



Documents de Travail

N° 2020/1 • Janvier 2020

ESTIMATION DE POTENTIELS DE COMMERCE PAR UN MODÈLE DE GRAVITÉ

Julien ARTHUR
Alisson DRAY
Léonardo PUPPETTO

ESTIMATION DE POTENTIELS DE COMMERCE PAR UN MODÈLE DE GRAVITÉ

Julien ARTHUR
Alisson DRAY
Léonardo PUPPETTO

Ce document de travail n'engage que ses auteurs. L'objet de sa diffusion est de stimuler le débat et d'appeler commentaires et critiques.

Julien ARTHUR était en poste à la Direction Générale du Trésor au Ministère de l'Économie et des Finances (France)

Alisson DRAY était en poste à la Direction Générale du Trésor au Ministère de l'Économie et des Finances (France)

Léonardo PUPPETTO est en poste à la Direction Générale du Trésor au Ministère de l'Économie et des Finances (France)

leonardo.puppetto@dgtresor.gouv.fr (+33 1 44 87 73 63)

Table des matières

Résumé.....	4
Abstract.....	4
Synthèse	5
Introduction.....	6
1. Les fondements théoriques des modèles de gravité : vers une approche « structurelle ».....	8
1.1 Les limites de l'équation de gravité à la Tinbergen	8
1.2 Anderson et Van Wincoop (2003) proposent un modèle de gravité fondé théoriquement, tenant compte de l'influence des marchés tiers sur les relations commerciales bilatérales	8
1.3 Des estimations avec effets fixes permettent de capter les termes de résistance multilatérale.....	10
2. Estimer une équation de gravité : MCO, moindres carrés non linéaires et estimateurs du pseudo maximum de vraisemblance	10
2.1. L'estimation MCO d'un modèle de gravité log-linéarisé sur des données de commerce est généralement biaisée	10
2.2. L'absence de flux de commerce entre de nombreux pays pose également problème dans une estimation par MCO.....	11
2.3. L'estimateur des moindres carrés non linéaires n'est généralement pas satisfaisant pour estimer les paramètres d'un modèle de commerce.....	12
2.4. Les estimateurs par pseudo maximum de vraisemblance (PML) basés sur des lois exponentielles linéaires permettent d'obtenir des estimations sans biais des paramètres d'intérêts	12
3. Le choix des données dans la construction du modèle de gravité	15
3.1. Les exportations sont intégrées en valeur afin de respecter la forme structurelle de l'équation	16
3.2. Le poids économique des partenaires est approché par leur produit intérieur brut (PIB).....	17
3.3. Les coûts de commerce bilatéral sont décrits par un ensemble de variables bien identifiées par la littérature.....	17
4. Les résultats obtenus sont en ligne avec la littérature	20
4.1. Validation des résultats et estimation retenue	20
4.1.1 On observe des effets très comparables entre les résultats issus des deux bases de données, permettant d'utiliser la base la plus large	20
4.1.2 L'utilisation d'effets fixes est essentielle pour l'estimation de l'effet des politiques commerciales	21
4.2. Les coefficients ne requièrent pas de transformation particulière et sont interprétables directement....	22
4.2.1 Interprétation des coefficients.....	22
4.2.2 Conversion en équivalents ad valorem	22
5. Le modèle se montre également très robuste au niveau sectoriel pour les biens et les services	22
5.1. Le modèle de gravité permet de calculer des « potentiels de commerce », c'est-à-dire des flux d'exportation théoriques à l'équilibre des équations estimées	22
5.2. Les estimations sectorielles des potentiels de commerce ont la même base économétrique que le modèle global	23
5.3. Les résultats des modèles de gravité sectoriels sont robustes et l'étude de la variation des coefficients entre secteurs apporte une information intéressante.....	25
Bibliographie	30

Résumé

Ce document de travail présente le modèle de gravité développé par la Direction générale du Trésor du Ministère de l'Économie et des Finances. Ce type de modèle, très répandu dans l'analyse des déterminants du commerce international, décrit la relation entre les flux commerciaux bilatéraux et des variables représentatives de forces d'attraction telles que la distance et la taille économique des partenaires, mais aussi l'existence de frontières communes, de systèmes juridiques similaires, d'anciens liens coloniaux ou encore d'un accord de libre-échange entre eux. Le modèle de gravité de la DG Trésor est estimé à un niveau agrégé (sur la base de flux de commerce bilatéraux annuels globaux entre 184 pays sur une période allant de 1949 à 2015) *via* la méthode du pseudo-maximum de vraisemblance de Poisson avec effets fixes, tenant ainsi compte des flux de commerce nuls et des termes de résistance multilatérale. L'approche agrégée étant susceptible de masquer des spécificités sectorielles, le modèle est également décliné à un niveau sectoriel (pour 11 secteurs d'échanges de biens et 3 secteurs d'échanges de services). Les flux d'exportations théoriques résultant de l'estimation par le modèle de gravité sont appelés potentiels de commerce. Leur analyse permet notamment : (i) de vérifier si les exportations d'un pays avec ses partenaires sont en ligne avec la structure gravitaire mondiale du commerce, une déviation par rapport à cet équilibre pouvant alors renseigner sur l'existence d'autres déterminants que ceux du modèle dans la relation commerciale bilatérale (ii) d'identifier les pays et les secteurs vers lesquels il existe des gains commerciaux potentiels dans l'optique d'aider à mieux cibler les politiques d'accompagnement du commerce extérieur et (iii) de simuler l'impact d'un choc tarifaire sur les flux commerciaux.

Abstract

This paper presents the gravity model developed by the Directorate General of the Treasury of the Ministry of Economy and Finance. This kind of model, which is widely used in the analysis of the features of international trade, describes the relationship between bilateral trade flows and variables representing forces of attraction such as the distance and economic size of the partners, but also the existence of common borders, similar legal systems, former colonial ties or a free trade agreement between them. The gravity model is estimated at an aggregate level (on the basis of aggregate annual bilateral trade flows between 184 countries over a period from 1949 to 2015) using the Poisson's pseudo-maximum likelihood method with fixed effects, thus taking into account zero trade flows and multilateral resistance terms. As the aggregate approach is likely to hide sectoral specific features, the model is also applied at a sectoral level (for 11 goods trade sectors and 3 services trade sectors). The theoretical export flows resulting from the gravity model's estimation are called trade potentials. Their analysis makes it possible in particular to : (i) to check whether a country's exports with its partners are in line with the global gravity structure of trade, a deviation from this equilibrium can then provide information on the existence of features other than those of the model in the bilateral trade relationship (ii) to identify the countries and sectors towards which there are potential trade gains with a view to helping to better target the foreign trade policy and (iii) to simulate the impact of a tariff shock on trade flows.

Synthèse

- **En économie un modèle de gravité décrit, par analogie avec la théorie newtonienne de la gravitation, le lien entre les flux commerciaux de partenaires et leur distance économique et géographique¹.** La première équation de gravité a été estimée par Tinbergen en 1969 et montrait une relation négative entre les exportations d'un pays et la distance le séparant de ses partenaires commerciaux. Des travaux plus récents ont permis d'asseoir les fondements microéconomiques des modèles de gravité et de mieux spécifier les équations à estimer. De nouvelles techniques d'estimation, initialement appliquées au domaine de la recherche médicale et de l'économie de la santé, ont également été mises en œuvre, permettant de corriger un biais d'estimation lié à l'utilisation d'estimateurs de type moindres carrés ordinaires (MCO) sur des données hétéroscédastiques log-transformées.
- **Le modèle élaboré par la DG Trésor repose sur une base de données recensant les échanges commerciaux entre plus de 180 pays depuis 1949** (soit plus de 640 000 observations). Conformément à la littérature, il intègre des effets fixes exportateur et importateur (les « termes de résistance multilatérale » introduits par Anderson et van Wincoop (2003)). Le modèle est estimé à un niveau global par pseudo-maximum de vraisemblance de Poisson avec effets fixes. Ce type d'estimateur permet d'assurer une contrainte de bouclage des données de commerce théoriques découlant du modèle, ce qui n'est pas le cas dans les estimations classiques par MCO.
- **Des équations au niveau sectoriel, comprenant 11 secteurs d'échanges de biens et 3 secteurs d'échanges de services, sont également estimées.** L'exhaustivité des bases de données utilisées permet d'obtenir des estimations statistiquement significatives et cohérentes avec l'intuition économique. Ce type d'application a jusqu'à présent été très peu étudié dans la littérature.
- **Afin de pouvoir éclairer des choix de politique économique, la notion de « potentiel de commerce » est introduite².** Pour une année donnée, le potentiel de commerce se comprend comme le niveau d'exportations qu'atteindrait un pays avec un ou plusieurs de ses partenaires si leur relation suivait exactement la spécification du modèle. L'analyse des potentiels de commerce présente des intérêts variés. Elle permet d'abord de vérifier si les exportations d'un pays avec ses partenaires sont en ligne avec la structure gravitaire mondiale du commerce. Une déviation par rapport à cet équilibre peut alors renseigner sur l'existence d'autres déterminants dans la relation commerciale bilatérale. Elle permet également d'identifier les pays vers lesquels il existe des gains commerciaux potentiels et d'éclairer les discussions sur la nature de la politique publique à mettre en œuvre pour les obtenir. Enfin, elle permet de constater et de simuler les effets de politiques commerciales sur les flux commerciaux.

¹ Ces modèles sont par exemple souvent utilisés pour évaluer les gains d'accords de libre-échange. Plus récemment, l'approche « gravitaire » a été mise en œuvre pour chiffrer les pertes sur le commerce et l'activité d'une sortie du Royaume-Uni de l'Union européenne (HM Treasury, 2016).

² Cf. Arthur J. et A. Dray (2017), « Potentiels de commerce : quelle stratégie pour le commerce extérieur ? », *Lettre Trésor-Éco* n° 212.

Introduction

Ce document de travail a pour objectif de décrire le modèle utilisé par la Direction générale du Trésor, de justifier les choix réalisés dans sa construction et de permettre la reproductibilité des travaux. Nous proposons également des pistes de réflexion liées aux méthodes d'estimation du modèle, à sa déclinaison au niveau sectoriel, à la variable décrivant la distance entre les pays, ou encore à la notion de « potentiel de commerce ».

Un modèle de gravité permet d'améliorer la compréhension des déterminants structurels des échanges commerciaux. L'équation de gravité décrit une relation entre la valeur de flux bilatéraux de commerce et un certain nombre de variables exogènes telles que la distance entre deux économies, leur taille, la présence d'une langue commune, etc. Sa simplicité, sa robustesse et sa polyvalence ont fait du modèle de gravité un outil très fréquemment utilisé pour analyser les déterminants du commerce international, par exemple l'évaluation ex-post ou ex-ante d'accords de libre échange ou la quantification de barrières commerciales non-tarifaires.

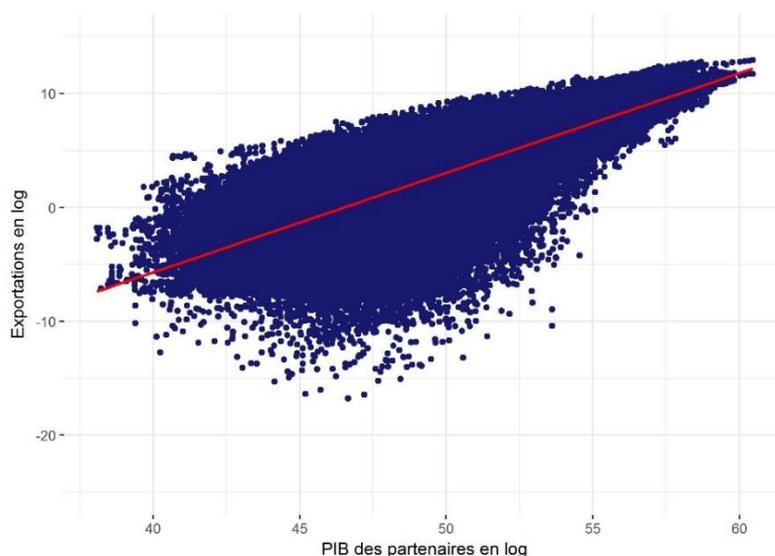
On parle de modèle de gravité par analogie à la théorie newtonienne de la gravitation. Le modèle propose en effet d'expliquer les flux bilatéraux de commerce par des « forces d'attraction » telles que la distance ou la taille économique des partenaires. Selon la littérature, les estimations de modèles de gravité sont celles qui sont les plus convaincantes au regard des hypothèses formulées *ex-ante* sur la relation théorique expliquant les flux de commerce par un ensemble de déterminants (Leamer et Levinsohn, 1995). Les fondements théoriques du modèle de gravité ont toutefois été développés plus récemment, avec notamment l'apparition de modèles dits « structurels » et de nouvelles méthodes d'estimation.

