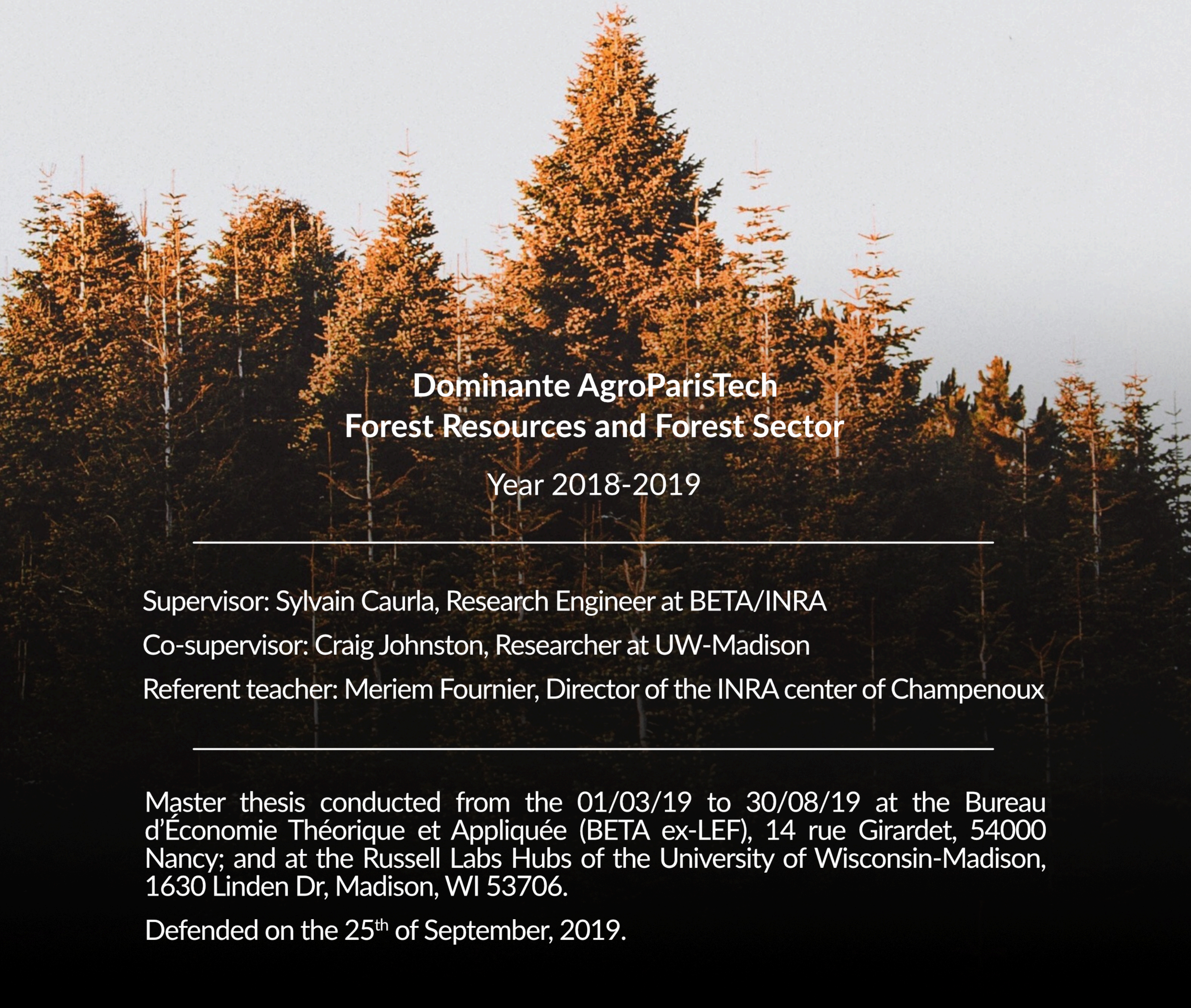
Forest Sector Modeling

A critical analysis from the perspective of the notions of space, place, and scale

---

Valentin MATHIEU



Dominante AgroParisTech
Forest Resources and Forest Sector

Year 2018-2019

Supervisor: Sylvain Caurla, Research Engineer at BETA/INRA

Co-supervisor: Craig Johnston, Researcher at UW-Madison

Referent teacher: Meriem Fournier, Director of the INRA center of Champenoux

---

Master thesis conducted from the 01/03/19 to 30/08/19 at the Bureau d'Économie Théorique et Appliquée (BETA ex-LEF), 14 rue Girardet, 54000 Nancy; and at the Russell Labs Hubs of the University of Wisconsin-Madison, 1630 Linden Dr, Madison, WI 53706.

Defended on the 25th of September, 2019.

Forest Sector Modeling

A critical analysis from the perspective of the notions of space, place, and scale

---

Valentin MATHIEU

Dominante AgroParisTech
Forest Resources and Forest Sector

Year 2018-2019

Supervisor: Sylvain Caurla, Research Engineer at BETA/INRA

Co-supervisor: Craig Johnston, Researcher at UW-Madison

Referent teacher: Meriem Fournier, Director of the INRA center of Champenoux

---

Master thesis conducted from the 01/03/19 to 30/08/19 at the Bureau d'Économie Théorique et Appliquée (BETA ex-LEF), 14 rue Girardet, 54000 Nancy; and at the Russell Labs Hubs of the University of Wisconsin-Madison, 1630 Linden Dr, Madison, WI 53706.

Defended on the 25th of September, 2019.

Abstract

This work aims to explore the relevance of new economic theories to improve the common economic structure of Forest Sector Models and, in particular, its spatial representation of wood flows. To do that, we rely on modern international trade theories and more especially on the work of P. Krugman and the New Economic Geography. We then gain interest in the Economic Geography approach that suggests *"geographical space"* shapes economic activity. Assuming that improving the representation of wood flows lies in better representing spatial features in economic theories, we define three notions that provide a complete comprehension of what implies *"geographical space"*, from a geographical perspective: space, place, and scale. We then derive major economic theories of international trade - those of P. Samuelson, P. Armington, P. Krugman, and M. J. Melitz - and confront the way they represent spatial features with those three geographical notions. We find that new economic theories qualitatively improve the spatial representation of wood flows. To assess what could be the impact of changing the economic structure of Forest Sector Models, we build two simple models describing Armington's and Krugman's equilibrium respectively that are explicitly comparable. Using FAOSTAT data as input, we run the model and test their sensitivity to different key parameters in Forest Sector Modeling. The preliminary results suggest shifting toward new economic theories may improve the way we model the forest sector.

Keywords: Forest Economics, Forest Sector Modeling, International Trade Theories, Economic Geography, Spatial Economics

Résumé

Ce travail a pour objectif d'explorer la pertinence des nouvelles théories économiques pour améliorer la structure économique des modèles de secteur forestier et, en particulier, la manière dont ils représentent les flux de bois. Pour cela, nous nous appuyons sur les théories économiques modernes du commerce international, notamment celle développée par P. Krugman et l'école de la Nouvelle Économie Géographique. Nous nous intéressons ainsi à l'approche de l'Économie Géographique, qui suppose que l'espace constitue le substrat de toute activité économique. Nous définissons ainsi trois notions géographiques - l'espace, le lieux et l'échelle - dans l'idée qu'une amélioration des structures théoriques existantes passe par une meilleure représentation des composantes spatiales. Après avoir développé quatre théories majeures du commerce international nous développons une analyse critique de leur représentation spatiale sous l'angle des trois notions géographiques. Cette analyse qualitative suggère que les théories du commerce international récentes présentent une meilleure représentation spatiale des flux de bois. Afin d'estimer l'impact d'un éventuel changement de structure économique des modèles de secteur forestier, nous construisons deux modèles simples décrivant respectivement la théorie développée par P. Armington et celle développée par P. Krugman en prenant garde de les rendre explicitement comparables. En utilisant les données statistiques de la FAO comme entrées des modèles, nous analysons leur sensibilité à différents paramètres clés de la modélisation du secteur forestier. Nos résultats suggèrent qu'un changement de structure économique peut améliorer la façon dont nous modélisons le secteur forestier.

Mots clés : Économie Forestière, Modélisation du Secteur Forestier, Théorie du commerce international, Économie Géographique, Économie Spatiale

AgroParisTech

Engagement de non plagiat

❶ Principes

- Le plagiat se définit comme l'action d'un individu qui présente comme sien ce qu'il a pris a autrui.
- Le plagiat de tout ou parties de documents existants constitue une violation des droits d'auteur ainsi qu'une fraude caractérisée
- Le plagiat concerne entre autres : des phrases, une partie d'un document, des données, des tableaux, des graphiques, des images et illustrations.
- Le plagiat se situe plus particulièrement à deux niveaux :
  - ➢ Ne pas citer la provenance du texte que l'on utilise, ce qui revient à le faire passer pour sien de manière passive.
  - ➢ Recopier quasi intégralement un texte ou une partie de texte, sans véritable contribution personnelle, même si la source est citée.

❷ Consignes

- Il est rappelé que la rédaction fait partie du travail de création d'un rapport ou d'un mémoire, en conséquence lorsque l'auteur s'appuie sur un document existant, il ne doit pas recopier les parties l'intéressant mais il doit les synthétiser, les rédiger à sa façon dans son propre texte.
- Vous devez systématiquement et correctement citer les sources des textes, parties de textes, images et autres informations reprises sur d'autres documents, trouvés sur quelque support que ce soit, papier ou numérique en particulier sur internet.
- Vous êtes autorisés à reprendre d'un autre document de très courts passages in extenso, mais à la stricte condition de les faire figurer entièrement entre guillemets et bien sur d'en citer la source.

❸ Sanction : En cas de manquement à ces consignes, la DEVE/le correcteur se réservent le droit d'exiger la réécriture du document sans préjuger d'éventuelles sanctions disciplinaires.

❹ Engagement :

Je soussigné (e) Valentin MATHIEU

Reconnaît avoir lu et m'engage à respecter les consignes de non plagiat

A NANCY le 11/09/19

Signature :

Cet engagement de non plagiat doit être inséré en début de tous les rapports, dossiers, mémoires.

Aknowledgement

After four years of studying at AgroParisTech, I realize that I could not have expected better supervision for my last internship as a master/engineering student. I truly want to thank my supervisor, Sylvain Caurla, for the opportunity he brings me, at a time when doors were closing one after another, to work at the BETA on questions I had in mind for a long time and to go to the US. Most importantly, I want to thank him for his trust and for the freedom he gave me during my master thesis. I also found such freedom in the co-supervision of Craig Johnston which I particularly thank for his kind welcome in the beautiful city of Madison and for all his ideas that have surely brightened my research work. I also thank Meriem Fournier, who has been my referent teacher for a while now, for her unfailing support and mentoring which I deeply value.

My gratitude goes to the BETA team too for their friendly welcome in the lab, their kindness, and their friendship. More especially, I quickly felt at home thanks to the colleagues with whom I shared my office: first Loïc, who kindly gave me his desk after he left; Etienne, my co-intern with whom I haunted the corridors of the BETA at night and during weekends, trying desperately to finish our reports; and of course May, for her eternal smile and kindness. I also especially thank Miguel Rivière and Antonello Lobianco that accepted to be part of the jury - and, by advance, I deeply apologize for the pain some parts of my report may cause.

I cannot mention the BETA without talking about the other team I integrated during my internship: the American team of the University of Wisconsin Madison. With Craig, I have been kindly welcomed by Jinggang Guo and Joseph Buongiorno. I am glad to have shared many pleasant moments talking with them at lunch or with a cup of coffee. I want to thank Jinggang for his friendship and the time we spend together - two months were too short. I also consider myself lucky to have met Joseph who shared many thoughts and stories that enlightened my time at the university and made me better understand Forest Sector Modeling. My experience in Madison would not have been the same without Matt who welcomed me in his house. Matt has shown me many secrets of Madison, taking me to bars, restaurants, on his boat... and has initiated me to the American way of life - so to speak. I am sure and I truly hope we will meet again.

This internship in the INRA and the travel to Madison would not have been possible without the precious help or Evelyne Lenel and Annick Demmangejaouen who helped me A LOT with the paperwork - and God knows that French paperwork can be close to hell...

I especially thank Christophe Boisseau and Jean-Yves Courtonne for their work on the comparison between FFSM simulations and SitraM data on which I based my work. I also thank Edward Balistreri and Thomas Rutherford for the work they did on modeling international trade theories and for their helpful answers to my emails. My thanks also go to Trevor Barnes for his papers about the history of Economic Geography and Regional Science and for the friendly emails we exchange.

This year has been vibrant, especially thanks to Holger Wernsdörfer and the RFF teaching team: I truly want to thank all of them for the quality of the education and training they provide and for the openness of the RFF training.

Not choosing to major in Forest Management has been hard in some perspectives, but I consider myself lucky to have met the amazing people of the RFF family and the master AETPF. I also want to thank my friends at the residence or who have left Nancy this year. Thank you to all of them, they have made this year magical.

I cannot imagine a better way to complete my studies at AgroParisTech and I truly feel accountable to all the people that directly or indirectly support me in my projects during all these years. Thank you again.

Note to the reader

The present report consists of a synthesis of the work that has been done during my master thesis. Even if it does not address all the work that has been done, it reflects quite well what we aimed to do while simultaneously trying to follow the writings and internship instructions given by AgroParisTech.

The reader will notice that this master thesis opens many doors without especially completing all the tasks it aimed to achieve. The six months limitation of the thesis did not permit to run all the experiences we envisioned for the last part, which then *only* presents some preliminary results. Furthermore, the pages limitation results in a missing part about the history of Economic Geography and of the notion of space in economics.

We recognize that there is more to do and are willing to complete this ongoing work after the master thesis. However, we also claim that this is not detrimental to the work we present in the this report, as the main goal of this master thesis was, in fact, to open up a range of new perspectives on Forest Sector Modeling. Of course, we leave the reader free to judge for him or herself whether what we present here reaches this goal or not.

Contents

List of Figures

List of Tables

01 Introduction

1 Introduction: From Forest Sector Models to spatial considerations

1.1 Contextual elements: wood flows representation in Forest Sector Models

Forest Sector Models are models that represent aggregated economic behaviors of the forest sector agents (Caurla, 2013) i.e. models that use economic theory to represent both forestry - the upstream segment - and forest industries - the downstream segment - and their interactions (Latta et al., 2013).

Their purpose is mainly to forecast how policies or economic shocks in timber and wood products markets may impact the forest sector evolution (Northway et al., 2013; Sjølie et al., 2015). In doing so, we can also define a Forest Sector Model as a mathematical model, based on equations, that requires computational tools to run simulations (Riviere et al., 2018).

This large definition encompasses a certain variety of models we mainly differentiate through three criteria (Caurla, 2013). First, we can discriminate Forest Sector Models by the way they represent temporal dynamics. We can then consider, on one hand, the "intertemporal models" that assume economic agents have perfect foresight of the future economic environment and simultaneously optimize their decisions; and, on the other hand, "recursive models" that assume economic agents have limited foresight or myopic behaviors. The first category is often associated with wood supply models, while the second usually represents both wood supply and demand. Second, we can differentiate partial equilibrium models with general equilibrium models. Partial equilibrium models only take into account the forest sector by representing other economic activities through exogenous parameters while general equilibrium models consider the forest sector in interaction with other economic sectors. Third, we can separate them by the geographical scale they represent. Indeed, a FSM usually refers to one particular geographical scale on which it is built: global, macro-regional, national, regional, and so forth.

The vast majority of Forest Sector Models are partial equilibrium models with a recursive dynamic (Riviere et al., 2018). At this stage, we choose to refer to recursive, partial equilibrium models when talking about FSM. One characteristics of these models is to be built on the same spatial price equilibrium theory, developed by P. Samuelson in 1952 (Caurla, 2013; Latta et al., 2013).[1] However, no previous work did question the accuracy and reliability of this theory with respect to observed wood flows. Our starting question can therefore be formulated

[1] We develop this theory and the ones that follow in Section 3.

as: does the economic theory that bases FSM well represent wood products flows and their dynamics? We will base our investigation on the French case of FFSM (French Forest Sector Model), a model that belongs to the category of recursive, national, partial equilibrium models (Caurla, 2012, 2013).

1.2 Questioning the relevance of wood flows modeling *à la* Samuelson

In 2018, Christophe Boisseau and Jean-Yves Courtonne from INRIA (the French national research institute for the digital science) conducted an unprecedented comparison between (1) data coming from the SitraM database and (2) the outputs of FFSM.

The SitraM database (French information system on freight) provides, among other things, the distance traveled by timber and wood products across old 22 French metropolitan regions as well as a composite index that states the tonnes of products per kilometer traveled. Boisseau and Courtonne processed these data to obtain flows between French regions. Figure 1 represents a map of the resulting main net flows in metropolitan France (left, Figure 1.1a) together with the results of wood flows from FFSM (right, Figure 1.1b).

The comparison is based on the years 2009 to 2015. Note that the comparison does not aim so much to compare the absolute values of both sources but rather to look for possible correlations.

Starting with a visual comparison of Figures 1.1a and 1.1b we can observe that the main wood products flows are quite different between SitraM data and FFSM simulations both in their initial point - the production region - and in their destination point - the consumption region. More especially, we observe that FFSM flows gravitate around one main demand location, which is the region Île de France, while SitraM flows are more sporadic. They do not seem, *a priori*, correlated. A statistical analysis corroborates this observation and shows that the two datasets present very low correlations. A heat map presents the statistical results of this analysis in Appendix A.

Three main hypotheses can be formulated to explain the gap between SitraM data and FFSM outputs. First, SitraM dataset may be biased for different reasons. For instance, the data collection may have incurred errors and some values may be inaccurate or wrong, and so forth. Second, referring to the modeling considerations of Richard Levins (Levins, 1966), FFSM may not be able to realistically represent trade flows, scarifying realism to generality and precision.[2] Third, the FFSM structure may not well represent wood products flows as it has not been initially built for it. While the issue of accuracy and reliability of SitraM database has already been raised (Courtonne, 2016), we chose to explore the

[2] A detailed description of Levins' thought about modeling is presented in Appendix B

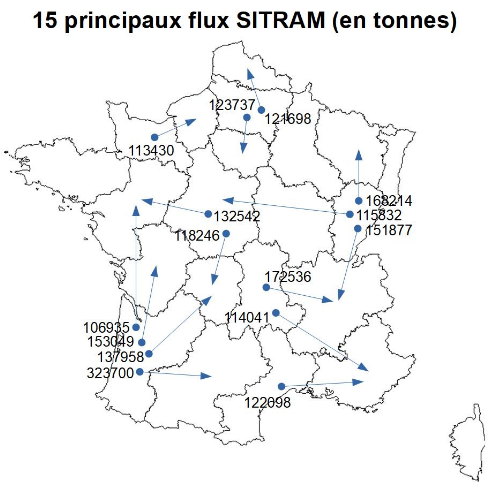


Figure 1.1a Map representing the 15 main timber and wood products flows, in Tons, considering SitraM data.

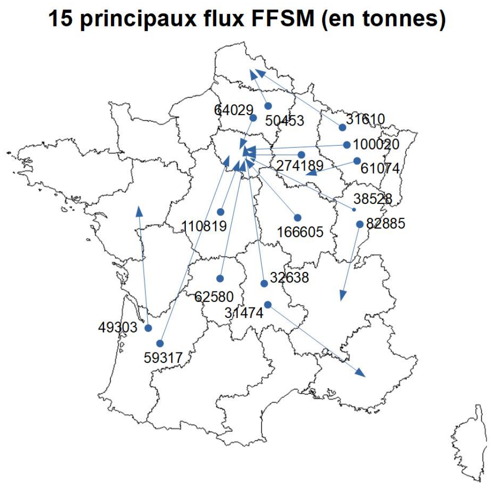


Figure 1.1b Map representing the 15 main timber and wood products flows, in Tons, considering FFSM outputs.

third hypothesis in this master thesis.

1.3 Seeking new theories to improve wood flows modeling in Forest Sector Models

Those considerations lead us to the following question: can we improve the way FSM represents wood products flows without changing their theoretical economic structure?

New economic theories of international trade have appeared since Samuelson's. Our work mention and will essentially focus on Armington's (1969), Krugman's (1980), and Melitz's (2003) theories.[3] In what follows, we will consider Samuelson's and Armington's framework as "early theories of international trade" while we will refer to Krugman's and Melitz's approaches as "modern theories of international trade". Among them, the theory P. Krugman has developed became famous after he was awarded the Nobel Price in 2008 and is known to be the basis of the New Economic Geography.

This makes a direct mention of the discipline of Economic Geography. Section 3 will get into an elaborate definition of this discipline which is, at first glance, not easy to define (Géneau de Lamarlière and Staszak, 2000). In particular, we will explore how Economic Geography capture the economic world diversity by including spatial factors that shape an economy (Coe et al., 2007) wondering if a better spatial representation is crucial to im-

3 Note that, while FSM do not fall back on Krugman's and Melitz's theories, some use Armington's theories, like FFSM.

prove the way FSM represent wood flows.

1.4 Research questions and goals of the study

From what precedes, we start form the assumption that any improvement in FSM that may not change the fundamental structure of the models consists in *“minor”* improvements and will not improve the way FSM represents trade flows. Furthermore, we suggest that improving the economic structure of those models involves to better represent spatial patterns of an economy in FSM.

