Repositorio Institucional de la Universidad Autónoma de Madrid

Opening the Interregional Trade "Black Box": The C-Intereg Database for the Spanish Economy (1995–2005)

Carlos Llano,[1] Almudena Esteban,[2] Julián Pérez,[3] and Antonio Pulido[3]

Abstract

Recent literature on border effect has demonstrated that national trade (intra- as well as interregional trade) tends to be more intense than international trade. Unfortunately, owing to the dearth of information on interregional economic relations, this important aspect of the economy has remained relatively ignored. In this article, the authors have described the methodology and main results of the largest estimation of Spanish interregional trade (1995–2005) carried out as a part of the C-Intereg project. The results obtained highlight the importance of the internal trade and the validity of the gravity model. Although the estimation focuses on the Spanish economy, the methodology can easily be applied to other European Union (EU) countries. In the upcoming years, this innovative database will be further developed in all its dimensions (space, time, and sectors) to serve as a promising

[1] Economic Analysis Department and L. R. Klein Institute-CEPREDE, Facultad de CC.EE y EE, Universidad Autónoma de Madrid, Madrid, Spain
[2] L. R. Klein Institute and CEPREDE, Madrid, Spain
[3] Applied Economics Department, L. R. Klein Institute and CEPREDE, Facultad de CC.EE y EE, Universidad Autónoma de Madrid, Madrid, Spain

Corresponding Author:
Carlos Llano, Economic Analysis Department and L. R. Klein Institute-CEPREDE, Facultad de CC.EE y EE, Módulo E-XIV, Universidad Autónoma de Madrid, Campus Cantoblanco, 28049 Madrid, Spain
Email: carlos.llano@uam.es

framework for the application of different techniques such as spatial interaction models or interregional input–output approaches.

Introduction

Despite the marked process of globalization and integration of organisation for economic cooperation and development (OECD) economies, the main bulk of economic transactions still continue to take place within the national boundaries. Indeed, a recent literature on *border effect* indicated that a pair of regions within a country tends to trade 10–20 times as much as an otherwise identical pair of regions across the countries.[1]

Another group of studies that focused on the United States and Japan (Hewings et al. 1998; Munroe and Hewings 1998; Hitomi et al. 2000) also found that within-country interregional trade is growing more rapidly than intraregional and international trade and that, in general terms, regions have become more closely linked. More recently, Jackson et al. (2006) noted that U.S. industries shipped approximately $7 trillion worth of goods interregionally in 1997 using the nation's highways, railroads, waterways, pipelines, and aviation systems (11 billion tons and 2.7 trillion ton-miles). This volume has increased by 18.8 percent (up to 14.5 percent and 9.9 percent for tons and ton-miles, respectively) since 1993. Hence, not only has the volume of interregional trade increased but trading patterns have also become more complex. The evidence from the European Union (EU) appears to confirm these findings. While the intensity of intra-European trade has grown in the recent decades (Van der Linden and Oosterhaven 1995; European Economy 1997, 2005), most of the trade still continues to be concentrated within each country (Oliver et al. 2003; Llano 2004a, 2004b; Ferreira 2008). As a result, although regional dependence on international trade may have increased in the majority of Europe's regions, one can expect most of the economic growth in a single region to be primarily explained by national causes (intraregional + interregional shocks) as opposed to the international ones.

Unfortunately, the absence of information regarding interregional trade makes it difficult to demonstrate such a hypothesis in the case of the majority of countries. Consequently, owing to this lack of data, a key area of economic activity remains largely unknown, thus limiting the possible analysis of the intersectoral and interregional linkages that account for the co-movements of regional cycles or interregional spillovers in terms of growth, employment, or productivity.

The size of this "black box" is enormous for most of the countries. For example, in the case of the Spanish economy that occupies a mid-position in terms of its

international openness ratio in the OECD context, over 80 percent of the national output is purchased by the country itself, with only 20 percent exported abroad. In addition, given the highly decentralized territorial organization of the country, it is possible to conclude that many strategic decisions regarding the competitiveness of the different regions are being taken without the required relevant information.

To solve this issue, a group of researchers from the L.R. Klein Institute-CEPREDE and eleven of the Spain's seventeen regional governments launched the C-Intereg project (www.c-intereg.es), set up as a permanent initiative focused on interregional economic relations in Spain and the rest of Europe. The first outcome of the project is the largest estimation, to date, of the Spanish interregional trade of goods (1995–2005), including information on the spatial origin and destination of flows (regions NUTS-2 and provinces NUTS-3), sixteen types of products, and six types of transport modes.

The main objective of the current article is to describe the methodology used in the estimation and analyze the main results obtained using the gravity model. In this regard, it is important to note that although the estimation focuses on the Spanish economy, this methodology can be easily applied to other EU countries. Indeed, a recent work has used a similar methodology to estimate interregional trade within Portugal as well as within the Iberian Peninsula as a whole (Ferreira 2008).

The rest of the article is structured as follows. In section on Interregional Trade: Estimation Methods and Previous Experiences, we have provided a brief review of the most common approaches used to estimate interregional trade at the international level as well as particularly in the case of Spain. The section on Estimating Spanish interregional trade in the C-Intereg Project presents the methodology used in the C-Intereg database, based on bilateral transport flows and regional prices. Finally, the interregional commodity flows for the period of 1995–2005 are analyzed in the Data Analysis section using the classical gravity model.

Interregional Trade: Estimation Method and Previous Experiences

Information concerning interregional commodity flows is scarce and incomplete in the majority of countries (Jackson et al. 2005, 2006). Therefore, interregional trade has to be estimated through nonsurvey techniques that are based on different hypotheses regarding the probability of interactions between a pair of points in space.

The current literature on international trade, transport economics, and spatial interaction models sets out several approaches to estimate commodity flows in space by taking into account a number of critical variables such as distance, industry specialization, infrastructure endowment, and transport connectivity. Gravity models, spatial choice, entropy maximizing paradigm, and neuronal network models are among the best-known methods (see Wilson 1970a, 1970b, 1973; Cesario 1975, 1976; McFadden 1981; Batten 1983; Haynes and Fotheringham 1984; Sen and

Table 1. Interregional Trade: Possible Approaches

Estimation Technique Used	Examples
Indirect estimation	
Gravity models	Sen and Smith 1995
Entropy maximizing paradigm	Batten 1983
Discrete choice; Neuronal networks . . .	Wilson 1970a; Cesario 1975; McFadden 1981; Nijkamp et al., 2004
Pool approach within an input–output model	Benvenuti et al. 1995
Direct estimation based on actual data	
Estimations based on transport flows	MRIO HERP (Polenske 1980; Hitomi et al. 2000). INTERTIO (Llano 2004a)
Estimations based on ad hoc surveys of producers and consumers	(Japanese Government 1974)

Smith 1995; Boyce and Bar-Gera 2003; Roy and Thill 2004; Nijkamp, Reggiani, and Tsang 2004).

Moreover, studies on interregional input–output (IO) modeling (Isard 1951, 1953; Moses 1955; Polenske 1980; Kim, Boyce, and Hewings 1983; Oosterhaven 1984; Miller and Blair 1985; Hewings et al. 1998; Hitomi et al. 2000; Liu and Vilain 2004; Kockelman et al. 2005; Ruiz and Kockelman 2006) and multiregional computable general equilibrium models and social accounting matrices (SAM; Stone 1961; Round 1995; Bröcker 1998; Jackson et al. 2006; Bröcker and Schneekloth 2006) also offer interesting approaches to estimate interregional trade flows, linking the IO and SAM frameworks to interregional transport flows (table 1).

The Estimation of Interregional Trade in Spain

Although several statistics with partial information on Spanish interregional trade are available (table 2), their consistency and comparability remains insufficient. Given these limitations, the main objective of the C-Intereg Project was to develop a permanent database with data on intra, interregional (and international) trade of all the Spanish regions, with different levels of sectoral and spatial disaggregation for both goods and services. The current database corresponds to the largest (1995–2005) estimation of interregional trade in goods at a regional level (Eurostat NUTS-2 level: eighteen Spanish regions or "Autonomous Communities") and the first at a provincial level (Eurostat NUTS-3 level: fifty-two Spanish provinces), including a detailed breakdown by the types of product (thirty categories of goods), transport mode (road, rail, ship, plane, pipe, and electricity system), and units (Euros, tons).

Table 2. Basic Statistics on Spanish Interregional Trade

	Accounting System	Statistical Sources
Top-down approach	Macroeconomic balance (Alcaide and Alcaide 2005) Regional input–output tables and regional accounts (Most of the Spanish regional input–output tables)	Interregional trade is deduced as the difference between regional inputs and outputs. Information based on *direct surveying of producers and/or consumers:* • *Available statistics*, such as Branch Industry, Surveys … (INE) • *Surveys designed "ad hoc"* for input–output table. Use of *interregional transport flow data*, with detailed information on the geographical origin and destination of flows: • National transport/freight statistics (volume). • Domestic prices using other information to value transport flows. *Fiscal information* concerning trade within Spain: • Fiscal declarations by companies and professionals with respect to the retransactions with other national agents (VAT and E-347 declarations).
Bottom-up approach	• Regional balance of payments (Parellada 1980; Oliver 1997; Pulido, López, and Llano 2000) • Complete interregional estimation (Llano 2001, 2004a; Oliver et al. 2003) Unavailable confidential statistics	

The methodology used for the estimates is based on previous works (Llano 2004a, 2004b; Pulido, López, and Llano 2000; Oliver 1997) and includes a number of improvements and extensions. It combines the most accurate data on Spanish transport flows of goods by transport modes (road, rail, ship, plane, pipe, and the Spanish electricity system) with additional information used to estimate fifty-two specific export price vectors, one for each province of origin, transport mode, and product type. The methodology also includes a process for debugging the original transport flow database that makes possible the identification and reallocation of multimodal transport flows and international transit flows (ITF), which may be originally hidden in the interregional flows. This procedure results in initial estimates of interregional trade flows in tons and current Euros, based on a combination of the transport and prices databases. Finally, a process of harmonization is applied to produce final figures in tons and Euros consistent with the figures for total output from the Spanish Industrial Survey and National Accounts. It should also be noted that at each stage up to the final aggregation into sixteen types of products and the final process of harmonization with the output figures, the methodology relies on the lowest level of disaggregation available. More specifically, the transport flows are based on classifications that range from 160 to 40 different types of products (depending on the four transport modes available) and price data on 11,000 different types of products.

Estimating Spanish Interregional Trade in the C-Intereg Project

The methodology used in the elaboration of the database is described in Figure 1 and can be divided into six steps described in the following sections.[2]

Harmonization of Transport Flows and Estimation of Unavailable Data

In this section, we have reviewed the situation of the Spanish statistics regarding the commodity transport flows and the transformations required for obtaining a full coverage of the movements accounted within the country during the period under study. Before going into further detail, it is important to bear in mind that for each year and transport mode, the goal is to estimate a set of origin and destination matrices (OD matrices) that capture all the deliveries in tons (Tn.) by the largest possible product disaggregation. For the purpose of clarity, the procedure applied for each transport mode is explained separately using a similar notation.