Un modèle de gravité simple peut s'appuyer sur un faible nombre de variables explicatives des flux de commerce, les déterminants principaux étant la taille économique des partenaires (approchée par le PIB) et la distance entre les pays (comme proxy des coûts de transport). Ces indicateurs étant facilement accessibles, il n'existe pas d'obstacles majeurs à la mise en œuvre d'une forme simple d'équation de gravité. De plus, l'amélioration continue de la disponibilité et de la qualité des données nécessaires, notamment pour les pays en développement, permet de renforcer la pertinence et la couverture du modèle.

L'importance prise par les modèles de gravité repose principalement sur le pouvoir explicatif élevé de ce type d'approche. La relation empirique entre le PIB ou la distance avec les échanges commerciaux entre deux pays est particulièrement stable dans le temps et donne, même à des niveaux faibles de raffinement, des résultats intéressants sur les flux commerciaux théoriques. Le coefficient de corrélation entre exportations et PIB des partenaires est significatif et supérieur à 50 %.

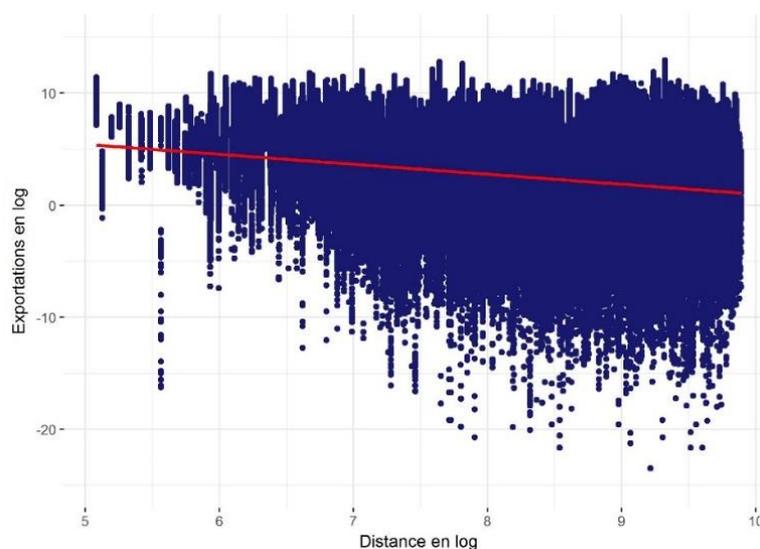
Utilisé dès le début du XX^e siècle dans l'analyse des flux migratoires (Ravenstein, 1885), le modèle de gravité s'applique aujourd'hui aux échanges de biens, de services et d'investissements directs à l'étranger. La littérature économique ayant recours à ce type d'estimation est particulièrement fournie. En 2008, Disdier et Head recensaient 103 publications et plus de 1 000 estimations d'équations de gravité. En 2013, Head et Mayer recensaient 159 publications et plus de 2 500 estimations. Récemment, les évaluations d'impact de la remise en cause de l'espace Schengen (Davis et Gift, 2014) ou de la sortie du Royaume-Uni de l'UE (HM Treasury, 2016) reposaient sur des équations de gravité.

Graphique 1 : Régression par MCO des exportations sur le PIB des partenaires commerciaux



Source : Données CHELEM, calculs DG Trésor.

Graphique 2 : Régression par MCO des exportations sur la distance géographique entre les partenaires commerciaux



Source : Données CHELEM, calculs DG Trésor.

Les résultats issus d'un modèle de gravité doivent être interprétés avec précaution. Les modèles décrits dans ce document de travail sont statiques. On parle de modèle statique car les effets simulés à partir du modèle estimé le sont toutes choses égales par ailleurs. Notamment, on ne tient pas compte de l'impact d'une variable explicative sur les autres, quand il existe, ainsi que de l'impact macroéconomique plus général, par exemple l'effet d'une hausse du commerce sur la productivité ou les prix, bien documentés par ailleurs dans la littérature. Le cadre de l'analyse est donc celui de l'**équilibre partiel**. En l'occurrence, il s'agit donc uniquement d'apprécier l'impact sur les exportations d'un pays vers un autre d'une modification des « coûts de commerce » (pris au sens large), toutes étant choses égales par ailleurs.

L'intégration de termes de résistance multilatérale dans les équations et les méthodes d'estimation récentes (cf. parties 1 et 2) permettent cependant d'assurer certaines contraintes de bouclage du modèle et un lien avec les théories d'équilibre général du commerce. Les résultats obtenus, à partir de données riches (partie 3) sont en ligne avec la littérature (partie 4). Ces résultats sont également déclinés à un niveau sectoriel (partie 5).

1. Les fondements théoriques des modèles de gravité : vers une approche « structurelle »

L'approche gravitaire a fait l'objet d'apports académiques importants depuis 1995. Deux formes d'équations de gravité sont présentées ci-dessous : une équation « intuitive », initialement mise en œuvre par Tinbergen (1962), et une spécification davantage structurelle, tirée de Anderson et van Wincoop (2003).

1.1 Les limites de l'équation de gravité à la Tinbergen

Le choix des variables intégrées à une équation de gravité a longtemps reposé sur la recherche *a priori* de facteurs influençant positivement ou négativement les flux commerciaux.

L'équation de gravité de Tinbergen (1962) est la première équation de gravité estimée. Elle mettait en relation les exportations Y du pays i vers le pays j , avec les PNB respectifs, la distance entre les partenaires commerciaux ($Dist$, en milliers de miles nautiques), une variable indicatrice renseignant sur la présence d'une frontière ($Front$) et une variable indicatrice de politique commerciale ($Pref$) selon que les biens du pays i recevaient un traitement préférentiel par le pays j :

$$\ln(Y_{ij}) = \alpha + \beta_1 \ln(PNB_i) + \beta_2 \ln(PNB_j) + \beta_5 \ln(Dist_{ij}) + \beta_6 Front_{ij} + \beta_7 Pref_{ij} + \varepsilon_{ij}$$

On parle de modèle de gravité « intuitif » dans la mesure où l'équation ne repose pas sur des fondements théoriques *a priori* mais sur les intuitions suivantes : (i) l'intégration d'une économie dans le commerce international est fonction de son poids économique (mesuré principalement par son PIB ou son PNB) et (ii) l'intensité des échanges entre deux économies est fonction de leur proximité géographique.

L'équation de Tinbergen a un pouvoir explicatif élevé, mais elle rencontre plusieurs limites :

- L'absence de variable décrivant les relations entre l'exportateur (l'importateur) et le reste du monde ne permet pas d'incorporer l'effet d'une hausse ou d'une baisse des coûts de commerce avec les pays tiers. Si le coût de commerce avec le reste du monde pour i et j augmente (par exemple dans le cas d'un accord de libre-échange entre i et j , voir Viner (1950)), les économies i et j devraient *a priori* commercer davantage entre elles.
- Une modification du niveau d'exportations peut intervenir sans modification des coûts de commerce bilatéraux par rapport au reste du monde. Par exemple une hausse des cours du pétrole peut se traduire par une baisse de l'ensemble des échanges bilatéraux, du fait d'une hausse généralisée des coûts de fret. De manière générale, l'absence d'effets fixe exportateur/importateur ne permet pas de rendre compte de la variation de facteurs transversaux ayant une influence sur l'ensemble des échanges commerciaux.

1.2 Anderson et Van Wincoop (2003) proposent un modèle de gravité fondé théoriquement, tenant compte de l'influence des marchés tiers sur les relations commerciales bilatérales

Le choix des données et des spécifications et méthodes d'estimation des équations de gravité reposent de plus en plus sur des fondements microéconomiques. Plusieurs approches théoriques sont compatibles avec le modèle de gravité, par exemple le modèle de concurrence monopolistique à la Krugman (Wei, 1996) ou les avantages comparatifs ricardiens (Eaton et Kortum, 2002).

Nous présentons ici le modèle d'Anderson et Wincoop (2003), qui tient compte théoriquement de l'influence des pays tiers sur la relation commerciale bilatérale entre deux économies i et j . Les « termes de résistance multilatérale » introduits par les auteurs permettent de rendre compte notamment de l'effet suivant : si le reste du monde présente de nombreuses barrières au commerce avec des économies i et j , alors ces deux économies auront davantage tendance à commercer entre elles qu'avec le reste du monde. Ce modèle permet également de répondre à la limite présentée *supra* sur l'impact d'un facteur exogène transversal qui affecterait les flux commerciaux bilatéraux. Au niveau économétrique, la mauvaise spécification du modèle peut notamment entraîner un biais de variables omises.

Anderson et Wincoop dérivent l'équation de gravité d'une fonction de demande à élasticité de substitution constante et biens différenciés par origine, reposant sur l'hypothèse d'Armington, (1969) de substituabilité imparfaite entre des biens locaux et importés. Dans ce cadre, la fonction d'utilité du consommateur (pays importateur, j) est la suivante,

$$(1) \quad \left(\sum_i \alpha_i^{\frac{1-\sigma}{\sigma}} c_{ij}^{\frac{\sigma-1}{\sigma}} \right)^{\frac{\sigma}{\sigma-1}}$$

Où, $\sigma > 1$ est l'élasticité de substitution constante entre les variétés de biens (dans le cadre de l'hypothèse d'Armington, les biens sont différenciés selon leur provenance), α est le paramètre de préférence ou la propension à consommer une variété de biens, et c la consommation du pays j provenant du pays i . Le consommateur maximise cette fonction d'utilité sous contrainte budgétaire.

Les auteurs montrent que les exportations Y_{ij} s'écrivent (cf. présentation du modèle détaillé en annexe) :

$$(2) \quad Y_{ij} = \frac{PIB_i PIB_j}{PIB} \left(\frac{\tau_{ij}}{\Phi_i \Phi_j} \right)^{1-\sigma}$$

Les exportations dépendent d'un terme lié à la taille économique relative des partenaires ($\frac{PIB_i PIB_j}{PIB}$), et d'un terme regroupant l'ensemble des coûts de commerce $\left(\frac{\tau_{ij}}{\Phi_i \Phi_j} \right)^{1-\sigma}$. τ_{ij} représente les barrières au commerce entre deux partenaires (liens historiques, langue commune, frontière commune, etc.) à l'instar de l'équation de Tinbergen. Deux nouvelles variables s'y greffent, appelées par les auteurs « termes de résistance multilatérale » : Φ_j représente ici la résistance multilatérale s'imposant à l'importateur, Φ_i celle de l'exportateur.

Ainsi la première partie de l'équation décrit un monde sans friction, alors que la seconde partie de l'équation éclaire sur les coûts de commerce bilatéraux (qui décrivent les barrières propres à la relation des partenaires) et les coûts multilatéraux (qui sont les barrières tenant compte de l'ensemble des coûts bilatéraux avec les pays tiers).

