Pound for Pound Export Diversi cation

James E. Anderson
Boston College
NBER

Yoto V. Yotov
Drexel University

December 20, 2020

Abstract

We propose a short-run model of the extensive margin of trade and deploy it to distinguish and quantify domestic and cross-border margins. Our empirical focus is on the domestic extensive marginof trade (domestic distribution of a product) and its importance for quantifying policy and globalization e ects on the international extensive margin of trade. We build a dataset that combines data on the domestic extensive margin and the standard international extensive margin. It reveals signi cant and intuitive variation in the domestic extensive margin across countries and over time. We quantify the extensive margin e ects of European Union (EU) integration, 2008-2018, and demonstrate that these e ects cannot be identi ed without the domestic extensive margin. We nd strong and highly heterogeneous e ects, both across countries and directionally.

JEL Classi cation Codes: F13, F14, F16
Keywords: Extensive Margin, Domestic Extensive Margin, Globalization, Gravity

---

The title of our paper was motivated by combat sports, where the term `pound for pound' is used to compare and rank ghters who compete in di erent divisions/categories based on their weight. Similarly, we propose `pound for pound export diversi cation' to take into account the fact that smaller and poorer economies produce less products. Pound for pound rankings and analysis on the extensive margin of trade adjust for this regularity to properly account for the relative performance of the `lighter weight' countries. We thank Joao Santos Silva for kindly sharing the original programming codes for the FLEX estimator. We thank Rebecca Freeman and Stephan-Alfons Nolte for very helpful advice and guidance to the data. We also

\Trade diversi cation is a national imperative for the Government of Canada. Over the next six years, starting in 2018-19, Canada's export diversi cation strategy will invest $1.1 billion to help Canadian businesses access new markets."

(Government of Canada, March 3, 2020)

\Increased diversi cation is associated with lower output volatility and greater macroeconomic stability [in low-income countries]. There is both a growth payo and a stability payo to diversi cation, underscoring the case for paying close attention to policies that facilitate diversi cation and structural transformation."

(IMF, March, 2014)

1 Introduction

introduction of the domestic extensive margin enables us to identify the effects of a number of policies whose impact is impossible to obtain within a properly specified empirical gravity model that only employs data on the external extensive margin. Specifically, with data on the domestic extensive margin we could identify the effects of (i) non-discriminatory export support policies, e.g., export subsidies, trade fairs, etc., (ii) non-discriminatory import protection policies, (iii) country-specific characteristics and policies, e.g. institutional quality, technical barriers to trade (TBT) etc., (iv) exchange rates, and (v) the effects of globalization on the extensive margin of trade. We also argue that the introduction of the domestic extensive margin may have implications for the estimates of bilateral trade policies, e.g., regional trade agreements (RTAs), membership in the World Trade Organization (WTO), etc.

On the theory side, we introduce the domestic extensive margin of trade in a short run structural gravity model that features dynamic adjustments of bilateral capacities by heterogeneous firms. The lens of the model allows focus on action on the extensive margin of international trade (new export destinations) and domestic trade (new products). Capital is sector- and destination-specific. Investment on the extensive margins is selected when the expected return exceeds the product of the opportunity cost of capital and an adjustment cost factor.[2] Since much of this capital is unobservable, its behavior is inferred by fixed effects modeled consistently with the theoretical implications of the model. A key implication of our model is that proper quantification of the international extensive margin (the set of partners any sector exports to) should also take into account the domestic margin of trade (the set of sectors with positive production).

The empirical analysis is based on a novel dataset that covers the extensive margin of trade in mining and manufacturing goods for 32 European countries over the period

[2]The baseline for our theory is the intensive margin short-run gravity model of Anderson and Yotov (2020). However, the general approach to model investment in bilateral trade links is also in the spirit of Arkolakis (2010), Head et al. (2010), Chaney (2014), Mion and Opromolla (2014), Sampson (2016), and Crucini and Davis (2016). Our innovations in relation to these papers are (i) developing the model on the extensive margin, and (ii) the focus on the domestic extensive margin.

2008-2018. The important and unique dimension of our dataset is the *domestic extensive margin*. The dataset is constructed from two original sources. Production data is taken from Eurostat's *Production Communautaire* (PRODCOM) database. Production data is combined with trade data from Eurostat's COMEXT data. The combination of PRODCOM and COMEXT allowed us to build an estimating sample that covers consistently constructed data on the external and the domestic extensive margins for 32 European countries and about

EU integration effects in a theory-consistent econometric specification without data on the domestic extensive margin. Second, from an econometric perspective, the focus on Europe (2008-2018) allows us to obtain estimates of the EU integration effects within a simple, flexible, and robust econometric specification with fixed effects only.[4] The fixed effects treatment is convenient because it enables us to obtain a series of EU integration estimates (across time and for individual countries) while, at the same time, the rich fixed effects structure of our model diminishes omitted variable and endogeneity concerns. Finally, the proposed application is interesting and relevant for its potential implications for export diversification strategies.

We rely on three different strands of the literature to specify our econometric model. First, the theory developed in this paper extends the CES structural gravity model to a closed form that features both domestic and cross-border extensive margins of trade. The model motivates our reduced-form empirical specification that identifies these margins. Second, the reduced form specification achieves identification with a rich set of fixed effects following recent developments in the empirical gravity literature on the intensive margin of trade. Third, the fixed effects representation of the theoretical model is estimated with the Santos Silva et al. (2014) FLEX estimator. FLEX is designed to consistently deal with the boundedness above and below of the extensive margin dependent variable. We also demonstrate the robustness of our main findings to the use of alternatives estimators including Tobit, OLS, and the Poisson Pseudo Maximum Likelihood (PPML) of Santos Silva and Tenreyro (2006, 2011). We show below that identification of the EU integration effects with the theory-consistent specification requires the use of data on the domestic extensive margin,

[4]Technically, we do have controls for membership in the World Trade Organization (WTO) and in Economic Integration Agreements (EIAs) in our main specifications. However, given the specifics of sample (i.e., covering only European economies) and the use of country-pair fixed effects, the estimates of the EIA and WTO covariates are identified of very few observations and the introduction of these variables does not affect our main results. For example, Montenegro is the only country from our sample that became a WTO member during the period of investigation (in 2012), while all the variation in the EIA covariate come from the trade agreements of very few countries including Bosnia and Herzegovina, Croatia, Montenegro, and Serbia. We capitalize on the fact that Montenegro is the only country that joined the WTO in our sample to demonstrate that the introduction of the domestic extensive margin also allows for identification of country-specific policy effects.

regardless of the estimator.

The empirical analysis starts with a benchmark specification that imposes common globalization effects across all countries in the sample. The main result from this analysis is that globalization has had a significant positive impact on the international extensive margin of trade relative to the domestic extensive margin for the European economies. Intensive integration processes within Europe are the natural explanation for this result. This result should be important from a policy perspective because there is plenty of anecdotal evidence that the impact of globalization on the intensive margin stalled during the years after the great recession. In contrast, our results indicate that the impact of globalization on the extensive margin during the same period has been economically strong and statistically significant. Our preferred specification implies that, on average across the countries in our

Our work complements and extends two strands of the literature. Most closely related is the literature on the extensive margin of trade. Melitz (2003), Helpman et al. (2008a) and Chaney (2008) are prominent examples of theoretical contributions to this literature, and Redding (2011) o ers an excellent survey of the related theoretical literature, the empirical challenges related to this research, and its implications for the extensive margin of trade. From an empirical and application perspective, see Hummels and Klenow (2005) for an important study on the extensive margin at the sector/product level, and Helpman et al. (2008a) for an in uential analysis of the extensive margin at the country level. Finally, from an estimation point of view, Santos Silva et al. (2014) summarize and extend the latest econometric developments in the estimation of the extensive margin of trade. Their FLEX estimator is used to obtain our main results. Our main innovations in relation to this literature are the modeling of the extensive margin in the short run and the introduction of the *domestic extensive margin*. As we demonstrate below, our contribution has implications for quantifying the e ects of various policies as well as for the measurement and the construction of indexes on the extensive margin of trade.[5]

The other branch of related literature includes papers that emphasize the importance of proper account for domestic trade ows on the intensive margin of trade. For example, Yotov (2012) uses domestic trade ows to resolve `the distance puzzle' in international trade. Ramondo et al. (2016) demonstrate that when domestic trade ows are taken into account, two other gravity literature puzzles are resolved: (i) that larger countries should be richer than smaller countries and (ii) that real income per capita increases too steeply with country size. Agnosteva et al. (2019) employ domestic trade ows to estimate heterogeneous domestic trade costs. Finally, Heid et al. (forthcoming) show that the use of domestic trade allows for identi cation of unilateral and non-discriminatory trade policies in intensive margin structural gravity models. Our contribution is that we o er a theoretical motivation and

[5]Thus, from a policy perspective, our contribution is related to a very large number of papers that study the impact of various determinants of the extensive margin of trade. Without an attempt to be exhaustive, for some excellent studies we refer the reader to Felbermayr and Kohler (2006), Berthou and Fontagne (2008), Cadot et al. (2011), Persson (2013), and Beverelli et al. (2015).

empirical evidence for the importance of the domestic extensive margin for quantifying the extensive margin of trade. Our methods open avenues for many extensions and applications, e.g., estimating the impact of country-speci c policies and characteristics (e.g., export promotion, institutional quality). We elaborate on some of these ideas in the concluding section of the paper.