In general terms, the final aim is to estimate all the transport flows in tons described by $F_{rs,t}^{i}$ elements that capture the flows F of product i in year t with origin in region r and destination in region s. Then, for each transport mode, we will define the equivalent elements $F_{rs,t}^{iR}$, $F_{rs,t}^{iT}$, $F_{rs,t}^{iS}$, $F_{rs,t}^{iA}$, and $F_{rs,t}^{iP}$, where the superscript "R" denotes the flows using road transportation, "T" denotes train, "S" denotes ship, "A" denotes aircraft, and "P" denotes pipe (table 3).
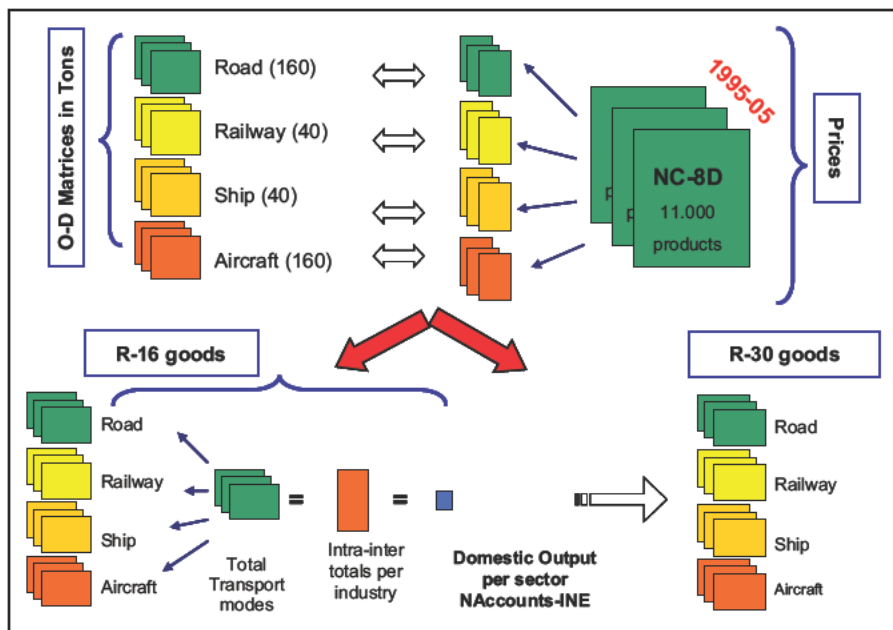
Figure I. C Intereg commodity database. Estimation process.

Transport flows by road ($F_{rs,t}^{iR}$): *Removing intramunicipal flows and ITF*. The Permanent Survey on Commodity Transport by Road (PSCTR), published by the Spanish Ministry of Public Works (Ministerio de Fomento), is one of the key sources of our original database on transport flows in tons (more than 80 percent of commodity flows in Spain are through road). Based on this key survey, for each year t and each type of product ($i = 160$), a large set of OD matrices of flows in tons were obtained. The main limitation of the original data is that from 2000 onward, the survey has been including the intramunicipal flows (IMF). Owing to the absence of this information for the period 1995–1999 and to assure the comparability of intraregional flows along the whole period (1995–2005), the IMF have been removed from the road transportation database.[3] In addition to this, as we will see in the section on First Debugging Procedure for Transport Flows in Tons, the original flows (OF) recorded by the PSCTR need to be depurated to prevent the inclusion of ITF hidden in the interregional deliveries. These processes are described in the following transformation:

$$F_{rs,t}^{iR} = OF_{rs,t}^{iR} - IMF_{rr,t}^{iRt} - ITF_{rs,t}^{iR} \qquad (1)$$

where $OF_{rs,t}^{iR}$ represents the original OD flows from the PSCTR survey, $IMF_{rr,t}^{iR}$ denotes the intramunicipal commodity flows included in the survey after 2000, and $ITF_{rs,t}^{iR}$ signifies the ITF moved by road (see the section on First Debugging Procedure for Transport Flows in Tons).

Table 3. Transport Statistics Used to Estimate Spanish Interregional Trade

Mode	Description and Main Features
Road	Permanent survey on goods transport by road (Encuesta Permanente de Mercancías por Carretera) • Source: Ministerio de Fomento (Spanish Ministry of Economic Development) • Product Disaggregation: 160 products (class. NSTR 3 digits) • Available since: 1995 onward Remarks: • Permanent survey (weekly basis) on the activity of a large sample of trucks in Spain: each trip includes origin–destination, type of product, volume, distance (km) . . . • Survey may also include international transit flows (ITF) moved from ports/air ports to final locations • It should be noted that the figures obtained from truck surveys may not be consistent with figures on production/purchases from firms/household surveys
Rail	RENFE statistics on complete wagon and container flows • Source: information from the Statistics Department of RENFE • Product Disaggregation: approximately forty categories (RENFE own classification) Remarks: • Every domestic flow recorded: high quality, low product detail. • No information on products transported by Container (30 percent of rail flows)
Ship	Statistics from Spanish Ports (Puertos del Estado) Indirect estimation of interregional flow matrices using optimization procedure based on: a) Tons loaded/unloaded by each Spanish port, kind of flow, and type of product. Source: *Statistical Yearbook*. Puertos del Estado. • Data: Annual. Twenty seven Spanish ports. • Product Disaggregation: forty products (Spanish Ports' own classification) b) Set of Spanish domestic flow matrices with ports of origin and destination 1989. Source: Domestic maritime flows by Origin and Destination 1989. Puertos del Estado: • Data: Annual. thirty eight largest Spanish ports (at that time) • Product disaggregation: fifty two products (CSTE)
Aircraft	O/D Matrices of domestic flows of goods by airport of Origin and Destination 1995. AENA. • Source: AENA and Ministerio de Fomento (Spanish Ministry of Public Works). • Data: Annual. Main Spanish Airports. • Product Disaggregation: None Remarks: • No information on sectoral disaggregation of domestic flows by air
Pipe	O/D matrix of oil flows using pipe 1995 • Source: CLH (main oil distributor in the Spanish market). 1993 • Product Disaggregation: None Remarks: • Indirect estimation using OD matrix obtained by Department of Economic Development and TEMA Consulting Group from CLH in 1993.

Transport flows by train ($F_{rs,t}^{iT}$): Assessing sectoral disaggregation to containers. Information on commodity flows by train is provided by the national rail network RENFE (a former national monopoly). Commodities are moved either in "complete wagon" (CW) or "containers" (CNT; 30 percent of tons were moved by CNT). Unfortunately, RENFE does not provide detailed information in relation to the types of products included in the containers. Hence, the allocation of these flows to specific products is done using the sectoral structure of railway flows that use "complete wagon," have the same region of origin and can be transported in containers. As not all types of products i can be transported in containers, we have used the superscript j to denote those types of products that are usually transported in this submode.[4] Accordingly, for each year t, the final amount of flows by railway from origin r to destination s of product i ($F_{rs,t}^{ir}$) is obtained from the following expression:

$$F_{rs,t}^{iT} = \text{CW}_{rs,t}^{i} + \left(\text{CNT}_{rs,t}^{i} \times \frac{\text{CW}_{r,t}^{i}}{\sum_{i=1}^{j} \text{CW}_{r,t}^{i}} \right) - \text{ITF}_{rs,t}^{iT}, \qquad (2)$$

where $\text{CW}_{rs,t}^{i}$ denotes the commodities of type i measured in tons, transported in "complete wagon" from region r to s, $\text{CNT}_{rs,t}^{i}$ is the total amount of commodities in tons transported in containers from region r to s and $\left(\dfrac{\text{CW}_{r,t}^{i}}{\sum_{i\,1}^{j} \text{CW}_{r,t}^{i}} \right)$ is the share of each commodities j that were transported in "complete wagon" (and can be transported in "containers") with origin in r, when compared with the total amount of commodities j transported in "complete wagon" with origin in that region r. Finally, $\text{ITF}_{rs,t}^{iT}$ captures the ITF moved by train, which, as in the case of commodity flows by road, also need to be depurated (see the section on Estimating Interregional Trade Prices).

Transport flows by ship ($F_{rs,t}^{iS}$): Estimation of OD matrices using updated margins and old OD matrices. Owing to the absence of an up-to-date set of OD matrices by ship with the required sectoral breakdown, these are estimated by means of a common approach used in the IO literature to update old intersectoral structures using new data for the margin totals. This procedure, based on previous works (Allen and Gossling 1975; Dijkman and Burgess 1994; Möhr et al. 1987), uses the long-scale optimization procedure (Fylstra et al. 1998; Lasdon and Smith 1992) to update the most recent OD matrices available (published by Spain's State Ports Authority—*Puertos del Estado*—in 1989) with new data on the volume of commodities loaded/unloaded in each port (1995–2005 *Statistical Yearbook*. Puertos del Estado[5]). This procedure can be briefly summarized in three steps:

1. Based on the available information for 1989, we built a set of 1989 OD maritime flows ($F_{rs,89}^{iS}$) that consider $r = 27$ possible origins and destinations (the main Spanish ports at that time) and $i = 46$ types of products (corresponding to the CSTE—Commodity Classification for Transport in Europe, that is, the classification used in the 1989 reference).

2. We built a parallel database with the annual data published for domestic maritime volumes loaded and unloaded in the main Spanish ports for each year (1995–2005) and the type of product ($i = 40$ categories). At this stage, it is important to note that for some years and types of products, the total amount of tons loaded (l) and unloaded (u) do not match. Given that in an OD matrix the sum along the rows must coincide with the sum along the columns, we have rescaled the largest vector to the total of the shortest one, using the following procedure.

$$
\begin{cases}
\text{If } \sum_s^{s=22} u_{s,t}^i < \sum_r^{r=22} l_{r,t}^i & \text{then} \quad l_{r,t}^i = l_{r,t}^i \times \dfrac{\sum_r^{r\,22} u_{s,t}^i}{\sum_r^{r\,22} l_{r,t}^i} \\[2ex]
\text{If } \sum_s^{s=22} u_{s,t}^i > \sum_r^{r=22} l_{r,t}^i & \text{then} \quad u_{s,t}^i = u_{s,t}^i \times \dfrac{\sum_r^{r\,22} l_{r,t}^i}{\sum_r^{r\,22} u_{s,t}^i}
\end{cases}
\tag{3}
$$

where $u_{s,t}^i$ is the original amount of commodity (i) unloaded (u) in region (s) and year (t) and $l_{r,t}^i$ is the original amount of commodity (i) loaded (l) in region (r) and year (t). Similarly, $u_{s,t}^i$ denotes the "new amount" of commodity (i) unloaded in region (s), while $l_{r,t}^i$ denotes the "new amount" of commodity (i) loaded in region (r) and year (t). As it is described in equation (3), the new amount is obtained by rescaling the larger with the lower total amount. Accordingly, for each year (t), we can obtain two new vectors of elements $u_{s,t}^i$ and $l_{r,t}^i$ that sum up the same amount of tons and maintain the structure of the original vectors. Therefore, for each year t and product i, these rescaled vectors are now ready to be used as margins for updating the old OD matrices.

