Climbing the Rungs of the Quality Ladder: FDI and Domestic Exporters in Romania

Online appendix

Matej Bajgar and Beata Javorcik

May 25, 2019

A Construction of variables

A.1 Unit values

We calculate unit value as the monetary value of an export flow divided by its physical quantity. The key question is which measure of physical quantity to use. We proceed as follows. First, we drop observations with export values equal to zero or with both measures of physical quantity equal to zero. Second, we use supplementary units as measure of physical quantity for all products which have supplementary units available for over 90% of the observations, and we use kilograms as measure of physical quantity for the remaining observations. Finally, we drop observations where the measure of physical quantity chosen for the given product is equal to zero. Using this procedure, we measure physical quantity in supplementary units for 34% observations and in kilograms for 66% observations, and we drop unit values for 0.2% observations.

Some values of unit values and their fluctuations within firm and product over time are too extreme to be likely to be true. We deal with the suspected outliers in three steps. First, to eliminate extreme levels of unit values, we demean each log unit value by the corresponding product-year-specific mean and then drop the 1% smallest and 1% largest observations. Second, to eliminate extreme unit value fluctuations over time within firm-product-destination combinations, we drop the 1% firm-product-destination combinations with highest variation of unit values over time. Finally, to eliminate extreme unit value year-to-year-changes, we drop the 1% largest year-to-year changes

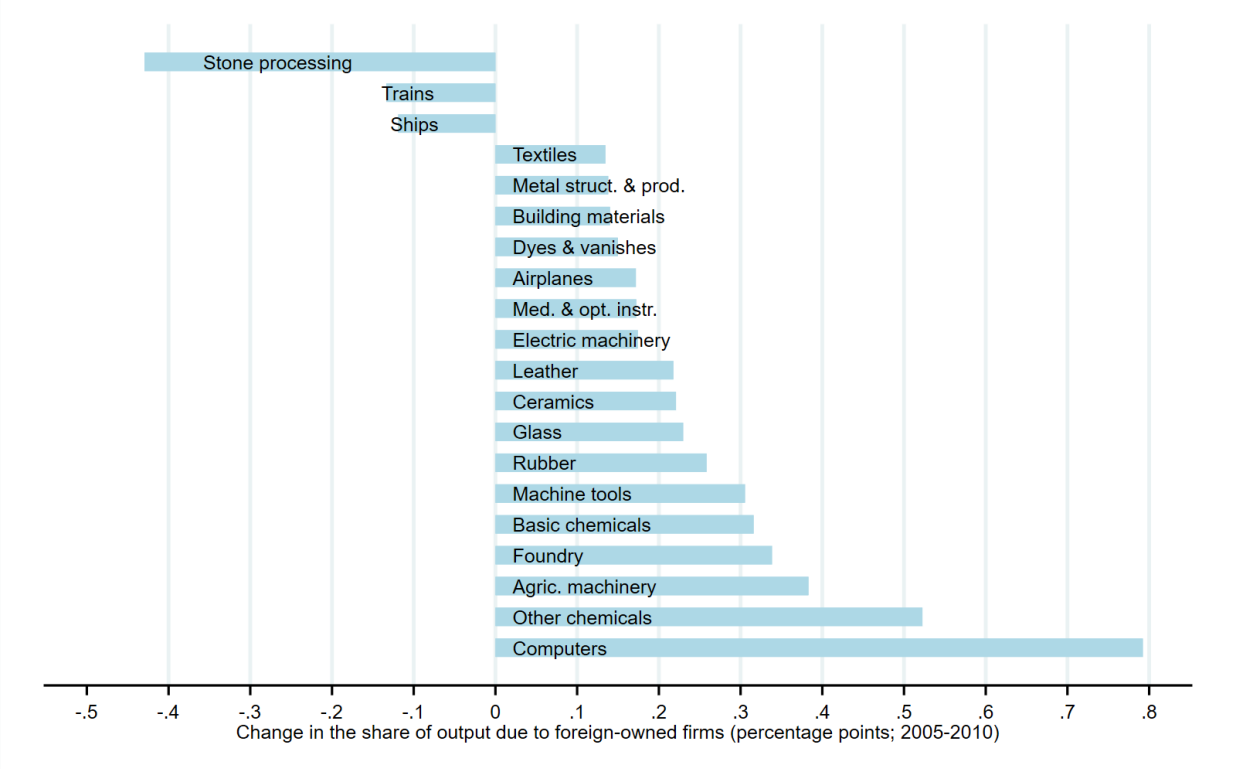
within a firm-product-destination combination.

