



**HAL**  
open science

## Proximity as a Source of Comparative Advantage

Elizaveta Archanskaia

► **To cite this version:**

| Elizaveta Archanskaia. Proximity as a Source of Comparative Advantage. 2013. hal-01070440

**HAL Id: hal-01070440**

**<https://hal-sciencespo.archives-ouvertes.fr/hal-01070440>**

Preprint submitted on 1 Oct 2014

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



Distributed under a Creative Commons Attribution - NoDerivatives | 4.0 International License



**SciencesPo.**

Department of Economics

*Discussion paper 2013-05*

# **Proximity as a Source of Comparative Advantage**

**Liza Archanskaia**

*Sciences Po Economics Discussion Papers*

# Proximity as a Source of Comparative Advantage \*

Liza Archanskaia<sup>†</sup>  
SciencesPo-OFCE

February 2013

## Abstract

This paper establishes that production unbundling has coincided with an inscreasing role of input costs in shaping the pattern of comparative advantage. I show that the wedge in the cost of the input bundle across countries in a multisectoral Ricardian model is given by a composite index of trade frictions incurred in sourcing inputs. As the cost share of inputs is sector-specific this wedge becomes source of comparative advantage whereby countries characterized by relatively high proximity to input suppliers specialize in sectors which use inputs more intensively. I find robust empirical evidence that the input cost channel has growing importance over 1995-2009. Nonetheless, consistently with the fundamental intuition of Ricardian models, the ranking of relative sectoral technology stocks still determines intersectoral specialization. Between 53-55% of intersectoral variation in relative sectoral exports is explained by technology while the input cost channel contributes 3 to 8% in the full sample, and 3 to 13% for the EU-15.

*Keywords:* Ricardian model, Intersectoral specialization, Trade costs

*JEL codes:* F10,F15

---

\*I am grateful to Thomas Chaney, Arnaud Costinot, Mathieu Crozet, Stefania Garetto, Christian Gormsen, James Harrigan, Jacques Le Cacheux, Philippe Martin, Thierry Mayer, Eduardo Morales and seminar participants at SciencesPo Economics department for advice and suggestions for improving this paper. The usual disclaimer applies.

<sup>†</sup>Email: archanskaia@gmail.com. Affiliation: PhD candidate at SciencesPo Economics department, affiliated with the French Economic Observatory (OFCE).

# 1 Introduction

## 1.1 What this paper does

This paper belongs to the strand of literature which, following the seminal work by Eaton and Kortum (2002, 2010) and Costinot (2009), seeks to identify the relative importance of technology, factor endowments, and trade costs in determining the pattern of specialization on world markets in many-good many-country Ricardian models. This approach allows defining a theoretically grounded measure of revealed comparative advantage, as in Chor (2010) and Costinot et al. (2012), and equips the researcher with a flexible tool to quantify the relative importance of fundamental country characteristics in determining the pattern of intersectoral specialization.

Costinot (2009) has developed a unifying framework which delivers a strong result in terms of intersectoral specialization when the primitives of the model which are technology and factor endowments are characterized by logsupermodularity. Specifically, if countries can be ranked in terms of a single characteristic, such as the quality of their institutions, and sectors can be ranked according to a sector-specific characteristic, such as their skill intensity, then if the primitives of the model are such that high-characteristic countries are relatively more likely to be endowed with factors which are relatively more productive in high-characteristic sectors, it is possible to deduce the pattern of specialization in terms of the ranking of relative sectoral output for any pair of countries.

Consequently, Costinot et al. (2012) have shown that the seminal Eaton and Kortum (EK) model extended to a multi-sector set-up with a finite number of sectors and an infinite countable number of differentiated varieties within each sector generates the stark prediction that the ranking of relative sectoral exports for any pair of countries on world markets can be predicted from the ranking of their relative sectoral technology stocks.<sup>1</sup>

The contribution of this paper is to show that with multistage production and trade in inputs, the proximity of the country to world technology defined as its ability to source least cost inputs worldwide, becomes a fundamental country characteristic which co-determines its pattern of comparative advantage, together with domestic technology and labor endowments.

We incorporate multistage production in the Costinot et al. (2012) model in the simplest possible way by assuming that output in each sector is produced using a bundle of inputs and labor. We incorporate a sector-specific

---

<sup>1</sup>This prediction is obtained under the assumption that bilateral trade costs contain a pair specific component common across sectors and a sector-specific component specific to the destination and common across exporters.

feature in the production function by assuming that sectors differ in the way they combine inputs and labor in production.

We show that the only component of the cost of the input bundle which varies across countries is a composite index of trade frictions which the country faces in sourcing inputs from all possible suppliers including itself. As the cost share of inputs is sector-specific, proximity to suppliers matters relatively more in sectors which use inputs relatively more intensively.

Consequently, the interaction of a sector-specific characteristic, the weight of inputs in gross output, with the trade cost magnification channel which works through trade in inputs, generates a ranking whereby countries characterized by high proximity to least cost inputs specialize in sectors which use inputs relatively more intensively, conditional on the distribution of domestic technology and labor endowments.

The magnifying effect of trade frictions in the context of cross-border vertical production segmentation has been studied by Yi (2010). In Yi (2010), this mechanism contributes to determining the co-location of the two production stages, inputs and assembly, and the extent of vertical specialization in countries' trade. In this paper we do not learn much about vertical specialization, but we gain mileage in the ability to separately identify the contributions of domestic technology and proximity to suppliers in determining the pattern of intersectoral specialization.

## 1.2 The Empirical Application

In the empirical analysis, we study the pattern of revealed comparative advantage of the main trade partners of the European Union in 1995-2009. We show how to bring the model to the data to quantify the relative weight of fundamental characteristics which determine the pattern of comparative advantage: domestic sectoral technology stocks, labor endowments by skill, and proximity to world technology. This additional component of comparative advantage which we refer to as the 'proximity mechanism' plays out through differences in the relative ease with which countries can source inputs from the best possible supplier of each variety in the world, interacted with the input intensity characteristic of the sector.

The empirical investigation proceeds in four steps. First, the model is used to derive a theoretically grounded measure of proximity to suppliers for each country which, brought to the data, is found to be very persistent overtime. It establishes a ranking of countries in our sample which reveals relatively high centrality of European countries, and of Central and Eastern European countries in particular, while non-European emerging economies such as China, Brazil, and Mexico are characterized by relatively low cen-

trality. Conceptually, in relative terms, the proximity characteristic is a summary statistic of locational comparative advantage because it captures the cost advantage conferred to the country through its ability to source the cheapest inputs worldwide, relatively to every other country in the world.

Second, we implement a fixed effects approach suggested by Costinot et al. (2012) to identify exporter-sector specific relative production costs which in the framework of our model contain four components: technology, wages, input costs, and exporter-specific trade costs which correspond to the trade restrictiveness the exporter faces on world markets.

Third, we project these relative production costs on the vectors of instrumented sectoral technology stocks and wages to identify the cost component unexplained by technology and factor endowments. In this step of the estimation we obtain the structural parameters of the model: the degree of dispersion in productivity and sectoral input intensities. Our preferred point estimates for the dispersion parameter, 6.7(.4) and 7.3(.5), are consistent with values obtained by previous studies.<sup>2</sup> Estimated input intensities are found to be strongly correlated with the share of expenditure on inputs in gross output computed at the sectoral level.

Fourth, we split the sample in two groups according to the proximity characteristic, and regress estimated residuals of relative sectoral production costs on relative proximity for each pair of exporters while interacting proximity with the input intensity of the sector. If the model correctly describes the pattern of production, the proximity mechanism should determine the pattern of intersectoral specialization conditional on domestic technology and labor costs.

We find robust empirical evidence that countries characterized by relatively high proximity to suppliers specialize in sectors which use inputs relatively more intensively. Further, we find that the proximity mechanism becomes a stronger predictor of relative sectoral rankings in the recent period (2002-2009).

Deardorff (2004) establishes a distinction between ‘global comparative advantage’ defined through relative labor requirements in production under frictionless trade and ‘local comparative advantage’ defined through relative labor requirements for production of landed goods under positive trade costs. Trade costs are paid in local labor inducing changes in relative labor requirements for landed goods.

An additional contribution of this paper is to check whether the pattern of specialization determined by local comparative advantage, i.e. the

---

<sup>2</sup>The preferred point estimate in Eaton and Kortum (2002) (resp. Costinot et al. (2012)) is 8.3 (resp. 6.5). Caliendo and Parro (2012) find 8.2 for manufacturing.

pattern of specialization under positive trade costs and trade in inputs, is different from the pattern which would prevail in a world with trade in inputs but without trade frictions. We decompose the variance of pairwise revealed comparative advantage rankings in the share due to domestic technology, factor endowments, and proximity. The proximity characteristic is indeed a summary statistic of trade frictions which co-determine the pattern of specialization under local comparative advantage by modifying expected sectoral production costs.

The main result is that the pattern of comparative advantage observed in a world with positive trade costs and trade in inputs conforms to the specialization pattern which would prevail at the intersectoral level in a frictionless world.<sup>3</sup> Consistently with the fundamental intuition of Ricardian models, the ranking of relative sectoral technology stocks determines the pattern of intersectoral specialization even under positive trade costs. Nonetheless, sector-specific cost differences induced by the proximity mechanism matter increasingly overtime.

### 1.3 Complementarity to Recent Studies

Our results are complementary to several recent empirical investigations of the mechanisms which shape the pattern of intersectoral specialization.

Harrigan and Evans (2005) and Harrigan (2010) provide empirical evidence for the US market on a demand-side mechanism which shapes countries' specialization on specific destination markets. In their framework, products can be ranked in terms of consumer preference for timely delivery and countries can be ranked in terms of distance to destination, with partners situated closeby characterized by their ability to provide timely delivery (or, alternatively, to provide it at a relatively lower cost). The model predicts that local partners will specialize in sectors where timely delivery is valued relatively more by consumers. In this paper we investigate a different but potentially complementary mechanism of intersectoral specialization driven by proximity to suppliers.

---

<sup>3</sup>Eaton and Kortum (2002) find that in 1990 the world was on a brink of a transition from a situation in which geography played a determining role in defining countries' specialization to the situation in which specialization would be driven by technology. These authors work with a one-sector economy, and define specialization as the labor share in manufacturing. In this paper, we describe specialization patterns within manufacturing. In conformity with the intuition of these authors, we find that specialization across manufacturing sectors is driven by technology even though trade frictions continue to play a non-negligible role.

Johnson and Noguera (2012a,c) document that production linkages in conjunction with proximity play an important role in shaping the pattern of bilateral trade. The authors find that the intensity of international production sharing in bilateral relationships, measured as the fraction of value added in gross exports, is increasing in proximity between source and destination. However, the authors do not investigate whether the extent of production sharing contributes to determining the pattern of countries' intersectoral specialization on world markets. Consequently, they do not check whether the extent of production sharing in each bilateral relationship can be summarized by a synthetic index of trade frictions incurred in sourcing inputs. The analysis conducted in this paper is complementary in that we provide a characterization of the aggregate effect of all bilateral production sharing relationships on the cost of the input bundle in each country.

Chor (2010) works in the framework of a multi-sector Ricardian model to quantify the relative importance of the channels which shape the pattern of intersectoral specialization by determining relative sectoral technology stocks. Consistently with Costinot (2009), the author specifies a functional form which determines sectoral technology stocks as a function of several complementarity mechanisms between country and sector characteristics.<sup>4</sup> In the empirical analysis, Chor (2010) identifies the relative contribution of these different dimensions of complementarity to determining the pattern of specialization. In this paper, we conduct a complementary decomposition exercise in that we provide evidence on the relative contribution of domestic technology and proximity to world technology in determining specialization without opening the black box of what technology is.<sup>5</sup>

Caliendo and Parro (2012) develop a multisector Ricardian model with multistage production and trade in inputs to identify the impact of tariff reductions due to NAFTA on changes in trade patterns and welfare of the US, Canada, and Mexico. Caliendo and Parro (2012) underscore the importance of trade in inputs and intersectoral input-output linkages in magnifying the gains from trade. Indeed, one of the key results of the paper is that welfare gains are 40% lower if this magnification mechanism is unaccounted for.

---

<sup>4</sup>Examples are: the interaction of institutional quality and skilled labor endowment with the technological complexity of the sector; the degree of development of financial markets and the degree of sectoral reliance on external financing.

<sup>5</sup>Chor (2010) looks at the impact of inputs on comparative advantage through the lens of incomplete contracts' theory. He finds that countries with high quality legal systems have a comparative advantage in sectors which use relationship specific inputs relatively more intensively with the idea that such sectors are more dependent on law enforcement efficiency. This mechanism is very different from the input intensity mechanism documented in this paper.



Further, the authors show that differences in sectoral input intensity and the degree of intersectoral linkages in production are crucial for understanding the differential impact of a given tariff reduction across sectors.

In this paper, we follow in the steps of Caliendo and Parro (2012) in pointing out the empirical relevance of explicitly accounting for multistage production and trade in inputs in studying a world characterized by production segmentation across borders and complex intersectoral production linkages. But instead of quantifying the magnification of the gains from trade following trade liberalization, here we focus on quantifying the contribution of sector-specific differences in the cost of the input bundle to determining the pattern of countries' specialization on world markets. Specifically, by making a simplifying assumption on the input-output structure, we show that the only component which drives a wedge in the cost of the input bundle across countries is the country-specific proximity characteristic.<sup>6</sup> It is this proximity characteristic which co-determines the ranking of relative sectoral exports because the wedge in the cost of inputs matters relatively more in sectors which use inputs more intensively. The empirical analysis we conduct in this paper thus focuses on a different dimension for which trade in inputs and intersectoral production linkages may matter, and is complementary to the analysis conducted by Caliendo and Parro (2012).

The paper is structured as follows. The next section outlines the model and derives the measure of proximity to suppliers while section 3 goes over the estimation procedure used in the empirical application. Section 4 gives details on the data we use and on results obtained in the estimation of model parameters. In section 5, we show how to bring the theoretically grounded measure of proximity to the data, discuss countries' ranking according to the proximity characteristic, and report results on the contribution of the proximity mechanism to determining intersectoral specialization. Section 6 conducts a variance decomposition of revealed comparative advantage (RCA) rankings across technology, labor endowments, and proximity to quantify the relative contribution of fundamental country characteristics to determining the pattern of comparative advantage. Section 7 concludes.

## 2 Stylized model

In substance this paper studies two questions. First, we ask under what circumstances the structure of trade costs combined with a sequential production process may constitute a source of comparative advantage. Second,

---

<sup>6</sup>The assumption is that the unit cost of the bundle of inputs is the same in all sectors while the cost share of inputs is sector specific.

we ask to what extent trade frictions contribute to determining the pattern of countries' intersectoral specialization on world markets.

To organize ideas, we use a many-good many-country Ricardian model developed by Costinot et al. (2012) and modified in this paper to incorporate sector-specific production features. We allow for a multistage production process and trade in both inputs and final goods as in the seminal one-sector model of Eaton and Kortum (2002). The main purpose of the model is to derive the microfounded proximity characteristic of each country and to show that this characteristic, interacted with sector-specific input intensity, co-determines the pattern of comparative advantage.

## 2.1 Model set-up

The set-up of the model follows Costinot et al. (2012).<sup>7</sup> There is a finite number of sectors  $k$ , and within each sector there is an infinite countable number of differentiated varieties  $\alpha \in A \equiv \{1, \dots, \infty\}$ . Varieties are produced in perfect competition, and the least cost producer of each differentiated variety supplies the market:

$$p_j^k(\alpha) = \min [c_{ij}^k(\alpha)]$$

where  $c_{ij}^k$  is the unit cost function. The production function of each variety is Cobb-Douglas in inputs and labor. Define the sectoral production cost component common to all varieties  $\omega_i^k$ :

$$\omega_i^k = \nu_i^{1-\zeta^k} P_i^{\zeta^k} \epsilon^k$$

where  $\zeta^k$  is the 'input intensity' characteristic of sector,  $\nu_i$  is the wage and  $P_i$  is the price of the input bundle.

Define the bilateral sector-specific trade friction  $\tau_{ij}^k$ . Assume further that labor productivity of variety  $\alpha$  in sector  $k$ ,  $z^k(\alpha)$ , is drawn from some productivity distribution.