3. Subsequently, for each year (1995 onward) and type of product ($i = 40$), using a Long-Scale Optimisation procedure (Fylstra et al. 1998; Smith and Lasdon, 1992), we obtained an OD flow matrix where the sums along the rows and columns coincide with the corresponding totals loaded and unloaded in each port, with regard to the year and type of product. The interactive procedure begins with the 1989 actual OD matrix for each type of product and concludes with a complete set of OD flow matrices for forty types of products from 1995 onward. In line with Möhr et al. (1987), this step can be described as a maximization procedure of the non-null values subject to the margin constrains.[6]

$$
F_{rs,t}^{iS} = \text{Max} \left[\sum_{r \in S} \sum_{s \in S} F_{rs,89}^{iS} \right],
\tag{4}
$$

Subject to:

$$\sum_{r \in S} F^{iS}_{rs,t} = l^i_{s,t} \quad (s = 1, 2, \ldots n),$$

$$\sum_{s \in S} F^{iS}_{rs,t} = u^i_{r,t} \quad (r = 1, 2, \ldots n),$$

where $F^{iS}_{rs,t}$ is an element on the new OD flow matrix by ship (S) estimated for year t (1995 onward), containing the flows in tons of product i moved from port r to port s; $F^{iS}_{rs,89}$ is an element on the old 1989 OD flow matrix by ship (S), containing the flows in tons of product i moved from port r to port s; $S = \{(r,s)|f_{rs} > 0\}$ represents the non-null flows traded between the twenty-seven main ports (r,s); $l^i_{r,t}$ is the vector of tons of product i loaded in the ports r in year t and $u^i_{s,t}$ is the vector of tons of product i unloaded in the port s in year t.

Transport flows by aircraft ($F^{iA}_{rs,t}$): Assessing the sectoral disaggregation of domestic OD matrices. Surprisingly, despite the existence of a single OD matrix of total domestic flows by aircraft in tons for each year ($DF^A_{rs,t}$), it has been proven to be impossible to obtain detailed data with respect to each sectoral disaggregation. Hence, for each year (1995–2005), the product disaggregation of this single matrix ($DF^A_{rs,t}$) has been deduced using the product specialization of international flights ($IF^{iA}_{r,t}$ for which data are available) for each airport of origin (r) and year (t). With this estimation, it is assumed that if a specific region has interregional exports of goods shipped by aircraft, then the sectoral structure of its domestic outflows to any of its destinations will be the same as those observed in international exports from the same airport. To increase the plausibility of such an assumption, we have only considered the international trade of Spanish regions with the nearest countries (France, Portugal, Germany, Italy, and Morocco), where the same competition structures in terms of transport modes operate. The flows are divided according to the 160 categories considered in the NSTR-3 digits used for road flows

$$F^{iA}_{rs,t} = DF^A_{rs,t} \times \frac{IF^{iA}_{r,t}}{\sum\limits_{i=1}^{i=160} IF^{iA}_{r,t}}, \tag{5}$$

where $F^{iA}_{rs,t}$ is an element on the new OD flow matrix by aircraft (A) in year t containing the flows in tons of product i moved from airport r to airport s. These sectoral flows are obtained by departing from the total amount of commodities in year t moved from airport r to airport s ($DF^A_{rs,t}$) as well as the sectoral structure of the international flows ($IF^{iA}_{r,t}$) for each commodity ($i = 160$) and airport of origin r in the percentage of the

total amount of commodities delivered by plane from the same origin $\sum_{i=1}^{i=160} \text{IF}_{r,t}^{iA}$.

Transport flows of refined oil products ($F_{rs,t}^{iP}$): Reallocation of oil products delivered by pipe. Owing to the special distribution process of the *refined oil products (gasoline, fuel-oil, kerosene, etc.)*, in which pipelines are used to approach the product to the final markets from distant refineries, the pipe bilateral flows should not be aggregated to the other transport flows but should only be used for the reallocation of some road and rail flows of the oil products.[7] From our point of view, if a region *n* with no refineries registers deliveries of oil products by road and train, it can be safely assumed that the exported product was originally produced in another region *r* with refinery, which is serving the region *n* by pipe. Consequently, the right allocation of the bilateral flows should first eliminate false exports attributed to the regions without refinery and reallocate them to the regions that originally produced the products and served them through the pipe.

To implement this, some manipulations of the OF are needed:

$$\text{OF}_{rs,t}^i = F_{rs,t}^{iR} + F_{rs,t}^{iT}.$$

$$F_{rs,t}^{iP} = \text{OF}_{rs,t}^i - \text{FR}_{rs,t}^i + \text{PF}_{rs,t}^i, \tag{6}$$

$$\text{PF}_{rs,t}^i = \text{FR}_{rs,t}^{i'} \times \frac{F_{rs,t=1993}^i}{\sum_{s=1}^{s=51(r\neq s)} F_{rs,t=1993}^i}, \tag{7}$$

where $F_{rs,t}^{iP}$ is an element of the final OD matrix of oil products moved by road and railway, which is obtained by the addition of three elements (equation 6):

a. The first one, $\text{OF}_{rs,t}^i$, denotes the original flows of oil products *i* moved by road ($F_{rs,t}^{iR}$) and train ($F_{rs,t}^{iT}$) in year *t* with origin in region *r* and destination in *s*;
b. The second, $\text{FR}_{rs,t}^i$, which is subtracted from the previous one, denotes a specific type of flow that fulfils the following requirements: (1) they are flows moved by road and railway of refined products *i* that could be transported by pipe and (2) they originated in regions (*r*) with no *refineries*. As we do not want to have regions without refinery exporting oil products to other regions, $\text{FR}_{rs,t}^i$ elements are thus eliminated from the road and railway database.
c. The third element, $\text{PF}_{rs,t}^i$, captures the $\text{FR}_{rs,t}^{i'}$ flows once they have been reallocated as exports from the regions with refineries, which are more likely to have originally produced the oil[8] and pumped it to the regions that, according to the original data, were exporting it by road and railway. As explained in equation (7), the reallocation of these flows is obtained by means of the only OD matrix available for the pipe, $F_{rs,t=1993}^i$, which was estimated for the year 1993.[9] Consequently, the reallocation of the "apparent exports" by road and truck of

regions without refinery to the regions with refinery is based on the percentages obtained from the most recent OD matrix available for flows of oil products through pipe.

Following these transformations, the element $F_{rs,t}^{iP}$ captures the final flows of a subgroup of refined oil products that could be transported by pipe, truck, and railway and whose regions of departure r have refineries. After this procedure, the obtained flows are then aggregated to the rest of the flows from the corresponding modes.

First Debugging Procedure for Transport Flows in Tons

Debugging interregional flows from ITF. Although all the information included in section on Harmonization of Transport Flows and Estimation of Unavailable Data refers to interregional flows in tons, the transported products may have an international origin/destination. Most of the estimations of the Spanish interregional trade do not tend to consider this complex issue (Ferreira 2008; Oliver 1997; Oliver et al. 2003). As a result, the international trade may be double-counted and the interregional trade of some specific regions may be overvalued. In the case of Spain, the highest risk in relation to ITF is associated with maritime-road and maritime-railway intermodal connections.[10] In earlier works, Llano (2001, 2004a) developed a meticulous approach to deal with this problem. Now, given the vast amount of data considered for the whole 1995–2005 period, the identification of international trade in transit along the Spanish roads is based on a more automatic procedure. The process of identification and elimination of ITF has focused solely on transport flows by road (which account for more than 90 percent of flows in Spain) and comprises the following three steps:

- For each year during the period 1995–2005, we obtained a database with the road flows between the fifty-two provinces and municipalities, where the country's twenty-seven main ports are located. This database contains flows for each of the 160 types of products considered in Spain's PSCTR.
- Using the comprehensive database of Spanish international trade (Spanish Tax Agency, www.aeat.es), we estimated the international flows that were exported/imported by ship and moved by road/railway from/to the point of origin/destination. The figures for each type of product and port were obtained by identifying, from the aforementioned database, flows matching the following restrictions: *Transport mode* = 1 (Ship); *Exports*: the province of *origin/destination* of the flow is different from the *Customs* point where the export/import is cleared (only Customs in the twenty-seven main Spanish Ports were considered).
- Next, by comparing the two databases mentioned earlier for each of the 160 types of products, we eliminated the flows from the PSCTR, which can be considered "ITF." By bearing in mind the different nature of the databases (the first is a survey based on a directory of transport companies, while the second is a register of all international trade operations), some limits were established to this

debugging process: (1) nonnegative figures may result: if the international flow is greater than the figure calculated in the PSCTR; the former can become 0 but not negative (2) for each product, the figures eliminated may be below 15 percent of the flow received (delivered) by the province from the same origin (destination).[11]

Allocation of multimodal flows. Another key limitation when estimating interregional trade using transport flows is the difficulty in identifying intermodal flows corresponding to the same transaction. When this issue is not appropriately taken into account by the used methodology, the growing complexity of current transport logistics may lead to an increasing bias in such estimations. In the case of Spain, the risk of intermodal connections is mainly associated with interregional trade between the mainland, the Islands (Balearics, Canary Islands), and Ceuta and Melilla (in Africa). However, owing to the existence of special fiscal regimes in some of these territories (Canary Islands, Ceuta, and Melilla), we have been able to access detailed fiscal information on interregional trade between the mainland and the Canary Islands, as well as with Ceuta and Melilla (Spanish Tax Authority, www.aeat.es). To our knowledge, this is the first time that such high-quality data on actual interregional trade between the Spanish mainland and the Islands has been included in a comprehensive estimation of interregional trade.

Estimating Interregional Trade Prices

In line with previous studies (Llano 2004a; Oliver et al. 2003; Oliver 1997; Pulido et al. 2000), we estimated interregional trade prices on the basis of *value/volume* relations deduced from detailed statistics contained in the Spanish Branch Surveys and International Trade databases. For each year, transport mode and type of product, we tried to estimate fifty-two export price vectors (one for each Spanish province, NUTS-3), denoted by $P_{r,t}^{im}$, at the lowest level of disaggregation, to capture price/quality differences among the regions for the same product. For the sake of simplicity, we used the superscript m to denote the four transport modes considered explicitly (R, T, S, and A).[12] Regional (NUTS-2) and national prices were used in the absence of provincial data. The estimation methodology comprises the following two steps:

a) Detailed estimation of prices (€/Tn.) for the base year 2000
   - *Industrial products*: The prices are deduced by dividing the data on value and volume published by the Survey of Industrial Products (INE-2000) for each region (NUTS-2) and with a high level of product disaggregation (4-digit NACE). Owing to the lack of information at a provincial level, all provinces within the same region share the same prices for each product, year, and transport mode.

