Can Business and Social Networks Explain the Border Effect Puzzle?*

Pierre-Philippe Combes[†] Miren Lafourcade[‡] Thierry Mayer[§]

October 7, 2002

Abstract

McCallum (1995) shows in an influential contribution that, even when controlling for the impact of bilateral distance and the region size, borders sharply reduce trade volumes between countries. We use in this paper data on bilateral trade flows between 94 French regions, for 10 industries and 2 years (1978 and 1993) to study the magnitude and variations over time of trade impediments, both distance-related and (administrative) border-related. We focus on assessing the role that business and social networks can play in shaping trade patterns and explaining the border effect puzzle.

Using a structural econometric approach, we show that intra-national administrative borders significantly affect trade patterns inside France. The impact is of the same order of magnitude as in Wolf (2000) for trade inside the United States. We show that more than 60% of these (puzzling) intra-national border effects can be explained by the composition of local labour force in terms of birth place (social networks) and by inter-plants connections (business networks). In addition, controlling for these network effects reduces the impact of transport cost on trade flows by a comparable factor. Thus, business and social networks that help reduce informational trade barriers are shown to be strong determinants of trade patterns and to explain a large part of the border puzzle.

JEL classification: F12, F15
Keywords: Gravity, border effects, networks.

*We thank Jean-Eric Thomas for having kindly made the trade flow data available to us. We are also grateful to Johannes Bröcker, Harry Flam, Laurent Gobillon, Keith Head and to seminar participants (ERWIT 2002, HWWA Workshop on border regions) for fruitful discussions. Part of this work has been performed when the second author was carrying out her PhD partly within the French Ministry of Transports Economic Department (SES-DAEI). The hospitality of this institution as well as financial support from SES-DAEI and from NATO (Combes' advanced fellowship grant) are gratefully acknowledged.

[†]CERAS-ENPC and Boston University. CNRS researcher also affiliated to CEPR (combes@enpc.fr, http://www.enpc.fr /ceras/combes/).

[‡]Université d'Evry Val d'Essonne (IUT - GLT Department) and CERAS-ENPC (lafourca@enpc.fr, http://www.enpc.fr/ceras/lafourcade/).

[§]Corresponding author: TEAM-Université de Paris I, 106-112, Bd de l'hôpital, 75647 Paris CEDEX 13, France. Also affiliated to CEPII, CERAS-ENPC and CEPR (tmayer@univ-paris1.fr, http://team.univ-paris1.fr/trombi/mayer/thierry.htm).

1 Introduction

It is one of the most widely accepted finding in international economics that distance matters in shaping the volumes of bilateral trade between nations. Trade flows fall with distance as shown by the myriad of papers using the gravity model.

A more recent finding, initiated by the work of McCallum (1995) is that, *in addition to the impact of distance*, borders seem to sharply reduce trade: For equal sizes and distances, regions trade much more between themselves if they are not separated by a national border. The initial work focuses on trade of Canadian provinces and the magnitude of that "border effect" can be summarized as following: Intra-national trade exceed international trade by a factor of about 20 in 1988 for given bilateral distance and size of regions. Several studies inspired by this first paper replicate the exercise for other countries and other periods (Wei, 1996, Helliwell, 1996 and 1997, Nitsch, 2000, Head and Mayer, 2000). The effect was found to be quite comparable across samples and always surprisingly large, so large that Obstfeld and Rogoff (2000) refer to the border effect as one of the "six major puzzles in international macroeconomics". Explaining this puzzle is now an important question in the research agenda as the estimated border effects in the literature represent a challenge to our view about the current level of trade integration.[1] The purpose of this paper is to contribute to the search of possible reasons explaining why borders still matter so much in trade.

The reasons why distance matters in international trade can be related to transaction costs, consisting mainly of transport costs, incurred when shipping a good. Distance is also related to the time elapsed before delivering the good, which represents additional costs when the product is perishable in nature or looses value after a short period of time[2]. Distance between countries is also correlated with the strength of cultural and informational linkages between them and those links have been shown to be important in bilateral trade volumes (see Rauch, 2001, for a review on this topic).

The reasons why borders still matter even when controlling for distance is more problematic. The literature has, to date, focused on four major explanations:

1. The first explanation is technical. Borders might appear to matter in trade because the estimated equation is mis-specified and/or the covariates used are imprecisely measured or badly constructed. The model specification explanation has been recently investigated by Anderson and Van Wincoop (forthcoming), whose work shows that estimating structural parameters from the theoretical gravity equation can reduce border effects. Head and Mayer (2002) focus on how mismeasurement in distances can also inflate the estimated border effect and propose a theory-based distance variable that reduces the estimated impact of borders.

2. The second and perhaps most straightforward explanation has to do with protection. If the countries in the sample considered still have significant (and not controlled for) formal barriers to trade such as tariffs or non tariff barriers, then the impact of those trade impediments is going to show up as a negative effect of the border on trade flows. Wolf (1997 and 2000) was the first to provide an indirect empirical test for the validity of the trade barriers explanation. The idea is that if national border effects are related to trade barriers, then those border effects should vanish when considering trade between and within regions inside

[1] Wolf (2000) for instance states in his conclusion that "The next research challenge is to further explore the causes of home bias" (p.561).

[2] Hummels (2001) and Evans and Harrigan (2002) provide estimates of the importance of time as a trade barrier.

a country. For that purposes, he uses trade flows between and within American states where the "standard" trade barriers are absent. He finds that US states borders have an impact that is less important than for international trade but still not negligible, suggesting that there exists a minimal level of market fragmentation even within a nation as integrated as the United States.

3. Transaction costs due to the use of different currencies have recently been proposed as a plausible explanation. The seemingly robust (although controversial) finding of Rose (2000) that monetary unions would triple bilateral trade flows, provides a potential cause for trade border effects. The fact that nations are almost by definition monetary unions could explain the seemingly excessive trade taking place inside their borders. Parsley and Wei (2001) and Taglioni (2001) provide some empirical support for this hypothesis, showing that exchange rate volatility explains a significant part of the border effect.

4. A last possible explanation has to do with home bias in consumer or firm preferences, which would lead to the following plausible explanation to the impact of borders: People may have a higher valuation for the goods produced locally simply because they are more familiar with them or because of "chauvinism". This increases the demand for these goods and consequently the observed intra-regional flows. In addition to the effect of distance, this creates a significant discontinuous drop in the flows when they cross the border. These Armington (1969) type home biased preferences can be easily introduced in monopolistic competition models to derive a structural specification of the gravity theory including border effects.

We propose in this paper a different explanation of border effects based on informational trade barriers. A recent strand of the literature surveyed by Rauch (2001) suggests that *business and social networks operating across borders* might help to alleviate some information problems and thus promote trade. Informational barriers make it difficult both for consumers to obtain relevant information on the goods produced abroad, and for foreign producers to learn the tastes of consumers or find efficient local retailers. Both mechanisms increase the transaction costs, and therefore the price of foreign goods, which has a negative impact on trade flows. Rauch (2001) also details how the reciprocal knowledge of trade partners reduces costly "opportunism" in business, networks being substitutes for contract enforcement laws, since "establishment of a moral community and collective punishment of cheaters are not mutually exclusive mechanisms for discouraging opportunistic behavior" (p. 1182). Empirical evidence in this direction is provided through the trade-creating effect of migrations (Gould, 1994, Head and Ries, 1998, Wagner et al., 2002) and business groups operating across national borders (Rauch and Trindade, 2002). We propose in this paper that the same mechanisms can contribute to explain why borders matter in shaping regional trade flows. We argue that these effects are not totally proportional to distance but would reduce trade flows in a discrete way, at the crossing of the borders.[3] Including business and social networks variables is shown to indeed reduce border effects. Finally, and in contrast to purely technical or exogenous explanations of border effects, our argument has important policy implications. For instance, increased mobility of firms and people in Europe may have a strong positive impact on trade flows through the reduction of the discrete negative impact of borders they induce.

All proposed explanations for the puzzle might of course play simultaneously, and disentangling them is a crucial step towards understanding which explanations are truly relevant. Our

[3]Note that Rauch (2001) opens his survey with a paragraph stating how border effects in trade can be explained by informational barriers.

work proceeds by neutralizing some possible explanations.

First, in order to try and neutralize the "technical" explanations of border effects, our estimated specification is rigorously derived from a monopolistic competition model of trade allowing for home biased preferences. This strongly links our estimations to the theoretical predictions. Moreover, we pay particular attention to the measurement and specification of the transaction cost and we make a clear distinction between its transport and information components. Second, we follow Wolf (2000) and focus on trade within a country, which cancels any possibility of trade policy effects as well as the monetary union explanation.

More precisely, we study border effects for 94 French administrative regions ("départements") and use data on trade flows at the industry level (10 industries) and for two years (1978 and 1993). We also benefit from a precise measure of inter-regional transport costs. We can therefore study the integration of French market over time and analyse the border effects for different industries.

We then investigate whether the remaining estimated border effect within France can be related to business and social networks. To that purpose, trade flows between two regions are related to the number of people working in a region who were born in the other. These bilateral stocks of migrants within France capture *social networks*. As emphasized by Rauch (2001), a large number of migrants from another area tends to promote trade because they keep active linkages with their networks at "home": "Immigrants know the characteristics of many domestic buyers and sellers and carry this knowledge abroad" (p.1184). Another explanation relies on the fact that migrants bring (at least partly) their tastes with them: "...the impact on bilateral trade of immigrants [...] reflect[s] immigrant taste for goods from their countries of origin" (p.1185). Gould (1994) also underlines that "the development of trust through immigrant contacts can decrease the costs associated with negotiating trade contracts and ensuring their enforcement". Besides social networks, Rauch (2001) claims that "foreign direct investment by one or more members of a domestic business has the same effect [as the migrant effects]" (p.1185). He also details more subtle effects of barriers to entry and collusion inside business groups strongly affecting trade patterns. Thus, we also consider these *business networks* by including in our set of explanatory variables the number of plants of each region that have a plant belonging to the same group[4] in the other region.