Ces termes de résistances s'écrivent :

$$\Phi_i = \left[\sum_k \frac{PIB_k}{PIB} (\Phi_k^{\sigma-1} \tau_{ki}^{1-\sigma}) \right]^{\frac{1}{1-\sigma}}$$

Elles dépendent des coûts de commerce bilatéraux de i et j avec l'ensemble des partenaires commerciaux $\tau_{k,[i,j]}$, des poids de l'ensemble des partenaires commerciaux ainsi que de leur résistance multilatérale. Les termes de résistance multilatérale $\Phi_{i,j}$ dépendent en outre des coûts de commerce bilatéraux entre des économies k, l , via les termes de résistance multilatérale $\Phi_k, k = 1 \dots n$. Ainsi, une variation des coûts de commerce entre deux économies k, l quelconques affectent la relation commerciale entre i et j .

Les résistances Φ sont des fonctions croissantes de τ : quand les coûts de commerce bilatéraux augmentent avec les pays tiers, les termes de résistances augmentent et les échanges commerciaux entre i et j augmentent également (on a $\frac{\partial y}{\partial \phi} > 0$), car les coûts de commerce relatifs entre i et j baissent.

La log-linéarisation de (2) donne l'équation suivante :

$$(3) \quad \ln(Y_{ij}) = -\ln(PIB) + \ln(PIB_i) + \ln(PIB_j) + (1 - \sigma)(\ln(\tau_{ij}) - \ln(\Phi_j) - \ln(\Phi_i)) + \varepsilon_{ij}$$

1.3 Des estimations avec effets fixes permettent de capter les termes de résistance multilatérale

Lorsque des termes de résistance multilatérale sont présents dans une équation de type gravitaire, le modèle est dit « structurel » du fait des fondements économiques associés à ce type de facteur, faisant dépendre la relation commerciale entre deux pays des coûts de commerce présents entre ces deux partenaires et le reste du monde. Ces termes de résistance ne sont pas directement observables mais il existe plusieurs méthodes permettant de les estimer.

Anderson et van Wincoop (2003) ont recours à des estimations par moindres carrés non linéaires et Head et Mayer (2014) à une méthodologie itérative conjuguant l'estimation par moindres carrés, effets fixes et l'imposition de contraintes sur les termes de résistance multilatérales (*structurally iterated least squares*). Bien que théoriquement adéquates, ces approches sont peu aisées à mettre œuvre.

Depuis Harrigan (1996), la littérature a principalement recours à des effets fixes pour prendre en compte les termes de résistance multilatérale. Deux types d'effets fixes sont utilisés :

- i. **Des effets fixes importateurs – années et exportateurs – années.** Ils ont l'avantage de contrôler la tendance de certains couples de pays à commercer de façon importante relativement à leur poids économique.
- ii. **Des effets fixes couple de pays,** qui permettent d'absorber les coûts de commerce invariants dans le temps entre deux partenaires, qu'ils soient observables ou non. Pour cette raison un certain nombre de coefficients tels que la distance, la langue commune ou encore le système juridique commun disparaissent des tables de résultats, et leurs effets sont intégrés au coefficient de l'effet fixe (*cf.* encadré 3 pour une discussion sur l'évaluation de l'impact des accords de libre-échange notamment).

2. Estimer une équation de gravité : MCO, moindres carrés non linéaires et estimateurs du pseudo maximum de vraisemblance

Dans cette partie, nous expliquons en quoi l'estimation d'une élasticité par MCO après log-linéarisation du modèle est généralement biaisée en présence d'hétéroscédasticité, une remarque qui s'applique à l'ensemble des modèles *Constant Elasticity of Substitution* (CES) log-linéarisés. Nous rappelons le principe d'estimation par pseudo maximum de vraisemblance et décrivons les propriétés de l'estimateur du pseudo maximum de vraisemblance de Poisson, désormais largement employé pour estimer les équations de gravité. Nous notons indifféremment en lettre minuscule les variables aléatoires et les observations.

2.1. L'estimation MCO d'un modèle de gravité log-linéarisé sur des données de commerce est généralement biaisée

Les paramètres de l'équation de gravité issus du modèle de Tinbergen sont fréquemment estimés par MCO sur le logarithme des données de commerce :

$$\begin{aligned} \ln(y_i) &= \beta x_i + \varepsilon_i \\ y_i &> 0, i = 1, \dots, n \\ \mathbb{E}[\ln(y_i)|x_i] &= \beta x_i \text{ avec } \mathbb{E}[\varepsilon_i|x_i] = 0 \end{aligned} \tag{1}$$

Cette transformation implique d'exclure les flux de commerce égaux à zéro de l'échantillon. En outre, déterminer l'impact de la variable x_i sur la variable y_i nécessite de retransformer le modèle. Le paramètre estimé par MCO ne donne en effet que l'impact au premier ordre d'une variation de x_i sur l'espérance du logarithme de y_i . À partir de (1) le modèle non transformé s'écrit :

$$\begin{aligned} y_i &= e^{\beta x_i + \varepsilon_i} \\ \mathbb{E}[y|x_i] &= e^{\beta x_i} \mathbb{E}[e^{\varepsilon_i}|x_i] \end{aligned} \tag{2}$$

On obtient ainsi les deux résultats suivants :

- l'impact d'une variation de x_i sur les flux de commerce n'est de manière générale pas égal à l'estimateur de β obtenu par les MCO sur la version log-linéarisée du modèle. Par l'inégalité de Jensen, on a en effet $\mathbb{E}[e^{\varepsilon_i}|X_i] \geq e^{\mathbb{E}[\varepsilon_i|X_i]} = 1$.
- si $\mathbb{E}[e^{\varepsilon_i}|x_i]$ ne dépend pas de x_i et peut être estimé, l'estimateur sans biais de l'impact au premier ordre est donné par $\frac{\partial \mathbb{E}[y_i|x_i]}{\partial x_i} = \hat{\beta} e^{\hat{\beta} x_i} e^{\varepsilon_i}$.

$\mathbb{E}[e^{\varepsilon_i}|x_i]$ dépend de x_i lorsqu'il y a présence d'hétéroscédasticité dans le modèle log-linéaire. Dans le cas étudié ici d'un modèle de commerce, la condition $y_i > 0$ permet de voir que les résidus ε_i sont corrélées aux variables explicatives. Quand y_i tend vers 0, $\beta x_i + \varepsilon_i$ tend nécessairement vers $-\infty$. Pour des valeurs élevées de y_i il n'y a cependant aucune contrainte sur le terme d'erreur et celui-ci présente *a priori* une variance élevée compte tenu de l'hétérogénéité observée dans les données de commerce³.

Encadré 1 : Impact de l'hétéroscédasticité sur les estimations MCO des paramètres d'un modèle log-linéarisé

On se place dans le cas d'un modèle transformé où les résidus suivent une loi normale avec présence d'hétéroscédasticité :

$$\begin{aligned} \ln(y_i) &= \beta x_i + \varepsilon_i \\ y_i &> 0, i = 1, \dots, n \\ \mathbb{E}[\ln(y_i)|x_i] &= \beta x_i \text{ avec } \varepsilon_i = \ln(v_i) \sim \mathcal{N}(0, \sigma(x_i)^2) \end{aligned} \quad (3)$$

Le modèle non-transformé s'écrit, avec v_i suivant une loi log-normale :

$$\mathbb{E}[y_i|x_i] = e^{\beta x_i} \mathbb{E}[e^{v_i}|x_i] = e^{\beta x_i + \frac{\sigma(x_i)^2}{2}}$$

L'impact au premier ordre de x_i sur y_i est proportionnel au terme $\beta + \sigma(x_i) \frac{d\sigma}{dx_i}$. L'estimateur de l'impact au premier ordre obtenu par MCO sur le modèle transformé est donc négativement biaisé si la volatilité des résidus augmente avec la valeur des observations, positivement biaisé sinon. L'hétéroscédasticité implique ici notamment que $\frac{\partial \ln \mathbb{E}[y_i|x_i]}{\partial x_i} \neq \frac{\partial \mathbb{E}[\ln(y_i|x_i)]}{\partial x_i}$, et l'estimation de β devrait donc être corrigée d'un terme liée à l'hétéroscédasticité pour obtenir un estimateur sans biais des paramètres d'intérêt.

2.2. L'absence de flux de commerce entre de nombreux pays pose également problème dans une estimation par MCO

Les flux nuls, particulièrement fréquents dans les bases de données de commerce international⁴, sont ignorés du fait de la transformation logarithmique. Toutefois, deux situations doivent être distinguées :

- Dans certains cas on peut considérer qu'il n'y a pas de perte d'information à la suppression de l'observation. Il s'agit des situations où les données manquantes résultent d'un défaut de construction de la base de données. Dans ce cas, on admet que les flux nuls sont distribués de façon aléatoire et leur

³ La matrice de variance-covariance des résidus n'est pas proportionnelle à la matrice identité.

⁴ Par exemple, la base de données DOTS du FMI contient sur un échantillon de 169 pays entre 1967 et 2014 environ 40 % de flux nuls.

suppression n'implique pas de perte d'information. *A priori* cette situation est rare et l'utilisation de sources alternatives de données apparaît pertinente.

- Dans d'autres cas, les valeurs nulles sont dues à un arrondi des valeurs faibles des flux ou encore à l'absence véritable de flux commerciaux entre deux pays. La situation est alors problématique puisque la spécification économétrique en elle-même occulte des informations importantes. La prise en compte de situations où deux pays n'échangent pas ajoute un pouvoir explicatif aux flux commerciaux. Deux stratégies permettent d'éviter ce problème : ajouter une constante (représentant un niveau minimal d'échange) ou estimer le modèle en niveau.

2.3. L'estimateur des moindres carrés non linéaires n'est généralement pas satisfaisant pour estimer les paramètres d'un modèle de commerce

Les exemples ci-dessus montrent que les élasticités d'un modèle de commerce sont généralement biaisées si elles sont estimées *via* les MCO sur la spécification log-linéaire. Manning et Mullahy (2001) étudient les performances d'un ensemble d'estimateurs sur des données à distributions variables en fonction de paramètres d'asymétrie et d'épaisseur. Dans cette partie nous nous plaçons dans le cas de données de commerce avec présence d'hétéroscédasticité.

La minimisation des termes d'erreur du modèle non transformé donne l'estimateur des moindres carrés non linéaires :

$$\hat{\beta} = \underset{\beta}{\operatorname{argmin}} \|y_i - e^{\beta x_i}\|^2$$

Les conditions du premier ordre s'écrivent :

$$\sum_{i=1}^n (y_i - e^{\beta x_i}) e^{\beta x_i} x_i = 0$$

L'efficacité de cet estimateur⁵ est généralement faible pour des données présentant des distributions asymétriques. Dans le cas de données de commerce, qui présentent des distributions très asymétriques avec valeurs extrêmes, l'estimateur des moindres carrés non linéaires fournit des estimations généralement peu précises des paramètres. Le critère étant la minimisation des carrés de l'erreur entre les données observées et les données estimées, la présence de valeurs extrêmes peut avoir tendance à mal estimer le modèle autour des observations les plus denses de l'échantillon (ce problème dit d'« *overfitting* » est normalement atténué en log-transformant les observations de la variable endogène, *modulo* les problèmes de biais des estimations soulevés ci-dessus).

2.4. Les estimateurs par pseudo maximum de vraisemblance (PML) basés sur des lois exponentielles linéaires permettent d'obtenir des estimations sans biais des paramètres d'intérêts

Le problème à résoudre consiste à estimer un paramètre d'une fonction non-linéaire sur des données hétéroscédastiques. La loi de chaque observation est *a priori* inconnue, mais le moment d'ordre 1, ici $e^{\beta x_i}$, est connu. Le modèle est dit spécifié au premier ordre. Hormis le fait que la variance de y_i tend vers 0 quand x_i tend vers 0, la forme de l'hétéroscédasticité est également inconnue.