Thus, we wonder if a shift from Samuelson’s spatial equilibrium theory to modern economic theories of international trade may improve FSM representation of timber and wood products flows by providing a better spatial representation of economic activities.

To do so, we aim to first clarify the vague notion of *"space"*, mentioned as *"geographical space"* in what follows, by precisely defining its three components, namely: space, place and, scale. Our work then derives the main economic theories of international trade and provide critical analysis of their spatial representation with regard to the three geographical notions. This will help us put some perspective on how economic models consider *"geographical space"* and, consequently, represent trade flows.

In the end, we will condense those goals and our work in general into the following question: how can we innovate in Forest Sector Modeling thanks to the contributions of Economic Geography?

Space, place, and scale

2 Space, place and scale: reconsidering economics in light with three key geographical concepts

It can be said that economists have historically put spatial aspects aside despite their importance in understanding economic mechanisms (Thisse and Walliser, 1998). This Section is crucial to assess where this path led the economic science in representing spatial features in international trade.

First and foremost, with regard to the concept of *"space"* that we used in the introduction, we can notice that we did not precisely define what it meant. Even if a common conception of *"geographic space"* was sufficient, it now becomes essential to bring a clear definition of this notion.

To do so, what follows defines three key geographical notions to assess spatial features of economic representations: space, place and scale. We suggest that, combining together, these three notions capture almost all the components that define *"geographic space"* and that they constitute a threefold analysis key to assess the conceptual works presented in the next Section.[4]

We base our definitions of space, place, and scale on *Economic Geography - A Contemporary Introduction* written by Coe, Yeung, and Kelly in 2007.

2.1 The notion of space as a physical component

In this Subsection we aim to move from the vague conception of *"space"* we have to its most fundamental component: physical distance and area. By doing so, this concept reasonably suggests that every economic activity takes place *"on the ground"*. Even if this proposition seems obvious, it is quite common to see economic processes discussed without considering how relative distance and location may contribute.

On the basis of the work of Coe et al., we deconstruct this notion in three interrelated elements:

a **Territoriality and shape.**

This element defines a territory by the contours that create its shape, as maps do. Inside those boundaries, jurisdiction on the ground defines a given territory as well as unevenness, might it be in infrastructure, in resources, and so forth. Figure 2.1a illustrates this concept of territoriality and shape.

b **Geographical location.**

This concept encompasses geographic coordinates as well as how a location relates to another. It brings two indi-

[4] Of course, other definitions and decompositions of the concept of *"geographic space"* may exist and be as relevant as the ones we provide here.

cations. First, a geographical location brings indication about physical properties that help understand the resource base of a territory, such as topography, climate, etc. Second, relative location in relation to others highlights spatial unevenness between locations and, in doing so, informs about the strategic significance of each location to global geopolitics. Figure 2.1b illustrates this notion of geographical location.

c **Flows across space.**

This element recognizes the significance of connections and relationships of (1) one given space with the rest of the world or (2) within one given territory. Different sorts of flows may be considered: commercial flows, migrations, information flows, etc. Taking account of these flows gives a better understanding of the international economic linkages that participate to shaping the economic world. Figure 2.1c gives a representation of this notion of connectivity.

Figure 2.1d combines Figure 2.1a, Figure 2.1b and Figure 2.1c by superimposition to illustrate the inter-relation between these four elements. It clearly appears how difficult it can be to only consider physical space in economic processes. And even so, this notion alone is not enough to have a complete definition of *"geographic space"*.

2.2 The notion of place as the specificity and uniqueness of a particular location

In their book, the authors state that while economics tend to formulate universal laws that apply equally to all contexts, geographers go in the opposite direction and tend to focus on the specificity and uniqueness of places. Therefore, they implicitly assume that economic processes may happen differently according to the context.

But place remains a vague notion and no clear definition naturally comes to mind. What is somehow clear is that place relates to somewhere in particular. The definition of place may then be the answer of the following question: what does shape economic processes at a particular location?

The answer is threefold. First, economic processes fundamentally set up in a specific economic, environmental, cultural, institutional and political context. Second, there are also shaped by the flows and relationships that link a specific place with the outside and that operate within that place. Third, they also depend on the past and present activities that happened on the place.

We thus hold a precise definition of the concept: a place is the unique intersection of all kind of contexts, all various relationships, and all past and present engagements of one specific location. Figure 3.2 illustrates this complex notion.

Thus, the notion of place intends to capture the richness and the complexity

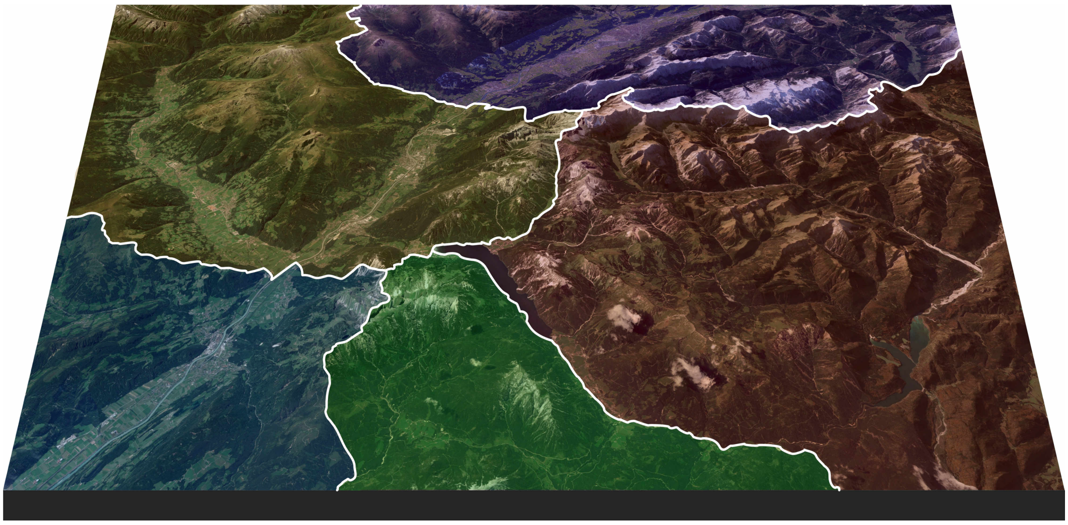


Figure 2.1a A representation of the concept of territoriality and shape: each territory is illustrated with different colors and shape.

Figure 2.1b A representation of the concept of geographical location: each location has specific coordinates and characteristics.

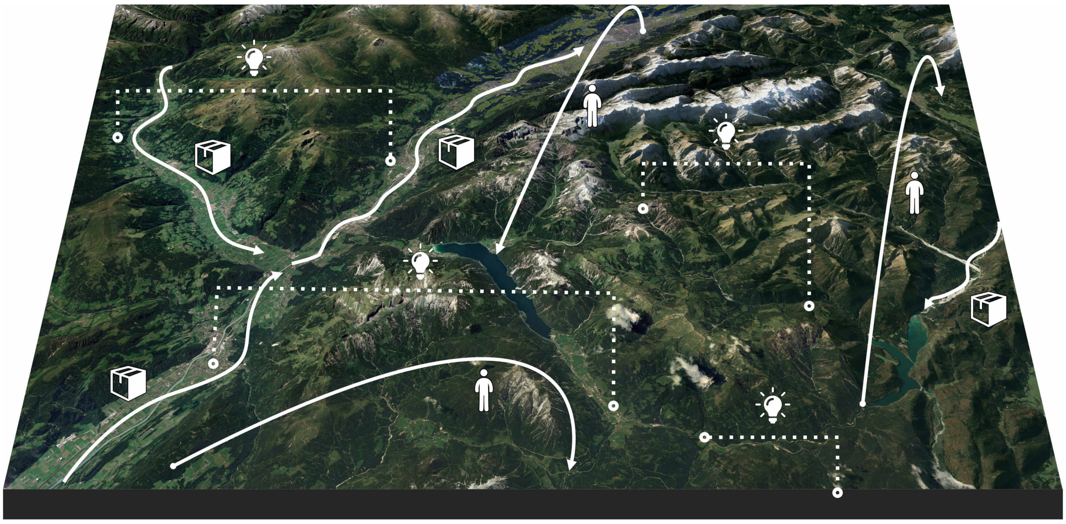


Figure 2.1c A representation of the concept of flows across space: flows of people (icon "person"), trade flows (icon "box"), and flows of ideas (icon "bulb") are represented.

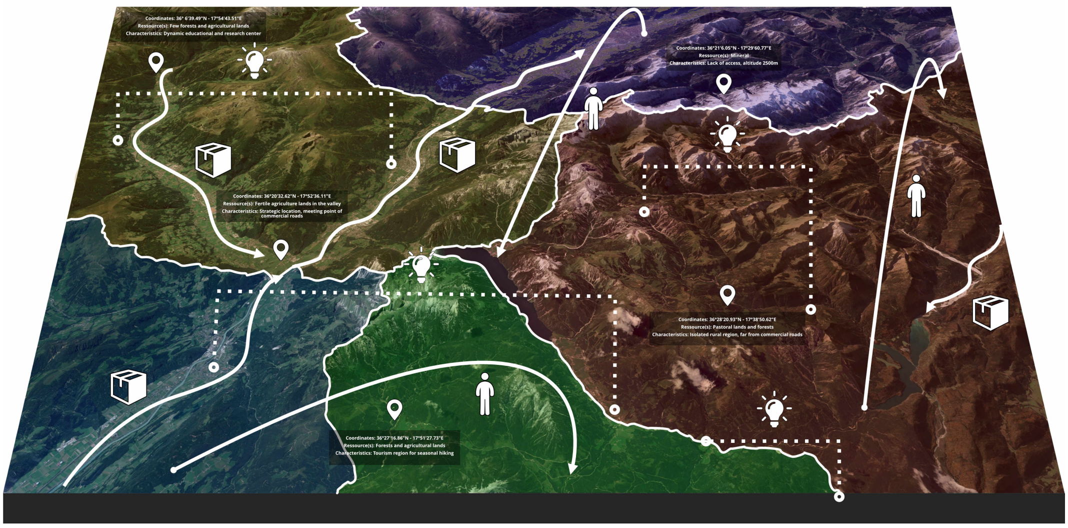


Figure 2.1d A synthesis of the three concepts that shows how difficult it is to simultaneously represent the three spatial notions.

Figure 2.1: A representation of the concept of space.

of one location to “take a more comprehensive view of the ‘economy’ that the one based on formal monetary exchange”.

2.3 The notion of scale as an organization of the notions of space and place

According to the previous definitions, a place can take all shapes and sizes. In other words, there are differently sized units that may meet the definitions of space and place. Therefore, we need the notion of scale to help us organize these two notions.

We define the notion of scale as the level or size of the place we consider and by the inter-connections that exist with the other levels. This definition reminds us that economic processes occur at different scales in a different way yet simultaneously, and that “one scale in isolation is inadequate to develop a comprehensive understanding” of the economy. Figure 2.2 gives an illustration of this notion.

Different typologies of scales exist. We, however, decide to use one that is close to the work of Coe et al.:

a **The global scale** which is the scale of the world, including all countries or regions.

b **The macro-regional** scale that includes several countries, like the European Union or North America.

c **The national scale** that coincides with the scale of a country like the USA or France.

d **The regional scale** that proceeds at a lower level than the national scale e.g. the Grand-Est region in France or the state of Wisconsin in the USA.

e **The local scale**, at a lower level, that may be, for instance, a subsection of a region - like the axis Metz-Nancy-Épinal - or a city like Madison.

We claim that with these three definitions we are well equipped to critically assess how economic theories represent spatial features.

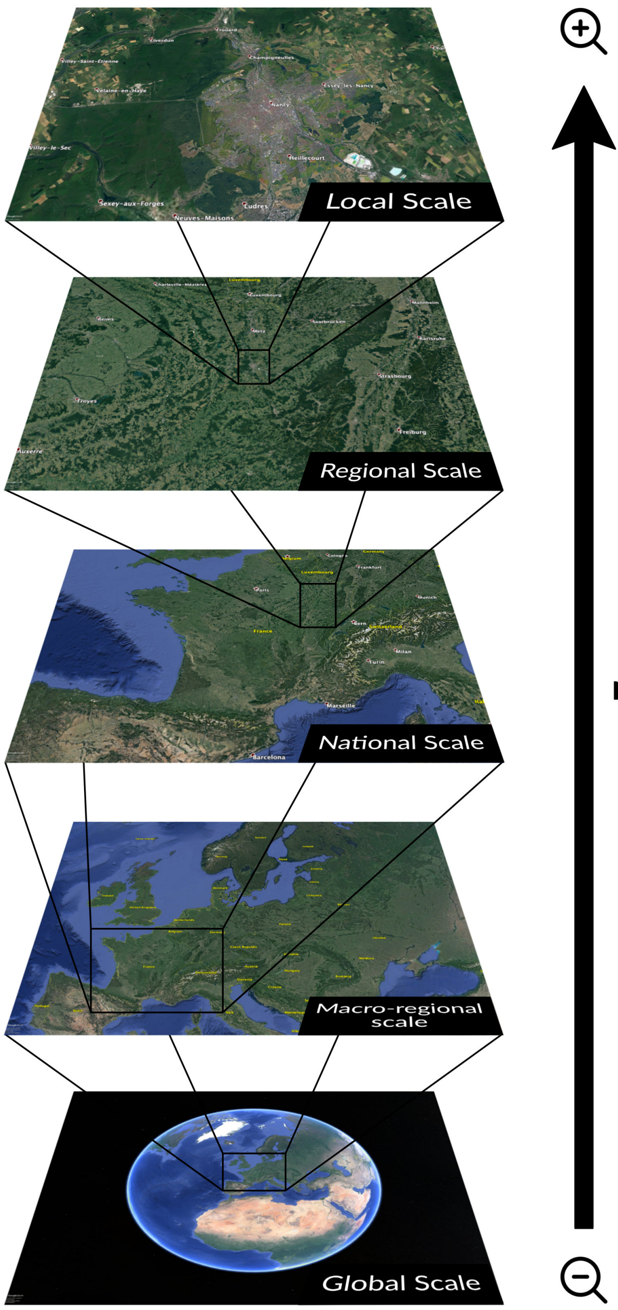


Figure 2.2: A representation of the concept of geographical scales. We can see that each scale is nested into another and that inter-connections exist between global, macro-regional, national, regional, and local scales.

Conceptual Analysis

3 A conceptual analysis of Forest Sector Models' economic structure from the perspective of the three notions of economic geography

Now that we have well defined the notions of space, place, and scale, we can use them as an analytical key to assess how FSM represent spatial features via their economic structure.

The motivation of this Section is twofold. First, we aim to present the principal international trade theories on which our work focuses on: namely Samuelson's, Armington's, Krugman's, and Melitz's theories.[5] We explore each framework detailing its main assumptions, its analytical expressions, and its implications. Table 3.1 summarizes all theories' main assumptions in page 41. Second, we study how each theory represents space from the perspective of the three geographical notions that we mentioned in Section 2. This conceptual analysis leads naturally to a comparison of each theories' spatial representation.

In what follows, the indexes r and s refer to regions of the world, while the indexes j and k represent a product in Armington's and Krugman's theories respectively. Mathematical demonstrations of Armington's and Krugman's approaches are available in Appendices C and D.

3.1 The spatial equilibrium of Samuelson (1952)

In this subsection, we present the theory developed by Samuelson in *Spatial Price Equilibrium and Linear Programming* written in 1952. Samuelson's framework will constitute a point of reference for the rest of this section.

3.1.1 Samuelson's conceptual framework and main assumptions: the emergence of spatial problems in international trade theories

Samuelson's theory aims to mathematically formulate the problem of the "communication of two spatially separate markets" in the context of perfect competition. In particular, Samuelson's framework converts this problem into a maximization problem. It formulates the problem as follows:

"[...] we are given at each of two or more localities a domestic demand and supply curve for a given product (e.g., wheat) in terms of its market price at that locality. We are also given constant transport

[5] In what follows, we consider Samuelson's and Armington's framework as "early theories of international trade" while we refer to Krugman's and Melitz's approaches as "modern theories of international trade".

costs (shipping, insurance, duties, etc.) for carrying one unit of the product between any two of the specified localities. What then will be the final competitive equilibrium of prices in all the markets, of amounts supplied and demanded at each place, and of exports and imports?" (Samuelson, 1952)

3.1.2 Analytical description of Samuelson's theory

While Samuelson's theory mainly uses graphic resolutions, we aim here to provide the key intuition behind this framework without multiplying the figures. We start with a simple two locations case and then generalize the approach to n regions.

The two locations case

We consider two regions of the world, region 1 and 2, that support two distinct markets for a given product, market 1 and 2 respectively. Without any trade, supply and demand in each market respectively meet in domestic price p_1 and p_2.

Let's assume than p_1 is lower than p_2. Thus, if we now suppose that products may flow between the two regions, trade will only flow from region 1, to region 2, since the pretrade price is lower in region 1. We then do not consider products flows from region 2 to region 1.

We also assume that when region 1, ships products to region 2, it occurs a transport cost of $\tau_{1,2}$ per unit of product shipped. With this specification, flows only occur if $p_2 - p_1$ exceeds $\tau_{1,2}$.

The same reasoning applies if the domestic price is lower in region 2. Samuelson summarizes all the equilibrium conditions as follows:

If $\quad p_2 > p_1 + \tau_{1,2}$,

then non-negative product flows from region 1, to region 2, noted $q_{1,2}$, may occur and, at the equilibrium, $q_{1,2} > 0$ implies $p_2 = p_1 + \tau_{1,2}$.

If $\quad p_2 < p_1 + \tau_{1,2}$ and if $p_1 < p_2 + \tau_{2,1}$,

then there is no trade between the two regions.

If $\quad p_1 > p_2 + \tau_{2,1}$,

then non-negative product flows from region 2, to region 1, noted $q_{2,1}$, may occur and, at the equilibrium, $q_{2,1} > 0$ implies $p_1 = p_2 + \tau_{2,1}$.

To formulate what precedes as a maximization problem, Samuelson considers the excess-supply functions of each region[6] and their differential. The area under the excess-supply curve of a given region gives us the social pay-off of this region, while the area under the differential between excess-supply curves measures the combined social pay-off of both region.

We obtain the equilibrium when the net excess-supply curve meets the transport costs curve. This coincides with a maximization of the "net social pay-off",

6 Samuelson defines the excess-supply function of a given region r, noted ES_r, as the lateral subtraction at every price between the demand curve and the supply curve.

noted NSP, and defined as the sum of all social pay-off of each region minus transport costs. This equilibrium derives the corresponding optimal level of exports.

A generalization to N regions

We here consider N regions of the world. We write the product quantities shipped from region r, to region s, as $q_{r,s}$ and the corresponding transport costs $\tau_{r,s}$. We define the social pay-off of a given region r, noted $S_r(q_r)$, as a function of the total exports of the region r, noted q_r; and as the area under the excess-supply function. We also express the transport cost as a function of the quantities shipped from a region r, to a region s, noted $q_{r,s}$, and note it as $t_{r,s}(q_{r,s})$.

We then derive the net social pay-off for n regions as:

$$NSP = \sum_{r=1}^{n} S_r(q_r) - \sum_{r<s}\sum t_{r,s}(q_{r,s}) \tag{3.1}$$

We also can summarize the previous conditions as follows:

If $-\tau_{r,s} \leq p_r - p_s \leq \tau_{r,s}$,

then both inequalities hold only if $q_{r,s} = 0 = q_{s,r}$.

If $p_s > p_r + \tau_{r,s}$,

then non-negative product flows from region r, to region s, noted $q_{r,s}$, may occur and, at the equilibrium, $q_{r,s} > 0$ implies $p_s = p_r + \tau_{r,s}$.

3.1.3 A critical analysis of how Samuelson's theory represents spatial features

With regards to space

Samuelson's framework mainly represents the notion of territoriality through two elements. First, by considering several regions that trade products. Second, by analytically adding transport costs that may partly reflect the physical distances between regions. However, there is no mention of heterogeneity within a given region. In its formulation, the approach does not take regions' boundaries and shape into consideration nor it does consider geographical location.

With regards to the flows representation, we notice that the approach considers flows of undifferentiated products between regions but not the flows that may occur within a given region.

With regards to place

With Samuelson's framework, we only approach the specificity of a given region through its domestic demand and supply and through all resultant variables (domestic price, excess-supply functions etc.). By doing so, this theory does not consider other contexts that shape economic activity but only the economic context.

We can assess that the relations between regions are, in a certain way, specified: flows have an orientation (from a given export region to a given import region), and conditions for existence (flows occur only if the differential between domestic prices exceed transport costs). Again, Samuelson's framework does not describe flows that occur within a region that may bring more specificity to this region.

With regards to scale

There is no direct mention of scale in Samuelson's paper. But, we can assess that the use of the vague term "localities" may apply to different spatial scale.

Figure 3.1 illustrates how Samuelson's theory represents the world.

3.2 The international trade formulated by Armington (1969): a distinction by place of production

We essentially base this Subsection on Armington's seminal paper *A Theory of Demand for Products Distinguished by Place of Production* written in 1969. We aim here to first present a comprehensive synthesis of his approach through an analytical development before examine its representation of space.

