# A re-evaluation of the impact of regional agreements on trade patterns: The EU case.

Lionel Fontagné* Thierry Mayer† Soledad Zignago‡

March 9, 2005

**Very preliminary and incomplete.**

**Abstract**

Not done yet

JEL classification: F12, F15
Keywords: Regional Integration, Border Effects, Gravity.

*CEPII. Email: fontagne@cepii.fr
†Université de Paris-Sud, CEPII, CERAS, and CEPR. Email: tmayer@univ-paris1.fr
‡CEPII and TEAM. Email: zignago@cepii.fr

# 1 Introduction

*A Wider Europe*. This March 2003 Communication from the Commission to the Council and the European parliament has launched a new European policy addressing the issues of prosperity, peace and security at the borders of an enlarged EU (actually, beyond EU25): the European Neighbourhood Policy (ENP).

According to Commissioner Benita Ferrero-Waldner, "The European Neighbourhood Policy is a new policy that invites our neighbours to the East and to the South to share in the peace, stability and prosperity that we enjoy in the European Union and which aims to create a ring of friends around the borders of the new enlarged EU" (European Commission Website).

Concerned countries are the following:. On the Eastern border: Belarus, Ukraine and Moldova. The Southern Caucasus countries are also concerned: Georgia, Armenia, and Azerbaijan. Lastly the countries of the Barcelona process: Morocco, Algeria, Tunisia, Libya, Egypt, Israel, Jordan, Lebanon, Syria, as well as the Palestinian Authority. This list does neither encompass current candidates (Bulgaria, Romania, Turkey, Croatia), nor the other Western Balkan states which are "potential candidates" according to the official terminology in Brussels. Accordingly, the cornerstone here is that the ENP is designed for those countries that "do not *currently* have a perspective for membership".

How new this policy is an important issue at stake. According to the Commissioner " The ENP offers the countries an increasingly close relationship with the EU, involving a significant degree of economic integration and a deepening of political co-operation in an approach founded on partnership, joint ownership and differentiation". However, economic integration and political co-operation had already been introduced one decade ago, with mitigated success, in the Barcelona framework. This is for the Mediterranean world. On the Eastern border of the enlarged EU, a very different experience will constitute the benchmark of this new policy; namely the May 2004 enlargement which is the coronation of a decade of reforms in the new member states. This is why, to some extent, this new policy might be interpreted as "New Wine in Old Wineskins" (Kelley, 2005): political, institutional and economic reforms, traded against a better access to the Internal Market or more generally "further integration and liberalization to promote the free movement of persons, goods, services and capital" (EU commission, 2003).

Interestingly, the Commission itself is inclined to accept an "Everything but institutions" (R. Prodi) menu, in which deep economic integration would be the core advance: internal market, access to EU policies, dismantling of trade barriers, adoption of EU norms and standards inter alia.

The general principle is to define Action Plans (AP) embedding the most sensitive issues (democracy, market economy, etc.) with each neighbour, according to a differentiation strategy. The first set of AP has been published in December 2004. For instance, the introduction of the 3-years AP for Ukraine is pointing out to the need to "support Ukraine's objective of further integration into European economic and social structures. [to] significantly advance the approximation of Ukrainian legislation, norms and standards to those of the European Union. It will also build solid foundations for further economic integration, including through joint efforts towards an EU-Ukraine Free Trade Area following Ukraine's accession to the WTO, on the basis of the adoption and implementation of economic and trade-related rules and regulations with the potential to enhance trade, investment and growth." Besides, concerns related to " Ukraine's efforts and concrete achievements in meeting commitments to common values", which are of course also political values, are also stressed.

To stick to the economic dimension of this new instrument, the issue at stake is: integrating economies on a preferential and regional basis will promote trade, but by how much? Does the

previous experience of European integration provide any evidence? Should we stick on a more systematic approach examining experiences of integration in the various regions of the world?

To be honest, nobody knows. And the problem is that everybody will answer ! We expect a large set of studies to be launched in the near future, drawing conclusions based on the extensive use of a tool having demonstrated its poor performances as far as the estimation of trade creation effects of regional agreements is concerned.

The purpose of this paper is accordingly to reconsider this question, based on new empirical evidence, an approach which is line with current developments of the literature. Rather than estimating the trade creation effects of a market integration between (e.g.) Ukraine and the EU, better start with the starters: how confident are we in the tool, and how to improve this confidence. This is not to deny the benefits of such an integration policy: on the contrary, recent evidence tends to boost its expected gains. However, even the recent literature is relying on incredibly poor data: GDP, (sometimes population) distance, a handful of covariates, and last but not least fixed (and sometimes random) effects. The sectoral dimension of the problem is completely hidden; this is generally the same for the level of tariffs; as well as the bulk of trade relationships of each economy, which are *internal*.

The influence of regional agreements on trade patterns has been the primary subject of the huge literature using gravity equations to study bilateral volumes of trade. There are dozens of papers in this literature dating back to the early seventies, that is soon after the gravity equation had been applied to international trade for the first time. The exceptional fit of a linear (in logs) simple regressions involving three core variables for which data are among the easiest to find clearly contributed to the initial and persistent success. Researchers investigated almost all possible preferential trading agreements (PTA), looking for positive deviations from the "norm of trade" given by gravity.

For a while, technical progress in this area has mostly consisted in adding covariates that were found to improve the fit of the regression, various measures of countries' geographical, historical and cultural proximity have been added to the set of explanatory variables, in an objective to provide better control for unobserved characteristics of the bilateral trade between two countries. This is indeed important for the assessment of the impact of PTA. For instance, most PTA involve a large proportion of contiguous countries. Omitting this variable, while putting an PTA dummy in a cross section regression can generate fallacious positive estimates on trade policy due to an omitted variable bias.

More recently, decisive progress has been made in the link between the empirical gravity equation and the different modelling frameworks available in trade theory. Many authors had tried to take into account the remoteness of trading partners, that is not only their bilateral distance, but also their location with respect to third countries. Recent work by several economists (surveyed in Feenstra, 2003) has helped to make sense of those ad hoc attempts, in deriving the gravity equation rigorously from trade theory and providing a clear description of what a correct econometric specification should be. Here again the matter is important for obtaining a valid estimate of the impact of a PTA. To take only an example, do Australia and New Zealand trade a lot with themselves because of their common remoteness from easy alternative source of products and markets to exports, or because of their trading arrangement? The improved link with theory enables to sort those out by radically improving the norm of trade to which exchanges within a PTA is compared.