L'approche *via* le pseudo maximum de vraisemblance (PML) consiste à estimer le paramètre d'intérêt β en maximisant la vraisemblance des observations, obtenue *via* une loi de même espérance (la loi est dite ajustée au premier ordre), mais qui n'est pas celle des observations. On appelle cette loi une « pseudo-loi ». Toutes les pseudo-lois ne fournissent pas des estimateurs convergents. Cependant, Gourieroux, Montfort et Trognon

⁵ Fonction *nl* de Stata.

(1984) ont montré que tout estimateur du PML basé sur une loi appartenant à la famille des lois dites exponentielle linéaire⁶ convergeait presque sûrement vers la vraie valeur du paramètre d'intérêt.

Pour des lois de la famille exponentielle linéaire, les conditions PML du premier ordre s'écrivent :

$$\sum_{i=1}^n \frac{\partial f}{\partial \beta}(x_i, \beta) \mathbb{V}(y_i | x_i)^{-1} (y_i - f(x_i, \beta)) = 0$$

Où $\mathbb{V}(y_i | x_i)$ est la variance de y_i sachant x_i calculée avec la pseudo-loi.

Silva et Teneyro (2006) examinent trois estimateurs PML, calculés avec des pseudos-lois de type Normale, Poisson et Gamma. Les propriétés de ces lois et des estimateurs associés sont rappelées dans le tableau 1, dans le cas d'un modèle à une variable exogène et un paramètre d'intérêt (les équations ont la même forme dans le cas vectoriel à k variables exogènes).

L'estimateur GPML a ainsi tendance à ne pas tenir compte des observations pour lesquelles x_i est élevé (présence d'un terme $e^{-\beta x_i}$ additionnel par rapport aux CPO PPML). À l'inverse, l'estimateur Normale-PML a tendance à surpondérer les observations où x_i est grand (présence d'un terme additionnel $e^{\beta x_i}$). L'estimateur PPML traite chaque observation de manière identique relativement aux deux autres estimateurs. Intuitivement, l'estimateur PPML ne discrimine donc pas au sein de l'échantillon des variables exogènes, ce qui semble particulièrement adapté aux données de commerce qui présentent des distributions fortement asymétriques, avec présence d'un grand nombre de valeurs extrêmes (flux nuls, ou flux bilatéraux très supérieurs à la moyenne de l'échantillon).

En simulant un grand nombre de jeux de données, Silva et Teneyro (2006) comparent ces trois estimateurs, en plus d'estimations conduites sur un modèle log-transformé avec les MCO. L'estimateur PPML ressort comme l'estimateur présentant les propriétés les plus intéressantes pour des paramètres d'intérêt présents dans des fonctions CES dès lors qu'il y a présence d'hétéroscédasticité. L'estimateur PPML est celui qui converge le mieux vers les vraies valeurs du paramètre. Sans hétéroscédasticité, les MCO sur données log-transformées demeurent la méthode d'estimation la plus appropriée.

Les travaux de Silva et Teneyro ont entraîné un changement de paradigme dans l'estimation de modèles de gravité, l'utilisation de l'estimateur PPML semblant désormais faire consensus au sein de la communauté académique. Head et Mayer (2014) recommandent ce type d'estimateur pour estimer des équations de gravité. Nous retenons donc cette approche dans le cadre de ce document de travail.

⁶ Ces lois ont une densité de la forme $l(y, m) = \exp(A(m) + B(y) + C(m).y)$. Il s'agit notamment des lois : gaussienne, poisson, gamma, binomiale, etc. Gourieroux, Montfort et Trognon (1984) fournissent les conditions du premier ordre et l'expression des variances des estimateurs PML associés à ces lois.

Tableau 1 : Estimateurs PML et conditions associées pour les pseudo-lois Gamma, Poisson et Normale ($f(x_i, \beta) = e^{\beta x_i}$)

Estimateur	Densité de la Pseudo-Loi ajustée au premier ordre ($\mathbb{E}[y_i x_i] = m_i = e^{\beta x_i}$)	Variance	CPO	Concavité ⁷
Gamma-PML (ci-après GPML)	$\frac{\exp(-\frac{ay_i}{m_i})y_i^{a-1}}{\Gamma(a)(\frac{m_i}{a})^a}, a > 0$	$\mathbb{V}[y_i x_i] = a^{-1}\mathbb{E}[y_i x_i]^2$	$\sum_{i=1}^n x_i(y_i - e^{\beta x_i})e^{-\beta x_i} = 0$	$\sum_{i=1}^n -ax_i^2 y_i e^{-\beta x_i}$
Poisson-PML (ci-après PPML)	$\frac{\exp(-m_i)m_i^{y_i}}{y_i!}$	$\mathbb{V}[y_i x_i] = \mathbb{E}[y_i x_i]$	$\sum_{i=1}^n x_i(y_i - e^{\beta x_i}) = 0$	$\sum_{i=1}^n -x_i^2 e^{-\beta x_i}$
Normale-PML (ci-après NPML)	$\frac{1}{\sqrt{2\pi}\sigma} \exp(-\frac{(x_i - m_i)^2}{2\sigma^2})$	$\mathbb{V}[y_i x_i] = \sigma^2$	$\sum_{i=1}^n x_i(y_i - e^{\beta x_i}) e^{\beta x_i} = 0$	$\frac{1}{\sigma} \sum_{i=1}^n x_i^2 (y_i e^{\beta x_i} - 2e^{2\beta x_i})$

Note : l'estimateur PML obtenu via une pseudo-loi Normale est le même que l'estimateur des moindres carrés non linéaires (les CPO sont les mêmes que celles présentées en 2.3). La pseudo vraisemblance n'est en général pas concave en β , contrairement aux deux autres estimateurs. La présence d'un maximum de la PML n'est donc pas garantie. Il est important de remarquer également que ces trois pseudo-lois permettent de tester trois types de spécification du lien entre variance et espérance : (i) la variance est proportionnelle à une constante (estimateur NPML) ; (ii) la variance est proportionnelle à l'espérance (estimateur PPML) ; (iii) la variance est proportionnelle au moment d'ordre 2 (estimateur GPML).

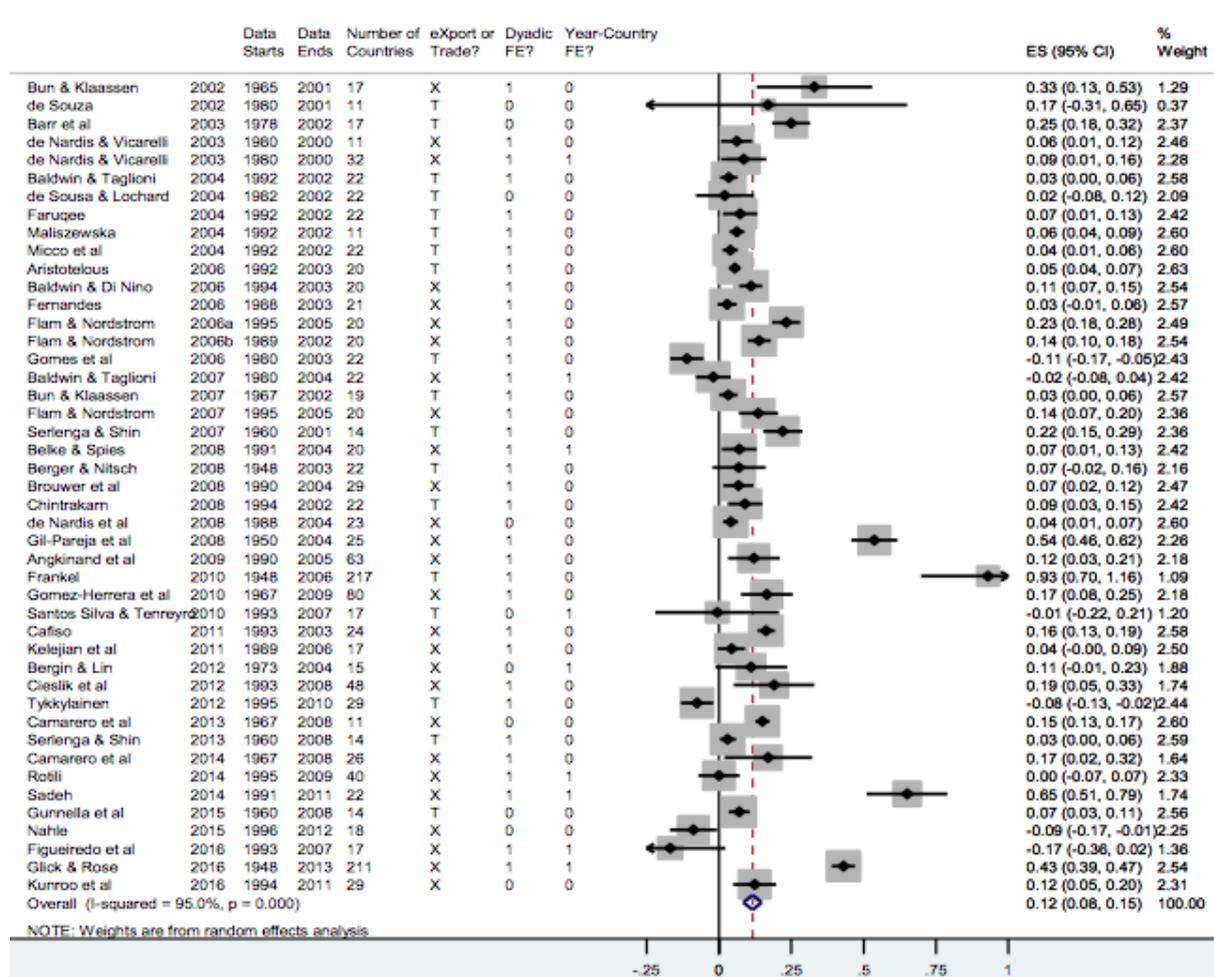
⁷ Un maximum existe si la dérivée seconde de la pseudo-vraisemblance est négative. Ce critère est respecté dans tous les cas pour les pseudo-lois Gamma et Poisson. La convergence de l'estimateur en utilisant une pseudo loi Normale n'est pas assurée dans tous les cas. Les conditions de convergence de ce dernier estimateur sont par ailleurs plus mauvaises (la variance des paramètres estimés est généralement plus élevée).

3. Le choix des données dans la construction du modèle de gravité

Baldwin et Taglioni (2006) ont résumé les trois principales erreurs pouvant être commises lors de l'estimation de modèles de gravité. Selon les auteurs, compte tenu des hypothèses exposées *supra*, un modèle de gravité structurel doit (i) être estimé en tenant compte des termes de résistance multilatérale, (ii) être estimé sur des données d'exportations ou d'importations (mais pas sur des échanges moyens ou totaux) ; (iii) exprimées en termes nominaux (étant déflatées par les termes de résistance).

En outre, une critique particulièrement forte adressée à l'utilisation de modèles de gravité réside dans la grande variété de résultats obtenus par la littérature dans l'analyse de politiques commerciales similaires. Rose (2016) recense les résultats de 45 estimations de l'effet de l'euro sur les exportations des pays membres de la zone. Les coefficients varient de -0,2 à près de 1.