The rest of the paper is structured as follows. Section 2 develops our theoretical model and then translates it into an econometric speci cation. Section 3 describes the data sources and our methods to construct the data. Section 4 reports and discusses our estimates of the impact of globalization and the results from a series of robustness experiments. Finally, Section 5 summarizes our contributions and ndings and points to a series of additional implications and extensions. The derivations of our theoretical model are in the Appendix.

2 Quantifying the Extensive Margin of Trade

Subsection 2.1 combines and extends three prominent strands of the trade literature to derive a short-run gravity theory on the extensive margin(s) of trade. Our key contributions in relation to the existing literature are the derivation of the *short-run extensive margin of international trade* and the introduction of the notion of *domestic extensive margin*. Subsection 2.2 capitalizes on a number of developments in the empirical literature on the extensive and the intensive margins of trade to translate our theory into an econometric speci cation.

1(tendluding)].

Mion and Opromolla (2014), Sampson (2016), Crucini and Davis (2016), and Anderson and Yotov (2020). Third, we account for action on the extensive margin of trade following Melitz (2003), Helpman et al. (2008a), Chaney (2008), and Redding (2011). The novelties are (i) our treatment of the extensive margin in the short run and (ii) the explicit account for the domestic extensive margin. Since the key building blocks of our theory are relatively standard in the literature, we relegate all derivations to the Appendix. This section summarizes our assumptions, presents the resulting model, and provides intuition behind each of its components with emphasis on the novel elements.

The world consists of many countries that produce their own product varieties (Armington, 1969) and trade with each other. Heterogeneous firms in each sector h and origin i allocate capital and labor to production and to distribution to a set of destinations j using Cobb-Douglas technology.[6] The capital becomes specific once allocated. Subsequently, the firms draw productivities from a Pareto distribution, demand shocks are realized and labor

Equation (1) can be decomposed into two structural terms. We label the first term `Structural Gravity' because, as famously demonstrated by Arkolakis et al. (2012), it can be derived from a very wide class of theoretical economic micro-foundations. The intuition behind this term is standard and simple, i.e., bilateral trade flows between two countries ($X^h_{ij;t}$) are proportional to the product of their sizes ($Y^h_{i;t}$ and $E^h_{j;t}$, output and expenditure, respectively) as a share of world output (Y^h_t), and inversely proportional to the trade frictions between them, which consist of direct bilateral trade costs ($t^h_{ij;t}$) and general equilibrium trade frictions captured by the multilateral resistances ($\Pi^h_{i;t}$ and $P^k_{j;t}$, outward and inward, respectively), and where $(1-\sigma^k)\theta^k$ is the trade elasticity, which is a function of the elasticity of substitution, σ^k, and another structural structural parameter, θ^k, which we define next.

We label the second term in equation (1) `*Short Run with Melitz Firms*' because it com-

of selection of heterogeneous rms. Selection results from the combination of xed labor costs for export and heterogeneous productivity draws by *ex ante* identical rms. The short run supply elasticity that is part of structural parameter ^{h} is a combination of diminishing labor productivity in shipping to destination j (parameterized as the inverse of capital's share parameter) and rm selection (parameterized by the shape parameter of a Pareto distribution).[10]

Our main concern in this paper is the application of this model to the extensive margin where j = n. This introduces a new term multiplying the right hand side of (1), (

switching on bilateral capacity when applied with international trade data only. We emphasize that (1) and (2) hold equally for international and domestic links, i.e., both for $i \neq j$ and for $8i = j$. With trade data that includes both domestic and international trade it is possible for xed e ects techniques to control for switching on bilateral capacity. Note that although domestic sales often precede exports, this is neither necessary nor universally observed.

Our theory and its application capture two distinct forms of the extensive margin of trade. First is the standard external (cross-border) margin of trade whereby the production and distribution for export changes. Second is the domestic margin $i = j$ where domestic distribution changes. This is what we call the `*Domestic Extensive Margin' (DEM)*, a key focus of our empirical analysis. The empirical analysis below demonstrates that proper econometric accounting for the domestic extensive margin may have signi cant implications for identifying the impact of a number of determinants of the external extensive margin. The lesson is consistent with the closed form model (1) but more broadly suggests the importance of simultaneously accounting for both extensive margins.

The domestic extensive margin could in principle be active in an already active sector; production need not imply domestic sales. This phenomenon is absent from the data in our 1btd form moneither n(e)-367(margin)-3n al

A key aspect of the short run structural gravity model (2) is its structural time invariance. This opens the door to exploit time variation in exogenous variables within the structural short run gravity model to empirically characterize the extensive margins of trade. There are two sources of time series variation in the two extensive margins. One is due to cyclic volatility of service in bilateral links. The other is secular change (growth or decline) in the number of markets served. Both sources of action on the extensive margins are potentially active and quantitatively important.[12] Both forms of the extensive margin are described by the simple selection mechanism of the heterogeneous rms embedded in speci cation (2).

2.2 From Theory to Empirics: Estimating the Extensive Margin

The lens of theoretical equation (2) focuses application on a corresponding econometric model. To this end, we proceed in three steps and rely on three di erent strands of the literature. First, we translate our theory into an econometric model, which is broadly consistent with other structural models on the extensive margin of trade, e.g., Helpman et al. (2008b). Second, following the recommendations of Santos Silva et al. (2014), we select their *FLEX* estimator to obtain our main results. Finally, guided by the empirical literature on the intensive and on the extensive margins of trade and by our key contribution (i.e., the introduction of the domestic extensive margin), we select the covariates in our empirical model.

We start by translating our theory into an econometric model.[13] Let $N_{ij;t}^{k}$ be an indicator equal to one when at least one rm exports k from i to j at time t. In order for this to be the case, there should be at least one rm in this sector that nds it pro table to produce and

learning how to produce and serve new sector/destinations plausibly takes place over time, inducing partial adjustment and correcting for mistakes. The treatment here abstracts from all such dynamic considerations to simplify focus on the essential static logic: entry requires a lower than eventually e cient capacity to raise next period returns above the opportunity cost of capital.

[12] Besedes and Prusa (2006) document the high volatility over time of 10 digit HS level bilateral US exports.

export, i.e., $\pi^k_{ij;t}(\%) > 0$. This implies that the probability for a given sector to be exported from origin i to destination j at time t is:

$$Pr(N^k_{ij;t} = 1 | x_{ij;t}) = Pr(\pi^k_{ij;t}(\%) > 0) = F^k(x^0_{ij;t}) \quad (3)$$

Letting $N_{ij;t} = \sum_k N^k_{ij;t}$ be the total number of sectors exported from i to j at time t, the previous expression implies:

$$E(N_{ij;t} | x_{ij;t}) = \sum_k Pr(N^k_{ij;t} = 1 | x_{ij;t}) = \sum_k F^k(x^0_{ij;t}) = N_{i;t} F(x^0_{ij;t}); \quad (4)$$

where $N_{i;t}$ is the total number of sectors available in origin i, and $F(x^0_{ij;t})$ (

distribution due to its low weight in the objective function. Following Santos Silva et al. (2014), we will estimate the model by Bernoulli pseudo-maximum likelihood, which is easy to implement and it is consistent under very general conditions, c.f., Papke and Wooldridge (1996).[15]

To demonstrate the robustness of our main results, we also experiment with three alternative estimators. First, we employ a double-bounded Tobit estimator. In addition, following the best current practices in the intensive margin gravity literature, we also experiment with the Poisson Pseudo Maximum Likelihood (PPML), which has the attractive properties of being a count multiplicative model, which can take into account the information contained in the zero observations in our sample. PPML established itself as the leading gravity estimator due to the seminal work of Santos Silva and Tenreyro (2006), and we refer the reader to Santos Silva and Tenreyro (2006) and Santos Silva and Tenreyro (2011) for excellent discussions of the attractive features of PPML for gravity estimations on the intensive margin of trade, and to Berthou and Fontagne (2008) for an application to the extensive margin of trade. Finally, despite its limitations in the current setting, i.e., inability to capture the behavior of the distribution at its bounds because the partial OLS effects are assumed to be constant, we also obtain robustness estimates with the OLS estimator. As demonstrated in the sensitivity analysis, our main results and conclusions are robust to the use of alternative estimators.