The unit cost of production for this variety is:

$$c_{ij}^k(\alpha) = \frac{\omega_i^k \tau_{ij}^k}{z^k(\alpha)}$$

The lower tier utility (production) function is CES. Sectoral price indices

---

<sup>7</sup>Details on the derivation are provided in the online appendix for the paper. We refer the reader to Costinot et al. (2012) for the benchmark sectoral set-up without multistage production.

are given by:

$$P_i^k = \left\{ \sum_{\alpha \in A} [p_i^k(\alpha)]^{1-\sigma} \right\}^{1/(1-\sigma)}$$

The upper tier utility (production) function is Cobb-Douglas. The overall price index (and cost of the input bundle) is:

$$P_i = \prod_{k=1}^K P_i^k \gamma^k \quad (1)$$

Expenditure on variety  $\alpha$  is:

$$x_i^k(\alpha) = \left( \frac{p_i^k(\alpha)}{P_i^k} \right)^{1-\sigma} \gamma^k Y_i$$

where  $Y_i$  is sum of labor income and expenditure on intermediates of all industries.

The key feature of the model is the assumption that the number of units of variety  $\alpha$  which can be produced with one unit of labor,  $z^k(\alpha)$ , is drawn from the Fréchet distribution with location parameter  $z$  and heterogeneity parameter  $\theta$ :

$$F_i^k(z) = \exp(-z/z_i^k)^{-\theta}$$

The fundamental sectoral productivity component is given by the expected sectoral productivity:

$$E [z_i^k(\alpha)] = z_i^k$$

This distributional assumption, as shown by Eaton and Kortum (2002), is tailored to separating out the stochastic cost component from the fundamental sectoral cost component by assuming that the draws  $z_i^k(\alpha)$  are iid across varieties and countries, and therefore the price of each variety is also iid across countries and varieties, just as the realization of least cost varieties across the set of potential suppliers. Using the strong law of large numbers for iid random variables and the continuous mapping theorem, it can be shown that the sectoral price index of effectively exported varieties is (see Costinot et al. (2012)):

$$E [p_j^k(\alpha)^{1-\sigma} | \alpha \in A_{ij}^k] = \Gamma c_{ij}^k^{-\theta} [\Phi_j^k]^{-(\theta+1-\sigma)/\theta}$$

where  $A_{ij}^k \equiv \{c_{ij}^k(\alpha) = \min_{i' \in I} [c_{i'j}^k(\alpha)]\}$  is the set of varieties for which source  $i$  is revealed least cost in destination  $j$ ,  $\Gamma$  is the Gamma function with the argument  $[(\theta + 1 - \sigma)/\theta]$  and  $\Phi_j^k$  is the least cost distribution of varieties across all exporters:  $\Phi_j^k = \sum_{i' \in I} [c_{i'j}^k]^{-\theta}$ .

The sectoral price index in the destination across all exporters is:

$$E [p_j^k(\alpha)^{1-\sigma}] = (P_j^k)^{1-\sigma} = \Gamma [\Phi_j^k]^{-(1-\sigma)/\theta} \quad (2)$$

It follows that bilateral sectoral exports are given by:

$$X_{ij}^k = \frac{[c_{ij}^k]^{-\theta}}{\sum_{i' \in I} [c_{i'j}^k]^{-\theta}} \gamma^k Y_j \quad (3)$$

In the set-up of Costinot et al. (2012), only labor is used in production, and it is assumed that bilateral trade costs are composed of a bilateral symmetric component and of a destination-sector specific component common across exporters  $\tau_{ij}^k = \tau_j^k * \tau_{ij}$ . In this case, the only exporter-sector specific cost component is the fundamental sectoral productivity parameter  $z_i^k$ . Consequently, the pattern of revealed comparative advantage for any pair of exporters on world markets is determined by the ranking of relative sectoral technology stocks:

$$\ln \left[ \frac{X_{ij}^k X_{i'j}^{k'}}{X_{ij}^{k'} X_{i'j}^k} \right] = \theta \ln \left[ \frac{z_i^k z_{i'}^{k'}}{z_i^{k'} z_{i'}^k} \right]$$

Thus, the model without multistage production and consequently no trade in inputs, delivers the result that the ranking of relative sectoral fundamental productivity determines the ranking of relative sectoral exports.

However, it is clear that the way one models sectoral costs has a direct incidence on the components which enter the theoretically grounded measure of revealed comparative advantage. In this paper, we incorporate multistage production and trade in inputs in the Costinot et al. (2012) model in the simplest possible way to quantify the contribution of technology separately from cost components linked to trade in inputs.

Output in each sector is produced using a bundle of inputs and labor. By analogy with the seminal EK set-up, intermediates are introduced by assuming that inputs from all sectors are combined to produce output in any sector, with the production function reproducing exactly the features of the expenditure function so that the cost of the input bundle is given by the overall price index  $P_{i,t} = \prod_{k=1}^K [P_{i,t}^k]^{\gamma^k}$ .

We incorporate a sector-specific feature in the production function by assuming that sectors differ in the way they combine inputs and labor in production. The cost share of inputs  $\zeta^k$  assumed sector-specific entails that differences in the cost of inputs matter relatively more in sectors which use inputs more intensively. Similarly, as the cost share of labor is sector-specific, the wage component also becomes sector-specific.<sup>8</sup>

We model bilateral trade costs as containing a bilateral symmetric component common across sectors  $\tau_{ij}$  and an exporter-sector specific component  $\tau_i^{E,k}$  common across destination markets.<sup>9</sup> This component captures trade frictions the exporter faces to get her products to any destination. We think of the pair-specific component as measuring trade costs independent of trade policy such as transport, coordination, and information costs. We think of the exporter-specific trade friction as determined by tariff and non-tariff barriers.<sup>10</sup>

Consequently, the pattern of revealed comparative advantage for any pair of exporters on world markets is no longer fully determined by the ranking of relative sectoral productivity. To see this, consider relative sectoral exports for a pair of exporters to some destination market:

$$\frac{X_{ij}^k}{X_{i'j}^k} = \left[ \frac{Y_i \nu_i^{1-\zeta^k} P_i^{\zeta^k} \tau_{ij}^k / z_i^k}{Y_{i'} \nu_{i'}^{1-\zeta^k} P_{i'}^{\zeta^k} \tau_{i'j}^k / z_{i'}^k} \right]^{-\theta}$$

In log terms, rescaling by the productivity heterogeneity parameter, and for a specific set of input intensity characteristics  $\zeta^k$ , the ranking of relative sectoral exports is given by a linear combination of four vectors: relative sectoral technology stocks, relative sectoral wages, relative sectoral input costs, and relative trade restrictiveness in exporters' access to world markets.

---

<sup>8</sup>In the empirical analysis, the wage component is sector-specific for an additional reason. It captures the skill composition in the sector (see 4).

<sup>9</sup>Waugh (2010) argues that this specification fits the data better than destination-specific components of trade frictions.

<sup>10</sup>In theory, the most favored nation principle should impede such differences across exporters, but in reality the complex structure of EU-15 trade policy, with multiple Preferential Trade Agreements (PTA) at different stages of implementation, and Generalized System of Preferences (GSP) tariffs granted to certain developing economies results in trade barrier variability across partners.

Indeed, for any two sectors and across destinations, we have:

$$\begin{aligned} \frac{1}{\theta} \left\{ \ln \left[ \frac{X_{ij}^k / X_{i'j}^k}{X_{ij}^s / X_{i'j}^s} \right] \right\} &= \underbrace{\left\{ \ln \frac{z_i^k}{z_{i'}^k} - \ln \frac{z_i^s}{z_{i'}^s} \right\}}_{TFP} + \underbrace{\ln \left[ \left( \frac{P_i}{P_{i'}} \right)^{\zeta^s - \zeta^k} \right]}_{INPUTS} \\ &+ \underbrace{\ln \left[ \left( \frac{\nu_i}{\nu_{i'}} \right)^{\zeta^k - \zeta^s} \right]}_{WAGES} + \underbrace{\ln \frac{\tau_i^{E,s} / \tau_{i'}^{E,s}}{\tau_i^{E,k} / \tau_{i'}^{E,k}}}_{EXPORT-COSTS} \end{aligned}$$

Suppose the input bundle is relatively cheap in country  $i$ . This cost advantage is increasing for country  $i$  in relative input intensity  $\zeta^k - \zeta^s$ . Consequently, this component of comparative advantage will push countries which benefit from relatively cheap inputs to specialize in relatively high input intensive sectors. However, how much this mechanism contributes to determining the pattern of overall sectoral exports is primarily an empirical question. Furthermore, it is no longer immediate that the pattern of intersectoral specialization is driven by relative technology stocks.

In the rest of this section we work with the cost of the input bundle to get a handle on the origin of differences in the cost of inputs. We show that the input-cost driven component of comparative advantage is fully determined by the structure of trade costs of the exporter with all of its potential inputs suppliers. We refer to this index of ‘distance to suppliers’ as the proximity characteristic of the exporter. We first derive the index of trade frictions which determines differences in input costs in a world where the least cost producer of the variety supplies the world market. We refer to this indicator as the proximity endowment of the country. We then show how to derive the proximity characteristic of the country in a world with bilateral trade frictions. In this more realistic case, the cost advantage conferred by the input cost component becomes an endogenous object. We solve this problem in section 5 by showing that the proximity endowment is a valid instrument for the endogenous time-varying proximity characteristic. We then use the instrumented proximity indicator to quantify the contribution in input costs to determining the pattern of comparative advantage.

Before proceeding to the derivation of the proximity characteristic, we underline that we could have adopted an alternative way of introducing inputs in the model by assuming that each sector sources inputs only from itself. The cost of the input bundle would then be sector-specific, and the index of trade frictions would also be sector-specific (see online appendix). We deliberately choose the set-up in which the only channel through which inputs may play a role in determining sector-specific production costs is through the

input intensity characteristic of the sector. Indeed, if the proximity mechanism is shown to co-determine the pattern of comparative advantage in the most restrictive set-up, our results would be providing a lower bound on the role of the input cost channel in determining the pattern of intersectoral specialization.

## 2.2 The input cost component of comparative advantage

### 2.2.1 Proximity endowment

In this subsection we make the simplifying hypothesis that there are no bilateral components of trade frictions. This restrictive hypothesis on the structure of trade costs is made for two reasons. First, it provides the intuition for how the proximity mechanism works in co-determining the pattern of comparative advantage. Second, this world is interesting *per se* because the measure of proximity is directly given by the inverse of a synthetic index of trade frictions. Thus, differences in the cost of inputs stem from a measure of proximity to suppliers which is independent of the distribution of sectoral market shares contingent on a specific trade equilibrium. It is in this sense that we refer to this proximity index as a measure of proximity endowment.

Assume that bilateral sectoral trade costs  $\tau_{ij}^k$  can be approximated by a destination-sector specific cost  $\tau_j^{M,k}$  and an exporter-specific cost component  $\tau_i^{E,k}$ . As previously, we think of trade costs on the export side as a measure of trade restrictiveness linked primarily to trade policy which the exporter faces in supplying her goods to world markets. We think of the destination-specific cost component as a synthetic measure of trade frictions linked to the structure of the trade network rather than to trade policy, expressed as a mark-up which multiplies the factory cost of any variety in reaching this market.

$$\tau_{ij}^k = \tau_j^{M,k} * \tau_i^{E,k}$$

Recall that the sectoral price index (see eqn.2) can be written:

$$P_j^k = \kappa [\Phi_j^k]^{-1/\theta}$$

where  $\kappa = [\Gamma(\frac{\theta+1-\sigma}{\theta})]^{1/(1-\sigma)}$ , and the sectoral price distribution parameter, given the assumption on trade costs, is given by:

$$\Phi_j^k = \sum_{n=1}^N \left[ \omega_n^k \tau_j^{M,k} \tau_n^{E,k} / z_n^k \right]^{-\theta}$$

Define  $\tilde{\Phi}^k$  the realized least cost distribution of varieties in sector  $k$  common across countries:

$$\tilde{\Phi}^k = \sum_{n=1}^N [\omega_n^k \tau_n^{E,k} / z_n^k]^{-\theta}$$

The country-specific distribution of least cost varieties can be written as a rescaled world distribution of least cost varieties:

$$\Phi_j^k = [\tau_j^{M,k}]^{-\theta} \tilde{\Phi}^k$$

The sectoral price index is a product of three components:

$$P_j^k = \kappa \tau_j^{M,k} [\tilde{\Phi}^k]^{-1/\theta} \quad (4)$$

Two of these components are common across countries: the world distribution of least-cost varieties, and the constant  $\kappa$ . The only country-specific component of the sectoral price index is the destination-specific trade cost component  $\tau_j^{M,k}$  which is an indicator of the ease with which country  $j$  gets access to the world distribution of least cost varieties in sector  $k$ .

Recall that the overall price index, which is also the cost of the input bundle, is a Cobb-Douglas price index computed across sectoral price indices. Plugging (4) in (1), the cost of the input bundle can also be written as a product of two components common across countries, and a country-specific index of trade frictions:

$$P_j = \underbrace{\prod_{s=1}^S [\tau_j^{M,s}]^{\gamma^s}}_{\text{SPECIFIC}} \underbrace{\kappa \prod_{s=1}^S [\tilde{\Phi}^s]^{\gamma^s}}_{\text{COMMON}} \quad (5)$$

The composite index of sectoral trade frictions  $\prod_{s=1}^S [\tau_j^{M,s}]^{\gamma^s}$  captures how difficult it is for the country to get access to the best world technology in sourcing inputs. We refer to the reciprocal of this index as the proximity endowment of country  $j$ :

$$PROX_j^M = \left\{ \prod_{s=1}^S [\tau_j^{M,s}]^{\gamma^s} \right\}^{-1} \quad (6)$$

Plugging (5) in the expression of the exporter-sector specific production cost  $\omega_j^k$ , we get:

$$\omega_j^k = \underbrace{\epsilon^k \kappa^{\zeta^k} \left\{ \prod_{s=1}^S [\tilde{\Phi}^s]^{\gamma^s} \right\}^{\zeta^k}}_{\text{COMMON}} \underbrace{\left\{ \prod_{s=1}^S [\tau_j^{M,s}]^{\gamma^s} \right\}^{\zeta^k} \left\{ \nu_j \right\}^{1-\zeta^k}}_{\text{SPECIFIC}} \quad (7)$$



Plugging this expression in the equation of relative sectoral exports for some pair of exporters, we obtain that the only component of the cost of the input bundle which contributes to determining the pattern of comparative advantage is the relative proximity to world technology which is sector-specific in as much as sectors differ in the cost share of inputs in production.

$$\begin{aligned}
\frac{1}{\theta} \left\{ \ln \left[ \frac{X_{ij}^k / X_{i'j}^k}{X_{ij}^s / X_{i'j}^s} \right] \right\} &= \underbrace{\left\{ \ln \frac{z_i^k}{z_{i'}^k} - \ln \frac{z_i^s}{z_{i'}^s} \right\}}_{TFP} + \underbrace{\ln \left[ \left( \frac{PROX_i}{PROX_{i'}} \right)^{\zeta^k - \zeta^s} \right]}_{INPUTS} \\
&+ \underbrace{\ln \left[ \left( \frac{\nu_i}{\nu_{i'}} \right)^{\zeta^k - \zeta^s} \right]}_{WAGES} + \underbrace{\ln \frac{\tau_i^{E,s} / \tau_{i'}^{E,s}}{\tau_i^{E,k} / \tau_{i'}^{E,k}}}_{EXPORT-COSTS} \quad (8)
\end{aligned}$$

The measure of countries' proximity endowment is the correct way to measure proximity to world technology in a world in which we can abstract from the actual distribution of best practice across countries in the world. To see why this is the case, go back to the sectoral market share equation which according to the model is the probability that a given source is least cost in a given sector. Making the hypothesis that the proximity endowment is given by the inverse of the index of destination-specific trade frictions in each sector implies that exporter-specific sectoral market shares are invariant across markets. Indeed, since destination-specific price distribution parameters are  $\Phi_j^k = [\tau_j^M]^{-\theta} \tilde{\Phi}^k$ , the sectoral market share equation simplifies to an expression independent of  $j$ :

$$\pi_{ij}^k = \frac{X_{ij}^k}{X_j^k} = \frac{[\omega_i^k \tau_i^{E,k} / z_i^k]^{-\theta}}{\tilde{\Phi}^k}$$

For this to hold, the distribution of best practice must be common across destinations in every sector:  $\pi_{ij}^k = \pi_{i'j}^k = \pi_i^k, \forall k$ . This corresponds to the assumption that there are no pair-specific trade frictions. Such a world shares with the frictionless world the feature that any supplier revealed least cost in a variety in some market would be revealed least cost in that variety in all markets.

However sectoral market shares are destination-specific in the data. Thus, ignoring differences in the actual distribution of best practice for different destinations must induce some degree of measurement error relatively to the true underlying proximity characteristic of the country. This is why in the next subsection we use the structure of the model to derive a theoretically

grounded measure of proximity which is given by a weighted  $l^\theta$ -norm of bilateral trade frictions in each sector, aggregated across sectors according to the Cobb-Douglas price index. We show that in a world with bilateral trade frictions the main insight remains intact in that the only cost component of inputs which contributes to the pattern of intersectoral specialization is the measure of relative proximity interacted with the input intensity of the sector.