- *Prices for agricultural products* are deduced from the National Survey on Agricultural Products (Eurostat and the Spanish Ministry of Agriculture, MAPA). Unfortunately, these prices are not available at regional level. Alternative sources (Eurostat) have been used for some specific products like energy.

At the same time, an alternative database has been estimated using information on international export prices at the lowest level of disaggregation (11,000 products; Combined Nomenclature, CN-8digits) for each year and transport mode at a provincial level. This database has been used as a reference to control for outliers in the previous estimation based on Branch Surveys.

b) Price projection for the remainder of the period (1995–2001; 2002–2005). Based on the price levels estimated for the base year 2000, the prices for the remaining years are estimated individually for each region, transport mode, and type of product using the available information on industrial and agricultural price indices at the lowest level of disaggregation (Eurostat). Owing to the lack of information at a regional level, for each type of product, the evolution of regional prices in 2000 followed the national growth rates for the remainder of the period. Before applying this procedure, we tested to verify whether the use of price indices could induce lower levels of instability than other alternative procedures (Oliver 1997; Pulido et al. 2000; Llano 2004a), such as inferring them directly from branch surveys or international trade databases for each year.

Translation of OD-Debugged Matrices from Tons into Monetary Units

Subsequently, the four sets of OD matrices of flows in tons by transport modes were translated into monetary units (Euros) by multiplying the corresponding four sets of price vectors estimated in the section on Translation of OD-Debugged Matrices from Tons into Monetary Units.[13] Accordingly, transport flows in tons $F^i_{rs,t}$ from region r to s of product i and year t were transformed into trade flows in Euros $M^i_{rs,t}$ with the same features, using specific export prices vectors, $P^{i\bullet}_{r,t}$.

All the OD matrices already valued in current Euros were then aggregated according to the largest common classification that allowed further comparisons and homogenizations (see the section on Final Screening of the R-16 OD Matrices of Goods in Euros).

$$M^i_{rs,t} = \left(F^{iR}_{rs,t} \times P^{iR}_{r,t} \right) + \left(F^{iT}_{rs,t} \times P^{iT}_{r,t} \right) + \left(F^{iS}_{rs,t} \times P^{iS}_{r,t} \right) + \left(F^{iA}_{rs,t} \times P^{iA}_{r,t} \right) \quad (8)$$

where $M^i_{rs,t}$ are the OD trade flows in monetary units, i denotes the type of product, r and s are the origin and destination regions, and t is the year. At this step, we used two different classifications: the main one, with sixteen types of product (R-16), and an extended one featuring thirty types (R-30), which maintains the corresponding

proportion with the first classification. Both these classifications (*R*-16 and *R*-30) were based on the NACE official classification and the specific classifications of the transport mode already considered.

Final Screening of the *R*-16 OD Matrices of Goods in EUROS

Finally, the trade flows for each of the sixteen types of products (*R*-16) were harmonized with the total output by region/province and industry obtained from different sources:

a) Industrial products (*R*3–*R*15)
   - For each of the fourteen OD matrices of industrial activities (excluding energy products), the intra- and interregional flows were harmonized with the available data published by the Spanish Industrial Survey (SIS). At the regional (NUTS-2) and provincial level (NUTS-3), the SIS provides information on total output per region r and industry i ($O^i_{r,t}$) as well as its geographical distribution by large markets: own region ($X\text{Intra}^i_{r,t}$); the rest of Spain ($X\text{Inter}^i_{r,t}$) and the rest of the world ($XRw^i_{r,t}$).

$$O^i_{r,t} = X\text{Intra}^i_{r,t} + X\text{Inter}^i_{r,t} + XRw^i_{r,t} \tag{9}$$

- Based on this information, for each year and industrial sector i, the "sums along rows of all the off-diagonal elements" in the OD matrix in Euros $M^i_{rs,t}$ were harmonized with the "Xinter," while the "on-diagonal" elements were harmonized with the "Xintra." Owing to the lack of information on the "amount of products consumed," no constraints were established for the "sums along columns." Thus, new OD matrices for "adjusted monetary flows" ($AM^i_{rs,t}$) were obtained for each year t and sector i, where the row totals for intra- and interregional flows coincided with the aggregated information from the SIS, while the bilateral structure was based on the whole set of OD matrices $M^i_{rs,t}$.

$$\text{Interregional } \textit{adjusted flows}: \ AM^i_{rs,t} = M^i_{rs,t} \times \frac{X\text{Inter}^i_{r,t}}{\sum\limits_{s=1}^{s=51(r\neq s)} M^i_{rs,t}} \tag{10}$$

$$\text{Intraregional } \textit{adjusted flows}: \ AM^i_{rr,t} = M^i_{rr,t} \times \frac{X\text{Intra}^i_{r,t}}{M^i_{rr,t}} \tag{11}$$

b) Agricultural products

- As the SIS does not provide information on agriculture, fishing, and forestry products ($R1$), the *adjusted flows* for this sector ($AM_{rs,t}^{i=R1}$) were obtained from the harmonization of intra $+$ interregional flows $\left(M_{rs,t}^{i=R1} + M_{rr,t}^{i=R1} \right)$ to the "*domestic output (nonexported abroad) of agricultural products*" provided by the National Accounts ($DO_t^{i=R1}$).[14] It should be noted here that the quality of this adjustment is weaker than in the previous case (industries), as it does not control the output produced and sold on the national market by each region/province.

$$AM_{rs,t}^{i=R1} = \left(M_{rs,t}^{i=R1} + M_{rr,t}^{i=R1} \right) \times \frac{DO_t^{i=R1}}{\displaystyle\sum_{s=1}^{s=52} \left(M_{rs,t}^{i=R1} + M_{rr,t}^{i=R1} \right)} \tag{12}$$

Special Cases

a) *R2—Mining and refinery*

- The SIS information with respect to the *mining and refinery* ($R2$) industry is biased owing to the presence of the headquarters of the big multinational companies in certain regions, particularly in Madrid. This effect is not so apparent in other industries. As a result, SIS data have to be reestimated to eliminate this "*headquarter effect*." The procedure uses the regional shares of Regional Accounts in terms of the gross value added (GVA) of this industry ($i = R2$).

$$AO_{r,t}^{i=R2} = O_{r,t}^{i=R2} \times \frac{GVA_{r,t}^{i=R2}}{\displaystyle\sum_{r=1}^{r=52} GVA_{r,t}^{i=R2}} \tag{13}$$

The *adjusted interregional totals* are $\quad AX\text{Inter}_{r,t}^{i=R2} = AO_{r,t}^{i=R2} \times \dfrac{X\text{Inter}_{r,t}^{i=R2}}{O_{r,t}^{i=R2}}$ (14)

The *adjusted intraregional totals* are $AX\text{Intra}_{r,t}^{i=R2} = AO_{r,t}^{i=R2} \times \dfrac{X\text{Intra}_{r,t}^{i=R2}}{O_{r,t}^{i=R2}}$ (15)

- Next, the OD matrices of the products corresponding to this industry are harmonized.[15]

Interregional *adjusted flows* : $AM_{rs,t}^{i=R2} = M_{rs,t}^{i=R2} \times \dfrac{AX\text{Inter}_{r,t}^{i=R2}}{\displaystyle\sum_{s=1}^{s=51(r\neq s)} M_{rs,t}^{i=R2}}$ (16)

$$\text{Intraregional } \textit{adjusted flows}: AM_{rr,t}^{i=R2} = M_{rr,t}^{i=R2} \times \frac{AX\text{Intra}_{r,t}^{i=R2}}{M_{rr,t}^{i=R2}} \qquad (17)$$

b) *R*16—Production and distribution of electricity, gas and water
The OD matrices of these products are estimated using the following methodology:

- As in the previous case, the SIS information concerning this industry requires correction to eliminate the "*headquarter effect.*" The procedure applied is the same as the previous one.
- Unlike the case of other commodities, there are no OD matrices in tons and Euros at the regional and provincial level for this industry. Consequently, the allocation of $AX\text{Inter}_{r,t}^{i=R16}$ to bilateral flows needs to be based on an alternative procedure:
  a. For each year, an OD matrix of electricity flows in GW/h is obtained from the main Spanish operator (*Red Eléctrica de España*). The matrices are at the regional (NUTS-2) but not at the provincial level (NUTS-3).
  b. To estimate the OD flows at the provincial level, we have built vectors of production $ep_{r,t}$ and consumption $ec_{s,t}$ of electricity by years t and provinces r and s. Then, given the system for electricity distribution in Spain and the fact that interregional electricity exports are always based on proximity, a set of OD matrices are obtained using a biproportional RAS procedure (Allen 1975; Möhr et al. 1987) based on the interprovincial distance matrix as a prior. Thus, for each year, through a biproportion, we can obtain a matrix $F_{rs,t}^{i=R16}$ that is nearest to another matrix D_{rs} (distance between the provinces in km) but with the row sums equal to $ep_{r,t}$ and the column sums equal to $ec_{s,t}$.
  c. Next, the OD matrices for electricity at provincial level (NUTS-3) are harmonized according to the actual OD matrices at regional level (NUTS-2) and are then used to estimate the bilateral flows in coherence with the information obtained from the SIS:

$$\text{Interregional } \textit{adjusted flows}: \ AM_{rs,t}^{i=R16} = M_{rs,t}^{i=R16} \times \frac{AX\text{Inter}_{r,t}^{i=R16}}{\displaystyle\sum_{s=1}^{s=51(r\neq s)} M_{rs,t}^{i=R16}} \qquad (18)$$

$$\text{Intraregional } \textit{adjusted flows}: AM_{rr,t}^{i=R16} = M_{rr,t}^{i=R16} \times \frac{AX\text{Intra}_{r,t}^{i=R16}}{M_{rr,t}^{i=R16}} \qquad (19)$$