3.2.1 Armington's conceptual framework and main assumptions: notions of variety and independence

In contrast to Samuelson's assumptions, Armington's approach assumes that products are differentiated not only by their kind but also by their place of production. Two regions of production thus produce two varieties of the same kind of product that are imperfect substitutes in demand. For instance, the theory considers milk produced in Wisconsin and French milk as two varieties of the same product, milk. The theory assumes a constant elasticity of substitution (CES) between any two varieties competing in any market.[7] In doing so, Armington's theory relaxes the assumption of perfect competition.

The theory thus considers each place of production as the supplier of one variety and as a source of demand for all other varieties. We also assume the preferences of a region for a variety of one given kind to be independent of its purchases of varieties of another kind. Thus, each demand for a particular product represents an isolated market. Armington's approach assumes that the size of one market is a function of money income and of the prices of the various products and affects the demand of a given product.

3.2.2 Analytical description of Armington's theory

We adapt Armington's main equations from Larch and Yotov (2016), and Allen and Arkolakis (2016).

Armington uses CES demand function assumption to derive a quantity index for a given product j, consumed in region s:

[7] Armington additionally assumes that this elasticity is the same for every market.

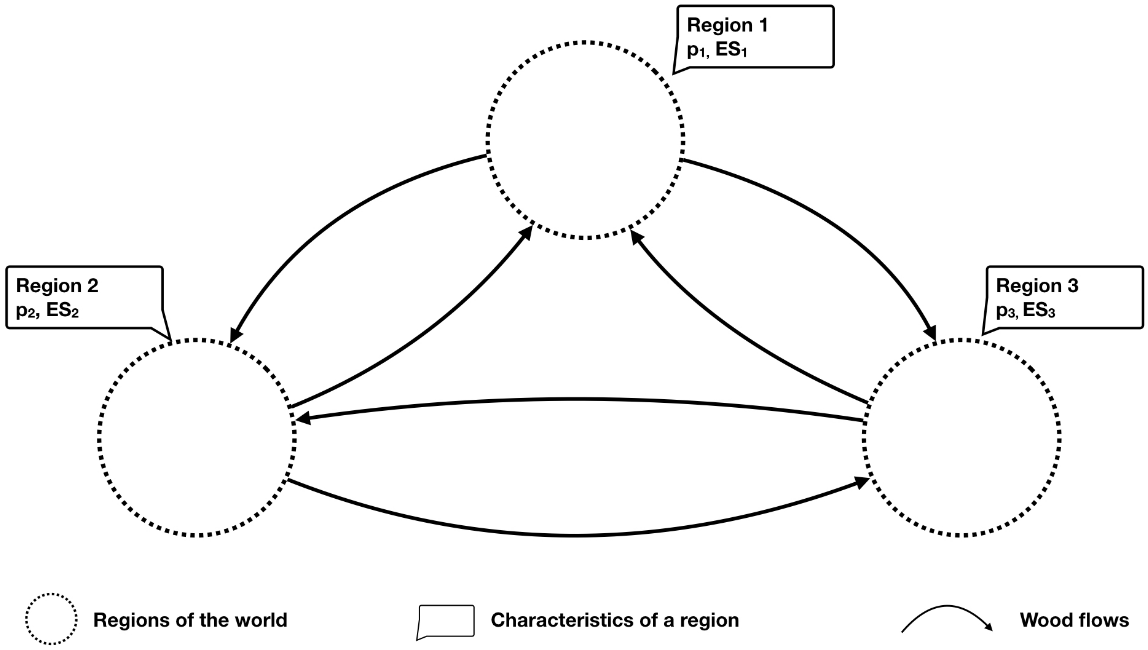


Figure 3.1: An illustration of how Samuelson's framework represents international trade in a world with three regions. Regions do not have clear boundaries nor form. Each region exports/imports wood products to/from the two other regions and has its own characteristics in price and in excess-supply function. Wood flows are not differentiated. To simplify, we do not represent transport costs.

$$M_{j,s} = \left[\sum_r (q_{j,r,s})^{\frac{\sigma - 1}{\sigma}} \right]^{\frac{\sigma}{\sigma - 1}} \tag{3.2}$$

where $M_{j,s}$ is the Armington's quantity index; $q_{j,r,s}$ the quantity consumed by region s, of a product j, produced in region r; and σ the constant elasticity of substitution.

We derive from equation (3.2) the Armington's price index of a region s, for a product j, as follows:

$$G_{j,s} = \left[\sum_r (p_{j,r,s})^{(1 - \sigma)} \right]^{\frac{1}{1 - \sigma}} \tag{3.3}$$

where $p_{j,r,s}$ is the delivered price of of a product j, produced in region r, that is shipped in a consumption region s.

It can thus be shown that the demand of a consumption region s, for a product j, produced in region r, has the following form:

$$q_{j,r,s} = \left(\frac{p_{j,r,s}}{G_{j,s}} \right)^{-\sigma} \times M_{j,s} \tag{3.4}$$

Note that in all these equations, regions r and s, can be a country, a group of countries and so forth, as Armington suggests in his paper.

3.2.3 A critical analysis of how Armington's theory represent spatial features

With regards to space

Armington's international trade theory does not represent territoriality: the regions considered are somehow homogeneous places that do not have any boundaries nor shape. Armington's formulation does not provide any indication about their locations and their relative significance. It does not include any physical distances that may incur transport costs. In a way, Armington's theory makes no physical distinction between the regions considered.

However, the flows between regions have qualitatively changed. Instead of considering flows of a product, like in Samuelson's framework, we now represent flows of varieties of one product. Thus, Armington's theory brings a more precise representation of the flows than Samuelson's. On the other hand, the theory does not represent flows that operate within the country.

With regards to place

Armington's approach gets closer to the specificity of each region by differentiating products by their place of production. Thus, one kind of product produced in a given place somehow becomes a specific variety of that place.

We can also assert that Armington's framework better describes the relationships that link regions between them, for two reasons. First, as previously mentioned, it qualitatively improves the representation of flows. Second, using the quantity and price indexes, Armington's equations imply that the variables describing one region's state depend on the other region's variables (equation (3.4)).

However, another formulation may improve the representation of these relationships in two ways: first, by relaxing the independence hypothesis and then considering the influence one market may have on another; second by representing the flows that occur within a given place.

Note that Armington's theory only deals with the economic sphere and does not explicitly mention any other contexts that shape economic processes, might it be environmental, political, and so forth. Moreover, past activities of a given place have no place in Armington's framework.

With regards to scale

In his paper, Armington explicitly mentions the notion of scale, suggesting that equations (3.4) may apply to a single country or a group of countries. By extension, it also applies to macro-regions, continents, and every scale, we may consider. Unlike Samuelson's theory, Armington's approach clearly includes a representation of the notion of scale.

However, we notice that no matter what scale we consider the theory analytically describes the economic process by the same equation. This seems counterintuitive, as we have previously suggested that economic processes occur differently at different scales. This is because Armington's approach assumes that economic processes are similar at each scale so that a single equation can encompass them.[8]

Figure 3.2 illustrates how Armington's theory approaches the world.

3.3 The core-periphery model of Krugman (1998): the basic model of the New Economic Geography

We base this Subsection on *The Spatial Economy: Cities, Regions and International Trade*, written by Fujita, Krugman, and Venables. Even if this book provides many details about the implications of Krugman's theory, we aim here to focus on the basic approach developed in its chapters 4 and 5. After the development of Krugman's theory, we will assess its representation of *"geographical space"*.

3.3.1 Krugman's conceptual framework and main assumptions: increasing returns under monopolistic competition

Krugman's theory consists in a spatial framework that assumes monopolistic competition and increasing returns of scale at the firm level. To do so, Krugman's approach spatializes the Dixit-Stiglitz model of monopolistic competition that provides its basis by adding locations of production and consumption and by including Samuelson's iceberg transport costs.

The approach considers two sectors:

1. An agricultural sector under perfect competition producing a unique and homogeneous product that can be shipped without incurring any transport costs.

[8] If economic processes are close from one scale to another, there is also no need to describe the inter-connections that relate each level.

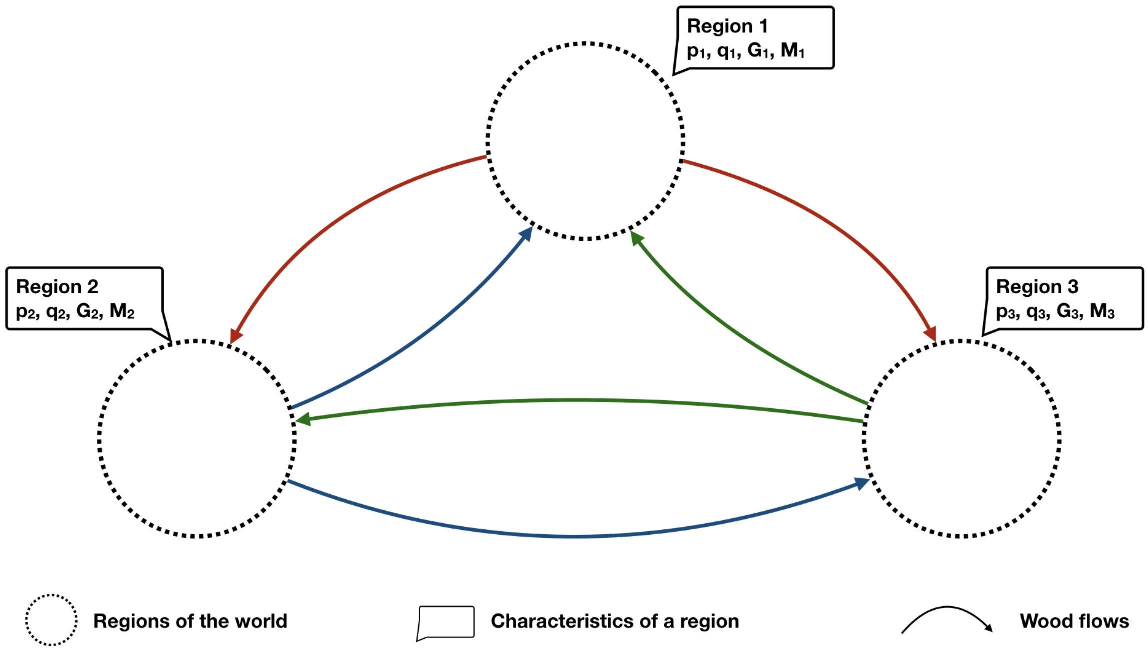


Figure 3.2: An illustration of how Armington's framework represents international trade in a world with three regions. Regions do not have clear boundaries nor form. Each region exports/imports wood products to/from the two other regions and has its own characteristics in price, quantities etc. Wood flows are differentiated by regions (we represent product differentiation by different colors).

2. A manufacturing sector under monopolistic competition that provides many varieties of manufactured products and that includes increasing return of scale and iceberg transport costs.

Each consumer purchases agricultural and manufactured goods. The expenditure share of manufactured products is constant, noted μ.

In contrast with Armington's approach, Krugman has built his theory at the firm level and assumes that products are differentiated at the firm level. Thus, each firm produces a variety of a given product which is an imperfect substitute of another variety produced by another firm. Using the Dixit-Stiglitz framework implies CES demand functions, like in Armington's.

All manufacturing firms use the same technology and then have the same kind of production function. Moreover, the quantity they produce and the associated price are the same between all firms that are in the same region.

3.3.2 Analytical description of Krugman's theory

A spatial version of the Dixit-Stiglitz model of monopolistic competition

Even if the theory considers two sectors, all the equations that follow refer to the manufactured sector. When a variable refers to the agricultural sector, it is noted with a superscript "A".

We apply the main results of the Dixit-Stiglitz model of monopolistic competition (developed in Appendix D) for a finite set of locations and we add Samuelson's iceberg transport costs.

Iceberg transport costs imply that when a unit of a product k, is shipped from a production region r, to a consumption region s, only a fraction of the original unit actually arrives: the rest melts on the way. We note $\tau_{k,r,s}$ the amount the product k, dispatched per unit received: it represents the iceberg transport costs.

Under this assumption, the delivered price in a region s, of a product k, produced in region r, noted $p_{k,r,s}$, is given by:

$$p_{k,r,s} = p_{k,r} \times \tau_{k,r,s} \tag{3.5}$$

where $p_{k,r}$ is the selling price of a product k, in region r, and $\tau_{k,r,s}$ is the iceberg transport cost incurred while shipping a product k, from a production region r, to a consumption region s.

Under those considerations, Krugman's price index takes the following form:

$$G_{k,s} = \left[\sum_r N_{k,r} \times (p_{k,r} \times \tau_{k,r,s})^{1-\sigma} \right]^{\frac{1}{1-\sigma}} \tag{3.6}$$

where $G_{k,s}$ is the Krugman's price index of a product k, consumed in region s; $N_{k,r}$ is the number of companies/number of varieties of a product k, produced in region r; and σ is the constant elasticity of substitution between any two pairs of varieties.

Under the Dixit-Stiglitz framework, the maximization of the consumer utility implies that the consumption demand in region s, for a product k, produced in region r, noted $q_{k,r,s}$, can be written as:

$$q_{k,r,s} = \mu \times Y_{k,s} \times (p_{k,r} \times \tau_{k,r,s})^{-\sigma} \times G_{k,s}^{(\sigma-1)} \tag{3.7}$$

where $Y_{k,s}$ is the income of region s generated by the sales of a product k.

To supply this level of demand for one region s, a firm in region r, has to ship $\tau_{k,r,s}$ times the amount $q_{k,r,s}$ because of iceberg transport costs. Summing across all consumption locations, we derive the total production of a single firm producing a variety in region r, noted $q_{k,r}^{tot}$, as follows:

$$q_{k,r}^{tot} = \mu \times \sum_s Y_{k,s} \times (p_{k,r} \times \tau_{k,r,s})^{-\sigma} \times G_{k,s}^{\sigma-1} \times \tau_{k,r,s} \tag{3.8}$$

All firms in all locations having the same technology, they involve the same fixed input F and marginal input requirement c. We assume the only input is labor. Profit maximization thus leads to the following formulation of the price $p_{k,r}$:

$$p_{k,r} \times \left(1 - \frac{1}{\sigma}\right) = c \times w_{k,r} \tag{3.9}$$

where $w_{k,r}$ is the wage rate for manufacturing workers that produce a product k, in region r.

Zero-profit condition then provides the equilibrium output of every firm, noted $q^*_{k,r}$:

$$q^*_{k,r} = F \times \frac{\sigma - 1}{c} \quad (3.10)$$

Considering what precedes, we can thus show that the wage equation is given by the following expression:

$$w_{k,r} = \left(\frac{\sigma - 1}{\sigma \times c}\right) \left[\frac{\mu}{q^*_{k,r}} \times \sum_s Y_{k,s} \times (\tau_{k,r,s})^{1-\sigma} \times G_{k,s}^{(\sigma-1)}\right]^{\frac{1}{\sigma}} \quad (3.11)$$

We now define the real wages of workers producing a product k, in region r, noted $\omega_{k,r}$, as the nominal wages deflated by the cost-of-living index, expressed as $G_{k,r}^{\mu} \times (p_r^A)^{(1-\mu)}$. We obtain the following formulation:

$$\omega_{k,r} = w_{k,r} \times G_{k,r}^{\mu} \times (p_r^A)^{(1-\mu)} \quad (3.12)$$

The core and periphery model

To derive Krugman's core-periphery model, we start by adding two normalisations that we present in Appendix D. We then assume that the world has $L_k^M = \mu$ manufacturing workers producing a product k, and $L^A = 1 - \mu$ farmers. We denote one region r, share of farmers world labor force by ϕ_r, and its share of workers world labor force producing a product k, by $\lambda_{k,r}$. We also assume that farmers wages are all equal to unit.

Under those assumptions, the Krugman's equilibrium is then defined as the simultaneous solution of the four following equations of:

The income

$$Y_{k,r} = \mu \times \lambda_{k,r} \times w_{k,r} + (1-\mu) \times \phi_r \quad (3.13)$$

The price index

$$G_{k,r} = \left[\sum_s \lambda_{k,s} \times (w_{k,s} \times \tau_{k,r,s})^{1-\sigma}\right]^{\frac{1}{1-\sigma}} \quad (3.14)$$

The nominal wages

$$w_{k,r} = \left[\sum_s Y_{k,s} \times (\tau_{k,r,s})^{1-\sigma} \times G_{k,s}^{\sigma-1}\right]^{\frac{1}{\sigma}} \quad (3.15)$$

The real wages

$$\omega_{k,r} = w_{k,r} \times G_{k,r}^{-\mu} \quad (3.16)$$

The model can then describes agglomeration by assuming that workers are moving to a region where the real wage is higher. The dynamics can be expressed by the following equation:

$$\dot{\lambda}_{k,r} = \gamma \times (\omega_{k,r} - \overline{\omega}) \times \lambda_{k,r} \quad (3.17)$$

where γ is a constant and $\overline{\omega}$ the average real wage generated by one product k, defined as follows:

$$\overline{\omega}_k = \sum_r \lambda_{k,r} \times \omega_{k,r} \quad (3.18)$$

3.3.3 A critical analysis of how Krugman's theory represent spatial features

With regards to space

Like Armington's, Krugman international trade theory does not explicitly represent territoriality: the regions considered do not have explicit boundaries, nor shape, nor physical heterogeneity. Even if Fujita et al. (1999) provides some extensions that include abstract locations and relative significances, the core theory previously presented makes no mention of it.

One great improvement in the representation of space is the inclusion of iceberg transport costs, for two reasons. First, the notion of transport costs closely relates to physical distance. Second, *"iceberg"* kind of transport costs is a black box that may implicitly take physical factors that impact transportation into consideration: the density of the road system, rivers to cross, mountains, and so forth. In doing so, Krugman's approach implicitly differentiates regions.

The flows between regions also qualitatively change compare to Armington's framework. Instead of considering flows at the region level, we now represent flows of varieties produced at the firm level. Thus, Krugman's theory defines international flows at a finer-grained scale. Surprisingly, while Krugman's framework partially describes the inner structure of a region through its number of companies, there is no mention of any intra-region relationships or trade.

With regards to place

Krugman's framework brings more specificity than Armington's to a given location by considering its number of firms. Differentiation by firms of production means that regions of the world now produce a *specific* set of varieties, which improves place representation.

In doing so, Krugman's work offers a better flows representation but remains somehow close to Armington's framework in two ways. First, it still approaches mutual dependence and influence between regions through quantity and price indexes. Second, there is still no representation of the flows that operate within a region.

Another improvement lies in the intersector relationships. Even if markets under monopolistic competition seem isolated, Krugman's approach considers two sectors and then explores the relationships that may exist between them.

Krugman's theory mainly deals with the economic sphere but can consider other contexts that shape economic processes through the notion of iceberg transport costs. As previously said, iceberg transport costs are a black box and thus may implicitly include other factors that affect transportation cost. We may think about, for instance, a difference of culture, a language barrier, a protectionist behavior, and so forth.

Moreover, through equation (3.17), Krugman's model includes a temporal dynamic that depends on the previous state of a region. Even though it mainly concerns the evolution of labor force share in different regions, this substantially improves the representation of space in international trade theories.

With regards to scale

We notice that there is no clear definition of a *"region"* in Krugman's work. We then can imagine a region of the world as a country, a group of countries, a continent, etc. We thus comment Krugman's scale representation in the same way we did for Armington's approach, assessing that Krugman's theory implicitly assumes economic processes are similar at each scale, which is counter-intuitive with regard to our geographical approach.

Thus, Krugman's theory goes further in space and place representations than Armington's framework, notably by using the "trick" of iceberg transport cost, as Krugman likes to say. We however still not reach full representation of the three complex notions of space, place, and scale. Figure 3.3 illustrates of how Krugman's theory describes the world.

3.4 The international trade formulated by Melitz (2003): including firms with different productivities

We aim here to provide the intuition behind the work of Melitz and then will not derive the analytical expressions of his theory. What follows is based on *The Impact of Trade on Intra-industry Reallocations and Aggregate Industry Productivity*, written by Melitz in 2003 and does not pretend to bring a complete statement of Melitz's theory.

3.4.1 Melitz's conceptual framework and main assumptions: heterogeneous firms under monopolistic competition

Melitz also bases his theory on the Dixit-Stiglitz model of monopolistic competition. But, instead of assuming that firms are homogeneous like Krugman's approach does, it assumes heterogeneity of productivity among firms. Moreover, this theory reproduces many empirical patterns revealed by micro-level studies like, for instance, the fact that "relatively more productive firms are much more likely to export" and somehow force "the least productive firms to exit".

In other words, Melitz's approach is an extension of Krugman's approach that gets closer to international trade empirical observations by relaxing the assumption of homogeneous firms.

3.4.2 A critical analysis of how Melitz's theory represent spatial features

Melitz's representation of the notions space, place and scale is close to Krugman's. However, by differentiating firms by their productivity Melitz's approach induces that while some productive firms will enter the export market, some least productive firms will produce only for the domestic market.