Our results show that administrative borders within France do indeed have a negative impact on trade. Our baseline estimates show that trade is in 1993 around six times lower between two non-contiguous regions than inside a region, for given size and distance. This magnitude is of the same order as results by Wolf (2000) for the United States. We find that the impact of borders declines over time, thus matching the trend over more economic integration within France. Importantly, we show that a substantial part, higher than 60%, of the trade border effect can be explained by both social and business networks. Finally, an industry-level analysis allows us to assess the relative importance of the network effects across sectors.

Before proceeding with structural econometrics, we present a graphical representation for three different samples (US states, European countries, and French regions), which is maybe the clearest way to present the twin effects of distance and borders on trade. Let m_{ij} denote the imports of location i from location j, d_{ij} the distance between the two, and Y_i the GDP in location i. The gravity model is, in its simplest form,

$$m_{ij} = GY_iY_j\,(d_{ij})^{-\varepsilon}\,, \tag{1}$$

[4] A group has a larger definition than a firm. For instance, all plants of Peugeot and Citroën belong to the same group called PSA.

where ε is a positive parameter and G is a constant. A convenient way to represent equation (1) is to graph $\ln\left(\frac{m_{ij}}{Y_i Y_j}\right)$ against $\ln(d_{ij})$. The result for the three different samples is represented in Figure 1.

The top graph in this figure uses a sample very close to the one used by Wolf (2000), that is, bilateral trade flows between and within US states in 1997 (Wolf used the 1993 data). The estimated coefficient for the distance is $\varepsilon = 1.15$ ($R^2 = 0.61$). The second one uses the sample in Head and Mayer (2000), that is bilateral trade flows between and within European countries in 1993. The estimated coefficient for the distance is $\varepsilon = 1.42$ ($R^2 = 0.58$). Last, the third one uses the original sample of this paper, that is bilateral trade flows (in volume) between and within French départements in 1993. The estimated coefficient for the distance is $\varepsilon = 1.73$ ($R^2 = 0.62$). In each of those panels, the circles represent flows between locations, and triangles represent flows within locations. We immediately can see that "internal flows" are much higher than the gravity prediction represented by the continuous straight line: Even accounting for the fact that flows inside a geographical unit cover a much lower distance than flows across geographical units, internal trade observations are large positive outliers in the gravity equation, which econometrically translates into the border effect for the three samples.

The rest of the paper proceeds as follows. In section 2, a theoretical model of trade under monopolistic competition is presented, in which business and social network effects work through transaction costs and heterogeneous tastes. This yields a directly estimable equation that bears some strong links with the gravity equation but is augmented such as making all variables justified by the theoretical model. The data we use are described at the end of section 2. Section 3 presents the estimations results and section 4 concludes.

2 A Model of home bias in national trade

We describe in this section the theoretical underpinnings of the specification of border and network effects we use. This modelling is inspired by the widely used trade model of monopolistic competition *à la* Dixit-Stiglitz-Krugman (Dixit and Stiglitz, 1977, Krugman, 1980), slightly modified to account for home bias in the consumers' preferences.

2.1 Consumption

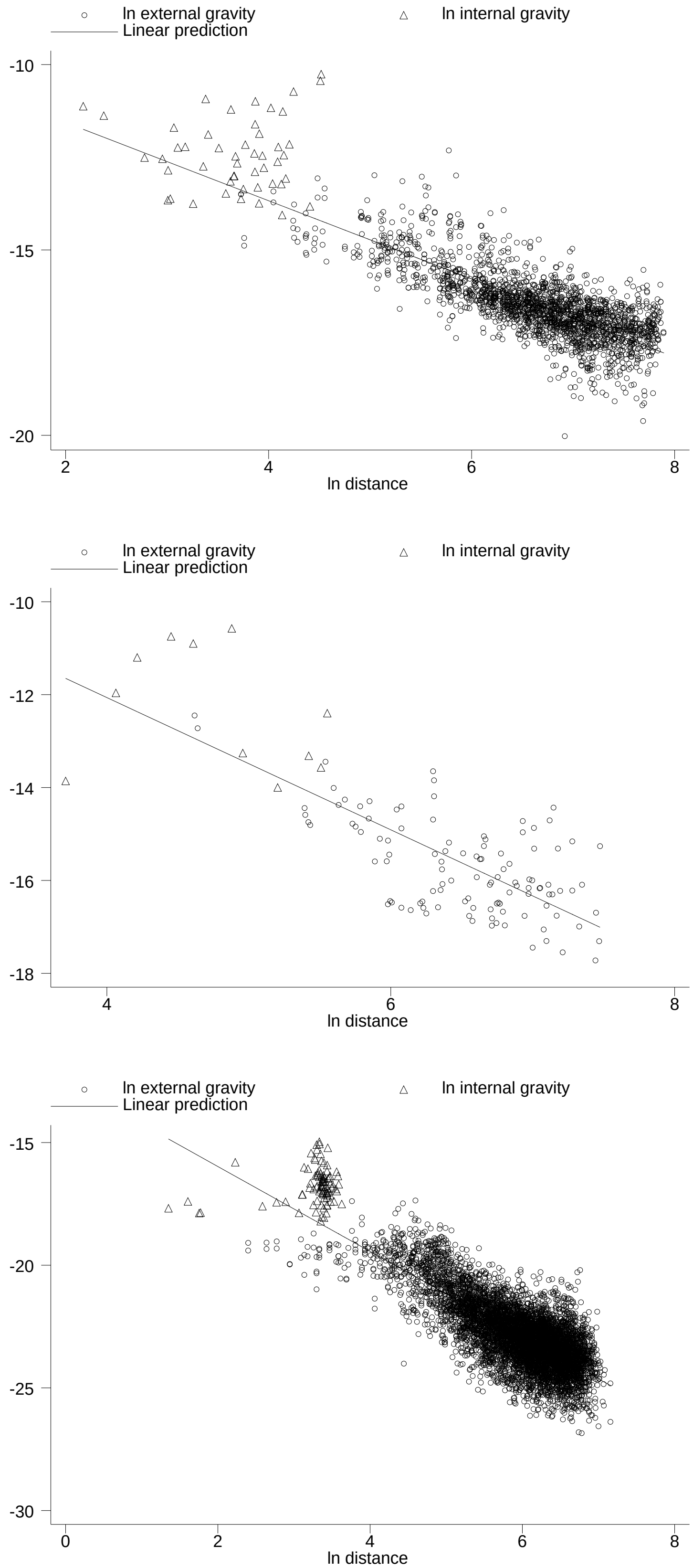
The representative consumer's utility in region i depends upon the consumption c_{ijh} of all varieties h produced in any region j. Varieties are differentiated with a constant elasticity of substitution (CES) but they do not enter symmetrically the utility function: A specific weight, a_{ij}, is attached to all varieties imported from region j. Let n_j denote the number of varieties produced in region j and N the total number of regions. The corresponding utility function is

$$U_i = \left(\sum_{j=1}^{N} \sum_{h=1}^{n_j} (a_{ij} c_{ijh})^{\frac{\sigma-1}{\sigma}} \right)^{\frac{\sigma}{\sigma-1}}, \tag{2}$$

where $\sigma > 1$ is the elasticity of substitution. Let m_{ij} denote the c.i.f. value of imports of region i from region j and p_{ij} the delivered price in region i of any variety produced in region j. Denoting by τ_{ij} the iceberg-type ad-valorem equivalent transaction cost between regions i and j and p_j the mill price in j, we have

$$p_{ij} = (1 + \tau_{ij})\, p_j. \tag{3}$$

Figure 1: Distance and Borders in Trade



It is then straightforward to obtain the following demand function, where $m_i = \sum_k m_{ik}$ is the whole expenditure of region i on the differentiated good varieties imported from all possible source regions (including i):[5]

$$m_{ij} = \frac{a_{ij}^{\sigma-1} n_j p_{ij}^{1-\sigma}}{\sum_k a_{ik}^{\sigma-1} n_k p_{ik}^{1-\sigma}} m_i. \tag{4}$$

The numerator of equation (4) links the bilateral flow to the size of the demand expressed by region i (m_i), the size of the considered industry in region j (n_j), the bilateral preference parameter (a_{ij}), the delivered price (p_{ij}), and the price index ($P_i \equiv \sum_k a_{ik}^{\sigma-1} n_k p_{ik}^{1-\sigma}$).

There are two major problems that remain to be solved in order to get an estimable equation. We must first deal with P_i which complicates estimation by introducing non linearity in unknown parameters. To do that, we use a convenient feature emphasized in Anderson *et al.* (1992) of CES demand functions, often called Independence of Irrelevant Alternatives (IIA)[6] due to the similarity with the logit model. In this type of demand functional form, the ratio of two bilateral trade flows to a same importing country depends only on the characteristics of the two corresponding exporting countries. Therefore it proves convenient to divide all bilateral trade flows by "imports from self".[7] Since our data correspond to trade volumes, we also use $m_{ij} = c_{ij} p_{ij}$ to obtain

$$\frac{c_{ij}}{c_{ii}} = \left(\frac{a_{ij}}{a_{ii}}\right)^{\sigma-1} \left(\frac{p_{ij}}{p_{ii}}\right)^{-\sigma} \left(\frac{n_j}{n_i}\right). \tag{5}$$

The second estimation problem relies on the fact that the number of varieties produced in region j, n_j, and the delivered prices, p_{ij}, are not observed. It is possible, however, to use the behavior of producers under monopolistic competition to get a correspondence with variables easier to observe, as regional production, wage, and transaction cost.

2.2 Production

As usual in this type of model, differentiation costs are considered sufficiently low to ensure that each variety is produced by a single firm. The production of each variety is subject to increasing returns to scale with a common technology among regions. Labor is the only input. Let f (g, respectively) denote the fixed (marginal, respectively) labor requirement for production, independent on the region. The needed quantity of labor in region j to produce a quantity q_j of a representative variety is therefore $l_j = gq_j + f$. If w_j denotes the wage rate in region j, the corresponding profit function is

$$\pi_j = p_j q_j - w_j(gq_j + f). \tag{6}$$

The Dixit-Stiglitz-Krugman model of monopolistic competition assumes that firms are too small to have a sizeable impact on the overall price index and on the regional income. This implies that the first order condition with respect to price gives a constant markup over marginal cost

[5]Note that, with a production function *à la* Ethier (1982), demand for inputs takes the same form, which is important as many industries we focus on are mainly producing inputs for other industries.

