8 Modelling orthodox and non-orthodox determinants of foreign trade

The number of studies devoted to the foreign trade of regions (that is, subnational spatial units) is increasing (Batabyal and Nijkamp, 2015; Bentivogli et al., 2018; Cassey et al., 2016; Cassey and Schmeiser, 2013; Dühr, 2005; Koenig, Mayneris, and Poncet, 2010; Kounetas and Napolitano, 2018; Nazarczuk and Umiński, 2018a). It, in particular, applies to medium and large-sized countries, where the extent of regional heterogeneity is large. One could give several reasons that drive attention to the regional level of analysis. They have recently been exemplified by Brodzicki and Umiński (2017) as treatment of a region as an SOE, increasing interdependence of regional economies due to the persistence of globalisation processes and an increasing role of the subsidiarity principle, transferring more power and responsibilities in governance to regional authorities. In related disciplines such as international business, researchers have begun to recognize the weakness of using country border-based data, as they may contain systematic biases and because important benefits and costs may be overlooked (Mudambi et al., 2018). In international trade studies, we consider as an additional reason in favour of using regional trade statistics – the accuracy of country bilateral trade data (Bhagwati, 1974; Makhoul and Otterstrom, 1998).

This is particularly important for the analysis of the determinants of trade flows. Between the causes of discrepancies in international trade statistics, there are structural differences in the compilation and statistical criteria. This is important because, as pointed out by Makhoul and Otterstrom (1998, p. 1604), “structural problems may reveal themselves as inconsistencies among countries with regard to trade coverage, recording of values, an indication of trading partners, time lags, variations in exchange rates, and changes in data collection systems”. In this research, focusing on (subnational) regions for two specific EU countries, we avoid structural differences, and we improve the accuracy of the data.

Our main contribution to the body of regional economics literature is that we go beyond standard determinants of trade intensity and aim at the identification of the impact of several new variables to explain regional trade, which we refer to as non-orthodox factors. To do so, we focus on regional trade statistics from regions in Poland and Spain to analyse their determinants of bilateral trade flows. In contrast to previous studies that construct the regional trade statistics for analysis,⁸ we focus on the regional trade statistics sourced from customs agencies.

⁸ For example Thissen, Lankhuizen, and Jonkeren (2015) construct a simulated trade matrix for trade between European regions.

Two countries are the subjects of the analysis, which is justified for several reasons. Poland and Spain are MS of the EU, represent similar levels of development, can be regarded as peripheral countries of the EU, having similar numbers and sizes of NUTS 2 regions. Spain has been a member of the EU since 1986 and represents old MS, while Poland entered the EU in 2004 and belongs to the group of new MS. Poland enjoys derogation from participation in the eurozone, Spain introduced the euro. Several differences between Poland and Spain can be identified. They relate to institutional quality and to historical factors that may still influence the patterns of regions' trade (Brodzicki, Jurkiewicz, Márquez-Ramos and Umiński, 2019).

We identify the determinants of the intensity of foreign trade of the Spanish and Polish NUTS 2 regions with all existing trade partners on a bilateral basis over the period 2005-2014. We do so within the framework of an augmented panel gravity model. Recently, Márquez-Ramos (2016b, 2016c) stated that the regional (subnational) context interacts with the national and the supranational context in a gravity model approach. In this research, we consider (orthodox and non-orthodox) factors determining bilateral trade flows at three geographic levels: regional, national, and supranational.

We estimate a gravity model and utilize a solid method of estimation to obtain unbiased estimates, that is, the Poisson Pseudo Maximum Likelihood (PPML) estimator (Santos Silva and Tenreyro, 2006) applied to a semi-mixed effects model. Methodologically speaking, in this approach, the pair effects are random to control for unobserved cross-sectional heterogeneity, and following Lombardía and Sperlich (2012), an extra fixed clustering variable is introduced. This furthermore applies that the gravity equation does not have to be log-linearized; thus there is no need for elimination of zero trade flows or the so-called zero-adjustment.

The application of the gravity approach in the regional analysis is driven by two main reasons. On the one hand, the regional perspective allows the capture of new determinants of trade flows, which cannot be assessed at the country level. These are, for instance, historical circumstances and the nature of the border effect, as shown by Brodzicki (2017b), infrastructural issues (Alamá-Sabater, Márquez-Ramos, Navarro-Azorín, and Suárez-Burguet, 2015) or the role of metropolises. On the other hand, an inquiry into the nature, performance, and competitiveness of regions' exports – possible with the use of a gravity approach – bears important information for regional authorities or any other agents engaged in regional development issues, e.g., related to export promotion at the regional level (Gil-Pareja, Llorca-Vivero, Martínez-Serrano, and Requena-Silvente, 2015).

We contribute to the body of empirical trade literature by bringing the analysis of trade in the regional context beyond a single country and verifying the hypotheses related to determinants of trade for Spain and Poland in order to check their universal nature.

8.1 Hypotheses

The research sets, as a principal aim, the identification of the role of relevant orthodox and non-orthodox determinants in the region’s foreign trade. They are presented in Table 8.1.