### 2.2.2 Proximity with bilateral trade frictions

In this subsection, we derive the indicator of countries' proximity to suppliers in a world with bilateral trade frictions.<sup>11</sup>

We go back to modelling trade costs as containing a bilateral component common across sectors  $\tau_{ij}$  and an exporter-sector specific component  $\tau_i^{E,k}$  common across destinations. The symmetric component picks up impediments to trade linked to physical features of the trade network such as transport costs.<sup>12</sup> The exporter-sector specific component corresponds to a synthetic indicator of trade restrictiveness the exporter faces on world markets.<sup>13</sup> This modelling of trade costs follows Waugh (2010).<sup>14</sup>

Recall that the sectoral market share equation is a probability measure  $\pi_{ij}^k$  which states the probability that country  $i$  is the least cost producer for country  $j$  across the spectrum of varieties in sector  $k$  (time subscripts are omitted to simplify notation):

$$\pi_{ij}^k = \frac{[\omega_i^k \tau_{ij} \tau_i^{E,k} / z_i^k]^{-\theta}}{\Phi_j^k}$$

Bring the bilateral trade cost component to the left hand side and sum across all suppliers to market  $j$  in sector  $k$  including domestic consumption of domestic varieties:

$$\sum_{n=1}^N \tau_{nj}^\theta \pi_{nj}^k = \frac{\sum_{n=1}^N [\omega_n^k \tau_n^{E,k} / z_n^k]^{-\theta}}{\Phi_j^k}$$

<sup>11</sup>See section 5 for details on how to compute this indicator in the data

<sup>12</sup>In the empirical application, it is assumed well approximated by bilateral distance  $dist_{ij}$ :  $\tau_{ij} = dist_{ij}^\rho$ , with  $\rho = 1$ .

<sup>13</sup>This indicator corresponds to 'one plus the uniform ad valorem equivalent tariff' which, if applied to all exports of this source towards world markets, would leave the level of aggregate exports unchanged. This is the definition of sectoral MA-OTRI indicators in Kee et al. (2009).

<sup>14</sup>It differs in the empirical application in that we assume a loglinear relationship between distance and bilateral trade costs. Further, we do not consider sharing a border and sharing a language in computing the bilateral component of trade costs.

Define  $\bar{\Phi}^k = \sum_{n=1}^N [\omega_n^k \tau_n^{E,k} / z_n^k]^{-\theta}$ . This sectoral price distribution parameter is common across countries. It summarizes the price distribution of best practice across varieties within sector  $k$ , inclusive of barriers linked to trade policy.

Rewrite the destination-specific parameter using the expression of world's best practice within the sector:

$$\Phi_j^k = \bar{\Phi}^k \left\{ \sum_{n=1}^N \tau_{nj}^\theta \pi_{nj}^k \right\}^{-1}$$

Using the fact that sectoral price indices are given by:

$$P_j^k = \kappa [\Phi_j^k]^{-1/\theta} \quad (9)$$

where  $\kappa = [\Gamma(\frac{\theta+1-\sigma}{\theta})]^{1/(1-\sigma)}$ , the sectoral price index is:

$$P_j^k = \kappa [\bar{\Phi}^k]^{-1/\theta} \left\{ \sum_{n=1}^N \tau_{nj}^\theta \pi_{nj}^k \right\}^{1/\theta} \quad (10)$$

The overall price index is  $P_j = \prod_{k=1}^K [P_j^k]^{\gamma^k}$ . Using (10), the cost of the input bundle in country  $j$  is:

$$P_j = \kappa \left\{ \prod_{k=1}^K [\bar{\Phi}^k]^{-\gamma^k/\theta} \right\} \left\{ \prod_{k=1}^K \left[ \sum_{n=1}^N \tau_{nj}^\theta \pi_{nj}^k \right]^{\gamma^k/\theta} \right\} \quad (11)$$

As previously, the overall price index specific to the destination is a product of world's best practice across sectors and of a proximity measure which states how far the country is from its suppliers.

Relative proximity is also a summary statistic of the relative cost of living for any two countries. The intuition is straightforward: the closer the country is to the best world technology, and the lower is its cost of living relatively to other countries. Consequently, relative real wages can be computed by adjusting the ratio of nominal wages by relative proximity while circumventing the problem of constructing actual price indices.

The proximity characteristic of the exporter in a world with bilateral trade frictions is given by the inverse of the weighted index of bilateral trade costs where sector-specific weights correspond to bilateral market shares within the sector:

$$\left[ \overline{PROX}_j^M \right]^{-1} = \prod_{k=1}^K \left\{ \sum_{n=1}^N \pi_{nj}^k \tau_{nj}^\theta \right\}^{\gamma^k/\theta} \quad (12)$$

This microfounded proximity indicator is a weighted  $l^\theta$ -norm of the vector of bilateral trade frictions in each sector, aggregated across sectors according to the Cobb-Douglas price index with exponents given by sectoral expenditure shares. Thus, (8) remains valid in the world with bilateral trade frictions in the sense that the only cost component of the input bundle which plays a role in co-determining intersectoral specialization is captured by the proximity characteristic. Four exporter-sector specific cost components determine the pattern of comparative advantage: technology stocks, export side trade costs, sectoral wages through their interaction with the input intensity of the sector, and relative proximity, also through its interaction with the input intensity of the sector.

The very important difference is that proximity is now an endogenous object. It differs from proximity endowment in that it weighs the components of bilateral trade costs by the effective weight of each supplier in the market. Even if bilateral components of trade frictions may be considered exogenous in that they are determined by slow-moving characteristics of the trade network (infrastructure, costs in the transport sector, coordination costs,...), market shares are contingent on a specific trade equilibrium. To solve this problem, we instrument the proximity characteristic with the indicator of proximity endowment in the empirical analysis (see sec.5).

### 3 Estimation procedure

In the empirical analysis we investigate whether the cost advantage conferred by the ability to source inputs at relatively lower cost leads to specialization of high proximity countries in sectors which use inputs relatively more intensively. The objective of the estimation procedure is to disentangle the contribution of proximity to world technology from the two other fundamental country characteristics which are technology and factor endowments.

The estimation procedure is based on the the gravity structure of trade in the formulation of the EK model at the sectoral level. The difference relatively to the sectoral imports' equation derived in Costinot et al. (2012) is that we allow for sector-specific components of the production cost  $\omega_i^k$  and for exporter-sector specific components of the trade cost  $\tau_{ij}^k$ :

$$X_{ij}^k = \frac{X_j^k [\omega_i^k \tau_{ij}^k / z_i^k]^{-\theta}}{\Phi_j^k} \quad (13)$$

where  $X_j^k = \gamma^k Y_j$  is expenditure on goods in sector  $k$  in country  $j$ .

In Costinot et al. (2012), under the assumption that sectoral trade costs  $\tau_{ij}^k$  are given by the product of a symmetric bilateral component  $\tau_{ij}$  which is

not sector-specific and a destination-specific sectoral component  $\tau_j^k$  common across trade partners, a fixed effects approach in the cross-sectional regression of bilateral sectoral exports on pair, destination- and source-sector dummies as in (14) allows retrieving directly the fundamental sectoral productivities  $z_i^k$  relatively to a benchmark country and industry.

$$X_{ij}^k = \exp \{ fe_{ij} + fe_j^k + fe_i^k + \omega_{ij}^k \} \quad (14)$$

where  $fe_{ij}$ ,  $fe_j^k$ , and  $fe_i^k$  are respectively pair, destination-sector, and exporter-sector fixed effects. The source-sector dummy  $e^{fe_i^k}$  raised to the exponent  $1/\theta$  corresponds to  $\left[ \frac{z_i^k}{z_i^s} / \frac{z_{i'}^k}{z_{i'}^s} \right]$  where  $z_i^s$ ,  $z_{i'}^k$ , and  $z_{i'}^s$  are normalized to 1.

If sectoral trade costs contain an exporter-sector specific component  $\tau_i^{k,E}$ , this approach allows computing relative fundamental sectoral productivity scaled by export-side trade costs  $z_i^k / \tau_i^{k,E}$ . Furthermore, if the production process combines labor and inputs differently across sectors, exporter-sector dummies capture sectoral components of wages and input costs contained in  $\omega_i^k$ .

Instead of fundamental productivities, we retrieve the reciprocal of the exporter-sector specific unit cost:  $z_i^k / \omega_i^k \tau_i^{k,E}$ . To recover fundamental sectoral technology, exporter-sector dummies need to be cleaned of exporter-sector specific components of producer costs which are wages, inputs', and trade costs.

The estimation procedure consists of three steps. First, we work in cross section, with  $t \in T = \{1995, \dots, 2009\}$ . Bilateral sectoral exports are regressed on pair, destination-sector, and source-sector fixed effects to separate out all exporter-sector specific determinants, i.e. to retrieve full fundamental sectoral production costs relatively to a benchmark country (the US) and industry (processed foods and beverages).

$$X_{ij,t}^k = \exp \{ fe_{ij,t} + fe_{j,t}^k + fe_{i,t}^k + \xi_{ij,t}^k \} \quad (15)$$

This step is identical to the estimation conducted in Costinot et al. (2012) to retrieve relative fundamental sectoral productivities under the assumption of no sector-specific characteristics other than technology stocks. The specific way in which we relax this assumption entails that the exporter-sector dummy is no longer  $fe_{i,t}^k = \theta \ln \left[ \frac{z_{i,t}^k}{z_{i,t}^s} / \frac{z_{i',t}^k}{z_{i',t}^s} \right]$  where  $z_{i,t}^s$ ,  $z_{i',t}^k$ , and  $z_{i',t}^s$  have been normalized to 1. Rather, the dummy corresponds to a comprehensive producer-cost component specific to the sector given by a combination of technology, wage, input, and trade-specific cost components, relatively to the benchmark country and industry for which this full sectoral production

cost had been normalized to 1.<sup>15</sup>

$$\widehat{fe}_{i,t}^k = \theta \ln(z_{i,t}^k) - \theta(1 - \zeta^k) \ln \nu_{i,t}^k - \theta \zeta^k \ln(P_{i,t}) - \theta \ln(\tau_{i,t}^{E,k}) \quad (16)$$

In the second step, we pool all data on estimated exporter-sector dummies  $\widehat{fe}_{i,t}^k$  in each year and sector, estimated for each exporter across EU-15 markets in cross-section, and we regress these dummies on instrumented sectoral wages ( $\widehat{\nu}_{i,t}^k$ ) and instrumented sectoral technology stocks ( $\widehat{z}_{i,t}^k$ ), controlling for the benchmark country component with year fixed effects  $fe_t$ .<sup>16</sup>

$$\widehat{fe}_{i,t}^k = \theta [\ln \widehat{z}_{i,t}^k - (1 - \zeta^k) \ln \widehat{\nu}_{i,t}^k] + fe_t + \lambda_{it}^k \quad (17)$$

This allows retrieving the structural parameters  $\theta$  and  $\zeta^k$  in a way consistent with the underlying model: the heterogeneity parameter of the productivity distribution is assumed constant across sectors and overtime, while input intensity characteristics are assumed sector-specific, common across countries, and overtime. We check that these parameters are precisely estimated, stable across variants of the instrumenting procedure, and consistent with previous studies. Further, we check that estimated  $\zeta^k$  parameters are strongly positively correlated with observed input intensity in our dataset.

This second step is needed to recover the input-cum-trade-policy cost component separately from determinants of specialization due to domestic technology stocks and factor endowments. Indeed, the residual cost component summarizes all channels through which trade costs may play a role in determining countries' specialization. To see why this is the case, focus on the residual of the second-step equation:  $\widehat{\lambda}_{i,t}^k = -\theta [\zeta^k \ln(P_{i,t}) + \ln(\tau_{i,t}^{E,k})]$ .

On the exports side, this residual contains the trade policy determined cost component  $\tau_{i,t}^{E,k}$  which captures how costly it is for the country to ship its products to world markets relatively to other exporters. On the imports side, the residual contains a vector of  $N$  pair-specific bilateral trade cost components  $\tau_{ni}$  which enter in the expression of the sectoral price index and which capture in a complex way how costly it is for country  $i$  to get access to inputs produced in all sources  $n$  including itself.

To see why this is the case, write the production cost  $\omega_j^k$  replacing the

---

<sup>15</sup>In the online appendix we report descriptive statistics on estimated relative production costs for countries of our sample, both in cross-section and overtime.

<sup>16</sup>In the data, labor is split in three skills. Skill-endowment of each sector is assumed fixed. Sector-skill specific wages are determined endogenously. Thus, wages have an intrinsic sectoral component, on top of the interaction with input intensity  $\zeta^k$ . See 4 for details.

cost of the input bundle by its value in (11):

$$\omega_j^k = \underbrace{\epsilon^k \kappa^{\zeta^k} \left\{ \prod_{s=1}^S [\overline{\Phi}^s]^{-\gamma^s/\theta} \right\}^{\zeta^k}}_{\text{sector-specific}} \underbrace{[\nu_j^k]^{1-\zeta^k} \left\{ \prod_{s=1}^S \left[ \sum_{n=1}^N \tau_{nj}^\theta \pi_{nj}^s \right]^{\gamma^s/\theta} \right\}^{\zeta^k}}_{\text{exporter-sector-specific}} \quad (18)$$

The exporter-sector specific cost contains two components: the wage, reconducible to the factor endowment and to the overall level of technological development of the exporter, and the proximity indicator which is a summary statistic of exporter's ease of access to world technology.<sup>17</sup>

Relative bilateral sectoral exports to market  $j$  for exporters  $i$  and  $i'$  are:

$$\ln \{ X_{ij,t}^k / X_{i'j,t}^k \} = \theta \left[ \ln \frac{z_{i,t}^k}{z_{i',t}^k} - (1 - \zeta^k) \ln \frac{\nu_{i,t}^k}{\nu_{i',t}^k} - \ln \frac{\tau_{ij} \tau_{i,t}^{E,k}}{\tau_{i'j} \tau_{i',t}^{E,k}} + \zeta^k \ln \left\{ \frac{\overline{PROX}_{i,t}^M}{\overline{PROX}_{i',t}^M} \right\} \right]$$

The residual  $\widehat{\lambda}_{i,t}^k$  obtained in the second step of the estimation procedure is a combination of the two trade cost characteristics of the country: the barriers it overcomes in getting access to world best practice on the supply side and the trade cost it must pay to get domestically produced varieties to world markets.

$$e^{\widehat{\lambda}_{i,t}^k} = \left\{ \prod_{s=1}^S \left[ \sum_{n=1}^N \tau_{nj}^\theta \pi_{nj}^s \right]^{\gamma^s/\theta} \right\}^{-\theta \zeta^k} \left[ \tau_{i,t}^{E,k} \right]^{-\theta} \quad (19)$$

Recall that exporter-sector dummies  $fe_{i,t}^k$  obtained in the first step of the estimation procedure (15) capture the sectoral cost component of the exporter relatively to a benchmark sector and country for which cost components are normalized to 1. It follows that relative exporter-sector dummies for any pair of exporters capture relative sectoral cost components for this pair of exporters, up to an exporter-year specific component which corresponds to the cost component of the exporter in the benchmark sector:<sup>18</sup>

$$fe_{i,t}^k - fe_{i',t}^k = \theta \left[ \ln \frac{z_{i,t}^k}{z_{i',t}^k} - (1 - \zeta^k) \ln \frac{\nu_{i,t}^k}{\nu_{i',t}^k} - \ln \frac{\tau_{i,t}^{E,k}}{\tau_{i',t}^{E,k}} + \zeta^k \ln \frac{\overline{PROX}_{i,t}^M}{\overline{PROX}_{i',t}^M} \right] + fe_{i,t} - fe_{i',t} + \xi_{ii',t}^k \quad (20)$$

<sup>17</sup>The destination-sector fixed effect of the first step now absorbs not only the sector-specific constant  $\epsilon_k$ , but also the two sector-specific components of the input cost common across exporters.

<sup>18</sup>The bilateral symmetric component of trade costs is absorbed by the pair-specific fixed effect in the first step of the estimation.

where  $fe_{n,t}$  for  $n = i, i'$  are exporter-year fixed effects.

In the empirical analysis, we use data on sectoral exports to each of EU-15 markets. According to our modelling of trade costs, the export side component  $\tau_{i,t}^{E,k}$  contains tariff and non tariff barriers which are exporter-sector specific and common across destination markets. This hypothesis adequately describes the underlying trade cost structure across EU-15 markets because the European Union is characterized by a unique external trade policy and a multiplicity of exporter-specific trade agreements, in particular with emerging economies. Thus, bilateral components capture plausibly symmetric barriers to trade such as transport costs while exporter-specific components capture source-specific trade restrictiveness linked to trade policy.