In doing so, Melitz better represents the inner flows of a region, qualitatively improves the overall representation of flows, and brings more regioanl speci-

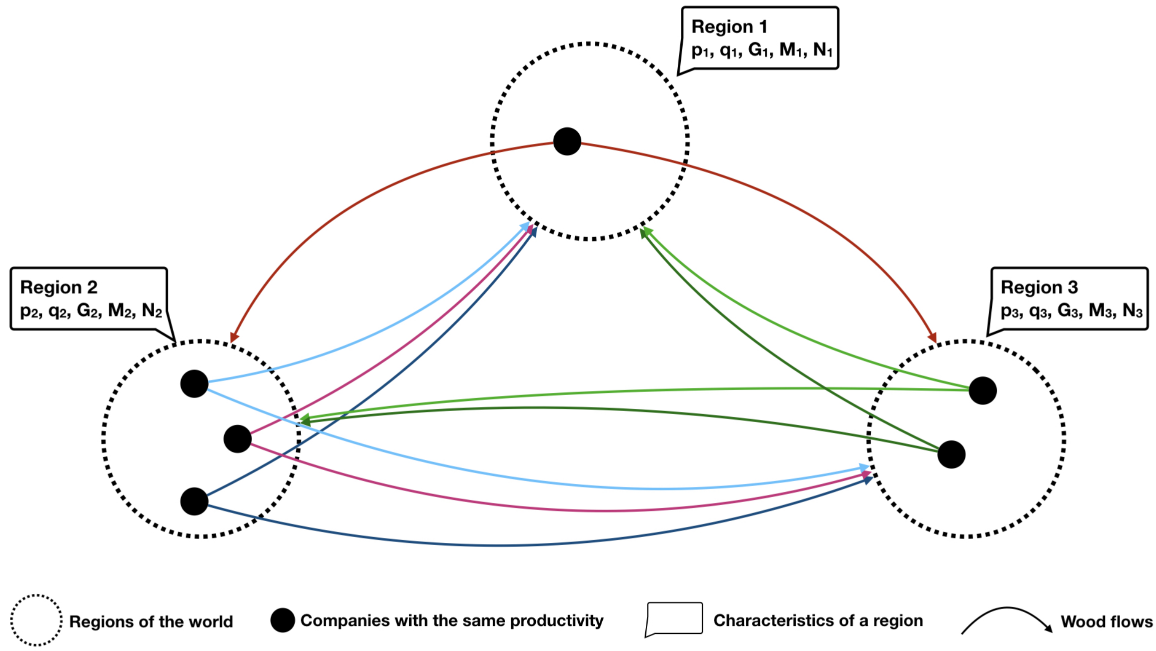


Figure 3.3: An illustration of how Krugman's framework represents international trade in a world with three regions. Regions do not have clear boundaries nor form. Each region has its own characteristics in price, quantities etc. Among those characteristics: a specific number of firms which have the same technology and productivity. Thus, firms that belong to the same region have the same characteristics in price, quantities etc. Wood flows are differentiated by firm (we represent product differentiation by different colors). Each firms of a given region exports wood products to the two other regions. To simplify, we do not represent transport costs.

ficity. Melitz's model also includes temporal dynamics, as Krugman's does.

It thus goes further than Krugman's framework in the representation of *"geographical space"*. Figure 3.4 represents Melitz's vision of the world.

3.5 Taking a step back to look at how international trade theories globally represent spatial features

Putting what precedes into perspective, we can assert that international trade theories have substantially improved the way they represent "geographical space" in economics. In particular, and from the assessment of how each theory represents its different components - namely space, place and scale - we can distinguish three main improvements

First, economists built modern international trade theories at a finer scale than early ones. This leads to a better specification of locations and an improved qualitative representation of flows. Second, they have included physical distance through the notion of transport costs, as

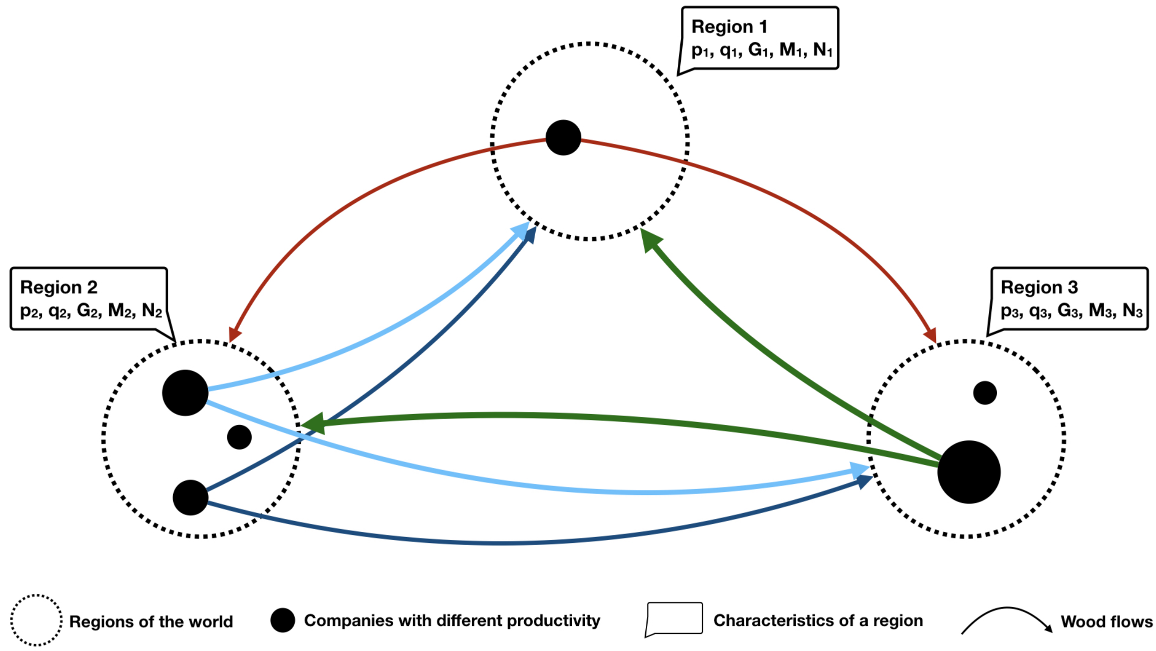


Figure 3.4: An illustration of how Melitz's framework represents international trade in a world with three regions. Regions do not have clear boundaries nor form. Each region has its own characteristics in price, quantities etc. Among those characteristics: a specific number of firms which have different technology. Firms that belong to the same region have the same environment but different productivity. Wood flows are differentiated by firm (we represent product differentiation by different colors). The most productive firms of a given region will export wood products to the two other regions, forcing the less productive firms to exit the export market. To simplify, we do not represent transport costs.

well as, yet implicitly, potential factors that may impact those costs, might it be physical factors or not. Third, as early as Samuelson's trade theory, they allow for considering different scales, but through a way that does not satisfies our conception of scale.

However, despite all the improvements we noted from the early theory of Samuelson to the New Economic Geography of Krugman, we still identify an incomplete representation of *"geographical space"* in economic theories. Why economic theories do seem to somehow neglect spatial representation?

A reason is that even if *"geographical space"* has been considered crucial by the founding fathers of economic theories, including spatial features in the existing theoretical frameworks is analytically difficult and lead to some methodological problems (Thisse and Walliser, 1998).

Another explanation lies in the dual nature of Economic Geography. Every discipline that aims to bridge two distinct disciplines, as Economic Geography does with economics and geography, may have a particular penchant for one side or the other. This duality can lead to different approaches. When Géneau de Lamarlière and Staszak attend to define what is Economic Geography in *Principes de géographie économique*, they propose two definitions of the discipline.

The first one relates to the *"economic side"* of Economic Geography and emphasizes the inter-relation that exists between the behavior of the rational economic agent, *Homo œconomicus*, and *"geographical space"*. This conception of Economic Geography mainly focuses on the spatial consequences economic agents may induce and how *"geographical space"* affects their behavior in return. This conception forms the basis of the New Economic Geography developed by P. Krugman.

The second one falls within the *"geographical side"* of Economy Geography. This approach aims to study an economy and focuses on, for instance, resources allocation, production localizations, the direction of trade flows, and so forth. To do so, it uses economics, but also history, sociology, ethnology and then considers a diversity of factors to explain economic processes. The notions we defined in the previous Section come from this conception of Economic Geography.[9]

This reason, in addition to explain the gap we have observed, indicates that Economic Geography encompasses a large spectrum of various methods, almost all gravitating around one goal: including spatial consideration in studying economic processes.

[9] In that respect, it is interesting to see that the French language makes a difference between *"la géographie économique"* (Economic Geography) and *"l'économie géographique"* ("Geographic Economics") depending on the approach considered, while the English language uses Economic Geography alone.

Assumptions	**Samuelson's theory**	**Armington's theory**	**Krugman's theory**	**Melitz's theory**
Economic environment	Perfect competition	Imperfect competition	Monopolistic competition	Monopolistic competition
Returns of scale	Constant returns of scale	Constant returns of scale	Increasing returns of scale	Increasing returns of scale
Product differentiation	No differentiation	Product differentiated by regions with a constant elasticity of substitution	Product differentiated by firms with a constant elasticity of substitution	Product differentiated by firms with a constant elasticity of substitution
Inclusion of transport costs	Fixed transport costs between two given regions	No transport costs	Iceberg transport costs *à la* Samuelson	Iceberg transport costs *à la* Samuelson
Exporters	Regions	Regions	Firms that have the same productivity	Firms that have different productivities
Importers	Regions	Regions	Regions	Regions

Table 3.1: A comparison of the main assumptions of international trade theories as they are explicitly formulated in their original framework.

Modelisation

4 Shifting from traditional economic theories to New Economic Geography's theory: towards a better representation of spatial features in Forest Sector Modelling

The previous Section of this report explored new theories of international trade to better represent spatial aspects in modeling economic processes. In particular, we showed that a shift from former economic theories - in particular, those that P. Samuelson and P. Armington formulated in 1952 and 1969 respectively - to more modern economic theories could improve the way FSM represents wood flows.

This Section therefore assumes that changing the core economic frameworks of FSM could help the modeler in mitigating, even perhaps overcoming, actual caveats that (s)he encounters. Besides, if it doesn't improve the models that much in that way, replacing the economic structure of FSM may open new perspectives for modeling the forest sector and more especially the wood trade.

To test those assumptions, we choose to focus on Armington's and Krugman's theories. There are two good reasons for doing so. First, a "classic" international trade theory like Armington's and a "modern" one like Krugman's well compare. Indeed, they are built on similar structures that make the comparison more tractable and intelligible.[10] Second, we previously showed that Krugman's theory brought a substantial theoretical improvement, while Armington's framework is somehow an enhancement of Samuelson's theory, and Melitz's framework a refinement of Krugman's theory.

The comparison will mainly rely on two GAMS codes (complete codes in GAMS language are provided in Appendix F) that compute simplified versions of Armington's and Krugman's equilibria using the same set of data and a similar set of parameters. These codes have been inspired by the seminal paper or Rutherford and Balistreri that adapts P. Armington's, P. Krugman's and M. Melitz's theories in a general equilibrium model (Balistreri and Rutherford, 2013)).[11]

[10] The frameworks of Samuelson's and Melitz's theories are so different that it would be hard to clearly identify what difference in the structure would make one theory react differently than the other.

[11] The equations essentially derive from sub-sections 3.2 and 3.3 and are detailed in next Subsection 4.1 with the calibration of the model and a quick description of the data used for the computation.

4.1 Toward comparable versions of Armington's and Krugman's models

4.1.1 Goal and scope

Our aim is to use the computational tool to project the effects of a change in the economic structure on the models' outputs. Since the theories introduced in sub-sections 3.2 and 3.3 do not directly compare, our objective is to adapt and slightly reformulate both of the theories in order to conduct a relevant comparison.

In this way, we first need to incorporate Samuelson iceberg costs in Armington's theory and to define explicitly an income equation. Then, we may readily draw a parallel between both theories' equations after a few manipulations and adjustments of Krugman's equations. In the end, we would be able to unambiguously read the change of structure inside the final sets of equations.

To exemplify our work, we will calibrate the model using real data from the publicly available FAOSTAT database.[12] The next Subsection is devoted to those practical aspects.

4.1.2 Integrating iceberg transport costs in Armington's theory to get closer to Krugman's theory

Krugman's theory significantly differs from Armington's by integrating iceberg transport costs. Since this assumption is likely to affect the equilibrium outcome of a theory, we want to include it in Armington's model to disentangle the effects of a change in structure from the effects of integrating iceberg costs *per se*.

We thus formulate a new version of Armington's theory that also assumes iceberg transport costs *à la* Samuelson. To do so, we stick to the definition of this type of transport cost we have formulated in Subsection 3.3: when one unit of a product, j, is shipped from a region, r, to another region, s, a part of it is "melting on the way" so that only $1/\tau_{j,r,s}$ unit of this product actually arrives in the consumption region, s. In other words, if the consumption region s, is about to receive one unit of a product j, produced in region r, then region r, should send $\tau_{j,r,s}$ units of this product to region s: it is the quantity of product dispatched per unit received (we suppose $\tau_{j,r,s} \geq 1$).

We can analytically formulate this definition as follows:

$$p_{k,r,s} = p_{k,r} \times \tau_{k,r,s} \tag{4.1}$$

where $p_{j,r,s}$ is the price at the consumption region, s, of a product, j, produced in region, r; $p_{j,r}$ is the selling price of a

[12] FAOSTAT data are available on `www.fao.org/faostat/en/#data`

product j, produced in the region r; and $\tau_{j,r,s}$ the iceberg transport costs of delivering a product j, produced in region r, to the consumption region s.

Analytically incorporating the iceberg transport costs assumption to Armington's previous set of equations is actually quite straightforward if we use equation (4.1).

Armington's price index formulation with iceberg transport costs

We naturally use equation (4.1) in the classic CES consumer price index expressed in equation (3.3). This leads to the following Armington's price index formulation that analytically reveals the iceberg transport costs:

$$G_{j,s} = \left[\sum_r (p_{j,r} \times \tau_{j,r,s})^{(1-\sigma)} \right]^{\frac{1}{1-\sigma}} \tag{4.2}$$

where $G_{j,s}$ is the price index of a product j, in region s, and σ the constant elasticity of substitution between any two pairs of products.

Armington's quantity index formulation

Adding the assumption of iceberg transport costs in Armington's theory change the quantity index equation since only $1/\tau_{j,r,s}$ unit of the quantities shipped from the production regions r, to the consumption region s, noted $q_{j,r,s}$. We then have the following formulation:

$$M_{j,s} = \left[\sum_r \left(\frac{q_{j,r,s}}{\tau_{j,r,s}} \right)^{\frac{\sigma-1}{\sigma}} \right]^{\frac{\sigma}{\sigma-1}} \tag{4.3}$$

where $M_{j,s}$ is the quantity index for a product j, in consumption region s.

An intuitive formulation of the income for Armington's theory

We can define the income as (1) the sum of the revenues in a region - that is to say, the sum of the sales to all other region times the selling price - or (2) the total expenditure - that is to say, the sum of the amount purchased in every region times the purchase price. Following the zero-profit assumption, these two formulations are strictly equivalent: a region should spend all its revenues in purchasing other products, and one region spends the exact amount it earns.

We choose to formulate the income of a region s, as its money expenditure, similarly to Armington (Armington, 1969) but this time with the assumption of iceberg transport costs using equation (4.1):

$$Y_{j,s} = \sum_r q_{j,r,s} \times (p_{j,r} \times \tau_{j,r,s}) \tag{4.4}$$

where $Y_{j,s}$ is the income of the region s.

An equation of Armington's quantity flows between regions that integrates iceberg transport costs

We integrate iceberg transport costs into equation (3.4) of the demand of a consumption region s, for a product j, produced in region r, in two steps. First, we use equation (4.1) to decompose the price $p_{j,r,s}$. Second, according to iceberg transport costs definition, a region r, has to ship $\tau_{j,r,s}$ times the amount needed to supply the demand of a region s, for a product j. The link between interregional quantity flows and iceberg transport costs then becomes explicit:

$$q_{j,r,s} = \left(\frac{p_{j,r} \times \tau_{j,r,s}}{G_{j,s}} \right)^{-\sigma} \times M_{j,s} \times \tau_{j,r,s} \tag{4.5}$$

With this set of four equations, we fully describe Armington's equilibrium with the iceberg transport costs assumption.

4.1.3 Adapting Krugman's theory to make it comparable with Armington's theory

Now that we have integrated the iceberg transport costs into Armington's equations we need to adapt Krugman's equations that are depicted in Subsection 3.3. This consists in two modifications. First, as we consider only one sector, we will omit the agricultural sector by assuming that μ equals unity. Second, we need to write down an equation for the Krugman's quantity index, and to use an income equation that is closer to the one we have for Armington's set of equations.

Krugman's price index formulation

We do not need to adjust the price index formulated in Subsection 3.3, and will keep defining it as follows:

$$G_{k,s} = \left[\sum_r N_{k,r} \times (p^f_{k,r} \times \tau_{k,r,s})^{(1-\sigma)} \right]^{\frac{1}{1-\sigma}} \tag{4.6}$$

where $G_{k,s}$ is the price index of a product k, at location s; $N_{k,r}$ the number of companies or number of varieties in a region r; $p^f_{k,r}$ the firm price of one of these varieties[13]; $\tau_{k,r,s}$ the iceberg transport costs incurred for shipping a good k, from region r, to consumption region s; and σ the constant elasticity of substitution.

Krugman's quantity index formulation

We then need to formulate an equation for the quantity index of a consumption region s. The Dixit-Stiglitz model, on which Krugman's theory is built, assumes CES formulations for the quantities index of consumption. This assumption is also shared by Armington's theory.

Following this, we define this index as a combination of all amount of goods consumed by region s, from every other region and more precisely from all companies from every other region, according to the framework defined by Krugman. We also remind that only a fraction $1/\tau_{k,r,s}$ of the amount of good k, produced by a firm in region r, and shipped to consumption region s, noted $q^f_{k,r,s}$, actually arrives. By analogy with the way Krugman's price index is defined, and knowing that the elasticity of substitution between every pairs of any product is constant, we then can fully express the quantity index of a region s, as follows:

[13] Remember that this price is common to every firm that are located in region r, so that the firm price actually is a "regional price."

$$M_{k,s} = \left[\sum_r N_{k,r} \times \left(\frac{q^f_{k,r,s}}{\tau_{k,r,s}} \right)^{\frac{\sigma-1}{\sigma}} \right]^{\frac{\sigma}{\sigma-1}} \tag{4.7}$$

where $M_{k,s}$ is the quantity index..

Adapting Krugman's income equation

Looking at the income equation that derives from Krugman's theory - formulated in Subsection 3.3 - the link with Armington's income equation (4.4) is not so clear. To draw an unambiguous parallel between the two sets of equations, we aim here to adapt Krugman's income equation.

Remember that there are mainly two ways to define the income of a region: (1) on the supply side, as the total revenues of one region; (2) on the demand side, as the total expenditures of one region. If we look at equation (3.13), we can see that Krugman has used the supply side definition of the income, formulated as the total production cost - that is to say, the labor force used by all the companies of the region times the wages. This is analytically equivalent to total revenues a region gets from the sales of its production.[14]

We propose a new expression of the Krugman's income we draw from Armington's income equation 4.4, using the demand side definition. The income at a location s, is then the total amount of firms production sold at every other location times its price:

$$Y_{k,s} = \sum_r (q^f_{k,r,s} \times N_{k,r}) \times (p_{k,r} \times \tau_{k,r,s}) \tag{4.8}$$

where $Y_{k,s}$ is the income, or total expenditure, of region s.

Krugman's inter-regional quantity flows equation

This last equation derives from the intermediate equation (3.7) described in Subsection 3.3. To keep our two sets of equation comparable, we do not need to use the total firm production equation (3.8) that is central in Krugman's framework.

We slightly adapt equation (3.7) of the consumption demand in a region s, for a product k, produced in region r, by formulating the income as the product of the quantity index and the price index. According to the definition of iceberg transport costs, a firm in region r, has to send $\tau_{k,r,s}$ times this amount to supply consumption demand in region s, for a product k. That property leads to an expression of the inter-regional firm flows that is quite similar to the Armington's one:

$$q^f_{k,r,s} = \left(\frac{p^f_{k,r} \times \tau_{k,r,s}}{G_{k,s}} \right)^{-\sigma} \times M_{k,s} \times \tau_{k,r,s} \tag{4.9}$$

Knowing that every firms of a region, r, are producing the same quantity $q^f_{k,r,s}$, it is straightforward to express the total quantity flows shipped from a production region r, to a consumption region s, of a

[14] A short demonstration of all forms the income equation can take and their equivalence is developed in Annex X, only using some definitions and the assumption of zero-profit.

product k, by multiplying by the number of firms in region r:

$$\begin{aligned} q_{k,r,s} &= N_{k,r} \times q^f_{k,r,s} \\ &= N_{k,r} \times \left(\frac{p^f_{k,r} \times \tau_{k,r,s}}{G_{k,s}} \right)^{-\sigma} \times M_{k,s} \end{aligned} \tag{4.10}$$

These equations are consistent to specify the Krugman's equilibrium and to compare the two theories. However, two additional equations are required if we want to observe any pattern of agglomeration in our model.