Table 8.1: The set of the testable hypotheses

Hypothesis	Description	Justification
1	The similarity in trading partners’ level of development (GDP per capita or TFP) positively contributes to the intensity of trade relations	According to the Linder hypothesis (Linder, 1961), we expect partners to trade more intensively, the more similar their level of development is. The similarity in TFP reflects correspondence in the supply side, particularly in the levels of technological sophistication and overall productivity. Similarity measured by GDP per capita reflects the proximity at the demand side.
2	Direct geographical neighbourhood (adjacency) has a positive impact on bilateral trade	The justification for this hypothesis is in the gravity model itself (adjacency is the closest kind of proximity). Also, as exporting is a risky and difficult activity for a happy few (Mayer and Ottaviano, 2008; Melitz and Redding, 2014), the adjacency reflects the presence of common land borders and the likelihood of the greater number of border crossings, facilitating trade. Also, the information gap about the partners in trade is lower, which reduces the overall risk and entry costs associated with exporting. A thorough interpretation of the adjacency role was provided by Hummels (1999), who associates the adjacency with lower direct trade costs.
3	Metropolitan character of a region positively contributes to the export intensity	Metropolises are perceived as nodes of the globalized economy. They contribute to foreign market risk reduction due to agglomeration externalities, associated with information spillovers. According to NEG and new trade theory, metropolises are characterised by a greater scale of economic activity, more diversified economic structure, and concentration of tradable sectors (Behrens, 2005) – home market effect. For Poland, Brodzicki and Umiński (2017) find that metropolises have a statistically significant and positive influence on foreign trade flows. However, if entrepreneurs’ (corporations) preferences are strongly focused on cost reduction (costs in metropolises are usually higher than elsewhere), exporters may prefer non-metropolitan locations.

continued **Table 8.1:** The set of the testable hypotheses

Hypothesis	Description	Justification
4	Quality of institutions positively affects bilateral trade	Institutions matter for regional development (Rodríguez-Pose et al., 2013). Their impact in the European context is more important than the first-nature geographic factors (Ketterer and Rodríguez-Pose, 2018). Álvarez, Rodríguez-Pose, and Zofío (2018) show their significance for shaping international trade flows between countries. Thus, we can postulate that greater institutional quality of both the region and its trade partner intensifies commercial links through the reduction of risk of business contracts. As shown by Do and Levchenko (2006), better contract enforcement, superior property rights, and better investors' protections enable agents to overcome frictions. Better institutions also reduce vulnerability to shocks, allowing for greater adaptive capacity (Bradley et al., 2017; Davoudi, Zauha, and Brooks, 2016).
5	Historical path-dependency affects the existing trade relations in the region-country framework	According to Eichengreen and Irwin (1998), persistence in trade relations stems from historical bonds between trading partners. Persistence reflects longstanding ties, i.e., for former Spanish (Martinez-Galarraga, 2014), Portuguese or French colonies (Head and Mayer, 2014) or former partitions of Poland (Brodzicki and Umiński, 2017). The persistence does not always have to facilitate trade – adverse effects due to past animosities can persist over prolonged periods as well.
6	FDI inflows have an impact on the intensity of bilateral trade	FDI can exert a substitution or complementary effect on trade, depending on the character of FDI. The substitutive relation between FDI and exports is expected if the main aim of FDI is tariff jumping (Blomström et al., 2002). In case of substantial differences in factor endowments between countries, Helpman (1984) and Helpman and Krugman (1985) predict that FDI flows create complementary trade. According to Jensen (2002), the positive effect of FDI can unambiguously be expected only in case of resource-seeking, vertical FDI. Estrin et al. (2008) point out that the nexus between FDI and trade is strongly dependent on the relative position of the subsidiaries within the MNE, which influences the probability of exporting and its intensity.

Source: Own elaboration.

8.2 Empirical strategy

It seems to be a consensus that a gravity model “works well” in its baseline specification, including distance and market size of trading partners. A thorough overview

of the evolution of the gravity concept, the theoretical underpinnings of the gravity equation, and its usage were presented by Head and Mayer (2014). Looking from this perspective, our approach corresponds with the supply-side derivation of the gravity model, which rests on the Ricardian comparative advantages, perceiving regions as parts of the global economy. Furthermore, as shown by Kepaptsoglou et al. (2010), the selection of other (non-orthodox) independent variables varies with the character of the research. In most empirical studies, the gravity approach is used to estimate, for instance, the consequence of monetary integration, customs unions implemented, export promotion carried on, etc., and the estimation strategy envisages the use of relatively few explanatory variables. This is due, basically, to the extended use of a strategy that consists of including several sets of fixed effects to avoid endogeneity in the gravity model at the country-level (Márquez-Ramos, 2016b).

Unfortunately, with this strategy, the identification of a number of key variables is not possible. Although there have been some methodological advances that allow identifying variables that by using a full battery of fixed effects could not be identified (Baier and Bergstrand, 2009; Heid, Larch, and Yotov, 2017), the approach we propose here is different. We aim at the identification of non-orthodox factors that influence trade with the purpose to help regions and actors responsible for export promotion at the regional level to increase export (or total trade) capabilities. This approach is consistent with the idea of using a region-to-country dataset to take into account the full structure of the gravity model. It can be treated as an alternative to the wide use of country-to-country statistics and several sets of fixed effects to control for potential omitted variable biases (Márquez-Ramos, 2016b). At the same time, this approach allows not only considering supranational factors (e.g., trade agreements), but also the region and country characteristics.

Our empirical strategy is the following. We construct a basic specification of an augmented gravity model and, in steps, add additional orthodox and non-orthodox variables to test the hypotheses. The analysis is conducted separately for total trade and exports of regions. Furthermore, we estimate the model jointly for all Spanish and Polish regions and then disjointedly, in a comparative manner, to identify discrepancies of reaction to various orthodox and non-orthodox factors investigated.

The factors included in the empirical analysis represent standard (or orthodox) determinants, such as the role of distance and size; and non-orthodox determinants, such as the differences in the level of development and technological sophistication, as measured by TFP, exchange rate variability and factors of structural character: metropolitan nature of some regions, cultural and institutional variation or trade-persistence (long-term nature historical factors, such as former colonies or partitions), quality of institutions or the impact of FDI.

To obtain unbiased results and to avoid heteroskedasticity and the zero trade flows adjustment, we utilize novel empirical approaches. An important problem in

modelling gravity and in particular in the case of regions' trade is the phenomenon of zero trade flows. The distribution of zero-trade flows is not random. They occur predominantly in relations with small, remote trade partners and are more likely when the analysis is applied to the region-to-country framework. The typical log-linearized model in the case of zero trade flows requires the necessary adjustments or treatments as missing data⁹.