[6]See, however, Lai and Trefler (2002) or Anderson and van Wincoop (forthcoming) for different approaches to the same issue involving non linear estimation techniques.

[7]Head and Mayer (2000) and Eaton and Kortum (2002) also use this property of the CES function to get their estimable trade equation.

$$p_j = \frac{\sigma}{\sigma - 1} g w_j. \tag{7}$$

Consequently, all varieties produced in region j have the same mill price. It is then straightforward to determine the equilibrium output of each firm, resulting from the free entry of firms and the corresponding zero profit condition

$$q_j = \frac{f(\sigma - 1)}{g}. \tag{8}$$

Hence, since we assume that firms share a common technology, they all have the same equilibrium output in all regions.

Let v_j denote the value of the total production in region j, we get $v_j = n_j p_j q_j$. Therefore, using (7) and (8), we get

$$\frac{n_j}{n_i} = \frac{v_j w_i}{v_i w_j}. \tag{9}$$

By the definition of the delivered prices (3) and using the pricing rule (7), equation (5) can be finally written, using equation (9), as

$$\frac{c_{ij}}{c_{ii}} = \frac{v_j}{v_i} \left(\frac{a_{ij}}{a_{ii}}\right)^{\sigma-1} \left(\frac{w_j}{w_i}\right)^{-(\sigma+1)} \left(\frac{1+\tau_{ij}}{1+\tau_{ii}}\right)^{-\sigma}. \tag{10}$$

2.3 The estimable specifications

In order to obtain an estimable specification, the final steps involve specifying the transaction cost (τ_{ij}) and the preference structure of consumers (a_{ij}).

We consider two different elements in the transaction cost: Physical transport costs, T_{ij} and information costs, I_{ij}. According to Gould (1994) and Rauch (2001), these information costs are first due to the fact that firms are not perfectly aware of the consumers' preferences and devote resources to acquire this information. Second, producers are not perfectly connected to (or have not perfect information on) the trade intermediaries that locally sell the goods to the consumers (households or firms). For instance, they are not aware of the location of stores or are not connected with the local retailers. All of these increase the transaction cost. Hence, we assume

$$1 + \tau_{ij} = T_{ij} I_{ij}. \tag{11}$$

For the transport cost, we assume the following structure

$$T_{ij} = (1 + t_{ij})^{\delta} \exp(-\theta(1 + t_{ij})^2), \tag{12}$$

where t_{ij} is the measure of transport cost between i and j we use. It incorporates both distance and time related elements of this cost. Parameters δ and θ are expected to be positive. In this case, this cost function embodies a standard feature of increasing returns in transport activities: The cost associated with transport reduces equilibrium trade flows, but the marginal cost of shipping a good decreases with distance.

For the information cost, we assume

$$I_{ij} = (1 + s_{ij})^{-\alpha_I}(1 + s_{ji})^{-\beta_I}(1 + b_{ij})^{-\gamma_I}(1 + b_{ji})^{-\rho_I} \exp\left(\varphi_I A_{ij} - \psi_I C_{ij}\right). \qquad (13)$$

A_{ij} and C_{ij} are dummy variables set to 1 when i is different from j, and when i and j are different but correspond to contiguous regions, respectively. We assume that the best information from which producers benefit corresponds to the market where they are located. Next, we consider two different geographic levels in the transmission of information. Our hypothesis is that the informational transaction cost is lower inside a region than between two non contiguous regions (A_{ij}), but higher between those than between two contiguous regions (C_{ij}).

Following the terminology used by Rauch (2001), we refer to the effects of s_{ij} and s_{ji} as *social networks*, and to the effects of b_{ij} and b_{ji} as *business networks*. s_{ij} and s_{ji} contain information costs related to the difficulty of establishing social networks between i and j. b_{ij} and b_{ji} are the corresponding variables intended to capture information flowing between i and j through business networks. Thus, we assume that producers in region j obtain some information on region i market thanks to the people employed in j that were born in i. For instance, these workers know where are located the region i consumers or retailers. We assume that this effect is larger, the larger the number of such workers in the region, and we use the variable s_{ji} to reflect that. Conversely, the higher the number of people in region i who were born in region j, the higher the probability that they act as importers for the region j producers, or that they are still connected to them and transmit some information on region i markets. This is the reason why we also introduce s_{ij}. Another reason for introducing these variables is the "opportunism" argument developed by Rauch (2001) and presented in the introduction, which also works both ways. When people know each other and belong to the same network, there is less need for using strong and costly contracts to protect both trading partners, which may be true for the network of people born in the same region. Note that the specification chosen by Gould (1994) also implies the presence of both variables we use, even if he only considers s_{ji}, having no data on s_{ij}.

We also consider by b_{ij} and b_{ji} the links between plants belonging to the same group. Clearly, mechanisms of the same nature as for employees may apply. Plants of region j have more information on the markets of region i where plants belonging to the same group are located (b_{ji}), and, the higher the number of these plants, the higher the probability of gathering such information (b_{ij}). Opportunistic behaviors should moreover be reduced between plants belonging to the same group. According to these intuitions, parameters α_I, β_I, γ_I, ρ_I, φ_I, and ψ_I are expected to be positive.

Consumers are assumed to have both deterministic and stochastic elements in their preferences, a_{ij}. We assume systematic preferences for local goods (produced in the region of consumption), for the goods produced in a contiguous region, and for the goods produced in the region where the consumer is born. This last effect is assumed to be increasing with the s_{ij} variable. Conversely, the higher the number of employees in region j born in region i, s_{ji}, the higher the probability the firms in region j produce goods corresponding to region i consumers' preferences on which they have more information: This corresponds to a higher a_{ij}. Comparable arguments apply for the plants, the reason why we also introduce the b_{ij} and b_{ji} variables here. Last, the random component in the preferences is denoted e_{ij}. Thus, α_a, β_a, γ_a, ρ_a, φ_a, and ψ_a being parameters that are expected to be positive, we assume that

$$a_{ij} = (1 + s_{ij})^{\alpha_a}(1 + s_{ji})^{\beta_a}(1 + b_{ij})^{\gamma_a}(1 + b_{ji})^{\rho_a} \exp[e_{ij} - \varphi_a A_{ij} + \psi_a C_{ij}]. \qquad (14)$$

Note that even if these variables play on both the preferences and the information part of the transaction cost, their effects are fundamentally different in both cases. In the former case, they correspond to exogenous effects directly affecting the consumers' preferences. In the latter case, they correspond to endogenous demand effects working in equilibrium only through the delivered price that depends on the transaction cost.

Replacing in equation (10) the different specifications we assume for the transaction cost (equations 11 to 13) and the preferences (equation 14), and using the notations $x \equiv \sigma x_I + (\sigma - 1)x_a$, for $x = \alpha$, β, γ, ρ, φ, and ψ, we obtain what we call the *odds specification*

$$\begin{aligned}\ln\left(\frac{c_{ij}}{c_{ii}}\right) &= \phi\ln\left(\frac{v_j}{v_i}\right) - (\sigma+1)\ln\left(\frac{w_j}{w_i}\right) - \sigma\delta\ln\left(\frac{1+t_{ij}}{1+t_{ii}}\right) + \sigma\theta\left[(1+t_{ij})^2 - (1+t_{ii})^2\right] \\ &\quad +\alpha\ln\left(\frac{1+s_{ij}}{1+s_{ii}}\right) + \beta\ln\left(\frac{1+s_{ji}}{1+s_{ii}}\right) + \gamma\ln\left(\frac{1+b_{ij}}{1+b_{ii}}\right) + \rho\ln\left(\frac{1+b_{ji}}{1+b_{ii}}\right) \\ &\quad -\varphi A_{ij} + \psi C_{ij} + \epsilon_{ij}. \end{aligned} \tag{15}$$

Unfortunately, not all structural parameters are identifiable. The information effects in the transaction cost cannot be distinguished from the preferences effects in α, β, γ, ρ, φ, and ψ. Thus, we estimate the total effect of each network variable, but not the separate effects working through preferences or transaction costs. $\epsilon_{ij} = (\sigma - 1)(e_{ij} - e_{ii})$ implies that errors are not independently distributed. This correlation is accounted for in the estimation through a robust clustering procedure, allowing for residuals of observations of a same importing region to be correlated.

Finally, note that the theoretical framework predicts $\phi = 1$. ϕ is a parameter introduced in order to give additional flexibility in the estimations. The results are virtually unaffected by this change to the model.[8]

We also estimate an alternative specification of the theoretical model, following Head and Ries (2001), which goes one step further in using the IIA property of the CES. The inverse "friction" function is defined in volume as

$$\Phi_{ij} = \sqrt{\frac{c_{ij}}{c_{ii}}\frac{c_{ji}}{c_{jj}}}. \tag{16}$$

Using equation (15), we obtain what we call the *friction specification*:

$$\begin{aligned}\ln(\Phi_{ij}) &= -\sigma\delta\ln\left(\frac{1+t_{ij}}{\sqrt{(1+t_{ii})(1+t_{jj})}}\right) + \sigma\theta\left[(1+t_{ij})^2 - \frac{1}{2}(1+t_{ii})^2 - \frac{1}{2}(1+t_{jj})^2\right] \\ &\quad +(\alpha+\beta)\ln\left(\sqrt{\frac{(1+s_{ij})(1+s_{ji})}{(1+s_{ii})((1+s_{jj})}}\right) + (\gamma+\rho)\ln\left(\sqrt{\frac{(1+b_{ij})(1+b_{ji})}{(1+b_{ii})(1+b_{jj})}}\right) \\ &\quad -\varphi A_{ij} + \psi C_{ij} + \varepsilon_{ij}. \end{aligned} \tag{17}$$

The friction specification has the advantage of being compatible with the strict version of the model implying $\phi = 1$. Importantly, it does not require data on regional values of production (v_i) and wages (w_i), which is a noticeable advantage considering the measurement errors and missing values often found in those series. A drawback of this setting is that less parameters can be identified than with the odds specification (15). Only $\sigma\delta$, $\sigma\theta$, $\alpha + \beta$, $\gamma + \rho$, φ, and ψ

[8]Estimations under the constraint that $\phi = 1$ are available upon request.

are identifiable. On the other hand, border effects are still identifiable and hence they can be compared to those obtained with the odds specification. Such comparisons are also possible for the products $\sigma\delta$ and $\sigma\theta$ and for the sums of the social $(\alpha+\beta)$ and business $(\gamma+\rho)$ network effects. Again, the autocorrelation introduced by the fact that $\varepsilon_{ij} = \frac{1}{2}(\epsilon_{ij} + \epsilon_{ji})$ is taken into account in estimation.

