Essay in Spatial Economic Analysis at the Firm-Level

Enza Maltese

14th February 2011

Contents

I Productivity and Spatial externalities 9

1 Estimating Productivity 10

1.1 Introduction . . . 10
1.2 Literature review . . . 12
1.3 Empirical Strategy . . . 20
1.3.1 Firm level models . . . 20
1.3.2 Productivity Aggregation . . . 28
1.4 Empirical Application . . . 31
1.4.1 Data description . . . 31
1.4.2 Results . . . 32
1.5 Conclusions . . . 34

2 Agglomeration Effects 38

2.1 Introduction . . . 38
2.2 Literature review . . . 41
2.3 Methodology . . . 47
2.3.1 Computing Spatial indices . . . 47
2.3.2 Empirical specification . . . 56
2.4 Results . . . 59
2.5 Conclusions . . . 63

II Does distance matter? 72

3 Spatial econometric application to Province-Level Productivity 73

3.1 Introduction . . . 73
3.2 Literature review . . . 74
3.3 Spatial Econometrics Methods . . . 77
3.3.1 Spatial panel . . . 84
3.4 Methodology and Results . . . 87
3.5 Conclusions . . . 92

4 Spatial economies within a General Equilibrium framework: New Economic Geography 97

4.1 Introduction . . . 97
4.2 Literature Review . . . 99
4.3 The Theoretical model . . . 105
4.3.1 Dixit-Stiglitz Model of Monopolistic Competition . . . 107
4.3.2 Core - Periphery . . . 120
4.4 Empirical model . . . 128
4.5 Empirical application . . . 130

4.5.1 Methodology . . . 130
4.5.2 Data Description. . . . 132
4.5.3 Results . . . 134
4.6 Conclusion . . . 137

A Statistical Appendix **152**

B Mathematical steps of NEG framework **154**
B.1 Dixit- Stiglitz Model . . . 154
B.2 No Black hole condition . . . 159
B.3 Changes in Manufacturing workers share effects . . . 161

Acknowledgments

I would like to express my sincere gratitude to my advisor Prof. Giorgio Fazio for the continuous support of my doctoral study and research, for his patience, motivation and enthusiasm. His guidance helped me in all the time of research and writing of this thesis. I could not have imagined having a better advisor and mentor for my Ph.D study.

I wish to thank Prof. Fingleton, who has advised my MSc dissertation, allowing to deepen my knowledge in New Economic Geography, topic that has been extended and has become a chapter of this work.

My sincere thanks also goes to Prof Chihwa Kao, Prof. Badi Baltagi and Prof. Stuart Rosenthal for welcome me in the Center for Policy Researchat Maxwell School in Syracuse University and for their meaningful suggestions.

I would like particularly to acknowledge the Board of Professors who organise and manage this Doctoral Program, and in particular Prof. Vincenzo Fazio and Fabio Mazzola who have coordinated it.

I wish to thank for helpful comments and useful discussions, the participants at the 2009 NARSC Conference in San Francisco and at the Department Seminar where I presented an extract from chapter two of this thesis.

I am indebted to my colleagues and friends Dott. Paolo Li Donni and Dott. Davide Piacentino, for providing a stimulating environment in which to learn and grow.

Last but not the least, I would like to thank my family and all the people I met during these years.

Introduction

In recent years, there has been an increase of interest on the role of geography and space for economic activity. This is also due to the emergence of the New Economic Geography that combines the traditional agglomeration theories into a general equilibrium framework. From the empirical standpoint, geography has been neglected for a long time because of the difficulty to combine time series with spatial effects.

We aim in this thesis to underline the role of space by considering its effects on productivity and wage rates at the level of micro-units, such as firms. We exploit this section to provide a brief overview of the evolution of how space has been considered in the literature and in order to clarify the structure of the overall work. Then, in the following chapters we try to go deeper and analyse empirically some of the predictions highlighted in the theory.

Space is an important factor for the economy and productive system; geography is relevant to the location choice of firms (industry) as well as the values of goods. According to the value theory proposed by Debreu (1959), where the spatial dimension is considered a possible state (of nature), the geographical dimension allows attributing different values to goods produced and/or sold in different locations. This idea has been underlined by Starlett (1974) who distinguishes between homogeneous and heterogeneous space. The former relates to firms facing the same convex production set and to consumers with the same preferences, but it is not neglected that good are traded and that transport is not costless.