The third and final step to complete our econometric setup is to explicitly define the covariates in our model. To this end, we rely on the numerous contributions to the empirical literature on the intensive and on the extensive margins of trade, as well as on our key contribution, i.e. the introduction of the domestic extensive margin. Taking into account the latest developments in the estimation of gravity equations (on the extensive and on the

[15]The refer the reader to Gourieroux et al. (1984) and Papke and Wooldridge (1996) for a discussion.

intensive margin), we start by de ning:

$$\exp\left(x^{0}_{ij;t}\right) = \exp\left(\ _{i;t} + \ _{j;t} + \ _{ij} + BIPOL_{ij;t\ \ 1}\right); \quad 8i \neq j \tag{5}$$

Equation (5) includes three sets of xed e ects. $_{i;t}$ and $_{j;t}$ are exporter-time and importer-time xed e ects, which would control for and absorb the multilateral resistance terms from our theoretical model, as well as any other country-speci c time-varying characteristics that may a ect the bilateral extensive margin, on the exporter and on the importer side, respectively. $_{ij}$ denotes a set of country-pair xed e ects, whose purpose is to account for all time-invariant bilateral determinants of the extensive margin of trade. Finally, $BIPOL_{ij;t}$ is a vector of time-varying bilateral determinants of trade, e.g., trade agreements, tari s, etc.

An important feature of all empirical papers on the extensive margin of trade, as captured by equation (5), is that, without exception, all of the existing extensive margin analyses are performed exclusively with international trade data only and without taking into account the domestic extensive margin.[16] As we demonstrate next, proper/theory-consistent account for the domestic extensive margin may have signi cant implications for estimating the impact of numerous determinants of the extensive margin of trade. To see this, note that once the domestic extensive margin is introduced, equation (5) becomes:

$$\exp\left(x^{0}_{ij;t}\right) = \exp\left(\ _{i;t} + \ _{j;t}\right.$$

of any non-discriminatory export support policies can be identified in the presence of the exporter-time fixed effects because the export support policies apply only to international and not to domestic trade. Specifically, $EXS_{i;t}$ is a vector of non-discriminatory export support policies, e.g., export subsidies, trade fairs, etc. We interact $EXS_{i;t}$ with $BRDR_{ij}$, which is an indicator variable for cross-border trade, equal to 0 for domestic trade. Thus, the resulting interaction, $EXS_{i;t}$ $BRDR_{ij}$, is time-varying and bilateral and, therefore, it can be identified in the presence of all fixed effects from (6).

The second new term in (6) is $IMP_{j;t}$ $BRDR_{ij}$, and it is constructed as an interaction between a vector of non-discriminatory import protection policies, $IMP_{j;t}$, and the international border dummy. Similar to the case of export support, the impact of any non-discriminatory import protection policies cannot be identified in the presence of importer-time fixed effects without the domestic extensive margin.

The third new term in (6) is $CNTRY_{j;t}$ $BRDR_{ij}$, and it is constructed as an interaction between a vector of country-specific characteristics and policies, e.g. institutional quality, technical barriers to trade (TBT) etc., $CNTRY_{j;t}$, and the international border dummy. Once again, the impact of such policies cannot be identified without the domestic extensive margin. The difference between this term and the directional (export and import policies) is that we can only identify the differential impact of such policies on international relative to internal trade, however not depending on the direction of trade flows, e.g., not on the impact of exports vs. imports.

The fourth new term in specification (6) is the exchange rate between i and j at t, $EXR_{ij;t}$. Even though exchange rates are bilateral their impact cannot be identified in gravity specifications with international trade data only due to perfect collinearity with the exporter time and importer time fixed effects. Once the domestic extensive margin is introduced, we can obtain estimates of the nonuniform/discriminatory impact of exchange rates on the external relative to the domestic extensive margin, because exchange rates do not vary domestically.

The fth new term in speci cation (6) is $\sum_{i;t} {}_{i;t} GLOB_{i;t} \quad BRDR_{ij}$, which denotes a set of time-invariant cross-border dummies $BRDR_{ij}$ interacted with origin-time globalization dummies $GLOB_{i;t}$. The motivation for the inclusion and emphasis on this term is twofold. First, from a practical perspective, the inclusion of the globalization dummies will enable us to address the challenge that we do not have data on the key variable of interest in (2), i.e., $_{in;t}^{h\ 1\ \ h}$. Thus, the country-time speci c globalization estimates that we will obtain in the empirical analysis will o er a exible and comprehensive/all-inclusive account for the dynamic evolution of the international bilateral links relative to the domestic extensive margin. Second, the inclusion of the time-varying border indicators would enable us to resolve the `the missing globalization puzzle', c.f., Coe et al. (2002), on the extensive margin of trade. In the empirical analysis we demonstrate that the e ects of globalization are

includes observations labeled as Con dential (:C), Estimated (:E), or Con dential/Estimated (:CE). These observations account for a total of approximately 20% of the original data. The observations labeled Con dential (:C) or Con dential/Estimated (:CE) account for more than 19%, while the Estimated (:E) observations were less than 1%. While the presence of con dential and/or estimated observations could have been potentially problematic for an analysis on the *intensive margin* of trade, they are not such a big concern in our case, where the focus is on the *extensive margin* and all we need to know is whether there is production or not in a given category. To take advantage of the information contained in the con dential and the estimated observations we proceed in three steps. First, we assign a value of one on the extensive margin for any estimated or con dential observations for which there were positive production values in the same category but in other years in the original data. Second, we assign a value of zero on the extensive margin for any estimated or con dential observations for which the non-missing production values in the same categories in all other years in the original data are zeroes. Finally, if the observations for all years for a given country and product category were classi ed as con dential and/or estimated, we assign a value of one on the extensive margin.

The last two steps in the construction of the domestic extensive margin are (i) to replace the positive reported production values with ones, and (ii) to sum them for each country and year in the sample. For consistent comparisons (since the number of possible products varies across years), we de ne our novel index of the *Domestic Extensive Margin (DEM)* as the ratio between the number of products actually produced by a given country in a given year, $D_{i;t}$, and the total number of possible products that could have been produced by the same country and in the same year, $N_{i;t}$:

$$DEM_{i;t} = \frac{D_{i;t}}{N_{i;t}}:$$

The domestic extensive margin indexes for all countries and all years in our sample appear

in Table 1. The total number of possible products are reported in the last row of the table. The last column of the table reports percentage changes for each country between the rst and the last year for which data are available. The exception is Serbia, for which the initial year for the percentage change in the last column is 2012. As can be seen from Table 1, the domestic extensive margin index for Serbia in 2011 is very di erent from the relatively stable indexes in the subsequent years. In combination with the fact that 2011 is the rst year for which Serbia was included in PRODCOM, we conclude that the 2011 data for Serbia are not reliable and, therefore, for the remainder of the analysis we treat the observations for Serbia in 2011 as missing.

Several interestist

First, we see that a number of countries have experienced an increase/improvement on the domestic extensive margin. The countries with the largest increases are Hungary, Netherlands, Lithuania, Slovenia, and Greece. Apart from the Netherlands, a possible explanation for such favorable ranking is that these countries have bene ted from their integration in the European Union. On the other side of the spectrum we nd Portugal, Croatia, Finland, Italy, and the United Kingdom. Finally, a third group of countries have not experienced signi cant changes on the domestic extensive margin. These countries include Norway, Spain, Germany, and Iceland. Interestingly, two of these countries (e.g., Germany and Spain) have very large indexes, while the other two countries (e.g., Norway and Iceland) are among the ones with the smallest indexes.