Specifications (15) and (17) are the two estimated equations in section (3). In each case, we estimate two borders effects. The *total border effect* corresponds to $\exp(\varphi)$. It corresponds to the ratio of intra-regional trade over inter-regional trade for non contiguous regions. The *local border effect* is given by $\exp(\varphi-\psi)$. It corresponds to the ratio of intra-regional trade over inter-regional trade for contiguous regions.

2.4 Data

The data needed consist of trade flows, regional production and wages, transport costs and bilateral measures of social and business networks.

Trade flows between and inside regions available for 1978 and 1993 come from the French Ministry of Transports database on industrial commodity flows. The source and construction method of these data are comparable to the U.S. Commodity Flow Survey (CFS) used in Wolf (2000) and Anderson and van Wincoop (forthcoming) for instance. They are available at a very detailed industry level. However, the observation number being sometimes low for some industries, we either fully aggregate the flows over all industries, or distinguish between 10 different industry aggregates. This trade flow dataset suffers from the same imperfection as the CFS concerning break loading and the way it treats final destination when the latter is a foreign country, and reciprocally for imports. While both database try to sort out flows that are only in transit in a region, a large amount of shipments to and from major ports is admitted to be in reality transit shipments. As a consequence, the corresponding region appears to be an excessive source of flows to other regions compared to their real production (and reciprocally as a destination). Using local GDPs as a proxy for regional production, v_i in equation (15), therefore yields an underestimated coefficient on this variable, even if it does not fundamentally affect the other variable effects.[9] We therefore use a different method for calculating the regional production. It is computed as the sum of the flows departing from the region, including the own region internal flow.

The theoretical model requires the use of a measure of transport costs between and within French regions. Whereas most studies investigating trade determinants use distance as a proxy for those costs, we follow a recent trend in the literature that uses newly available data on actual transport costs (see for instance Hummels, 1999, Limão and Venables, 2001, or Combes and Lafourcade 2001). Our dataset for 1978 and 1993 provides the cost for a truck to join pairs of French regions. The calculation of this cost is based on the real road network. This is a generalized transport cost in the sense that it includes both a cost per kilometer (gas, tolls,...) depending on the road type and a time opportunity cost (drivers' wages, insurance,...). This measure therefore accounts for distance-related and time-related components of transport costs, the latter being shown to be very important by Hummels (2001). As an implication, our transport cost measure fluctuates significantly across time, -38% on average between 1978 and 1993. This contrasts with the physical distance often used in empirical trade studies. The reader may find more details on the way this cost is computed and some descriptive statistics on its components, variations across time and impact on the French regional development patterns in Combes and Lafourcade (2001 and 2002).

[9]Results using GDP are available upon request.

Since no intra-regional direct transport cost data exist for France, we estimate those. We first regress transport costs on real road distances and then apply estimated coefficients to internal distances in order to get the corresponding internal transport costs. The internal distance is obtained using a simple geographical approximation. Each region is approximated as a disk upon which all production concentrates at the center and consumers are uniformly distributed throughout a given proportion of the total area. We let this proportion vary in order to investigate the impact of different agglomeration patterns of consumers inside the regions. When the consumers are located uniformly on the total area, we get $d_{ii} = 0.376\sqrt{A} = \frac{2}{3}R$, where A is the region area and R the corresponding radius if the region were a disk. When the consumers are located uniformly on $\frac{1}{16}$ of the total area surrounding the center, which better suits the observed concentration of population in France, the internal distance formula is $d_{ii} = 0.094\sqrt{A} = \frac{1}{6}R$.

Production costs are approximated by the average wage per employee in the given year and region, which is computed using surveys from the French National Institute of Statistics (INSEE).

The social network variables correspond to the number of people working in the destination region who were born in the origin region (and the reverse). The business network variables correspond to the number of plants located in the destination region belonging to a group which has at least one other plant located in the origin region (and the reverse). The data needed for construction of those latter variables are available only for 1993. Both types of network variables are also computed using INSEE surveys.

Table 1 gives summary statistics for the data we use. The first line clearly shows that inter-regional trade flows are much smaller than intra-regional ones, the former being on average equal to less than 1% of the latter. Even the highest inter-regional trade flow represents 86% of the corresponding intra-regional flow only. Inter-regional transport costs are on average much higher (8.9 times) than intra-regional ones. This seems to be the most straightforward explanation for the differences in inter- and intra-regional trade flows. Indeed, the simple correlation between these two variables is also high, equal to -0.38 (see Table 2). On average, the number of migrants represents less than 2% of the number of people working in the region where they were born. There are about 10 extremely high figures (above one) for this variable, which all correspond to people who were born in a département of the Parisian area and still work in in a département this area. The difference between the inter-regional plant connections and the intra ones is lower, the ratio being around $\frac{1}{3}$. These large gaps are a first indication that networks could be important in explaining the level of trade between regions. This is confirmed by the strong positive correlations (see Table 2) between network variables and the relative trade flows, with a higher correlation for social networks.

Table 1: Summary Statistics

Variable	Mean	Sdt. Dev.	Min	Max
rel. flows, $\frac{f_{ij}}{f_{ii}}$	0.009	0.029	0	0.864
rel. transport costs, $\frac{t_{ij}}{t_{ii}}$	8.914	4.035	1.073	23.602
rel. migration from origin, $\frac{s_{ij}}{s_{ii}}$	0.013	0.043	0	1.154
rel. migration from destination, $\frac{s_{ji}}{s_{ii}}$	0.019	0.172	0	8.812
rel. plant connections from origin, $\frac{b_{ij}}{b_{ii}}$	0.277	0.151	0	0.833
rel. plant connections from destination, $\frac{b_{ji}}{b_{ii}}$	0.373	0.411	0	7.333

Note: Statistics are computed on the observations where $i \neq j$.

Table 2: Correlation matrix

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
rel. flows (1)	1	0.09*	0.01	-0.38*	0.49*	0.31*	0.10*	0.00
rel. production (2)		1	0.29*	0.06*	-0.10*	-0.03	-0.41*	-0.36*
rel. wages (3)			1	0.02	-0.15*	0.04*	-0.39*	-0.52*
rel. transport costs (4)				1	-0.28*	-0.12*	-0.07*	-0.03*
rel. migration from origin (5)					1	0.64*	0.28*	0.28*
rel. migration from destination (6)						1	0.08*	0.05*
rel. plant connections from origin (7)							1	0.73*
rel. plant connections from destination (8)								1

Note: * denotes significantly different from 0 at the 1%level.

Finally, we notice that social network variables are also highly correlated with the relative transport cost. Figure 2 helps understanding this correlation and, more generally, the spatial patterns of network variables. The left-hand side maps correspond to the social network variable $\frac{s_{ij}}{s_{ii}}$ and the right-hand side to the business one $\frac{b_{ij}}{b_{ii}}$, in 1993. Each pair of maps corresponds to one of the region of destination including the three largest French cities: Paris (top pair), Rhône (Lyon) (middle pair) and Bouches-du-Rhône (Marseille) (bottom pair). For each map, the highest class, for which the plotted variable is equal to 1, is drawn in black and only includes the region to which the map refers, which allows to locate it.

The top left map shows that the number of migrants to Paris is larger from those regions that are either not too far from Paris (North, North-West of France), or that are large even if more remote (the départements hosting Bordeaux, Lyon, and Marseille for instance). This gravity pattern also clearly emerges for Rhône and Bouches-du-Rhône. The effect of distance is strong again, and large regions as Paris or Nord appear as major sources of migrants. Regarding business networks, the distance effect is less striking. The impact of the size of the origin region, however, is still clear, the spatial pattern of business network being quite similar independently from the destination region. Levels change, however. This conclusion is confirmed by the high correlations between the business variables and the relative production for instance (see Table 2).

3 Results

3.1 Border effects without network variables

We now proceed to the estimation of the two specifications derived from the theoretical model. We begin with the odds specification (15) that uses the bilateral trade flows relative to the trade with self as dependent variable. Table 3 presents the results of the odds specification without social nor business network variables. Our purpose here is to investigate the impact of internal distance computation and contiguity variable inclusion on the estimated border effect, and choose a benchmark specification used later for our main purpose, the assessment of the impact of networks on trade.

For each year, estimations are computed with two definitions of the internal distance, one assuming that consumers are uniformly located on the entire region area (columns (1)) and the other on only $\frac{1}{16}$ of this area (columns (2)). Column (3) tests the impact of the inclusion of the contiguity variable.