Assume that exporter-specific sectoral trade costs  $\tau_{i,t}^{E,k}$  are well approximated for each exporter by the component  $\tau_{i,t}^E$  which is time-varying and common across sectors. In relative terms, this common component is a summary statistic of the relative trade restrictiveness faced by a pair of exporters on EU-15 markets. To give an example, this component could be common to all countries of Central and Eastern Europe, and in relative terms it would indicate the degree of preferential access of the CEECs to EU-15 markets relatively to a specific non-European emerging economy, such as China or Mexico. If this assumption holds, then pair fixed effects capture this common component of trade barriers in the first step of the estimation.

Suppose this relative common component is multiplied by a stochastic component  $l_{i,t}^k/l_{i',t}^k$ , distributed lognormal with mean 1. Assuming that the stochastic component is statistically independent of the regressors in (20) allows rewriting relative pairwise RCA rankings as a function of three complementary components: relative technology stocks, relative sectoral wages, and relative proximity. These three components fully account for the microfounded measure of revealed comparative advantage for a given pair of exporters on world markets, up to a relative exporter-year fixed effect  $fe_{i,t} - fe_{i',t}$  and a stochastic component captured by the residual  $\xi_{ii',t}^k$  which comprises the stochastic component  $\ln(l_{i,t}^k/l_{i',t}^k)$ .

$$fe_{i,t}^k - fe_{i',t}^k = \theta \left[ \ln \frac{z_{i,t}^k}{z_{i',t}^k} - (1 - \zeta^k) \ln \frac{\nu_{i,t}^k}{\nu_{i',t}^k} + \zeta^k \ln \frac{\overline{PROX}_{i,t}^M}{\overline{PROX}_{i',t}^M} \right] + fe_{i,t} - fe_{i',t} + \xi_{ii',t}^k \quad (21)$$

The residual component of RCA rankings illustrates that conditional on the distribution of technology and wages, intersectoral specialization within a pair is determined by the relative proximity characteristic interacted with



the input intensity of the sector. The proximity index is a summary statistic of the input component of the cost advantage conferred to the country by the ease of its access to best technology worldwide in sourcing inputs. It becomes a source of comparative advantage at the intersectoral level because the input component of production costs matters relatively more in sectors which use inputs intensively.

$$\frac{e^{\lambda_{i,t}^k}}{e^{\lambda_{i',t}^k}} = \left[ \frac{\widehat{PROX}_{i,t}^M}{\widehat{PROX}_{i',t}^M} \right]^{\theta \zeta^k} \left[ \frac{l_{i',t}^k}{l_{i,t}^k} \right]^\theta \quad (22)$$

The residual component contains all information on the channel through which relatively high proximity countries specialize in sectors which use inputs relatively more intensively, as measured by  $\zeta_k$ . Given our assumption on export side trade costs, a simple test of the role played by the proximity mechanism in co-determining the pattern of specialization is a straightforward implementation of (22).

Consequently, in the third step of the estimation procedure, we use the proximity ranking of countries to test for the role of relative input costs in determining the pattern of comparative advantage. We split the sample in 2 groups according to the proximity characteristic (see sec.5): the EU15 and the CEECs constitute the ‘high’-proximity group, while non-European developed and developing countries constitute the ‘low’-proximity group.

We rescale estimated residuals  $\widehat{\lambda}_{i,t}^k$  of (17) by the estimated heterogeneity parameter  $\widehat{\theta}$ , and compute all pairwise combinations of sectoral annual residuals  $(1/\widehat{\theta})(\widehat{\lambda}_{i,t}^k - \widehat{\lambda}_{i',t}^k)$  where  $i \in H$  are countries of the high proximity group, and  $i' \in L$  are countries of the low proximity group.

The indicator of sectoral relative proximity is computed as the log of the relative proximity characteristic for each pair, instrumented with relative proximity endowment. Instrumented relative proximity is then interacted with the input intensity characteristic of the sector  $\widehat{\zeta}^k$  estimated in the second step:  $\widehat{\zeta}^k \ln(\widehat{PROX}_{i,t}^M / \widehat{PROX}_{i',t}^M)$ . We estimate (23) on data pooled for all years. We include exporter-year fixed effects  $\{fe_{i,t} - fe_{i',t}\}$  to control for characteristics of the benchmark sector for each exporter and year.

$$\frac{1}{\widehat{\theta}} \left[ \widehat{\lambda}_{i,t}^k - \widehat{\lambda}_{i',t}^k \right] = \beta_0 + \beta_1 \ln \left\{ \left( \frac{\widehat{PROX}_{i,t}^M}{\widehat{PROX}_{i',t}^M} \right)^{\widehat{\zeta}^k} \right\} + fe_{i,t} - fe_{i',t} + \eta_{ii',t}^k \quad (23)$$

The coefficient of interest is  $\beta_1$ : according to the model,  $\beta_1$  should be positive and close to 1.

## 4 Data and Estimation of Model Parameters

### 4.1 The Data

#### 4.1.1 Exporter-sector dummies

To obtain the ranking of relative sectoral exports on EU-15 markets (step 1 of the estimation), we use the COMEXT database. COMEXT provides exhaustive information on bilateral trade flows for each country of the EU-15 with each other country in the world at the 8-digit level (CN classification). We use data on total imports to identify the set of EU-15 main trading partners in 1995-2010, defined as the set of countries which make up at least 1% of total EU-15 imports in more than one year in the period under study.<sup>19</sup>

As the model is silent about countries' endowments of primary goods, we restrict attention to categories classified as manufacturing. We use the CN8-BEC correspondence to drop inputs produced from raw gas, petroleum, coal, and nuclear fuel. We construct a correspondence from the CN8 to the 4- and 2-digit NACE 1.1 and ISIC Rev.3 classifications where manufacturing corresponds to sectors 15 – 36 at the 2-digit level.<sup>20</sup> In the estimation, we exclude energy products (sector 23) to be consistent with dropping energy inputs, and tobacco products (sector 16) for which data is patchy. This leaves 20 sectors at the 2-digit level (see tab.??).

In the online appendix we provide descriptive statistics on exporter-sector dummies estimated at the 2-digit level in 1995-2010. It underlines the persistence in country-specific relative sectoral rankings. It discusses changes in the pattern of revealed comparative advantage at the bilateral level and by partner type. In Costinot et al. (2012), these rankings would correspond to the ranking of fundamental sectoral productivities while in this paper the ranking results from technology, factor, input, and export-side trade cost components specific to the sector and exporter.

---

<sup>19</sup> See tab.?. In practice, we include all members of the European Union, excluding Cyprus and Malta but including Croatia.

<sup>20</sup>There are 121 active 4-digit codes in ISIC Rev.3, and a bit more in NACE 1.1. There are minor discrepancies between NACE 1.1 and ISIC Rev.3 at the 4-digit level, mainly because NACE 1.1 is a more detailed classification. There are no discrepancies at 2-digit in the sense that CN8 products are classified within the same 2-digit category in both classifications.

### 4.1.2 TFP and wages

To estimate the parameters of the model (step 2 of the estimation), we need information on technology stocks and labor costs. We construct these components using the World Input Output Database (see Timmer (2012)) which provides harmonized information on gross output, workforce, hourly wages, expenditure on inputs and labor, nominal investment and real capital stocks by sector for all but six countries of our sample.<sup>21</sup>

Sectoral total factor productivity is constructed by fitting a Cobb-Douglas production function while allowing factor shares to vary by country and sector. In logs, TFP is given by the residual of real gross sectoral output  $Y_i^k$  from which we subtract the contribution of three production factors which are inputs  $I$ , labor  $H$ , and capital  $K$ , weighted by their respective income shares  $\beta_{f,i}^k$ , with  $f = \{I, H, K\}$ :

$$\ln(\bar{z}_i^k) = \ln Y_i^k - \beta_{I,i}^k \ln I_i^k - \beta_{H,i}^k \ln H_i^k - \beta_{K,i}^k \ln K_i^k \quad (24)$$

Real gross sectoral output and real expenditure on inputs are obtained by deflating the corresponding nominal values by output and input deflators provided in the WIOD. Labor use is taken directly from WIOD as the total number of hours worked in the sector. Obtaining real capital expenditure is more tricky. WIOD provides information on nominal capital expenditure, nominal investment, and a deflator for nominal investment. We approximate the use of capital in the production process by predicting nominal capital expenditure by deflated sectoral investment.<sup>22</sup> Income shares of production factors are computed as ratios of nominal expenditure on inputs, labor costs, and capital to the nominal value of sectoral output.

We use two measures of sectoral wages. The first is the hourly wage in the sector  $\nu_i^k$  obtained by dividing total labor compensation by total hours worked. The second measure is obtained by adjusting observed labor costs for human capital accumulated through education.

The adjustment consists in rescaling hourly skill-specific wages by a proxy of worker efficiency as in Eaton and Kortum (2002). Skill-specific wages  $\nu_{edu,i}^k$  are given by the average hourly wage rescaled by the ratio of the cost share of the skill in total costs  $\omega_{edu,i}^k$  to the time share of the skill in total hours worked  $\bar{\omega}_{edu,i}^k$ :

$$\nu_{edu,i}^k = \frac{\omega_{edu,i}^k}{\bar{\omega}_{edu,i}^k} \nu_i^k$$

<sup>21</sup>Croatia, Norway, Switzerland, Malaysia, Singapore, and Thailand are the six countries absent from the WIOD. We have not been able to find an alternative data source for these countries.

<sup>22</sup>Real GFCF explains 67% of variation in nominal capital expenditure.

The efficiency adjustment is implemented by multiplying skill-specific wages by an exponential function which argument is the average number of years of schooling for the skill  $S_{edu}$  multiplied by the return to education  $g = .06$ .<sup>23</sup>

$$\bar{\nu}_{edu,i}^k = \nu_{edu,i}^k e^{-gS_{edu}}$$

In WIOD, hourly wages are reported for low ( $l$ ), medium ( $m$ ), and high ( $h$ ) skilled workers. We use the International Standard Classification of Education correspondence of skills to educational attainment to define  $S = \{8, 13, 18\}$  for  $edu = \{l, m, h\}$ , respectively.<sup>24</sup>

The sectoral efficiency-adjusted hourly wage  $\bar{\nu}_i^k$  corresponds to a weighted average of skill-specific efficiency-adjusted wages  $\bar{\nu}_{edu,i}^k$  as in (25).<sup>25</sup>

$$\bar{\nu}_i^k = \sum_{edu} \frac{\omega_{edu,i}^k}{\bar{\omega}_{edu,i}^k} \bar{\nu}_{edu,i}^k \quad (25)$$

### 4.1.3 Level of aggregation

WIOD reports data for 13 manufacturing sectors instead of the 20 sectors obtained at the 2-digit level (compare tab.?? and tab.??). Since we only have data on measured TFP and hourly wages for 13 sectors, we reestimate exporter-sector dummies at this level of aggregation, and work with 13 manufacturing sectors in the second and third steps of the estimation.

In four cases, this higher level of aggregation corresponds to pooling data on production of processed inputs and final output for a specific industry. This is the case in the textile (sectors 17 and 18), paper (sectors 21 and 22), metal products (sectors 27 and 28), and transport industries (sectors 34 and 35). For these four industries the higher level of aggregation may actually improve consistency with the production structure considered in the model.

For two industries, this higher level of aggregation introduces a discrepancy between trade and production data. Thus, the WIOD pools together food manufacturing with the tobacco industry while the latter is dropped from trade data because of mediocre data quality. The second discrepancy is due to aggregation of miscellaneous manufacturing (36) with the recycling

---

<sup>23</sup>We follow Eaton and Kortum (2002) in using 6% as the return to education. This estimate is reported as conservative in Bilal and Klenow (2000).

<sup>24</sup>UNESCO, ISCED 1997, reedited 2006.

<sup>25</sup>In practice, we adjust the number of hours worked by skill, and compute efficiency-adjusted wages as the ratio of total cost to the efficiency-adjusted number of hours worked.

industry (37), the latter being absent from trade data. We assume this discrepancy to be relatively minor because the common component is likely to be representative of gross sectoral output.

The problematic aspect of data aggregation in WIOD is the pooling of data on computer manufacturing, electrical and audiovisual equipment, and medical-optical precision equipment (sectors 30 – 33) into a single industry. Tab.?? shows that these four sectors vary significantly in the share of value added (VA) in gross output.<sup>26</sup> The precision equipment industry has the highest VA share in manufacturing (.41) while the computer and office machinery is relatively input intensive with the share of VA at just over .25 of gross output.<sup>27</sup>

The level of aggregation may impact the ranking of sectors in terms of input intensity. It may also play against our assumption that inputs' share in gross output is a sector-specific characteristic common across countries. Indeed, even if the underlying production functions have common factor shares at a relatively fine level of disaggregation, the sectoral mix of subsectors' input intensity is likely to be country specific. In particular, measured sectoral input intensity at the WIOD level becomes an endogenous object if high(low) proximity countries tend to specialize in high(low) proximity subsectors within each industry. If this is the case, the pattern of intersectoral specialization is relatively less determined by the proximity mechanism. Consequently, working at a higher aggregation level is likely to make it more difficult to pick up the working of the proximity mechanism at the intersectoral level.

In the online appendix, we use production and value added at the 4-digit level in ISIC Rev.3 reported in UNIDO INDSTAT4 to gauge the sensitivity of measured input intensities to the level of data aggregation. For the main economies of the EU-15, the assumption of common factor shares is best borne out in the data at the 4-digit level. But if we focus on the ordinal ranking of sectors as a function of inputs' weight in the production function, this ranking is found to be relatively stable across countries even at higher aggregation levels. Similarly, in the WIOD, the ranking of measured sectoral input intensities is strongly correlated across countries. Consequently, the assumption of common sectoral input intensities used in the estimation of

---

<sup>26</sup>The table reports input intensity in manufacturing at the 2-digit level for the main economies of the EU-15. The indicator is computed as  $1 - VA/PROD$  using [UNIDO INDSTAT](#).

<sup>27</sup>This example illustrates that the intensity of inputs' use in production is not equivalent to the ranking of sectors according to technological complexity. Similarly, tab.?? shows that there is no one-to-one mapping from the share of inputs in production to the share of inputs in total imports which we refer to as sectoral input intensity in trade.

the model is consistent with the data in as much as it captures an ordinal ranking of sectors as a function of input intensity.

## 4.2 Estimation of the model

### 4.2.1 Motivation of an instrumental variables estimator

The second step of the estimation is the crucial point of the analysis in which we obtain structural parameters of the model and construct the residual component of exporter-sector specific production costs. This residual is used to evaluate the role of the proximity mechanism in co-determining the pattern of comparative advantage in the final step of the estimation.

The first reason for implementing an instrumental variables estimator is to ensure that this residual vector is orthogonal to variation in measured TFP and hourly wages which is attributable to domestic technology and labor endowments. This is because the two fundamental country-sector characteristics other than proximity which determine the pattern of specialization according to the model are technology stocks and factor endowments.

The second reason is the need to obtain consistent estimates of model parameters which may be hindered by errors-in-variables in measured TFP and hourly wages. Furthermore, joint determination of sectoral exports with non-instrumented TFP and wages cannot be excluded. Isolating the variation in measured TFP and hourly wages determined by fundamental country characteristics helps stymie both sources of bias.

The estimation is conducted in two stages. In the first stage sectoral TFP and wages are regressed on a common set of instruments to identify the variation in measured TFP and hourly wages explained by domestic technology and labor endowments. Both characteristics are sufficiently slow-moving to be considered exogenous to a given trade equilibrium.

In the second stage we project estimated exporter-sector dummies on the space defined by the vectors obtained in the first stage which are instrumented sectoral wages ( $\widehat{\mathcal{V}}_{i,t}^k$ ) and instrumented sectoral TFP ( $\widehat{\mathcal{Z}}_{i,t}^k$ ) while allowing the coefficient on wages to be sector-specific. This is done to identify the component of RCA rankings which is orthogonal to variation in TFP and sectoral wages picked up by technology stocks and labor endowments:

$$\widehat{fe}_{i,t}^k = \theta \ln \widehat{\mathcal{Z}}_{i,t}^k - \theta(1 - \zeta^k) \ln \widehat{\mathcal{V}}_{i,t}^k + fe_t + \lambda_{it}^k \quad (26)$$

Consequently, the residual of the reduced form model specified in (26),  $\lambda_{it}^k$ , is the variable which should be used in the third step of the estimation to identify the contribution of the proximity mechanism to determining the com-

ponent of RCA rankings unexplained by the two other fundamental country-sector characteristics.