This interpretation has led to the development of the spatial impossibility theorem, which states that in economies characterised by a finite number of consumers, locations and firms, it is not possible to achieve any equilibrium if transport costs are introduced. This is due to the idea that economic activity is perfectly divisible and consumers do not have any reason to distinguish among locations, so that the only possible equilibrium is autarky. Recognising the indispensable role of transport costs and the unrealistic autarky assumption, economic theory needs to relax some assumptions such as homogeneous space and admit no-convex production sets to have a good picture of the real economy. Introducing geography and distance into economic models requires new assumptions, as a consequence of the combination of the spatial impossibility theorem and the aim to explain the spatial pattern of the economic activities. These assumptions are:

- heterogeneous space (due to the uneven distribution of technologies, natural resources and the existence of point of transport);
- production and consumption externalities (positive effects generated by specialisation and diversity attracting a large number of agents);
- monopolistic competition (firms' choices are motivated by profit maximisation and produce differentiated products, i.e. commodities produced in different locations are considered as different), so firms are not price takers, but they charge a mark-up on their marginal cost in order to obtain positive profits.

Relaxing the perfect competition assumption, we can incur in scale economies. An important distinction about scale economies is between internal and external. In particular, internal scale economies are associated to large firms, which are able to produce optimising inputs employed; instead, external scale economies are more relevant for small and medium firms, which are here under analysis. Firms try to exploit economies that are generated outside them and inside a local setting where they are localised. This is the concept of externalities that we mainly aim to analyse. Externalities allow, even small firms, to exhibit increasing return to scale. [1] Three economic fields have taken interest in space: international trade theory, regional and urban economics. In particular, within International Trade theory, a spatial economic framework can be developed exploiting trade theory, based on local productivity differences. Exponents of this strand of literature are Ricardo (1821) and Heckscher (1919)- Ohlin (1933).

Ricardo (1821)has proposed a theory that links land use to its relative fertility, this concept has been expanded later to other goods and factors, arguing that countries have the same endowments but hold different technologies. If a country could produce everything more efficiently than another country, it would specialise in what it was best at producing and trade with other nations.

Instead, the Heckscher-Ohlin theory focuses on comparative advantages across locations, sustaining that countries have the same technologies but have different endowments, the international immobility of factors leads to different prices if economies are closed. Instead, trade leads countries

[1] Moreover, the New Economic Geography framework has considered the economy as composed by two main sectors, one exhibiting increasing returns to scale and the other exhibiting constant returns to scale.

to specialise in the production of the good whose relative endowment is higher. Within regional and urban economics, several definitions of space have been proposed. In particular localisation theories have used a physical/geographic concept in terms of distance and transport costs. Localisation theories focus on three main aspects: 1) firms localisation decisions, which follow transport cost minimisation and agglomeration economies assuming a point market(see Weber (1909) for an example; 2) spatial competition focusing on how the market is allocated among producers and assuming profits maximisation, spatial homogeneous demand and point firms, an approach has been followed by Hotelling (1929) and Losch (1954); 3) production areas focusing on location equilibrium derived by balancing two factors, transport and land cost, and assuming a point final market and spatial spread supply, the approach has been followed by Alonso (1964).

The meaning of space has evolved together with the emergence and the spreading of new theories. For example, local development theories consider the space as a container of relationships, which can be established among economic agents representing the driving force of growth. Relevant factors to achieve these linkages is the geographical proximity, so they become easier if the geographical area analysed is quite small.

Local development theories have led to the formulation of the concept of diversified-relational space, which has reached the maximum acknowledgement during the 1970s, together with the district theories, where space is considered as a container of externalities, fed by the decrease in transaction costs due proximity, and implying an improvement of firms productivity through a cumulative process.

Diversified-relational space refers mainly to the social relations, which can be established in a geographical area, such as a region or a city. The occurrence of these relations drives firms to cluster in order to benefit of these positive externalities.

These phenomena has been analysed in pioneer works by Marshall (1890),Jacobs (1969) and Porter (1990), who have investigated different aspects of agglomeration. In particular, Marshall (1890), observing the concentration of cutlery and hosier industries in the regions of Sheeld and Northampton, noted how these areas were not characterised by rich natural endowments, so he recognised that there must be other factors which can cause these phenomena. Marshall has sustained that , industrial districts are places where there exists an unconscious learning process.

