KIEL CENTRE FOR GLOBALIZATION

KCG Working Paper

Big Data Analytics and Exports - Evidence for Manufacturing Firms from 27 EU Countries

Joachim Wagner

No. 28 | September 2023

KCG Working Paper No. 28 | Sep. 2023

[This version: September 18, 2023]

Big Data Analytics and Exports - Evidence for Manufacturing Firms from 27 EU Countries

Joachim Wagner

Abstract: The use of big data analytics (including data mining and predictive analytics) by firms can be expected to increase productivity and reduce trade costs, which should be positively related to export activities. This paper uses firm level data from the Flash Eurobarometer 486 survey conducted in February – May 2020 to investigate the link between the use of big data analytics and export activities in manufacturing enterprises from the 27 member countries of the European Union. We find that firms which use big data analytics do more often export, do more often export to various destinations all over the world, and do export to more different destinations. The estimated big data analytics premia for exports are statistically highly significant after controlling for firm size, firm age, patents, and country. Furthermore, the size of these premia can be considered to be large. Successful exporters tend to use big data analytics.

Keywords: Big data analytics, exports, firm level data, Flash Eurobarometer 486

JEL Classification: D22, F14

Joachim Wagner
Leuphana University Lueneburg and Kiel Centre
for Globalization
D-21314 Lüneburg
Germany
joachim.wagner@leuphana.de

Acknowledgement: The firm level data used in this study are taken from the Flash Eurobarometer 486 and can be downloaded free of charge after registration at http://www.gesis/eurobarometer. Stata code used to generate the empirical results reported in this note is available from the author.

1 Motivation

Digital technologies like artificial intelligence, cloud computing, the use of robots to automate processes, or big data analytics, are more and more widely applied by innovative firms. However, comprehensive empirical evidence on the links between the use of digital technologies and various dimensions of firm performance seems to be lacking. A case in point is the role of big data analytics (e.g., data mining and predictive analytics) for export activities of firms. In their comprehensive discussion of artificial intelligence (AI) and international trade Goldfarb and Trefler (2018, p. 1) state that "even to the extent that progress has been made in understanding the impact of AI, we remain largely uninformed about its international dimensions. This is to our great loss." [1]

This note contributes to the literature by looking at differences in exports between manufacturing enterprises from 27 member countries of the European Union that use or do not use big data analytics. We expect these differences to be positive for firms that use big data analytics for two reasons:

First, the use of big data analytics (including data mining and predictive analytics) by firms can be expected to increase productivity. According to a large empirical literature that uses firm level data from many different countries productivity and export activities in firms are positively related (Ferencz, López-González and García 2022, p. 12; see Wagner 2007 for a survey of the empirical literature).

Second, big data analytics can be expected to reduce trade costs (Ferencz, López-González and García 2022, p. 12). The use of data mining and predictive analytics allows firms to do comprehensive research on competitors and customers on foreign markets faster and at lower costs. Furthermore, it can help to improve predictions in future changes in consumer demand there (Meltzer 2018, p. 2).

To anticipate the most important result we find that firms which use big data analytics do more often export, do more often export to various destinations all over the world, and do export to more different destinations. The estimated big data analytics premia for exports are statistically highly significant after controlling for firm size, firm age, patents, and country. Furthermore, the size of these premia can be considered to be large. The take-home message, therefore, is that successful exporters tend to use big data analytics.

The rest of the paper is organized as follows. Section 2 introduces the data used and discusses the export activities that are looked at. Section 3 reports results from the econometric investigation. Section 4 concludes.

2 Data and discussion of variables

The firm level data used in this study are taken from the Flash Eurobarometer 486 survey conducted in February – May 2020. Note that while the data were collected at the start of the COVID-19 pandemic, the data on export activities relate to the year 2019, the year before the pandemic. We use data for firms from the 27 member states of the European Union in 2020 (i.e., firms from the UK are

[1] See Ferencz, López-González and García (2022), Goldfarb and Trefler (2018) and Meltzer (2018) for a discussion of various aspects of the relations between artificial intelligence and international trade.

are no longer included in the sample). The sample covers 2,355 firms from manufacturing industries (included in NACE section C); the numbers of firms by country are reported in the appendix table.