Even with those progresses, one has to think about the reference group to which members of a PTA are compared when doing such an exercice. Indeed, the endogeneity of the membership decision has been recently demonstrated by Baier and Bergstrand (2004), and the choice of the control group to evaluate the impact of the PTA is not an easy task. Furthermore, results of this literature exhibit

disconcerting variance, with widely varying estimates among studies, troubling ranking of trade creating agreements, and other characteristics that contrast dramatically with the stability of the core gravity estimates.

We envision to solve this reference group issue using two different methods. First, as already done in the border effect literature, we will consider international trade inside a PTA not with respect to trade with third countries solely, but will compare those two flows to a third one: trade within countries, which seems to be the natural benchmark of a perfectly integrated economy. Second, we will apply score matching techniques, recently introduced in the trade literature. This involves a two step procedure, where in the first step the probability of becoming a member of the PTA is actually modelled, and in a second step, the trade volumes of actual PTA members are compared to a counterfactual consisting of the most proximate countries possible, that is the one that were most likely to enter the considered PTA but did not.

# 2 The myriad of problems associated with estimating the impact of PTAs on trade patterns

The considerable literature estimating the trade impact of PTAs rely mostly on a dummy variable introduced in a gravity-type equation explaining the log of bilateral imports of country $i$ from country $j$ ($M_{ij}$) of the following type:

$$\begin{aligned} \ln M_{ij} &= \alpha + \beta_1 \ln Y_i + \beta_2 \ln Y_j + \beta_3 \ln d_{ij} \\ &\quad + \beta_4 \mathbf{C}_{ij} + \beta_5 \mathrm{EU}_{ij} + \epsilon_{ij}, \end{aligned} \tag{1}$$

where the $Y$ terms (designating the GDPs of the two trade partners) and $d_{ij}$ (measuring the distance separating them) represent the "core" of the gravity equation. The core gravity alone explains most of the variance in bilateral exports. The $\mathbf{C}_{ij}$ term represents a vector of control covariates, intended to refine the econometric exercice, in general through a better specification of transaction costs (most frequent variables are common language, contiguity and variables indicating colonial links). Income per capita is also often introduced to account for endowment differences. Last, researchers introduce a dummy variable for the PTA under investigation, the European Union in this example ($\mathrm{UE}_{ij}$), taking the value of one if $i$ and $j$ are both member countries. The exponential of the coefficient on this variable is interpreted as the deviation from the trade norm due to the PTA: Everything else equal, $\exp(\beta_5) = \frac{M_{ij}}{M_{ik}}$, where $k$ is another country that does not belong to the EU.

The results of such estimations are not as clear as what would be expected. Take as an example the deepest and longest PTA experience, the EU. The evidence on the extent and evolution of European trade integration using gravity estimates is in fact *surprisingly negative*. Start by considering widely cited work by Jeffrey Frankel and co-authors. Table 2 in Frankel et al. (1995) and Table 4.2 in Frankel (1997) are very representative of their results for European trade integration. The least that can be said is that they are not very conclusive at first sight. The coefficients on the EU dummy variables start to be significant in the mid-eighties and a typical estimate in 1990 is 0.5, that is two member countries of the EC trade that year "only" 65% more ($\exp(.5) = 1.65$) than two otherwise similar country (based on a 63 times 63 matrix of total trade representing 88.7% of world trade in 1992). Even worse, when adding common language the coefficient is reduced by half and becomes insignificant when adding the general level of openness of European countries. Haveman and Hummels (1998)

have very similar results for trade flows over the 1980-1992 period, with coefficients at 0.585 in 1980, 0.825 in 1986 and 0.406 in 1989. Soloaga and Winters (2001) compare a large number of regional agreements with the goal of identifying the evolution of regional agreement dummies over time and particularly around the periods of bloc formation/revival. Here again the sample is very large with 58 importing countries over the 1980-1996 period. *The EC dummy is found to be consistently and very significantly negative.* Worse, the economic magnitude of the effect seems to be very large. The coefficient starts at -1.78 in 1980 and is at -0.80 in 1996 (which at least shows some "integration" over time).

Some estimates are more reasonable, but the point here is that PTA coefficients using this standard method are not very stable, and sometimes contradict expectations based on other empirical evidence on the actual trade liberalization taking place inside the different PTAs. The question we ask in this paper is how much of the problems in those results come in fact from flaws in the standard gravity equation estimation, in terms of econometric specification, and proper identification of the impact of PTAs. We start by describing the different issues, and then turn to the empirics.

## 2.1 How to tackle the endogeneity problem

Clearly, the presence of a free trade agreement between two countries is not exogenous to their characteristics. Their geographical location (Are they adjacent? Are they remote from the bulk of other markets?), their economic size (Does each country offer to its counterpart a wide or a limited variety of goods and services?), a common history, a common culture, a common set of political institutions, are among the main characteristics impacting the probability of signing a PTA. One may object that the intensity of their actual bilateral trade relationship is also a key determinant: This is where the issue becomes intricate!

The endogeneity problem is common in international trade: Papers addressing the impact of non-tariff barriers on trade flows, for instance, will often find a positive impact at a sufficiently disaggregated level, because of the unobservable (e.g. political economy) variables (Trefler, 1993). Regarding bilateral trade equations, Baier and Bergstrand (2004) identify various sources of endogeneity.

The measurement error is potentially a key issue. There is a first difficulty associated with the time dimension of the PTA (When does it start? Are the effects delayed? Have economic agents anticipated this integration ?) and a second one associated with the very nature of the agreement (How deep is the integration?). In both cases, a dummy used in cross section will hardly be used with confidence. The simultaneity bias, associated with the fact that exports and imports are endogenous to trade flows is not a key concern: What we tackle is bilateral flows, not total flows, and only the net exports are at stake. Accordingly, instrumenting does not change significantly the results (Frankel, 1997). Omitted variables are numerous: They range from institutions to domestic regulations. Baier and Bergstrand instrument PTAs in various ways without solving the problem as long as the panel dimension of the data is not mobilized. Their interpretation of such result is that instrumental variables are in fact correlated with trade flows. Accordingly, cross-sectional estimates of the impact of PTAs should be avoided. Using panel data, they get more convincing and larger effects, which are definitively positive.