Krugman's additional equations to observe an agglomeration pattern[15]

To model economic agglomeration, we first need to make the firm price an endogenous variable of our model. As shown in Subsection 3.3, the initial formulation Krugman's theory is already considering the firm price as endogenous and formulates it through equation (3.9) as:

$$p^f_{k,r} = c \times w_{k,r} \times \frac{\sigma}{\sigma - 1} \tag{4.11}$$

where c is the marginal cost of production and $w_{k,r}$ the nominal wage of region r.

We also need to introduce the regional wages, which are the driving force of agglomeration, according to Krugman. This new variable can directly derive from previous Subsection 3.3, equation (3.11):

$$\begin{aligned} w_{k,r} = \frac{\sigma - 1}{\sigma \times c} \times \\ \left[\frac{1}{q^f_{eq}} \times \sum_s Y_{k,s} \times (\tau_{k,r,s})^{1-\sigma} \times (G_{k,s})^{\sigma-1} \right]^{\frac{1}{\sigma}} \end{aligned} \tag{4.12}$$

where q^f_{eq} is the equilibrium output of any active firm defined by equation (3.10).

4.1.4 Drawing a parallel between the two theories that highlights the change of structure

The two sets of equations are summarized in Table 4.2.

This table clearly highlights that Krugman's theory is an Armington's theory at a finer-grained scale, the scale of the firm. We can directly read the change of structure in the equations: all Krugman's equations use firm-level variables and are weighted by the number of firms.

In other words, Krugman's approach considers regions of the world as an aggregation of firms that have similar properties like production costs, prices, etc. This is quite convenient for our comparison analysis with Armington's formulation of international trade. In this way, we can use the same region-agglomerate variables to carry out our comparison and have a full comprehension of how we model the change of economic structure.

[15] We do not present any results on agglomeration patterns in the present report since the study of economic agglomeration is still at its premises.

Variable	Armington's formulation	Krugman's formulation
Price index $G_{\cdot,s}$	$G_{j,s} = \left[\sum_r (p_{j,r} \times \tau_{j,r,s})^{(1-\sigma)}\right]^{\frac{1}{1-\sigma}}$ (4.2)	$G_{k,s} = \left[\sum_r N_{k,r} \times (p^f_{k,r} \times \tau_{k,r,s})^{(1-\sigma)}\right]^{\frac{1}{1-\sigma}}$ (4.6)
Quantity index $M_{\cdot,s}$	$M_{j,s} = \left[\sum_r \left(\frac{q_{j,r,s}}{\tau_{j,r,s}}\right)^{\frac{\sigma-1}{\sigma}}\right]^{\frac{\sigma}{\sigma-1}}$ (4.3)	$M_{k,s} = \left[\sum_r N_{k,r} \times \left(\frac{q^f_{k,r,s}}{\tau_{k,r,s}}\right)^{\frac{\sigma-1}{\sigma}}\right]^{\frac{\sigma}{\sigma-1}}$ (4.7)
Income $Y_{\cdot,s}$	$Y_{j,s} = \sum_r q_{j,r,s} \times (p_{j,r} \times \tau_{j,r,s})$ (4.4)	$Y_{k,s} = \sum_r (q^f_{k,r,s} \times N_{k,r}) \times (p_{k,r} \times \tau_{k,r,s})$ (4.8)
Inter-regional flows $q_{\cdot,r,s}$	$q_{j,r,s} = \left(\frac{p_{j,r} \times \tau_{j,r,s}}{G_{j,s}}\right)^{-\sigma} \times M_{j,s} \times \tau_{k,r,s}$ (4.5)	$q_{k,r,s} = N_{k,r} \times \left(\frac{p^f_{k,r} \times \tau_{k,r,s}}{G_{k,s}}\right)^{-\sigma} \times M_{k,s} \times \tau_{k,r,s}$ (4.10)

Table 4.2: Two comparable sets of four equations that describe Armington's and Krugman's equilibria.

4.1.5 Additional constrain to the models

While we have fully specified Armington's and Krugman's equilibria in what precedes, it appears that our models may collapse[16] when we slightly change the value of a given parameter.

It is hard to analytically explain this response. We have assumed the models collapse because we did not constrain them enough and that too many degrees of freedom may lead to collapsing. As this problem does not allow us to use the outputs of the models for comparison, and as we want them to keep the same orders of magnitude than the real data, we choose to fix the quantity index variable value, $M_{\cdot,s}$, to its initial value, $M^0_{\cdot,s}$.[17]

The reason we choose this variable is for its place in the models' structure. As we can see on Table 4.2 and on Figure 4.1 that illustrates the links between variables and parameters, the quantity index only affects the quantities shipped between regions and depends on these same quantities and on the three parameters on which we focus our analysis. By fixing its value, we thus make sure the quantities shipped between two regions and then the all system will not collapse and that a sensitivity analysis on the different parameters is still possible.

16 By collapsing, we mean that all the outputs of our model tend to 0 or to the lower boundaries we put on variables.

17 Even if we fix the quantity index value, some collapses may happen when variables reach their lower boundary. This occurs for "extreme" parameters values and does not compromise our analysis.

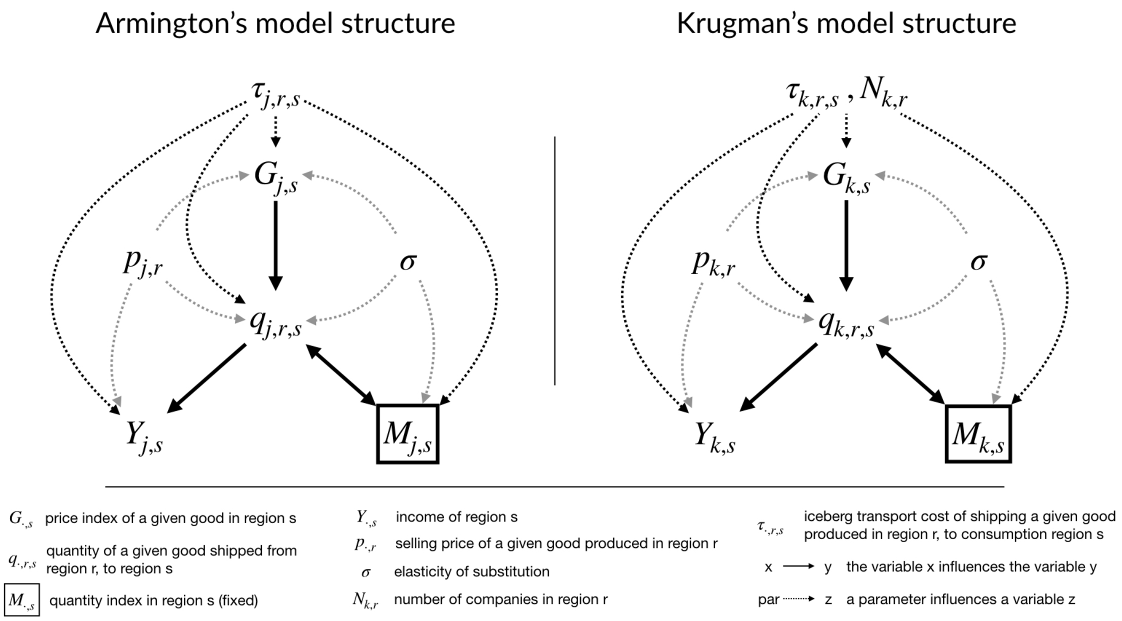


Figure 4.1: A representation of the structure of Armington's and Krugman's models. By looking at the links between variables and parameters, we can see the role of the quantity index as a pivotal variable. First, it only connects with the quantities shipped between regions and then will not *"freeze"* all variables when fixed. Second, as this variable directly links with the elasticity of substitution, the iceberg transport costs and the number of companies in the case of Krugman's model, fixing it will let the system being sensitive to those parameters.

4.2 A joint calibration of the models based on the FAOSTAT database

We aim to work on a concrete case. Thus, we use data from the publicly available database of the FAO named FAOSTAT: `http://www.fao.org/faostat/en/#data`. We describe this database and how we process it in what follows.

For the sake of clarity in the comparison, we also try to compute very close calibrations for both models. The details of those calibration methods follow the description of the FAOSTAT data.

4.2.1 Description of the FAOSTAT data and its processing

The FAOSTAT database gathers a huge data amount with regards to different domains, might it be the global population, environmental indicators, or even information about a particular sector. With regards to the forest sector, the database contains two main domains: Forestry Production and Trade; and Forestry Trade Flows.

These two domains will provide data for all countries and territories in the world - when available - on roundwood, primary wood, and paper products and for each year from at least 1997 to 2017.[18] Forestry Production and Trade "contains data on the production and trade" in those products and provides five elements: the Production Quantity, the Export Values and Quantities and the Import Values and Quantities; while Forestry Trade Flows “contains data on the bilateral trade flows” and only gives four elements: Export and Import Values and Quantities.

To compute both models, we need to know the amount of wood, for one specific item, that is shipped from one region to another for all regions considered. We also need the quantities self-consumed by a region. To do so, we use the Export and Import Values and Quantities elements from Forestry Trade Flows, that is meant to describe bilateral trade flows. We collect the Production Quantity element from the domain Forestry Production and Trade.

Although we could have based our study on other products that are common between the two domains, we choose to only focus on one joint item - the industrial roundwood from the year 2017 - for two reasons. First, considering different wood products would have complicated our comparison, as every wood product is somehow linked to another by an input-output relation. Including several products requires estimating input-output coefficients and brings an unnecessary complexity to the models. Second, considering only one wood product is already enough for our analysis since we do not aim to study the sensibility of the theories with regards to the number of products considered.

While the data is collected at a national scale, the FAO has aggregated na-

[18] Last check in August 2019.

tional data at a sub-continental or at a continental scale. We can then use from the same database national-scale, sub-continental-scale, and continental-data as input for our models.[19]

At each scale we produce a matrix of international flows from elements of the two domains. Export regions - indexed i - are placed in rows and import regions - indexed j - in columns. Each matrix field contains the export quantity of industrial roundwood shipped from export region, i, to import region, j. We then create a second matrix based on the first one. In this new matrix, we add to first diagonal values the self-consumption quantity of the region, given by the following expression:

$$AC_i = PROD_i - \sum_j EXP_{i,j} \tag{4.13}$$

where AC_i is the self-consumption quantity of the export region i; $PROD_i$ is the production quantity of the export region i; and $EXP_{i,j}$ the export quantity shipped from export region i, to import region j.

We also build third matrix, that fills the fields with the export value from export region i, to import region j. Note that these export values do not consider the self-consumption of a given region and only refer to export quantities.

When processing FAOSTAT data, it is important to take some precautions. First, FAO data make a distinction between different regions of China and considers China, Hong-Kong, Macao, Taïwan and China mainland, to be precise. We recommend to only select China when building the matrix. Second, we have also noticed, when using sub-continental and continental data, that the FAO made some adjustments when aggregating the national-level data. Sadly, the FAO does not detail the adjustment methods used. We then decide to get rid of the adjustment field (named "Others (adjustment)") to keep an unambiguously way to agglomerate national data at sub-continental and continental levels (the continental-level data being a sum of sub-continental-level data, being itself a sum of national-level data). This is still consistent with our goal to compare the two theories since we aim not to describe perfectly the "true" international wood flows but to have initial values that are close to reality.

Even with those precautions, some inconsistencies may still come out. It appears that in some case the total export quantities are greater than the total production of the country, even if we take the imports into consideration. In this case, we assume that the Production Quantity declared by the region is inaccurate and decide to adjust it so that the self-consumption is equal to zero.

19 In what precedes, only continental-data will be used to produce preliminary results. However, we consider to use other scales for future experiences.

4.2.2 Calibration of the two models at sub-continental and continental scales data

The calibrations methods are similar for the two models. The only difference lies in additional parameters that Krugman's theory requires. Note that what follows only applies for the sub-continental and the continental scale.

We assume FAOSTAT data represent an equilibrium and use them as benchmark values for the calibration. We denote all initial values that a particular variable takes by adding a superscript "0".

First, we use the matrix of export and self-consumption quantities previously defined as the benchmark value of the quantity of a product j or k, shipped from the export region r, to consumption region s, noted $q^0_{\cdot,r,s}$; and we use the second matrix of the export values to calibrate the corresponding price $p^0_{\cdot,r,s}$. Thus $q^0_{\cdot,r,r}$ is the self-consumption benchmark value, and the price at a particular location r.[20]

First, we use the matrix of export quantities previously defined as the benchmark value of the quantity of a product j or k, shipped from the export region r, to consumption region s, noted $qv^0_{\cdot,r,s}$; and we use the second matrix of the corresponding export values - noted $v^0_{\cdot,r,s}$ to calibrate the delivered price of a given good shipped from production region r, to consumption region s, noted $p^0_{\cdot,r,s}$.

Secondly, in order to calibrate the selling price of a given product in region r, noted $p^0_{\cdot,r}$, for Armington's and Krugman's models[21], we assume that there are no iceberg transport costs when goods are shipped inside a given region *i.e.* $\tau_{\cdot,r,r} = 1$. Using the previous calibration of $p^0_{\cdot,r,s}$ and equation (4.1), the selling price of a given good in region r, is simply the ratio between the export value and the export quantity in region r, that is to say: $v^0_{\cdot,r,r}/qv^0_{\cdot,r,r}$. Note that this calibration of the selling price is only possible at the sub-continental and continental scales, as the first diagonal of the two matrix considered is not empty.[22]

Combining the calibrations of $p^0_{\cdot,r,s}$ and $p^0_{\cdot,r}$, we directly derive of the iceberg transport costs of a given product shipped from region r, to region s, noted $\tau_{\cdot,r,s}$, from equation (4.1):

$$\tau_{\cdot,r,s} = \frac{p^0_{\cdot,r,s}}{p^0_{\cdot,r}} \tag{4.14}$$

Additional parameters have to be added to calibrate Krugman's model,

[20] Note that Krugman's approach defines firm-level price by region so that the price calibration is identical between the two models

[21] Note that Krugman's approach defines firm-level price by region so that the price calibration is identical between the two models

[22] Indeed, as the sub-continental and continental scales are aggregations of countries, the first diagonal contains values different that may differ from 0. Indeed, some European countries may export products to other European countries. It then appears that "Europe is exporting to Europe", for instance. This property helps us calibrating the price

among them: the initial number of companies of a region, r - $N^0_{k,r}$ - and the marginal cost of production - c. We arbitrarily fix the number of companies of a given region and change these numbers during some experiences.[23] The marginal cost value is defined as a unit.

Krugman's model uses a firm-level quantity variable that requires calibration as well. To do so, we use equations (4.9) and (4.10) that lead us to the firm-level quantity benchmark value:

$$q^{f,0}_{k,r,s} = \frac{q^0_{k,r,s}}{N^0_{k,r}} \tag{4.15}$$

Lastly, all other variables benchmark values are given by previous equations: the quantity index benchmark value, $M^0_{\cdot,s}$, is given by equations (4.3) and (4.7); the price index benchmark value, $G^0_{\cdot,s}$, derives from equations (4.2) and (4.6); the initial income, $Y^0_{\cdot,s}$, is defined by equations (4.4) and (4.8); and the wages initial values, $w^0_{k,r}$, follow equation (4.12).

4.3 Are new international trade theories less sensitive to some parameter changes? A sensitivity analysis of the two economic structures

4.3.1 Assumptions on the sensitivity of both theory and objective of the experience

We here want to test the following assumption:

Krugman's theory is less sensitive to some key parameters of Forest Sector Models that are, today, difficult to precisely estimate. These parameters encompass the elasticity of substitution; the transport costs and the number of firms per region. Moving from an economic structure *à la* Armington to a framework *à la* Krugman could thus make those models less sensitive regarding these parameters.

Note that, if the elasticity of substitution and the transport costs[24] may be quite difficult to estimate, the number of firms may be in some cases well informed according to the region considered. Even though, we still want to assess the sensitivity of Krugman's model to this parameter as it remains complicated to get this information in some location of the world and as we may argue that the sensitivity

[23] The number of companies in a given region may also change during the computation, especially if we make endogenous the agglomeration phenomenon. However, we do not present such agglomeration pattern in the present report.

[24] By transport costs, we mean the total cost of a transaction, not only due by the physical distance but also due to cultural differences, language barriers, etc. Furthermore, this conception has led P. Krugman to consider Samuelson iceberg transport costs that are quite convenient to analytically write such transport costs.

of Krugman's theory to this parameter is itself interesting. We, of course, do not consider a sensitivity analysis of this last parameter for Armington's theory.

4.3.2 How to test the sensitivity to the elasticity of substitution, the transport costs and the number of firms per region

To test the previous assumption, we compute the Armington's and Krugman's equilibria for different parameters values, *ceteris paribus*, and compare the values of the following output variables: the price and quantity indexes of each region; the income of each region; and the trade flows between regions.

We conduct the sensitivity analysis to the elasticity of substitution and the transport costs parameters for the two models as follows:

1. Regarding **the elasticity of substitution**, we collect the output values of the two models by varying the elasticity of substitution from 1.2 to 6, by 0.2 steps, all other parameters being equal (we fix the number of companies to 10 for every region in Krugman's model).

2. Regarding **the transport costs**, we decrease/increase the transport cost by 10%, 20%, 30%... 100% and collect the output values for comparison.

3. The sensitivity analysis to **the number of firms** will concern only Krugman's model. We first change the number of companies keeping it the same for all regions. We run our model for three different number of firms: 10, 50 and 100; for different values of elasticity of substitution: 2, 4, and 6. We then look into the impact an increase of this number in only one region may have and thus repeat the experience but this time by changing the number of firms in only one region.

4.3.3 Preliminary results: models' sensitivity to the elasticity of substitution

We here only present some preliminary results for the sensitivity analysis on the elasticity of substitution for aggregate variables, namely: the price index, the quantity index, and the regional income. Figure 4.2 presents those results.

We can first notice that all curves for all variables have the same trend, which confirms that our two models have similar structures and are explicitly comparable.

With regards to the regional price index (Figure 4.2a and Figure 4.2b), we can see that, for the two theories, this variable reaches a threshold when the elasticity of substitution is increasing. We observe that this threshold is, for each continent, lower in the Krugman's model. When looking at the inflection point of the logistic curve, we also note the second derivative changes its algebraic sign for higher values of σ in the case of Krugman's model. Moreover, the maximum slope observed in both case for all continents is always higher in Armington's model. For those reasons, we can assert that Krugman's price index is less sensitive to the elasticity of substitution than Armington's.

With regards to the quantity index (Figure 4.2c and Figure 4.2d), we also observe that, in each case, curves reach a threshold when σ is increasing. This threshold is higher for all continents in Krugman's framework, which means the magnitude of change due to a variation of σ is lower than in Armington's case as the quantity index is decreasing when the elasticity of substitution is increasing. The initial slopes of the quantity index curves are also higher in the case of Armington's model and the curves become "flat" for lower values of σ than with Krugman's model. That also suggests a lower sensitivity to the elasticity of substitution in the case of Krugman's quantity index.

Regional income curves are quite similar (Figure 4.2e and Figure 4.2f). The major difference relies on the higher incomes, those of ASIA and EU continents, which present lower levels in the case of Krugman's theory. It then appears than Krugman's framework makes the higher levels of income less sensitive to σ.

Those preliminary results confirm our initial intuition that Krugman's theory would be less sensitive to parameters than Armington's and would then, in a certain way, improve the way we model the forest sector. However, we need more results, especially on the wood flows levels between regions, to conclude. For the moment, these results mostly verify there is no computational problem in the model we previously built.

4.4 Modeling perspectives

The previous subsection ended with preliminary results that suggest there is a need for continuing our modeling work. This section aims to present complementary experiences we are willing to conduct in the future. Beyond these experiences, we are aware that we need a deeper analysis of the results - especially on the quantities shipped from one region to another - to reach our goals.

4.4.1 A need for refining the sensitivity analysis on parameters

In addition to the preliminary results presented in the previous section, we need to run more experiences to fully assess the specific sensitivity of each international trade theory. More especially, we need to analyze the sensitivity of each model to iceberg transport cost values and the sensitivity of Krugman's model to the number of companies.

We also do not aim to only focus on sensitivity analysis *ceteris paribus* - like we did for the sensitivity analysis on sigma for instance - but want to test the sensitivity of the models on one parameter changing the values of others, like we aimed to do for the model's sensitivity to the number of companies.

Using different scales as input may also indirectly reveal in part the models' sensitivity toward change of scale.

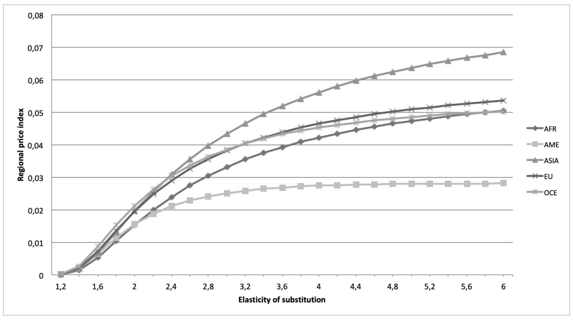


Figure 4.2a Armington's price index sensitivity to the elasticity of substitution.

