984)	-62.12*** (0.254)	-67.74*** (3.279)	-30.75*** (12.12)
<i>CF residual</i>	0.0864 (0.165)	-3.668*** (0.0415)	-7.937*** (0.00362)	-8.602*** (0.125)	-11.13*** (0.0336)	-12.10*** (0.429)	-6.644*** (1.694)
<i>Const</i>	2.193 (3.180)	79.46*** (0.815)	182.3*** (0.0714)	182.6*** (2.421)	283.3*** (0.642)	296.7*** (8.287)	192.4*** (31.54)
<i>N. obs</i>				470			
<i>R&D_Sp</i>				0.157 (0.195)			

Standard errors in parentheses; * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. CQIV censored quantile instrumental variable; CF residual control function residual from first stage. Omitted categories are firms located in the South operating in the Supply Dominated sector. A dummy for the year 1994 included but now shown in all regressions. Estimates performed with Stata 14 Software

Declarations

Conflict of interest On behalf of all authors, the corresponding author states that there is no conflict of interest.

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

References

- Aghion, P., Bergeaud, A., Lequien, M., & Melitz, M. J. (2018). The impact of exports on innovation: Theory and evidence. NBER Working Papers 24600.
- Altomonte, C., & Aquilante, T. (2012). The EU-EFIGE/Brugel-Unicredit Dataset. Working Papers, Bruegel 753.
- Altomonte, C., Aquilante, T., Bekes, G., & Ottaviano, G. I. P. (2013). Internationalization and innovation of firms: Evidence and policy. *Economic Policy*, 28(76), 663–700.
- Altomonte, C., Aquilante, T., & Ottaviano, G. I. P. (2012). The triggers of competitiveness: the EFIGE cross-country report. *Blueprints, Bruegel* 738.
- Antonelli, C., & Crespi, F. (2013). The “Matthew effect” in R&D public subsidies: The Italian evidence. *Technological Forecasting and Social Change*, 80, 1523–1534.
- Antonietti, R., & Cainelli, G. (2011). The role of spatial agglomeration in a structural model of innovation, productivity and export: A firm-level analysis. *The Annals of Regional Science*, 46(3), 577–600.
- Arnold, J. M., & Hussinger, K. (2010). Exports versus FDI in German manufacturing: Firm performance and participation in international markets. *Review of International Economics*, 18(4), 595–606.
- Aw, B. Y., Roberts, M. J., & Xu, D. Y. (2011). R&D investment, exporting, and productivity dynamics. *American Economic Review*, 101(4), 1312–1344.
- Banca d'Italia. (2019). Economie regionali. L'economia delle regioni italiane. Dinamiche recenti e aspetti strutturali. N. 22.
- Basile, R. (2001). Export behavior of Italian manufacturing firms over the nineties: The role of innovation. *Research Policy*, 30(8), 1185–1201.
- Baum, C. F., Caglayan, M., & Talavera, O. (2016). R&D expenditures and geographical sales diversification. *The Manchester School*, 84(2), 197–221.
- Benfratello, L., & Razzolini, T. (2008). Firms productivity and internationalization choices: evidence for a large sample of Italian firms. In L. Piscitello & G. D. Santangelo (Eds.), *Multinationals and local competitiveness*. Franco Angeli.
- Bernard, A. B., Jensen, J. B., Redding, S. J., & Schott, P. K. (2012). The empirics of firm heterogeneity and international trade. *Annual Review of Economics*, 4(1), 283–313.
- Berry, S., & Haile, P. (2016). Identification in differentiated products markets. *Annual Review of Economics*, 8, 27–52.
- Bonfiglioli, A., Crinò, R., & Gancia, G. (2018). Betting on exports: Trade and endogenous heterogeneity. *The Economic Journal*, 128(609), 612–651.
- Bottasso, A., & Piccardo, C. (2013). Export activity and firm heterogeneity: A survey of the empirical evidence for Italy. *Economia e Politica Industriale*, 40(4), 27–61.
- Brancati, R., Marrocu, E., Romagnoli, M., & Usai, S. (2018). Innovation activities and learning processes in the crisis: Evidence from Italian export in manufacturing and services. *Industrial and Corporate Change*, 27(1), 107–130.