It is also possible that the estimated impact of the PTAs will be influenced by an inadequate specification since, for instance, countries very isolated because of their physical geography will have at the same time few neighbors to sign PTAs with and high price indexes which reduces their imports.

We will therefore in a first step compare simple gravity PTA estimates with coefficients that deal properly with the specification issues of the gravity equation.

## 2.2 Assessing properly the effects of PTAs

There are a lot of issues to tackle.

First we must rely on *longitudinal data* and hence exploit the panel dimension of our data set. We will consider 105 importers declaring their imports from 135 exporters. We are using mirror data in order to complete our data set. Accordingly, CIF *Mij* imports of the 30 non-reporting countries are inferred from the corresponding *Xji* reported by the 105 reporters; The corresponding value is inflated by a 10 percent CIF/FOB factor, in line with the DOTS. Accordingly, we avoid using mirror data: imports by non-declaring countries will not be taken into account. The period covered is 1976-2000 (annual data).

Second, we consider that relying on *sectoral data* is a prerequisite. Instead of using GDPs, we will rely on industry level data for production and for consumption within 26 ISIC industries. In total, we rely on a 135x134x26x25 potential trade flows, minus zero values, due to computation capacities. Notice also that certain reporting countries have missed to report their imports in certain years; accordingly, 105 reporters is a maximum. By sake of comparison Baier and Bergstrand (2004) rely on 96x95x5 potential observations minus the zeros. The two differences are that they rely on a panel with only 5 periods (1960, 70,..,2000) and have less zeros since they rely on total exports, rather than on sectorial exports.

Third, instead of relying on a single dummy for all PTAs, it is important to tackle separately the impact of each PTA, since *not all PTAs will have the same effects*, in particular in absence of information as their exact content. we identify 7 different regional agreements: EU, NAFTA, ASEAN, MERCOSUR, Andean Community, CACM. The huge differences among these agreements regarding the level of economic integration of their members is obvious.

Fourth, we must take into account not only the existence of a PTA, but also the level of tariff barriers. Here, we will rely on bilateral tariff data based on an aggregation of HS6 tariff data for some 5,000 products towards our 26 industries. Weights are based on world imports for the corresponding products.

In addition to these prerequisites, econometric issues have to be tackled regarding i) the issue of constraining the elasticity of trade flows to production to one; and ii) the use of fixed (vs random) effects by pair of countries in order to control for all unobservable determinants of PTAs.

## 2.3 The right benchmark for integration

The *international* trade flows are not sufficient to gauge international markets integration. The measure of the degree of international fragmentation of market is therefore by nature linked to the assessment of the impact of national borders. In Europe, this has been best summarized as the whole idea of the *Single* Market, which explicitly states its goal to be the abolition of the economic significance of national borders, in particular as far as prices are concerned.

In order to make that assessment, one needs to consider not only international trade flows but also flows of goods *inside* each country and see how they compare. To do a comparison between internal and external trade flows *caeteris paribus*, the gravity approach may be adapted (McCallum (1995) and Wei (1996)are the seminal papers). Indeed, even in the absence of information regarding trade flows between sub-national regions, the *total* volume of trade occurring within a country is

equal to the overall production of the country minus its total exports. This observation can then be inserted in a bilateral trade equation, together with all the international flows.

The border effects methodology offers important advantages for the study of regional integration. For a lot of issues, the border effect measure is a useful methodology because it captures *all* impediments to trade related to the existence of the national borders, through their impact on trade flows. Accordingly, all hidden determinants associated with internal regulations impacting trade flows will be controlled for.

Also, border effects are more informative in the study of the evolution of trade barriers. In a traditional gravity equation, using a dummy variable for trade taking place inside a given PTA, how should we interpret a rise in the coefficient on this dummy variable? Using the traditional Vinerian interpretation of regional integration, this rise can first come from trade creation; the rise can however also come from a trade diversion. The gravity equation in its most traditional form (and even in more elaborated forms like Fukao et al., 2003, recent paper) find it hard to differentiate among the two causes, whereas border effects methodology enables to track a potential fall in the surplus of trade taking place inside countries, and therefore separate trade creation from trade diversion effect. John Romalis (2002) provides an intermediate approach, where a bilateral trade equation of US imports is first run, and US imports from self are then used to compute trade diversion effects of NAFTA and CUSFTA.

We will therefore compare the border effect methodology that takes an appropriate benchmark of almost perfect integration to identify the impact PTAs, to a more traditional use of the gravity equation. Remind however that in both cases we will rely first on panel data and second on sectoral data. This combination is new even in the context of the traditional gravity estimation.

Let us fist explain why relying on a border effect methodology is justified from a theoretical point of view.

## 2.4 Beyond endogeneity: The link with theory

In order to better address the issues raised in the previous section, the key issue of the empirical specification of the gravity equation has to be reconsidered. All empirical specifications should be guided (at least crudely) by theoretical analysis. This is how we will proceed below.

Complete reviews of how trade theory can yield gravity-type predictions are to be found in Feenstra (2003) and Anderson and van Wincoop (2004). One of the most convincing theoretical foundation of the gravity model is the Dixit-Stiglitz-Krugman model of trade under monopolistic competition and iceberg trade costs. It yields the following equation:

$$M_{ij} = n_j p_{ij} c_{ij} = n_j p_j^{1-\sigma} \tau_{ij}^{1-\sigma} E_i P_i^{\sigma-1}, \tag{2}$$

where $n_j$ designate the number of varieties and $p_j$ the mill price in $j$, $E_i$ and $P_i = (\sum_k n_k p_k^{1-\sigma} \tau_{ik}^{1-\sigma})^{1/(1-\sigma)}$ are the expenditure and the "price index" in $i$, respectively.

We can see from (2) that trade costs influence demand more when there is a high elasticity of substitution, $\sigma$. Following Head and Mayer (2000), we take the ratio of $m_{ij}$ over $m_{ii}$, country $i$'s imports from itself, the $\mu_i Y_i P_i^{\sigma-1}$ term then drops and we are left with relative numbers of firms, relative preferences, and relative costs in $i$ and $j$:

$$\frac{m_{ij}}{m_{ii}} = \left(\frac{n_j}{n_i}\right) \left(\frac{a_{ij}}{a_{ii}}\right)^{\sigma-1} \left(\frac{p_j}{p_i}\right)^{1-\sigma} \left(\frac{\tau_{ij}}{\tau_{ii}}\right)^{\sigma-1}. \tag{3}$$

We will refer below to this specification as explaining "relative imports".