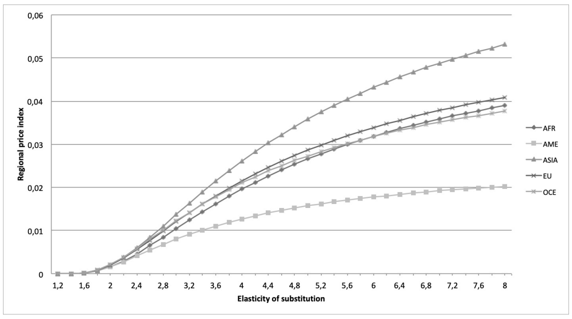


Figure 4.2b Krugman's price index sensitivity to the elasticity of substitution.

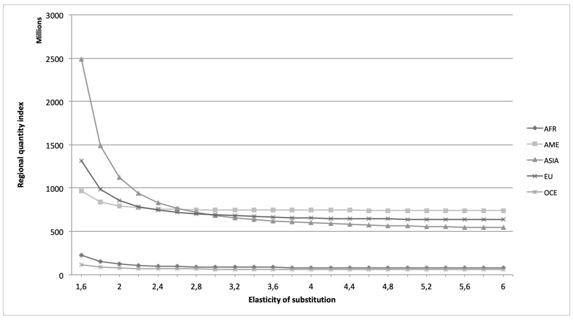


Figure 4.2c Armington's quantity index sensitivity to the elasticity of substitution.

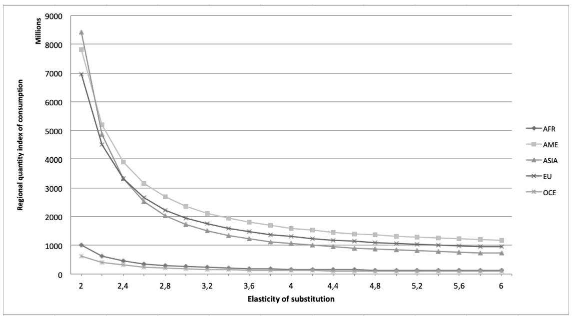


Figure 4.2d Krugman's quantity index sensitivity to the elasticity of substitution.

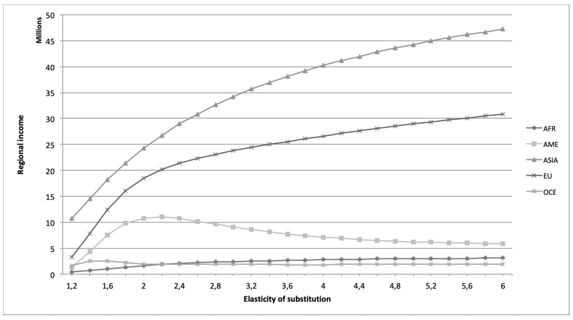


Figure 4.2e Armington's regional income sensitivity to the elasticity of substitution.

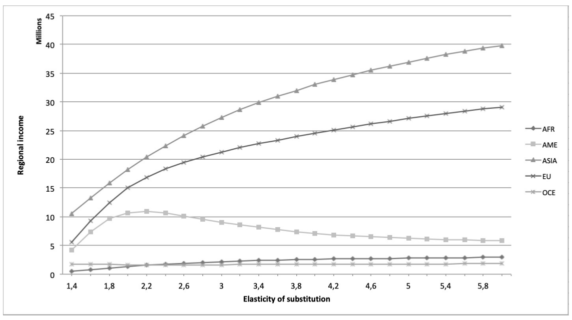


Figure 4.2f Krugman's regional income sensitivity to the elasticity of substitution.

Figure 4.2: Sensitivity analysis results on the elasticity of substitution: toward less sensitive models by shifting to Krugman's theory.

4.4.2 Theories' sensitivity toward change of spatial scale: a possible path to multi-scale models

Usually, a FSM is built to answer questions at a specific spatial scale, could it be global, macro-regional, national, regional, or even local. Therefore, changing the scale one model has been developed for may result in wicked results. For instance, a model that initially considers national and regional scales would not not be able to run at a global scale or describe a local economic phenomenon. Not being able to consider other spatial scales does not disqualify it to answer the question. But what this property tells us is that, in some ways, the spatial scale will highly contribute to shaping the validity domain of a model.

Nevertheless, we can think of a multi-scale model that can overcome this bias by generating consistent outputs for different spatial scales.

This implies that such a model would generate outputs consistent for every scale: outputs at a sub-scale that, when aggregated to an upper scale, would be close to those we obtain at the upper scale. For instance, this kind of model would equally use national data or continental data, and the outputs generated at the continental scale would be very close, after proper agglomeration, to the aggregated national scale outputs.

We are aware that modeling the forest sector at a particular scale captures, in a way, economic processes that belong to this scale. It may even be reasonably expected that economic processes occur differently depending on the spatial scale considered and that the full economic ecosystem results in a complex combinaison of each of them. Building a multi-scale model would then mean that, whatever the spatial scale consider, the algorithm could bridge the gap between the different scales and consider all of them in computing the equilibrium.

It would probably be unlikely to say that a change in the economic structure of present FSM would lead to a multi-scale model. However, in doing so, we can still expect to improve sensitivity to changes of spatial scales. We again suppose that Krugman's theory would bring a better resilience to changes of scales than Armington's theory, as the first observation of this Subsection has already suggested.

Using FAOSTAT database appears to be quite convenient for testing this assumption: continental and sub-continental scales are built on the national data collect of the FAO through a data agglomeration process. That means we can easily link one scale to another by simply aggregating the national scale data into sub-continental or continental scale data.

Each model would compute each set of data: continental, sub-continental and national data sets. For each scale of data used as input we collect the outputs. Then we aggregate the national scale outputs at the continental level and we do the same for the sub-continental scale outputs. The aggregation protocol is the same that the one used by the FAO: it consists in regrouping countries in the corresponding continent and subcontinents in the corresponding continent too.

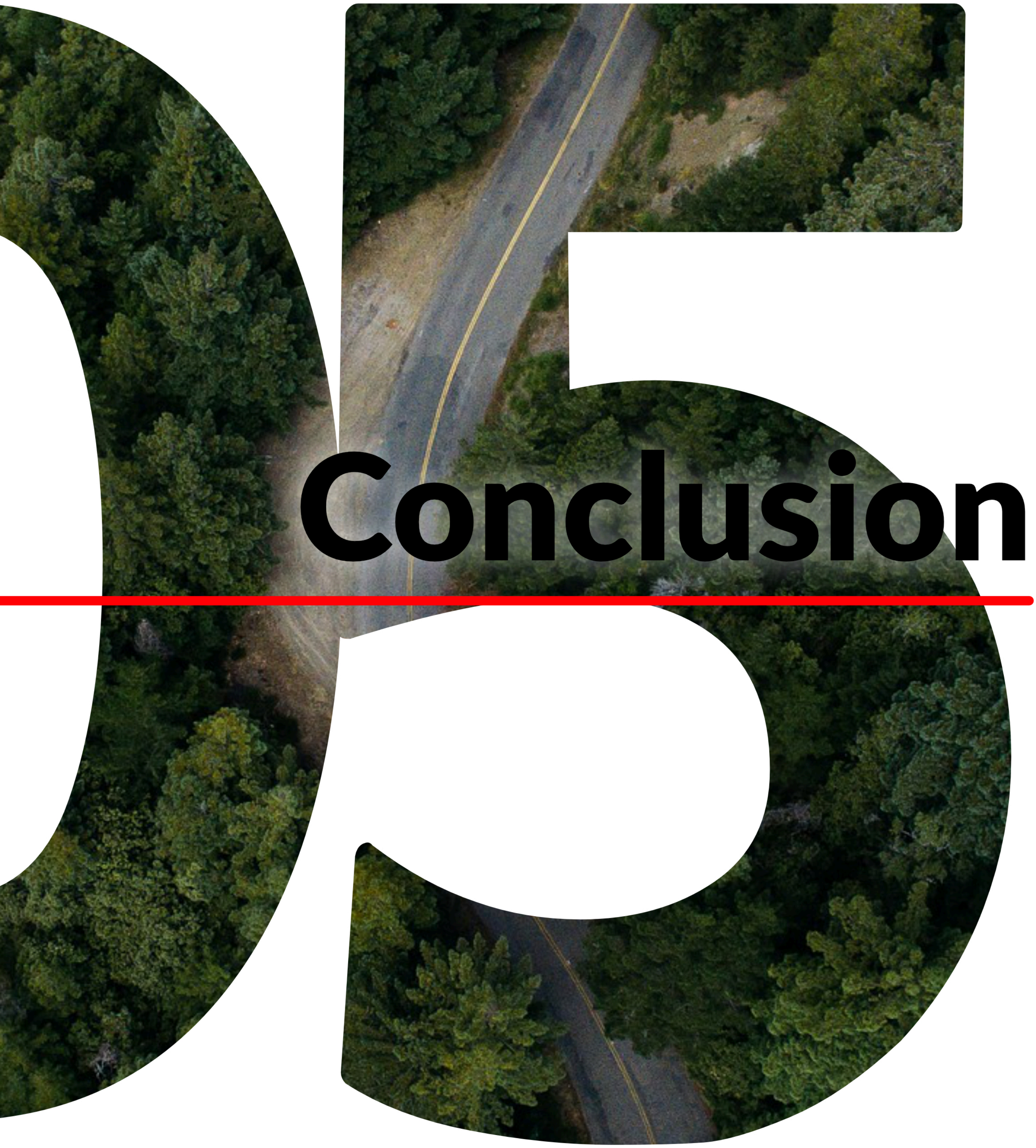
4.4.3 Simulating an agglomeration effect with our simplified Krugman's model: toward a dynamic representation of economic activity?

We also want to use equations (4.11) and (4.12) into Krugman's model to simulate agglomeration patterns. Using these two equations, it would be possible to update the number of companies per region after the economic equilibrium is reached. A recursive loop would then allow to represent an economic dynamic across regions and to test if economic agglomeration depicted by Krugman will lead to more *"realistic"* models.

4.4.4 Further improvements

Finally, we want to improve our work in three ways. First, we want to make the selling price of a given region endogenous in our models. Second, we want to develop a simple method to calibrate the number of companies per region based on real data. Third, we want to include Krugman's framework in FFSM to derive our analysis on a more developed model than the ones presented in this section.

05 Conclusion



5 Conclusion: some research leads to innovate in Forest Sector Modeling

Through the development of international trade theories and their comparison thanks to computational tools, we conceptually showed that changing the economic structure of FSM may substantially improve the way we model wood flows. But instead of settling for this result, our work intended to explain why we came to this conclusion by getting back to economic fundamentals.

This led us to the key notion of *"geographical space"* that attracted the founding fathers of economics but that has remained neglected by economics. Geographical perspective helped us link the qualitative improvements modern theories brought in representing trade flows with their better consideration of spatial features. Nevertheless, our study also shows that including space in the actual economic framework incurs an accumulation of conceptual and technical issues and still seems at its premises.

The reader of this report should keep in mind that what precedes remains highly conceptual and that other interesting findings may lie in a deeper theoretical analysis. Indeed, our work does not pretend to be exhaustive and economic theories we did not consider may bring more material and sharp our analysis. Going beyond the economic structure of FSM may also be fertile: FSM' structures include other modules that may play a role in representing *"geographical space"*, like land-use modules that certainly bring further spatial considerations.

Furthermore, modeling the forest sector does not, of course, rely on theory but also on the available data and its processing, and on the statistical methods used to match the theoretical structure with those data. This threefold motivation reminds us that we essentially focused our work on one side of the triangle: the theory. In this respect, that may in part explain why including modern and sophisticated economic theories in FSM actual framework face some sticking points. If it is difficult, yet manageable, to get a good proxy of the number of companies for each region considered, it seems intractable to differentiate them by their productivity like Melitz's theory does. This perspective invites us to explore the two other sides of the triangle we did not consider here.

At this stage, I remind what Joseph Buongiorno, a pioneer in Forest Sector Modeling, shared with me once. He thought we generally rely too much on economic theory while modeling the forest sector. Thus, it may be the time to change the economic paradigm on which we build our models. At the time of supercomputers that can solve intractable analytical problems, of advanced mathematics that base complexity theory, and advanced Geographic Information System (GIS), modeling spatial features might not be as unmanageable as economists may think.

Bibliography

Treb Allen and Costas Arkolakis. Elements of advanced international trade. *Graduate Trade Notes*, 2016.

Paul S Armington. A Theory of Demand for Products Distinguished by Place of Production. *International Monetary Fund Staff Papers*, 16(1):159–178, March 1969.

Edward J Balistreri and Thomas F Rutherford. Computing General Equilibrium Theories of Monopolistic Competition and Heterogeneous Firms. In Peter B Dixon and Dale W Jorgenson, editors, *Handbook of Computable General Equilibrium Modeling*, pages 1513–1570. Elsevier, 2013.

Sylvain Caurla. *Modélisation de la filière forêt-bois française - Évaluation des impacts des politiques climatiques*. PhD thesis, AgroParisTech, Laboratoire d'Économie Forestière, January 2012.

Sylvain Caurla. Une typologie et une histoire des modèles économiques de secteur forestier. *Revue Forestière Française*, 65:163–181, September 2013.

Neil M Coe, Henry W. C. Yeung, and Philip F Kelly. *Economic Geography : A Contemporary Introduction*. Blackwell Publishing, January 2007.

Jean-Yves Courtonne. *Environmental assessment of territories through supply chain analysis: biophysical accounting for deliberative decision-aiding*. PhD thesis, Université Grenoble Alpes, June 2016.

Masahisa Fujita, Paul Krugman, and Anthony J Venables. *The Spatial Economy: Cities, Regions, and International Trade*. The MIT Press, 1999.

Isabelle Géneau de Lamarlière and Jean-François Staszak. *Principes de géographie économique*. Editions Bréal, 2000.

Mario Larch and Yoto V Yotov. General equilibrium trade policy analysis with structural gravity. *CESifo Working Paper Series*, 2016/08:1–87, July 2016.

Gregory S Latta, Hanne K Sjølie, and Birger Solberg. A review of recent developments and applications of partial equilibrium models of the forest sector. *Journal of Forest Economics*, 19(4):350–360, December 2013.

Richard Levins. The Strategy of Model Building in Population Biology. *American Scientist*, 54(4):421–431, December 1966.

Marc J Melitz. The Impact of Trade on Intra-Industry Reallocations and Aggregate Industry Productivity . *Econometrica*, 71(6):1695–1725, November 2003.

Steven Northway, Gary Q Bull, and John D Nelson. Forest Sector Partial Equilibrium Models: Processing Components. *Forest Science*, 59(2):151–156, April 2013.

Miguel Riviere, Sylvain Caurla, et al. Integrating non-timber objectives into bio-economic models of the forest sector: a review of recent innovations and current shortcomings. Technical report, Bureau d'Economie Théorique et Appliquée, UDS, Strasbourg, 2018.

Paul A Samuelson. Spatial Price Equilibrium and Linear Programming. *The American Economic Review*, 42(3):283–303, June 1952.

Hanne K Sjølie, Greg S Latta, Erik Trømborg, Torjus F Bolkesjø, and Birger Solberg. An assessment of forest sector modeling approaches: conceptual differences and quantitative comparison. *Scandinavian Journal of Forest Research*, 30(1):60–72, January 2015.

Jacques-François Thisse and Bernard Walliser. Is space a neglected topic in mainstream economics? *Recherches Ecnomiques de Louvain*, 64(1):11–22, January 1998.

Appendices

Appendix A Statistical Analysis conducted by Boisseau and Courtonne (2018)

Figure A.1 shows one result of Boisseau and Courtonne's statistical work. We can see correlations between FFSM outputs en SitraM data are low, no matter what year is considered and no matter whether we consider tonnes or tonnes per kilometer traveled. This support the first observation made in the introduction by comparing FFSM simulations main wood flows in France with SitraM data main wood flows.

Figure A.1: Heatmap of correlations between FFSM simulations in SitraM data at different years, considering tonnes or tonnes per kilometers.

Appendix B Levins' considerations in modeling, a compromise between generality, precision and reality

We aim here to develop the modeling considerations of Richard Levins, also known as the *"Levins' triangle"* represented in Figure B.1 (Levins, 1966) and to apply them to Forest Sector Models. Even if Levins' work explicitly refers to the field of biology, his thoughts on modeling remain general and can apply to other fields, like economics.

Levins' thinking relies on a reasonable statement: might it be biological or economic, we study a complex system that is impossible to completely describe by a brute approach, for three reasons. First, because of missing data that may prevent us from estimating the several parameters we need. Second, because those complex systems lead to intractable analytical problems.[25] Third, the results would be as complex as the system to study and would have no meaning for us. Thus, it is then crucial "to simplify the models in a way that preserves the essential features of the problem."

Doing so requires simplifications that keep us from working with "manageable models which maximize generality, realism, and precision." Therefore, a modeler may consider different strategies:

1. Sacrifice generality to realism and precision. Those are the *"empirical models,"* that reduce the number of parameters to those relevant to a particular case and provide accurate predictions.

2. Sacrifice realism to generality and precision. Those are the *"analytical models"* which rely on general equations that are quite unrealistic but still bring precise results. Modelers using this kind of model expect that a small deviation from reality results in small deviation in the simulations.

3. Sacrifice precision to generality and realism. Those are more fundamental models we call *"mechanistic models."* They bring qualitative results and do not specify the mathematical form of an equation.

From Levins' typology and what we describes in the introduction, we suggest Forest Sector Models mostly belong to the *"analytical models"* category. Thus, they somehow fundamentally sacrifice realism to generality and precision. This calls for getting back

[25] This may not be as valid as when Levins wrote his ideas since computer science has progressed and can now solve problems considered intractable at that time.

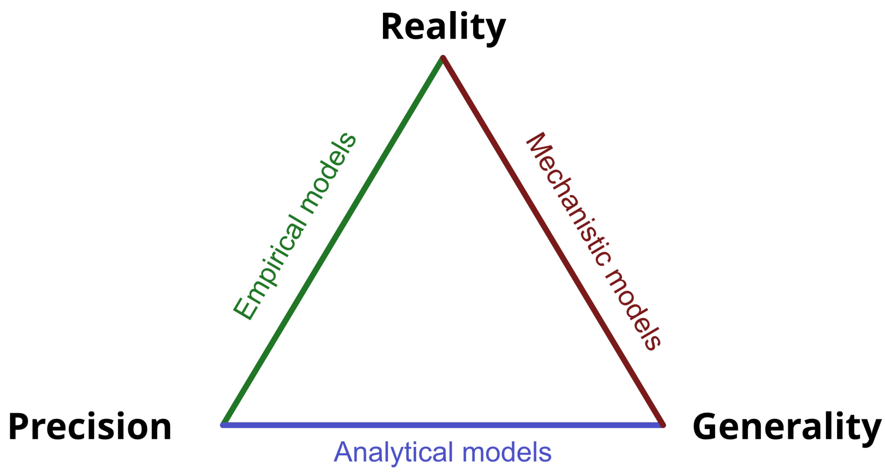


Figure B.1: The Levins' triangle, an illustration of the compromise a modeler has to do between reality, precision, and generality in modeling.

to modeling fundamentals in economics and to explore if FSM' are able to represent trade flows and their dynamics.

Appendix C A development of Armington's theory based on *A Theory of Demand for Products Distinguished by Place of Production* (1969)

We adapt Armington's main equations from Larch and Yotov (2016), and Allen and Arkolakis (2016).

In what follows, the indexes r, s and t represent regions of the world. The index j, refers to a given product.

C.1 Reminder: conceptual framework and main assumptions

We assume that products are differentiated not only by their kind but also by their place of production. Two regions of production thus produce two varieties of the same kind of product that are imperfect substitutes in demand.

Each place of production is considered as the supplier of one variety but also as a source of demand for all other varieties. The preferences of a region for a variety of one given kind are assumed to be independent of its purchases of varieties of another kind. Each demand for a particular product thus represents an isolated market.

Given this assumptions of independence, Armington's theory defines a quantity index for the consumption of the product j, in region s, noted $M_{j,s}$, as a function of all the quantities of a product j, produced in region r, consumed by the region s, noted $q_{j,r,s}$:

$$M_{j,s} = \phi_j \ (q_{j,r,s}) \tag{C.1}$$

It also defines a price index of a product j, in a consumption region s, noted $G_{j,s}$, as a function of the delivered prices associated with a product j, shipped from its production region r, to the consumption region s, noted $p_{j,r,s}$.

Armington adds the condition that the demand for $M_{j,s}$ "must be the same as the value of ϕ_j implied by all demands for products in the j^{th} market". This conditions implies that:

$$G_{j,s} = p_{j,r,s} \div \frac{\partial \phi_j}{\partial q_{j,r,s}} \tag{C.2}$$

To ensure that $G_{j,s}$ respects the condition of independence with regard to $M_{j,s}$, we assume that ϕ_j is a linear and homogeneous function.

C.2 Armington's analytical development assuming a constant elasticity of substitution in each market

The theory uses a constant elasticity of substitution (CES) demand function to express the Armington's quantity index analytical expression for a given product j, consumed in region s:

$$M_{j,s} = \phi_j \ (q_{j,r,s}) = \left[\sum_r (q_{j,r,s})^{\frac{\sigma-1}{\sigma}} \right]^{\frac{\sigma}{\sigma-1}} \tag{C.3}$$

where σ the constant elasticity of substitution.