Specifically, we use the estimation of a semi-mixed effect model with the PPML estimator. The estimator can be applied to a nonlinear multiplicative model (Santos Silva and Tenreyro, 2011; Westerlund and Wilhelmsson, 2011).

The functional form of the gravity model is based on the PPML approach. Time fixed effects are included. Trade-partners specific effects η_{ij} are estimated as random effects. We account for the importance of multilateral trade resistance by the inclusion of pair-specific effects, which enables us to capture the overall level of trade protection (Anderson and van Wincoop, 2003).

The PPML approach in a gravity context was extended by Savasci (2011), who suggested the use of PPML, where the pair effects were introduced as random, to control for unobserved cross-sectional heterogeneity. The problem in the estimation of the proposed version of the model is potential misspecification, due to the assumption of independence with the random effects. As a solution to the problem, Proença, Sperlich, and Savaşçı (2015) proposed a method which relaxes the strict assumption of the random effects model but entails more restrictions than the standard fixed-effects model, that is through the introduction of a fixed clustering variable (a semi-mixed effects model, SMEM). SMEM includes a nonparametric proxy in a parametric multilevel model to filter out potential mean dependency between the random effects and covariates (Lombardía and Sperlich, 2012). The inclusion of a nonparametric filter does not degrade the estimation of the effects of time-invariant variables – that is time, and other fixed effects can be included in the specification of the model. The resulting model is a semi mixed-effects model, in the sense that it still has a residual random effects component. The SMEM approach has recently been applied by Brodzicki and Umiński (2017) for the case of foreign trade of Polish regions. The SMEM model has the following form (Proença et al., 2015):

$$T_{ijt} = \exp[\ln\alpha_0 + \beta_1 \log(Y_{it}) + \beta_2 \log(Y_{jt}) + \gamma \log(Z_{ij}) + \alpha_t + \eta_{ij}] \varepsilon_{ijt}$$

where Z_{ij} is the distance between the trade partners and η_{ij} are random effects.

⁹ The solution could be the application of PPML estimators. Santos Silva and Tenreyro (2006) argued that the standard logarithmic transformation in the gravity model might not be an appropriate approach to estimate elasticities. It is because trade models containing multiplicative errors do not satisfy the assumption of the homoskedasticity of the error term due to the dependency between the error term of the transformed log-linear model and the regressors.

The dependent variables in our empirical model are the value of total trade (*ttrade*) in thousands of EUR between a given region *i* and a given trade partner *j* in a year *t* and the value of exports (*exports*) in thousands of EUR from a given region *i* to a given trade partner *j* in a year *t*.

In the baseline specification, we control for the physical distance between trade partners, as measured by the log of distance in kilometres (*ln_distance*) and their sizes, as measured by the log of real GDP (*ln_rGDP_r*; *ln_rGDP_c*) in million USD. We control for the similarity in the levels of economic development, as measured by the log of absolute difference in real GDP per capita between a given region, and a given trade partner (*diff_y*). Because Hypothesis 1 states that trading partners trade more intensively the more similar their level of development is; thus, the coefficient on the variable is expected to be negative.

The data for regional institutional quality from the University of Gothenburg is available for two years only. To obtain a panel, we acknowledged their regional variation within Poland and Spain and took into account their evolution for both countries from the Global Governance Indicators database (*r_inst_quality*). Furthermore, we performed principal component analysis on the spectrum of institutional quality variable¹⁰, provided in the Global Governance Indicators dataset to obtain a single institutional proxy for countries (*c_inst_quality*).

In the calculation of TFP for regions, we utilized the following approach. We took the data for TFP for Poland and Spain from PWT 9.0. Acknowledging that TFP is a prime source of variation in real GDP per capita, we took information on the value of real GDP per capita in a region, relative to country mean as a rough approximation of a region's TFP. Using the information on regions TFP and trade partners TFP we calculated the difference in TFP measure (*diff_tfp*) as the log of the absolute difference between the two values. In order to address potential collinearity with the log of GDP per capita, we drop the difference in GDP per capita in model specifications incorporating differences in TFP levels.

To account for the metropolitan status of a region we consulted the ESPON (Dühr, 2005) study on metropolitan areas in the EU, constructing a dummy variable for metropolitan areas (*metro*) and in addition, we took into account the Metropolitan European Growth Areas (MEGA) classification thus creating dummy variables for MEGA 1, 2, 3 and 4 regions (*mega1*, *mega2*, *mega3*, *mega4*, respectively). MEGA 1 currently have a global impact, while MEGA 4 are potential metropolises with European significance. Table 8.2. describes the variables used together with the character of the variable (orthodox versus non-orthodox) as well as the geographical scale of spatial variation (regional, national, or supranational) and the expected direction of impact.

10 These included: voice and accountability, political stability and absence of violence, government effectiveness, regulatory quality, rule of law as well as control of corruption.

Table 8.2: Description of the key variables

Variable	Description	Character of the variable	Geographic scale of spatial variation	Presumed impact on the dependent variable	Hypothesis number
ln_distance	Log of distance in km between region's and trade partner's capital cities	orthodox	regional and national	negative	*
ln_rgdp_c	log of real GDP of the partner country	orthodox	national	positive	*
ln_rgdp_r	Log of real GDP of a region	orthodox	regional	positive	*
diff_y	Difference in the level of development as measured by the log of absolute difference in real GDP per capita	orthodox	regional and national	negative	H1
diff_tfp	Difference in the level of technological sophistication as measured by the log of absolute difference in the TFP levels	non-orthodox	regional and national	negative	H3
adjacency	Adjacency between a region and a partner country	orthodox	regional	positive	H2
access_2_sea	Access to sea by a region	orthodox	regional	positive	*
c_land-locked	Landlockedness of the partner country – CEPII	orthodox	national	negative	*
mega1	Region having MEGA 1 status in accordance with the ESPON MEGA classification	non-orthodox	regional	positive	H3
mega3	Region having MEGA 3 status in accordance with the ESPON MEGA classification	non-orthodox	regional	positive	H3
mega4	Region having MEGA 4 status in accordance with the ESPON MEGA classification	non-orthodox	regional	positive	H3
metro	Metropolitan status of the region in accordance with the ESPON MEGA classification (MEGA 1, 2, 3 or 4)	non-orthodox	regional	positive	H3
eu28	Dummy variable for partner country in the EU	orthodox	national and supranational	positive	*
r_inst_qual	Regional institutional quality – QoG	non-orthodox	regional	positive	H4