In Table 3, the line "Total Border" gives the coefficient used to calculate the total border effect: -2.24 in column 1978(2) for instance means that in 1978, inter-regional flows between

Figure 2: Relative number of Migrants, s_{ij}/s_{ii} (left) - Relative number of Plants Connections, b_{ij}/b_{ii} (right), for Paris (top), Rhône (middle) and Bouches-du-Rhône (bottom), in 1993

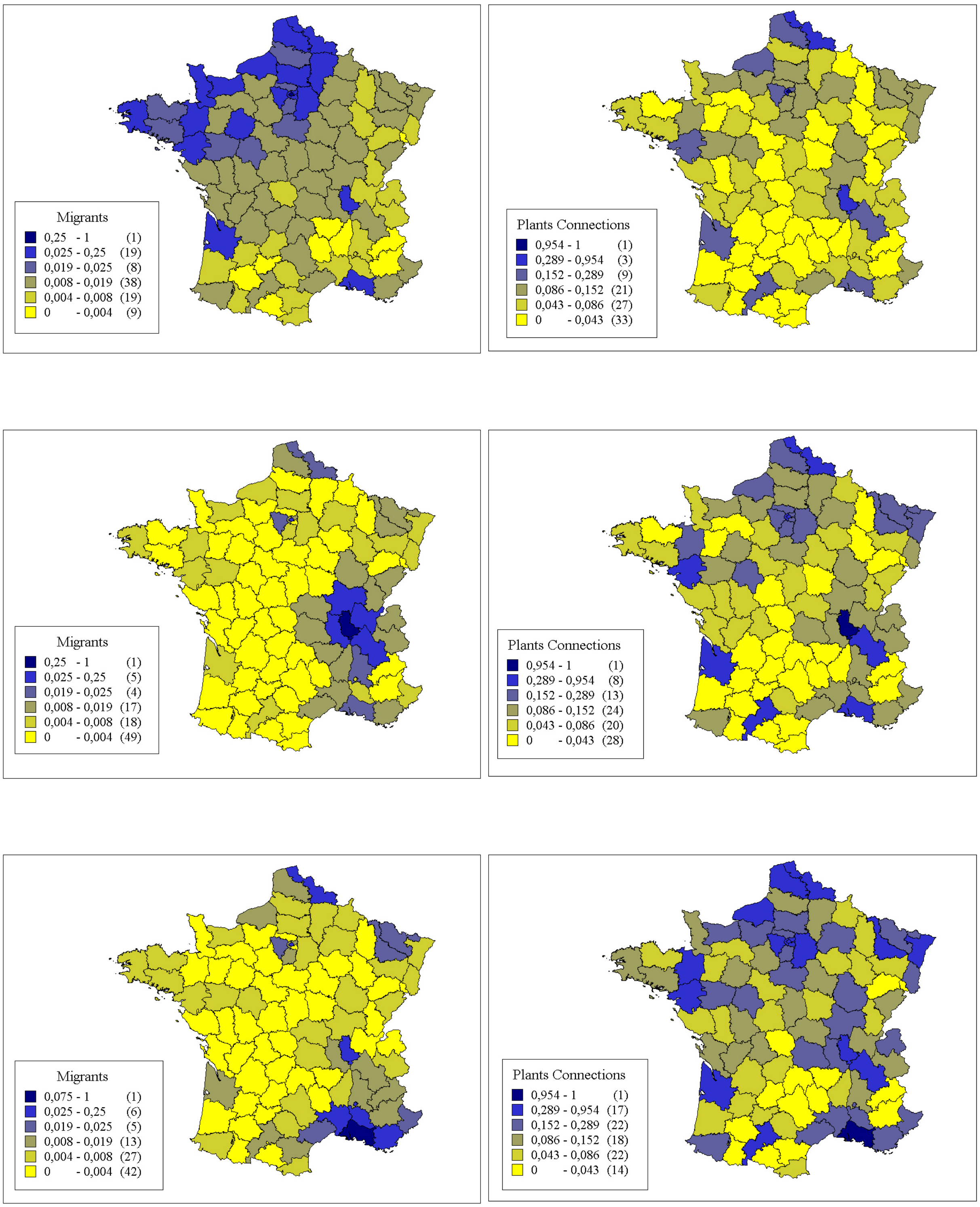


Table 3: Odds specification - Aggregate - 1978 and 1993

	Dependent Variable: $\ln(c_{ij}/c_{ii})$					
Model :	1978(1)	1978(2)	1978(3)	1993(1)	1993(2)	1993(3)
ln rel. production	0.60[a]	0.58[a]	0.58[a]	0.58[a]	0.55[a]	0.56[a]
	(0.08)	(0.08)	(0.08)	(0.07)	(0.06)	(0.06)
ln rel. transport costs (2/3R)	-1.91[a]			-2.00[a]		
	(0.18)			(0.17)		
transport costs square term (2/3R)	0.7e-08			0.7e-08		
	(0.5e-08)			(1.1e-08)		
ln rel. wages	-2.12[b]	-1.52	-1.61[c]	-2.59[a]	-1.99[a]	-2.06[a]
	(1.03)	(0.93)	(0.89)	(0.49)	(0.43)	(0.41)
ln rel. transport costs (1/6R)		-2.37[a]	-2.92[a]		-2.31[a]	-2.79[a]
		(0.10)	(0.10)		(0.11)	(0.10)
transport costs square term (1/6R)		1.8e-08[a]	2.8e-08[a]		2.6e-08[a]	4.8e-08[a]
		(0.3e-08)	(0.3e-08)		(0.8e-08)	(0.8e-08)
Total border	-3.65[a]	-2.24[a]	-1.20[a]	-2.95[a]	-1.84[a]	-0.99[a]
	(0.20)	(0.15)	(0.13)	(0.19)	(0.16)	(0.13)
Contiguity	1.40[a]	1.06[a]		1.09[a]	0.88[a]	
	(0.11)	(0.07)		(0.11)	(0.08)	
N	7935	7935	7935	7491	7491	7491
R^2	0.383	0.404	0.395	0.408	0.422	0.416
RMSE	1.591	1.564	1.575	1.536	1.518	1.526

Note: Robust standard errors in parentheses with [a], [b] and [c] respectively denoting significance at the 1%, 5% and 10% levels.

two non contiguous regions are exp(2.24) = 9.4 times lower than intra-regional ones. Column 1978(1) shows that the definition of the internal distance significantly affects the level of this border effect, a now well-established result (see Wei, 1996, and Helliwell and Verdier, 2001, for instance). However, in the current paper, we are not primarily interested in this issue, but in the way border effects are affected by networks. And indeed, we find this impact to be very similar across definitions of the internal distance chosen.

As expected, removing the contiguity variable decreases the border effect which corresponds now to the average ratio of intra-regional trade flows over all inter-regional flows. The contiguity variable permits to distinguish between two different kinds of border effects. Indeed, apart from the fact that inter-regional flows between two non contiguous regions are 9.4 times stronger than intra-regional ones, column 1978(2) also shows that inter-regional flows between two contiguous regions are $\exp(2.24-1.06) = 3.2$ times lower than intra-regional ones, while they are themselves $\exp(1.06) = 2.9$ times higher than those between two not contiguous regions. Thus, the total border effect can be decomposed as: $9.4 = 3.2 \times 2.9$. By contrast, the estimation that does not include the contiguity dummy (column 1978(3)) only shows that inter-regional flows are on average exp(1.20) = 3.3 times lower than intra-regional ones. Our central results concerning the impact of networks are largely independent of whether we include a contiguity dummy or not.

Transport costs impede trade flows, in the expected convex way. Using the assumption that consumers are not spread over the whole area of the region makes both the transport cost coefficient stronger and the quadratic one more significant (comparing columns 1978(2) and 1978(1)). On the other hand, the transport cost coefficients are simultaneously lower than those obtained when the contiguity dummy variable is not included. Thus, the estimation corresponding to columns 1978(2) leads to intermediate levels of both border and transport costs effects. Moreover, the global quality of regressions is slightly higher under this assumption, the R^2 being higher. We therefore choose to work with this "intermediate scenario" specification, even if the choice of the benchmark specification is rather innocuous on our results. The network variable effects for 1978 in the next section is compared to the benchmark estimation corresponding to column 1978(2).[10]

According to theory, the relative production coefficient should be equal to 1. This not the case here, as often in this kind of estimations. Even if the impact of the production value is a bit low, it is still largely positive: Relative trade increases with the relative size of the trading regions. Moreover, the derived border and transport effects (which are the variables of primary interest) are quite comparable whether the coefficient on production is imposed to be unitary or not. Finally, note that the estimates for the relative wages are also a bit low compared to what theory predicts. This is also an usual result in the empirical literature estimating price elasticities using trade flows. Recent studies by Erkel-Rousse and Mirza (2002), Hanson (1998), Head and Ries (2001) or Lai and Trefler (2002) for instance, find higher estimates of σ but have to rely on different estimation techniques and/or different type of data. Compared with existing studies estimating price elasticities in a gravity-like equation, our levels for this parameter are actually fairly high. Moreover, as will appear in the next section, introducing network effects increases this estimated value.

Similar comments can be drawn from the estimations on 1993 data. A noticeable difference is that the level of the total border effect drops from 9.4 in 1978 to $\exp(1.84) = 6.3$ in 1993 in the benchmark specification. The local border effect also decreases from 3.3 to 2.6. Those results suggest a process of economic integration within France. Gradual integration is also observed at the European level by Nitsch (2000) or Head and Mayer (2000) over the same period. Note that this fall in border effects comes *in addition* to the fall in our transport cost measure across time (-38% between 1978 and 1993). Last evidence of integration, we also observe a decrease in the estimated transport cost coefficient, simultaneously becoming more convex. The other estimates, corresponding to coefficients on relative production and wages, are stable across time. The predictive power of the model is slightly higher in 1993.

An intermediary result is therefore that in 1993, intra-national administrative borders within France seem to affect trade with an impact of a magnitude similar to the one Wolf (2000) finds for trade inside the United States in 1993.

We now turn to the estimation of the friction specification given in equation (17). Table 7 gives the estimation for 1978 and 1993 (columns 1978(1) and 1993(1), respectively), without network variables. Even if the model is now more constrained, all estimates are similar to the benchmark odds specification, for both years. Of primary interest, implied border effects are noticeably close. The total border effect is equal to 10.1 in 1978 and to 6.9 in 1993 in the friction specification, and the contiguity effects are exactly the same in the friction and the (benchmark) odds specifications, for both years. The transport cost estimates are also very similar in both estimations.

These consistent results can be viewed as robustness evidence, despite important differences in the variables included in the regressions. This implies that possible mis-specification or omitted variable problems are likely to be of secondary importance in our results.

[10]We will have the same intermediate scenario specification (contiguity included and smaller internal distance) used as a benchmark for 1993.

3.2 Border effects with network variables in the odds specification

Our purpose is now to study the impact of business and social networks on border effects. We first estimate different variants of the odds specification in order to isolate the social networks effects from the business ones. Results are given in Tables 4 and 5 for 1978 and 1993, respectively. The only available network variable in 1978 is based on social network (migrations), whereas business network variables can also be computed for 1993. For both tables, migration variables are introduced one by one in columns (1) and (2). Column (3) gives the results when the two migration variables are simultaneously introduced. Columns (4) to (6) in Table 5 proceed in the same way for the business network variable in 1993. Column (7) presents the 1993 estimation including all network variables.