- Bratti, M., & Felice, G. (2012). Are exporters more likely to introduce product innovations? *The World Economy*, 35(11), 1559–1598.
- Braymen, C., Briggs, K., & Boulware, J. (2011). R&D and the export decision of new firms. *Southern Economic Journal*, 78(1), 191–210.
- Bugamelli, M., Cipollone, P., & Infante, L. (2000). L'internazionalizzazione delle imprese italiane negli anni Novanta. *Rivista Italiana Degli Economisti*, 5(3), 349–386.
- Bugamelli, M., Fabiani, S., Federico, S., Feletigh, A., Giordano, C., & Linarello, A. (2018). Back on track? A macro-micro narrative of Italian exports. *Italian Economic Journal*, 4(1), 1–31.
- Bustos, P. (2011). Trade liberalization, exports and technology upgrading: evidence on the impact of MERCOSUR on Argentinean firms. *American Economic Review*, 101(1), 304–340.
- Cassiman, B., & Golovko, E. (2011). Innovation and internationalization through exports. *Journal of International Business Studies*, 42(1), 56–75.
- Castellani, D. (2002). Export behaviour and productivity growth: evidence from Italian manufacturing firms. *Review of World Economics/Weltwirtschaftliches Archiv*, 138(4), 605–628.
- Chernozhukov, V., Fernandez-Val, I., & Kowalski, A. E. (2015). Quantile regression with censoring and endogeneity. *Journal of Econometrics*, 186(1), 201–221.
- Chernozhukov, V., Fernandez-Val, I., & Kowalski, A. E. (2019). Censored quantile instrumental-variable estimation with Stata. *Stata Journal*, 19(4), 768–781.
- Criscuolo, C., Haskel, J. E., & Slaughter, M. J. (2010). Global engagement and the innovation activities of firms. *International Journal of Industrial Organization*, 28(2), 191–202.
- D'Angelo, A. (2012). Innovation and export performance: A study of Italian high-Tech SMEs. *Journal of Management and Governance*, 16(3), 393–423.
- Damijan, J. P., Kostevc, C., & Polanec, S. (2010). From innovation to exporting or vice versa? *The World Economy*, 33(3), 374–398.
- Dhingra, S. (2013). Trading away wide brands for cheap brands. *American Economic Review*, 103(6), 2554–2584.
- Dosi, G., Grazzi, M., & Moschella, D. (2015). Technology and costs in international competitiveness: From countries and sectors to firms. *Research Policy*, 44(10), 1795–1814.
- Fasso, C. (2018). Export-led innovation: The role of export destinations. *Industrial and Corporate Change*, 27(1), 149–171.
- Fernández-Mesa, A., & Alegre, J. (2015). Entrepreneurial orientation and export intensity: Examining the interplay of organizational learning and innovation. *International Business Review*, 24(1), 148–156.
- Frazzoni, S., Mancusi, M. L., Rotondi, Z., Sobrero, M., & Vezzulli, A. (2011). Relationship with banks and access to credit for innovation and internalization of SMEs. In: *L'Europa e oltre. Banche e imprese nella nuova globalizzazione* (pp. 93–114). editore Bancaria Editrice.
- Greenaway, D., Guariglia, A., & Kneller, R. (2007). Financial factors and exporting decisions. *Journal of International Economics*, 73(2), 377–395.
- Grossman, G. M., & Helpman, E. (1991). *Innovation and growth in the global economy*. MIT Press.
- Grossman, G. M., & Helpman, E. (2015). Globalization and growth. *American Economic Review*, 105(5), 100–104.
- Haddad, M., Lim, J. J., Pancaro, C., & Saborowski, C. (2013). Trade openness reduces growth volatility when countries are well diversified. *Canadian Journal of Economics/Revue Canadienne D'économique*, 46(2), 765–790.
- Haller, S. A. (2012). Intra-firm trade, exporting, importing, and firm performance. *Canadian Journal of Economics/Revue Canadienne D'économique*, 45(4), 1397–1430.
- Halpern, L., & Muraközy, B. (2012). Innovation, productivity and exports: The case of Hungary. *Economics of Innovation and New Technology*, 21(2), 151–173.
- Impullitti, G., & Licandro, O. (2018). Trade, firm selection and innovation: The competition channel. *Economic Journal*, 128, 189–229.
- Lileeva, A., & Trefler, D. (2010). Improved access to foreign markets raises plant-level productivity... for some plants. *The Quarterly Journal of Economics*, 125(3), 1051–1099.
- Love, J. H., & Roper, S. (2015). SME innovation, exporting and growth: A review of existing evidence. *International Small Business Journal*, 33(1), 28–48.
- Mancusi, M. L., Vezzulli, A., Frazzoni, S., Rotondi, Z., & Sobrero, M. (2018). Export and innovation in small and medium enterprises: The role of concentrated bank borrowing. *Economica*, 85(337), 177–204.
- Melitz, M. J., & Ottaviano, G. I. P. (2008). Market size, trade, and productivity. *Review of Economic Studies*, 75(1), 295–316.