Graphique 3 : Résultats des estimations de l'effet de l'euro sur les échanges commerciaux



Source : Rose (2016), Université de Berkeley

Selon l'auteur, il est important d'utiliser le panel le plus large possible (le plus grand nombre de pays et d'années disponibles) à partir d'une équation de gravité structurelle afin de capter le maximum d'informations. En l'occurrence il relève que l'impact de l'euro sur le commerce est biaisé vers le bas lorsque les petits pays sont évincés de l'échantillon. Suivant les recommandations de Rose, deux bases de données sont construites et utilisées pour tester la robustesse du modèle au choix des données :

- Un modèle « super-mondial » incluant les données de 189 pays depuis 1949 (modèle I)
- Un modèle mondial : avec des données harmonisées couvrant la période 1967-2014 et 84 pays (modèle II)

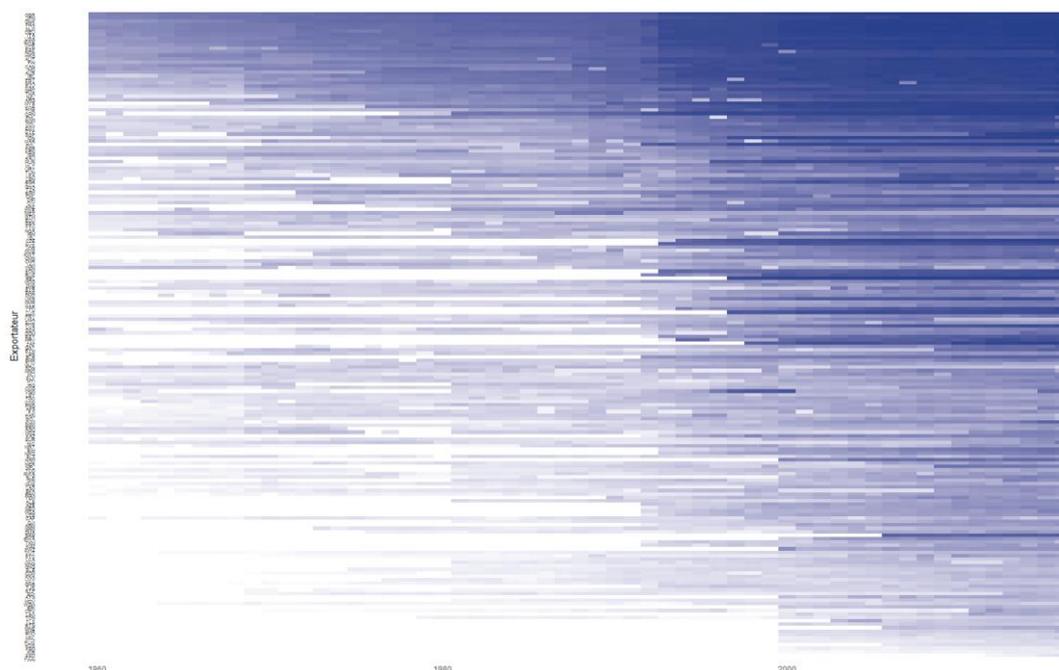
Le modèle intègre l'ensemble des variables classiquement retenues par la littérature. Il s'agit des données d'exportations en valeur, des données décrivant la taille des partenaires (PIB), ainsi que celles décrivant les coûts de commerce (τ_{ij}), telles que la distance, les droits de douanes, les caractéristiques communes des partenaires, ou encore les mesures unilatérales et non discriminatoires.

3.1. Les exportations sont intégrées en valeur afin de respecter la forme structurelle de l'équation

Les deux bases de données contiennent uniquement des données d'exportations, tel que recommandé par Baldwin et Taglioni (2008). Nous intégrons ces données en valeur afin de ne pas introduire de déflateur exogène.

Le modèle I utilise des flux commerciaux en valeur issus de la base DOTS du FMI complétés par les données COMTRADE (ONU) et des données recueillies par Fouquin et Hugot (2016). Il s'agit de données Franco à Bord (FAB). La couverture temporelle et géographique des données FMI permet une analyse large en accord avec les conclusions de Rose, ainsi qu'une mise à jour simple permettant d'inclure des données très récentes (jusqu'en 2015). La base comprend 189 pays depuis 1949, pour environ 640 000 observations complètes. À l'instar des autres bases de données de commerce international, les données manquantes sont principalement concentrées sur les premières années du panel et certains pays ont été exclus du fait d'une absence totale de données⁸.

Graphique 4 : Couverture de la base de données du modèle I



Note de lecture : pour chaque exportateur (en ordonnée), chaque année (en abscisse) l'existence et le montant des exports entre 1960 et 2016 est représenté par une échelle de couleur allant du blanc (flux nul ou absent) au bleu foncé (flux important).

Source : Calculs DG Trésor, données FMI DOTS et ONU COMTRADE.

Le modèle II repose sur les données issues de l'harmonisation réalisée par le CEPII dans la base CHELEM à des fins de test de robustesse. La couverture géographique et temporelle, de 1967 à 2014 pour 84 pays, est plus réduite mais permet tout de même d'obtenir des résultats satisfaisants. Le principal avantage de cette base est l'absence de flux nuls.

⁸ Il s'agit notamment du Kosovo, des Iles Marshall, Palaos, du Soudan du Sud, Saint Marin et de la Micronésie.

3.2. Le poids économique des partenaires est approché par leur produit intérieur brut (PIB)

Dans l'optique de capter la production des agents économiques résidant du pays exportateur (et non de nationalité du pays exportateur) le produit intérieur brut (PIB) est préféré au Produit national brut (PNB). Les données de PIB en dollars courant proviennent de la base IFS du FMI et de la base WDI de la Banque Mondiale. Nous complétons ces données avec la base de Hugot et Fouquin (2016) pour les observations les plus anciennes. Les mises à jour du modèle seront réalisées à partir des données IFS couvrant l'intégralité des pays étudiés.

L'ajout d'une variable continue décrivant la taille des populations des pays permet de prendre en considération le PIB par tête. Nous utilisons la base WDI de la Banque mondiale pour les données démographiques des pays sur l'ensemble de la période.

3.3. Les coûts de commerce bilatéral sont décrits par un ensemble de variables bien identifiées par la littérature

Deux types de variables sont intégrés à ce vecteur : d'une part des variables continues, et d'autre part des variables indicatrices. Le rôle des variables indicatrices est de décrire les déterminants de long terme des coûts bilatéraux de commerce. Nous reprenons les variables traditionnellement incorporées aux modèles de gravité : données relatives à l'existence de frontière terrestre commune, de langue officielle ou ethnique commune, de liens coloniaux, de système juridique commun, de monnaie commune, d'appartenance à l'OMC et d'accord de libre-échange.

La distance entre les partenaires commerciaux est géodésique mais pourrait à terme être remplacée par une distance logistique.

Les coûts de commerce bilatéraux sont intuitivement une fonction croissante de la distance entre deux économies. La plupart des études mettant en œuvre des modèles de gravité ont recours à des données de distances géodésiques, mesurant des distances « à vol d'oiseau » calculées à partir des longitudes et latitudes de deux points d'intérêt. Ce type de distance est en général construit comme la distance moyenne entre les plus grandes villes d'un pays et celles de ces partenaires. Nous utilisons ici pour les modèles I et II la variable de distance pondérée par la population de la base de données CEPII Distances⁹.

Encadré 2 : Une piste de recherche dans l'évaluation de l'impact commercial de la distance serait de considérer une variable de distance logistique

La distance géodésique souffre de plusieurs limites :

- i. La trajectoire directe n'est jamais la route réellement suivie par les marchandises.
- ii. Elle ne prend pas en compte le poids de chaque mode de transport dans le commerce international alors même que le coût associé à chacun est différent et a évolué de façon différente au XX^e siècle.
- iii. À distance géodésique égale, certains pays peuvent être structurellement soumis à des coûts de transport plus importants que d'autres, en fonction des spécialisations sectorielles et de la situation géographique (pays enclavés par exemple, dont les coûts de transport peuvent être fonction des coûts de transport terrestre et maritime). Une partie de cet effet peut être capté par les effets fixes alors que l'on souhaiterait qu'elle le soit par le coefficient de distance.

Malgré ces limites, la distance géodésique utilisée dans la majorité des modèles de gravité présente un pouvoir explicatif relativement fort.

⁹ Thierry Mayer & Soledad Zignago (2011), "Notes on CEPII's distances measures: The GeoDist database," *CEPII Working Paper* 2011-25.

Des distances logistiques entre pays ont également été calculées. Cette variante a pour objectif de raffiner la variable de distance pour les pays européens où 70 % à 75 % des échanges commerciaux sont transportés par la route et où les distances géodésiques semblent moins pertinentes.

Afin de se rapprocher d'une définition logistique de la distance parcourue par les échanges, on peut utiliser les distances routières et maritimes qui constituent la plus grande partie des modes de transport de biens.

Les distances routières entre les 5 plus grandes villes des 28 pays européens ont été mesurées (API de Google), puis une moyenne simple des distances entre ces villes a été calculée afin d'obtenir des distances routières par couple de pays (D_{ij}^r). Nous avons également utilisé les données de distances maritimes (D_{ij}^m) de la base de données de Fouquin et Hugo (2016). Nous réalisons une simple pondération de la distance bilatérale, par un coefficient m représentant le poids en année t (données de mode de transport Eurostat).

$$D_{ij} = m^r D_{ij}^r + m^m D_{ij}^m$$

Dans un premier temps nous considérons la distance comme une constante et utilisons les moyennes des ratios de mode de transport et de distance maritimes. Une régression simple des échanges sur la distance logistique permet de constater une meilleure corrélation qu'avec la distance géodésique.

La généralisation de cette logique devra aboutir à considérer la distance comme variant dans le temps au sein des couples de pays notamment du fait de l'ouverture de nouveaux passages maritimes et de la variation du poids de chaque mode de transport.

Les droits de douane bilatéraux sont tirés de Hugot et Fouquin (2016) ; la base construite intègre également des données sur les communautés de frontière, de langue, de passé colonial, de système juridique et d'accords commerciaux.

Les droits de douane ont longtemps été une barrière importante au commerce, et l'actualité récente met en exergue l'importance d'une bonne prise en compte de ce paramètre. Toutefois, peu de modèles de gravité incorporent les droits de douane bilatéraux du fait de la difficulté à rassembler ce type de données. Notre base de données intègre ici les droits de douane traités par Hugot et Fouquin (2016). Ces données contiennent plus de 8 000 observations sur les ratios de droits de douane bilatéraux à l'importation. Elles couvrent 410 paires de pays sur toute la période d'estimation.

La variable frontière commune a pour objectif d'indiquer si les partenaires partagent une frontière terrestre. Les données sont issues de la base GEO du CEPII.

La variable indicatrice de langue commune vise également à approcher les coûts de commerce. L'hypothèse sous-jacente étant un niveau plus faible des coûts de transaction entre deux pays parlant la même langue. Nous utilisons les données de Head (2010). La variable indicatrice est active si les deux pays ont au moins un langage commun parlé par au moins 9 % de leur population.

Une indicatrice d'appartenance à l'Union européenne vaut 1 si les deux partenaires sont membres de l'UE l'année t . À noter que l'utilisation d'une variable indicatrice implique que les gains liés à une entrée dans une union commerciale sont symétriques aux pertes associées à une sortie (cf. encadré 3).

Encadré 3 : Les accords de libre-échange et les unions monétaires sont vraisemblablement endogènes aux liens commerciaux existant entre les pays

Rose (2000) estime à l'aide d'une équation de gravité que les membres d'une union monétaire échangent en moyenne de 3 à 4 fois plus. Ce fort impact des unions commerciales et douanières a fait l'objet d'une abondante littérature, dans la mesure où l'endogénéité n'était pas traitée dans ces études. La question est ainsi de savoir si des accords commerciaux ou monétaires sont conclus entre des pays ayant déjà une forte propension à commercer entre eux. *A priori*, de tels accords peuvent rarement être considérés comme entièrement exogènes. Il semble donc important d'étudier les déterminants

des ALE ou des unions monétaires, ce qui semble avoir été peu le cas jusqu'à présent, ainsi que le rappellent Baier et Bergstrand (2007).

Baier et Bergstrand (2007) estiment que l'impact des ALE est sous-estimé. Un biais d'estimation important proviendrait entre autres de variables omises. S'ils présentent des gains à l'intégration commerciale, deux pays seraient d'autant plus incités à signer un accord de libre-échange qu'ils présentent des niveaux de régulation interne importants. Ce type d'accord s'accompagne généralement de mesures plus larges de libéralisation et de renforcement de la compétitivité, et les ALE comprennent un volet non-tarifaire de plus en plus important.