Table 4: Odds specification with network effects - Aggregate - 1978

	Dependent Variable: $\ln(c_{ij}/c_{ii})$		
Model :	(1)	(2)	(3)
ln rel. production	0.48^a	0.42^a	0.42^a
	(0.09)	(0.08)	(0.08)
ln rel. wages	-0.38	-1.46^c	-1.16
	(1.06)	(0.87)	(1.10)
ln rel. transport costs	-2.06^a	-1.92^a	-1.91^a
	(0.12)	(0.11)	(0.11)
transport costs square term	$1.5e\text{-}08^a$	$1.4e\text{-}08^a$	$1.4e\text{-}08^a$
	(0.3e-08)	(0.3e-08)	(0.3e-08)
ln rel. migrations from origin	0.19^a		0.05
	(0.03)		(0.05)
ln rel. migrations from destination		0.27^a	0.22^a
		(0.04)	(0.06)
Total border	-1.81^a	-1.67^a	-1.65^a
	(0.18)	(0.20)	(0.20)
Contiguity	0.91^a	0.87^a	0.86^a
	(0.08)	(0.09)	(0.09)
N	7935	7935	7935
R^2	0.424	0.432	0.433
RMSE	1.537	1.526	1.525

Note: Robust standard errors in parentheses with ^{a}, ^{b} and ^{c} denoting significance at the 1%, 5% and 10% levels, respectively.

One of the main conclusion to be drawn from those tables is that both network variables have a strong negative impact on the estimated border effects in the two years considered. The reduction in the estimated border effect reaches 62.6% in 1993 when all network variables are considered. This is summarized in Table 6. The first line in this table computes the total border effect variation implied by each estimation of the odds specification. For instance, the first figure in this line means that when only the origin social network variable is introduced, the total border effect in 1978 varies by $\exp(1.81 - 2.24) - 1 = -35.0\%$, 2.24 and 1.81 being the estimate from Table 3, column (2), and from Table 4, column (1), respectively. The other figures of the line are similarly obtained with the other estimations.

When introduced simultaneously, social network variables have a strong impact on the total border effect which decreases by 44.5% and 56.3% in 1978 and 1993, respectively. When only one social network variable is introduced, variations of comparable magnitude are obtained. Note also that the effect of migrations from destination are stronger than those from origin, both in 1978 and 1993. The origin effect even loses statistical significance when both variables are introduced

Table 5: Odds specification with network effects - Aggregate - 1993

	Dependent Variable: $\ln(c_{ij}/c_{ii})$						
Model :	(1)	(2)	(3)	(4)	(5)	(6)	(7)
ln rel. production	0.41^a	0.41^a	0.37^a	0.21^b	0.35^a	0.30^a	0.26^a
	(0.06)	(0.06)	(0.06)	(0.08)	(0.07)	(0.07)	(0.07)
ln rel. wages	-1.01^a	-2.60^a	-1.89^a	-3.91^a	-2.78^a	-3.13^a	-2.80^a
	(0.34)	(0.44)	(0.46)	(0.51)	(0.45)	(0.56)	(0.54)
ln rel. transport costs	-1.92^a	-1.83^a	-1.73^a	-2.11^a	-2.06^a	-2.06^a	-1.75^a
	(0.16)	(0.14)	(0.15)	(0.11)	(0.11)	(0.11)	(0.16)
transport costs square term	1.3e-08	1.1e-08	0.8e-08	$1.5\text{e-}08^c$	1.3e-08	1.2e-08	0.4e-08
	(0.9e-08)	(0.8e-08)	(0.8e-08)	(0.8e-08)	(0.8e-08)	(0.8e-08)	(0.9e-08)
ln rel. migrations from origin	0.23^a		0.13^a				0.06
	(0.04)		(0.05)				(0.05)
ln rel. migrations from destination		0.29^a	0.22^a				0.14^a
		(0.05)	(0.05)				(0.05)
ln rel. plant connections from orig.				0.57^a		0.15	0.05
				(0.06)		(0.11)	(0.11)
ln rel. plant connections from dest.					0.70^a	0.57^a	0.53^a
					(0.07)	(0.13)	(0.13)
Total border	-1.30^a	-1.15^a	-1.02^a	-1.36^a	-1.26^a	-1.25^a	-0.86^a
	(0.16)	(0.18)	(0.18)	(0.16)	(0.16)	(0.16)	(0.19)
Contiguity	0.67^a	0.61^a	0.56^a	0.87^a	0.88^a	0.87^a	0.69^a
	(0.08)	(0.09)	(0.08)	(0.07)	(0.07)	(0.07)	(0.08)
N	7491	7491	7491	7491	7491	7491	7491
R^2	0.436	0.44	0.443	0.45	0.456	0.456	0.463
RMSE	1.5	1.495	1.49	1.482	1.473	1.473	1.464

Note: Robust standard errors in parentheses with ^{a}, ^{b} and ^{c} denoting significance at the 1%, 5% and 10% levels, respectively.

Table 6: Border, Transport Cost, and Network Effect Variations (odds specification)

	Social						Business			Soc. & Bus.
	Orig.		Dest.		Tot.		Orig.	Dest.	Tot.	Tot.
	1978	1993	1978	1993	1978	1993	1993	1993	1993	1993
Border (variation, %)	-35.0	-42.1	-43.4	-49.9	-44.5	-56.3	-38.3	-44.3	-45.0	-62.6
Transport (variation, %)	-50.0	-49.7	-63.3	-57.8	-63.8	-64.4	-23.7	-28.1	-28.5	-58.7
Network (level)	2.9	3.7	4.5	5.2	4.7	7.3	2.5	3.6	3.6	9.4

simultaneously in 1978, probably due to the fairly high correlation between the variables (see Table 2), even if its sign is still correct. Both effects are significant in 1993 when simultaneously introduced. Last, the implied border effect reductions are stronger in 1993 than in 1978.

These are evidence that social networks matter in shaping trade flows in a way that is largely discontinuous and not proportional to distance, which explains a large part of the border effect. Moreover, this phenomenon increases over time and is stronger as regards the migrations from destination than from origin. People now located in the origin region are more able to export to the destination region because they know both the destination consumer's preferences and the local retailers for instance. The fact that people migrating in the reverse direction carry their preferences to the destination region, thus increasing the imports from their origin region, has a significant but relatively smaller influence.

Business networks also play an important role in shaping trade and reducing border effects. It can be noticed in Table 5 that business networks variables exhibit a highly significant positive influence on bilateral trade. A 10% rise in this variable increases the ratio of bilateral trade to self trade by values as high as 7%. When introduced separately, the origin business variable decreases the total border effect by 38.3%, the destination business one by 44.3%. Similarly to migrations, the destination effect is stronger. This is confirmed by the fact that not only the level of estimates is higher for the destination networks when both are introduced simultaneously, but the origin effect also becomes non significant. As for social networks, both the local border effect and the contiguity one significantly decrease, but the former in higher proportions.

When all network variables are introduced (in the 1993 regression), the total border effect is drastically reduced by 62.6%. Even if the origin effect of both the social and the business variable are not significant in this regression, probably due to the correlations between variables (see Table 2), all estimates have the correct sign. Moreover, the border effect reduction is larger than in any of the other estimations. Hence, we find that 1) all variables separately affects trade 2) their influences are at least partially orthogonal.

Our central focus is to measure the impact of networks on the border effect puzzle, which we have just presented. However, another striking conclusion stems from our estimations. Not only does the discontinuous effect of borders sharply drop when network variables are introduced, but so does the impact of the transport cost. With the estimated coefficients, the impact of the transport cost on relative trade flows can first be computed as the average of the relevant term in equation (15):

$$\widehat{\sigma\delta} \ln\left(\frac{1+t_{ij}}{1+t_{ii}}\right) - \widehat{\sigma\theta}\left[(1+t_{ij})^2 - (1+t_{ii})^2\right]. \qquad (18)$$

This effect averages at 166.9 in 1978 and 123.4 in 1993, when network effects are not considered. This means that the difference between inter-regional and intra-regional transport costs causes, on average, inter-regional trade flows to be 123.4 times lower than internal ones in 1993. Thus, transport costs still largely impede trade, even when border effects are taken into account. Moreover, the impact of transport costs is on average much stronger than the impact of borders.

Next, the second line of Table 6 reports the variations of these transport cost impediments when network effects are introduced. Noticeably, the impact of the transport cost declines by amounts of comparable magnitude to the border effect reduction. For instance, when we introduce both social network effects (business ones, respectively), the decline is equal to 64.4% (28.5%, respectively) in 1993. Moreover, transport cost are significantly more influenced by social networks than by business networks, contrary to border effects on which both effects have the same magnitude. For these reasons, and due to the correlation between network variables, the impact of all the network variables simultaneously introduced appears to be slightly lower, -58.7%, than the one due to the social networks only. Finally, we once more observe that destination networks have a stronger effect on the impact of transport costs.

Thus, both the standard trade impediments (the impact of transport cost) and borders are strongly reduced by networks. We infer that, thanks to their correlation with the network variables, these variables proxy for effects others than physical cost or trade barriers, namely preferences or information costs, in basic gravity or border effect estimations. When network variables are introduced, border and transport cost variables capture a much less important share of the preferences or information cost effects, which is the reason why their impact decreases.

Besides, it is possible to compute the impact of networks on relative trade. This is reported in the last line of Table 6, which computes the average of the relevant term in equation (15),

$$-\widehat{\alpha}\ln\left(\frac{1+s_{ij}}{1+s_{ii}}\right)-\widehat{\beta}\ln\left(\frac{1+s_{ji}}{1+s_{ii}}\right)-\widehat{\gamma}\ln\left(\frac{1+b_{ij}}{1+b_{ii}}\right)-\widehat{\rho}\ln\left(\frac{1+b_{ji}}{1+b_{ii}}\right). \tag{19}$$

The first figure in the last line of Table 6 means that the differences across regions in the migrants number relative to the people working in the region they are native from (the origin social network variable) make, on average, the inter-regional trade flows 2.9 times lower than the internal ones in 1978 and 3.7 times in 1993. Both social network variables acting simultaneously decrease inter-regional trade flows by 4.7 times in 1978 and 7.3 times in 1993. The business effects, evaluated in 1993, makes the inter-regional trade flows 3.6 times lower and, finally, all networks effects together make them 9.4 times lower than inter-regional ones. These figures clearly reflect the large substitution of network effects to transport cost and border ones in the explanation of the inter-regional trade flows.