This section presented the *Domestic Extensive Margin* index. The accompanying analysis revealed wide heterogeneity in the *DEM* indexes across the countries in our sample as well as signi cant variation of *DEM* over time. This variation is useful for identi cation of heterogeneous EU integration e ects on domestic and international margins below, controlling for size e ects and multilateral resistance e ects consistent with the structural gravity model.

3.2 Matching the Domestic & International Extensive Margins

We rely on the COMEXT database to construct the international extensive margin of trade. According to the o cial Eurostat web site *COMEXT is Eurostat's reference database for detailed statistics on international trade in goods"*, and the dataset o ers very detailed statistics according to the Combined Nomenclature (CN) classi cation system.[19] We follow the standard method to construct the extensive margin, i.e., rst, we assign values of one to the positive product-level ows in COMEXT, and then we sum them for each pair-year combination. The result is a time-varying bilateral variable, which is de ned as the number of products exported from i to j at year t. The structure of COMEXT, in combination with

[19]We accessed the historical version of the Comext data at https://ec.europa.eu/eurostat/estat-navtree-portlet-prod/BulkDownloadListing?sort=1&dir=comext.

the design of PRODCOM, presented several opportunities to construct and experiment with alternative estimating samples. We describe those opportunities and our choices next.

For each reporting country COMEXT includes separate data on exports and on imports. Based on this information, we construct and experiment with three alternative measures of the extensive margin of trade. First, following most of the literature on the intensive margin of trade ows, our main extensive-margin variable is constructed based on the average between the import and the export ows in COMEXT. Alternatively, we also construct a sample, where we start with the reported export values and we replace the missing exports observations with the corresponding non-missing import values. We call this sample the *\Exporter-based Sample"*, and we experiment with it in the sensitivity analysis. Similarly, we also construct a measure where we start with the reported import values and we replace the missing import observations with the corresponding non-missing export values. We call this sample the *\Importer-based Sample"*. As we demonstrate later, estimates obtained with the three alternative samples are very similar to each other.

The key novelty of our analysis is the introduction of the *the domestic extensive margin*. Thus, it is very important for our purposes to construct a consistent correspondence between the domestic and the international extensive margins. To this end, we bene ted tremendously from the fact that the two main underlying databases (PRODCOM and COMEXT) that we used to build our estimating samples were designed to be consistent with each other by construction. Speci cally, as noted in the PRODCOM user guide, *\[b]efore data collection could begin, it was necessary to draw up a common list of products to be covered ... As PRODCOM statistics have to be comparable with external trade statistics, which are based on the Combined Nomenclature (CN), there had to be a close relationship between the two nomenclatures."* We took advantage of the close matching and existing concordances between the PRODCOM and the CN classi cations to construct consistent estimating samples that cover both the domestic and the international extensive margins.[20]

While the matching between PRODCOM and CN was intended to be very close by design, *"it was felt by the PRODCOM committee that there were instances where the CN classification gave too much detail in how it broke down products within a specific category, but equally instances when it did not give enough detail to meet the needs of the likely end users of PRODCOM data."* (p.6, PRODCOM Guide). As a result, the matching between the PRODCOM classification and the Combined Nomenclature includes one-to-one matches, many CN to one PRODCOM matches, one CN to many PRODCOM matches, and many CN to many PRODCOM matches. There was also a small fraction of products of the PRODCOM categories that did not have a match in the Combined Nomenclature.[21]

As expected, an investigation of the matching patterns between PRODCOM and CN reveals that most of the cases are one-to-one matches and the second largest share includes many CN to one PRODCOM matches. In combination, these two types of matches cover between 78.5% and 100% of the PRODCOM product categories for which there is a CN match.[22] Therefore, we constructed and experimented with two alternative estimating samples based on the underlying product matching and coverage between the PRODCOM and the CN classifications. The first sample includes only the products for which we have one-to-one matching. The number of products that we cover this way is around 2000 in each year of our sample. We label this sample the *"Conservative-product sample"*, and we use it in the robustness analysis. The second sample, which is the one used in the main analysis, is our *"Extended-product sample"* because it covers all products for which we have one-to-one or one-to-many matching between PRODCOM and CN, i.e., in the latter case there are multiple CN products corresponding to a single PRODCOM category. The number of products in this extended sample varies between 3276 and 3513, thus covering almost all (between 93% and 100%) possible products in the original PRODCOM classification for which there was a

&IntCurrentPage=4.

[21] Specifically, the fraction of PRODCOM products that could not be matched to the CN classification varies between 6.57%, in 2016, and 9.58% in 2010.

[22] Specifically, they cover more than 97% of the PRODCOM products in 2008 and 2009, 100% of the PRODCOM products in 2010 and 2011, more than 80% of the PRODCOM products between 2012 and 2016, and 78.5% of the PRODCOM products in 2017 and 2018.

CN match.

Based on the *"Extended-product sample"* used to obtain our main results, the last column of Table 1 reports the percentage change in the total number of exported products for each country in our sample during the period 2008-2018, i.e., on the international extensive margin (IEM) of trade. While we use bilateral international extensive margin data in our econometric analysis, the percentage changes in the total number of exported products that we report here are informative for at least two reasons. First, according to the indexes in the last column of Table 1, the countries in our sample can be classified in three distinct and sufficiently large groups: (i) We see some countries that experienced a very significant increase in the international extensive margin between 2008 and 2018, e.g., Montenegro and North Macedonia, followed by Croatia and Hungary; (ii) The second group of countries did not experience significant change in the number of exported products between 2008 and 2018. Some examples include, Lithuania, Bosnia and Herzegovina and Bulgaria; (iii) Finally, a number of countries saw a decrease in the number of exported products during the period of investigation. The decrease is moderate and occurs mostly in developed countries, e.g., France, Germany, UK, Austria. Based on these results, we draw the intuitive conclusion that the countries that have benefited the most on the international extensive margin of trade are smaller and poorer European economies, while the larger and more developed countries have actually contracted the number of products that they export.

Second, comparisons between the percentage changes in the last two columns of Table 1 reveal some interesting patterns of the relationship between the evolution of the international and the domestic extensive margins of trade. These patterns motivate the econometric analysis below that identifies the relative impact of globalization and European integration on the international relative to the domestic margin of trade. The percentage changes in the last two columns of Table 1 reveal the following patterns: (i) Faster positive growth on the external margin and slower positive growth on the domestic margin, e.g., North Macedonia; (ii) No change on the external margin but an increase on the domestic margin, e.g.,

Finally, inspection of COMEXT revealed that there were export data for Cyprus, Luxembourg, and Malta, even though, as discussed earlier, these countries were not present in PRODCOM. We also noticed that for many countries the number of exported products in COMEXT exceeded the number of domestically produced products in PRODCOM. This motivated us to implement an alternative procedure to construct the domestic margin of trade, which further demonstrated the robustness of our main findings. Specifically, for each year-country combination we constructed the domestic margin of trade as the total number of products exported by this country to any other country in the world. The implicit assumption that we make when implementing this procedure is that any product that is produced in a given country is exported to at least one trading partner. We view this assumption as plausible for two reasons. First, because our sample covers mining and manufacturing (and not agriculture and especially services, where localized consumption is a more significant problem for trade). And second, because usually every country declares exports to its most closely related partner in almost every category.

The proposed procedure also has several important advantages. First, by construction, it ensures that the number of internationally-traded products will always be smaller and in rare cases equal (i.e, when the maximum number of products exported to a specific trading partner is the same as the total number of exported products). Second, it will enable us to construct the domestic margin of trade for all 35 countries and all years that are covered in the original PRODCOM database. Third, the procedure allows for the construction of the domestic extensive margin based only on international trade data. Thus, in principle, it can be used to construct the domestic extensive margin for a very large number of countries, as long as the underlying international trade data are available for all pairs. This is not the case in our sample, because COMEXT does not include trade between non-EU countries. Therefore, we only experiment with a sample that covers the original 35 countries from PRODCOM and the extended product list.

margin. Finally, we also control for additional time-varying bilateral variables (e.g., economic integration agreements, EIAs, and membership in the world trade organization, WTO). These control variables come from the dynamic gravity database of the U.S. International Trade Commission, c.f., Gurevich and Herman (2018). We do note, however, that given the speci cs of our sample (i.e., covering only European economies) and the use of country-pair xed e ects, the estimates of the EIA and WTO covariates would be identi ed of very few observations. For example, all the variation in the EIA covariate could come from the trade agreements of very few countries including Bosnia and Herzegovina, Croatia, Montenegro, and Serbia. Similarly, Montenegro is the only country in our sample that became a WTO member during the period of investigation (in 2012). We will capitalize on this in order to demonstrate that the introduction of the domestic extensive margin will enable us to identify country-speci c WTO e ects.

the external/cross-border extensive margin differentially relative to the domestic extensive margin. Even though the set of country-year-specific globalization dummies does not allow us to identify the effects of specific policies, we find their use appropriate to capture the integration processes in Europe. From a methodological perspective, note that none of these effects could be identified without the use of observations on the domestic extensive margin.