Table 4. Spatial Distribution of Trade (Intraregional, Interregional, and International) *All goods, R1–R16. Millions of Euros*

| Average 1995–2005 | Exports to | | | Imports from | | Balance | | Openness[a] |
	Intraregional (1)	Spain (2)	World (3)	Spain (4)	World (5)	Spain (6)=(2 − 4)	World (7)=(3 − 5)	Ratio (2+3+4+5)/(1+2+3)
Andalusia	15,836.3	19,832.1	9,786.4	20,318.4	10,560.5	−486.3	−774.1	133%
Aragon	4,642.2	10,933.3	5,487.0	13,329.1	5,049.2	−2,395.8	437.8	165%
Asturias	4,261.5	5,037.0	1,644.9	5,214.0	1,874.6	−176.9	−229.6	126%
Balearic Islands	2,027.7	735.5	868.3	4,353.7	1,483.3	−3,618.3	−615.0	205%
Canary Islands	3,300.0	1,657.2	718.0	8,019.7	3,252.9	−6,362.5	−2,534.8	240%
Cantabria	1,297.7	3,183.1	1,306.8	4,387.8	1,476.8	−1,204.7	−169.9	179%
Cast.a-Leon	9,318.3	15,450.4	7,241.0	17,409.4	7,189.9	−1,959.1	51.1	148%
Cast.a-La Mancha	3,523.5	12,373.4	1,844.8	12,200.6	3,189.9	172.8	−1,345.1	167%
Catalonia	37,817.7	42,616.6	31,229.9	25,254.1	44,801.7	17,362.5	−13,571.8	129%
Valencia	14,479.7	20,416.7	14,423.8	22,179.3	12,082.7	−1,762.6	2,341.1	140%
Extremadura	1,631.6	2,574.9	764.3	4,860.0	393.6	−2,285.1	370.7	173%
Galicia	7,246.2	12,061.3	7,618.8	8,229.1	8,116.0	3,832.2	−497.2	134%
Madrid	12,814.2	23,211.1	12,376.1	25,020.7	36,641.4	−1,809.6	−24,265.3	201%
Murcia	2,737.7	6,622.3	3,044.7	7,346.9	3,718.1	−724.5	−673.5	167%
Navarre	2,250.5	6,661.0	4,162.9	6,476.3	3,334.2	184.7	828.8	158%
Basque country	8,965.3	17,771.5	10,482.5	14,969.8	9,134.5	2,801.7	1,348.0	141%
La Rioja	754.9	2,930.6	781.1	3,377.3	597.7	−446.7	183.4	172%
Ceuta and Melilla	6.2	108.9	56.8	1,230.6	330.9	−1,121.7	−274.1	172%
Total	132,911.2	204,176.8	113,838.3	204,176.8	153,227.8	0.0	−39,389.5	1005%

Source: Interregional trade is obtained from the C-Intereg database; international trade from Spanish Customs (AEAT).
[a] Openness ratio: As trade does not include services, openness is calculated by the following expression instead of the commonly used one: $(X + M)/GDP$.

Data Analysis

In this section, we have analyzed the main results obtained, first using a descriptive approach and then by means of a simple econometric model based on the classical gravity model.

Descriptive Analysis: Identifying the Main Flows

As it can be expected, the level of domestic trade (intraregional + interregional) of goods in Spain during the period 1995–2005 is clearly higher than that for international trade (see table 4). This is true for both exports and imports, for the economy as a whole and for the majority of Spanish regions.

Regarding the balances obtained, only five regions demonstrated positive balances for interregional trade, when compared with seven for international trade. The highest interregional surplus is found in Catalonia, followed by Galicia, the Basque Country, Navarre, and Castilla-La Mancha. In some cases, the sign of the balance for the domestic market is observed to be the opposite of that for the foreign market. For example, while Catalonia, Galicia, and Castilla-La Mancha showed positive balances in the national market and deficits in the international one, the reverse was true for regions like Aragon, Valencia, Castilla-Leon, Extremadura, and La Rioja. The openness ratios are found to be much higher than the ratio obtained when only international trade is considered. The highest values have been observed in the smallest regions in terms of size, namely, Madrid, La Rioja, Canary Islands, Balearic Islands, and Cantabria.

Despite the dominance of intranational trade (intraregional + interregional trade), the evolution of international flows is more dynamic (see Figure 2): while international imports and exports increased by 169 percent and 122 percent, respectively, between 1995 and 2005, interregional and intraregional trade grew by just 82 percent and 80 percent, respectively. Consequently, although domestic trade still continues to account for the main bulk of activity in the Spanish economy, the process of globalization and EU integration appear to be gradually shifting the natural economic relations away from the regional neighbors toward foreign countries. Indeed, in some of the most important regions like Catalonia and Madrid, although interregional exports continue to be higher than their international counterparts, the opposite is true in the case of imports.

Moreover, consistent with the previous findings in the case of the United States and Japan (Jackson et al. 2006), figure 2 shows how the evolution of interregional trade is more dynamic (79 percent) than intraregional trade (73 percent). This result points to the increasing integration of Spanish regions and may also be related to the fragmentation process of the production chain (Freenstra 1998; Jones and Kierzkowski 2005) and the *hollowing out* process described in the literature (Hewings et al. 1998; Munroe and Hewings 1998, Hitomi et al. 2000).
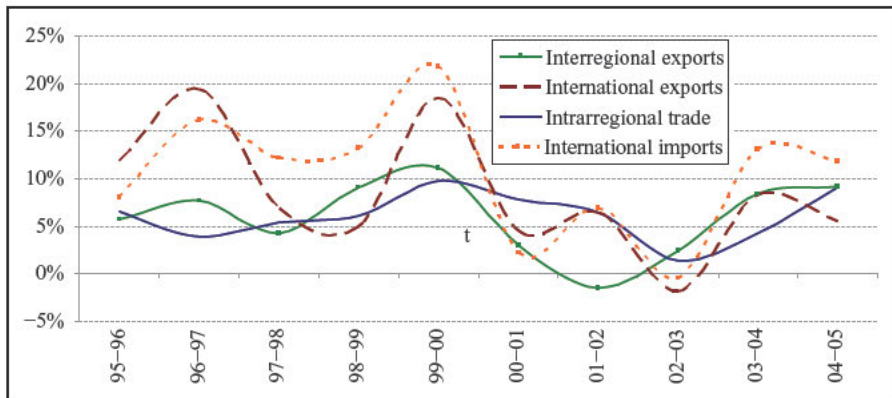
Figure 2. Evolution of international and domestic trade in goods (1995–2005). *Growth rates of trade in current prices.* Millions of Euros. All types of goods (R1–R16).

Using the relative share of each bilateral flow of total domestic trade, we identified the strongest trade flows during the period. Based on the average figures for the period 1995–2005 (see tables A1 and A2 in the Appendix), the most intense intra- and interregional flows are found to be usually associated with highly industrialized and heavily populated regions, such as Catalonia, Madrid, Valencia, and Andalusia. It is also interesting to note the existence of intense and stable commodity flows connecting regions located far away from each other (e.g., Catalonia–Madrid, Catalonia–Andalusia; figure 3).

In addition to the main intra- and interregional flows, we may want to identify those representing the highest levels of concentration of inflows or outflows for each region. These coefficients can help us identify the main suppliers and clients for each region. In this regard, it is particularly interesting to note that most of the biggest export and import shares are registered between contiguous regions. In figure 4, for example, we have shown the highest bilateral import shares in 2005, with the regions colored according to the regional axes of development usually considered in regional analysis of Spain. Furthermore, the stability of the main clients and suppliers for each region throughout the period indicates the presence of important interregional intersectoral linkages between them (see tables A1 and A2 in the Appendix). In our earlier researches (Pérez et al. 2009; Llano 2004a), we have already shown how these linkages act as transmission chains for interregional spillovers in terms of growth, employment, and productivity. Soon, we expect to show further evidence in this respect, based on this extended database.

Moreover, the tendency toward stable and intense flows between the strongest economies (figures 3 and 5 and tables A1 and A2 in the Appendix), together with
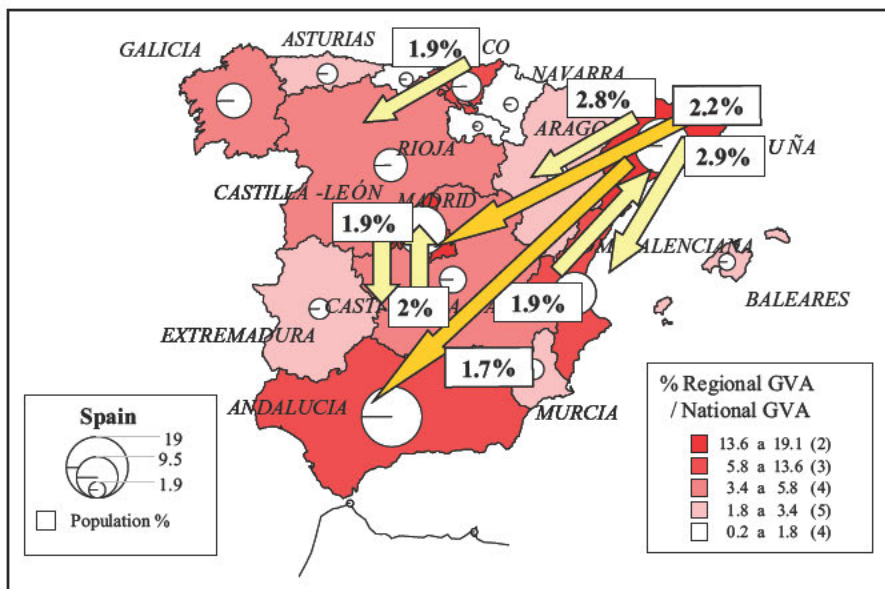
Figure 3. Strongest interregional flows in 2005. All goods, R1–R16. % of total interregional trade in Millions of Euros.

the high market shares found between contiguous regions (figure 4), point to the classical gravity model as the most appropriate candidate to explain the intensity of bilateral trade among the Spanish regions. This hypothesis has been tested in the next section.

Testing Interregional Trade Using the Classical Gravity Equation

The gravity equation has been widely and successfully used to analyze the intensity of bilateral trade between pairs of nations and regions. In its most basic formulation, bilateral trade between two geographic areas is directly proportional to their economic sizes and inversely proportional to the distance between them. More refined specifications of the model may include additional variables to capture other effects, such as accessibility, sectoral specialization, historical and cultural inertias, or network associations of all kinds, which would also condition the direction and intensity of the flows.