3.3 Border effects with network variables in the friction specification

As a second step and in a robustness test purpose, we estimate for both years the corresponding network effects in the friction specification, which is also computed industry by industry in the second part of this section. The results of the friction regressions on aggregated data are given in Table 7. As for the odds specification, Table 8 reports the variations of the implied border and transport cost effects, and the corresponding level of network effects.

The impact of networks in shaping trade flows and reducing border effect is even larger in this specification than in the odds one.

Border effects are reduced by 52.6% in 1978 and 60.4% in 1993 using social networks alone. Regarding the 1993 estimation, business networks alone have a higher effect on trade flows than when the odds specification is used. The implied reduction in the border effect is similar to the social networks, equal to 57.6%.

The total border effect is reduced by 70.8% when both social and business network effects are considered. The remaining total border effect in France in 1993 is equal to 2 (exp(0.7)). Although this could seem to be still a large number for a country as integrated as France, it does not seem totally unreasonable, and even looks very small compared to what is found in the literature. Estimates of border effects on international trade for the same period in Europe are between 10 and 15 (see Nitsch, 2000 and Head and Mayer, 2000 for instance). Inside the United States, Wolf (1997, 2000) finds a baseline border effect of 4.39, whereas Nitsch (2002) estimates in a recent paper an intra-German border effect of 2.2, very close to our final figure. These residual border effects may be attributed to pure home preference bias as implied by our theoretical model. It is also possible that "technical" questions raised by Anderson and van Wincoop (forthcoming) or Head and Mayer (2002) contribute to a part of the remaining border effect.

Table 7: Friction specification with network effects - Aggregate - 1978-1993

	Dependent Variable: $\ln(\Phi_{ij})$					
Model :	1978(1)	1978(2)	1993(1)	1993(2)	1993(3)	1993(4)
ln rel. transport costs	-2.29[a]	-1.64[a]	-2.22[a]	-1.54[a]	-1.93[a]	-1.60[a]
	(0.12)	(0.12)	(0.11)	(0.11)	(0.10)	(0.10)
transport costs square term	1.7e-08[a]	1.1e-08[a]	2.3e-08[b]	0.2e-08	0.7e-08	-0.2e-08
	(0.4e-08)	(0.4e-08)	(1.1e-08)	(0.9e-08)	(1e-08)	(0.9e-08)
ln rel. migrations		0.37[a]		0.40[a]		0.22[a]
		(0.03)		(0.05)		(0.04)
ln rel. plant connections					0.95[a]	0.80[a]
					(0.10)	(0.09)
Total border	-2.31[a]	-1.56[a]	-1.93[a]	-1.00[a]	-1.07[a]	-0.70[a]
	(0.18)	(0.19)	(0.16)	(0.21)	(0.18)	(0.21)
Contiguity	1.06[a]	0.81[a]	0.88[a]	0.53[a]	0.86[a]	0.67[a]
	(0.10)	(0.09)	(0.08)	(0.09)	(0.08)	(0.08)
N	3709	3709	3413	3413	3413	3413
R^2	0.467	0.521	0.511	0.544	0.573	0.581
RMSE	1.249	1.185	1.182	1.141	1.105	1.094

Note: Robust standard errors in parentheses with [a], [b] and [c] respectively denoting significance at the 1%, 5% and 10% levels.

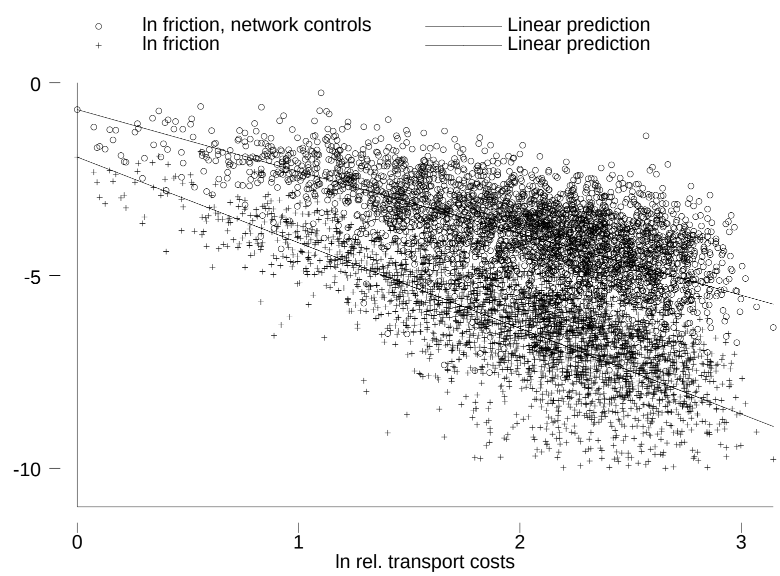
Table 8: Border, Transport Cost, and Network Effect Variations (friction specification) (%)

	Social		Business	Soc. & Bus.
	1978	1993	1993	1993
Border (variation, %)	-52.6	-60.4	-57.6	-70.8
Transport (variation, %)	-76.2	-69.1	-27.9	-59.7
Network (level)	8.1	9.6	6.7	16.8

The decrease of the transport cost impact is also striking, with the same feature than the one observed in the odds estimations, the effect of business network being smaller than the social network one. All network variables together decrease the transport impact by 59.7%. The corresponding network impact on relative trade flow is significantly larger than those obtained in the odds estimation. The gap between the inter-regional social networks and the internal ones causes the inter-regional trade flows to be 8.1 times lower than the internal ones in 1978, 9.6 times in 1993, 6.7 times for the business networks alone and 16.8 times when all network effect are simultaneously introduced, in 1993.

Last, the improvement in the overall fit of the regression when taking into account all network variables is a further indication that those effects are important in shaping trade volumes inside France, which translates for instance into a rise of the R^2 from 0.51 to 0.58 in 1993. Those results can be summarized in a plot against the relative transport cost of two modified versions of equation (17) where all right-hand side variables (with corresponding estimated coefficients) are passed on the left-hand side except for the intercept and the linear transport cost term. Figure 3, in which circles (crosses, respectively) correspond to the friction specification with (without, respectively) network controls, sums up the main conclusions of our study.

Figure 3: Gravity predictions with and without network controls



Following the inclusion of network variables, we observe that: (*i*) the regression line is higher than the prediction without network controls, and crosses the vertical axis closer to the origin, meaning that border effects almost disappear, (*ii*) the slope with respect to transport also decreases in absolute value, meaning that the transport cost impact is reduced, (*iii*) points are less dispersed around the regression line, meaning that the fit of the regression globally improves.

Overall, sections 3.1, 3.2 and 3.3 provide results that all point to a positive answer to the question asked in the title of this paper. We now proceed to detailing that answer industry by industry.

3.4 Border effects by industry in the friction specification

The level of aggregation of the data we use possibly masks the fact that different industries may be more or less affected by border effects. Furthermore it might be interesting, in terms of economic implications, to determine whether the social and business effects on trade are more or less relevant depending on the industry. The same magnitude being obtained for border effects in the friction specification as in the odds specifications, we use the former to investigate these effects for each industry separately. The friction specification has the advantage that production and wage variables are not needed, which is valuable as they are specially difficult to obtain at the industry level.

Tables 9 and 10 present the derived border effects for each industry when controlling for social networks only, in 1978 and 1993, respectively. The induced falls in trade impediments stemming from this control are given in the right-hand side of the tables (columns "Var. Bord. (%)" and "Var. Trans. (%)" for the border and the transport cost effects, respectively). Column "Soc." gives the elasticity of the social network variable, and column "Bord." the remaining total border effect. The last line of the table presents the average values stemming from the pooled estimation on all industries. Table 11 gives the corresponding results when controlling for both social and business networks in 1993.

Table 9: Friction specification - Industries - Social networks - 1978

Industry	Obs.	Soc.	Bord.	Var. Bord. (%)	Var. Trans. (%)
Miscellaneous	1557	0.4*	2.9*	-45.3	-77.8
Food and beverages	1316	0.3*	4.1*	-35.9	-66.0
Agriculture	1020	0.3*	5.3*	-33.7	-68.9
Wood and paper	1289	0.3*	4.4*	-31.2	-63.3
Extraction, steel, construction	1758	0.2*	17.2*	-26.5	-55.2
Machinery	1409	0.2*	5.6*	-26.3	-58.2
Chemicals	757	0.1*	5.0*	-13.8	-42.3
Transport	431	0.1	1.9	-13.6	-34.9
Rubber products	159	0.1	4.8	-11.1	-23.4
Textile and clothing	123	0.0	3.4*	3.0	-19.4
Pooled	9819	0.3*	4.4*	-34.3	-69.4

Note: * denotes significance at the 5%level.

The pooled regression exhibits an average reduction of the total border effect that amounts to 34.3% in 1978 and 47.5% in 1993 due to social networks, and 55.7% in 1993, due to the adding of business controls. These are other evidence of the network role that increases over time despite the French ongoing decrease in transport costs over the period. The average remaining total border effect, with a value of 2.7 in 1993, is low and quite similar to the one obtained on aggregated data. The average elasticities of trade frictions with respect to social and business variables are also comparable and always fairly high.

When performed on each industry separately, specific estimates by industry are obtained. In Tables 9, 10, and 11 industries are ranked by decreasing impact of networks on the total border effect. The highest reductions in border effects are obtained for the food related industries, namely the food and beverage industry and agriculture. However, other industries such as chemicals or extraction, steel, and construction also experience a fairly high reduction, despite the well known high transport costs that firms bear for shipping these materials.[11] The decrease in the transport

[11]Recall that the differences in transport costs sensitivity across goods arising from different transportability for

Table 10: Friction specification - Industries - Social networks - 1993

Industry	Obs.	Soc.	Bord.	Var. Bord. (%)	Var. Trans. (%)
Miscellaneous	1486	0.5^*	1.8	-62.5	-72.4
Agriculture	808	0.5^*	3.1^*	-61.7	-71.6
Food and beverages	1669	0.3^*	1.8^*	-48.6	-56.7
Machinery	1022	0.3^*	3.8^*	-47.2	-63.0
Extraction, steel, construction	1465	0.3^*	8.7^*	-42.4	-54.6
Chemicals	611	0.2	5.0^*	-31.5	-43.3
Textile and clothing	116	0.2	4.3^*	-24.6	-25.7
Wood and paper	1192	0.1	3.5^*	-16.7	-24.4
Transport	533	0.1	3.9^*	-13.3	-20.8
Rubber products	163	-0.2	10.1^*	50.7	65.5
Pooled	9065	0.3^*	3.2^*	-47.5	-61.1

Note: * denotes significance at the 5%level.