A finding that $_{i;t} > 0$ implies that there is a *relative* increase in the international extensive margin relative to the domestic extensive margin. In principle, a positive estimate of $_{i;t}$ may reflect several scenarios, e.g., (i) faster growth on the external margin and slower growth on the domestic margin; (ii) no change on the external margin but a decrease on the domestic margin, (iii) growth on the international extensive margin and no change on the domestic extensive margin; (iv) growth on the international extensive margin and decrease on the domestic extensive margin; (v) decrease on the international extensive margin and faster decrease on the domestic extensive margin, etc. Based on the DEM and the IEM indexes and their relationship that we discussed in the data section, we saw examples of each of those scenarios and we will return to them when we interpret our results. In sum, what we can identify is the effects of globalization/European integration on the international *relative* to the domestic extensive margin. Finally, we note that, due to perfect collinearity with the country-pair fixed effects, we have to omit the border estimate for one year for each

domestic extensive margin observations we obtain estimates of the impact of globalization on

the marginal e ect of the globalization estimate in 2018, which captures the total impact of globalization during the period of investigation. The marginal e ect is 270.797 (std.err. 38.050), which means that, on average, the number of internationally traded products increased by about 271 relative to the number of domestically traded products during the period of investigation, or about 7.2 percent of the total number of possibly traded products in 2018.

Another important result from column (1) of Table 2 relates to our estimate of the impact of WTO. We remind the reader that Montenegro is the only country in our sample, which joined the WTO during the period of investigation. Thus, from a policy perspective, the positive and signi cant estimate that we obtain (0.390, std.err. 0.054) implies that WTO membership has bene ted trade diversi cation for this country.

The rest of the columns in Table 2 o er estimates from a series of sensitivity experiments designed to test the robustness of our main ndings. Broadly, we split our robustness checks in two categories: (i) alternative estimators, which are reported in panel B of Table 2; and (ii) alternative samples, which are reported in panel C of Table 2. Speci cally, the results in columns (2) to (4) of panel B are obtained with the Tobit, the PPML, and the OLS estimators, respectively. We also experiment with the following alternative estimating samples: (i) the \Exporter-based Sample", as de ned in the Data section, in column (5);

stand out. First, we note that, overall, the results in column (11) con rm our main conclusions. However, second, we notice that the monotonic increase in the globalization estimates is violated in 2017, where the estimate on GLOB_2017 is still statistically signi cant but smaller as compared to the estimates on GLOB_2016. Inspection of the underlying domestic margin data reveals some unusual patterns. Speci cally, as illustrated in Figure 1, which graphs the yearly percentage changes in the domestic extensive margin for all countries in our sample, there are four unusual spikes in 2017, which are for Cyprus, Malta, Iceland, and Montenegro. Column (12) of Table 1 reproduces the results from column (11) but without the outliers. The monotonically increasing pattern of the globalization estimates is restored.

Finally, the estimates in the last column of Table 2 are obtained with an estimating sample that does not include the domestic extensive margin. Consistent with our main argument and contribution related to the bene ts and importance of properly accounting for the theory-consistent domestic extensive margin of trade, the estimates in column (13) reveal that without the *DEM* observations we cannot identify neither the globalization e ects that we are after nor the country-speci c impact of WTO on Montenegro's extensive margin. The only covariate whose e ects we can still identify is the bilateral EIA variable. It should be noted, however, that even though our EIA estimates in column (13) and the main results in column (1) of Table 2 are very similar, this does not necessarily need to be the case, as EIAs may have a di erential impact on the domestic vs. the external extensive margin.[27]

4.2 Country-speci c Globalization E ects

Consistent with the theory, the main speci cation allows for di erential, country-speci c e ects of globalization. Thus we employ $\sum_{i;t} {}_{i;t} GLOB_{i;t} BRDR_{ij}$, where the globalization estimates, ${}^{\wedge}_{i;t}$, now vary not only for each year but also for each country in our sample. Due to perfect collinearity with the country-pair xed e ects, we need to drop one border estimate for each country and our choice are the country e ects for 2008. Thus, all other country-

[27] We believe that a detailed analysis of the impact of globalization on the domestic extensive margin would be interesting and informative, but it is beyond the scope of this paper.

speci c globalization estimates should be interpreted as deviations from the corresponding border e ect for the same country in 2008 and, by construction, the estimates for 2018 would capture the total (cumulated) e ects during the period 2008-2018. The results appear in Table 3.

The main implication of the estimates in Table 3 is wide heterogeneity of estimated globalization e ects, mostly statistically signi cant. To facilitate discussion, we focus on the cumulative e ects for 2018 from the last column of Table 3, and we plot them in Figure 2. The gure enables us to group the countries in our sample in four categories.

The rst group includes ve countries, Austria, Belgium, Sweden, France and Germany { the only countries for which we obtain negative globalization estimates for 2018. In fact,

European integration for development and inequality.

To test the robustness of our results, we reproduce the results from Table 3 with alternative estimators and with alternative samples. For clarity and simplicity of exposition, we do not report all results but, instead, we focus on one representative country from each of the three groups that we identi ed in Figure 2, and we present our ndings graphically. Specifically, we chose Sweden, Ireland, and Portugal. Figure 3 visualizes the estimates that we obtain with the four alternative estimators. Panel A presents our main estimates from Table 3, which are obtained with the *FLEX* estimator. The estimates in Panel B are obtained with the Tobit estimator. PPML estimates appear in Panel C. Finally, the results in Panel D are obtained with the OLS estimator. Based on the estimates in Figure 3, we conclude that our main ndings about the (heterogeneous) impact of globalization on the extensive margin are robust to the use of alternative estimators.

The results in Figures 4 and 5 are obtained with alternative estimating samples. In particular, Panel A of Figure 4 visualizes estimates from our \Exporter-based sample" as described in the Data section. Panel B instead uses the \Importer-based sample" . The results in Panel C are based only on the positive observations in the main sample. The estimates in Panel D are obtained with three-year interval data. The estimates in Panel E use the \Conservative-product" sample. Finally, the results in Panel F are obtained with the \Extended-country" sample. Based on the estimates in Figure 4, we conclude that our main ndings about the (heterogeneous) impact of globalization on the extensive margin are robust to the use of these alternative estimating samples.

Figure 5 reports estimates that are based on two samples with alternative de nitions of the domestic extensive margin. Speci cally, the estimates in Panel A of Figure 5 are obtained with our main sample, where the DEM is constructed directly from the raw PRODCOM data, while the results in Panel B are obtained from a sample where the DEM is constructed as the total number of products that are exported based on the COMEXT export, as described in Section 3. Two main ndings stand out from Figure 5. First, we see that the estimates and

their evolution over time is comparable, between the two panels, for Ireland and for Portugal. However, second, the evolution of the globalization estimates for Sweden is quite di erent. The natural explanation for this result is, of course, the di erence in the construction of the domestic extensive margin. Comparison between the evolution of DEM for Sweden depending on the construction method reveals that the number of products that Sweden produces has fallen in both cases. However, the decrease is almost three times larger (i.e., by 188 vs. 65 products) in the export-based DEM sample. This explains the di erence between the two panels and points to the importance of proper measurement of the domestic extensive margin.

4.3 On the Heterogeneous Impact of EU Membership

We conclude the empirical analysis with an investigation of the extensive margin e ects of European integration on the new EU members. Given the period of investigation, we focus on three countries, including Bulgaria and Romania, which both joined in 2007 (the year before the start of our sample), and Croatia, which joined in 2013. Even though the sample of new EU members is small, we nd the analysis instructive both from a methodological and from a policy perspective. In order to emphasize some important aspects of our speci cations and corresponding estimates, we develop the analysis sequentially, in three speci cations. The estimates are presented in Table 4. Each of the three panels in Table 4 reports estimates from a di erent speci cation. The dependent variable is always the number of products sold from exporter i to importer j, including domestic sales, and all estimates are obtained with the Flex estimator of Santos Silva et al. (2014). All speci cations include exporter-time,

2, we now introduce a set of border dummies for trade between the three new EU members and the old EU members. For brevity, we only report the estimates of the globalization

extensive margin in order to challenge the standard assumption in the trade literature that, before exporting a given product, firms are already necessarily selling this product domestically. This idea is motivated by anecdotal evidence that points to alternative scenarios, e.g., where some products are simultaneously offered for sale on the domestic and on the foreign markets, or even when products are first exported and only then they are sold domestically. We believe that, in combination with theory, our new dataset that combines the international and the domestic extensive margin can provide interesting insights in this direction.