In the survey firms were asked in question Q23_5 whether they introduced *Big Data Analytics* (e.g., *Data Mining* and *Predictive Analytics*). Firms that answered in the affirmative are classified as users of big data analytics. Descriptive evidence is reported in Table 1, showing a share of 13.8 percent of firms with big data analytics.

In the empirical study we look at various measures of export activity of firms:[2]

First, firms were asked in question Q11_1 whether they exported any goods (or not) in 2019. Firms are classified as exporters or non-exporters based thereon. Descriptive evidence is reported in Table 1, showing a share of 64.5 percent of exporters.

[Table 1 near here]

Second, firms were asked in questions Q11_2 to Q11_8 whether they exported goods in 2019 to the following destinations: Other EU countries; other European countries outside the EU (including Russia); North America; Latin America; China; other countries from Asia and the Pacific; countries from the Middle East and Africa. Descriptive evidence is reported in Table 1, showing that 61.8 percent of firms exported to countries from the EU, while 29.2 percent exported to other European countries. The other destinations follow with shares between some 10 percent and about 16 percent. Exporters to each destination are investigated separately.

Third, from the evidence reported for exports to the seven destinations mentioned for each exporting firm the number of different destinations exported to is calculated. The share of firms by number of export destinations is reported in Table 2. Not surprisingly, most exporters serve one or two destinations only, but there are quite some firms that export to more (or even all) destinations.

[Table 2 near here]

In the empirical investigation of the link between the use of big data analytics and exports we control for three firm characteristics that are known to be positively linked with exports: firm age (measured in years, based on the answer given to question Q1), firm size (measured as the number of employees – excluding the owners - at the time of the survey; see question Q2A), and whether the firms has a patent or a patent application pending (see question Q9_6).[3] Descriptive statistics are again reported in Table 1.

[2] To the best of my knowledge (based on a Google Scholar search for "Flash Eurobarometer 486" performed on September 17, 2023) the data used in this note have not been used to investigate the links between exports and the use of bog data analytics before. Note that all measures looked at here refer to extensive margins of exports; information on intensive margins (share of exports in total sales) are not available in the data used.

[3] Given that these variables are included as control variables only, we do not discuss them in detail here. Suffice it to say that numerous empirical studies show a positive link between these firm characteristics and export performance.

Furthermore, in the empirical investigations the country of origin of the firms is controlled for by including a full set of country dummy variables.

3 Testing for big data analytics premia in export activities

To test for the difference in the types of export activities listed in section 2 between firms that do and do not use big data analytics, and to document the size of these differences, an empirical approach is applied that modifies a standard approach used in hundreds of empirical investigations on the differences between exporters and non-exporters that has been introduced by Bernard and Jensen (1995, 1999). Studies of this type use data for firms to compute so-called exporter premia, defined as the ceteris paribus percentage difference of a firm characteristic - e.g. labour productivity - between exporters and non-exporters. These premia are computed from a regression of log labour productivity on the current export status dummy and a set of control variables:

(1) $\ln LP_i = a + ß\ Export_i + c\ Control_i + e_i$

where i is the index of the firm, LP is labour productivity, Export is a dummy variable for current export status (1 if the firm exports, 0 else), Control is a vector of control variables, and e is an error term. The exporter premium, computed from the estimated coefficient ß as 100(exp(ß)-1), shows the average percentage difference between exporters and non-exporters controlling for the characteristics included in the vector Control (see Wagner (2007) for a more complete exposition of this method).

Here we look at differences between firms that do and that do not use big data analytics (instead of differences between exporters and non-exporters) and are interested in the existence and size of big data analytics premia in export activities (instead of exporter premia in various forms of firm performance like productivity). For export activities that are measured by dummy variables (the decision to export or not, and the decision to export to one of the seven export destinations listed in section 2) the empirical model is estimated by Probit instead. Therefore, (1) becomes (2)

(2) $Indicator_i = a + ß\ Big\ Data\ Analytics_i + c\ Control_i + e_i$

where i is the index of the firm, Indicator is a dummy variable for the use or not of a type of export activity, Big Data Analytics is a dummy variable for the use of big data analytics by the firm (1 if the firm uses it, 0 else), Control is a vector of control variables (that consists of measures of firm age, firm size, and patents, and dummy variables for countries), and e is an error term. The big data analytics premium is computed as the estimated average marginal effects of the big data analytics dummy variable.