To estimate (3), we need to specify more fully the model. The first step is to use the supply side characteristics of the monopolistic competition model. Firms producing $q_j$ in country $j$ employ $l_j$ workers in an IRS production function $l_j = F + \gamma q_j$, where $F$ is a fixed (labour) costs, and $\gamma$ the inverse productivity of firms. Profits are $\pi_j = p_j q_j - w_j(F + \gamma q_j)$, with $w_j$ the wage rate in $j$. Using the pricing equation, together with the free entry condition, we get the equilibrium output of each representative firm, $q_j = \frac{F(\sigma-1)}{\gamma}$. With identical technologies, $q_j \equiv q$, $\forall j = 1..N$ and noting $v_j$ the value of production for the considered industry in $j$, $v_j = qp_j n_j$, and we get the first substitution to be made in equation (3):$\frac{n_j}{n_i} = \frac{v_j}{v_i}\frac{p_i}{p_j}$.

Finally, functional forms for trade costs ($\tau_{ij}$) and preferences ($a_{ij}$) have to be specified in order to get an estimable equation.

- Trade costs are a function of distance ($d_{ij}$, which proxies for transport costs) and "border-related costs", which can consist of tariffs and/or broadly defined NTBs (quantitative restrictions, administrative burden, sanitary measures...). We note the *ad valorem* equivalent of all border-related costs $\text{brc}_{ij}$:

$$\tau_{ij} \equiv d_{ij}^{\delta}(1 + \text{brc}_{ij}).$$

  We want to control for the impact of membership of Preferential Trading Arrangements (PTAs) and we observe some of the actual protection taking place between importing and exporting countries (bilateral applied tariffs). We assume the following structure for border-related costs, which vary across country pair and depend on the *direction* of the flow for a given pair:

$$1 + \text{brc}_{ij} \equiv (1 + t_{ij})(\exp[\eta E_{ij} + \theta \text{PTA}_{ij}])$$

  In this specification, $t_{ij}$ denotes the *ad valorem* bilateral tariff. $\text{PTA}_{ij}$ is a dummy variable set equal to 1 when $i(\neq j)$ and $j$ belongs to a regional integration agreement and $E_{ij}$ is the intercept. We expect $\theta > 0$ to be the lowest of those parameters, which will be true if, all national borders impose transaction costs, with the minimum burden of those costs being between PTA members.

- Preferences have a random component $e_{ij}$, and a systematic preference component for goods produced in the home country, $\beta$. Sharing a common language, a common border, a common colonizer is assumed to mitigate this *home bias*. For the sake of simplicity we only introduce in the equation the effect of sharing a common language, but in the estimations we introduce more variables of bilateral affinity as the contiguity, the colonial links and the common history.[1]

$$a_{ij} \equiv \exp[e_{ij} - (\beta - \lambda L_{ij})(E_{ij} + \text{PTA}_{ij})].$$

  $L_{ij}$ is set equal to one when two different countries share the same language. When $L_{ij}$ switches from 0 to 1, home bias changes from $\beta$ to $\beta - \lambda$.

[1]The "same country" variable sets to one if the two countries were or are the same state or the same administrative entity for a long period. The "colonial link" dummy refers to countries that have ever had a colonial link. The "common colonizer" dummy equals to one if countries have had a common colonizer after 1945.

We obtain an estimable equation from the monopolistic Krugman (1980) competition equation with home bias. In its more general form, the estimated equation in the next sections will be:

$$\begin{aligned}\ln\left(\frac{m_{ij}}{m_{ii}}\right) &= -(\sigma-1)[\beta+\eta]+\ln\left(\frac{v_j}{v_i}\right)-\sigma\ln\left(\frac{p_j}{p_i}\right)-(\sigma-1)\ln(1+t_{ij}) \\ &\quad -(\sigma-1)\delta\ln\left(\frac{d_{ij}}{d_{ii}}\right)+(\sigma-1)\lambda L_{ij} \\ &\quad -(\sigma-1)[\theta-\eta]\mathrm{PTA}_{ij}+\epsilon_{ij}, \end{aligned} \tag{4}$$

with $\epsilon_{ij}=(\sigma-1)(e_{ij}-e_{ii})$.

The constant of this regression ($-(\sigma-1)[\beta+\eta]$) gives the border effect of international trade for countries *that do not belong to a PTA*. Accordingly, we expect a very large estimated value for this parameter. It includes both the level of protection of the importing country ($\eta$) and the home bias of consumers ($\beta$). The coefficient on $\mathrm{PTA}_{ij}$ gives the additional volume of trade generated by the agreement.

## 2.5 Results: Gravity plus sectoral panel data

The first step is to run simple gravity estimates using our panel of sectoral data.

Instead of relying on GDPi and GDPj as proxies for demand in $i$ and supply capacity in $j$, we take benefit of the sectoral dimension of our data to use the information on demand in sector $k$ of country $i$, and on supply of country $j$ in the same sector. Using such information will leave us more confident with the use of a disaggregated gravity equation. *Except in column 3 of Table 1, we keep internal flows* in the regressions by sake comparison with results using the border effect methodology in the next section. The surplus of intra-national trade is captured by the *Border* dummy.

In the first column of Table 1, we have a simple gravity equation explaining imports of country $i$ from provider $j$ in sector $k$ in year $t$, on the basis of production in $j$, consumption in $i$, transport cost between $i$ and $j$ proxied by their distance, plus additional variables generally used in such exercises: Namely contiguity, common language, common colonizer, colonial link, and whether jurisdictions i and j belong or not to the same country. As compared to traditional gravity equations, a new variable is introduced (Border), which takes the value 1 when an external flow is considered and 0 otherwise. We lastly introduce a proxy taking the value 1 if a PTA is signed between i and j and 0 otherwise.

All parameter estimates have the expected sign and are significant at the 1 percent level, with the exception of the variable "same country". The magnitude of the coefficients on production and consumption is much below the unit elasticity that was expected. The value of the parameter for the dummy PTA is 0.94, corresponding to a 155 percent increase in trade for countries having signed an agreement.

In column 2, we simply split the variable PTA into various PTAs. The parameters are not significantly changed. The interesting and expected result is the very large difference in the value of the parameter estimate for PTAs, which are all significant at the 1 percent level. It ranges from 0.48 (Andean Community) to 2.36 (NAFTA), discouraging from using a single dummy for all kind of PTAs.