The productivity dispersion parameter  $\theta$  is directly given by the coefficient on instrumented TFP, while sector-specific input intensities  $\zeta^k$  are computed from the coefficient on instrumented sectoral wages using estimated  $\theta$ .<sup>28</sup>

#### 4.2.2 Which instruments?

Sectoral workforce  $L_i^k$  constitutes a logical instrument for hourly wages  $\nu_i^k$  because sectoral wages are decreasing in labor endowment.<sup>29</sup> The information on the number of persons engaged in the sector is directly provided in WIOD.

Efficiency adjusted wages  $\bar{\nu}_i^k$  are instrumented with efficiency-adjusted sectoral workforce  $\bar{L}_i^k$ . We compute the number of persons engaged by skill  $L_{edu,i}^k$ , and adjust skill-specific labor by the human capital of the worker:

$$\bar{L}_{edu,i}^k = L_{edu,i}^k e^{gS_{edu}}$$

The adjusted labor force is the sum of efficiency-adjusted labor by skill:

$$\bar{L}_i^k = \sum \bar{L}_{edu,i}^k$$

Sectoral technology stocks are modelled as a function of capital accumulation and R&D activity.<sup>30</sup> Accordingly, we use two sets of instruments for measured TFP. In the first specification, sectoral TFP is instrumented with real sectoral capital stocks and R&D personnel. Data on real capital stocks is provided by the WIOD in 1995-2007. Data on the full time equivalent number of persons employed to conduct R&D activity is reported in ANBERD (see below). The caveat is the restriction of the estimation window to 1995-2007. The advantage is the ability to implement standard tests on instrument validity given that the equation is overidentified.

In the second specification we use nominal R&D expenditure as the indicator of R&D activity. We consider that R&D expenses are mostly incurred to finance investment and employment of R&D personnel. Consequently, we first deflate sectoral expenditure on R&D by regressing it on real investment

---

<sup>28</sup>The empirical counterparts of structural input intensities are the country-year specific income shares of inputs in gross output  $\beta_{i,t}^{k,I}$ .

<sup>29</sup>Labor endowments by skill in each sector are considered predetermined by making the hypothesis that sector-specific human capital impedes labor movement across sectors in the short term. The sector-specific mix of skills is taken as given. See the online appendix for details.

<sup>30</sup>See Eaton and Kortum (1999, 2002).



and R&D personnel.<sup>31</sup> Measured TFP is instrumented with predicted R&D expenditure. In this specification the estimation window is extended to 2009 because real investment data is reported in WIOD in 1995-2009.

The bottleneck is the availability of data on R&D activity (see online appendix . Time series data on R&D personnel and nominal R&D expenditure for all developed and a subset of emerging economies are taken from the 2011 edition of OECD ANBERD.<sup>32</sup> For China, we compiled sectoral data on R&D personnel and nominal R&D expenditure in 1995-2009 using the Yearbook Database of China Data Online.<sup>33</sup> Bulgaria, Brazil, India, Indonesia, Lithuania, Latvia, and Russia were dropped because of lacking data on R&D expenditure and personnel.<sup>34</sup> This leaves 26 countries in the second step of the estimation.

### 4.2.3 Estimated parameters

To estimate this model we need instrumented sectoral hourly wages and instrumented TFP. Consequently, in the first stage we run 13 regressions in which measured TFP and hourly sectoral wages are regressed on a common set of instruments which include R&D personnel and real capital stocks in (I) and (II) (deflated R&D expenditure in (III) and (IV)) together with the workforce of each of the 12 sectors. In (I) and (III) sectoral workforce is efficiency-adjusted. In (II) and (IV) we use raw data on hourly wages and number of persons engaged in the sector.

Tab.1 reports results of the second stage (see App.?? for results of the first stage). Reported values of Kleibergen-Paap rk LM and Cragg Donald Wald F statistics attest that instruments pass respectively the underidentification and weak identification tests across specifications.<sup>35</sup> As the equation is overidentified in the first two specifications, we report the result of the test of overidentifying restrictions (Hansen J statistic). The joint null that instruments are uncorrelated with the error term and correctly excluded from the estimation is not rejected at conventional significance levels.

---

<sup>31</sup>The estimated coefficient on R&D personnel is .92(.009), and .23(.01) on real investment. The two variables explain 87% of observed variation in nominal R&D expenditure.

<sup>32</sup>Downloaded in July 2012 from [OECD ANBERD](#).

<sup>33</sup>See [Yearbook Database](#). The data is reported in html and pdf formats in China Statistical Yearbook on Science and Technology (1996-2008), China Statistics Yearbook on High Technology Industry (2002, 2003, 2007), and in the chapter ‘Education, Science, and Technology’ of China Statistical Yearbook (2007-2011).

<sup>34</sup>Only data on nominal R&D expenditure is available for Russia in ANBERD.

<sup>35</sup>The underidentification test rejects the null that the matrix of reduced form coefficients is not full rank.



Table 1: Second stage: Estimated parameters

	(I)	(II)	(III)	(IV)
<i>TFP</i>	7.258*** (0.506)	6.718*** (0.431)	7.842*** (0.524)	7.280*** (0.448)
<i>WAGE</i>	-1.343*** (0.212)	-1.388*** (0.145)	-1.610*** (0.211)	-1.583*** (0.149)
WAGE 19	1.090*** (0.292)	0.558*** (0.138)	1.226*** (0.274)	0.640*** (0.131)
WAGE 20	-1.265*** (0.178)	-0.793*** (0.101)	-1.136*** (0.163)	-0.727*** (0.095)
WAGE 21 – 22	-1.471*** (0.156)	-0.959*** (0.091)	-1.365*** (0.143)	-0.910*** (0.085)
WAGE 24	-0.522*** (0.158)	-0.354*** (0.092)	-0.339** (0.153)	-0.250*** (0.091)
WAGE 25	-0.520*** (0.154)	-0.332*** (0.089)	-0.410*** (0.144)	-0.274*** (0.085)
WAGE 26	-0.840*** (0.142)	-0.527*** (0.083)	-0.767*** (0.131)	-0.498*** (0.078)
WAGE 27 – 28	-0.240 (0.156)	-0.142 (0.091)	-0.078 (0.149)	-0.054 (0.089)
WAGE 29	-1.447*** (0.142)	-0.924*** (0.083)	-1.351*** (0.131)	-0.882*** (0.078)
WAGE 30 – 33	-1.158*** (0.151)	-0.750*** (0.089)	-1.058*** (0.141)	-0.702*** (0.085)
WAGE 34 – 35	-0.466*** (0.179)	-0.339*** (0.099)	-0.261 (0.169)	-0.219** (0.093)
WAGE 36 – 37	-1.392*** (0.177)	-0.836*** (0.099)	-1.270*** (0.162)	-0.778*** (0.093)
<i>Obs</i>	4196	4196	4833	4833
Hansen J	0.711	1.167		
Hansen J p-val	0.399	0.280		
Kleibergen-Paap rk LM	363.5	519.8	396.4	526.1
Cragg Donald Wald F	51.96	66.63	53.72	65.89

2-step GMM estimation. Depvar is estimated exporter-sector dummy:  $\widehat{f}e_{i,t}^k$ .

Regressors are logs of instrumented TFP and sectoral wages. Wages are efficiency adjusted in (II);(IV). The coefficient on WAGE corresponds to elasticity for sector 17 – 18.

For every other sector: elasticity given by sum of coef. WAGE and coef. of sector.

Estimates robust to an arbitrary form of heteroskedasticity.\*\*\* p<0.01, \*\* p<0.05, \* p<0.1

Years: 1995-2007 for (I)-(II); 1995-2009 for (III)-(IV). Year fixed effects included.

The parameters of the model are precisely estimated across the four specifications. The range of point estimates for the heterogeneity parameter is  $\theta = \{6.7, 7.8\}$  with a standard error of about 0.5.<sup>36</sup> The assumption of sector-specific coefficients on hourly wages is not rejected by the data.

Tab.2 reports implied sector specific factor shares together with the mean value of sectoral income shares observed in the data for the EU-15. There

<sup>36</sup>The point estimate for  $\theta$  is 4.5 when the production function is assumed common across sectors. See app.??.

Table 2: Sectoral factor share of inputs

	DATA	(I)	(II)	(III)	(IV)
<b>17-18</b>	0.68	0.82	0.79	0.79	0.78
<b>19</b>	0.72	0.97	0.88	0.95	0.87
<b>20</b>	0.67	0.64	0.68	0.65	0.68
<b>21-22</b>	0.63	0.61	0.65	0.62	0.66
<b>24</b>	0.69	0.74	0.74	0.75	0.75
<b>25</b>	0.65	0.74	0.74	0.74	0.74
<b>26</b>	0.62	0.70	0.71	0.70	0.71
<b>27-28</b>	0.66	0.78	0.77	0.78	0.78
<b>29</b>	0.64	0.62	0.66	0.62	0.66
<b>30-33</b>	0.66	0.66	0.68	0.66	0.69
<b>34-35</b>	0.76	0.75	0.74	0.76	0.75
<b>36-37</b>	0.65	0.62	0.67	0.63	0.68

Col. "D" reports income share of inputs for the EU-15 in WIOD (mean in 1995-2009).

Col. (I)-(IV) report factor shares of inputs computed from estimated coefficients.

are clear discrepancies between the two sets of parameters. In particular, the variance is much higher in estimated parameters. Nonetheless, the positive correlation between the data and the values implied from estimation is strong.

## 5 The proximity mechanism and the pattern of intersectoral specialization

To test for the presence of the proximity mechanism in the data, we need to define an empirical counterpart to the indicator of proximity to suppliers. As shown in sec.2.2.2, the microfounded proximity indicator is a weighted  $l^\theta$ -norm of the vector of bilateral trade frictions in each sector, aggregated across sectors according to the Cobb-Douglas price index with exponents given by sectoral expenditure shares:

$$\left[ \overline{PROX}_{i,t}^M \right]^{-1} = \prod_{s=1}^S \left\{ \sum_{n=1}^N \pi_{ni,t}^s \tau_{ni}^\theta \right\}^{\gamma^s / \theta} \quad (27)$$

Furthermore, as this indicator is contingent on a given trade equilibrium, we need to define a valid instrument. We argue that such an instrument is given by the empirical counterpart to the measure of proximity endowment which, as shown in sec.2.2.1, is an index of trade frictions independent of the distribution of market shares:

$$\left[ PROX_i^M \right]^{-1} = \prod_{s=1}^S \left[ \tau_i^{M,s} \right]^{\gamma^s} \quad (28)$$

We first construct both indicators for countries of our sample in 1995-2009, and show that the proximity characteristic is highly persistent and strongly correlated with the indicator of proximity endowment. We then test whether locational comparative advantage, defined as the log difference in instrumented proximity interacted with the input intensity characteristic of the sector, contributes to determining intersectoral specialization in the residual component of RCA rankings.

## 5.1 The proximity characteristic

The proximity characteristic is constituted by four components: bilateral trade frictions, bilateral sectoral market shares, expenditure shares in each sector, and a parameter measuring the dispersion of productivity.

According to our modelling of trade costs, bilateral trade frictions pick up impediments to trade linked to physical features of the trade network, such as information and transport costs. As is common in the literature, we consider that a satisfactory approximation to this symmetric cost component is bilateral distance  $dist_{ij}$ :  $\tau_{ij} = dist_{ij}^\rho$ , with  $\rho = 1$ .<sup>37</sup>

As benchmark for  $\theta$ , we use point estimates obtained in specifications (I) and (II): respectively 7.26 and 6.72 (see sec.4). Results are not sensitive to taking alternative values in the range of conventional values for this parameter:  $\theta \in [4.5, 8.5]$ .

Sectoral expenditure shares  $\gamma_{j,t}^k$  are constructed using data on total output, exports, and imports. Sectoral expenditure is  $X_{j,t}^k = PROD_{j,t}^k - EXP_{j,t}^k + IMP_{j,t}^k$ , with  $EXP$  total exports, and  $IMP$  total imports. Expenditure shares are given by  $\gamma_{j,t}^k = X_{j,t}^k / \sum_{s=1}^S X_{j,t}^s$ . Output data is taken from the WIOD database (see sec. 4). Total sectoral exports and imports are obtained from the COMTRADE database where we take information on world exports and imports for each country of our sample at the ISIC Rev.3 nomenclature at the 4-digit level, and we aggregate this data to the level of 13 manufacturing sectors to be consistent with output data provided by the WIOD. In practice, we use sectoral expenditure shares for the EU-12 in each year, and we check that results are not sensitive to the assumption of common sectoral expenditure shares across countries:  $\gamma_{j,t}^k = \gamma_t^k$  where  $\gamma_t^k$  is the expenditure share on sector  $k$  in year  $t$  for the EU-12.<sup>38</sup>

<sup>37</sup>The ranking of countries according to the proximity characteristic is not sensitive to picking an alternative value of  $\rho$ .

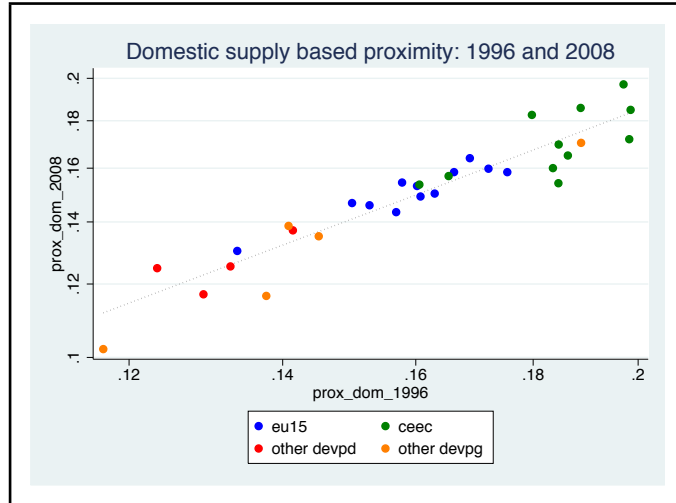
<sup>38</sup>The EU-12 are the EU-15 with the exclusion of Luxembourg, Belgium, and Netherlands for which output data is inconsistent with data on total exports and imports. Expenditure shares for the EU-12 are persistent overtime, and results are not sensitive to assuming  $\gamma_t^k = \gamma^k$  in 1995-2009.

Sectoral bilateral market shares  $\pi_{ij,t}^k = X_{ij,t}^k / X_{j,t}^k$  are constructed using bilateral imports data at the sectoral level together with data on sectoral output.<sup>39</sup>

The main bottleneck of this exercise is linked to obtaining plausible measures of domestic expenditure on domestic production  $X_{jj,t}^k = PROD_{j,t}^k - EXP_{j,t}^k$ . For 20% of observations, consumption of domestic varieties is negative. Data on domestic market share in adjacent years is used to adjust observations with negative values of domestic consumption. If domestic consumption is also negative in adjacent years, we use the median value of domestic market share across the years for which domestic market share is positive. Data adjustment is done in a way which leaves the original data on output and total imports unchanged. We redefine domestic market share  $\tilde{\pi}_{jj,t} = \pi_{jj,t'}$  using information in the adjacent year (or the median). We then redefine domestic consumption as  $\tilde{X}_{jj,t}^k = \tilde{\pi}_{jj,t} * X_{j,t}^k$ . Finally, we adjust the value of total sectoral exports to be consistent with adjusted domestic market shares given data on output and total imports.

$$\widetilde{EXP}_{j,t}^k = \frac{PROD_{j,t}^k - \tilde{\pi}_{jj}(PROD_{j,t}^k + IMP_{j,t}^k)}{(1 - \tilde{\pi}_{jj})}$$

Figure 1: Proximity characteristic in 1996 and 2008



<sup>39</sup>Data on bilateral imports from each of their trading partners for countries in the sample is taken from the COMTRADE database.

The proximity characteristic cannot be constructed for 10 countries of our sample due to lacking or inconsistent output data (see table ??). Fig.1 plots the proximity characteristic for the remaining set of countries in 1996 and 2008.<sup>40</sup> There is substantial variability across countries in proximity to suppliers. The sample is split in four subgroups according to the ranking of proximity: the EU-15, the CEECs, non-European developed, and non-European developing economies.<sup>41</sup>

Figure 2: Microfounded proximity (subset of countries)

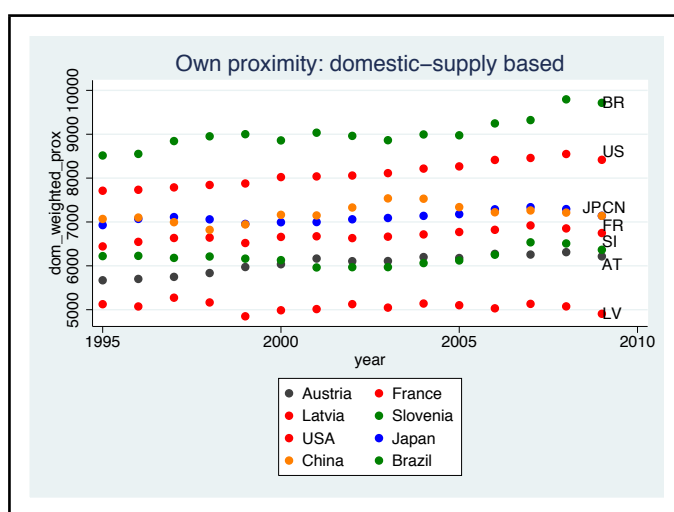


Fig.2 plots the indicator of distance to suppliers measured in kilometers, i.e. the reciprocal of the proximity characteristic, for two countries of each subgroup in 1995-2009.<sup>42</sup> For each of the four subgroups we plot the measure of distance to suppliers for the most and the least distant country from least cost inputs.<sup>43</sup>

The persistence of the proximity characteristic is not surprising. Bilateral trade costs are time invariant by construction. This conforms to the assumption that bilateral components of trade costs correspond to persis-

<sup>40</sup>For readability, proximity indicators are rescaled by  $10^3$  in this graph.