Based on various contributions from the literature on empirical applications of the gravity model to international and interregional trade (Anderson and van Wincoop 2003; Baldwin and Taglioni 2006; Helpman et al. 2007; Cheng and Wall 2005; Egger 2005; LeSage and Pace 2008a), equation (20) suggests a general
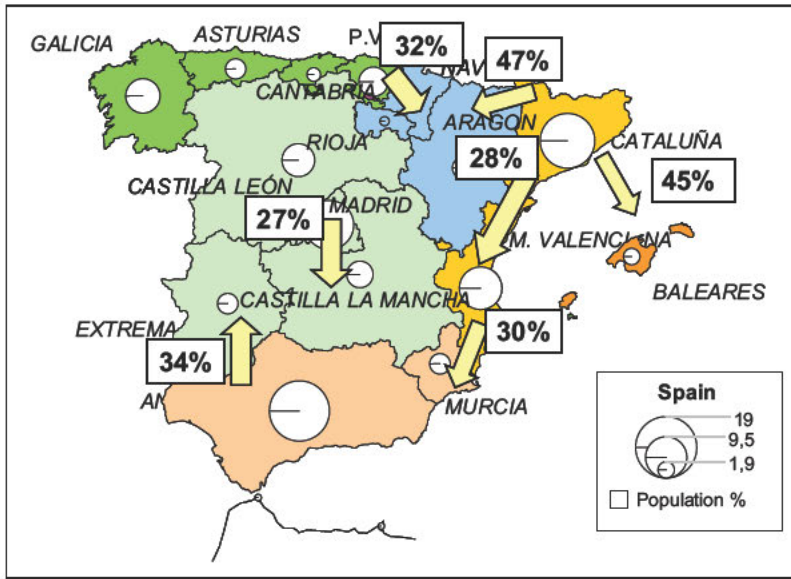
Figure 4. The largest import shares in the Spanish interregional trade. 2005. *All goods,* RI–RI6. % of an inflow on the total interregional imports of a region. Millions of Euros.

specification that includes a group of variables for the origin and destination regions, stocked respectively in matrix Xo_r (capacity of emission) and Xd_s (power of attraction).

$$\ln Y_{rs} = \beta_0 + \beta_1 \ln Xo_r + \beta_2 \ln Xd_s + \beta_3 \ln \text{dis}_{rs} + \varepsilon_{rs}, \tag{20}$$

where subscript r indicates an exporting Spanish region and s an importing Spanish region; Y_{rs} denotes the exports from region r to s, expressed in current Euros and dis_{rs} is the distance between region r and s. Table 5 presents all the variables tested: the variables in *boldface italics* are the ones finally included in the Xo_r and Xd_s groups of variables in the extended specifications of the model; variables in *pale gray* have been discarded or included in other variables. To analyze the general relations during the period (1995–2005), the model has been applied using average figures, both for trade flows and regressors. It should be noted here that the analysis is exclusively focused on total bilateral trade, leaving sector-specific flows for further analysis. All the variables are expressed in logs. The models were estimated by the ordinary least square (OLS) procedure, using public domain routines programmed in Matlab (LeSage and Pace 2009). In this regard, several articles have discussed the most appropriate method for estimating gravity models, both for cross-sectional data (Anderson and van Wincoop 2003; Baldwin and Taglioni 2006; Helpman et al. 2007) and panel data specifications (Cheng and Wall 2005; Egger 2005). As the goal of our model is to evaluate the coherence of the database in the light of the gravity

Table 5. Variables Included in the Gravity Model

Variable	Description
Pop	Population by regions NUTS 2 in each year. INE (National Institute of Statistics)
Gdp	GDP by regions in current prices for each year. INE
Income	Per capita income by region (GDP/POP). INE
Size	Size by regions NUTS 2 in km^2. INE
Export	International exports of goods by regions for each year. AEAT (Tax authority)
Import	International imports of goods by regions for each year. AEAT
Acc min	Minimum driving time by truck from one region to the nearest one. 2002. EUROSTAT (Regio DB)
Infras	This variable attempts to capture the relative situation of each region in terms of transport infrastructure endowment relative to its size: $$Infras_r = (Km^2\_motorway_r + Km^2\_rail_r)/Km^2\_surface_r$$
Remot	This variable is taken from Kimura et al. (2006): REMOTENESS, defined as the log of relative distance of region r in year y: $$Remot_r = \log \left[\frac{1}{\sum_r \frac{gdp_r}{\sum_r \frac{gdp}{dist_r}}} \right]$$
Distance	Physical distance in km between the capitals of provinces. INE
Kstock	Capital Stock by regions. For each year. FBBVA IVIE
Motorway	km^2 of highways by regions. EUROSTAT (Regio DB)
Rail	km^2 of railways by regions. EUROSTAT (Regio DB)
M2retail	Size of retailing centres by regions (sq. metres) in each year. La Caixa Bank (*Statistical Yearbook*)
Tripsintra	Number of trips by HGV trucks within the region in each year. EUROSTAT (Regio DB)
Hempl	Number of employees in High Technology industries by regions for each year. EUROSTAT (Regio DB)

model, we adopted the most basic method of estimation (OLS), which might serve as a baseline for further development (LeSage and Pace 2008; LeSage and Llano 2008a, 2008b).

Before analyzing the results obtained with the gravity model, it is interesting to observe the stability and persistence of the OD flows throughout the period (1995–2005). To this end, figure 5 plots the distance and the OD flows in millions of Euros between the Spanish regions (NUTS-2), both expressed in log-scale and ranked according to the 1995 OD flows. Note that if the ascending ranking of the flows during the period 1996–2005 had coincided with the one observed in 1995, then the shape of the resulting lines would have evolved in parallel. Although this is far from being the case, as it can be observed in figure 5, we can still appreciate a great stability in the flows. In fact, from 1995, the deviation of the most intense flows (upper
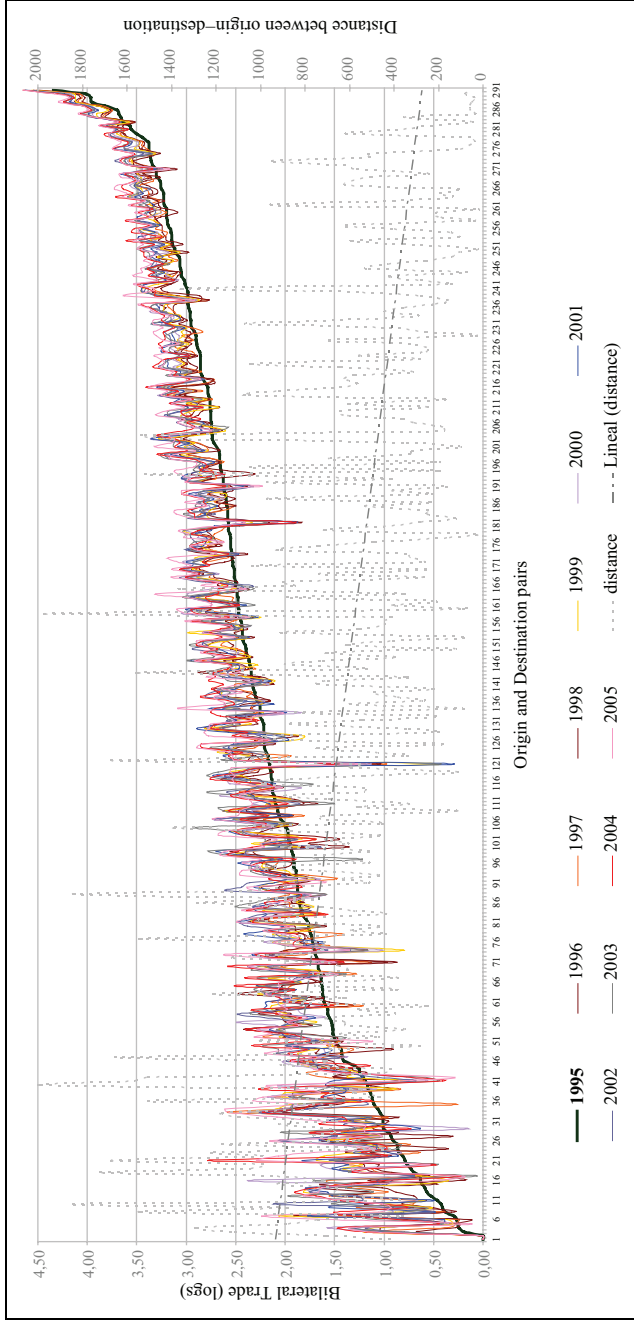
Figure 5. Ranked interregional bilateral flows for the period 1995–2005 and distance (km). Log scale. All OD flows and distance are ranked according to the 1995 flow order.

right side of the figure) during subsequent years has been much lower than that in the weakest OD flows (bottom left side of the figure). This tendency indicates the stability of the intensity of the flows and the persistence of the OD structure throughout the period. Furthermore, when the distance between the OD pairs is also ranked according to the 1995 ranking, we can obtain an unstable distance curve that captures the presence of strong flows between noncontiguous regions. In spite of the irregular shape of the distance curve, the associated tendency line clearly shows a negative slope. Consequently, we may expect that this negative relation between the intensity of trade and distance between the producer and consumer will be captured with a negative sign by the gravity model.

Next, we have focused on the results obtained using the basic specification of the model. Based on table 6, interregional flows in logs are relatively well explained ($R^2 = .89$) by distance as well as origin and destination gross domestic product (GDP). The variables are highly significant and their signs coincide with the expected ones according to previous studies in the literature (positive in the case of GDP and negative in the case of distance).

By taking the most basic specification as a baseline, we tested nine more specifications, including some of the variables used in previous studies focusing on international and interregional trade of goods and services. In model 2, following Kyriacos (2006), the basic specification is completed with the size of the origins and destinations in terms of *population*. The results obtained show no significant relation between the intensity of the bilateral flows and the population of the origin and destination regions.

In model 3, the population is substituted by the *per capita income* of the origin and destination regions (O income and D income). Here, the significance of this variable is higher for the origin than for the destination. Furthermore, the different signs obtained point to an inverse relationship between the intensity of the flows and the relative wealth of the origins and destinations, given that the flow intensity increases with the prosperity of the origins and decreases with that of the destinations.

When income is substituted by the actual *size* of each pair of regions (model 4), only the size of the destination regions appears to condition the flow intensity.

Models 5–7 include some new variables that attempt to capture the relation between the intensity of bilateral flows, accessibility, and transport infrastructure endowment. The variables included in model 5 (O acc min, D acc min) consider the minimum time that a truck spends in traveling from each region to the nearest one. According to the results obtained, the intensity of the flows increases when the origins are more accessible, whereas the accessibility of destinations does not appear to have a significant influence. In models 6 and 7, the accessibility variables are substituted by a measure of relative infrastructure endowment (O infras; D infras) or the *remoteness index* used in previous works analyzing international trade (Kimura and Lee 2006). In the case of model 6, although only the destination variable is significant, it is important to highlight the negative relations obtained,

both for origins and destinations, between infrastructure endowment and the intensity of the bilateral flow. This somewhat surprising result can be owing to the definition of the variable in relative terms, which might tend to reduce the relative importance of such infrastructures in certain larger regions (such as Catalonia, Valencia, the Basque Country, or Andalusia) that are also the main origin of interregional trade flows. However, a relatively high endowment of infrastructures in small regions may also coincide with strong interregional inflows (i.e., Madrid, Navarre, La Rioja, Murcia, etc.). In the case of model 7, the relation between the flow intensity and the remoteness index is also negative and significant, both for origins and destinations. In this case, the results obtained match those that can be expected, as the intensity of the flows decreases when the origin and destination regions are located far from the regions with the largest GDPs. In contrast, the main drawback of this specification is the reduction in the relevance of O gdp and D gdp, which seems to compete with the *remoteness* variables.

In model 8, we have explored the relation between the intensity of interregional bilateral trade and the importance of the origin and destination regions in terms of international goods trade. The results obtained for these four new variables (O export, O import, D export, D import) require a more detailed analysis. First, it should be stressed that all but O import are significant. Second, the signs obtained reveal interesting information with respect to the interrelation between national and foreign markets. For example, the negative coefficient obtained in the case of O imports and D imports suggests some kind of competition between the markets, as high interregional commodity outflows in origin and destination regions correlate negatively with high international imports arriving to these regions. Conversely, the positive sign obtained for O exports and D exports implies that strong interregional outflows and inflows are associated with high levels of international exports, both in origin and destination regions. This result, perhaps, captures the spatial dimension of the typical intermediate-final value-chain in the exporting sectors (Hitomi et al. 2000).