continued **Table 8.2:** Description of the key variables

Variable	Description	Character of the variable	Geographic scale of spatial variation	Presumed impact on the dependent variable	Hypothesis number
c_inst_quality	Partner country institutional quality – result of principal component analysis on World Bank governance indicators	non-orthodox	national	positive	H4
history	Dummy variable for common historical ties between a region and a given trade partner (including partitions and colonies)	non-orthodox	regional and national	positive/negative	H5
pl_p_common	Dummy for a region in a common partition with a foreign country	non-orthodox	regional	positive/negative	H5
pl_p_deu	Dummy for Polish region in the former Prussian partition	non-orthodox	regional	positive/negative	H5
pl_p_rus	Dummy for Polish region in the former Russian partition	non-orthodox	regional	positive/negative	H5
pl_p_aus	Dummy for Polish region in the former Austro-Hungarian partition	non-orthodox	regional	positive/negative	H5
ln_no_fdi_cum_za	Log of the cumulated number of FDI in the region from a trade partner – zero adjusted (FDI inflow)	non-orthodox	regional and national	positive/negative	H6
ln_no_fdi_area_za	Log of the cumulated number of FDI in the region from a trade partner per square km – zero adjusted (FDI inflow)	non-orthodox	regional and national	positive/negative	H6
ln_value_esp_reg_fdi_abroad_za	Log of the value of Spanish regions FDI in foreign trade partners – zero adjusted (FDI outflow)	non-orthodox	regional and national	positive/negative	H6

Source: Own elaboration based on a review of the theoretical and empirical literature. * denotes variables of a baseline gravity equation, including distance, market size, and trade costs.

Regarding the data panel, it has been constructed for the trade of 16 NUTS 2 level regions of Poland (voivodships) and 19 NUTS 2 (autonomous regions) level regions of Spain with all possible trade partners (countries). The panel covers the period from 2005 to 2014 and is unbalanced due to missing observations. The dataset has been acquired from a number of sources. The trade data for Polish and Spanish regions

have been obtained from the Polish Customs Chamber (Izba Celna) and retrieved from the Spanish DataComex database (<http://datacomex.comercio.es>), according to the headquarters of exporters companies.

The main source of auxiliary data for countries is the Penn World Tables PWT 9.0. (Feenstra, Inklaar, and Timmer, 2015). The main source of auxiliary data for regions is the Quality of Government EU Regional Dataset (Charron et al., 2016). The information on the institutional quality of countries has been obtained from the 2016 version of the Worldwide Governance Indicators of the World Bank (Kaufmann, Kraay, and Mastruzzi, 2011). The data on FDI inflows by trade partner into regions has been obtained from PAIIZ for Poland and from <http://datainvex.comercio.es>, and from the Spanish Ministry of Economy and Competitiveness. The distance between trade partners has been calculated in kilometres, between the capital cities of the regions and the partner countries, with the use of “as the crow flies” method, using the information on their longitude and latitude. Additional geographical information has been acquired from CEPII (Mayer and Zignago, 2011; Melitz and Toubal, 2014).

8.3 Empirical results

Our empirical results for Polish and Spanish regions, taken jointly, are presented in Tables 8.3 and 8.4.

Tables 8.5 and 8.6 provide separate estimates for Polish and Spanish regions' exports. We have to stress that the general fit of the augmented gravity model is high. The model explains approx. 75 to 80% of the variation in the observed variables. The dummy for EU28 is the clustering variable used in the semi-mixed effect method of estimation.

According to our results, the impact of distance is negative and above one (in absolute terms). In addition, the impact of the size of the region and the size of trade partners is statistically significant and positive. The magnitude of the impact is generally higher for the size of the region than the size of the trade partner. The negative impact of distance is more pronounced for exports than for total trade (compare Table 8.4 to 8.3). The magnitude for exports is higher for more peripheral Polish regions than more central Spanish regions (Tables 8.5 and 8.6).

In all estimated models, the impact of dissimilarity in the levels of economic development is statistically significant and negative (except for several specifications for Polish regions, not accounting for the institutional quality). Thus, partners of a similar level of development (as measured by GDP per capita), in accordance with the Linder hypotheses, trade more. Also, the impact of the difference in TFP levels between the region and its trade partner is statistically significant and negative. It is worth noting that the magnitude of the impact is several times higher for Polish regions' than Spanish regions' exports, depicting significantly higher elasticity of reaction. Thus, hypothesis H1 has been positively verified.

The adjacency, as expected (H2), has a positive impact on bilateral trade flows. Its impact is visibly more important for Spanish regions. This could be due to the importance that France represents for Spanish exports. Trade flows originating from Spain go mainly to European destinations, and the trade is principally done using road transport. Therefore, the frontier with France is highly important, as is the main way to cross the Pyrenees Mountains (Márquez-Ramos, 2015). Gallego and Llano (2017), commenting on the limited number of road crossings with France in the Pyrenees, speak of dam-like effects on trade. Poland borders seven countries and has numerous border crossings. This, in turn, is in line with the result of Martincus (2010) for Brazilian regions, which suggest that trade openness favours location in regions closer to the largest neighbour trading partner.