<sup>41</sup>Two countries behave differently from their subgroup: Russia is characterized by relatively high and Great Britain by relatively low proximity.

<sup>42</sup>See online appendix for exhaustive information on each subgroup overtime.

<sup>43</sup>There are two exceptions to this rule. In the EU-12 group we do not plot Great Britain as the most distant country because France is more representative of countries in the group. In the non-EU developing group we do not plot Russia as the least distant country because China is more representative of countries in the group.

tent characteristics of the trade network such as transport and information costs, as captured by bilateral distance. Moreover, the weighting function is defined by sectoral bilateral market shares. The market share is the probability that a specific source is revealed least cost in the sector. The fundamental country characteristic which determines this probability is the expected sectoral productivity. Levchenko and Zhang (2012) show that in half a century the distribution of sectoral technology stocks undergoes substantial changes. It is however likely that in the short time span studied in this paper this distribution is relatively stable.

The proximity characteristic is an endogenous object which depends on the distribution of sectoral market shares and consequently on domestic technology. As we seek to identify the contribution of trade frictions to determining the pattern of intersectoral specialization separately from domestic technology, we construct an indicator of proximity endowment which, by definition, is independent of the distribution of market shares. Proximity endowment is used as instrument for the proximity characteristic.

The second reason for instrumenting the proximity characteristic with an indicator of proximity endowment is measurement error which stems from the assumption that destinations do not differ in their restrictiveness to foreign supply. Distance to suppliers is underestimated by more for countries which are relatively closed to foreign supply because both domestic market share and the indicator of proximity need to be rescaled by the trade restrictiveness index (TRI) of the destination.<sup>44</sup>

## 5.2 The proximity endowment

It is not immediate how to construct the empirical counterpart of proximity endowment. As we seek to capture impediments to trade linked to physical features of the trade network such as transport or information costs, we consider that the fundamental component of proximity endowment is given by the physical location of the country relatively to all of its potential trade partners.

If geographical location is key, then one way to measure this endowment is to compute the length of the bilateral distance vector for this country with all of its potential suppliers including the country itself. This provides a time-invariant indicator of countries' centrality under the assumption that bilateral distance is a sufficiently good proxy of trade impediments other than trade policy. This assumption is difficult to verify without data on transport

---

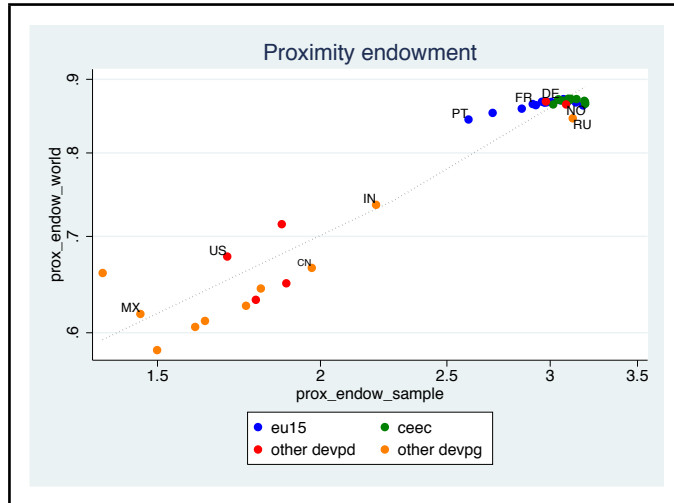
<sup>44</sup>See online appendix for details. This appendix reports TRI in manufacturing constructed by Kee et al. (2009) for 2008 to underscore variability in the data.

costs, but it is widely used in empirical applications.

$$[PROX_i^M]^{-1} = \left[ \sum_{n=1}^N dist_{in}^2 \right]^{0.5}$$

We compute two measures of proximity endowment. First, we measure proximity endowment as the reciprocal of the  $l^2$ -norm of the distance vector while restricting the number of potential input suppliers to the 42 countries of the sample. Second, we measure proximity endowment as the reciprocal of the  $l^2$ -norm of the distance vector with the 224 countries in the world on which internal and bilateral distance data are available (Mayer and Zignago (2011)). Fig.3 illustrates that countries' ranking is not sensitive to restricting the number of potential trade partners to the 42 countries of our sample.

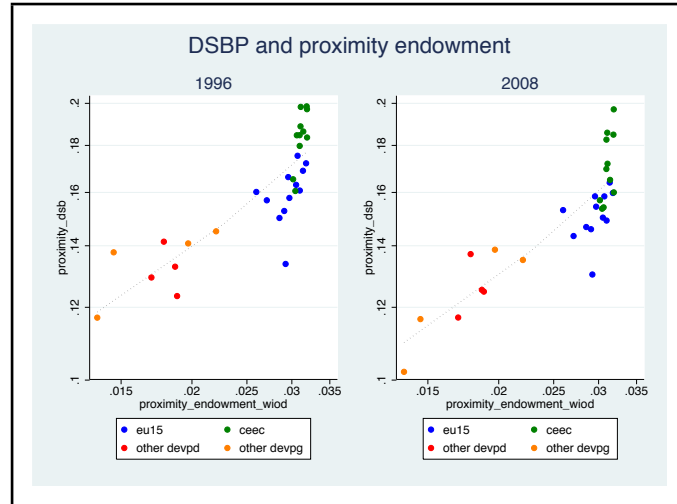
Figure 3: Proximity endowment



Finally, we check whether this time-invariant indicator is a good predictor of the proximity characteristic computed in the previous subsection. As illustrated in fig.4, the two proximity measures are strongly correlated. The measure of proximity endowment obtained when the number of potential suppliers is restricted to the 42 countries of our sample explains 70% of the variation in weighted proximity.<sup>45</sup> The caveat, as is clear from the graph, is that the instrument does a relatively poor job in capturing variation in proximity within subgroups.

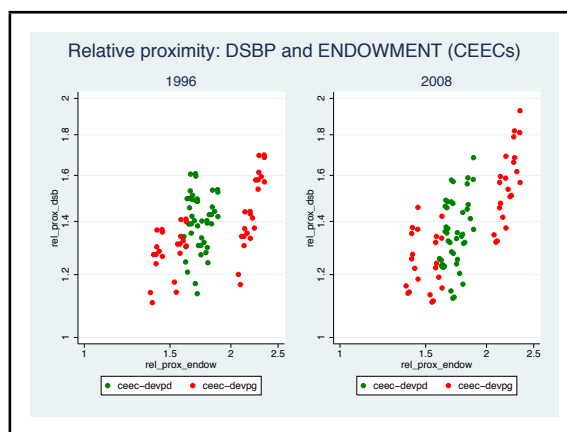
<sup>45</sup>Proximity endowment computed using the full set of potential trade partners captures less than 60% of variation in the time-varying proximity characteristic.

Figure 4: Proximity endowment and microfounded proximity



The ranking of countries in terms of relative proximity is invariant to using the indicator of proximity endowment or the time-varying proximity characteristic. As shown in fig.5 and 6, both measures allow splitting the sample in the high-proximity group which includes the EU-15 and the CEECs, and a low-proximity group which includes non-European emerging and developed economies.

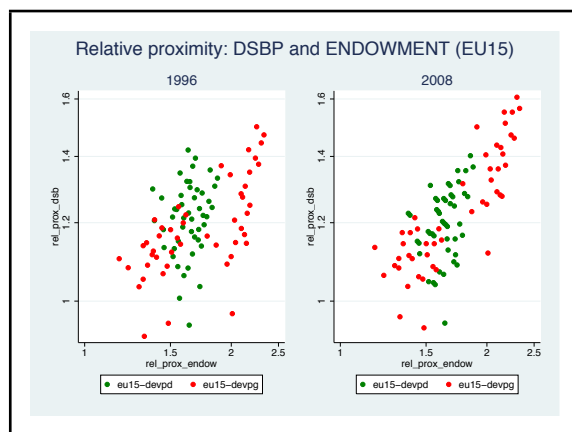
Figure 5: Relative proximity for the CEECs



However, the magnitude of the proximity gap differs across the two in-



Figure 6: Relative proximity for the EU15



dicators for a subset of countries. In particular, China and India appear relatively more distant from world technology when the indicator of proximity endowment is used. This discrepancy may be driven by the overestimation of the time-varying proximity characteristic in markets which are relatively closed to foreign supply (see online appendix for the formal presentation of this argument).

This subsection has shown that relative proximity should drive a persistent wedge in relative producer costs across countries of our sample. These cost differences become source of comparative advantage because the difference in production costs due to differences in the cost of the input bundle is most pronounced in sectors which use inputs relatively more intensively, i.e. in sectors where the weight of inputs in total costs is relatively high. In the next subsection, we check whether this mechanism is active in the data by testing whether high-proximity countries specialize in input-intensive sectors.

### 5.3 Proximity in intersectoral specialization

The proximity mechanism captures sector-specific differences in production costs which arise because countries differ in their ability to source inputs from the best technology worldwide. In this subsection, we check whether the cost advantage conferred by the ability to source inputs at relatively lower cost leads to specialization of high proximity countries in sectors which use inputs relatively more intensively. We find that once we control for technology stocks and factor endowments, the proximity mechanism indeed determines the ranking of relative sectoral exports across country pairs.

The indicator of sectoral relative proximity is computed as the log of the relative proximity characteristic for each pair instrumented with relative proximity endowment. The proximity characteristic is instrumented because we focus on the component of proximity which does not hinge on a specific distribution of market shares. In particular, we want to make sure that the indicator of proximity is orthogonal to domestic technology. Instrumented relative proximity is then interacted with the input intensity characteristic of the sector  $\widehat{\zeta}^k$  obtained in the second step of the estimation:  $\widehat{\zeta}^k \ln(\widehat{PROX}_{i,t}^M / \widehat{PROX}_{i',t}^M)$ .

The sample is split in two groups according to the proximity characteristic, with the EU15 and the CEECs in the ‘high’ and the non-European developed and developing countries in the ‘low’ proximity group. As we have already controlled for differences in relative technology stocks and relative wages, the level of development of the country should not be a source of within-group heterogeneity. However, as shown in the online appendix, differences in domestic market openness to foreign supply may bias the microfounded measure of relative proximity in cross-section and overtime. Consequently, we also split each country group in two subgroups by crossing the proximity criterion with the criterion of domestic market openness.

We work with the residuals of exporter-sector dummies obtained in the second-step of the estimation procedure. We rescale estimated residuals  $\widehat{\lambda}_{i,t}^k$  by the estimated heterogeneity parameter  $\widehat{\theta}$ , and compute all pairwise combinations of sectoral annual residuals  $(1/\widehat{\theta})(\widehat{\lambda}_{i,t}^k - \widehat{\lambda}_{i',t}^k)$  where  $i \in H$  are countries of the high proximity group, and  $i' \in L$  are countries of the low proximity group.

We estimate (29) on data pooled for all years. Exporter-year fixed effects  $\{fe_{i,t} - fe_{i',t}\}$  are included to control for characteristics of the benchmark sector for each exporter and year.

$$\frac{1}{\widehat{\theta}} \left[ \widehat{\lambda}_{i,t}^k - \widehat{\lambda}_{i',t}^k \right] = \beta_0 + \beta_1 \ln \left\{ \left( \frac{\widehat{PROX}_{i,t}^M}{\widehat{PROX}_{i',t}^M} \right)^{\widehat{\zeta}^k} \right\} + fe_{i,t} - fe_{i',t} + \eta_{ii',t}^k \quad (29)$$

The coefficient of interest is  $\beta_1$ : according to the model,  $\beta_1$  should be positive and close to 1. In practice, there are potentially multiple sources of measurement error in the data which we have sought to eliminate via instrumenting. Nonetheless, it is prudent to focus on the sign rather than on the absolute value of the estimated coefficient.

Table 3 shows that the proximity mechanism contributes to determining intersectoral specialization in the residual component of RCA rankings in the way predicted by the model, with high proximity countries producing relatively more for world markets in sectors which use inputs relatively more intensively:  $\beta_1$  is positive and significant across specifications. In col.(1)-(4), this result is obtained when countries are grouped in 2 bins according to the proximity characteristic while in col.(5)-(6) we additionally split low proximity countries in two groups according to the initial trade restrictiveness of their domestic markets.

Table 3: Proximity mechanism in the residual component of RCA rankings

	<i>all</i> (I)	<i>all</i> (I)	<i>all</i> (IV)	<i>all</i> (IV)	<i>both-to-devpd</i> (I)	<i>both-to-devpg</i> (I)
<i>relprox * inpint</i>	0.689*** (0.064)	0.375*** (0.093)	1.255*** (0.100)	0.658*** (0.152)	1.288*** (0.101)	0.176** (0.078)
<i>recent</i>		0.585*** (0.126)		1.033*** (0.200)		
Obs	17748	17748	20097	20097	8883	8865
$R^2$	0.674	0.674	0.665	0.665	0.541	0.776
Recent FE		YES		YES		

Depvar is rescaled relative residual component of the exporter-sector dummy:  $1/\theta [\widehat{\lambda}_{i,t}^k - \widehat{\lambda}_{i',t}^k]$ .

In (I) instruments are R&D personnel, real capital stocks, efficiency-adjusted workforce.

In (IV) instruments are deflated R&D expenditure and raw data on workforce.

'*relprox \* inpint*' is log of relative proximity interacted with sectoral input intensity.

'*recent*' is interaction between proximity and the 2001-2007(9) subperiod.

Exporter-year fixed effects are included in each specification.

col.1-4: countries split in two groups according to proximity ranking.

col.5-6: EU15 and CEECs to resp. developed and developing.

Col.(1) and (3) show that results are robust to the set of instruments used in the second step of the estimation. In col.(1)-(2) the set of instruments is R&D personnel, real capital stocks, and efficiency-adjusted sectoral workforce while in col.(3)-(4) it is deflated R&D expenditure and sectoral workforce unadjusted for efficiency. Results are qualitatively similar in (II) and (III) (not shown). In col.(5) and (6) it is shown that the proximity mechanism plays out strongly relatively to countries with low centrality and low levels of trade restrictiveness. Evidence in favor of the proximity mechanism is weaker relatively to countries with low centrality and high trade restrictiveness. The proximity mechanism becomes a stronger predictor of the ranking of sectoral exports in the recent period as shown in col.(2) and (4).

Table 4 shows more evidence on the proximity mechanism by splitting the sample of countries in four subgroups. The mechanism is present in the

data in the majority of specifications.

Table 4: Proximity mechanism in the residual component by subgroup

	(I)	(II)	(III)	(IV)
<i>eu15-to-devpd</i>	1.379***	2.359***	1.344***	2.263***
<i>std-error</i>	(0.134)	(0.224)	(0.125)	(0.207)
<i>nb-obs</i>	5541	5541	6399	6399
<i>R</i> <sup>2</sup>	.485	.483	.477	.476
<i>ceec-to-devpd</i>	1.151***	2.242***	0.890***	1.712***
<i>std-error</i>	(0.156)	(0.259)	(0.142)	(0.233)
<i>nb-obs</i>	3342	3342	3894	3894
<i>R</i> <sup>2</sup>	0.517	0.518	0.507	0.506
<i>eu15-to-devpg</i>	0.165	0.356**	0.254**	0.520***
<i>std-error</i>	(0.105)	(0.177)	(0.103)	(0.171)
<i>nb-obs</i>	5529	5529	6100	6100
<i>R</i> <sup>2</sup>	0.742	0.740	0.741	0.738
<i>ceec-to-devpg</i>	0.191*	0.623***	0.127	0.489***
<i>std-error</i>	(0.113)	(0.188)	(0.108)	(0.178)
<i>nb-obs</i>	3336	3336	3704	3704
<i>R</i> <sup>2</sup>	0.782	0.779	0.784	0.780

(I)-(IV) differ in the set of instruments for TFP and hourly wages.