A.2 Industry classification

The SBS data define the main activity of each firm in terms of 4-digit NACE industries. Data for years 2005-2007 use NACE (rev. 1.1) and data for years 2008-2010 NACE (rev. 2). In order to make the SBS data compatible with our input-output table, we first convert all observations to NACE rev. 1.1. There is no unambiguous concordance between rev. 1.1 and rev. 2, so we derive our own concordance from the data so as to maximize the continuity in NACE within firms across time. For each observation from 2008 or later, we define pre-revision NACE as the NACE of the same firm in 2007 (or earlier if we do not observe the firm in 2007). We create a concordance where each 4-digit NACE rev. 2 code corresponds to the most common pre-revision NACE among observations with the same 4-digit NACE rev. 2 code. Then, we convert NACE for years 2008 and later to NACE rev. 1.1 using this concordance. Validity of this method is supported by the fact that on the 2-digit level where our subsequent analysis takes place, the number of firms which change NACE between 2007 and 2008 is very similar to the corresponding numbers for other years. Once we know the 4-digit NACE (rev 1.1) for each firm in each year, we use a concordance table to convert the NACE industries into the Romanian industrial classification in which the input-output table is defined. Finally, to ensure that variation over time in FDI presence in different industries is not driven by firms changing their reported industrial affiliation, we make the affiliation time-invariant by setting it to the mode value for each firm. This procedure modifies NACE for less than 4% of firm-year observations.

B Additional tables and figures

Figure B.1
Industries with the largest changes in foreign presence (2005-2010)



Notes: Only industries with an increase or a decrease of more than 10 percentage points are included.

Table B.1

FDI inflows into central and eastern European countries (percent of GDP)

	2003	2004	2005	2006	2007	2008	2009	2010	2011	2012	Average
Bulgaria	10.1	13.4	13.6	23.5	29.4	19.0	7.0	3.2	3.4	3.7	15.9
Czech Republic	2.2	4.4	9.0	3.7	5.8	2.9	1.5	3.1	1.1	5.4	4.3
Estonia	9.4	8.0	20.6	10.7	12.4	7.3	9.6	8.4	1.2	6.7	11.5
Hungary	2.6	4.2	7.0	6.1	2.9	4.1	1.6	1.7	4.2	10.6	3.9
Latvia	2.7	4.6	4.4	8.4	8.1	3.8	0.4	1.6	5.1	3.5	4.4
Lithuania	1.0	3.4	3.9	6.0	5.1	4.1	-0.0	2.2	3.4	2.0	3.6
Poland	2.1	5.1	3.4	5.7	5.5	2.8	3.0	3.0	3.7	0.7	3.9
Romania	3.7	8.5	6.5	9.3	5.8	6.8	2.9	1.8	1.3	1.3	5.5
Slovakia	8.9	9.6	6.5	10.4	5.4	5.2	-0.0	2.0	2.2	3.1	4.9
Slovenia	1.0	2.4	1.6	1.7	3.2	3.6	-1.3	0.8	2.0	0.3	1.6

UNCTAD. Underlined years appear in our data. Column Average represents the average over that period, 2005-2010.