Pour traiter les biais de sélection inhérents aux accords de libre échange (deux pays peuvent signer un ALE parce qu'ils commercent déjà beaucoup), Wolf et Ritschl (2011) se placent dans un contexte d'expérience naturelle avec un groupe de test (l'échantillon « traité ») comprenant les échanges commerciaux entre pays appartenant à des unions monétaires formées après la chute de l'étalon or. Le groupe de contrôle est constitué des données de commerce préalable à la chute de l'étalon or. L'impact diffère sensiblement des estimations initiales de Rose. Frankel (2010) utilise également une évaluation de type expérience naturelle. Avec l'adoption de l'euro, les pays de la zone franc ont subi un choc que l'on peut considérer exogène d'unification monétaire avec l'ensemble des pays de la zone euro, alors que l'union monétaire ne concernait auparavant que la zone franc. L'introduction de l'euro aurait entraîné une hausse d'environ 50 % des échanges entre les pays de la zone franc et la zone euro.

Glick et Rose (2002) et Head et al. (2010) estiment l'équation sur des données de panel et introduisent des effets fixes dyadiques, c'est-à-dire par paires de pays. Ces effets fixes sur les données en coupe, autorisés à varier chaque année (estimateur *within*), permettent de contrôler des effets liés à la relation bilatérale et variant dans le temps, limitant ainsi les risques de biais relevant de variables omises. Nous retenons cette approche dans nos estimations (cf. *infra*). Si ce type d'effets fixes permet de forcer l'estimation sur une constance identique par paire de pays, indépendamment du fait que les deux économies aient un accord de libre-échange, l'estimation peut souffrir de problème de surestimations. Les écarts-types du paramètre d'appartenance à une union monétaire apparaissent fréquemment sous-évalués, si bien que la robustesse du paramètre doit être considérée avec précaution (une approche par expérience naturelle, tel qu'explicitée ci-dessus, permettrait de résoudre ce problème de surestimation dans le cadre d'une étude spécifiquement dédiée à l'étude de l'impact des ALE sur le commerce, ce qui n'est pas l'objet des travaux présentés ici).

Lochard (2005) relève la présence d'un biais d'agrégation dans l'équation de Rose, lié au traitement homogène d'unions monétaires ayant peu à voir en pratique, par exemple la sélection d'unions monétaires multilatérales, unilatérales (cas du Panama ou d'autres économies d'Amérique Latine ayant adopté le dollar américain), ou encore la prise en compte d'anciennes colonies ayant conservé la monnaie de leur colonisateur – si ces différents types d'unions monétaires ont des effets différents sur le commerce, leur traitement homogène biaise les résultats en ce qui concerne les unions monétaires multilatérales telles que la zone euro.

Enfin Wolf et Ritschl (2011) soulignent l'importance du choix des données dans la mesure de ce type d'effet. Ces auteurs soulignent ainsi que les échantillons retenus par Rose, de même que de nombreux autres travaux de ce courant de la littérature, couvrent essentiellement des périodes de disparition d'unions monétaires, du fait de la prise d'indépendance de certains pays de l'échantillon. Ceci implique que la disparition des unions monétaires est traitée symétriquement à la formation des unions monétaires, ce qui n'est *a priori* pas justifié, les unions ayant une forte probabilité d'être démantelées des suites d'une crise monétaire ou politique ayant déjà entraîné une réduction des échanges.

Les estimations les plus récentes établissent un impact moyen de l'ordre de 50 % pour les ALE (cf. les estimations et la revue de littérature de Head et Mayer (2014)). L'impact de l'ALENA est généralement plus élevé que celui de l'UE.

4. Les résultats obtenus sont en ligne avec la littérature

Le modèle est estimé dans sa forme multiplicative :

$$y_{ij} = \exp[\alpha_0 + \alpha_1 \ln(PIB_i) + \alpha_2 \ln(PIB_j) - \alpha_3 \ln(Dist_{ij}) + \tau_{ij} + \Phi_j + \Phi_i] \varepsilon_{ij}$$

Où y_{ij} représente le flux commercial du pays i vers le pays j , PIB_i le produit national brut du pays i , $Dist_{ij}$ la distance entre le pays i et le pays j , τ_{ij} les termes de résistance bilatérale représentés par des variables indicatrices de langue commune, monnaie commune, etc., ε_{ij} le terme d'erreur, $\alpha_{0,1,2,3}$ les paramètres d'intérêt, Φ_i est le terme de résistance multilatéral de l'exportateur et Φ_j un effet spécifique au pays de destination (effet spécifique de l'importateur).

4.1. Validation des résultats et estimation retenue

Nous validons les résultats obtenus à l'aide de trois critères. Ils sont évalués au regard de :

- (i) leur cohérence en fonction de la base de données utilisée ;
- (ii) l'impact du recours aux effets fixes ;
- (iii) la méthode d'estimation utilisée en comparaison avec la littérature.

4.1.1 On observe des effets très comparables entre les résultats issus des deux bases de données, permettant d'utiliser la base la plus large

Les coefficients estimés à partir des deux bases de données décrites *supra* sont très proches. Suivant les recommandations de Rose (2016), nous retenons la base I « super mondiale » pour sa couverture spatiale et temporelle plus large et sa reproductibilité¹⁰. C'est donc cette base qui est utilisée pour les estimations présentées dans la suite de ce document.

Tableau 2 : Comparaison des coefficients selon la base utilisée (PPML et MCO) avec effets fixes directionnels

	MCO		PPML	
	Base I	Base II	Base I	Base II
Distance	-1,281*** (-47,50)	-1,539*** (-88,94)	-0,875*** (-125,63)	-0,818*** (-123,90)
Frontière commune	0,144 (1,32)	0,506*** (6,45)	0,345*** (23,09)	0,506*** (29,76)
Langue Commune	0,595*** (9,69)	0,432*** (12,28)	0,200*** (11,18)	0,295*** (16,37)
Colonisateur Commun	0,755*** (6,56)	0,764*** (16,60)	0,645*** (15,27)	0,433*** (13,43)
Système juridique commun	0,619*** (15,68)	0,378*** (16,04)	0,242*** (22,21)	0,164*** (15,30)
UE	-0,453*** (-6,85)	-0,406*** (-4,96)	0,343*** (17,89)	0,297*** (17,64)
ALENA	0,235 (0,71)	-0,0316 (-0,11)	0,690*** (22,31)	0,136*** (32,52)
Nombre d'observations	244 236	638 329	244 236	638 373
R ²	0,7883	0,7288	0,9021	0,8918

Données : base I « super mondiale » FMI et Comtrade, base II chelem (Cepii)
Estimations : DG Trésor.

Les résultats obtenus par les deux modèles sont cohérents avec la littérature. Les paramètres estimés en PPML sont en général plus faibles que ceux obtenus via l'estimateur des MCO (cf. tableau 3 ci).

¹⁰ Toutes les données sont en accès libre.

Tableau 3 : Comparaison avec les coefficients estimés par la littérature*Modèles de gravité structurels issus de la littérature, présentés dans Head et Mayer (2014)*

Coefficients	Médiane	Moyenne	Ecart-type	# d'osb.
PIB exportateur	0,86	0,74	0,45	31
PIB importateur	0,67	0,58	0,41	29
Distance	-1,14	-1,1	0,41	328
Frontière commune	0,52	0,66	0,65	266
Langue commune	0,33	0,39	0,29	205
Lien colonial	0,84	0,75	0,49	60
Accord de libre échange	0,28	0,36	0,42	108
Union Européenne	0,19	0,16	0,5	26
ALENA	0,53	0,76	0,64	17
Monnaie commune	0,98	0,86	0,39	37

Source : Head et Mayer (2014).

4.1.2 L'utilisation d'effets fixes est essentielle pour l'estimation de l'effet des politiques commerciales

Dans un premier temps, nous comparons les coefficients obtenus avec les deux méthodologies avec et sans effets fixes. L'estimation en MCO avec effets fixes directionnels (cf. tableau 3) présente des résultats satisfaisants pour les variables de distance, de langues, frontières et système juridique commun. Toutefois les variables Union Européenne et accord de libre-échange ne ressortent avec le signe attendu qu'une fois ajoutés des effets fixes couples de pays (cf. tableau 4). Ces estimations confirment les résultats de la littérature. Cette dernière spécification est donc retenue dans le cadre du modèle de gravité de la DG-Trésor.

Tableau 4 : Comparaison des coefficients estimés en PPML avec et sans effets fixes (directionnels et couples de pays)

	Sans effets fixes	Effets fixes exportateurs	Effets fixes importateurs	Effets fixes imp/exp	Effets fixes couples pays
Distance	-0,691*** (-18,94)	-0,770*** (-64,97)	-0,819*** (-72,92)	-0,818*** (-100,25)	
PIB exportateur	0,781*** (51,38)		0,824*** (165,33)		
PIB importateur	0,785*** (41,65)	0,823*** (213,17)			
Langue Commune	0,487*** (5,61)	0,555*** (17,28)	0,312*** (9,77)	0,295*** (19,23)	
Frontière commune	0,616*** (5,96)	0,506*** (22,69)	0,562*** (26,52)	0,502*** (29,60)	
Colonisateur Commun	0,772*** (5,13)	0,642*** (18,93)	0,711*** (22,84)	0,433*** (14,12)	
Système juridique	0,0432 (0,52)	0,114*** (6,06)	0,0890*** (4,88)	0,164*** (15,93)	
Union Européenne	-0,0601 (-0,77)	0,127*** (5,83)	0,00529 (0,18)	0,297*** (14,91)	0,355*** (10,38)
R ²				0,8918	0,9923
Nombre d'observations	638 373	638 358	638 349	638 373	638 370

Estimations : DG Trésor.

4.2. Les coefficients ne requièrent pas de transformation particulière et sont interprétables directement

4.2.1 Interprétation des coefficients

L'interprétation des coefficients du modèle estimé avec le pseudo-maximum de vraisemblance de Poisson est la même que pour une estimation en moindres carrés ordinaires.

Les coefficients des variables indicatrices doivent être transformés afin de les interpréter comme des élasticités. On retiendra donc :

$$(1) \quad e = \exp(\alpha) - 1$$

Ainsi, toutes choses égales par ailleurs, si $\alpha_{RTA} = 0,2$, alors l'effet positif moyen d'un accord de libre échange sur le niveau des exportations entre deux pays membres est de 22,1 %.

4.2.2 Conversion en équivalents *ad valorem*

L'équivalent *ad valorem* d'une mesure de politique commerciale traduit son impact sur le volume des échanges en pourcentage de droits de douane. Il s'agit d'évaluer le niveau de droits de douanes nécessaire permettant d'aboutir à une diminution égale du volume d'exportations. Si le modèle de gravité contient des données relatives aux droits de douanes, on l'obtient par :

$$(2) \quad EAV = e^{\alpha_{RTA}/\alpha_{TARIFF}} - 1$$

En l'absence de données sur les droits de douanes, ou à des fins de contrôle, on peut également utiliser l'élasticité de substitution de l'équation structurelle de gravité. On trouve l'équivalent *ad valorem* par :

$$(3) \quad EAV = e^{\alpha_{RTA}/-\sigma} - 1$$

Dans cette situation, il conviendra d'utiliser des valeurs de σ issues de la littérature.

5. Le modèle se montre également très robuste au niveau sectoriel pour les biens et les services

5.1. Le modèle de gravité permet de calculer des « potentiels de commerce », c'est-à-dire des flux d'exportation théoriques à l'équilibre des équations estimées

Nous définissons la notion de potentiels de commerce comme des flux d'exportations « théoriques » résultant de l'estimation d'un modèle de gravité.

Ils reflètent pour chaque couple de pays le niveau théorique d'exportations compte tenu de la structure du commerce mondial pour une période donnée. Cet indicateur n'apporte pas d'information sur un niveau « optimal » d'exportations, mais plutôt sur un niveau *a priori* accessible d'exportations au regard des déterminants du commerce exhibés dans les estimations ci-dessus.