Years: 1995-2007 for (I)-(II); 1995-2009 for (III)-(IV).

Exporter-year fixed effects are included in each specification.

Robust standard errors in parentheses.\*\*\* p<0.01, \*\* p<0.05, \* p<0.1

This subsection has shown that locational comparative advantage, defined as the gap in the relative distance to suppliers interacted with the input intensity characteristic of the sector, contributes to determining the pattern of intersectoral specialization for a given pair of exporters. In the next section we use the structure of the model to quantify the contribution of the proximity mechanism to determining the pattern of comparative advantage relatively to the contribution of domestic technology and labor endowments. In particular, we check whether the pattern of specialization predicted by relative sectoral technology stocks is modified by the proximity mechanism.

## 6 Decomposition of comparative advantage

### 6.1 Does proximity matter?

We pool data on residuals of exporter-sector dummies obtained in the second step of the estimation and construct pairwise relative sectoral residuals for high relatively to low proximity countries. We regress these relative residuals on sectoral indicators of relative proximity to check which share of the

variance in the residual component of sectoral exports is explained by the proximity mechanism. Throughout this section the time-varying proximity characteristic is instrumented with proximity endowment.<sup>46</sup> The proximity mechanism explains 18-20% of the variance in the residual component of relative sectoral exports (see tab.5).

Table 5: Fraction of variance attributable to the proximity mechanism

	<i>all</i> (I)	<i>all</i> (II)	<i>all</i> (III)	<i>all</i> (IV)
<i>relprox * inpint</i>	2.777*** (0.282)	3.381*** (0.336)	2.583*** (0.255)	3.043*** (0.297)
$R^2$	<b>0.178</b>	<b>0.200</b>	<b>0.181</b>	<b>0.196</b>
Obs	17,748	17,748	20,097	20,097

Depvar is rescaled relative residual component of the exporter-sector dummy.

'*relprox \* inpint*' is log of relative proximity interacted with sectoral input intensity.

(I)-(IV) differ in the set of instruments for TFP and hourly wages.

We report the coefficient on relative proximity and the fraction of explained variance.

We check that the fraction of variance attributable to proximity is not explained by the intercorrelation of proximity with technology and wages by computing the coefficient of partial determination between proximity and the ranking of relative exports. This statistic measures the fraction of the variance in the residual component of relative sectoral exports attributable to the component of the proximity vector unexplained by technology and wages.<sup>47</sup>

The proximity vector is nearly orthogonal to vectors of instrumented TFP and wages. Consequently, the variation in the ranking of relative sectoral exports attributable to differences in trade frictions incurred in sourcing inputs is not reducible to the two other characteristics which according to the model should determine the pattern of specialization (domestic technology and factor endowments). The coefficient of partial determination between proximity and relative sectoral exports is 15-17% in the full sample (see tab.6).

We also check whether the contribution of proximity changes overtime by computing the coefficient of partial determination in cross section. Fig.7

<sup>46</sup>The non-instrumented measure is contingent on a given trade equilibrium because it weighs the components of the distance vector with observed market shares. The instrument corresponds to the indicator of proximity in a world where the most efficient supplier of each variety is unique across destinations, i.e. in a world without bilateral components of trade frictions (see sec.2.2).

<sup>47</sup>For some variable  $y$ , the coefficient of partial determination measures the fraction of variance in the residuals of  $y$  wrt  $x_i$  for  $i \neq j$  explained by residuals of  $x_j$  wrt  $x_i$ .

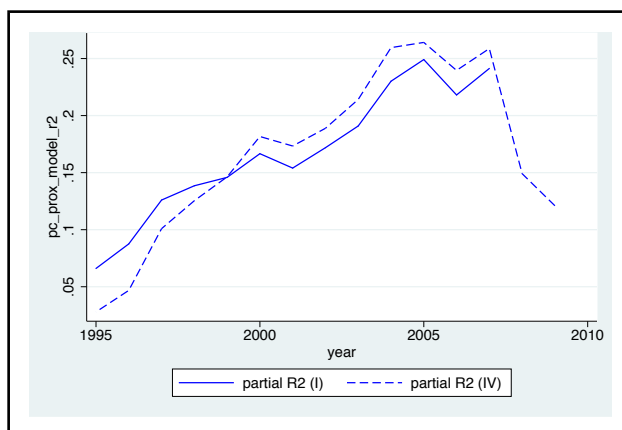
Table 6: Coefficient of partial determination (proximity, all years)

	<i>all</i> (I)	<i>all</i> (II)	<i>all</i> (III)	<i>all</i> (IV)
<i>resid - relprox</i>	2.601*** (0.305)	3.180*** (0.363)	2.446*** (0.283)	2.907*** (0.330)
$R^2$	<b>0.154</b>	<b>0.173</b>	<b>0.154</b>	<b>0.169</b>
Obs	17,748	17,748	20,097	20,097

Depvar is rescaled relative residual component of the exporter-sector dummy. 'resid - relprox' is the component of the vector of relative sectoral proximity orthogonal to instrumented TFP and wages. (I)-(IV) differ in the set of instruments for TFP and hourly wages.

documents that the fraction of unexplained variation in the ranking of relative sectoral exports attributable to the proximity mechanism increases fourfold in 10 years, exceeding 20% in 2000-2007. This result is not sensitive to the set of instruments used in the second step of the estimation (see solid line for specification (I) and dashed line for specification (IV) in fig.7).

Figure 7: Coefficient of partial determination (proximity, annual)



The proportion of variance attributable to proximity plunges in 2008-2009. The reduction in the weight of the proximity mechanism in co-determining the pattern of specialization is consistent with evidence that trade networks linked to production fragmentation across borders were severely hit in the aftermath of the financial crisis. Eaton et al. (2011) find that more than 80% of the decline in world trade in this period is due to the reduction in demand for manufactures (durable goods). If production of these goods is fragmented

across borders, the decrease in demand leads to a reduction in inputs' trade which magnifies the reduction in trade relatively to the reduction in GDP. This chain of events reduces world trade as a share of GDP and concomitantly reduces the importance of proximity to suppliers in determining the pattern of specialization.

## 6.2 Variance decomposition of RCA rankings: the intersectoral component

We have documented that the proximity mechanism plays a non negligible role in explaining the variation in relative sectoral exports. However, it could be that relative proximity succeeds in capturing the fraction of variance linked to features which vary by the exporter pair but does not contribute to the pattern of intersectoral specialization. We check this hypothesis by decomposing the variance of the intersectoral component of revealed comparative advantage (RCA) rankings across technology, factor endowments, and proximity. We control for characteristics of the exporter which do not vary at the intersectoral level by including exporter-year fixed effects in the specification.

Recall that relative exporter-sector dummies capture the components of sectoral costs which determine the pattern of comparative advantage together with the cost component of each exporter in the benchmark sector. We use the estimated parameters of the model and the instrumented components of TFP, wages, and proximity together with exporter-year fixed effects  $fe_{i,t}$  to compute the contribution of these three characteristics to determining the pattern of intersectoral specialization:

$$\begin{aligned} \frac{1}{\widehat{\theta}} \left( \widehat{fe}_{i,t}^k - \widehat{fe}_{i',t}^k \right) &= \alpha_0 + \alpha_1 \ln \left[ \frac{\widehat{z}_{i,t}^k}{\widehat{z}_{i',t}^k} \right] + \alpha_2 \ln \left\{ \left[ \frac{\widehat{v}_{i,t}^k}{\widehat{v}_{i',t}^k} \right]^{-(1-\widehat{\zeta}^k)} \right\} + \\ &\alpha_3 \ln \left\{ \left[ \frac{\widehat{PROX}_{i,t}^M}{\widehat{PROX}_{i',t}^M} \right]^{\widehat{\zeta}^k} \right\} + fe_{i,t} + fe_{i',t} + \xi_{ii',t}^k \end{aligned} \quad (30)$$

Estimated exporter-sector dummies  $\widehat{fe}_{i,t}^k$  and estimated productivity heterogeneity parameters  $\widehat{\theta}$  are used to compute rescaled relative sectoral exporter-sector dummies for high proximity countries  $i$  relatively to low-proximity countries  $i'$ :  $1/\widehat{\theta} \left( \widehat{fe}_{i,t}^k - \widehat{fe}_{i',t}^k \right)$ ,  $\forall k, i, i'$ .

Instrumented sectoral TFP is used to compute relative technology stocks:  $\left(\widehat{z}_{i,t}^k/\widehat{z}_{i',t}^k\right)$ . Instrumented sectoral wages and estimated input intensity parameters are used to compute the relative sectoral cost component linked to the use of labor in production:  $\left(\widehat{\nu}_{i,t}^k/\widehat{\nu}_{i',t}^k\right)^{-(1-\widehat{\zeta}^k)}$ . Instrumented micro-founded proximity and estimated input intensity parameters are used to compute the relative sectoral cost component linked to the use of inputs:  $\left(\widehat{PROX}_{i,t}/\widehat{PROX}_{i',t}\right)^{\widehat{\zeta}^k}$ .

According to the model, a regression of rescaled relative exporter-sector dummies on technology, wages, proximity, and exporter-year fixed effects should produce  $\alpha_1 = \alpha_2 = \alpha_3 = 1$ . In tab.7 we report the results of estimating (30) in the four main specifications. Consistently with the underlying model, the null hypothesis of coefficients' equality cannot be rejected in most specifications. But estimated coefficients are statistically different from 1. Col.(5) and (6) report the standardized regression coefficients for specifications (I) and (IV) respectively. To underline the relative importance of TFP and proximity, notice that one standard deviation in TFP increases relative exports by 2.5 (resp.1.9) standard deviations while 10(resp.5) standard deviations of proximity are needed to produce the same result.

Table 7: The intersectoral component of RCA rankings

	<i>all</i> (I)	<i>all</i> (II)	<i>all</i> (III)	<i>all</i> (IV)	$\beta$ -coef (I)	$\beta$ -coef (IV)
<i>tfp</i>	2.143*** (0.110)	2.105*** (0.107)	2.124*** (0.111)	1.994*** (0.107)	2.50	1.94
<i>wage</i>	1.981*** (0.112)	1.919*** (0.109)	2.291*** (0.120)	2.178*** (0.117)	2.32	2.08
<i>proximity</i>	1.668*** (0.160)	2.964*** (0.274)	1.642*** (0.156)	2.861*** (0.265)	0.24	0.40
$R^2$	0.731	0.731	0.731	0.726		
Obs	17,748	17,748	20,097	20,097		

Depvar is rescaled relative exporter-sector dummy:  $1/\theta [fe_{i,t}^k - fe_{i',t}^k]$ .

(I)-(IV) refer to alternative instrumenting procedures for technology and wages.

Col.(5)-(6) report standardized regression coefficients. Years: 1995-2007 for (I)-(II);

1995-2009 for (III)-(IV). Exporter-year fixed effects are included in each specification.

Standard errors are clustered by pair.\*\*\*  $p < 0.01$ , \*\*  $p < 0.05$ , \*  $p < 0.1$

Tab.8 reports the coefficient of partial determination between technology, wages, proximity and relative sectoral exports for the full sample and by subgroup in the specification with exporter-year fixed effects. Technology largely outweighs the proximity mechanism in explaining the residual



variation in relative sectoral exports at the intersectoral level. The relatively minor contribution of proximity may be in part an artefact of the simplifying assumption that all sectors use the same input bundle.

Table 8: Coefficient of partial determination: full sample and subgroups

	<b>tfp</b>	<b>wage</b>	<b>proximity</b>
<b>all</b>	<b>0.28</b>	<b>0.24</b>	<b>0.04</b>
<b>both-to-devpd</b>	<b>0.25</b>	<b>0.19</b>	<b>0.02</b>
<i>eu15-devpd</i>	0.26	0.17	0.03
<i>ceec-devpd</i>	0.26	0.23	0.00
<b>both-to-devpg</b>	<b>0.32</b>	<b>0.30</b>	<b>0.04</b>
<i>eu15-devpg</i>	0.34	0.33	0.06
<i>ceec-devpg</i>	0.27	0.26	0.02

Our objective is to quantify the fraction of the variance attributable to proximity out of total explained variance at the intersectoral level. We compute this as the ratio of semipartial  $r^2$  of proximity to the sum of semipartial  $r^2$  for instrumented TFP, wages, and proximity. This sum defines the fraction of the variance in relative sectoral exports unexplained by exporter-year fixed effects and uniquely associated with these three variables.<sup>48</sup> By construction, the fraction of variance unexplained by exporter-year fixed effects but associated with more than one of these three regressors is excluded. Total intersectoral variance is defined in this restrictive way to avoid doublecounting.

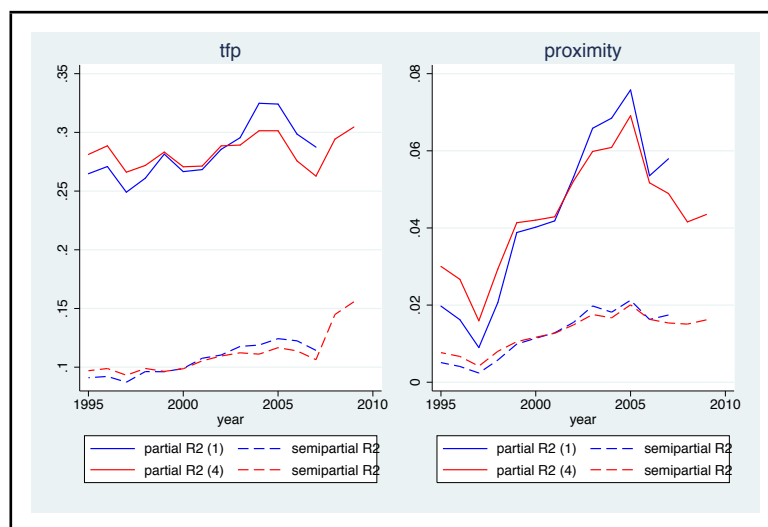
When we pool data for all years in the full sample, total explained variance at the intersectoral level uniquely associated with instrumented TFP, wages, and proximity is .20 in (I) (.23 in (IV)), i.e. about 30% of total explained variance. Domestic technology corresponds to 53% (resp. 47%) of this total while just 6% (resp.5%) is reconducible to proximity.

When we look at the sequence of cross sections the proximity mechanism is found to play an increasing role in the intersectoral component of RCA rankings. Fig.8 reports the squared partial and semipartial correlation for domestic technology and proximity for specifications (I) and (IV). Both specifications are reported to check the sensitivity of results to the set of instruments for TFP and wages used in the second step of the estimation.

The coefficient of partial determination illustrates that the fraction of the residual variance attributable to technology remains stable at slightly

<sup>48</sup>The semipartial  $r^2$  reports the fraction of total variance in relative sectoral exports explained by each regressor residualized with respect to all the other regressors.

Figure 8: Partial and semipartial  $r^2$  in cross section: full sample



less than 30% in 1995-2007 while the fraction attributable to the proximity mechanism increases from 2 to about 7%.

If we restrict attention to total explained variance at the intersectoral level, the fraction attributable to the proximity mechanism increases by 5 percentage points in 1995-2007 in (I), from 3 to 8%, while the fraction attributable to technology increases from 53 to 55%.<sup>49</sup>

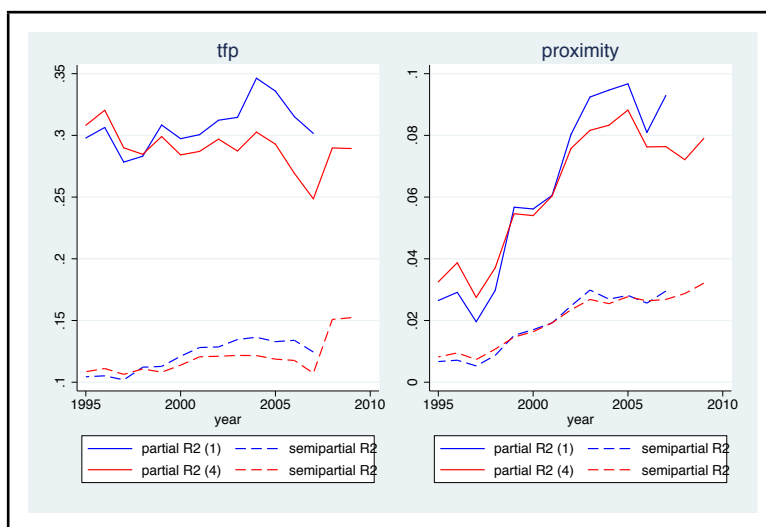
We nuance these findings by conducting the variance decomposition in cross section by subgroups. Fig.9 reports the squared partial and semipartial correlation for domestic technology and proximity for specifications (I) and (IV) for the EU-15 relatively to low proximity countries. Results for this subsample broadly replicate our findings for the full sample. Total explained variance at the intersectoral level increases from .20 to .23 in 1995-2007 in (I). The fraction of this total attributable to technology increases by 2 percentage points from 53 to 55%. Over the same period the contribution of proximity to explaining the intersectoral component of RCA rankings increases by 10 percentage points, from 3 to 13%.<sup>50</sup>

Fig.10 graphs the contribution of technology to explaining the variation in relative sectoral exports separately for the EU-15 and for the emerging

<sup>49</sup>The sum of semipartial  $r^2$  for instrumented TFP, wages, and proximity increases from .17 to .21 in (I), and from .20 to .22 in (IV). In (IV) the contribution of proximity increases from 4 to 7%.