We then can derivate this expression by the quantity shipped from a given production region:

$$\frac{\partial M_{j,s}}{\partial q_{j,r,s}} = \frac{\partial}{\partial q_{j,r,s}} \left[\left(\sum_r (q_{j,r,s})^{\frac{\sigma-1}{\sigma}} \right)^{\frac{\sigma}{\sigma-1}} \right] = \frac{\sigma}{\sigma-1} \times \frac{\sigma-1}{\sigma} \times (q_{j,r,s})^{\frac{-1}{\sigma}} \times \left(\sum_r (q_{j,r,s})^{\frac{\sigma-1}{\sigma}} \right)^{\frac{1}{\sigma-1}} \tag{C.4}$$

that leads to the following expression:

$$\frac{\partial M_{j,s}}{\partial q_{j,r,s}} = \left(\frac{M_{j,s}}{q_{j,r,s}} \right)^{\frac{1}{\sigma}} \tag{C.5}$$

We can then write:

$$\frac{\partial M_{j,s}/\partial q_{j,r,s}}{\partial M_{j,s}/\partial q_{j,t,s}} = \left(\frac{q_{j,t,s}}{q_{j,r,s}} \right)^{\frac{1}{\sigma}} = \frac{p_{j,r,s}}{p_{j,t,s}} \tag{C.6}$$

This yields to the following relation between shipped quantities and associated prices:

$$q_{j,t,s} = q_{j,r,s} \times \left(\frac{p_{k,r,s}}{p_{k,t,s}} \right)^{\sigma} \tag{C.7}$$

We then use equation (C.7) in equation (C.3) and comes to the following result:

$$M_{j,s} = q_{j,r,s} \times \left[\sum_t \left(\frac{p_{j,r,s}}{p_{j,t,s}} \right)^{\sigma-1} \right]^{\frac{\sigma}{\sigma-1}} \tag{C.8}$$

that can also be expressed as:

$$q_{j,r,s} = M_{j,s} \times \left[\sum_t \left(\frac{p_{j,r,s}}{p_{j,t,s}} \right)^{\sigma-1} \right]^{\frac{\sigma}{1-\sigma}} \tag{C.9}$$

Combining equation (C.5), equation (C.1) and equation (3.3), we then can express the price index as:

$$G_{j,s} = p_{j,r,s} \times (q_{j,r,s})^{\frac{1}{\sigma}} \times (M_{j,s})^{-\frac{1}{\sigma}} \tag{C.10}$$

Using the formulation of $M_{j,s}$ given by equation (C.8), we then derive the Armington's price index of a consumption region s, for a product j, as follows:

$$G_{j,s} = p_{j,r,s} \times (q_{j,r,s})^{\frac{1}{\sigma}} \times q_{j,r,s}^{-\frac{1}{\sigma}} \times \left[\sum_t \left(\frac{p_{j,r,s}}{p_{j,t,s}} \right)^{\sigma-1} \right]^{\frac{1}{1-\sigma}} \tag{C.11}$$

$$\Rightarrow G_{j,s} = p_{j,r,s} \times \left[\sum_t \left(\frac{p_{j,r,s}}{p_{j,t,s}} \right)^{\sigma-1} \right]^{\frac{1}{1-\sigma}} \tag{C.12}$$

$$\Rightarrow \quad G_{j,s} = \left[\sum_r (p_{j,r,s})^{(1-\sigma)} \right]^{\frac{1}{1-\sigma}} \tag{C.13}$$

We then turn around equation (C.12):

$$\left(\frac{p_{j,r,s}}{G_{j,s}} \right)^{\sigma} = \left[\sum_t \left(\frac{p_{j,r,s}}{p_{j,t,s}} \right)^{\sigma-1} \right]^{\frac{\sigma}{\sigma-1}} \tag{C.14}$$

Combining equation (C.14) with equation (C.9) thus gives us the demand of a consumption region s, for a product j, produced in region r:

$$q_{j,r,s} = \left(\frac{p_{j,r,s}}{G_{j,s}}\right)^{-\sigma} \times M_{j,s} \tag{C.15}$$

Note that in all these equations, regions r, s and t, can be a country, a group of countries and so forth, as Armington suggests in his paper.

Appendix D A development of Krugman's theory based on *The Spatial Economy: Cities, Regions and International Trade* (1999)

In what follows, the indexes r, and s represent regions of the world. Indexes i, j and k, refer to a given product.

D.1 Reminder: conceptual framework and main assumptions

Krugman's theory intends to develop a spatial framework assuming monopolistic competition and increasing returns of scale at the firm level. To do so, Krugman's approach spatializes the Dixit-Stiglitz model of monopolistic competition that provides its basis by adding locations of production and consumption and by including Samuelson's iceberg transport costs between them.

The approach considers two sectors:

1. An agricultural sector under perfect competition that produces a unique and homogeneous product that can be shipped without incurring any transport costs.
2. A manufacturing sector under monopolistic competition that provides many varieties of manufactured products and that includes increasing return of scale and iceberg transport costs.

Each consumer purchases agricultural and manufactured goods.

All manufacturing firms use the same technology and then have the same kind of production function. Moreover, the quantity they produce and the associated price are the same between all firms that are in the same region.

D.2 The Dixit-Stiglitz model of monopolistic competition

We first consider economic activity without referring to any location. We also assume that there is a large number of manufactured products so that we can represent the products range as continuous.

The utility function of a consumer is given by the following Cobb-Douglas function:

$$U = M^{\mu} \times A^{1-\mu} \tag{D.1}$$

where M is the composite index of consumption of manufactured products; A is the composite index of consumption of the agricultural good; and μ is the constant expenditure share over manufactured products.

Referring to the Dixit-Stiglitz framework, we define the quantity index of manufactured products over a continuous range of varieties by the following constant elasticity of substitution (CES) function:[26]

$$M = \left[\int_0^n m(i)^{\frac{\sigma-1}{\sigma}} \, \mathrm{d}i \right]^{\frac{\sigma}{\sigma-1}} \tag{D.2}$$

where $m(i)$ is the consumption of one variety of manufactured product; n is the number of available varieties; and σ is the constant elasticity of substitution.

The budget constraint is given by the following formulation of the income, Y:

$$Y = A \times p^A + \int_0^n p(i) \times m(i) \, \mathrm{d}i \tag{D.3}$$

where p^A and $p(i)$ are respectively the price of the agricultural good and the price of a given variety of manufactured products i.

To maximize its utility subject to budget constraint, the consumer first tries to minimize the cost of attaining M:

$$min \quad \int_0^n p(i) \times m(i) \, \mathrm{d}i \quad s.t. \quad M = \left[\int_0^n m(i)^{\frac{\sigma-1}{\sigma}} \, \mathrm{d} \right]^{\frac{\sigma}{\sigma-1}} \tag{D.4}$$

To solve this minimisation problem, we use the Lagrangian method:

$$L = \int_0^n p(i) \times m(i) \, \mathrm{d}i - \lambda \times \left[\int_0^n m(i)^{\frac{\sigma-1}{\sigma}} \, \mathrm{d}i - M^{\frac{\sigma-1}{\sigma}} \right] \tag{D.5}$$

$$\Rightarrow \quad \frac{\partial L}{\partial m(i)} = p(i) - \lambda \times \frac{\sigma - 1}{\sigma} \times m(i)^{-\frac{1}{\sigma}} = 0 \tag{D.6}$$

$$\Rightarrow \quad m(i) = \left(\frac{p(i) \times \sigma}{\lambda \times (\sigma - 1)} \right)^{-\sigma} \quad and \quad m(j) = \left(\frac{p(j) \times \sigma}{\lambda \times (\sigma - 1)} \right)^{-\sigma} \tag{D.7}$$

[26] We slightly adapt the formulation from the one Krugman provides in his initial work.

$$\Rightarrow \quad \frac{m(i)}{m(j)} = \left(\frac{p(i)}{p(j)}\right)^{-\sigma} \tag{D.8}$$

$$\Rightarrow \quad m(i) = m(j) \times \left(\frac{p(j)}{p(i)}\right)^{\sigma} \tag{D.9}$$

Including this previous expression in the initial constraint we obtain:

$$M^{\frac{\sigma-1}{\sigma}} = \int_0^n m(i)^{\frac{\sigma-1}{\sigma}}\,\mathrm{d}i = \int_0^n m(j)^{\frac{\sigma-1}{\sigma}} \times \left(\frac{p(j)}{p(i)}\right)^{\sigma-1}\mathrm{d}i = m(j)^{\frac{\sigma-1}{\sigma}} \times p(j)^{\sigma-1} \times \int_0^n p(i)^{\frac{1}{\sigma-1}}\,\mathrm{d}i \tag{D.10}$$

$$\Rightarrow \quad m(j) = \frac{p(j)^{-\sigma}}{\left[\int_0^n p(i)^{1-\sigma}\,\mathrm{d}i\right]^{\frac{\sigma}{\sigma-1}}} \times M \tag{D.11}$$

We then have:

$$m(j) \times p(j) = p(j)^{1-\sigma} \times \frac{1}{\left[\int_0^n p(i)^{1-\sigma}\,\mathrm{d}i\right]^{\frac{\sigma}{\sigma-1}}} \times M \tag{D.12}$$

$$\Rightarrow \quad \int_0^n m(j) \times p(j)\,\mathrm{d}j = \int_0^n p(j)^{1-\sigma}\,\mathrm{d}j \times \frac{1}{\left[\int_0^n p(i)^{1-\sigma}\,\mathrm{d}i\right]^{\frac{\sigma}{\sigma-1}}} \times M \tag{D.13}$$

noting that: $\int_0^n p(j)^{1-\sigma}\,\mathrm{d}j = \int_0^n p(i)^{1-\sigma}\,\mathrm{d}i$, we then derive the following expression:

$$\int_0^n m(j) \times p(j)\,\mathrm{d}j = \left[\int_0^n p(j)^{1-\sigma}\,\mathrm{d}j\right]^{\frac{1}{1-\sigma}} \times M \tag{D.14}$$

The first order condition finally leads to:

$$m(i) = \left[\frac{p(i)}{G}\right]^{-\sigma} \times M \tag{D.15}$$

The maximization problem of the consumer then consists in choosing M and A so that:

$$max \quad U = M^{\mu} \times A^{1-\mu} \qquad s.t. \quad Y = G \times M + p^{A} \times A \tag{D.16}$$

We again use the Lagrangian method:

$$L = M^{\mu} \times A^{1-\mu} - \lambda \times (G \times M + p^{A} \times A - Y) \tag{D.17}$$

$$\Rightarrow \frac{\partial L}{\partial M} = \mu \times M^{\mu-1} \times A^{1-\mu} - \lambda \times G = 0 \quad \Rightarrow \quad \lambda = \frac{\mu \times M^{\mu-1} \times A^{1-\mu}}{G} \tag{D.18}$$

$$\Rightarrow \frac{\partial L}{\partial A} = (1-\mu) \times M^{\mu} \times A^{-\mu} - \lambda \times p^{A} = 0 \tag{D.19}$$

which yields to the following consumer demand functions:

For agriculture:

$$A = (1-\mu) \times \frac{Y}{p^{A}} \tag{D.20}$$

For manufactures:

$$M = \mu \times \frac{Y}{G} \tag{D.21}$$

which implies that:

$$m(i) = \mu \times Y \times p(i)^{-\sigma} \times G^{\sigma-1} \tag{D.22}$$

D.3 Considering multiple locations and iceberg transport costs

D.3.1 Consumer behaviour

We now apply the previous results of the Dixit-Stiglitz model of monopolistic competition to a finite set of locations and add Samuelson's iceberg transport costs.

Iceberg transport costs imply that when a unit of a product k, is shipped from a production region r, to a consumption region s, only a fraction of the original unit actually arrives: the rest melts on the way. We note $\tau_{k,r,s}$ the amount the product k, dispatched per unit received: it represents the iceberg transport costs.

Under this assumption, the delivered price in a region s, of a product k, produced in region r, noted $p_{k,r,s}$, is given by:

$$p_{k,r,s} = p_{k,r} \times \tau_{k,r,s} \tag{D.23}$$

where $p_{k,r}$ is the selling price of a product k, in region r and $\tau_{k,r,s}$ is the iceberg transport cost incurred while shipping a product k, from a production region r, to a consumption region s.

Under those considerations, Krugman's price index takes the following form:

$$G_{k,s} = \left[\sum_r N_{k,r} \times (p_{k,r} \times \tau_{k,r,s})^{1-\sigma} \right]^{\frac{1}{1-\sigma}} \tag{D.24}$$

where $G_{k,s}$ is the Krugman's price index of a product k, consumed in region s; $N_{k,r}$ is the number of companies/number of varieties of a product k, produced in region r; and σ is the constant elasticity of substitution between any two pairs of varieties.

Under the Dixit-Stiglitz framework, the maximization of the consumer utility implies that the consumption demand in region s, for a product k, produced in region r, noted $q_{k,r,s}$, can be written as:

$$q_{k,r,s} = \mu \times Y_{k,s} \times (p_{k,r} \times \tau_{k,r,s})^{-\sigma} \times G_{k,s}^{(\sigma-1)} \tag{D.25}$$

where $Y_{k,s}$ is the income of region s generated by the sales of a product k.

To supply this level of demand for one region s, a firm in region r, has to ship $\tau_{k,r,s}$ times the amount $q_{k,r,s}$ because of iceberg transport costs. Summing across all consumption locations, we derive the total production of a single firm producing a variety in region r, noted $q_{k,r}^{tot}$, as follows:

$$q_{k,r}^{tot} = \mu \times \sum_s Y_{k,s} \times (p_{k,r} \times \tau_{k,r,s})^{-\sigma} \times G_{k,s}^{\sigma-1} \times \tau_{k,r,s} \tag{D.26}$$

D.3.2 Producer behaviour

All firms in all locations having the same technology, they involve the same fixed input F and marginal input requirement c. We assume the only input is labor.

We can then express the labor input in region r, noted $l_{k,r}$, required to produce a quantity $q_{k,r}^{tot}$ of a product k, as:

$$l_{k,r} = F + c \times q_{k,r}^{tot} \tag{D.27}$$

As labor is the only input for production, the profit of a firm generated by the sales of a product k, produced in region r, noted $\pi_{k,r}$, is given by:

$$\pi_{k,r} = p_{k,r} \times q_{k,r,s}^{tot} - w_{k,r} \times (\ F + c \times q_{k,r}^{tot}\ ) \tag{D.28}$$

where $w_{k,r}$ is the wage rate for manufacturing workers that produce a product k, in region r.

Profit maximization implies that:

$$\frac{\partial \pi_{k,r}}{p_{k,r}} = q_{k,r,s}^{tot} + \frac{\partial q_{k,r,s}^{tot}}{\partial p_{k,r}} \times p_{k,r} - \frac{\partial q_{k,r,s}^{tot}}{\partial p_{k,r}} \times w_{k,r} \times c = 0 \tag{D.29}$$

$$\Rightarrow \quad q_{k,r,s}^{tot} + \frac{\partial q_{k,r,s}^{tot}}{\partial p_{k,r}} \times (p_{k,r} - w_{k,r} \times c) = 0 \tag{D.30}$$

noting that:

$$\frac{\partial q_{k,r,s}^{tot}}{\partial p_{k,r}} = \frac{\partial}{\partial p_{k,r}} \left[\mu \times \sum_s Y_{k,s} \times (p_{k,r})^{-\sigma} \times (\tau_{k,r,s})^{-\sigma} \times G_{k,s}^{\ \sigma-1} \times \tau_{k,r,s} \right] = -\sigma \times \frac{1}{p_{k,r}} \times q_{k,r,s}^{tot} \tag{D.31}$$

we easily derives the following formulation of the price $p_{k,r}$:

$$p_{k,r} \times \left(1 - \frac{1}{\sigma}\right) = c \times w_{k,r} \tag{D.32}$$

Giving this pricing rule, the profits now has the following formulation:

$$\pi_{k,r} = w_{k,r} \times \left[\frac{q_{k,r,s}^{tot} \times c}{\sigma - 1} - F \right] \tag{D.33}$$

We assume free entry and exit of firms. That leads to a zero-profit condition that provides the following equilibrium results:

the equilibrium output of every firm, noted $q_{k,r}^{*}$:

$$q_{k,r}^{*} = F \times \frac{\sigma - 1}{c} \tag{D.34}$$

and the equilibrium labor input, noted $l_{k,r}^{*}$:

$$l^*_{k,r} = F + c \times q^*_{k,r} = F \times \sigma \tag{D.35}$$

Therefore, if $L_{k,r}$ is the total number of manufacturing workers that produce a product k, in region r, we can then define the number of firms producing a product k, in region r as:

$$N_{k,r} = \frac{L_{k,r}}{l^*_{k,r}} = \frac{L_{k,r}}{F \times \sigma} \tag{D.36}$$

D.3.3 An equation of the nominal wages

Thus, the zero-profit condition implies that firms produce $q^*_{k,r}$. We note that the firm equilibrium output, $q^*_{k,r}$, also follows equation (D.26):

$$q^*_{k,r} = \mu \times \sum_s Y_{k,s} \times (p_{k,r} \times \tau_{k,r,s})^{-\sigma} \times G_{k,s}^{\sigma-1} \times \tau_{k,r,s} \tag{D.37}$$

We can then turn around this equation to express the firm price $p_{k,r}$ as follows:

$$(p_{k,r})^{\sigma} = \frac{\mu}{q^*_{k,r}} \times \sum_s Y_{k,s} \times (\tau_{k,r,s})^{1-\sigma} \times G_{k,s}^{(\sigma-1)} \tag{D.38}$$

Using equation (D.32), we directly derive the wage equation:

$$w_{k,r} = \left(\frac{\sigma-1}{\sigma \times c}\right) \left[\frac{\mu}{q^*_{k,r}} \times \sum_s Y_{k,s} \times (\tau_{k,r,s})^{1-\sigma} \times G_{k,s}^{(\sigma-1)}\right]^{\frac{1}{\sigma}} \tag{D.39}$$

We now define the real wages of workers producing a product k, in region r, noted $\omega_{k,r}$, as the nominal wages deflated by the cost-of-living index, expressed as $G_{k,r}^{\mu} \times (p_r^A)^{(1-\mu)}$. We obtain the following formulation:

$$\omega_{k,r} = w_{k,r} \times G_{k,r}^{\mu} \times (p_r^A)^{(1-\mu)} \tag{D.40}$$

D.4 A simplification thanks to some normalizations

We choose units such as: $c = \frac{\sigma-1}{\sigma}$ and $F = \frac{\mu}{\sigma}$.

It then follows that:

From equation (D.32): $p_{k,r} = w_{k,r}$

From equation (D.34) and (D.35): $q^*_{k,r} = l^*_{k,r} = \mu$

From equation (D.36): $N_{k,r} = \frac{L_{k,r}}{\mu}$

Using those normalizations, the price index and wage equations now have the following expression:

$$G_{k,r} = \left[\frac{1}{\mu} \times \sum_{s} L_{k,s} \times (w_{k,s} \times \tau_{k,r,s})^{1-\sigma} \right]^{\frac{1}{1-\sigma}} \tag{D.41}$$

$$w_{k,r} = \left[\sum_{s} Y_{k,s} \times (\tau_{k,r,s})^{1-\sigma} \times G_{k,s}^{\sigma-1} \right]^{\frac{1}{\sigma}} \tag{D.42}$$

D.5 The core and periphery model

To derive Krugman's core-periphery model, we now assume that the world has $L_k^M = \mu$ manufacturing workers producing a product k, and $L^A = 1 - \mu$ farmers. We denote one region r, share of farmers world labor force by ϕ_r, and its share of workers world labor force producing a product k, by $\lambda_{k,r}$. We also assume that farmers wages are all equal to unit.

Under those assumptions, the Krugman's equilibrium is then defined as the simultaneous solution of the four following equations of:

The income

$$Y_{k,r} = \mu \times \lambda_{k,r} \times w_{k,r} + (1 - \mu) \times \phi_r \tag{D.43}$$

The price index

$$G_{k,r} = \left[\sum_{s} \lambda_{k,s} \times (w_{k,s} \times \tau_{k,r,s})^{1-\sigma} \right]^{\frac{1}{1-\sigma}} \tag{D.44}$$

The nominal wages

$$w_{k,r} = \left[\sum_{s} Y_{k,s} \times (\tau_{k,r,s})^{1-\sigma} \times G_{k,s}^{\sigma-1} \right]^{\frac{1}{\sigma}} \tag{D.45}$$

The real wages

$$\omega_{k,r} = w_{k,r} \times G_{k,r}^{-\mu} \tag{D.46}$$

The model can then describes agglomeration by assuming that workers are moving to a region where the real wage is higher. The dynamics can be expressed by the following equation:

$$\dot{\lambda}_{k,r} = \gamma \times (\omega_{k,r} - \overline{\omega}) \times \lambda_{k,r} \tag{D.47}$$

where γ is a constant and $\overline{\omega}$ the average real wage generated by one product k, defined as follows:

$$\overline{\omega}_k = \sum_r \lambda_{k,r} \times \omega_{k,r} \tag{D.48}$$

Appendix E A development of the income equation

We can define the income as (1) the sum of the revenues in a region - that is to say, the sum of the sales to all other region times the selling price - or (2) the total expenditure - that is to say, the sum of the amount purchased in every region times the purchase price. Following the zero-profit assumption, these two formulations are strictly equivalent: a region should spend all its revenues in purchasing other products, and one region spends the exact amount it earns.