Les potentiels estimés par le modèle de gravité présenté ci-dessus tiennent compte à la fois de déterminants « bilatéraux » (touchant à la relation spécifique entre deux pays) et « multilatéraux », captés dans les effets fixes importateurs et exportateurs notamment, qui permettent de tenir compte des effets de résistance multilatérale.

Compte tenu des variables explicatives retenues (distance, PIB, et autres variables introduites dans la régression, cf. section 3), les potentiels de commerce ne sont affectés qu'indirectement par les déterminants internes de la performance à l'export du pays (fiscalité, change, compétitivité de manière générale, etc.), dont les effets peuvent être captés par les effets fixes.

Un écart au potentiel de commerce peut amener à s'interroger sur l'existence d'avantages comparatifs particuliers sur un secteur ou une géographie donnée (cas d'un niveau d'exportations supérieur au potentiel), ou à l'inverse de la présence de barrières au commerce insoupçonnées (cas d'un niveau inférieur au potentiel). Même s'ils constituent un outil précieux de politique économique, les potentiels de commerce ne doivent donc pas être interprétés comme des objectifs d'exportation en tant que tel.

En matière de politique publique, l'analyse des potentiels de commerce permet d'identifier les pays vers lesquels des gains sont possibles plus rapidement et permettent de savoir où orienter les efforts de soutien structurel à l'export compte tenu des marges de manœuvre disponibles.

5.2. Les estimations sectorielles des potentiels de commerce ont la même base économétrique que le modèle global

Les modèles de gravité sont traditionnellement utilisés avec des données d'exportations agrégées. Afin de mener une analyse à un niveau plus fin, nous estimons également les équations à un niveau sectoriel. Les modèles sectoriels sont estimés de la même manière que le modèle global *via* la méthode du pseudo-maximum de vraisemblance de Poisson.

Avec l'approche sectorielle, la variable endogène de l'estimation correspond au flux de commerce bilatéral dans le secteur analysé¹¹. Une équation de gravité est ainsi estimée pour chaque secteur. Les variables exogènes sont les mêmes que celles documentées *supra*.

Les secteurs sont sélectionnés en fonction de leur avantage comparatif révélé (ACR), un indicateur calculé par le CEPII renseignant sur la performance relative des secteurs exportateurs d'une économie. Dans le cadre de cette étude et dans la perspective de formulation de recommandations de politique de soutien aux exportations, nous ne retenons que les secteurs à ACR positifs. Ce choix suppose que les conclusions qui pourraient être tirées d'une telle analyse n'ont pas vocation à être transversales, qu'elles concernent des secteurs déjà performants (ainsi il n'est pas envisagé d'utiliser ce type d'approche pour des secteurs structurellement peu performants de l'économie française), et qu'elles ne sont pas de nature à modifier les ACR des secteurs concernés. Il s'agit donc avant tout d'identifier certains secteurs et pays avec lesquels des marges de progression semblent possibles à l'export, par exemple *via* un meilleur accompagnement du tissu exportateur, la conclusion de négociations commerciales afin d'augmenter des potentiels de commerce, ou encore l'identification de marchés potentiellement porteurs à l'export.

Via le critère de sélection basé sur les ACR, 11 secteurs touchant au commerce de biens sont retenus, représentant 70 % des exportations de biens français, et 50 % du commerce mondial. Les services sont également intégrés à l'analyse. Comme nous ne disposons pas d'ACR pour le secteur des services, nous retenons les trois principaux secteurs de services exportés par la France, représentant plus de 70 % des exportations françaises de service.

Les 14 secteurs résultant de cette sélection sont présentés dans le tableau 5.

¹¹ Données Cepii, entre 8 400 et 224 000 observations selon le secteur étudié.

Tableau 5 : Secteurs des biens et services pour lesquels une équation de gravité est estimée

Aéronautique et spatial
Produits chimiques
Produits alimentaires
Pharmacie
Sidérurgie
Instruments médicaux et de précision optique
Papier, pâte de papier, carton, édition
Machines
Equipements et appareils électiques
Automobile
Prestations intellectuelles aux entreprises ¹²
Tourisme
Transport

La présence de flux nuls, plus nombreux sur les échanges bilatéraux au niveau sectoriel qu'au niveau global (en moyenne, 16 fois plus de flux nuls selon les secteurs), renforce l'intérêt du recours à l'estimateur PPML. En effet, lorsque des petits pays exportent des montants faibles au niveau global, les exportations ne concernent généralement que quelques secteurs spécifiques. Beaucoup de pays n'exportent aucun produit dans certains secteurs, et l'information sous-jacente aux flux nuls ne serait pas prise en compte via une estimation par MCO par exemple. L'estimateur PPML permet au contraire la prise en compte de l'information contenue dans ces flux d'exportations sectorielles nuls.

Les 13 équations suivantes sont estimées par PPML avec des effets fixes exportateur-année et importateur-année :

- $X_{ij}^{Alimentation} = \exp[\alpha_0 + \alpha_1 \ln(Dist_{ij}) + \alpha_2 FrontièreCommune + \alpha_3 LangueCommune + \alpha_4 Colonie + \alpha_5 SystemeJuridique + \alpha_6 UE + \alpha_7 Alena + \alpha_8 Autre accord commercial] + \varepsilon_{ij}$
- $X_{ij}^{Aéronautique} = \exp[\alpha_0 + \alpha_1 \ln(Dist_{ij}) + \alpha_2 FrontièreCommune + \alpha_3 LangueCommune + \alpha_4 Colonie + \alpha_8 Autre accord commercial] + \varepsilon_{ij}$
- $X_{ij}^{Appareils electriques} = \exp[\alpha_0 + \alpha_1 \ln(Dist_{ij}) + \alpha_2 FrontièreCommune + \alpha_3 LangueCommune + \alpha_4 Colonie + \alpha_5 SystemeJuridique + \alpha_7 Alena + \alpha_8 Autre accord commercial] + \varepsilon_{ij}$
- $X_{ij}^{Automobile} = \exp[\alpha_0 + \alpha_1 \ln(Dist_{ij}) + \alpha_2 FrontièreCommune + \alpha_4 Colonie + \alpha_5 SystemeJuridique + \alpha_7 Alena + \alpha_8 Autre accord commercial] + \varepsilon_{ij}$
- $X_{ij}^{Chimie} = \exp[\alpha_0 + \alpha_1 \ln(Dist_{ij}) + \alpha_2 FrontièreCommune + \alpha_3 LangueCommune + \alpha_4 Colonie + \alpha_5 SystemeJuridique + \alpha_6 UE + \alpha_8 Autre accord commercial] + \varepsilon_{ij}$
- $X_{ij}^{Instruments médicaux et optique} = \exp[\alpha_0 + \alpha_1 \ln(Dist_{ij}) + \alpha_2 FrontièreCommune + \alpha_4 Colonie + \alpha_6 UE + \alpha_7 Alena + \alpha_8 Autre accord commercial] + \varepsilon_{ij}$

¹² Services juridiques, de comptabilité, conseil en gestion et relations publiques, publicité/étude de marché/sondages, R&D, architecture, ingénierie, services scientifiques, services agricoles, location-exploitation, services liés au commerce, autres services fournis aux entreprises (services d'emploi notamment).

- $X_{ij}^{Machines} = \exp[\alpha_0 + \alpha_1 \ln(Dist_{ij}) + \alpha_2 Fronti\grave{e}reCommune + \alpha_3 LangueCommune + \alpha_4 Colonie + \alpha_6 UE + \alpha_7 Alena + \alpha_8 Autre\ accord\ commercial] + \varepsilon_{ij}$
- $X_{ij}^{Papier} = \exp[\alpha_0 + \alpha_1 \ln(Dist_{ij}) + \alpha_2 Fronti\grave{e}reCommune + \alpha_3 LangueCommune + \alpha_4 Colonie + \alpha_5 SystemeJuridique + \alpha_6 UE + \alpha_7 Alena + \alpha_8 Autre\ accord\ commercial] + \varepsilon_{ij}$
- $X_{ij}^{Pharmacie} = \exp[\alpha_0 + \alpha_1 \ln(Dist_{ij}) + \alpha_2 Fronti\grave{e}reCommune + \alpha_3 LangueCommune + \alpha_4 Colonie + \alpha_6 UE] + \varepsilon_{ij}$
- $X_{ij}^{Siderurgie} = \exp[\alpha_0 + \alpha_1 \ln(Dist_{ij}) + \alpha_2 Fronti\grave{e}reCommune + \alpha_3 LangueCommune + \alpha_4 Colonie + \alpha_5 SystemeJuridique + \alpha_6 UE + \alpha_7 Alena + \alpha_8 Autre\ accord\ commercial] + \varepsilon_{ij}$
- $X_{ij}^{Transports} = \exp[\alpha_0 + \alpha_1 \ln(Dist_{ij}) + \alpha_2 Fronti\grave{e}reCommune + \alpha_5 SystemeJuridique + \alpha_8 Autre\ accord\ commercial] + \varepsilon_{ij}$
- $X_{ij}^{Voyages} = \exp[\alpha_0 + \alpha_1 \ln(Dist_{ij}) + \alpha_2 Fronti\grave{e}reCommune + \alpha_3 LangueCommune + \alpha_4 Colonie + \alpha_5 SystemeJuridique] + \varepsilon_{ij}$
- $X_{ij}^{Services\ aux\ entreprises} = \exp[\alpha_0 + \alpha_1 \ln(Dist_{ij}) + \alpha_2 Fronti\grave{e}reCommune + \alpha_5 SystemeJuridique + \alpha_8 Autre\ accord\ commercial] + \varepsilon_{ij}$

5.3. Les r sultats des mod les de gravit  sectoriels sont robustes et l' tude de la variation des coefficients entre secteurs apporte une information int ressante

Les r sultats des 14 estimations (13 sectorielles et une totale) sont pr sent s dans le tableau 6.

Les estimations au niveau sectoriel se montrent robustes et permettent une analyse approfondie des potentiels de commerce sectoriels. Les R^2 sont compris entre 0,82 et 0,95 selon le secteur et les coefficients varient entre les secteurs de mani re coh rente.

Ainsi, on peut constater que tous les accords commerciaux (UE, ALENA ou autre) n'ont aucun impact sur les services de voyages alors qu'ils ont un impact fort sur l'explication des  changes de biens. La distance a tr s peu d'impact sur les exportations a ronautiques alors qu'elle a un impact tr s fort sur les autres biens et services  chang s. Le fait d' tre membre de l'Union europ enne a beaucoup plus d'impact sur les exportations dans le secteur de l'alimentation (agriculture et agroalimentaire) que sur les autres secteurs.

La variation des coefficients selon le secteur permet  galement de constater la robustesse du mod le au niveau sectoriel. Par exemple le coefficient de la distance est en moyenne plus  lev  dans les secteurs de biens que dans les services. Au sein des biens, les exportations de produits pharmaceutiques d pendent moins de la distance que d'autres secteurs de biens d' quipement plus co teux   transporter par rapport   la valeur du bien, par exemple dans le secteur de la sid rurgie. Les services de voyages d pendent positivement et fortement d'une langue commune.