Table B.2

Exported product quality and FDI: Samples from differenced specifications

	(1) Full sample	(2) Sample D1	(3) Sample D2	(4) Sample D3	(5) Sample D4
Upstream FDI (s,t-1)	0.463***	0.373***	0.376***	0.510***	0.825***
	(0.113)	(0.105)	(0.122)	(0.135)	(0.165)
Downstream FDI (s,t-1)	0.594**	0.373	0.648**	0.528	0.741
	(0.293)	(0.275)	(0.320)	(0.351)	(0.497)
Own FDI (s,t-1)	0.168	0.160	0.176**	0.388***	0.516***
	(0.110)	(0.110)	(0.084)	(0.110)	(0.134)
R-squared	0.068	0.061	0.089	0.126	0.187
N	146760	74907	49823	33532	18562
Upstream = 0.463 (p-value)		0.393	0.476	0.727	0.029
Upstream FDI (s,t-1)	0.543***	0.437***	0.483***	0.496***	0.976***
	(0.137)	(0.127)	(0.145)	(0.190)	(0.248)
Downstream FDI (s,t-1)	0.645*	0.365	1.022**	0.717	1.004
	(0.367)	(0.338)	(0.428)	(0.477)	(0.633)
Own FDI (s,t-1)	0.259	0.194	0.268	0.385*	0.201
	(0.174)	(0.166)	(0.168)	(0.200)	(0.209)
R-squared	0.018	0.016	0.022	0.026	0.049
N	146760	74907	49823	33532	18562
Upstream = 0.543 (p-value)		0.404	0.679	0.804	0.083

*Notes: *** 1%, ** 5%, * 10%.* Standard errors, reported in parentheses, have been clustered on industry-year combinations. In panel A, the dependent variable is the logarithm of export unit values. In panel B, the dependent variable is the quality estimated as in Khandelwal *et al.* (2013) assuming 2-digit-HS-specific σ_s based on estimates of Broda and Weinstein (2006). Observations are defined at the firm-product-destination-year level. The specification includes firm-product-destination and region-year fixed effects and industry-region linear time trends. Columns 2-5 respectively restrict the sample to the observations that serve for calculating the 1st-4th differences used to obtain the estimates reported in columns 2-5 of Table 2.

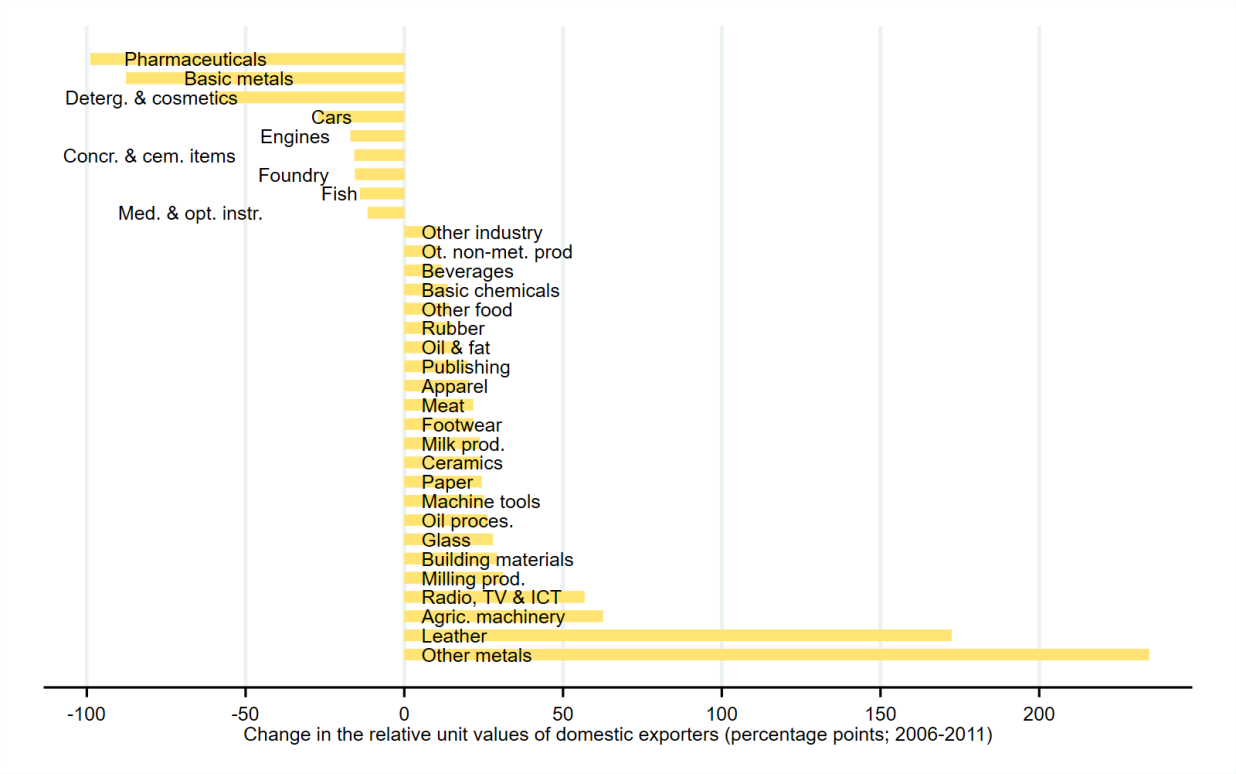
Figure B.2

Evolution of foreign presence in individual industries (2005-2010)