To perform the empirical analysis we constructed a dataset covering the domestic (and international) extensive margin for the European economies. We see significant potential benefits from expanding the dataset to cover all possible countries in the world. For example, one clear advantage of such database would be that it will include the poorer and less-developed economies, where export diversification and the extensive margin are particularly important. We believe that the creation of such extended dataset is feasible and, in fact, significantly easier and more reliable as compared to a corresponding dataset on the intensive

ratio between the sum of the number of exported products and the number of imported products divided by the number of domestically produced products. We see this index as the extensive margin counterpart of the standard Openness to Trade (OTT) index that is

References

Agnosteva, Delina E., James E. Anderson, and Yoto V. Yotov , \Intra-national Trade Costs: Assaying Regional Frictions," European Economic Review, 2019, 112, 32{50.

Anderson, James E. , \A Theoretical Foundation for the Gravity Equation," American Economic Review, 1979, 69 (1), 106{116.

_ and Eric van Wincoop , \Gravity with Gravitas: A Solution to the Border Puzzle," American Economic Review, 2003, 93 (1), 170{192.

Anderson, James E. and Yoto V. Yotov , \Short Run Gravity," Journal of International Economics, 2020, 126

Gurevich, Tamara and Peter Herman , \The Dynamic Gravity Dataset: 1948-2016," USITC Working Paper 2018-02-A, 2018.

Yotov, Yoto V. , \A Simple Solution to the Distance Puzzle in International Trade," Economics Letters, 2012, 117 (3), 794{798.

Table 1: The Domestic Extensive Margin, 2008-2018

Country Name	ISO	2008	2009	2010	2011	2012	2013	2014	2015	2016	2017	2018	% DEM	% IEM
Austria	AT	0.51	0.51	0.51	0.51	0.51	0.51	0.51	0.51	0.52	0.52	0.52	1.69	-6.32
Bosnia and Herzegovina	BA	.	.	.	0.19	0.19	0.19	0.19	0.19	0.20	0.19	0.19	3.31	-1.05
Belgium	BE	0.49	0.49	0.48	0.48	0.49	0.49	0.49	0.49	0.49	0.50	0.50	2.63	-6.31
Bulgaria	BG	0.48	0.48	0.48	0.48	0.48	0.48	0.49	0.49	0.49	0.49	0.49	3.01	

Table 2: Globalization/European Integration and the Extensive Margin of Trade, 2008-2018

	A. Main	B. Alternative Estimators			C. Alternative Samples								
	FLEX	TOBIT	PPML	OLS	EXPRT	IMPRT	PSTV	INTLRV	PROD	CNTRY	DEMEXP	NOOUTL	NO DEM
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)	(13)
BRDR_2009	-0.009	-4.560	-0.000	-0.012	0.006	-0.026	-0.009	-0.027	-0.010	-0.011	0.017	0.017	
	(0.009)	(7.514)	(0.006)	(0.011)	(0.011)	(0.008)	(0.009)	(0.009)	(0.009)	(0.011)	(0.011)	(0.011)	
BRDR_2010	0.072	42.212	0.034	0.040	0.082	0.062	0.072		0.068	0.070	-0.046	-0.045	
	(0.025)	(14.843)	(0.011)	(0.018)	(0.020)	(0.035)$^{+}$	(0.025)		(0.025)	(0.022)	(0.034)	(0.033)	
BRDR_2011	0.117	71.679	0.051	0.067	0.133	0.104	0.117		0.122	0.106	0.064	0.065	
	(0.030)	(17.690)	(0.014)	(0.025)	(0.025)	(0.040)	(0.030)		(0.030)	(0.027)	(0.034)$^{+}$	(0.034)$^{+}$	
BRDR_2012	0.154	102.959	0.067	0.076	0.175	0.141	0.154	0.146	0.171	0.142	0.292	0.289	
	(0.034)	(19.107)	(0.016)	(0.031)	(0.033)	(0.040)	(0.034)	(0.042)	(0.033)	(0.031)	(0.043)	(0.042)	
BRDR_2013	0.199	129.159	0.088	0.129	0.214	0.191	0.199		0.224	0.189	0.326	0.323	
	(0.040)	(22.612)	(0.019)	(0.035)	(0.039)	(0.045)	(0.040)		(0.039)	(0.037)	(0.044)	(0.044)	
BRDR_2014	0.229	145.953	0.102	0.150	0.255	0.212	0.229		0.256	0.221	0.366	0.362	
	(0.043)	(24.266)	(0.021)	(0.036)	(0.043)	(0.047)	(0.043)		(0.043)	(0.040)	(0.048)	(0.047)	
BRDR_2015	0.287	177.271	0.129	0.204	0.319	0.265	0.287	0.274	0.320	0.279	0.409	0.405	
	(0.047)	(26.322)	(0.025)	(0.040)	(0.049)	(0.051)	(0.047)	(0.052)	(0.048)	(0.044)	(0.048)	(0.047)	
BRDR_2016	0.325	198.697	0.147	0.255	0.367	0.296	0.325		0.370	0.318	0.483	0.477	
	(0.052)	(28.392)	(0.029)	(0.047)	(0.055)	(0.055)	(0.052)		(0.054)	(0.049)	(0.049)	(0.048)	
BRDR_2017	0.429	254.843	0.186	0.314	0.456	0.422	0.429		0.495	0.422	0.399	0.492	
	(0.072)	(40.638)	(0.037)	(0.053)	(0.066)	(0.095)	(0.072)		(0.081)	(0.071)	(0.077)	(0.066)	
BRDR_2018	0.456	268.345	0.200	0.339	0.479	0.454	0.456	0.466	0.525	0.449	0.579	0.574	
	(0.065)	(37.192)	(0.038)	(0.057)	(0.065)	(0.082)	(0.065)						

Figure 1: Export-based DEM Indexes. Percentage Changes, 2008-2018

Notes: This figure visualizes the country-specific percentage changes in the domestic extensive margin indexes that are constructed based on export data. See text for discussion and further details.

Table 3: Country-specific Globalization Effects on the Extensive Margin of Trade, 2008-2018

ISO	2009	2010	2011	2012	2013	2014	2015	2016	2017	2018
AT	0.004	0.019	-0.018	-0.060	-0.057	-0.045	-0.055	-0.092	-0.133	-0.133
	(.013)	(.016)	(.021)	(.028)*	(.029)*	(.035)	(.036)	(.039)*	(.046)**	(.047)**
BA				0.149	0.199	0.301	0.300	0.288	0.470	0.389
				(.059)*	(.064)**	(.066)**	(.065)**	(.073)**	(.112)**	(.097)**