In the joined sample of Polish and Spanish regions, the impact of access to a sea by a region is positive and statistically significant. It also holds for Spanish regions taken separately. The impact on Poland's regions exports is positive, however, not statistically significant. It could be related to the more pronounced role of the maritime transport for Spanish trade flows. In Q1 of 2017, the gross weight of seaborne goods amounted to 113.4 M tonnes in Spain and only 18.3 M tonnes in the case of Poland (according to the *mar_go* database of EUROSTAT). It should be stressed, however, that the impact of landlockedness of the trade partner is negative and statistically significant in all specifications considered.

The metropolitan status of a region in accordance with the ESPON classification boosts its total trade and exports in the joined sample of Polish and Spanish regions, even if we control for historical factors. If we control for the variation in the significance in the metropolitan area (a division of MEGA 1 to 4), the elasticities differ. They are positive for total trade (with the highest magnitude of the impact for MEGA3 – or weak European metropolitan areas: Mazowieckie (Warsaw), Pais Vasco (Bilbao) and Valencia, and positive and robust for exports of MEGA3 and MEGA4 (potential metropolitan areas with the European significance) regions. On the other hand, if the analysis is conducted for Polish and Spanish exports disjointedly, the coefficient on the metro variable is not statistically significant for Poland and is negative for Spain. At the same time, the impact of MEGAs is generally negative, *ceteris paribus*. From the country-level perspective, the nexus between the metropolitan status and exports requires further investigation.

In each model, from the sixth specification onwards, we control for regional governance quality and partners' institutional quality. The results are robust with the positive impact of institutional quality of the trade partner only. The lack of statistical significance of the coefficient on the regions' institutional quality could stem from the lower dispersion of the variable, compared to the one observed for countries. Further research on this relationship is recommended. H4, therefore, has been only partially verified.

In the next specifications, we test for the impact of historical ties, *ceteris paribus*. It includes former colonies of Spain and the relatively recent impact of partitions of

Poland. The impact of history is generally positive in the joined sample, supporting the notion of trade persistence. It is further analysed in separate models for Poland and Spain and indicates a distinct difference. Spanish regions, in general, export more to former Spanish colonies (Table 8.6). However, due to the data character that we use, we cannot observe regional variation in this case. For Poland, the impact varies between regions and the former empire that a given region had belonged to, before 1918. The impact differs clearly between former Prussian, Russian and Austro-Hungarian regions of Poland (Table 8.5). The findings are in line with the results reported by Brodzicki and Umiński (2017).

In the final set of specifications, we control for the impact of FDI inflows and outflows from regions. We introduce the information on a cumulated number of FDI (zero-adjusted) and the relative number of FDI per km², for a given trade partner in a given region. The impact of both on total trade in the joined sample is negative; however, statistically significant only in the first case. The impact on exports is positive, however not statistically significant for the first variable, and negative, for the second variable.

The impact for exports of Polish regions (Table 8.5) of both variables is positive, although not statistically significant. In the case of exports of Spanish regions (Table 8.6), it is positive for a cumulated number of investors and negative for the area-adjusted variable. In the case of Spain, due to data availability, we can additionally account for the value of FDI outflows by Spanish regions that have a positive impact. It is in line with the notion that FDI flows are complementary with regional trade flows and in line with results obtained for other countries (see, for example, Leichenko and Erickson (1997) for the US).

8.4 Conclusions

Although the gravity approach has been intensively used in international economics, the determinants of regional export activity still require thorough investigation. Both international and regional aspects are involved. The studies of a similar kind have been conducted in some countries and their regions; however, our objective is to verify both orthodox and non-orthodox determinants for regions in two EU member states.

The empirical results have proven most of the theoretical postulates related to factors potentially determining trade relations in a “region-to-country” framework. In particular, the paper has identified the determinants of the intensity of foreign trade of the Spanish and Polish NUTS 2 regions with all existing trade partners on a bilateral basis over the period 2005-2014. The augmented gravity panel model has been estimated with the use of a novel and superior method of estimation, namely the semi-mixed effects by PPML.

Table 8.3: Estimates for the value of total trade of Polish and Spanish regions

VARIABLES	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
ln_distance	-1.085*** (0.0132)	-1.094*** (0.0139)	-1.102*** (0.0137)	-1.069*** (0.0144)	-1.065*** (0.0146)	-1.052*** (0.0142)	-1.010*** (0.0133)	-1.025*** (0.0149)	-1.011*** (0.0133)
ln_rgdpc	0.856*** (0.00725)	0.834*** (0.00769)	0.837*** (0.007)	0.830*** (0.00766)	0.830*** (0.00751)	0.820*** (0.00819)	0.816*** (0.00816)	0.831*** (0.0104)	0.819*** (0.00862)
ln_rgdpr	1.091*** (0.0105)	1.075*** (0.0110)	1.079*** (0.0106)	1.011*** (0.0151)	0.979*** (0.0241)	1.021*** (0.0153)	1.039*** (0.0159)	1.083*** (0.0216)	1.047*** (0.0180)
diff_y	-0.142***	-0.142***		-0.162***	-0.156***	-0.153***	-0.172***	-0.183***	-0.177***
adjacency	(0.0204)	(0.0209)		(0.0208)	(0.0200)	(0.0220)	(0.0229)	(0.0237)	(0.0234)
access_2_sea	0.241*** (0.0455)	0.207*** (0.0466)	0.171*** (0.048)	0.208*** (0.0467)	0.233*** (0.0470)	0.214*** (0.0470)	0.176*** (0.0477)	0.193*** (0.0479)	0.182*** (0.0476)
c_landlocked		0.0706*** (0.0236)	0.072*** (0.0234)	0.0582*** (0.0239)	0.0689*** (0.0248)	0.0658*** (0.0243)	0.0721*** (0.0253)	0.0774*** (0.0252)	0.0604*** (0.0266)
		-0.266*** (0.0255)	-0.287*** (0.0253)	-0.283*** (0.0253)	-0.282*** (0.0250)	-0.305*** (0.0257)	-0.242*** (0.0272)	-0.258*** (0.0274)	-0.246*** (0.0271)
mega1					0.263*** (0.0649)				
mega3					0.347*** (0.0497)				
mega4					0.170*** (0.0355)				
eu28	0.262*** (0.0262)	0.266*** (0.0272)	0.297*** (0.0260)	0.301*** (0.0260)	0.308*** (0.0253)	0.255*** (0.0259)	0.321*** (0.0260)	0.331*** (0.0261)	0.325*** (0.0263)
diff_tfp			-0.067*** (0.009)						

continued **Table 8.3:** Estimates for the value of total trade of Polish and Spanish regions