Table 11: Friction specification - Industries - Social and Business networks - 1993

Industry	Obs.	Soc.	Bus.	Bord.	Var. Bord. (%)	Var. Trans. (%)
Miscellaneous	1486	0.4^*	0.5^*	1.3	-72.9	-67.2
Agriculture	808	0.4^*	0.1	3.0^*	-63.0	-70.5
Food and beverages	1669	0.2^*	0.4^*	1.4	-60.0	-49.2
Extraction, steel, construction	1465	0.2^*	0.4^*	6.9^*	-54.3	-44.3
Machinery	1022	0.3^*	0.2	3.6^*	-50.0	-60.5
Chemicals	611	0.3^*	-0.2	5.5^*	-24.7	-49.7
Wood and paper	1192	0.1	0.2	3.3^*	-21.4	-19.5
Transport	533	0.1	0.1	3.7^*	-17.8	-18.7
Textile and clothing	116	0.6	-1.0	6.3^*	10.5	-79.4
Rubber products	163	-0.1	-0.2	11.4^*	70.1	43.9
Pooled	9065	0.3^*	0.3^*	2.7^*	-55.7	-55.3

Note: * denotes significance at the 5%level.

cost effect is often larger than the decrease in the border effect. The ranking made according to the strongest variation in transport cost effect is the same as the one made according to the strongest variation in the border effect. As a conclusion, networks similarly affect both kinds of trade impediments, in all industries.

Only the lowest ranked industries in terms of border effect variation may experience an increase of the border effect. However, those are industries in which network effects are not significant, probably due to the small number of observations. This lack of observations also probably explains why both network variables, which are moreover positively correlated, are not often simultaneously significant when both introduced in 1993.

Rankings are almost the same in all considered periods and whether business networks are also considered or not. The textile and clothing industry (the rank of which significantly increases) and the wood and paper industry (for which it decreases) are two exceptions. Moreover, these variations are reduced when business network effects are also introduced. Therefore, differences across industries regarding the impact of network effect can be considered as fairly stable.

4 Conclusion

Obstfeld and Rogoff (2000) present the border effect on trade as a major puzzle to international macroeconomics. The understanding of this puzzle has recently progressed in several directions. Abstracting from methodological issues, those advances have mostly focused on providing explanations for the large impact of *political* (international) borders.

The monetary union explanation, as any type of specifically international transaction costs explanation like tariff or non tariff barriers, cannot however be the only reason why borders matter in trade flows. Indeed, as was first shown by Wolf (1997, 2000), *administrative* (intra-national) borders also seem to significantly impede trade.

We tried in this paper to investigate an explanation for those intra-national border effects: If the existence of social and business networks promotes trade as emphasized by Rauch (2001), this could explain the existence of border effects inside a country as networks are presumably much more dense and easier to maintain on short distances and therefore inside administrative borders.

We have shown that intra-national administrative borders significantly matter in trade patterns inside France with an impact of the same order of magnitude that Wolf (2000) finds for trade inside the United States. However, more than 60% of these intra-national border effects can be explained by the employment composition in terms of birth place (social networks) or by inter-plants connections (business networks). The two types of networks taken separately make the estimate of border effects fall by around 50%. When controlling for both type of networks, a French region is estimated to trade only twice more with itself than with a non adjacent region of similar size and distance. Moreover, we have also shown that networks reduce the impact of transport cost on inter-regional trade by an amount of comparable magnitude, around 60%. Last, these effects emerge in almost all industries, with slightly different but stable intensities.

References

Anderson, S.P., A. De Palma and J.-F. Thisse, 1992, *Discrete Choice Theory of Product*

instance are taken into account since the transport cost parameters now also vary across industries.

Differentiation. Cambridge : MIT Press.

Anderson J. and E. van Wincoop, forthcoming, "Gravity with Gravitas: A Solution to the Border Puzzle", *American Economic Review.*

Armington, 1969, "A Theory of Demand for Products Distinguished by Place of Production", *IMF Staff Papers*, 16, 159-178.

Combes, P.P., and M. Lafourcade, 2001, "Transport Cost Decline and Regional Inequalities: Evidence from France", CEPR Discussion Paper, 2894.

Combes, P.P., and M. Lafourcade, 2002, "Transport Costs Effects on Core-Periphery Patterns of Market Potential: France, 1978-1998", CERAS mimeo.

Dixit, A. and J. Stiglitz, 1977, "Monopolistic Competition and Optimum Product Diversity", *American Economic Review* 67(3): 297-308.

Eaton, J. and S. Kortum, 2002, "Technology, Geography and Trade" *Econometrica*, 70(5): 1741-1780.

Erkel-Rousse, H. and D. Mirza, 2002, "Import Price-elasticities: Reconsidering the Evidence", *Canadian Journal of Economics*, 35(2): 282-306.

Ethier, W., 1982, "National and International Returns to Scale in the Modern Theory of International Trade", *American Economic Review*, 72: 389-405.

Evans, C. and J. Harrigan, 2002, "Distance, Time, and Specialization", Federal Reserve Bank of New York, mimeo:
http://www.ny.frb.org/rmaghome/economist/harrigan/pubs.html.

Gould, D., 1994, "Immigrant Links to Home Country: Empirical Implications for US Bilateral Trade Flows", *Review of Economics and Statistics*, 76: 302-316.

Hanson, G., 1998, "Market Potential, Increasing Returns, and Geographic Concentration", NBER Working Paper, 6429. Revised 2001 version available at:
http://www-irps.ucsd.edu/faculty/gohanson/Uscnty.pdf

Head, K. and T. Mayer, 2000, "Non-Europe : The Magnitude and Causes of Market Fragmentation in Europe", *Weltwirtschaftliches Archiv*, 136(2): 285-314.

Head, K. and T. Mayer, 2002, "Illusory Border Effects: How internal geography affects external trade volumes", CEPII Working paper, 2002-1:
http://www.cepii.fr/anglaisgraph/workpap/pdf/2002/wp02-01.pdf.

Head, K. and J. Ries, 2001, "Increasing Returns Versus National Product Differentiation as an Explanation for the Pattern of US-Canada Trade", *American Economic Review* 91(4): 858-876.

Helliwell, J., 1996, "Do National Borders Matter for Quebec's Trade?", *Canadian Journal of Economics*, 29(3): 507-522.

Helliwell, J., 1997, "National Borders, Trade and Migration", *Pacific Economic Review*, 3(3): 165-185.

Helliwell J. and G. Verdier, 2001, "Measuring Internal Trade Distances: A New Method Applied to Estimate Provincial Border Effects in Canada," *Canadian Journal of Economics*, 34(4):1024-1041.

Hummels, D., 1999, "Towards a Geography of Trade Costs", Krannert School of Management mimeo:
http://www.mgmt.purdue.edu/faculty/hummelsd/research/toward/Toward3.pdf.

Hummels, D., 2001, "Time as a Trade Barrier", Krannert School of Management mimeo:
http://www.mgmt.purdue.edu/faculty/hummelsd/research/time3b.pdf.

Krugman, P., 1980, "Scale Economies, Product Differentiation, and the Pattern of Trade", *American Economic Review*, 70: 950-959.

Lai, N. and D. Trefler, 2002, "The Gains from Trade with Monopolistic Competition: Specification, Estimation, and Mis-Specification", NBER Working Paper, 9169.

Limão, N. and A. Venables, 2001, "Infrastructure, Geographical Disadvantage, Transport Costs and Trade", *World Bank Economic Review*, 15(3): 451-479.

McCallum, J., 1995, "National Borders Matter: Canada-US Regional Trade Patterns", *American Economic Review*, 85:615-23.

Nitsch, V., 2000, "National Borders and International Trade: evidence from the European Union", *Canadian Journal of Economics*, 22(4): 1091-1105.

Nitsch, V., 2002, "Border Effects and Border Regions: Lessons from the German Unification", mimeo Bankgesellschaft Berlin.

Obstfeld, M. and K. Rogoff, 2000, "The Six Major Puzzles in International Macroeconomics: Is There a Common Cause?", NBER Working Paper, 7777.

Parsley, D. and S-J. Wei, 2001, "Explaining the Border Effect: The Role of Exchange Rate Variability, Shipping Costs and Geography" *Journal of International Economics*, 55(1): 87-105.

Rauch, J., 2001, "Business and Social Networks in International Trade", *Journal of Economic Literature*, 39: 1177-1203.

Rauch, J. and V. Trindade, 2002, "Ethnic Chinese Networks in International Trade" *Review of Economics and Statistics*, 84(1):116-130.

Redding S. and A. Venables, 2001, "Economic Geography and International Inequality", London School of Economics mimeo:
http://econ.lse.ac.uk/staff/ajv/winc.pdf.

Rose A.K., 2000, "One money, One market: the Effect of Common Currencies on Trade", *Economic Policy*, 15(30): 7-46.

Taglioni D., 2001, "Evidence on the Relationship between Exchange Rates Variability and Trade Barriers : a Role for Volatility in Explaining Border Effects?", University of Geneva mimeo.

Wagner, D., Head, K. and J. Ries, 2002, "Immigration and the Trade of Provinces", forthcoming in *Scottish Journal of Political Economy.*

Wei, S-J., 1996, "Intra-National Versus International Trade : How Stubborn Are Nations in Global Integration?", NBER Working Paper, 5531.

Wolf, H.C., 1997, "Patterns of Intra- and Inter-State Trade", NBER Research Working Paper, 5939.

Wolf, H.C., 2000, "Intranational Home bias in Trade", *Review of Economics and Statistics*, 82(4): 555-563.