In column 3, we replicate the previous estimation, excluding however the border variable, in accordance with our neglect of internal flows. Results are not significantly modified.

In columns 4 and 5, we proceed with external as well as internal trade flows for each country. Carefully notice that we do not rely on equation (4) at this stage, but simply estimate a classical

Table 1: Simple Gravity

| | Dependent Variable: Ln Imports | | | | |
|---|---|---|---|---|---|
| Model : | (1) | (2) | (3) | (4) | (5) |
| intcpt | -0.42$^a$ | -0.30$^a$ | -5.48$^a$ | -0.93$^a$ | -8.52$^a$ |
| | (0.03) | (0.03) | (0.03) | (0.07) | (0.02) |
| Border | -5.28$^a$ | -5.31$^a$ | | -4.85$^a$ | -5.20$^a$ |
| | (0.02) | (0.02) | | (0.05) | (0.02) |
| PTAs | 0.94$^a$ | | | | |
| | (0.01) | | | | |
| Contiguity | 0.74$^a$ | 0.74$^a$ | 0.70$^a$ | 0.81$^a$ | 0.53$^a$ |
| | (0.01) | (0.01) | (0.01) | (0.03) | (0.01) |
| Common Language | 0.57$^a$ | 0.64$^a$ | 0.65$^a$ | 0.79$^a$ | 0.78$^a$ |
| | (0.01) | (0.01) | (0.01) | (0.02) | (0.01) |
| Colonial Link | 0.79$^a$ | 0.76$^a$ | 0.75$^a$ | 0.73$^a$ | 0.38$^a$ |
| | (0.01) | (0.01) | (0.01) | (0.03) | (0.01) |
| Same Country | -0.01 | 0.07$^a$ | 0.06$^b$ | 0.36$^a$ | 0.51$^a$ |
| | (0.02) | (0.02) | (0.02) | (0.05) | (0.02) |
| Common Colonizer | 0.55$^a$ | 0.50$^a$ | 0.50$^a$ | 0.82$^a$ | 1.36$^a$ |
| | (0.01) | (0.01) | (0.01) | (0.03) | (0.01) |
| Production | 0.78$^a$ | 0.77$^a$ | 0.77$^a$ | 0.81$^a$ | 1 |
| | (0.00) | (0.00) | (0.00) | (0.00) | |
| Consumption | 0.57$^a$ | 0.56$^a$ | 0.56$^a$ | 0.53$^a$ | 1 |
| | (0.00) | (0.00) | (0.00) | (0.00) | |
| Distance | -0.85$^a$ | -0.84$^a$ | -0.87$^a$ | -0.80$^a$ | -0.99$^a$ |
| | (0.00) | (0.00) | (0.00) | (0.01) | (0.00) |
| EU | | 1.41$^a$ | 1.36$^a$ | 1.09$^a$ | 0.53$^a$ |
| | | (0.01) | (0.01) | (0.02) | (0.01) |
| NAFTA | | 2.36$^a$ | 2.34$^a$ | 2.11$^a$ | 0.72$^a$ |
| | | (0.07) | (0.07) | (0.09) | (0.08) |
| ASEAN | | 1.49$^a$ | 1.46$^a$ | 1.48$^a$ | 1.14$^a$ |
| | | (0.05) | (0.05) | (0.09) | (0.06) |
| MERCOSUR | | 1.16$^a$ | 1.16$^a$ | 0.75$^a$ | 0.96$^a$ |
| | | (0.11) | (0.11) | (0.14) | (0.12) |
| Andean Comm. | | 0.48$^a$ | 0.48$^a$ | 0.24$^a$ | 1.06$^a$ |
| | | (0.06) | (0.06) | (0.08) | (0.07) |
| CACM | | 1.69$^a$ | 1.67$^a$ | 1.52$^a$ | 2.63$^a$ |
| | | (0.08) | (0.08) | (0.15) | (0.09) |
| Ln(1+Tariff) | | | | -2.79$^a$ | |
| | | | | (0.05) | |
| N | 1250552 | 1250552 | 1221579 | 245667 | 1250552 |
| $R^2$ | 0.531 | 0.533 | 0.487 | 0.559 | 0.272 |
| RMSE | 2.506 | 2.5 | 2.524 | 2.451 | 2.725 |

Note: Standard errors in parentheses: $^a$, $^b$ and $^c$ represent respectively statistical significance at the 1%, 5% and 10% levels. The reported standard errors take into account the correlation of the error terms for a given importer.

gravity equation, tentatively using more detailed data that it is generally done. The value of this dummy variable is negative and very large: crossing a border dramatically reduces trade. The parameter estimates are very similar with the ones obtained without the internal trade flows, including the one concerning PTAs. In column 4, we split the dummy variable for PTAs into our 6 dummies; We also introduce tariffs. We record a reduction in the value of the parameter for Border and for PTAs, with the exception of ASEAN. The elasticity on tariffs makes sense. Lastly, we constrain the elasticity on production and consumption to one, as expected in a theoretical model, in column 5. The hierarchy of positive trade impacts of PTAs is now profoundly affected.

With the exception of the estimated equation reported in column 5, one might argue that *taking into account* what we considered as the right benchmark, namely *internal flows*, *does not fundamentally change the picture*. The estimation of the trade impact of regional arrangements remains highly unstable. In particular, as exemplified in column 4 of Table 1, the estimated impact of the European internal market is surprisingly limited: should we conclude that NAFTA and the EU are associated with such large differences? This is because columns 4 and 5 use the right data, but with the *wrong specification*. The right specification is the one given by the theoretical modelled referred to above. One needs to introduce price indices, and since those prices are not observable, we must turn to relative imports in order to drop this term. Accordingly, production, consumption and price levels have to be dropped and replaced by relative values. This will be done in the next section.

## 2.6 Results: Border effects plus panel data

We now turn to the estimated equations corresponding to our theoretical benchmark. Results reported in table 6 correspond to estimations without intercept (as referred to above, but including a border effect. Production, consumption and distances are now considered in relative terms in order to drop the price index term. We keep Contiguity, Common language, Colonial link, Same country, and then proceed as previously. In column 1 we use a single dummy for all PTAs. In column 7 we disentangle PTAs. In column 8 we add information on bilateral tariffs. In column 9 we constrain the elasticity on the relative production to be unity instead of .75 as estimated otherwise.

All covariates bear the expected sign. The PTA variable in column 6 will not be interpreted here, since we know that it should be disentangled. relative prices and distances also highlight the expected sign.