For the number of export destinations, (1) becomes (3)

(3) $number_i = a + ß\ Big\ Data\ Analytics_i + c\ Control_i + e_i$

where i is the index of the firm, number is the number of export destinations, Big Data Analytics is a dummy variable for the use of big data analytics by the firm (1 if the firm uses it, 0 else), Control is a vector of control variables (that consists of measures of firm age, firm size, and patents, and dummy variables for countries), and e is an error term. The big data analytics premium is the estimated coefficient ß; it shows the average difference between firms that use and do not use big data analytics, controlling for firm age, firm size, patents, and country of origin of the firm.

Results are reported in Tables 3 - 5. The big picture that is shown is crystal clear: Firms that use big data analytics are more often exporters, do more often export to any of the different destinations, and do export to a larger number of destinations. All estimated big data analytics premia are statistically highly significant ceteris paribus after controlling for firm age, firm size, patents, and country of origin of the firms.[4] Furthermore, the size of these premia can be considered to be large – the estimated marginal effects reported in Table 3 and Table 4 are in the order of magnitude of ten percent, and from Table 5 we see that the average difference in the number of destinations exported to is +0.701 in favour of firms that use big data analytics (with an average value of 1.544 destinations for all firms).

[Tables 3 – 5 near here]

However, it is an open question (that is asked the same way when exporter premia are discussed; see Wagner 2007) whether these premia are due to self-selection of more export active firms into the use of big data analytics or whether these premia are the effect of using big data analytics.

4 Concluding remarks

This paper demonstrates that the use of big data analytics is positively related to export activities of firms from manufacturing industries. Big data analytics premia are large for all types of export activities looked at here. Does this study imply that in order to be successful in export markets, firms should use big data analytics? Or that using big data analytics will help the firms to be successful as an exporter? This is an open question (that is asked the same way when exporter premia are discussed) because we do not know whether these premia are due to self-selection of exporting firms into the use of big data analytics, or whether they are the effect of using big data analytics. This issue cannot be investigated with the cross-section data at hand. To answer this important question longitudinal data for firms are needed that cover several years and that include a sufficiently large number of firms that switch the status between using big data analytics or not over time (in both directions). To the best of my knowledge such data are not available as of today. Let's collect it!

[4] Note that all control variables have the expected positive sign and all are highly significant statistically.

References

Bernard, A. B., J. B. Jensen, and R. Z. Lawrence (1995). Exporters, Jobs, and Wages in U.S. Manufacturing: 1976-1987. *Brookings Papers on Economic Activity: Microeconomics* 67–119.

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

Ferencz, J., J. López-González, and I. O. García (2022). Artificial Intelligence and International Trade: Some Preliminary Implications. *OECD Trade Policy Paper* 260. OECD Publishing, Paris.

Goldfarb, A. and D. Trefler (2018). AI and International Trade. *National Bureau of Economic Research Working Paper* 24254.

Meltzer, J. P. (2018). The impact of artificial intelligence on international trade. *Center for Technology Innovation at Brookings.*

Wagner, J. (2007). Exports and Productivity: A survey of the evidence from firm level data. *The World Economy* 30 (1): 5–32.

Table 1: Descriptive statistics

Variable	Mean	Std. Dev.	Min	Max
Big Data Analytics (Dummy; 1 = yes)	0.138	0.345	0	1
Exporter (Dummy; 1 = yes)	0.645	0.478	0	1
Export Destination (Dummy-Variables; 1 = yes)				
- EU-countries	0.618	0.486	0	1
- Other Europe	0.292	0.455	0	1
- North America	0.157	0.364	0	1
- Latin America	0.099	0.298	0	1
- China	0.109	0.311	0	1
- Other Asia	0.138	0.345	0	1
- Middle East, Africa	0.132	0.339	0	1
Number of Export Destinations	1.544	1.857	0	7
Firm Age (years)	29.03	23.43	0	170
No. of Employees	91.63	269.11	1	5000
Patent (Dummy; 1 = yes)	0.120	0.325	0	1
No. of Firms in Sample	2,355			

Source: Own calculation based on data from Flash Eurobarometer 486.