Beyond this mystical approach, three main sources of agglomeration externalities have been identified by Marshall, and rediscovered later by Arrow and Romer:

- Inputs market's externalities: cluster of producers of a given industry on a given region arises an incentive for input suppliers to locate in the neighbourhood, implying that producers can share specialised services, cut transport costs yielding more efficient purchases of inputs.
- Labour market's externalities: industrial clusters do not refer only to production, but also to the clustering of workers, and in particular specialised workers, leading to the creation of a specialised labour market, where firms can draw specific skills.
- Knowledge externalities: industrial clusters favours the exchange of information and knowledge, creating a favourable environment for technological and knowledge spillovers.

Often, these three kinds of industry externalities are recalled as localisation economies, as opposed to Jacobs' urbanisation economies.

Jacobs (1969)has proposed her idea about agglomeration, mainly referred to the role of the city. She believed that the 'city' is dynamic and acts as a living organism and she has proposed that city should be dense, full of diversity and complexity instead of the formulated and consolidated urban planning since those characteristics facilitate the vitality to maintain the ecology of the city and make the city alive. She has introduced the concept of cross-fertilisation, that refers to the spreading of knowledge across different industries in a given territory.

Porter's idea has been often neglected in agglomeration studies, probably because his focus is on management and organisational economics, but he has also followed innovation theories. Porter has focused on the role of competition, increasing when agglomeration occurs, sustaining that competition favours innovation improving firms' performance. In this thesis, we mainly dwell on three concepts of space. First, how space is seen within local development theories, empirically testing the three theories proposed by Marshall, Jacobs and Porter. Then, we proceed to consider the evolution of the concept of space identifying space with distance. Since positive effects arise through proximity, they tend to decrease as distance increases. Moreover, we have proceeded to analyse the effect of transport costs, by taking the appropriate transformation of distance. The analysis related to the role of distance follows two approaches. First, we exploit new tools such as spatial econometric models, which allow examining how proximity can affects firms' productivity. The other approach originates from the NEG framework and focuses on the role of transport cost, within a General Equilibrium Model, where externalities are measured by Market Potential.

Applying Spatial econometric tools to measure how province-level productivity can affect the productivity of neighbouring provinces, does not require additional assumptions related to the theory, but only assumptions underlying the econometric estimator.

This thesis consists in two parts, each divided in two chapters. The first part carries on the traditional approach to study agglomeration economies, applying it on micro panel data. In particular, chapter 1 engages firm level production function estimation and Total Factor Productivity measurement, the second chapter focuses on the external economies measurement and their effects on firms productivity. The second part centres on the role of distance. In particular, in chapter 3 we replicate the analysis developed in chapter 2 exploiting new tools provided by spatial econometric . Finally, in chapter 4, distance is transformed in transport costs and agglomeration forces are analysed in a General Equilibrium framework.

Part I

Productivity and Spatial externalities

Chapter 1

Estimating Productivity

1.1 Introduction

In economic growth theory, productivity is considered as an engine of growth. Therefore, its importance cannot be underestimated in empirical growth regressions and in the investigation of convergence. However, the correct measurement of productivity is not straightforward and remains a contentious task. Since one of the objectives of this work is to analyse the impact of spatial economies on productivity, we think it is important to first address the issue of productivity measurement.

In many empirical studies, productivity has been investigated looking at the "efficiency" of human resources, i.e. labour productivity. While this measure of productivity is rather simple to determine, since its computation requires only labour and output, it can also be considered as partial. A more complete measure of productivity, permitting to take into consideration the contribution of all the inputs employed in the production process is Total Factor Productivity (TFP, hereafter). This feature of TFP, together with the pervasiveness of the role of spatial externalities, makes it particularly attractive for our empirical investigation.

TFP can be considered as a proxy of technological progress, a key variable in in the neoclassical growth accounting framework. This framework determines the growth of output as the sum of inputs growth, such as capital and labour, and identifies Total Factor Productivity as the portion of output growth that can not be explained by the growth in inputs. An increase in TFP, keeping input factors constant, can be interpreted as an improvement of the efficiency to combine inputs and determines a shift in the production frontier. However, given its residual nature, TFP can be considered as "unobservable". [1]

Productivity measurement has received a lot of attention, in particular in order to understand income differences across countries/regions. Indeed, earlier studies on productivity have been conduced at the macro-level, such as the country/regional level. In these studies, the economy wide output is compared to the economy wide stock of capital and labour. Clearly, this is done after the aggregation of micro level data, since only at the micro level there exists a real production function. The optimal level of aggregation to measure productivity represents an important issue in the productivity literature.