Interestingly, the border effect is very large. There are explanations to be given here. First, we do consider the border effect among countries that do not belong to a PTA. Accordingly, the effect is maximized. Second, this is a worldwide estimation, incorporating countries such as Tanzania and Pakistan, in addition to OECD members: the limited amount of numerous South-South trade flows does not come out as a surprise to us. Third, the individual border effects between pair of countries are simply averaged in these pooled estimates, notwithstanding the economic size of the economic partners. Fourth, constraining the supply elasticity to unity significantly reduces the parameter estimate. Fifth, the border effect between Tanzania and Pakistan bears the same weight as between the U.S. and Japan. The level of border effects estimated here is roughly only 20 times what would be observed within an integrated region such as the EU. Also, taking the tariff equivalent of such border effect would lead to reasonable magnitudes. Last but not least, they might be a bias in estimating border effects, but what is more interesting is their evolution over time.

Now, we get in column 7 a positive impact of the EU on trade flows which is much more in line with what should be expected. In particular, it is at least not inferior to the one estimated for the NAFTA, and it is larger than for the ASEAN or for the MERCOSUR, for instance.

Table 2: Border Effects

| | Dependent Variable: Ln Rel. Imports | | | |
|---|---|---|---|---|
| Model : | (6) | (7) | (8) | (9) |
| Border | -6.43$^a$ | -6.42$^a$ | -5.52$^a$ | -5.74$^a$ |
| | (0.03) | (0.03) | (0.06) | (0.03) |
| PTAs | 1.75$^a$ | | | |
| | (0.02) | | | |
| Contiguity | 1.22$^a$ | 1.26$^a$ | 1.13$^a$ | 0.88$^a$ |
| | (0.02) | (0.02) | (0.03) | (0.02) |
| Common Language | 0.24$^a$ | 0.38$^a$ | 0.54$^a$ | 0.44$^a$ |
| | (0.01) | (0.01) | (0.03) | (0.01) |
| Colonial Link | 1.16$^a$ | 1.06$^a$ | 1.01$^a$ | 1.12$^a$ |
| | (0.02) | (0.02) | (0.04) | (0.02) |
| Same Country | 0.24$^a$ | 0.40$^a$ | 0.61$^a$ | 0.26$^a$ |
| | (0.03) | (0.03) | (0.05) | (0.03) |
| Common Colonizer | 0.73$^a$ | 0.68$^a$ | 1.15$^a$ | 0.72$^a$ |
| | (0.02) | (0.02) | (0.07) | (0.03) |
| Rel. Production | 0.72$^a$ | 0.72$^a$ | 0.76$^a$ | 1 |
| | (0.00) | (0.00) | (0.01) | |
| Rel. Prices | -0.32$^a$ | -0.33$^a$ | -0.15$^a$ | -0.95$^a$ |
| | (0.01) | (0.01) | (0.02) | (0.01) |
| Rel.Distance | -0.45$^a$ | -0.46$^a$ | -0.53$^a$ | -0.70$^a$ |
| | (0.01) | (0.01) | (0.02) | (0.01) |
| EU | | 2.57$^a$ | 1.88$^a$ | 2.38$^a$ |
| | | (0.02) | (0.04) | (0.02) |
| NAFTA | | 2.54$^a$ | 1.99$^a$ | 2.29$^a$ |
| | | (0.06) | (0.07) | (0.06) |
| ASEAN | | 2.00$^a$ | 1.81$^a$ | 1.97$^a$ |
| | | (0.08) | (0.15) | (0.08) |
| MERCOSUR | | 1.02$^a$ | 0.43$^a$ | 1.01$^a$ |
| | | (0.07) | (0.10) | (0.08) |
| Andean Comm. | | 0.24$^a$ | -0.32$^a$ | 0.10 |
| | | (0.07) | (0.09) | (0.07) |
| CACM | | 2.42$^a$ | 1.98$^a$ | 2.48$^a$ |
| | | (0.08) | (0.15) | (0.08) |
| Ln(1+Tariff) | | | -4.62$^a$ | |
| | | | (0.25) | |
| N | 1311754 | 1311754 | 274761 | 1311754 |
| $R^2$ | 0.44 | 0.445 | 0.489 | 0.226 |
| RMSE | 2.716 | 2.703 | 2.607 | 2.804 |

Note: Standard errors in parentheses: $^a$, $^b$ and $^c$ represent respectively statistical significance at the 1%, 5% and 10% levels. The reported standard errors take into account the correlation of the error terms for a given importer.

Introducing tariffs seriously erodes the positive impact of the PTAs we consider here.

# 3 Are PTAs exogenous?

Another source of concern is the interpretation of the effect of a PTA dummy variable. Signing a preferential trade arrangement or a free trade zone agreement is not an event that should be modelled as an exogenous shock happening to certain pairs of countries and not to others. PTAs are of course not exogenous, and depend on determinants that potentially also affect trade and are sometimes unobservable to the econometrician. The omitted variable bias is an important question, which has been neglected altogether until quite recently. It raises questions about controlling for those unobservables, potentially finding instruments for the signature of PTAs and thinking about what a reasonable counterfactual is.

## 3.1 Omitted variables

The first problem relates to the unobservable characteristics that make a PTA likely between two countries. When implementing empirically their hypothesis regarding unobservable characteristics, Baier and Bergstrand (2004) indeed find that their list of covariates explain an important part of the PTA decision. Some of those determinants are not a real issue for our purposes, because they can be observed very easily. For instance, bilateral distance in their first probit regression, explains more than half of the PTA decision. It also explains trade patterns, but is controlled for in the trade equation, alleviating the endogeneity issue.

It is however quite likely that there other simultaneous determinants of PTAs and trade that are not that so easy to observe. Bilateral trust between the two partners is an example of a variable that affect both PTA signature and trade flows (see Guiso et al., 2005) and is almost impossible to observe. For a certain proportion of those unobservable, the bilateral structure is likely to be fixed over time, which suggests to *introduce fixed effects for pairs of trade partners*, as done in Carrère (2005). This specification allows to purge the coefficient on PTAs from bilateral fixed characteristics of countries that affect both PTAs and trade flows.

Therefore, we also as a third step, introduce country-pair effects, in order to identify the trade-impact of PTAs entirely on the time dimension, that is either comparing trade flows before and after the initial signature, or before and after entries in the PTA.