- Melitz, M. J., & Redding, S. J. (2014). Heterogeneous firms and trade. *Handbook of International Economics*, 4, 1–54.
- Monreal-Pérez, J., Aragón-Sánchez, A., & Sánchez-Marín, G. (2012). A longitudinal study of the relationship between export activity and innovation in the Spanish firm: The moderating role of productivity. *International Business Review*, 21(5), 862–877.
- Nassimbeni, G. (2001). Technology, innovation capacity, and the export attitude of small manufacturing firms: A logit/tobit model. *Research Policy*, 30(2), 245–262.
- OECD/Eurostat. (2018). *Oslo Manual 2018: Guidelines for collecting, reporting and using data on innovation. The measurement of scientific, technological and innovation activities* (4th ed.). OECD Publishing.
- Orts, V., & Marti, J. (2018). Firm heterogeneity and export activity of European firms: A quantile analysis. *Revista De Economía Mundial*, 49, 79–102.
- Pavitt, K. (1984). Sectoral patterns of technical change: Towards a taxonomy and a theory. *Research Policy*, 13(6), 343–373.
- Powell, D., & Wagner, J. (2014). The exporter productivity premium along the productivity distribution: Evidence from quantile regression with nonadditive firm fixed effects. *Review of World Economics*, 150(4), 763–785.
- Secchi, A., Tamagni, F., & Tomasi, C. (2016a). Financial constraints and firm exports: Accounting for heterogeneity, self-selection, and endogeneity. *Industrial and Corporate Change*, 25(5), 813–827.
- Secchi, A., Tamagni, F., & Tomasi, C. (2016b). Export price adjustments under financial constraints. *Canadian Journal of Economics/revue Canadienne D'économique*, 49(3), 1057–1085.
- Sterlacchini, A. (2001). The determinants of export performance: A firm-level study in Italian manufacturing. *Review of World Economics*, 137(3), 450–472.
- Van Beveren, I., & Vandenbussche, H. (2010). Product and process innovation and firms decision to export. *Journal of Economic Policy Reform*, 13(1), 3–24.
- Wagner, J. (1995). Exports, firm size, and firm dynamics. *Small Business Economics*, 7(1), 29–39.
- Wagner, J. (2006). Export intensity and plant characteristics: What can we learn from quantile regression? *Review of World Economics*, 142(1), 195–203.
- Wagner, J. (2012). Exports, R&D and productivity: A test of the Bustos-model with German enterprise data. *Economics Bulletin*, 32, 716–719.
- Wakelin, K. (1998). Innovation and export behavior at the firm level. *Research Policy*, 26(7–8), 829–841.
- Zhou, Y., & Song, L. (2016). International trade and R&D investment: Evidence from Chinese manufacturing firms. *China & World Economy*, 24(1), 63–84.

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.