Finally, models 9 and 10 attempt to combine the best variables tested in the previous modes. In the case of model 9, only four of the eight variables included are significant (O gdp, O exports, D gdp, and Distance), whereas in model 10, all the variables are clearly significant and the explanatory power reaches its highest value ($R^2 = .908$). According to this final specification, the intensity of the bilateral flows is positively correlated with the origin and destination GDP and international exports in the origins but decreases with the remoteness of the destinations and the distance between them.

In relation to the distance variable, it is interesting to note that in all the specifications, Log distance is always significant, with negative coefficients between -1.06 and -1.13. This result indicates that, on an average, an increase of one unit in the distance between two points in Spain is associated with a decrease of one unit in the intensity of the flow.[16]

Conclusion

According to the first law of geography and more recent research concerning "border effect," one can expect to find higher levels of integration between regions within a country than between countries themselves. Paradoxically, the available information on economic interactions between countries is more extensive and detailed than between regions within the same country.

In this article, we have explored the main features of the C-Intereg database for Spanish interregional trade. Beginning with a brief review of the literature on interregional trade, we have described the methodology used in the estimation of the C-Intereg database, which is based on the use of transport flows and prices at the lowest level of disaggregation. Subsequently, we have provided an overview of the spatial structure of the Spanish interregional trade, which allows us to identify the most intense flows during the period of 1995–2005. Finally, the data set has been analyzed in further detail using ten alternative specifications of the classical gravity model.

In the immediate future, we hope to develop the C-Intereg project in all its dimensions (time, space, and sectors) to produce estimates for the upcoming years and for the new types of commodities and services. It is also expected that this database will eventually serve as a promising framework for the application of different techniques such as spatial interaction models or interregional IO approaches (LeSage and Llano 2008a, 2008b; Requena and Llano 2010; Ghemawat, Llano, and Requena 2010; Hewings et al. 2008; Sonis, Hewings, and Llano 2007; Pérez et al. 2009; Artal, Llano, and Requena 2009, Artal et al. 2010).

Notes

1. McCallum (1995), Helliwell (1996, 1998), Anderson and van Wincoop (2003), Freenstra 2002, Okubo (2004), and Gil et al. (2005).

2. For the sake of simplicity, although we used the general concept of "regions," the meth odology was applied to the fifty two Spanish provinces (Nuts 3) and then aggregated to the eighteen Spanish regions (Nuts 2). In addition, the method relies, at each successive step, on the highest level of disaggregation for each transport mode.

3. It should be noted that, in some cases, IMF do not correspond to economic transactions but to the capillary distribution of commodities stocked and distributed within each municipal ity for final consumption. In this respect, the elimination of IMF may be an advantage rather than a drawback, because it helps to prevent the double counting of intraregional flows. In contrast, the elimination of the intramunicipal flows (IMF) will tend to decrease the intra provincial trade, and therefore, the border effect at this spatial level.

4. Due to the absence of information on the specific commodities transported in the contain ers, and considering that the elimination of this sub mode will lead to a decrease of 30 percent of flows by railway, we have suggested the methodology described above. Based on the RENFE's qualitative information, we were informed that some commodities such as raw materials, non transformed agricultural products, energetic products, minerals, and so on, are not transported in containers. From this information we inferred that, for each origin, the containers will contain the same (j) categories that were able to be trans ported in containers and were present in the production and exporting structure by railway of each exporting province. This assumption is based on the idea that if a province is spe cialized in producing and exporting commodities in "complete wagon," similar firms within this province would deliver them in containers, with the limitation always on the nature of the product and the possibility to be transported in this mode.

5. Puertos del Estado: http://www.puertos.es/es/estadisticas/index.html.

6. The problem could also be described as the minimization of the difference between the total sum of bilateral flows in each new OD matrix and the old ones, subject to the same marginal constraints (Allen and Gossling 1975; Dijkman and Burgess 1994).

7. At this point, it is helpful to consider the distribution structure of oil products. In Spain, most of the refineries are located in the coast, near the big ports where the imported crude is unloaded and refined. There is just one refinery located in an inner region (Castilla La Mancha). Some refined products are pumped through the pipes to the main points of con sumption, where they are stocked in tanks and redistributed by trucks and trains. Trucks and trains can also departure from the refineries themselves to serve close markets. Based on this structure, we consider that pipe flows should be consider as "transit flows" from the refineries to the tanks. However, the pipe flows serve to reallocate other flows of refined oil products that are apparently delivered from regions with no refineries.

8. As exports with origin in regions without refinery become imports from the regions with refinery that served them, the transformed OD matrix $FR^{i}_{rs,t}$ is denoted as $FR^{i'}_{rs,t}$.

9. A 1993 OD matrix calculated by THEMA Consulting Group and the Ministry of Economic Development used data from CLH (Compañia Logística de Hidrocarburos, S.A. the main Spanish oil distributor). Owing to the liberalization of the oil distribution sector in Spain, it has not been possible to obtain more recent information on OD flows after 1993.

10. In Spain, most of the international trade with non European countries involves shipping as the transport mode. Trade by road tends to use international freight services, which are not included in our database.

11. This threshold is based on the information available with respect to the importance of trucks/railways for the interconnection of ports and "production" and "consumption" spots for international flows. As the comparability of the two databases is limited, the 15 percent threshold avoids a complete elimination of flows between each pair of provinces.

12. Note that the pipe (P) was just used to reallocate apparent exports by road and truck.

13. It should be noted that here, the OD matrices of refined oil products originally transported by trucks and trains and reallocated according to the pipe information are finally included in the original transport modes. Accordingly, there are no OD matrices of pipe.

14. Note that as National Accounts do not include the total output per year, in some years $DO_t^{i\ R1}$ must be deduced from the gross value added ($GVA_t^{i\ R1}$) and the international exports ($EXP_t^{i\ R1}$) in year t and the total output of the agricultural sector ($GO_t^{i\ R1}$) published in the last input output table ($t = 2000$): $DO_t^{i\ R1} = GVA_t^{i\ R1} \times \frac{GO_{t=2000}^{i=R1}}{GVA_{t=2000}^{i=R1}} \quad EXP_t^{i\ R1}$

15. It should be noted that we have used $s = 51$, because there are fifty two provinces. Additionally, it should be taken into account that interregional flows of some refined products moved by road and railway are subsequently reallocated according to the interprovincial pipe network.

16. Note that although trade flows are in Euros and distance in kilometers, all variables in the model are included in the logs.

Appendix

Table A1. Ranking of the Main Intraregional Flows (% Over Intraregional + Interregional Trade). *All Goods, R1–R16. Millions of Euros*

	1995		2000		2005	
	Origin = Destination	%	Origin = Destination	%	Origin = Destination	%
1	Catalonia	11.63	Catalonia	11.13	Catalonia	10.12
2	Andalusia	4.75	Andalusia	4.18	Andalusia	4.86
3	Valencia	4.23	Valencia	4.12	Valencia	4.43
4	Madrid	4.03	Madrid	3.89	Madrid	3.45
5	Castilla León	3.07	Castilla León	2.74	Basque Country	2.78
6	Basque Country	2.30	Basque Country	2.55	Castilla León	2.62
7	Galicia	2.16	Galicia	2.32	Galicia	2.18
8	Aragon	1.44	Asturias	1.34	Aragon	1.34
9	Asturias	1.32	Aragon	1.30	Castilla La Mancha	1.14
10	Castilla La Mancha	1.02	Castilla La Mancha	0.89	Asturias	1.11
	INTRA + INTER	263,396	INTRA + INTER	349,310	INTRA + INTER	442,272

Source: Own compilation based on C-Intereg database.

Table A2. Ranking of the Main Interregional Flows (% of Total Interregional Trade). All Goods, R1–R16. Millions of Euros

	1995			2000			2005		
	Origin	Destination	%	Origin	Destination	%	Origin	Destination	%
1	Cataluña	Valencia	4.03	Cataluña	Valencia	3.88	Cataluña	Valencia	2.94
2	Cataluña	Madrid	3.11	Cataluña	Aragon	3.00	Cataluña	Aragon	2.80
3	Cataluña	Aragon	3.11	Cataluña	Madrid	2.62	Cataluña	Madrid	2.21
4	Valencia	Cataluña	2.67	Valencia	Cataluña	2.44	Cast a-La Mancha	Madrid	2.02
5	Cataluña	Andalucia	2.35	Cast a-La Mancha	Madrid	1.81	Valencia	Cataluña	1.96
6	Madrid	Andalucia	1.87	Cataluña	Andalucia	1.81	C. Madrid	Cast a-La Mancha	1.95
7	Cataluña	Basque Country	1.84	Cataluña	Cast a-León	1.73	Cataluña	Cast a-León	1.82
8	Valencia	Andalucia	1.77	Madrid	Andalucia	1.71	Aragon	Cataluña	1.74
9	Cast a-La Mancha	Madrid	1.76	Aragon	Cataluña	1.69	Cataluña	Andalucia	1.73
10	Madrid	Cataluña	1.71	Basque Country	Cast a-León	1.67	Andalucia	Madrid	1.67
	Total Inter		149,597	Total Inter		215,335	Total Inter		272,062

Source: Own compilation based on C-ntereg database.

References

Alcaide, J., and P. Alcalde. 2005. Balance Económico Regional (Autonomías y Provincias). Años 2000 y 2004. Fundación de las Cajas de Ahorros Confederadas (FUNCAS). Gabinete de Estadística Regional. Madrid, Diciembre 2005.

Allen, R. I. G., and W. F. Gossling, (eds.). 1975. *Estimating and projecting input output coef ficients*. Input Output Publishing Company, London.

Anderson, J. E., and E. van Wincoop. 2003. Gravity with gravitas: A solution to the border puzzle. *American Economic Review* 93:170 92.

Artal, A., C. Llano, and F. Requena. 2009. Factor productivity differences and missing trade problems in a regional HOV model. *Papers in Regional Science.* doi: 10.1111/j.1435 5957.2009.00266.x.

———. 2010. The factor content of bilateral trade: Evidence from Spanish regions. *Journal of Applied Economics.*(IN PRESS).

Baldwin, R., and D. Taglioni. 2006. Gravity for dummies and dummies for gravity equations. NBER. Working Paper 12516. http://www.nber.org/papers/w12516

Batten, D. F. 1983. *Spatial analysis of interacting economics*. The Netherlands: Kluwer Nijh off Publishing.

Benvenuti, S., D. Martellato, and C. Raffaelli. 1995. INTEREG: A twenty region Input Output model for Italy. *Economic System Research* 7(2):101 16 AN: 0367929.

Boyce, D. E., and H. Bar Gera. 2003. Validation of multiclass urban travel forecasting models combining origin destination, mode and route choices. *Journal Regional Science* 43:517 40.

Bröcker, J. 1998. Spatial effects of Transeuropean networks: Preliminary results from a spa tial computable general equilibrium analysis. Diskussionsbeiträge aus dem Institut für Wirtschaft und Verkehr der Technischen Universität Dresden 4/98, Dresden: TU Dresden.