Column 1 in table 3 is now imply replicating the column 1 in table 2 by sake of comparison with additional estimates. In column 2, we introduce random effects and in column 3 random effects. Egger (2000) advocates in favor of the rejection of the random effects in the gravity model. However, we have here an additional industry dimension what makes it more justifiable to test such specification. Interestingly, it permits to keep covariates which have only a bilateral (and no time) dimension, such as Contiguity for instance. But actually, such choice does not significantly changed the results. Now, the value of the parameter estimate on PTAs is dramatically reduced. Using fixed effects does not change this observation.

We conclude that on average, a large part of the trade effects of PTAs are due to unobservable characteristics of countries entering in such agreements. *Instead of considering that on average trade is 5.7 times larger within a RAT, we find only a 2.4 increase when controlling for such unobservable country-pairs characteristics.*

The same is true for the bulk of our sample: in the EU case, this difference is particularly striking. Neglecting the bilateral fixed (or random) effects would lead to overestimate trade effects by a factor

Table 3: Fixed effects dyads

| | Dependent Variable: Ln Rel. Imports | | | | | |
|---|---|---|---|---|---|---|
| Model : | OLS | RE | FE | OLS | RE | FE |
| Border | -6.43$^a$ | -6.79$^a$ | -7.58$^a$ | -6.42$^a$ | -6.74$^a$ | -7.55$^a$ |
| | (0.03) | (0.05) | (0.00) | (0.03) | (0.05) | (0.00) |
| Rel. Production | 0.72$^a$ | 0.86$^a$ | 0.89$^a$ | 0.72$^a$ | 0.86$^a$ | 0.89$^a$ |
| | (0.00) | (0.00) | (0.00) | (0.00) | (0.00) | (0.00) |
| Rel. Prices | -0.32$^a$ | -0.64$^a$ | -0.62$^a$ | -0.33$^a$ | -0.64$^a$ | -0.62$^a$ |
| | (0.01) | (0.01) | (0.01) | (0.01) | (0.01) | (0.01) |
| Rel.Distance | -0.45$^a$ | -0.49$^a$ | | -0.46$^a$ | -0.50$^a$ | |
| | (0.01) | (0.01) | | (0.01) | (0.01) | |
| Contiguity | 1.22$^a$ | 1.46$^a$ | | 1.26$^a$ | 1.52$^a$ | |
| | (0.02) | (0.10) | | (0.02) | (0.10) | |
| Common Language | 0.24$^a$ | 0.26$^a$ | | 0.38$^a$ | 0.30$^a$ | |
| | (0.01) | (0.05) | | (0.01) | (0.05) | |
| Colonial Link | 1.16$^a$ | 1.36$^a$ | | 1.06$^a$ | 1.31$^a$ | |
| | (0.02) | (0.10) | | (0.02) | (0.10) | |
| Same Country | 0.24$^a$ | 0.74$^a$ | | 0.40$^a$ | 0.76$^a$ | |
| | (0.03) | (0.16) | | (0.03) | (0.16) | |
| Common Colonizer | 0.73$^a$ | 0.68$^a$ | | 0.68$^a$ | 0.67$^a$ | |
| | (0.02) | (0.07) | | (0.02) | (0.07) | |
| PTAs | 1.75$^a$ | 0.88$^a$ | 0.87$^a$ | | | |
| | (0.02) | (0.01) | (0.01) | | | |
| EU | | | | 2.57$^a$ | 0.80$^a$ | 0.76$^a$ |
| | | | | (0.02) | (0.02) | (0.02) |
| NAFTA | | | | 2.54$^a$ | 1.32$^a$ | 1.31$^a$ |
| | | | | (0.06) | (0.08) | (0.08) |
| ASEAN | | | | 2.00$^a$ | 1.05$^a$ | 1.03$^a$ |
| | | | | (0.08) | (0.06) | (0.06) |
| MERCOSUR | | | | 1.02$^a$ | 1.29$^a$ | 1.29$^a$ |
| | | | | (0.07) | (0.15) | (0.15) |
| Andean Comm. | | | | 0.24$^a$ | 1.40$^a$ | 1.41$^a$ |
| | | | | (0.07) | (0.07) | (0.07) |
| CACM | | | | 2.42$^a$ | 0.69$^a$ | 0.65$^a$ |
| | | | | (0.08) | (0.09) | (0.09) |
| N | 1311754 | 1311754 | 1321917 | 1311754 | 1311754 | 1321917 |
| $R^2$ | 0.44 | 0. | 0.182 | 0.445 | 0. | 0.181 |
| RMSE | 2.716 | . | 2.286 | 2.703 | . | 2.287 |

Note: Standard errors in parentheses: $^a$, $^b$ and $^c$ represent respectively statistical significance at the 1%, 5% and 10% levels. The reported standard errors take into account the correlation of the error terms for a given importer.

6.

# 4 The right counterfactual

The last method of introducing country-pair fixed effects does not however entirely solve the omitted variables problem. Indeed, some of the characteristics that raise both bilateral trade and the probability of signing a PTA might be moving together over time. In order to control for this type of simultaneity bias, one needs to build the appropriate counterfactual, that is comparing the change in trade flows following the PTA not with the average non-member pairs, but with the pair that resembles the most the one that signed the agreement. This is what is done by matching techniques. Those start by estimating the probability of signing an PTA, then match each "treated" dyad with the dyad that was most likely to be treated but was not, and last compare their respective trade flows, before and after the treatment.

– To be completed –

# 5 References


Aitken, N. (1973), "The Effect of the EEC and EFTA on European Trade: A Temporal Cross-Section Analysis", *American Economic Review* 63(5):881-92.

Anderson, J. and E. van Wincoop (2003a), "Gravity with Gravitas: A Solution to the Border Puzzle", *American Economic Review* 93(1):170-192.

Anderson, J. and E. van Wincoop (2004), "Trade Costs" *Journal of Economic Literature*, 42(3).

Baier, S. and J. Bergstrand (2004), "Do Free Trade Agreements Actually Increase Members' International Trade", mimeo October.

Bouët, A., Y. Decreux, L. Fontagné, S. Jean, and D. Laborde (2004), "A consistent, *ad-valorem* equivalent measure of applied protection across the world: The MAcMap-HS6 database", CEPII Working Paper # 22.

Chen, N. (2003), "Intra-national versus international trade in the European Union: why do national borders matter?", forthcoming in the *Journal of International Economics*.