In Subsection 4.1 we choose to formulate the income of a region s, as its money expenditure, similarly to Armington (Armington, 1969), for Armington's and Krugman's theories. That leads to the following expression of the income:

For Armington: $Y_{j,s} = \sum_r q_{j,r,s} \times (p_{j,r} \times \tau_{j,r,s})$ (4.4)

For Krugman: $Y_{k,s} = \sum_r (q^f_{k,r,s} \times N_{k,r}) \times (p_{k,r} \times \tau_{k,r,s})$ (4.8)

where $Y_{\cdot,s}$ is the regional income; $q_{j,r,s}$ and $q^f_{k,r,s}$ the quantity of a good j or k, shipped from region r, to region s, respectively for Armington's theory and Krugman's approach; $N_{k,r}$ the regional number of companies in Krugman's framework; $p_{\cdot,r}$ the regional price of a good produced in region r; and $tau_{\cdot,r,s}$ the iceberg transport costs of shipping a product j or k, from region r, to region s.

We may think of another formulation, expressing $Y_{\cdot,s}$ as the product of the quantity index and the price index of the region, s. This would be analytically correct and equivalent to the formulation we choose. Indeed, remember that:

$$q_{j,r,s} = \left(\frac{p_{j,r,s}}{G_{j,s}}\right)^{-\sigma} \times M_{j,s} \tag{E.1}$$

and that:

$$G_{j,s} = \left[\sum_r (p_{j,r,s})^{(1-\sigma)}\right]^{\frac{1}{1-\sigma}} \tag{E.2}$$

where $M_{j,s}$ is the quantity index of a product j, in region s and $G_{j,s}$ is the price index of a product j, in region s.

Using equations (E.1) and (E.2) in equation (4.4) we have:

$$Y_{j,s} = \sum_r q_{j,r,s} \times (p_{j,r} \times \tau_{j,r,s})$$
$$= \sum_r (p_{j,r} \times \tau_{j,r,s})^{1-\sigma} \times (G_{j,s})^{\sigma} \times M_{j,s}$$
$$= (G_{j,s})^{\sigma} \times M_{j,s} \times \sum_r (p_{j,r} \times \tau_{j,r,s})^{1-\sigma} \quad \text{(E.3)}$$

According to equation (E.2), we notice that $\sum_r (p_{j,r} \times \tau_{j,r,s})^{1-\sigma} = (G_{j,s})^{1-\sigma}$. We then have the following equation:

$$Y_{j,s} = \sum_r q_{j,r,s} \times (p_{j,r} \times \tau_{j,r,s}) = M_{j,s} \times G_{j,s} \quad \text{(E.4)}$$

A quite similar demonstration lead to the same result with Krugman's framework. We can then express the income as the product of the two indexes. However, equations (4.4) and (4.8) seem more convenient to visualize the change of structure between the two sets of equations and the place of iceberg transport costs.

We can also consider the income as the total revenues of one region (the "supply-side definition" of the income). In the case of Krugman's theory, that leads to the following expression:

$$Y_{k,s} = \sum_r (q^{f}_{k,s,r} \times N_{k,s}) \times (p_{k,s} \times \tau_{k,s,r}) \quad \text{(E.5)}$$

Now, remember from Subsection 3.3 that:

$$q^{tot}_{k,s} = \sum_r q^{f}_{k,s,r} \times \tau_{k,s,r}) \quad \text{(E.6)}$$

Using this previous equation and the zero-profit condition at the firm level[27], we have:

[27] This condition implies that the selling price of a firm is equal to its production cost which is, in Krugman's framework, the cost of labor. Appendix D provides a decomposition and development of this cost (see "Producer behavior").

$$
\begin{aligned}
Y_{k,s} = \sum_r (q^f_{k,s,r} \times N_{k,s}) \times (p_{k,s} \times \tau_{k,s,r}) \\
= N_{k,s} \times p_{k,s} \times \sum_r q^f_{k,s,r} \times \tau_{k,s,r}) \\
= N_{k,s} \times p_{k,s} \times q^{tot}_{k,s} \\
= N_{k,s} \times (F + c \times q^{tot}_{k,s}) \times w_{k,s} \quad \text{(E.7)}
\end{aligned}
$$

By extending the zero-profit condition to the regional level, income definitions on the supply side and on the demand size are analytically equivalent.

Appendix F GAMS codes of Armington's and Krugman's theories

F.1 GAMS code for Armington's model

```
$Title Armington Trade Equilibrium from A Theory of Demand for Products
Distinguished by Place of Production

*Valentin MATHIEU, AgroParisTech / INRA (valentin.mathieu@inra.fr).
*August 2019

*Copy past the good list of regions during the Setup

*List of continents
*AFR,AME,ASIA,EU,OCE

*List of sub-continents
*AUSNZ,CARI,CAMER,CASIA,EAFRI,EASIA,EEURO,MELA,MAFRI,
NAFRI,NAMER,NEURO,POLY,SAMER,SEASIA,SAFRI,SASIA,SEURO,WAFRI,WASIA,WEURO

Set
r          countries or regions     /AFR,AME,ASIA,EU,OCE/
j          goods                    /IR/;
Alias (r,s);

Parameters
        sig          elasticity of substitution  /3/,

        v0(j,r,s)    benchmark value of quantities of a product j
        shipped from region r to region s,

        qv0(j,r,s)   benchmark quantities of a product j
        shipped from region r to region s without autoconsumption,

        p0(j,r,s)    benchmark delivered price of a product j
        shipped from region r to region s,

        q0(j,r,s)    benchmark quantities of a product j
        shipped from region r to region s with autoconsumption
```

```
        G0(j,r)       benchmark price index in region r,

        M0(j,r)       benchmark quantity composite index of the
        consumption of a product j in region r,

        Y0(j,r)       benchmark total spending of region r
        or regional income,


        tau(j,r,s)    iceberg transport cost factor
        ;

*===Import Data from Spreadsheet===*

*Build gdxxrw instruction file
$onecho > importQ.txt

*Choose the type of Dataset you want to import
*Pay attention to the cells imported if you modify the xlsx file

*For Continents
*IR
par=q0 rng=Matrix_Q_IR!A10:G15 Rdim=2 Cdim=1
set=r rng=Matrix_Q_IR!B11:B15 Rdim=1 Cdim=0
set=s rng=Matrix_Q_IR!C10:G10 Rdim=0 Cdim=1

*For Sub-Continents
*IR
*par=vx0 rng=Matrix_Q_IR!A26:W47 Rdim=2 Cdim=1
*set=r rng=Matrix_Q_IR!B27:B47 Rdim=1 Cdim=0
*set=s rng=Matrix_Q_IR!C26:W26 Rdim=0 Cdim=1

$offecho

$onecho > importV.txt

*Choose the type of Dataset you want to import
*Pay attention to the cells imported if you modify the xlsx file

*For Continents
```

```
*IR
par=v0 rng=Matrix_V_IR!A10:G15 Rdim=2 Cdim=1
set=r rng=Matrix_V_IR!B11:B15 Rdim=1 Cdim=0
set=s rng=Matrix_V_IR!C10:G10 Rdim=0 Cdim=1

*For Sub-Continents
*IR
*par=v0 rng=Matrix_V_IR!A26:W47 Rdim=2 Cdim=1
*set=r rng=Matrix_V_IR!B27:B47 Rdim=1 Cdim=0
*set=s rng=Matrix_V_IR!C26:W26 Rdim=0 Cdim=1

$offecho

$onecho > importQV.txt

*Choose the type of Dataset you want to import
*Pay attention to the cells imported if you modify the xlsx file

*For Continents
*IR
par=qv0 rng=Matrix_Q_IR!A1:G6 Rdim=2 Cdim=1
set=r rng=Matrix_Q_IR!B2:B6 Rdim=1 Cdim=0
set=s rng=Matrix_Q_IR!C1:G1 Rdim=0 Cdim=1

*For Sub-Continents
*IR
*par=qv0 rng=Matrix_Q_IR!A1:W47 Rdim=2 Cdim=1
*set=r rng=Matrix_Q_IR!B2:B22 Rdim=1 Cdim=0
*set=s rng=Matrix_Q_IR!C1:W1 Rdim=0 Cdim=1

$offecho

*Choose the good way to import the file : in GAMSStudio files ,
just have to change the file 's name

$call gdxxrw.exe C:\Users\vlmathieu\Documents\GAMSStudio\
Armington_Krugman_Models\Data\Continents.xlsx @importQ.txt
$gdxin Continents.gdx

$load q0
$gdxin
```

```
$call gdxxrw.exe C:\Users\vlmathieu\Documents\GAMSStudio\
Armington_Krugman_Models\Data\Continents.xlsx @importV.txt
$gdxin Continents.gdx

$load v0
$gdxin

$call gdxxrw.exe C:\Users\vlmathieu\Documents\GAMSStudio\
Armington_Krugman_Models\Data\Continents.xlsx @importQV.txt
$gdxin Continents.gdx

$load qv0
$gdxin

p0(j,r,s) = v0(j,r,s)/qv0(j,r,s);
tau(j,r,s) = p0(j,r,s) / p0(j,r,r);

*Sensitivity analysis on tau
tau(j,r,s) = 1*(p0(j,r,s) / p0(j,r,r));
tau(j,r,r) = 1;

G0(j,s) = sum(r, (p0(j,r,r)*tau(j,r,s))**(1-sig))**(1/(1-sig));
M0(j,s) = sum(r, (q0(j,r,s)/tau(j,r,s))**((sig-1)/sig))**(sig/(sig-1));
Y0(j,s) = sum(r, q0(j,r,s)*(p0(j,r,r)*tau(j,r,s)));

DISPLAY v0, qv0, q0, p0, G0, M0, Y0, tau;

*========================*


*===Equation System===*

Positive Variables
        q(j,r,s)    Quantity of a good j shipped from region r
        to region s,

        G(j,r)      Price index of a good j in region r,

        M(j,r)      Composite index of the consumption of a good j
        in region r,
```

```
        Y(j,r)          Total spending of region r - or
        total income in region r
        ;

Parameter
        p(j,r)          Price of a product j in region r;

Equations
        PI(j,s)         Price index of product j in region s,
        QI(j,s)         Quantity index of product j in region s,
        QS(j,r,s)       Quantity shipped of j from region r
                        to region s,
        INCOME(j,r)     Income or total spending of region r
        ;

PI(j,s)..
    sum(r, (p(j,r)*tau(j,r,s))**(1-sig))**(1/(1-sig)) -
    G(j,s) =e= 0;

QI(j,s)..
    sum(r, (q(j,r,s)/tau(j,r,s))**((sig-1)/sig))**(sig/(sig-1)) -
    M(j,s) =e= 0;

QS(j,r,s)..
    (p(j,r)*tau(j,r,s)/G(j,s))**(-sig)*M(j,s)*tau(j,r,s) -
    q(j,r,s) =e= 0;

INCOME(j,s)..
     sum(r, q(j,r,s)*(p(j,r)*tau(j,r,s))) -
     Y(j,s) =e= 0;

model A_1 /PI.G, QI.M, QS.q, INCOME.Y/;

*Set the level values and check for benchmark consistency
G.l(j,r)        = G0(j,r);
q.l(j,r,s)      = q0(j,r,s);
Y.l(j,r)        = Y0(j,r);
p(j,r)          = p0(j,r,r);

*Fix the quantity index level to its initial value
```

```
M.fx(j,r)      = M0(j,r);

*Set some boundaries
G.lo(j,r)       = 1e-6;
q.lo(j,r,s)     = 1e-6;
Y.lo(j,r)       = 1e-6;

Solve A_1 using MCP;

DISPLAY G.l, M.l, p, q.l, Y.l, tau;

*========================*



*===Result Export in a Excel file===*

execute_unload "test.gdx" G.l, M.l, q.l, Y.l, p, tau;

execute 'gdxxrw.exe test.gdx o=test.xlsx var=G.l rng=G!a1
var=M.l rng=M!a1 par=p rng=p!a1 var=q.l rng=q!a1
var=Y.l rng=Y!a1 par=tau rng=tau!a1'

*=======================================*
```

F.2 GAMS code for Krugman's model

```
$Title Krugman Trade Equilibrium from The Spatial Economy: Cities,
Regions, and International Trade

*Valentin MATHIEU, AgroParisTech / INRA (valentin.mathieu@inra.fr).
*July 2019

*Copy past the good list of regions during the Set

*List of continents
*AFR,AME,ASIA,EU,OCE

*List of sub-continents
*AUSNZ,CARI,CAMER,CASIA,EAFRI,EASIA,EEURO,MELA,MAFRI,NAFRI,
NAMER,NEURO,POLY,SAMER,SEASIA,SAFRI,SASIA,SEURO,WAFRI,WASIA,WEURO
```

```
Set
r          countries or regions     /AFR,AME,ASIA,EU,OCE/
k          goods                    /IR/;
Alias (r,s);

Parameters

         sig           elasticity of substitution  /3/,


         v0(k,r,s)     benchmark value of quantities of a product k
         shipped from region r to region s,

         qv0(k,r,s)    benchmark quantities of a product k
         shipped from region r to region s without autoconsumption,

         pf0(k,r,s)    benchmark delivered price of a product k
         shipped from region r to region s,

         q0(k,r,s)     benchmark quantity of product k
         shipped from region r to region s with autoconsumption,

         qf0(k,r,s)    benchmark quantity of a variety of the good k
         shipped from a firm in region r to region s with autoconsumption,


         G0(k,r)       benchmark price index in r,

         M0(k,r)       benchmark quantity composite index of
         the consumption of a product k in region r,

         Y0(k,r)       benchmark total spending of region r
         or regional income,

         N0(k,r)       benchmark number of firms
         producing product k in region r,

         tau(k,r,s)    iceberg transport cost factor
         ;
```

```
*===Calibration===*

Table N0(k,r) initial number of companies
            AFR        AME        ASIA        EU        OCE

IR      10.000     10.000     50.000     10.000     10.000
;

*N0(k,r) = 50;

*===Import Data from Spreadsheet===*

*Build gdxxrw instruction file
$onecho > importQ.txt

*Choose the type of Dataset you want to import
*Pay attention to the cells imported if you modify the xlsx file

*For Continents
*IR
par=q0 rng=Matrix_Q_IR!A10:G15 Rdim=2 Cdim=1
set=r rng=Matrix_Q_IR!B11:B15 Rdim=1 Cdim=0
set=s rng=Matrix_Q_IR!C10:G10 Rdim=0 Cdim=1


*For Sub-Continents
*IR
*par=vx0 rng=Matrix_Q_IR!A26:W47 Rdim=2 Cdim=1
*set=r rng=Matrix_Q_IR!B27:B47 Rdim=1 Cdim=0
*set=s rng=Matrix_Q_IR!C26:W26 Rdim=0 Cdim=1

$offecho

$onecho > importV.txt

*Choose the type of Dataset you want to import
*Pay attention to the cells imported if you modify the xlsx file

*For Continents
*IR
```

```
par=v0 rng=Matrix_V_IR!A10:G15 Rdim=2 Cdim=1
set=r rng=Matrix_V_IR!B11:B15 Rdim=1 Cdim=0
set=s rng=Matrix_V_IR!C10:G10 Rdim=0 Cdim=1

*For Sub-Continents
*IR
*par=v0 rng=Matrix_V_IR!A26:W47 Rdim=2 Cdim=1
*set=r rng=Matrix_V_IR!B27:B47 Rdim=1 Cdim=0
*set=s rng=Matrix_V_IR!C26:W26 Rdim=0 Cdim=1

$offecho

$onecho > importQV.txt

*Choose the type of Dataset you want to import
*Pay attention to the cells imported if you modify the xlsx file

*For Continents
*IR
par=qv0 rng=Matrix_Q_IR!A1:G6 Rdim=2 Cdim=1
set=r rng=Matrix_Q_IR!B2:B6 Rdim=1 Cdim=0
set=s rng=Matrix_Q_IR!C1:G1 Rdim=0 Cdim=1

*For Sub-Continents
*IR
*par=qv0 rng=Matrix_Q_IR!A1:W47 Rdim=2 Cdim=1
*set=r rng=Matrix_Q_IR!B2:B22 Rdim=1 Cdim=0
*set=s rng=Matrix_Q_IR!C1:W1 Rdim=0 Cdim=1

$offecho

*Choose the good way to import the file : in GAMSStudio files ,
just have to change the file 's name
$call gdxxrw.exe C:\Users\vlmathieu\Documents\GAMSStudio\
Armington_Krugman_Models\Data\Continents.xlsx @importQ.txt
$gdxin Continents.gdx

$load q0
$gdxin

$call gdxxrw.exe C:\Users\vlmathieu\Documents\GAMSStudio\
```

```
Armington_Krugman_Models\Data\Continents.xlsx @importV.txt
$gdxin Continents.gdx

$load v0
$gdxin

$call gdxxrw.exe C:\Users\vlmathieu\Documents\GAMSStudio\
Armington_Krugman_Models\Data\Continents.xlsx @importQV.txt
$gdxin Continents.gdx

$load qv0
$gdxin

qf0(k,r,s) = q0(k,r,s)/N0(k,r);
pf0(k,r,s) = v0(k,r,s)/qv0(k,r,s);
tau(k,r,s) = pf0(k,r,s)/pf0(k,r,r);

*Sensitivity analysis on tau
tau(k,r,s) = 1*(pf0(k,r,s)/pf0(k,r,r));
tau(k,r,r) = 1;

G0(k,s) = sum(r, N0(k,r)*(pf0(k,r,r)*tau(k,r,s))**(1-sig))**(1/(1-sig));
M0(k,s) = sum(r, N0(k,r)*(qf0(k,r,s)**((sig-1)/sig)))**(sig/(sig-1));
Y0(k,s) = sum(r, (q0(k,r,s)*N0(k,r))*(pf0(k,r,r)*tau(k,r,s)));

DISPLAY v0, qv0, q0, qf0, pf0, G0, M0, Y0, N0, tau;

*======================*


*===Equation System===*

 Positive Variables
        qf(k,r,s)       Quantity of a variety of a good k shipped
                        from a firm in region r to region s,
        q(k,r,s)        Quantity of a good k shipped from region r
                        to region s,
        G(k,r)          Composite price index of a good k in region r,
        M(k,r)          Composite index of the consumption of a good k
                        in region r,
        Y(k,r)          Total spending of region r - or total income
```

```
                      in region r
        ;

Parameters
        pf(k,r)       Firm price of a product k in region r,
        N(k,r)        Number of firms producing k in region r;

Equations
        PI(k,s)       Price index of product k in region s,
        QI(k,s)       Quantity index of product k in region s,
        QfS(k,r,s)    Quantity shipped of k from region r to
                      region s,
        INCOME(k,r)   Income or total spending of region r
        ;

PI(k,s)..
    sum(r, N(k,r)*(pf(k,r)*tau(k,r,s))**(1-sig))
    **(1/(1-sig)) -
    G(k,s) =e= 0;

QI(k,s)..
    sum(r, N(k,r)*((qf(k,r,s)/tau(k,r,s))**((sig-1)/sig)))
    **(sig/(sig-1)) -
    M(k,s) =e= 0;

QfS(k,r,s)..
    (pf(k,r)*tau(k,r,s)/G(k,s))**(-sig)*M(k,s)*tau(k,r,s) -
    qf(k,r,s) =e= 0;

INCOME(k,s)..
     sum(r, (qf(k,r,s)*N(k,r))*(pf(k,r)*tau(k,r,s))) -
     Y(k,s) =e= 0;


model K_1 /PI.G, QI.M, QfS.qf, INCOME.Y/;

*Set the level values and check for benchmark consistency
G.l(k,r)        =    G0(k,r);
qf.l(k,r,s)     =    qf0(k,r,s);
Y.l(k,r)        =    Y0(k,r);
pf(k,r)         =    pf0(k,r,r);
```

```
N(k,r)          =    N0(k,r);

*Fix the quantity index level to its initial value
M.fx(k,r)       =    M0(k,r);

*Set some boundaries
G.lo(k,r)       = 1e-6;
qf.lo(k,r,s)    = 1e-6;
Y.lo(k,r)       = 1e-6;

Solve K_1 using MCP;

q.l(k,r,s)=qf.l(k,r,s)*N(k,r);

DISPLAY G.l, M.l, pf, qf.l, q.l, Y.l, tau;

*=========================*



*===Result Export in a Excel file===*

execute_unload "test.gdx" G.l, M.l, pf, qf.l, q.l, Y.l, tau;

execute 'gdxxrw.exe test.gdx o=test.xlsx var=G.l rng=G!a1
var=M.l rng=M!a1 var=pf rng=pf!a1 var=qf.l rng=qf!a1
var=q.l rng=q!a1 var=Y.l rng=Y!a1 par=tau rng=tau!a1'

*========================================*
```