Figure 2: Globalization and the Extensive Margin. Country-speci c E ects, 2018

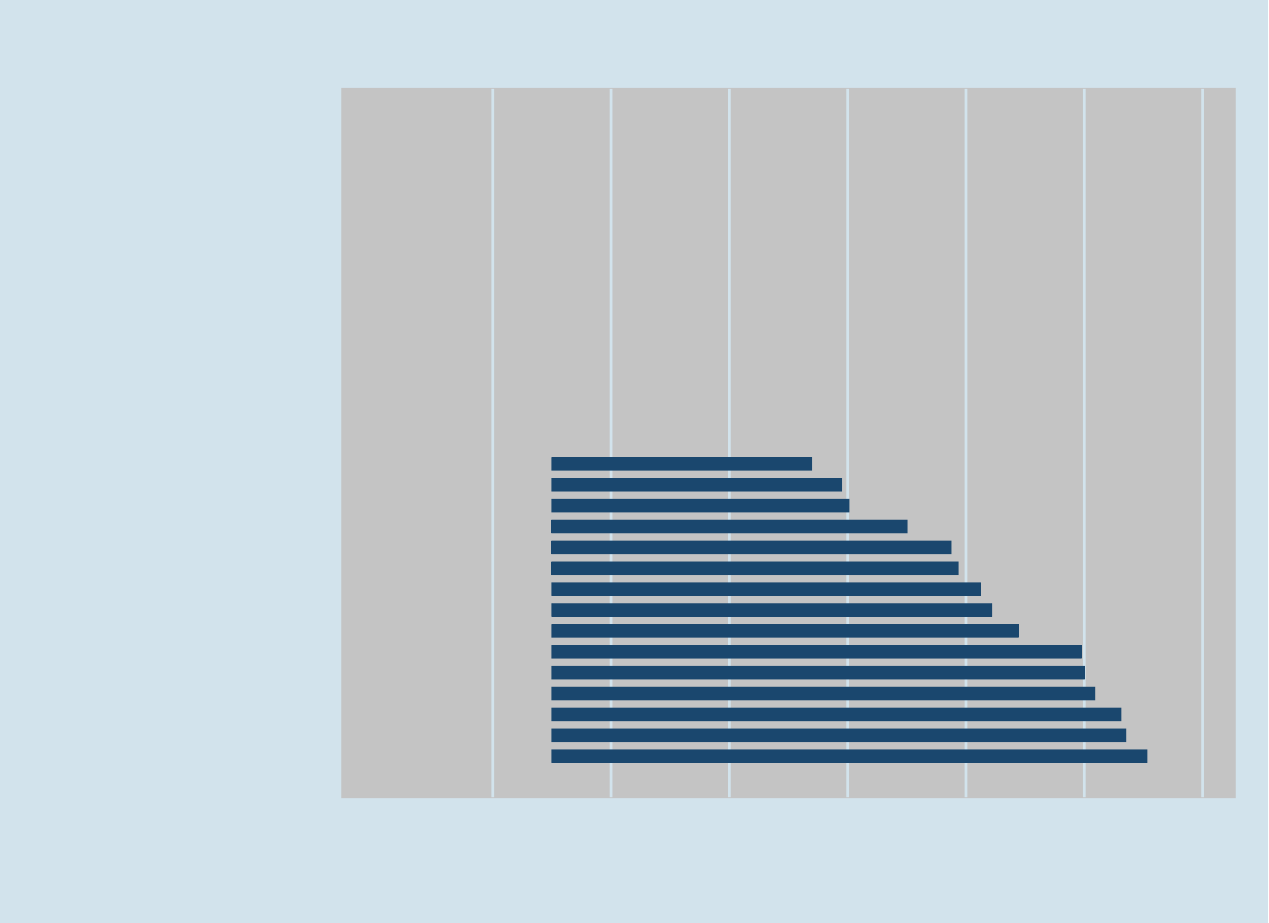
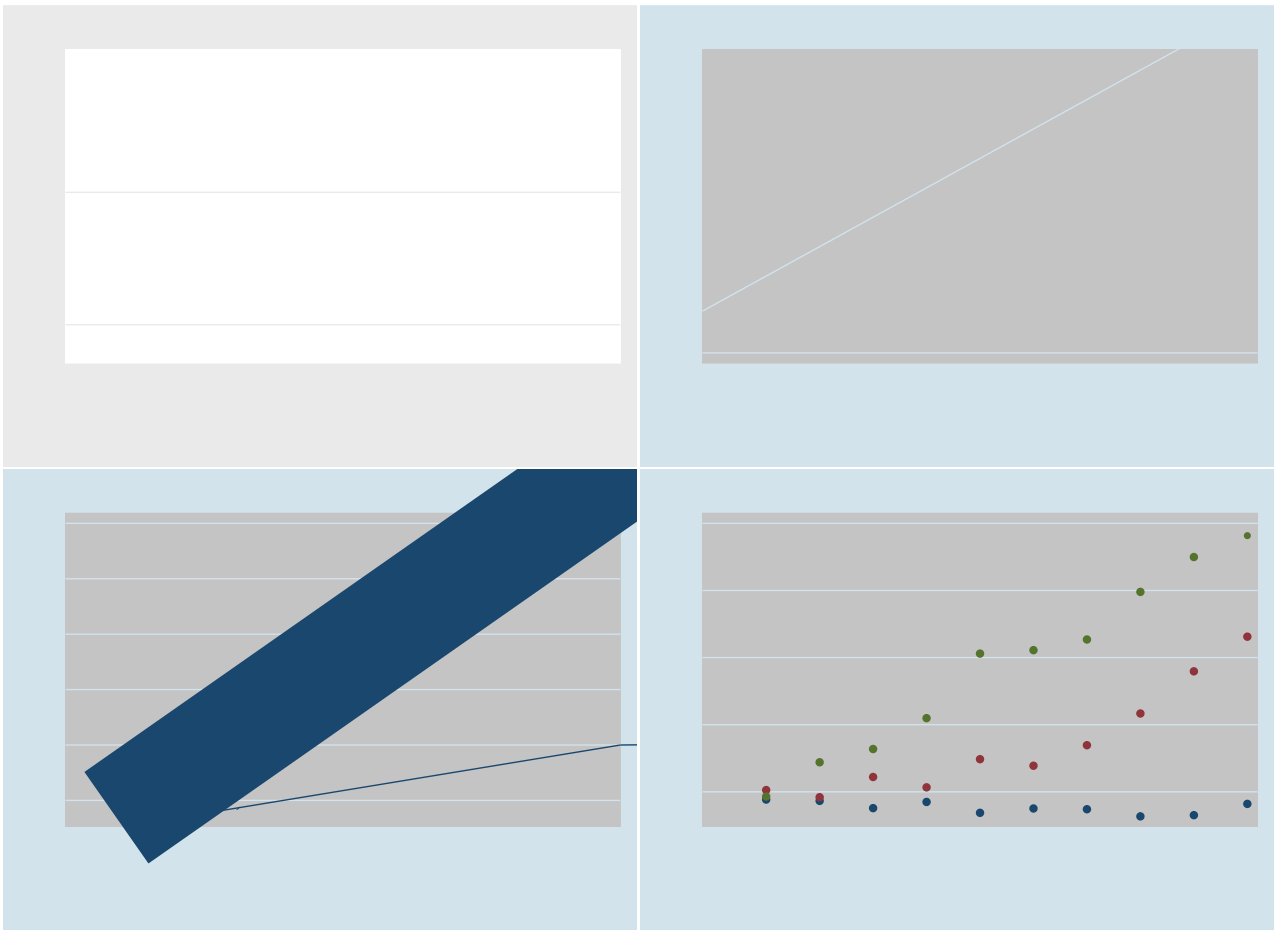


Figure 3: Country-speci c Estimates, Robustness: Alternative Estimators

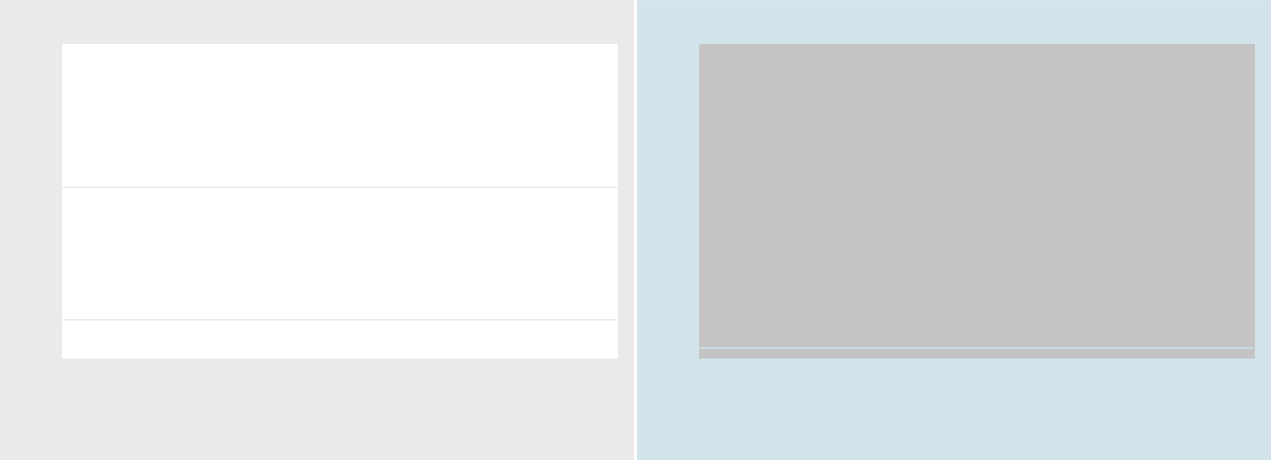


Notes: This gure visualizes the country-speci c estimates of the globalization e ects on the extensive margin for a selected group of countries including Sweden, Ireland, and Portugal. All estimates are obtained from equation (7), with exporter-time, importer-time, and pair xed e ects in a panel setting for all years and all countries in the sample. The di erence between the four panels in the gure are due to the use of alternative estimators. Speci cally, Panel A visualizes our main estimates, which are obtained with the Flex estimator of Santos Silva et al. (2014). The estimates in Panel A are in fact those from Table 3. The estimates in Panel B are obtained with the Tobit estimator. PPML estimates appear in Panel C. Finally, the results in Panel D are obtained with the OLS estimator. See text for further details.