Disdier, A-C. and K. Head (2003), "Exaggerated Reports of the Death of Distance: Lessons from a Meta-Analysis", presented at the ETSG 2003 meetings in Madrid.

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

European Commission (2003), *The Internal Market - Ten Years without Frontiers*, available at

Evenett, S. and W. Keller (2003), "On Theories Explaining the Success of the Gravity Equation", *Journal of Political Economy*, 110(2): 281-316.

Feenstra, R. (2003), *Advanced International Trade: Theory and Evidence*, Princeton: Princeton University Press.

Fontagné L., and M. Mimouni (2001), "L'environnement, nouvel obstacle au commerce de produits agricoles et alimentaires", *Economie Internationale*, (87):63-87.

Frankel, J.A., 1997, *Regional Trading Blocs* Washington: Institute for International Economics.

Frankel, J., E. Stein et S-J. Wei, 1995, "Trading Blocs and the Americas: The Natural, the Unnatural, and the Supernatural." *Journal of Development Economics* 47(1):61-95.

Fukao, K., T. Okubo and R. Stern (2001), "An Econometric Analysis of Trade Diversion Under NAFTA", *North American Journal of Economics and Finance* 14:3-24.

Guiso, L., P. Sapienza and L. Zingales (2005), "Cultural Biases in Economic Exchanges", CEPR Discussion Paper #4837.

Haveman, J., U. Nair-Reichert and J. Thursby (2003), "How effective are trade barriers? An empirical analysis of trade reduction, diversion and compression", *The Review of Economics and Statistics* 85(2): 480-485.

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

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.

Krugman, P.R. (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.

McCallum, J. (1995), " National Borders Matter: Canada-US Regional Trade Patterns", *American Economic Review* 85:615-623.

Nitsch, V. (2000), "National Borders and International Trade: Evidence from the European Union." *Canadian Journal of Economics* 33(4) : 1091-1105.

Romalis, J. (2002), "NAFTA's and CUSFTA's Impact on North American Trade" mimeo, University of Chicago GSB.

Rose, A.K. (2000), "One Money, One Market: Estimating the Effect of Common Currencies on Trade" *Economic Policy* 30:9-45.

Soloaga I. and A. Winters (2001), "Regionalism in the Nineties: What Effect on Trade?", *North American Journal of Economics and Finance* 12:1-29.

Trefler D. (1993), "Trade Liberalization and the Theory of Endogenous Protection: An Econometric Study of U.S. Import Policy", *The Journal of Political Economy* 101:138-160.

Wagner, D. (2003), "Aid and Trade–An Empirical Study," *Journal of the Japanese and International Economies*, 17(2):153-173.

Wei, S-J. (1996), "Intra-National Versus International Trade: How Stubborn Are Nations in Global Integration?" National Bureau of Economic Research Working Paper # 5531.

# A Data requirements

The needed data involves bilateral trade and production figures in a compatible industry classification for developed and developing countries. Those come from the CEPII Trade and Production database, which is an extension of the Alessandro Nicita and Marcelo Olarreaga at the World Bank.[2] This dataset (available at http://www.cepii.fr/anglaisgraph/bdd/TradeProd.htm) compiles data for developing and developed countries at the ISIC rev2 3-digit industry level over the period 1976-2000.

The relative prices are captured by the price level of GDP expressed relative to the United States. This data comes from the Penn World Tables v.6.1.[3]

As can be seen in equation (4), we need measures of distances between ($d_{ij}$) *and within* ($d_{ii}$) countries for the countries in the sample. Two potential problems arise: How to define internal distances of countries and how to make those constructed internal distances consistent with "traditional" international distances calculations? The second question is in fact crucial for obtaining a correct estimate of the border effect. Any overestimate of the internal / external distance ratio will yield to a mechanic upward bias in the border effect estimate. We have developed a new database of internal and external distances[4], which uses city-level data in the calculation of the distance matrix to assess the geographic distribution of population inside each nation. The basic idea is to calculate distance between two countries based on bilateral distances between cities weighted by the share of the city in the overall country's population. This procedure can be used in a totally consistent way for both internal and international distances, which solves the problems highlighted above. The database also contains the contiguity, common language, colonial relationship and common colonizer variables used here.

Tariffs can be measured at the bilateral level and for each product of the HS6 nomenclature in the TRAINS database from UNCTAD. We base our investigation on weighted averages of bilateral tariffs obtained from TRAINS. Those tariffs are aggregated from Jon Haveman's treatment of TRAINS data (UTBC Database[5]) in order to match our ISIC rev2 industry classification using the world imports as weights for HS6 products. The obtained variable is a rather crude measurement of protection, when compared for instance with a dataset recently made available (called Market Access Map, MacMap, see Bouët et al., 2004) that takes into account the complex system of bilateral preferences across countries in the world at a detailed product level. This type of data however lacks any consistent time coverage which is an important issue here. Even in manufactured goods and between industrialized countries, tariffs are not negligible and vary quite substantially across industries and countries combinations.

Besides tariffs, there are other obstacles to trade imposed by governments at the border in order to protect national industries and that will be captured by the border effects in the above regressions. Those non-tariff barriers (NTBs), for which tariff equivalent are difficult to compute, take a myriad of different forms, from traditional border formalities and administrative harassment to more sophisticated sanitary and phyto-sanitary measures. For a given HS6 category, the NTB variable is

[2]The original data comes principally from United Nations sources, the COMTRADE database for trade and UNIDO industrial statistics for production.

[3]We also experiment with a more detailed –but more incomplete, and maybe more noisy– variable of relative wages by industry. The Dixit-Stiglitz behavior of profit maximizing firms yields the well-known fixed markup over marginal costs ($p_j = \frac{\sigma}{\sigma-1}\gamma w_j$), which gives us $\frac{p_j}{p_i} = \frac{w_j}{w_i}$. The relative wages come from UNIDO and consist of the industry's wage bill divided by the number of employees. In the end, the results are slightly better with the global price variable and we therefore present results only with this one.

[4]Available at http://www.cepii.fr/anglaisgraph/bdd/distances.htm

[5]http://www.eiit.org/Protection/extracts.html

set equal to 1 if at least one of the underlying tariff lines in that category is subject to a NTB, and 0 otherwise. As for tariffs data, this information on NTBs comes from Jon Haveman's treatment of TRAINS data and is then aggregated to match with the 3-digit ISIC rev2 classification by calculating a frequency index.