Awareness of the limits of the aggregate production function estimation, together with the increased availability of longitudinal micro-level data and the development of micro-founded economic theory, have increased the interest in productivity measured at the micro-level.

Productivity analysis at the micro level aims to analyse the smaller units of the supply side of the economic system, i.e. firms or plants. The aggregate productivity is obtained as a sum (or weighted sum) of the productivity of the micro units. An important point is that productivity can be measured at different levels of aggregation connected with each other and aggregate productivity grows if the productivity of every component grows, or it grows the share of the component with higher productivity. Indeed, it is fundamental to implement an appropriate aggregation procedure that takes into consideration not only productivity, but also the role of each agent in the market.

A "bottom-up" approach to calculating aggregate TFP, , where the bottom level is represented by firms/plants, entails many advantages. First, it allows taking into account the heterogeneity of firms. Second, it considers the reallocation of input and output factors and firms' turnover, so relevant in the understanding of the determination of TFP.

Many factors can determine firms' heterogeneity. Among the others, managerial capabilities, uncertainty about the innovation process (both in terms of product and production technology)

[1] Many definitions have been provided for TFP, but, probably, the most common is the one proposed by Griliches (1979) defining TFP as "a measure of our ignorance". This definition identifies TFP with Solow's residual, that is determined from the production function estimation and it represents everything that input factors, such as labour and capital, are not able to explain of output.

and its diffusion consumer tastes, capital vintage,[2] and localization.

Every cause of heterogeneity represents an important aspect of productivity, but we here want to focus particularly on the heterogeneity caused by the location choice of firms. Specifically, firms localization decisions can arise from two main kinds of advantages; the so called first and second nature advantages. The former are due to the physical endowment of a place , that may decrease the costs faced by firms, the latter are related to human relationships and the exchange of information. We do not want to neglect first nature advantages, that we will introduce as control variables, but we want to focus on second nature advantages. Indeed, according to the agglomeration theories, that will be discussed later , we think that firms' choice to locate close to each other can favour the exchange of knowledge and improve firms productivity.

In order to achieve a consistent measure of TFP, we need to appropriately estimate the production function. In order to do this, we start from the investigation of the micro-level productivity, comparing different approaches to estimate the firm level production function, and then we proceed to aggregate the estimated firms' productivity. For comparison, we also estimate aggregate production functions and highlight the main differences and potential biases in the results.

This chapter is organized as follows: in section 2 we provide a review of the literature on production function estimation from the studies on productivity at macro level to the recent works on micro level productivity. In this section, we discuss the aggregation bias issue and some of the proposed solutions to this problem. In section 3, we report a brief data description, in section 4 we describe the most common methodologies adopted to estimate the firm level production function. Section 5 reports on the results obtained and section 6 concludes.

1.2 Literature review

There is an ample reference literature on TFP both at the theoretical, including the growth models developed since the 1950s, and at the empirical levels.

The neoclassical theory attributed an important role to technological progress, as the engine of growth, overcoming the traditional hypothesis that it is capital accumulation that determines growth.

[2] This feature is important since productivity growth can be affected by the placing of new technologies. Indeed, not all firms are able to innovate and capital can be subject to different rates of obsolescence and senescence.

Originally, growth theories aimed at analysing differences in per capita income across countries, sustaining that variables, such as capital and output, grow on their balanced path in line with their steady state equilibrium and that every economy will achieve the balanced growth path. In particular, the initial conditions of these variables do not matter, but economies facing the same exogenous variables will achieve the same steady state; this idea implies the concept of convergence.

The inability of neoclassical growth theory to predict some of the stylized facts of growth has raised criticism. In particular, it has been highlighted how capital and output growth do not necessarily decline over time and, further, that traditional models are built on aggregated date and do not consider individual behavior.

Some of the drawbacks that characterized the neoclassical growth theory have been overcome by the endogenous growth theory that mainly focuses on the models proposed by Lucas (1988), based on human capital, and by Romer (1986a)and Solow (1994), based on the assumption of endogenous technology.

Endogenous growth theories argue that technology requires the development of know-how through workers' experience and the availability to share and spread knowledge. This idea is formalized by Arrow (1962) as *learning by doing* based on the following points: knowledge is accumulated over time through practice and increases the production possibilities; knowledge is considered as a public good, i.e. it is non-rival and may be considered as a positive externality. Each firm can benefit of technology developed by other firms and in the overall economy.