VARIABLES	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
metro				0.199*** (0.0342)		0.177*** (0.0351)	0.158*** (0.0364)	0.143*** (0.0368)	0.159*** (0.0361)
r_inst_quality						-0.114** (0.0501)	-0.121** (0.0507)	-0.119** (0.0500)	-0.115** (0.0504)
c_inst_quality						0.0332*** (0.00641)	0.0307*** (0.00633)	0.0368*** (0.00587)	0.0315*** (0.00626)
history							0.387*** (0.0427)	0.382*** (0.0418)	0.384*** (0.0425)
ln_no_fdi_ cum_zs								-0.0248*** (0.00844)	
ln_no_fdi_ area_zs									-4.066 (3.101)
Constant	4.498*** (0.174)	5.056*** (0.192)	4.797*** (0.176)	5.507*** (0.204)	5.803*** (0.261)	5.366*** (0.213)	4.863*** (0.217)	4.321*** (0.277)	4.758*** (0.241)
Observations	39,118	38,377	38,377	38,377	38,377	38,098	36,321	36,321	36,321
R-squared	0.781	0.777	0.778	0.777	0.780	0.778	0.780	0.782	0.780
No of param- eters	7	9	10	10	12	12	13	14	14
Log likelihood	5.260e+12	5.160e+12	5.140e+12	5.130e+12	5.110e+12	5.050e+12	4.830e+12	4.820e+12	4.830e+12

Source: Own estimation in STATA 15. Semi-mixed effects model estimated with PPML. Robust standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.01. Dependent variable – the value of total trade in USD.

Table 8.4: Estimates for exports of Polish and Spanish regions

VARIABLES	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
ln_distance	-1.237*** (0.0133)	-1.259*** (0.0154)	-1.262*** (0.0150)	-1.242*** (0.0161)	-1.247*** (0.0159)	-1.190*** (0.0151)	-1.161*** (0.0148)	-1.152*** (0.0168)	-1.167*** (0.0150)
ln_rgdp_c	0.821*** (0.00728)	0.790*** (0.00784)	0.794*** (0.00767)	0.787*** (0.00784)	0.790*** (0.00795)	0.748*** (0.00744)	0.739*** (0.00737)	0.731*** (0.00992)	0.751*** (0.00771)
ln_rgdp_r	0.988*** (0.0102)	0.948*** (0.0111)	0.957*** (0.0113)	0.907*** (0.0162)	0.966*** (0.0241)	0.929*** (0.0161)	0.944*** (0.0166)	0.921*** (0.0231)	0.982*** (0.0183)
diff_y	-0.164*** (0.0168)	-0.157*** (0.0174)		-0.173*** (0.0162)	-0.167*** (0.0161)	-0.174*** (0.0167)	-0.190*** (0.0170)	-0.183*** (0.0174)	-0.214*** (0.0176)
adjacency	0.277*** (0.0436)	0.203*** (0.0478)	0.180*** (0.0480)	0.204*** (0.0483)	0.208*** (0.0504)	0.231*** (0.0483)	0.211*** (0.0484)	0.202*** (0.0482)	0.238*** (0.0508)
access_2_sea		0.222*** (0.0262)	0.224*** (0.0263)	0.213*** (0.0270)	0.196*** (0.0276)	0.226*** (0.0273)	0.227*** (0.0284)	0.224*** (0.0289)	0.176*** (0.0292)
c_landlocked		-0.346*** (0.0318)	-0.369*** (0.0311)	-0.360*** (0.0317)	-0.359*** (0.0321)	-0.420*** (0.0314)	-0.377*** (0.0322)	-0.368*** (0.0322)	-0.393*** (0.0321)
mega1					-0.0511 (0.0675)				
mega3					0.0823* (0.0493)				
mega4					0.113*** (0.0388)				
eu28	0.415*** (0.0247)	0.414*** (0.0273)	0.456*** (0.0281)	0.434*** (0.0264)	0.437*** (0.0261)	0.272*** (0.0284)	0.312*** (0.0277)	0.309*** (0.0278)	0.331*** (0.0277)
diff_tfp			-0.0562*** (0.00878)						
metro				0.124***		0.0787**	0.0594	0.0678	0.0647

continued **Table 8.4:** Estimates for exports of Polish and Spanish regions

VARIABLES	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
r_inst_quality				(0.0389)		(0.0396)	(0.0407)	(0.0427)	(0.0396)
						-0.121**	-0.128**	-0.129**	-0.104*
c_inst_quality						(0.0595)	(0.0609)	(0.0611)	(0.0595)
						0.105***	0.104***	0.101***	0.107***
history						(0.00480)	(0.00478)	(0.00511)	(0.00469)
						0.226***	0.226***	0.232***	0.211***
ln_no_fdi_cum_zs							(0.0454)	(0.0458)	(0.0451)
							0.0135		
ln_no_fdi_area_zs								(0.00909)	-21.89***
Constant	6.378*** (0.133)	7.352*** (0.156)	6.989*** (0.158)	7.649*** (0.175)	7.035*** (0.249)	7.446*** (0.180)	7.178*** (0.185)	7.459*** (0.261)	6.686*** (0.205)
Observations	47,378	46,373	46,373	46,373	46,373	45,875	43,456	43,456	43,456
R-squared	0.774	0.766	0.765	0.764	0.765	0.772	0.769	0.770	0.771
No of parameters	7	9	10	10	12	12	13	14	14
Log likelihood	2.360e+12	2.250e+12	-1.129e+12	2.250e+12	2.240e+12	2.130e+12	2.070e+12	2.070e+12	2.050e+12

Source: Own estimation in STATA 15. Semi-mixed effects model estimated with PPML. Robust standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1. Dependent variable – the value of exports in USD.