Figure 4: Country-speci c Estimates, Robustness: Alternative Samples

Notes: This gure visualizes the country-speci c estimates of the globalization e ects on the extensive margin for a selected group of countries including Sweden, Ireland, and Portugal. All estimates are obtained from equation (7), with exporter-time, importer-time, and pair xed e ects in a panel setting for all countries in the sample. The di erence between the six panels of the gure are due to the use of alternative estimating samples. Speci cally, Panel A visualizes estimates from our \Exporter-based sample" as described in the Data section. Panel B instead uses the \Importer-based sample" . The results in Panel C are based only on the positive observations in the main sample. The estimates in Panel D are obtained with three-year interval data. Panel E uses the \Conservative-product" sample. Finally, the results in Panel F are obtained with the \Extended-country" sample. See text for further details.

Figure 5: Country-speci c Estimates, Robustness: DEM De nition



Notes: This gure visualizes the country-speci c estimates of the globalization e ects on the extensive margin for a selected group of countries including Sweden, Ireland, and Portugal. All estimates are obtained from equation (7), with exporter-time, importer-time, and pair xed e ects in a panel setting for all countries in the sample. The di erence between the two panels of the gure are due to the de nition/construction of the domestic extensive margin. Speci cally, the estimates in Panel A are our main estimates, which are obtained with a DEM measure based on production data from PRODCOM, while the results in

Table 4: European Integration and the Extensive Margin for New EU Members

Year	A. Symmetric	B. Asymmetric		C. Country-speci c					
				Bulgaria (2007)		Romania (2007)		Croatia (2013)	
		Imp.EU	Exp.EU	Imp.EU	Exp.EU	Imp.EU	Exp.EU	Imp.EU	Exp.EU
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
2009	0.00200	-0.0270	0.0330	-0.00100	-0.192	0.0770	0.0560	-0.185	0.369
	(.015)	(.054)	(.059)	(.061)	(.099)+	(.069)	(.066)	(.088)*	(.095)**
2010	0.0260	-0.0200	0.0760	0.0400	-0.0230	0.0950	0.0370	-0.217	0.314
	(.019)	(.057)	(.058)	(.067)	(.094)	(.071)	(.076)	(.089)*	(.104)**
2011	0.0260	0.395	-0.339	0.499	-0.472	0.469	-0.273	0.168	-0.395
	(.024)	(.082)**	(.084)**	(.141)**	(.121)**	(.154)**	(.143)+	(.104)	(.141)**
2012	0.0290	0.378	-0.322	0.393	-0.407	0.332	-0.228		
	(.033)	(.069)**	(.077)**	(.107)**	(.113)**	(.115)**	(.127)+		
2013	0.0430	0.386	-0.297	0.407	-0.318	0.407	-0.197	0.188	-0.520
	(.026)	(.06)**	(.061)**	(.118)**	(.095)**	(.133)**	(.104)+	(.1)+	(.122)**
2014	-0.00100	0.355	-0.356	0.477	-0.286	0.306	-0.230	0.109	-0.707
	(.034)	(.062)**	(.067)**	(.102)**	(.092)**	(.102)**	(.115)*	(.107)	(.116)**
2015	0.0370	0.356	-0.282	0.401	-0.235	0.307	-0.198	0.175	-0.561
	(.034)	(.056)**	(.064)**	(.099)**	(.095)*	(.087)**	(.118)+	(.113)	(.112)**
2016	0.0740	0.401	-0.251	0.393	-0.226	0.324	-0.198	0.294	-0.466
	(.035)*	(.059)**	(.062)**	(.097)**	(.097)*	(.079)**	(.119)+	(.124)*	(.103)**
2017	0.136	0.428	-0.161	0.410	-0.120	0.311	-0.149	0.425	-0.305
	(.04)**	(.073)**	(.067)*	(.092)**	(.105)	(.1)**	(.137)	(.165)**	(.106)**
2018	0.160	0.424	-0.104	0.389	-0.0860	0.335	-0.0750	0.343	-0.295
	(.042)**	(.067)**	(.066)	(.103)**	(.103)	(.096)**	(.123)	(.137)*	(.11)**

Notes: This table reports estimates of the impact of European integration on the extensive margin of trade for the three most recent EU members in our sample, i.e. Bulgaria (2007), Romania (2007), and Croatia (2013). Each panel of the table reports estimates from a di erent speci cation. The dependent variable in each speci cation is

k_j to each destination j. Distribution presupposes production denoted `destination' 0. (Firms have identical per- rm capital k_j because prior to receiving their productivities, all rms are identical.) As the period of analysis opens each rm draws a Hicks-neutral productivity scalar% After the productivities are drawn, rms hire labor at wage rate w to produce and distribute the good, equating w to the value of marginal product of labor for production and for distribution to each destination.

Index the rms by their productivity draws % The pro t of rm %on sales toj using variable labor

The aggregate value sellers of trade shipped to destination j is given by integrating the value of marginal product of variable labor over rms % Use (9) in (10) and simplify to yield the supply function:

$$X_j = Ak_j M_j^{k_j}$$

Short run gravity is obtained by solving the market clearing equation for $w_i^{(1-\sigma)}$. First replace $P_j^{(1-\sigma)} E_j$ with $E_j P_j^{(1-\sigma)}$ in (14). Then sum (14) over j and solve:

$$w_i^{(1-\sigma)} = \frac{Y_i/Y}{\sum_j (t_{ij}/P_j)^{(1-\sigma)} (E_j/Y) \Phi_{ij}^{\sigma-1}} = \frac{Y_i/Y}{\Pi_i^{(1-\sigma)}}: \quad (15)$$

Here $\Phi_{ij} = U_{ij} M_{ij} k_{ij} / \sum_j U_{ij} M_{ij} k_{ij}$. Note that $\Phi_{ij} = \phi_{ij} U_{ij}/U_i$ where $\phi_{ij} = K_{ij}/K_i$, the ex ante capital share and $U_i = \sum_j U_{ij} \phi_{ij}$, the average utilization rate of capital. Thus $\Pi_i^{(1-\sigma)} = \sum_j (t_{ij}/P_j)^{(1-\sigma)} (E_j/Y) \Phi_{ij}^{\sigma-1}$ is the sellers multilateral resistance in the heterogeneous firms case.

Substitute the right hand side of (15) for $w_i^{(1-\sigma)}$ in (14). The result is short run gravity for the heterogeneous firms case.

$$X_{ij} = \frac{Y_i E_j}{Y} \left(\frac{t_{ij}}{\Pi_i P_j} \right)^{(1-\sigma)} \left(\frac{U_{ij}}{U_i} \right)^{\sigma-1}$$

returns may di er because utilization rates U_{ij} may di er in this allocation. [33] Fully e cient ex ante investment requires equal utilization rates $U_{ij} = U_i$) $k_{ij} = {}_i a_{ij}$; $8i; j$; ${}_i > 1$: ${}_i > 1$ is required for the marginal product of capital to be positive, since e ectively the period-by-period xed cost of labor absorbs a portion of capital in order to a ect ex post utilization. Formalizing the implications:

Proposition 1

$k_{ij} = {}_i a_{ij}$, ${}_i > 1$ and ${}_{ij} = s_{ij}$ is necessary and su cient for ex ante e ciency.

With ex ante e cient investment, long run gravity obtains. In (16), fully e cient investment ${}_{ij} = X_{ij} = Y_i$) implies a solution to (16) as if the exponent $= 1$ and multilateral resistance ${}_i ; P_j$

The proof follows from

$$s_{in} = {}_{in} = {}_{in} = \frac{E_n}{Y} \quad \frac{t_{in}}{{}_i P_n} \quad {}^{(1\ )} \quad (\ {}_{in}$$