Table 2: Share of Firms by Number of Export Destinations

Number of Export Destinations	Number of Firms	Percent
0	835	35.46
1	700	29.72
2	338	14.35
3	150	6.37
4	100	4.25
5	73	3.1
6	68	2.89
7	91	3.86
Total	2,355	100

Source: Own calculation based on data from Flash Eurobarometer 486

Table 3: Estimation results, Part I: Exporter vs. Non-Exporter

Dependent variable: Exporter (Dummy; 1 = yes)
Method: Probit

Variable	Coefficient	p-value	Marginal effect	p-value
Big Data Analytics (Dummy; 1 = yes)	0.386	0.000	0.122	0.000
Firm Age (years)	0.0045	0.001		
No. of employees	0.0010	0.000		
Patent (Dummy; 1 = yes)	0.720	0.000		
Country (26 Dummy variables)	included			
Constant	included			
No. of firms	2,355			

Source: Own calculation based on data from Flash Eurobarometer 486

Table 4: Estimation results, Part II: Exporter by Destination

Dependent variable: Exporter (Dummy; 1 = yes)

Method: Probit

Variable	Coefficient	p-value	Marginal effect	p-value
EU countries				
Big Data Analytics	0.395	0.000	0.129	0.000
Firm Age	0.005	0.001		
No. of employees	0.001	0.000		
Patent	0.730	0.000		
Other Europe				
Big Data Analytics	0.511	0.000	0.163	0.000
Firm Age	0.007	0.000		
No. of employees	0.0006	0.000		
Patent	0.705	0.000		
North America				
Big Data Analytics	0.409	0.000	0.095	0.000
Firm Age	0.006	0.000		
No. of employees	0.0004	0.000		
Patent	0.751	0.000		
Latin America				
Big Data Analytics	0.523	0.000	0.097	0.000
Firm Age	0.005	0.001		
No. of employees	0.0005	0.000		
Patent	0.596	0.000		
China				
Big Data Analytics	0.532	0.000	0.100	0.000
Firm Age	0.007	0.000		
No. of employees	0.0005	0.000		
Patent	0.615	0.000		
Other Asia				
Big Data Analytics	0.501	0.000	0.109	0.000
Firm Age	0.006	0.000		
No. of employees	0.0006	0.000		
Patent	0.654	0.000		
Middle East, Africa				
Big Data Analytics	0.529	0.000	0.113	0.000
Firm Age	0.007	0.000		
No. of employees	0.0005	0.000		
Patent	0.664	0.000		
No. of firms	2,355			

Note: All empirical models include 26 country dummy variables plus a constant

Source: Own calculations based on data from Flash Eurobarometer 486

Table 5: Estimation results, Part III: Number of Export Destinations

Dependent variable: Number of export destinations for exporters

Method: OLS

Variable	Coefficient	p-value
Big Data Analytics (Dummy; 1 = yes)	0.717	0.000
Firm Age (years)	0.011	0.000
No. of employees	0.0007	0.000
Patent (Dummy; 1 = yes)	0.956	0.000
Country (26 Dummy variables)	included	
Constant	included	
R-squared	0.278	
No. of firms	2,355	

Note: Estimated standard errors are clustered at the level of the 27 countries

Source: Own calculations based on data from Flash Eurobarometer 486

Appendix

Appendix: Number of Firms by Country

Country	Number of Firms	Percent
Austria	86	3.65
Belgium	81	3.44
Bulgaria	97	4.12
Cyprus	33	1.40
Czech Republic	94	3.99
Germany	74	3.14
Denmark	75	3.18
Estonia	99	4.20
Spain	137	5.82
Finland	88	3.74
France	101	4.29
Greece	111	4.71
Croatia	136	5.77
Hungary	117	4.97
Ireland	30	1.27
Italy	149	6.33
Lithuania	64	2.72
Luxembourg	25	1.06
Latvia	75	3.18
Malta	21	0.89
Netherlands	55	2.34
Poland	101	4.29
Portugal	93	3.95
Romania	102	4.33
Sweden	75	3.18
Slovenia	130	5.52
Slovakia	106	4.50
Total	2,355	100.00

Source: Own calculations based on data from Flash Eurobarometer 486