Given our empirical focus, we do not discuss in depth the theoretical literature on TFP, and, instead, we prefer to analyse the most commonly adopted methodologies to determine TFP at the macro and micro level. At the macro level, according to Islam (1999), the common denominator to compute TFP at the country level is represented by the the growth accounting approach. This approach has been developed in three directions: 1) Time series ; 2) Cross-section and 3) Panel regression. The time series growth accounting approach has been commonly used to measure the TFP of industrialized countries, since they hold the longest tradition in data collection, and therefore longer temporal series. This approach allows investigating how each factor affects the aggregate output growth rate. The starting point is a neoclassical production function, such as $Y(t) = A(t) \cdot F[K(t), L(t)]$, where $A(t)$ represents technological progress, $Y(t)$, $K(t)$ and $L(t)$ are, respectively, output, capital and labour input factors. In order to identify productivity from the production function equation, it is standard to take the log of the production function and then differentiate with respect to time. The assumption of perfect competition allows to refer to the Euler equations, which imply that the input costs will be equal to the input marginal products, respectively capital marginal product $\frac{A \cdot F_K \cdot K}{Y}$ indicated by $\alpha(t)$ and labour marginal product $\frac{A \cdot F_L \cdot L}{Y}$ indicated by $(1 - \alpha(t))$. Perfect competition, together with the assumption of constant return to scale, implies that the sum of labour and capital marginal product, is equal to one.

The growth rate of technological progress, our measure of TFP, is the growth rate of output minus the weighted sum of the growth rate of inputs factors. Otherwise, empirically we can define TFP as a residual obtained by estimating the production function and subtracting the fitted values from actual output.

As stressed by Jorgenson and Griliches (1967), this procedure can be enhanced by paying attention to the qualitative improvements of inputs factors that should be measured dis-aggregating the information related to the quantitative and qualitative characteristics of inputs, i.e. qualitative aspects of labor inputs can be determined through education, years of schooling and working experience. The qualitative information of capital can be obtained disaggregating capital for species. However, this approach presents two drawbacks: it assumes technology as exogenous and it does not allow explaining how the change in inputs and TFP growth are correlated to public policies, technology and preferences.

The cross–sectional Growth Accounting approach has been proposed by Hall and Jones (1996), who aim to apply the time series growth accounting approach to cross sectional data. Also in this case, the starting point is a production function that varies across countries. This approach introduces in the production function the concept of human capital, defined as an exponential function of the number of workers[3]

The main advantages of this approach are that it does not require a production function form and it does not face time varying econometric estimation issues. The main disadvantages are represented by the issue of countries' order, as including /excluding a country could have an important effect since each country has a different weight, and, furthermore, by the initial assumptions on the capital stock and its depreciation.

Finally, the Panel Regression approach has been proposed in order to overcome some of the drawbacks of the previous approaches and also to explain some empirical regularities in countries' growth. It develops from the more recent cross-sectional approach to measure TFP. Technology is

[3]The production function is represented by $Y_i = A_i F(K_i, H_i)$ where H is the stock of human capital defined as $H_i = e^{\phi(S)} L_i$

described as Cobb-Douglas and is characterised by the exponential growth of technological progress. Taking the log of the production function, the exponential growth of technological progress allows separating the initial level of technology with respect to its growth rate, and assuming that every country faces the same growth rate of technology, the relative TFP between two countries will be constant and equal to the initial level of technology of those countries[4]. The panel approach has been adopted in Islam (1995) using both Chamberlain (1982)Chamberlain (1984) Minimum Distance estimator and using the covariance estimator (based on the fixed effects model).

Recognising the limits of analysing the production function at the aggregated level, in recent year we have seen the emergence of a new direction in the empirical analysis, consisting in estimating the production function, and related productivity, at the firm level.

According to Lutero (2010), the aggregation issue may be studied from different points of view: temporal aggregation, contemporaneous aggregation (across variables in order to build composite indices) and longitudinal or spatial aggregation, with respect to the geographical space.

Dealing with disaggregated data implies advantages, such as more powerful and efficient estimation (because of the greater availability of observations and consequent greater degrees of freedom), more complete information, and better forecasts compared to the macro models.