<sup>50</sup>In (IV) the contribution of proximity increases from 4 to 12% in 1995-2007.

Figure 9: Partial and semipartial  $r^2$ : EU15



economies of Central and Eastern Europe relatively to low proximity countries. The fraction of residual variance attributable to technology is just .17 for the CEECs in 1995, much lower than the corresponding statistic for the EU-15. However, by 2009 almost 40% of the residual variation in relative sectoral exports is attributable to the intersectoral variation in technology for the CEECs. The corresponding statistic for the EU-15 is almost ten percentage points lower.

Over the same period the fraction of total explained variance at the intersectoral level increases from .11 to .26 for the CEECs relatively to low proximity countries. This means that the share of the variance uniquely associated with intersectoral determinants of comparative advantage increases from 15 to 36% of total explained variance. The contribution of technology to explaining the intersectoral component of RCA rankings is roughly stable at 52-53% of the total while the share attributable to proximity is reduced from 4 to 2%. These findings indicate that differential technology upgrading across sectors is likely to have increasingly shaped CEECs' specialization pattern on world markets.

As a final check we split the sample of low proximity countries in two subgroups according to the level of trade restrictiveness of their domestic markets. For the EU-15 the contribution of the proximity mechanism increases in both subsamples: from 5 to 10% (resp. from 2 to 10%) in the subsample of non European developed (resp. developing) economies.

Figure 10: Partial and semipartial  $r^2$ : technology

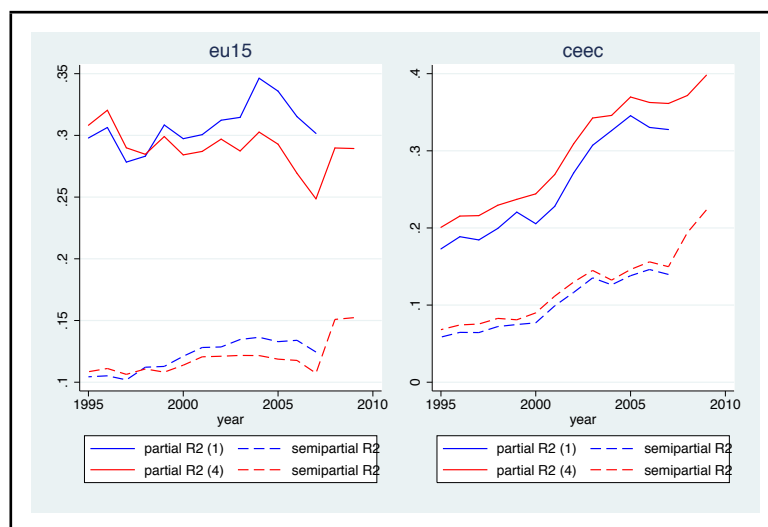


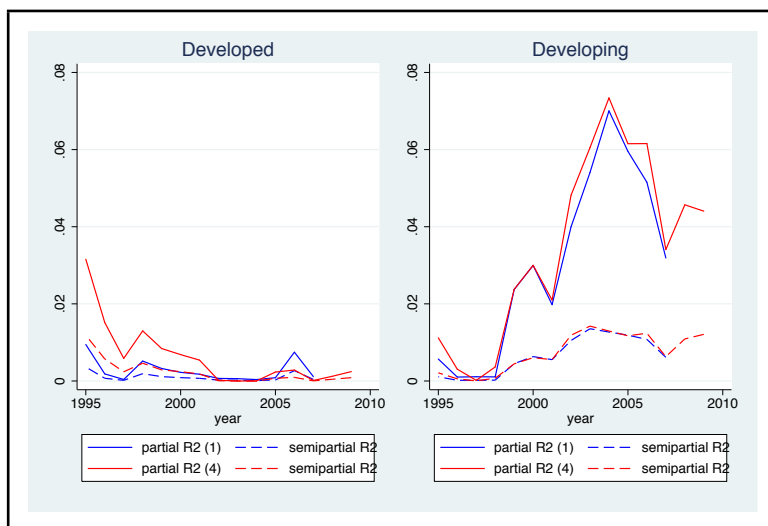
Fig.11 graphs the results for the CEECs. The fraction of the variance uniquely associated with proximity is about nil for the CEECs relatively to non-European developed economies. On the other hand, the proximity mechanism plays an increasing role in explaining the intersectoral variation in RCA rankings of the CEECs relatively to non-European emerging economies. Total variance uniquely associated with instrumented TFP, wages, and proximity increases from .07 to .20 in this subsample over 1995-2004. The fraction attributable to proximity increases from 1.5 to 6.4%, but is subsequently reduced to 2.8% by 2007.<sup>51</sup>

This section has documented that the proximity mechanism explains a small share of the variation in relative sectoral exports at the intersectoral level relatively to the variation attributable to technology and labor endowments. We think of this result as establishing a lower bound on the impact of the input cost channel as a consequence of our simplifying assumption that all sectors use the same input bundle in production.

We find robust empirical evidence that the importance of the proximity mechanism is increasing overtime. This result is consistent with a rapidly growing empirical literature which, starting with the seminal paper by Hummels et al. (2001), has documented increasing internationalization of the

<sup>51</sup>The reduction in sample size is reflected in results' sensitivity to the instrumenting procedure. In specification (IV) the contribution of proximity is stable at 2.3-2.6% while the contribution of technology is stable at 49% in 1995-2007.

Figure 11: Partial and semipartial  $r^2$ : proximity (CEECs)



production process.<sup>52</sup> Indeed, we expect the input cost component of comparative advantage to matter relatively more in shaping the pattern of intersectoral specialization when the pattern of production and trade costs makes it optimal for producers to increasingly segment production across borders.

This section also provides empirical evidence on the importance of technological upgrading in shaping the pattern of intersectoral specialization for emerging economies. We find that the sectoral ranking of relative technology stocks explains an increasing fraction of the total variation in relative sectoral exports for the emerging economies of Central and Eastern Europe. For developed economies the share attributable to technology is stable overtime.

Finally, we find that the fit of the model to the data is satisfactory. The full model explains about 3/4 of total variation in relative sectoral exports. Roughly a third of this total is uniquely associated with the intersectoral variation of the three explanatory variables which according to the model should determine the pattern of intersectoral specialization. Domestic technology explains between 53-55% of the variation in the intersectoral component of RCA rankings while the cost advantage conferred to high-proximity countries by the ability to source inputs relatively more cheaply explains between 3 and 8% of the intersectoral variation. For countries of the EU-15 the contribution of the proximity mechanism increases by 10 percentage points in 1995-2007,

<sup>52</sup>The non exhaustive list is Daudin et al. (2011), Johnson and Noguera (2012c,a,b), Stehrer (2012) and Borowiecki et al. (2012).

from 3 to 13% of total explained variance at the intersectoral level.

### 6.3 Does technology determine the pattern of intersectoral specialization under positive trade costs?

We check whether the pattern of intersectoral specialization observed in our sample of countries is different from what it would be if differences in input costs driven by trade frictions played no role in co-determining fundamental exporter-sector specific production costs.

In col.(1) of table 9 we report the sign of the relationship between overall RCA rankings and the proximity mechanism. In col.(2) we report the sign of the relationship between the intersectoral component of RCA rankings predicted by technology, wages, and proximity with the proximity mechanism. Col.(3)-(4) report the sign of the relationship with the proximity mechanism for (resp.) rankings of relative technology  $\hat{z}_{i,t}^k / \hat{z}_{i',t}^k$  and relative costs  $\hat{\omega}_{i,t}^k / \hat{\omega}_{i',t}^k$ . Finally, col.(5) recalls the sign of the relationship between the residual intersectoral component of RCA rankings and the proximity mechanism obtained in the third step of the estimation.

We document that the proximity mechanism contributes to determining the pattern of comparative advantage through the sectoral cost component  $\omega_{i,t}^k$  in most specifications. Furthermore, the proximity mechanism is always picked up in the residual intersectoral component of RCA rankings orthogonal to instrumented TFP and wages.

Nonetheless, the proximity mechanism does not inflect the pattern of intersectoral specialization driven by domestic technology. Indeed, whenever relative domestic technology rankings covary positively with proximity, the intersectoral component of RCA rankings picks up a positive link with the proximity mechanism. This is the case in col.(2) for the pattern of specialization of European countries relatively to non-European developed economies. However, whenever this is not verified, the intersectoral component of RCA rankings negatively covaries with relative sectoral proximity. In particular, this is the case in col.(2) for the specialization pattern of European countries relatively to non-European emerging economies.

We conclude that the proximity mechanism shapes the pattern of intersectoral specialization conditional on a given distribution of technology and labor endowments but it does not inflect the ranking of relative sectoral exports predicted by relative technology stocks. Consistently with the fundamental intuition of Ricardian models, the pattern of comparative advantage is determined by the ranking of relative sectoral technology stocks, even under positive trade costs and trade in inputs.

Table 9: The intersectoral component of RCA rankings

	overall RCA (1)	predicted RCA (2)	relative TFP (3)	relative cost (4)	relative residual (5)
all	–	–	–	+	+
both-to-devpd	+	+	+	?	+
both-to-devpg	–	–	–	+	+
eu15-to-devpd	+	+	+	–	+
ceec-to-devpd	+	+	+	+	+
eu15-to-devpg	–	–	–	+	+
ceec-to-devpg	–	–	–	+	+

In (1) depvar is rescaled relative exporter sector dummy:  $1/\hat{\theta}(\widehat{fe}_{it}^k - \widehat{fe}_{i't}^k)$ .

In (2) depvar is the ranking of relative exports predicted by instrumented TFP, wages, and proximity.

In (3) depvar is relative instrumented TFP:  $\ln(\widehat{z}_{it}^k/\widehat{z}_{i't}^k)$ .

In (4) depvar is relative sector-specific production cost:  $\ln(\widehat{\omega}_{it}^k/\widehat{\omega}_{i't}^k)$ .

In (5) depvar is rescaled relative residual of sectoral exports:  $1/\hat{\theta}(\widehat{\lambda}_{it}^k/\widehat{\lambda}_{i't}^k)$ .

## 7 Conclusion

It has become common knowledge that the process of production is increasingly fragmented internationally. In this paper we have investigated two questions: whether production unbundling has become a new source of comparative advantage and whether it has modified the determination of countries' specialization pattern on global markets. We answer 'yes' to the first, and 'no' to the second question.

Our main result is that production unbundling has coincided with an increasing role of input costs in shaping the pattern of comparative advantage in 1995-2009. But we also find that in our sample of 36 developed and emerging economies the ranking of relative sectoral technology stocks continues to determine the overall pattern of revealed comparative advantage just as in the benchmark multisectoral Ricardian world with bilateral trade frictions but no sector specific production characteristics.

We have shown that the only component in the cost of the input bundle which varies across countries is given by a composite index of trade frictions incurred in sourcing inputs from all potential suppliers. Conceptually, in relative terms, the proximity characteristic is a summary statistic of locational comparative advantage because it captures the cost advantage conferred to the country through its ability to source the cheapest inputs worldwide, rel-

atively to every other country in the world.

Relative proximity is also a summary statistic of the relative cost of living for any pair of countries. Indeed, we show that the overall price index can be decomposed in an index which captures the realized distribution of least cost technology in the world and an index of trade frictions which is country specific. Consequently, the closer the country is to the best world technology, and the lower is its cost of living relatively to other countries. A complementary result of the paper is that relative real wages can be computed by adjusting the ratio of nominal wages by relative proximity while circumventing the problem of constructing actual price indices.

As the cost share of inputs is sector-specific the wedge in the cost of inputs becomes source of comparative advantage. The model predicts that once we control for domestic technology and labor endowments, we should find that countries characterized by relatively high proximity to suppliers specialize in sectors which use inputs relatively more intensively. We present robust empirical evidence confirming this prediction in the data.

The input cost channel explains 15-20% of the residual variation in relative sectoral exports, but just 6% of total explained variation in the intersectoral component of RCA rankings if data is pooled in 1995-2009. In annual cross sections between 53-55% of the total variation in the intersectoral component of relative sectoral exports is attributable to technology while the contribution of the input cost channel increases from 3 to 8% in the full sample, and from 3 to 13% for EU-15 countries. The input cost channel is not only active at the intersectoral level, but acquires increasing importance overtime.

This line of research can be pursued in two directions. First, it would be interesting to investigate which type of shocks to the distribution of technology or to the structure of trade costs would be needed to inflect the pattern of comparative advantage given our result that the characteristics which determine intersectoral specialization are very slow moving. Second, it would be interesting to improve the mapping of the model to the actual sectoral structure of input sourcing to compute sector-specific theory-based indices of trade frictions. This more realistic production structure will help test our assertion that the results presented in this paper correspond to the lower bound of the true contribution of the input cost channel to the pattern of countries' specialization on world markets.



## References

- Bils, M. and P. Klenow (2000). Does schooling cause growth? American Economic Review 90, 1160–1183.
- Borowiecki, M., B. Dachs, D. Hanzl-Weiss, S. Kinkel, J. Poschl, M. Sass, T. C. Schmall, R. Stehrer, and A. Szalavetz (2012). Global value chains and the eu industry. WIIW Research Report (383).
- Caliendo, L. and F. Parro (2012). Estimates of the trade and welfare effects of *nafta*. NBER Working Paper Series (18508).
- Chor, D. (2010). Unpacking sources of comparative advantage: A quantitative approach. Journal of International Economics 82, 152–167.
- Costinot, A. (2009). An elementary theory of comparative advantage. Econometrica 77(4).
- Costinot, A., D. Donaldson, and I. Komunjer (2012). What goods do countries trade? a quantitative exploration of ricardo’s ideas. Review of Economic Studies 79, 581–608.
- Daudin, G., C. Riffart, and D. Schweisguth (2011). Who produces for whom in the world economy? Canadian Journal of Economics 44(4), 1403–1437.
- Deardorff, A. (2004). Local comparative advantage: trade costs and the pattern of trade. University of Michigan Discussion papers (500).
- Eaton, J. and S. Kortum (1999). International technology diffusion: theory and measurement. International Economic Review 40(3), 537–570.
- Eaton, J. and S. Kortum (2002). Technology, geography, and trade. Econometrica 70(5), 1741–1779.
- Eaton, J. and S. Kortum (2010). Technology in the Global Economy: A Framework for Quantitative Analysis. Unpublished manuscript.
- Eaton, J., B. Neiman, S. Kortum, and J. Romalis (2011). Trade and the global recession. NBER Working Paper Series (16666).
- Harrigan, J. (2010). Airplanes and comparative advantage. Journal of International Economics 82, 181–194.
- Harrigan, J. and C. Evans (2005). Distance, time, and specialization: Lean retailing in general equilibrium. The American Economic Review 95(1).

- Hummels, D., J. Ishii, and K.-M. Yi (2001). The nature and growth of vertical specialization in world trade. Journal of International Economics 54(1), 75–96.
- Johnson, R. C. and G. Noguera (2012a). Accounting for intermediates: Production sharing and trade in value added. Journal of International Economics 86, 224–236.
- Johnson, R. C. and G. Noguera (2012b). Fragmentation and trade in value added over four decades. NBER Working Paper Series (18186).
- Johnson, R. C. and G. Noguera (2012c). Proximity and production fragmentation. American Economic Review: Papers and Proceedings 102(3), 407–411.
- Kee, H. L., A. Nicita, and M. Olarreaga (2009). Estimating trade restrictiveness indices. The Economic Journal 119(1), 172–199.
- Levchenko, A. A. and J. Zhang (2012). Comparative advantage and the welfare impact of european integration. University of Michigan Discussion papers (626).
- Mayer, T. and S. Zignago (2011). Notes on cepii’s distance measures (geodist). CEPII working paper series (25).
- Stehrer, R. (2012). Trade in value added and the value added in trade. WIIW working paper series (81).
- Timmer, M. (Ed.) (2012, April). The World Input Output Database (WIOD): Contents, Sources, and Methods. Unpublished manuscript.
- Waugh, M. (2010). International trade and income differences. American Economic Review 100(5), 2093–2124.
- Yi, K.-M. (2010). Can multistage production explain the home bias in trade? American Economic Review 100(1), 364–393.