Table 8.5: Estimates for exports of Polish NUTS 2 regions

VARIABLES (1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)	(13)
ln_dis- tance	-1.298*** (0.0149)	-1.320*** (0.0154)	-1.320*** (0.0154)	-1.313*** (0.0151)	-1.216*** (0.0173)	-1.242*** (0.0242)	-1.247*** (0.0241)	-1.240*** (0.0248)	-1.253*** (0.0239)	-1.214*** (0.0176)	-1.234*** (0.0238)	-1.240*** (0.0240)
ln_rgdp_c	0.793*** (0.00974)	0.751*** (0.00969)	0.762*** (0.00951)	0.752*** (0.00918)	0.738*** (0.0104)	0.749*** (0.0129)	0.749*** (0.0127)	0.744*** (0.0129)	0.749*** (0.0123)	0.734*** (0.0107)	0.743*** (0.0133)	0.747*** (0.0128)
ln_rgdp_r	1.252*** (0.0189)	1.270*** (0.0211)	1.218*** (0.0202)	1.716*** (0.0511)	1.127*** (0.0353)	1.137*** (0.0370)	1.126*** (0.0366)	1.120*** (0.0369)	1.118*** (0.0366)	1.115*** (0.0355)	1.119*** (0.0374)	1.129*** (0.0366)
diff_y	0.172*** (0.0269)	0.164*** (0.0280)		0.142*** (0.0277)	-0.189*** (0.0473)	-0.168*** (0.0510)	-0.175*** (0.0488)	-0.173*** (0.0495)	-0.167*** (0.0489)	-0.183*** (0.0475)	-0.160*** (0.0550)	-0.163*** (0.0515)
adjacency	0.0821 (0.0727)	0.0880 (0.0750)	0.0638 (0.0747)	0.114* (0.0644)	0.158** (0.0705)	0.188*** (0.0713)	0.183** (0.0718)	0.199*** (0.0732)	0.182** (0.0728)	0.150** (0.0704)	0.182** (0.0718)	0.185*** (0.0711)
access_2_ sea	0.0506 (0.0701)	0.0489 (0.0696)		0.264*** (0.0629)	0.000161 (0.0681)	0.0105 (0.0700)	0.00177 (0.0685)	0.00421 (0.0688)	-0.00344 (0.0688)	-0.00785 (0.0679)	0.00608 (0.0703)	0.0103 (0.0699)
c_land- locked	-0.359*** (0.0346)	-0.370*** (0.0323)	-0.359*** (0.0346)	-0.424*** (0.0334)	-0.345*** (0.0354)	-0.342*** (0.0381)	-0.364*** (0.0373)	-0.373*** (0.0397)	-0.346*** (0.0396)	-0.313*** (0.0357)	-0.337*** (0.0374)	-0.339*** (0.0379)
eu28	0.345*** (0.0372)	0.404*** (0.0372)	0.412*** (0.0379)	0.402*** (0.0376)	0.479*** (0.0355)	0.357*** (0.0394)	0.344*** (0.0397)	0.345*** (0.0397)	0.339*** (0.0397)	0.355*** (0.0391)	0.361*** (0.0521)	0.348*** (0.0420)
diff_tfp					-0.0106*** (0.0117)							

continued **Table 8.5:** Estimates for exports of Polish NUTS 2 regions

VARIABLES (1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)	(13)
metro			0.0446 (0.0420)		0.0817* (0.0421)	0.0787* (0.0440)	0.0844* (0.0433)	0.0941** (0.0434)	0.0947** (0.0434)	0.0978** (0.0421)	0.0914** (0.0451)	0.0854* (0.0442)
mega3				-1.031*** (0.102)								
mega4				-0.352*** (0.0522)								
r_inst_					-0.0898	-0.0863	-0.0868	-0.0894	-0.0863	-0.0905	-0.0850	-0.0962
quality												
c_inst_					(0.0629)	(0.0661)	(0.0650)	(0.0651)	(0.0652)	(0.0625)	(0.0665)	(0.0653)
quality					0.123***	0.117***	0.119***	0.127***	0.120***	0.129***	0.111***	0.115***
history					(0.0107)	(0.0113)	(0.0112)	(0.0116)	(0.0108)	(0.0114)	(0.0160)	(0.0118)
						-0.113*					-0.110*	-0.115*
						(0.0606)					(0.0601)	(0.0611)
pl_p_							-0.127**					
common							(0.0606)					
pl_p_deu								-0.139**	-0.154**			
								(0.0703)	(0.0692)			
pl_p_rus								0.230***		0.262***		
								(0.0757)		(0.0713)		
pl_p_aus										-0.637***	-0.627***	

continued
Table 8.5: Estimates for exports of Polish NUTS 2 regions

VARIABLES	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)	(13)
ln_no_fdi_ cum_zs										(0.0644)	(0.0642)	0.0203	
ln_no_fdi_ area_zs												(0.0309)	109.4
													(168.6)
Observa- tions	21,566	20,930	20,930	20,930	20,930	20,768	19,776	20,768	20,768	20,768	20,768	19,776	19,776
R-squared	0.786	0.776	0.776	0.777	0.820	0.781	0.770	0.769	0.768	0.767	0.783	0.769	0.770

Source: Own estimation in STATA 15. Semi-mixed effects model estimated with PPML. Robust standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1. Dependent variable – the value of exports in USD. The constant term always positive and statistically significant removed from the presentation for clarity.