Notes: The x-axis represents time and the y-axis represents the share of output due to foreign-owned firms.

Figure B.3

Industries with largest changes in unit values relative to EU15 (2006-2011)



Notes: Bars represent medians (across products exported by each industry) of the ratio of unit values of exports by domestically owned Romanian firms to unit values of exports by EU15 countries. Information on Romanian exports comes from Romanian customs data described in Section 2 of the paper; information on EU15 exports comes from Eurostat. Each product is assigned to the industry with the highest number of exporter-destination pairs for that product. Only 8-digit-CN products exported by both Romania and EU15 in all years and only industries with increase or decrease of more than 10 percentage points are included.

Table B.3
Exported product quality and FDI: Dynamic specification

	UV		Khandelwal et al. (2013)	
	(1)	(2)	(3)	(4)
	FE	GMM	FE	GMM
L.Quality	-0.095***	0.198***	-0.103***	0.106
	(0.027)	(0.052)	(0.025)	(0.075)
Upstream FDI (s,t-1)	0.333***	0.359***	0.397***	0.370**
	(0.091)	(0.106)	(0.122)	(0.147)
Downstream FDI (s,t-1)	0.492*	0.369	0.179	-0.029
	(0.256)	(0.262)	(0.340)	(0.402)
Own FDI (s,t-1)	0.053	0.057	0.046	-0.016
	(0.083)	(0.091)	(0.122)	(0.128)
R-squared	0.079		0.025	
N	62851	30723	62851	30723
Hansen		0.99		0.45

*Notes: *** 1%, ** 5%, * 10%.* Standard errors, reported in parentheses, have been clustered on industry-year combinations. In columns 1 and 2, the dependent variable is the logarithm of export unit values. In columns 3 and 4, the dependent variable is quality estimated as in Khandelwal *et al.* (2013) assuming HS 2-digit-specific σ_s based on estimates of Broda and Weinstein (2006); see Section 2 of the paper for more detail. Columns 1 and 3 are based on a dynamic fixed-effects estimator. Columns 2 and 4 are based on a difference GMM estimator by Arellano and Bond (1991) using the orthogonal deviations transformation (Arellano and Bover, 1995) and the two-step procedure, with standard errors incorporating the Windmeijer (2005) finite-sample correction. Following Roodman (2009), instrument count is reduced by collapsing the instrument matrix. Observations are defined at the firm-product-destination-year level. Both specifications include firm-product-destination and region-year fixed effects.

References

Arellano, M. and Bond, S. (1991). 'Some tests of specification for panel data: Monte Carlo evidence and an application to employment equations', *Review of Economic Studies*, vol. 58(2), pp. 277–97.

Arellano, M. and Bover, O. (1995). 'Another look at the instrumental variable estimation of error-components models', *Journal of Econometrics*, vol. 68(1), pp. 29 – 51.

Broda, C. and Weinstein, D.E. (2006). 'Globalization and the gains from variety', *The Quarterly Journal of Economics*, vol. 121(2), p. 541.

Khandelwal, A.K., Schott, P.K. and Wei, S.J. (2013). 'Trade liberalization and embedded institutional reform: evidence from Chinese exporters', *American Economic Review*, vol. 103(6), pp. 2169–95.

Roodman, D. (2009). 'How to do xtabond2: an introduction to difference and system GMM in Stata', *Stata Journal*, vol. 9(1), pp. 86–136.

Windmeijer, F. (2005). 'A finite sample correction for the variance of linear efficient two-step GMM estimators', *Journal of Econometrics*, vol. 126(1), pp. 25 – 51.