In spite of the advantages of using micro level data, many economists still prefer using aggregated data because they are more easily available, simpler to analyse, allowing more parsimonious analysis and faster statistical process.

On the other hand, the interest in aggregation bias is due to the increased availability of longitudinal dataset together with the developments in econometric tools, such as micro-econometric panel data models or dynamic factorial analysis, which allow taking into consideration information embedded in micro data into an aggregate representation.

Referring more closely to the longitudinal aggregation, we need to recall the seminal contributions of Leontief (1947), Theil (1954) ,Grunfeld and Griliches (1960) and Zellner (1962), on the aggregation bias issue. This issue can be formalised as follows: assuming β is the true value of the parameter analysed, it has been shown that the relation between β value, and parameters obtained performing an aggregated estimator β_A and disaggregated estimators β_D, can be expressed as follow:

[4] I.e. $\frac{A_{it}}{A_{jt}} = \frac{A_{i0} \cdot e^{gt}}{A_{j0} \cdot e^{gt}} = \frac{A_{i0}}{A_{j0}}$

$$\left|plim\left(\widehat{\beta}_D - \beta\right)\right| < \left|plim\left(\widehat{\beta}_A - \beta\right)\right|;$$

denoting that the asymptotic bias is larger in the aggregated estimator.

Albeit the seminal papers on aggregation bias are not recent, the aggregation issue is still debated. In particular, the recent works focus on the aggregation rule, on how to find the suitable one and the link between micro and macro data, as in the case of Lintunen et al. (2009).

Aggregation bias has been considered a very important topic within the productivity and production function literature. Production function aggregation has been criticized because only firms (plants) deal with input factors allocation in the production process. Indeed, firms production function indicates the optimal quantity to produce in order to maximize profits. This thought has been supported by Klein (1964), who sustains that the aggregate production function should be a mirror of the technological relation of the micro production function, since the latter gives the maximum amount that can be produced for a given level of inputs, synthesizing the dynamics of production. Instead, the aggregate production function returns the amount produced, without considering inputs allocation.

A detailed treatment has been proposed by Felipe and Fisher (2003), who consider two kinds of aggregation: one across different kinds of capital employed in the production process and the other one across firms (plants), representing the aggregation that we are interested in. Felipe and Fisher provide a survey of the existing literature and give suggestions to applied economists on how to tackle aggregation issues.

They start recalling the famous Cambridge-Cambridge debate on the definition of capital, since the debate on the aggregate production function has concentrated for a long time on the aggregation of capital. This debate includes controversies between Robinson and her colleagues of Cambridge, UK, and Samuelson and Solow of Cambridge, US. They argue that capital should be considered, on one side, as a financial fund easily transferable to a different activity, and on the other as a collection of production factors.

Robinson criticises this distinction, sustaining that capital is a collection of heterogeneous factors, and that aggregate variables can be compared only if they are measured through monetary units. Some conditions, therefore, need to be satisfied before proceeding to aggregate, in order to avoid the problems due to heterogeneity.

These conditions have been presented as theorems. In particular, we can recall three main theorems, proposed by Leontief (1947), Nataf (1976) and Gorman (1959). *Leontief theorem* provides the sufficient and necessary conditions, which allow a production function to be considered as aggregated. This theorem states that aggregation can be performed only in the case that the marginal rate of substitution across variables included in the aggregate production function is independent from the variables not included in the aggregation. For example, considering the aggregation across capital types, this can be computed only if these are independent from labour.

Nataf theorem stresses that capital aggregation in homogeneous groups is not the only issue and another important point is represented by the aggregation of the production function. Nataf sustains that the aggregate production function is possible only if the micro production function are additively separable in capital and labour (for example CD or CES production functions). Finally, *Gorman theorem* states that the marginal rate of substitution between inputs must be the same across firms to proceed to aggregate inputs.

Other conditions to be satisfied in order to obtain a correct aggregation are proposed by Felipe and Fisher (2003), who criticises Nataf stating that even the aggregate production function needs to satisfy the efficiency condition implying the maximization of output for given inputs. Furthermore, Fisher states that the problem of aggregation does not concern only capital, but also labour and output.Fisher also argues that aggregation does not create any problem when capital and labour are homogeneous across firms, but this condition does not seem realistic since firms are generally heterogeneous.[5] Yet, the main problems still pertain to the capital input, since labour can more easily be considered homogeneous. So, Fisher sustains that is not a good practice to aggregate capital, and it is even more unsuitable when we deal with a large number of firms.