Table 8.6: Estimates for exports of Spanish NUTS 2 regions

VARIABLES	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
ln_distance	-1.085*** (0.0224)	-1.060*** (0.0242)	-1.085*** (0.0238)	-1.063*** (0.0244)	-1.057*** (0.0247)	-1.108*** (0.0238)	-1.214*** (0.0240)	-1.161*** (0.0282)	-1.219*** (0.0246)	-1.153*** (0.0252)
ln_rgdp_c	0.795*** (0.00985)	0.758*** (0.00990)	0.784*** (0.00969)	0.759*** (0.00984)	0.756*** (0.00989)	0.748*** (0.00940)	0.779*** (0.0104)	0.738*** (0.0154)	0.784*** (0.0108)	0.735*** (0.0130)
ln_rgdp_r	0.950*** (0.0139)	0.887*** (0.0143)	0.890*** (0.0143)	1.002*** (0.0340)	0.942*** (0.0429)	1.020*** (0.0348)	1.032*** (0.0355)	0.955*** (0.0422)	1.042*** (0.0364)	0.958*** (0.0340)
diff_y	-0.312*** (0.0240)	-0.349*** (0.0232)		-0.351*** (0.0233)	-0.342*** (0.0231)	-0.230*** (0.0240)	-0.159*** (0.0259)	-0.163*** (0.0257)	-0.166*** (0.0267)	-0.153*** (0.0260)
adjacency	0.488*** (0.0473)	0.378*** (0.0490)	0.370*** (0.0491)	0.375*** (0.0475)	0.392*** (0.0460)	0.367*** (0.0474)	0.304*** (0.0487)	0.269*** (0.0474)	0.312*** (0.0508)	0.309*** (0.0470)
access_2_sea		0.475*** (0.0298)	0.468*** (0.0299)	0.527*** (0.0355)	0.550*** (0.0361)	0.531*** (0.0351)	0.529*** (0.0355)	0.549*** (0.0347)	0.502*** (0.0376)	0.540*** (0.0347)
c_landlocked		-0.556*** (0.0599)	-0.534*** (0.0620)	-0.553*** (0.0593)	-0.558*** (0.0586)	-0.637*** (0.0490)	-0.567*** (0.0468)	-0.548*** (0.0455)	-0.571*** (0.0472)	-0.550*** (0.0461)
eu28	0.509*** (0.0347)	0.530*** (0.0375)	0.684*** (0.0355)	0.525*** (0.0372)	0.533*** (0.0372)	0.323*** (0.0391)	0.295*** (0.0370)	0.270*** (0.0359)	0.297*** (0.0371)	0.297*** (0.0375)
mega1					-0.125 (0.104)					
mega3					-0.166** (0.0818)					
mega4					-0.277*** (0.0836)					
diff_tfp			-0.0554*							

Table 8.6: Estimates for exports of Spanish NUTS 2 regions

continued

VARIABLES	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
metro			(0.0114)	-0.259*** (0.0770)		-0.276*** (0.0770)	-0.291*** (0.0778)	-0.317*** (0.0775)	-0.279*** (0.0772)	-0.319*** (0.0746)
r_inst_quality						0.152* (0.0901)	0.152 (0.0926)	0.159* (0.0913)	0.155* (0.0930)	0.134 (0.0917)
c_inst_quality						0.0868***	0.113***	0.0984***	0.114***	0.103***
history						(0.00683)	(0.00668)	(0.00740)	(0.00673)	(0.00692)
ln_no_fdi_cum_z							1.257*** (0.0448)	1.141*** (0.0555)	1.266*** (0.0453)	1.081*** (0.0515)
ln_no_fdi_area_z								0.0514*** (0.0115)	-6.009* (3.618)	0.0196***
ln_value_esp_reg_fdi_abroad_z										
Observations	25,812	25,443	25,443	25,443	25,443	25,107	23,680	23,680	23,680	23,680
R-squared	0.794	0.811	0.810	0.815	0.818	0.820	0.821	0.825	0.820	0.824

Source: Own estimation in STATA 15. Semi-mixed effects model estimated with PPML. Robust standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1. The constant term always positive and statistically significant removed from the presentation for clarity. Full results available upon request.

As the gravity model has been proven to be a useful and robust tool to estimate the nature of the region-to-country trade – it should be more intensively used by regional authorities to assess the impact of various policy scenarios and predict their consequences.

The model has been jointly estimated for all Spanish and Polish regions and then in the second step disjointedly, in a comparative manner. The orthodox or standard determinants of trade flows, such as distance and market size, prove the robustness of the application of the gravity approach to the “region-to-country” trade framework. The impact of other standard variables such as adjacency is in line with theoretical postulates. Polish and Spanish regions trade more with partners at a similar level of development in line with the Linder hypothesis. As regards non-orthodox factors, the similarity in productivity levels as measured by TFP, the institutional quality of trading partner, the metropolitan status of the region, and historical linkages influence the bilateral trade flows. Therefore, we agree with Linders, Burger, and van Oort (2008), that more elaborate variety of tangible and intangible dimensions of transactional distance are crucial to correctly interpret differences between trade partners.

The export potential of a region can be boosted by the attraction of FDI, thus proving the importance of effective regional promotional activity and raising high investment attractiveness. As we have shown, the similarity in the level of development boosts bilateral trade. It paradoxically could direct export promotion policies towards more dissimilar partners. However, it could be highly region-specific and requires further investigation, also embracing the structural factors. Having shown that adjacency matters, promotional activity should be more pronounced on non-border regions.