Bröcker, J., and N. Schneekloth. 2006. European transport policy and cohesion: An assess ment by CGE analysis. *Scienze Regionali* 5:47 70.

Cesario, F. J. 1975. Linear and nonlinear regression model of spatial interaction. *Economic Geography* 51:69 77.

———. 1976. Alternative models for spatial choice. *Economic Geography* 52:363 73.

Cheng, I. H., and J. H. Wall. 2005. Controlling for heterogeneity in gravity models of trade and integration. *Federal Reserve Bank of St. Louis Review* 87:49 63.

Dijkman, H., and A. Burgess. 1994. Estimating origin destination matrices from incomplete data. *Traffic Engineering and Control*.

Egger, P. 2005. Alternative techniques for estimation of cross section gravity models. *Review of International Economics* 13:881 91.

European Economy. 1997. The European Union as a world trade partner, European Commis sion. Directorate General for Economic and Financial Affairs. No. 3.

———. 2005. The EU Economy: 2005 review. Rising international economic integration. Opportunities and Challenges. European Commission. Directorate General for Economic and Financial Affairs. No. 6.

Ferreira, R. 2008. *El Comercio Interregional Iberico*. Ed. Associaciào de Desenvolvimento Regional. Instituto Politécnico de Portalegre, Portalegre.

Freenstra, R. 1998. Integration trade and disintegration of production in global economy. *Journal of Economic Perspectives* 12 (Fall) 12:31 50.

Freenstra, Robert C. 2002. Border effects and the gravity equation: consistent methods for estimation. *Scottish Journal of Political Economy* 49:491 506.

Fylstra, D., L. Lasdon, J. Watson, and A. Waren. 1998. Design and use of the microsoft excel solver. *Interfaces* 28:29 55.

Ghemawat, P., C. Llano, and F. Requena. 2010. Competitiveness and interregional as well as international trade: The case of Catalonia. *International Journal of Industrial Organiza tion*. doi: 10.1016/j.ijindorg.2010.03.013.

Gil, S., R. Llorca, J. A. Martnez, and J. Oliver. 2005. The border effect in Spain. *The World Economy* 28:1617 31.

Haynes, K. E., and A. S. Fotheringham. 1984. *Gravity and spatial interaction models*. News bury Park, CA: SAGE.

Helliwell, J. F. 1996. Do national borders matter for Quebec's trade? *Canadian Journal of Economics* 29:507 22.

———. 1998. *How much do national borders matter?* Washington, DC: Brookings Institu tion Press, pp. xxv + 155.

Helpman, E., M. Melitz, and Y. Rubinstein. 2007. Estimating trade flows: Trading partners and trading volumes. NBER Working Paper 12927.

Hewings, G., C. Llano, and M. Sonis. 2008. Trade integration and sectoral specialisation: The evolution of the sectoral structure of the Spanish regions. Mimeo.

Hewings, G. J. D., M. Sonis, J. Guo, P. Israilevich, and G. Schindler. 1998. The hollowing out process in the Chicago economy, 1975 2015. *Geographical Analysis* 30:217 33.

Hitomi, K., Y. Okuyama, and G. J. D. Hewings. 2000. The role of interregional trade in gen erating change in the interregional Leontief inverse in Japan, 1980 1990. *Economic Sys tem Research* 12:515 37.

Isard, W. 1951. Interregional and regional input output analysis: A model of space economy. *Review of Economics and Statistics* 33:318 28.

———. 1953. Regional commodity balances and interregional commodity flows. *American Economic Review* 43:167 80.

Jackson, R., Walter Schwarm, and Yasuide Okuyama. 2005. An evaluation of interregional trade estimation methods. *Journal of Regional Analysis and Policy* 36:84 93.

Jackson, R., Walter Schwarm, Yasuhide Okuyama, and Samia Islam. 2006. A method for con structing commodity by industry flow matrices. *Annals of Regional Science* 40:909 20.

Japanese Government. 1974. Administrative Management Agency et al. The 1974 Input Out put Table o Japan. Tokyo, 1974.

Jones, R. W., and H. Kierzkowski. 2005. International fragmentation and the new economic geography. *North American Journal of Economics and Finance* 16:1 10.

Kim, T. J., D. E. Boyce, and G. J. D. Hewings. 1983. Combined input output and commodity flow models for interregional development planning: Insight from a Korean application. *Geogrical Analysis* 15:330 42.

Kimura, F., and H. H. Lee. 2006. The gravity equation in international trade in services. *Review of World Economics* 142:599 614.

Kockelman, K. M., L. Jin, Y. Zhao, and N. Ruiz Juri. 2005. Tracking land use, transport, and industrial production using random utility based multizonal input output models: Applica tions for Texas trade. *Journal of Transportation Geography* 13:275 86.

Kyriacos, A. 2006. Are there differences across countries regarding the effect of currency unions on trade? Evidence from EMU. *Journal of Common Market Studies* 44:17 27.

Lasdon, L. S., and S. Smith. 1992. Solving large space nonlinear programs using GRG. *ORSA Journal on Computing* 4:2 15.

LeSage, J. P. and R. K. Pace. 2009. Introduction to spatial econometrics, CRC, Press/Taylor & Francis Group.

LeSage, J. P., and C. Llano. 2008a. A spatial interaction model with spatially structured origin and destination effects. *Journal of Regional Science*. (IN PRESS).

. 2008b. Forecasting spatially dependent origin and destination commodity flows. Paper presented at the South Regional Science Association, Washington, DC.

LeSage, J. P., and R. K. Pace. 2008. Spatial econometric modeling of origin destination flows. *Journal of Regional Science* 48:941 67.

Liu, N. L., and P. Vilain. 2004. Estimating commodity inflows to a substate region using input output data: Commodity flow survey accuracy tests. *Journal of Transportation Sta tistics* 7:23 37.

Llano, C. 2001. Economía espacial y sectorial: el comercio interregional en el marco input output. Doctoral Thesis.

. 2004a. Economía espacial y sectorial: el comercio interregional en el contexto de un modelo Multirregional para la economía española. Instituto de Estudios Fiscales. Minis terio de Economía y Hacienda. Colección Investigaciones: N° 1/2004.

. 2004b. The interregional trade in the context of a multirregional input output model for Spain. *Estudios de Economía Aplicada* 22:1 34.

McCallum, J. 1995. National borders matter: Canadian U.S. regional trade patterns. *American Economic Review* 85:615 23.

McFadden, D. 1981. Econometric models of probabilistic choice. In *Structural analysis of discrete data with econometric application*, eds. C. F. Mansky and D. McFadden. Cambridge, MA: MIT Press.

Miller, R., and P. Blair. 1985. *Input output analysis: Foundation and extensions*. Englewood Cliffs, NJ: Prentice Hall.

Möhr, M., W. H. Crown, and K. R. Polenske. 1987. A linear programming approach to solving infeasible RAS problems Texto completo disponible; *Journal of Regional Science* 27(4): 587 603.

Moses, L. N. 1955. The stability of interregional trading patterns and input output analysis. *American Economic Review* 45:803 32.

Munroe, D., and G. Hewings. 1998. The role of intraindustry trade in interregional trade in the Midwest of the US. Discussion Paper. REAL. University of Illinois, Urbana Champaign.

Nijkamp, P., A. Reggiani, and Wai Fai Tsang. 2004. Comparative modelling of interregional transport flows: Applications to multimodal European freight transport. *European Journal of Operational Research* 155:584 602.

Okubo, T. 2004. The border effect in the Japanese market: A gravity model analysis. *The Japanese and International Economies* 18:111.

Oliver, J. 1997. en La Balança de Pagaments de Catalunya: Una aproximació als fluxos econòmics amb la resta dÉspanya i l'estranger (1993 1994). Institu dÉstudis Autonòmics. Generalitat de Catalunya.

Oliver, J., J. Luria, A. Roca, and J. Pérez. 2003. en La apertura exterior de las regiones en España: Evolución del comercio interregional e internacional de las Comunidades Autón omas. 1995 1998. Institu dÉstudis Autonòmics. Generalitat de Catalunya. Ed Tirant lo Blanch. Valencia.

Oosterhaven, J. 1984. A family of square and rectangular interregional input output tables and models. *Regional Science & Urban Economics* 14:565 82.

Parellada, M. 1980. Los flujos comerciales entre Cataluña y el Resto de España (1975). *Revista Económica de Banca Catalana* 58:1 11.

Pérez, J., M. Dones, and C. Llano. 2009. An interregional impact analysis of the EU Structural Funds in Spain (1995 1999). *Papers in Regional Science* 88:509 30.

Polenske, K. R. 1980. *The U.S. multiregional input output accounts and model*. Lexington, MA, USA: D.C. Heath and Company.

Pulido, A., A. López, and C. Llano. 2000. La Balanza Comercial de la Comunidad de Madrid en sus relacionaes con el Resto de España (1995 1998). en Mella J. M. y Sanz B., La Balanza de Pagos de la Comunidad de Madrid (1995 1998). Biblioteca Civitas de Econ omía y Empresa. Colección Economía.

Requena, F., and C. Llano. 2010. The border effects in Spain: An industry level analysis. *Empirica*. doi: 10.1007/s10663 010 9123 6.

Round, J. I. 1995. A SAM for Europe: Social accounts at the regional level revisited. In *Social and demographic accounting*, eds. G. J. D. Hewings and M. Madden, 15 40. Cambridge, UK: Cambridge University Press.

Roy, J. R. and J. C. Thill. 2004. Spatial interaction modelling. *Papers in Regional Science* 83:339 61.

Ruiz, N., and K. Kockelman. 2006. Evaluation of the Trans Texas corridor proposal: Appli cation and enhancements of the random utility based multiregional input output model. *Journal of Transportation Engineering* 132:531 9.

Sen, A., and T. E. Smith. 1995. *Gravity models of spatial interaction behavior*. Heidelberg: Springer Verlag.

Smith, S., and L. Lasdon. 1992. Solving large sparse nonlinear programs using GRG. *ORSA Journal on Computing* 4:2 15.

Sonis, M., G. Hewings, and C. Llano. 2007. Trade integration and sectoral specialisation: The evolution of the sectoral structure of the Spanish regions. Working Paper REAL, August 2007.

Stone, R. 1961. Social accounts at the regional level. In *Regional Economic Planning*, eds. Walter Isard and J. H. Cumberland, 263 269. Paris: Organization for European Economic Cooperation.

Van der Linden, J. A., and J. Oosterhaven. 1995. European community intercountry input output relations: Construction method and main results for 1965 1985. *Economic System Research* 7:249 69.

Wilson, A. G. 1970a. Interregional commodity flows: Entropy maximizing approaches. *Geo graphical Analysis* 2:255 82.

    . 1970b. *Entropy in urban and regional modelling*. London: Pion Limited.

    . 1973. Further developments of entropy maximizing transport models. *Transportation and Planning Techniques* 1:183 93.