Tableau 6 : Résultats des estimations sectorielles par PPML

	Contrib à l'explication des exports	Tous biens	Alimentation	Aéronautique	Appareils Electriques	Automobile	Chimie	Instruments médicaux & optique	Machines	Papier	Pharmacie	Siderurgie	Transports	Voyages	Services aux entreprises
Distance	63%	-0.818***	-0.734***	-0.0758**	-0.822***	-0.709***	-0.897***	-0.599***	-0.666***	-0.967***	-0.406***	-0.986***	-0.676***	-1.137***	-0.472***
Langue	4.5%	0.295***	0.295***	0.0729*	0.155***	ns	0.152***	ns	0.251***	0.425***	0.148***	0.222***	ns	0.413***	ns
Frontière	6%	0.502***	0.438***	1.094***	0.0636***	0.359***	0.543***	0.568***	1.266***	1.272***	1.354***	0.455***	0.337***	0.633***	0.264***
Colonisateur	0.3%	0.433***	0.977***	1.638***	0.737***	1.410***	0.297***	0.273***	0.329***	0.346***	0.319***	0.945***	ns	0.378**	ns
Système juridique	2.4%	0.164***	0.306***	ns	0.307***	0.438***	0.0941***	ns	ns	0.205***	ns	0.202***	0.270***	0.209***	0.423***
UE	6.6%	0.297***	0.764***	ns	ns	ns	0.426***	0.201***	0.340***	0.428***	0.365***	0.533***	ns	ns	ns
Autre accord com. régional	8.6%	0.312***	0.320***	0.283***	0.271***	1.060***	0.372***	0.421***	0.457***	0.109*	ns	0.342***	0.159*	ns	0.447***
ALENA	8.8%	0.136***	0.561***	ns	1.076***	1.019***	ns	0.0945**	0.156***	0.292***	ns	0.219***	ns	ns	ns
N		638 373	224 112	105 438	198 806	180 238	218 648	195 974	215 407	199 183	166 739	163 692	9 055	8 399	8 572
R²		0.89	0.88	0.90	0.93	0.94	0.93	0.85	0.92	0.95	0.82	0.89	0.82	0.92	0.90

Effets fixes : importateur-année ; exportateur-année

ns : non significatif

* p<0.05 ** p<0.01 *** p<0.001

Estimations DG Trésor.

Annexe – le modèle d'Anderson et van Wincoop

Anderson et van Wincoop dérivent l'équation de gravité d'une fonction de demande à élasticité de substitution constante et biens différenciés par origine, reposant sur l'hypothèse d'Armington (1969) de substituabilité imparfaite entre des biens locaux et importés. Dans ce cadre, la fonction d'utilité du consommateur (pays importateur, j) est la suivante :

$$(1) \quad \left\{ \sum_i \alpha_i^{\frac{1-\sigma}{\sigma}} c_{ij}^{\frac{\sigma-1}{\sigma}} \right\}^{\frac{\sigma}{\sigma-1}}$$

Où, $\sigma > 1$ est l'élasticité de substitution constante entre les variétés de biens (dans le cadre de l'hypothèse d'Armington, les biens sont différenciés selon leur provenance), α est le paramètre de préférence ou la propension à consommer une variété, et c la consommation du pays j provenant du pays i . Le consommateur maximise cette fonction d'utilité sous contrainte budgétaire :

$$(2) \quad \sum_i p_{ij} c_{ij} = E_j$$

La dépense E du pays j équivaut à la somme de ses consommations c depuis tous les pays i (j étant compris) au prix des exportations (p). Ainsi, les exportations (Y_{ij}) du pays i vers le pays j sont égales à la part (π_{ij}) des dépenses totales du pays j (E_j) allouées à i :

$$(3) \quad Y_{ij} = \pi_{ij}^{1-\sigma} E_j$$

Cette part de dépenses est fonction de la propension (α_i) du pays j à consommer des biens en provenance de i , de l'indice prix des exportations du pays i (p_i) et des coûts de commerce bilatéraux (τ_{ij}) rapportés à l'indice de prix domestique (Φ_j). Dans la mesure où le consommateur maximise son utilité en fonction de ses préférences et des prix domestiques et locaux, on a¹³ :

$$(4) \quad Y_{ij} = \left(\frac{\alpha_{ij} p_i \tau_{ij}}{\Phi_j} \right)^{(1-\sigma)} E_j$$

L'apport des auteurs est de démontrer la nécessité d'inclure dans l'équation des variables capables de contrôler pour les coûts relatifs de commerce (avec les pays tiers) non seulement pour l'exportateur mais aussi pour l'importateur.

Si on considère Φ_j comme un indice de prix domestique (la résistance commerciale de l'importateur) permettant d'apprécier le niveau de protection commerciale imposé par l'importateur dans ses relations avec l'ensemble des pays tiers, alors il s'exprime comme la somme de la propension à consommer de j (ou marqueur de préférence CES), de l'indice de prix des exportations et des coûts de commerce bilatéraux depuis l'ensemble des exportateurs (i), soit :

$$(5) \quad \Phi_j = \left[\sum_i (\alpha_{ij} p_i \tau_{ij})^{1-\sigma} \right]^{\frac{1}{1-\sigma}}$$

¹³ Il s'agit de noter ici que α_{ij} peut être interprété comme un droit de douane car il détermine la propension à consommer la variété d'un bien, et est appliqué au prix des exportations. Ainsi lors de l'estimation de l'équation de gravité les données de droits de douanes n'intégreront pas le vecteur de coût bilatéral de commerce τ_{ij} .

Selon Anderson et Wincoop, la propension du pays i à exporter ($\alpha_i p_i$), est fonction de sa production (PIB_i/PIB) rapportée à la production mondiale et la résistance commerciale à laquelle l'exportateur fait face au niveau mondial (Φ_i) :

$$(6) \quad \alpha_i p_i^{1-\sigma} = \frac{PIB_i/PIB}{\Phi_i^{1-\sigma}}$$

Cette résistance multilatérale est égale à la somme des résistances bilatérales pondérées par le niveau global de protection du partenaire (Φ_j) :

$$(7) \quad \Phi_i^{1-\sigma} = \sum_j \left(\frac{\tau_{ij}}{\Phi_j} \right)^{1-\sigma} \frac{E_j}{PIB}$$

En remplaçant 8 dans 7 on obtient

$$(8) \quad \Phi_j = \sum_i \left(\frac{\tau_{ij}}{\Phi_i} \right)^{1-\sigma} \frac{PIB_i}{PIB}$$

Les auteurs considèrent qu'à l'équilibre les revenus (PIB_i) d'un pays i sont équivalents à la somme de ses exportations vers tous ses partenaires (y compris i) ou à la somme de la part des dépenses de tous les partenaires j allouées à i donc :

$$(9) \quad PIB_i = \sum_j Y_{ij} = \sum_j \pi_{ij} E_j = \sum_j \left(\frac{\alpha_{ij} p_i \tau_{ij}}{\Phi_j} \right)^{(1-\sigma)} E_j$$

On considère également, à l'équilibre, que les dépenses (E_j) du pays j sont équivalentes à la somme de ses importations depuis tous ses partenaires (y compris j). On retrouve donc (5).

Les auteurs reprennent naturellement le terme de résistance bilatérale τ_{ij} qui représente l'ensemble des barrières au commerce entre deux partenaires (liens historiques, langue commune, frontière commune etc.) à l'instar du modèle de gravité intuitif - équation (1).

Ainsi, les trois composantes des coûts de commerce sont la résistance bilatérale (τ_{ni}), la résistance multilatérale de l'importateur (Φ_j) et la résistance multilatérale de l'exportateur (Ω_i). On obtient finalement¹⁴ :

$$(10) \quad Y_{ij} = \frac{PIB_i PIB_j}{PIB} \left(\frac{\tau_{ij}}{\Phi_i \Phi_j} \right)^{1-\sigma}$$

¹⁴ La forme présentée ici est une version simplifiée de l'équation d'Anderson et van Wincoop. Cette dernière peut être déclinée au niveau sectoriel avec k le secteur d'intérêt et σ_k l'élasticité de substitution constante entre les secteurs, soit : $Y_{ij}^k = \frac{Y_i^k Y_j^k}{Y^k} \left(\frac{\tau_{ij}^k}{\Phi_i \Phi_j} \right)^{1-\sigma_k}$

L'équation (10) est particulièrement transparente et peut être décomposée en deux. D'une part, les exportations dépendent de la taille économique relative des partenaires ($\frac{PIB_i PIB_j}{PIB}$); d'autre part elles dépendent des coûts de commerce $\left(\frac{\tau_{ij}}{\Phi_i \Phi_j}\right)^{1-\sigma}$.

Ainsi, en ne dérivant que la première partie de l'équation, on décrit un monde sans friction alors que la seconde partie de l'équation éclaire sur les coûts de commerce bilatéraux (qui décrivent les barrières propres à la relation des partenaires) et les coûts multilatéraux (qui sont les barrières compte tenu du niveau de frictions avec les pays tiers).

L'équation log-linéarisée sera donc la suivante :

$$(11) \quad \ln(Y_{ij}) = -\ln(PIB) + \ln(PIB_i) + \ln(PIB_j) + (1 - \sigma)(\ln(\tau_{ij}) - \ln(\Phi_j) - \ln(\Phi_i)) + \varepsilon_{ij}$$

Bibliographie

- Anderson J. E. & E. Van Wincoop (2003), "Gravity with gravitas: a solution to the border puzzle", *American economic review*, 93(1), 170-192.
- Anderson J. E., Larch M. & Y. Yotov (2015), "Estimating general equilibrium trade policy effects: GE PPML", *MIMEO*.
- Anderson J. E., Larch M. & Y. V. Yotov (2015), "Growth and trade with frictions: A structural estimation framework (No. w21377)". National Bureau of Economic Research.
- Baier S. et J. Bergstrang (2002), "On the endogeneity of international trade flows and free trade agreements", *Manuscript*.
- Baier S. et J. Bergstrang (2005), "Do free trade agreements actually increase Members' international trade?", *Manuscript*.
- Baldwin, Taglioni, (2006), "Gravity for dummies and dummies for gravity equations", *NBER*.
- Da Silva J.S. et S. Tenreyro (2006), "The log of Gravity", *The review of Economics and Statistics*.
- Davis, Gift, The positive effects of the Schengen agreement on European trade.
- Eaton et Kortum, (2002), Technology, geography and trade.
- Egger P., Larch M., Staub K. E., R. Winkelmann (2011), "The Trade Effects of Endogenous Preferential Trade Agreements", *American Economic Journal: Economic Policy* 3, August.
- Fally (2014), "Structural gravity and fixed effects" [publié en 2015 apparemment].
- Feenstra R. (2003), "Advanced international trade: Theory and Evidence", Princeton, NJ: Princeton University Press.
- Fontagné L., Mayer T. et S. Zignago (2005), "A re-evaluation of the impact of regional trade agreements on trade", *Draft Version*, CEPII.
- Gourieroux C., Monfort A. & A. Trognon (1984), « Pseudo maximum likelihood methods: Theory », *Econometrica: Journal of the Econometric Society*, 681-700.
- Head K. et T. Mayer, (2013), "Gravity Equations: Toolkit, Cookbook, Workhorse." *Handbook of International Economics*, Vol. 4, eds. Gopinath, Helpman, and Rogoff, Elsevier.
- Head K., T. Mayer J. Ries (2010), "The erosion of colonial trade linkages after independence" *Journal of International Economics*, 81(1):1-14. (formerly CEPII discussion paper # 2008-27).
- Piermartini et Yotov, (2016), "Estimating Trade policy effects with structural gravity".
- Rose A. K. (2000), "One money, one market: the effect of common currencies on trade", *Economic policy*, 15(30), 08-45.
- Rose A. K. (2016), "Why Do Estimates of the EMU Effect on Trade Vary So Much?", CEPR Discussion Paper No. 11532.
- Sevestre P. (2002), « Économétrie des données de panel », Dunod.
- Taglioni et De Benedict, (2011) "The gravity in international trade".
- Tinbergen Jan (1962), "Shaping the World Economy. Suggestions for an International Economic Policy", *The Twentieth Century Fund*, New York.
- Viner J. (1950). The Customs Union Issue. Carnegie Endowment for International Peace. *New York*.
- Wei S.J. (1996), "Intra-National Versus International Trade: How Stubborn are Nations in Global Integration?" *NBER Working Paper*, No. 5531.
- Nikolaus et Ritschl, O. Albrecht (2011), "Endogeneity of Currency Areas and Trade Blocs Evidence from a natural Experiment Wolf".