The literature on micro level productivity has not experienced substantial improvements from the theoretical point of view, and it is largely based, like for the macro level, on the growth accounting approach. At the micro level, most improvements are related to the tools applied to obtain consistent estimates of inputs elasticities, such as more efficient estimators provided by the econometric literature.

Departing from the approach followed at the macro level, based on the Ordinary Least Squares (OLS) estimator, studies that analyse the micro level production function have recognized some

[5] In our case, this can be particularly the case since we deal with a large number of firms.

drawbacks that affect this estimator and have applied alternatives such as the Instrumental Variables (IVs, hereafter), the Fixed Effects Estimator, the Generalized Method of Moments (hereafter, GMM), and semi-parametric approach à-la-Olley and Pakes and Levisohn and Petrin. These estimators allow overcoming issues such as endogeneity and simultaneity due to the likely correlation between productivity and input choice, together with the inability, to distinguish firms' heterogeneity.

The Olley and Pakes and the Levisohn and Petrin approaches represent an important turning point in production function estimation at the micro level. Olley and Pakes (1996)(OP), in particular, aim to analyse the telecommunication equipment industry parameters during a period of restructuring and want to investigate the relationship between productivity and firm entry and exit. Since they are interested in the evolution of the industry productivity, they first estimate the production function at the plant level and later use these results to analyze changes in productivity. They recognize the relevance of firm entry and exit into the industry and consider productivity as the main determinant of firm turnover. However, the firm's exit strategy, even if productivity is smaller than an indifference threshold, may depend on capital stock. Their estimator is based on the assumption that there is one unobservable state variable, i.e. productivity, that determines firm behavior and is determined by the observed state variables, i.e. the factor inputs and investment, used as a proxy for productivity. These assumptions allow accounting for the simultaneity issue. Using this estimation procedure, they are able to take into account the impact of firms' turnover, and via the semi parametric approach obtain an estimation of the production function parameters consistent for endogeneity between factor inputs and productivity.

Levinsohn and Petrin (2003)(LP) extend the approach of OP, constructing an estimator that allows overcoming the potential serial correlation between inputs and unobserved specific productivity shocks. In particular, they show that intermediate inputs can solve the issue of simultaneity. However, in place of investment, that can assume zero values, they introduce intermediate inputs, that are usually positive. In their work, they analyze 6665 Chilean manufacturing firms observed yearly from 1979 to 1986, covered in the Chilean manufacturing census of all firms with at least 10 employees, and focus on the four largest industries at three digit-level (Metal, Textiles, Food and Wood products). The intermediate inputs are chosen in their work on the grounds of data availability, i.e. they choose the variables with less zero values (missing values) in the dataset. They then compare their estimator with a simple OLS, OP and System GMM,4 and conclude that parameter sizes do differ depending on the estimation approach.

Criticism to Olley and Pakes (1996) (and LP) approach has been moved by Ackerberg et al. (2006) and Wooldridge (2009), but neither of the alternative approaches have received, to the best of our knowledge, empirical application. In particular, Ackerberg et al. (2006), have stressed how both OP and LP may suffer of collinearity problems, due to the dependence of labour on productivity and capital, underline the importance of the choice of timing and propose a new procedure that departs from the Olley and Pakes (1996) and LP because no coefficients will be estimated in the first stage of estimation. In contrast, input coefficients are simultaneously estimated in the second stage. However, the first step will still be important to extrapolate the untransmitted error ϵ_{ij} from the production function.

Wooldridge (2009)has proposed an alternative approach to the control variable approach of OP and LP, judged inefficient since he sustains that that method ignores the contemporaneous correlation in the errors across the equations that compose the model and does not efficiently account for serial correlation or heteroskedasticity in the errors. Following the same assumptions of LP, Wooldridge sustains that it is possible to obtain the same moment conditions proposed by LP through the GMM estimator. Wooldrige enumerates a number of advantages in its approach: it overcomes the identification issue, faced by LP in the first step, since GMM allows to obtain both capital and labor parameter estimates efficiently; it is robust for correlation and heteroskedasticity and takes into account the cross equation correlation in order to improve in efficiency; GMM joint system estimation does not require bootstrapping to obtain fully robust standard errors and it is computationally simpler than LP. It efficiently uses the moment conditions implied by the OP and LP assumptions and uses the cross-equation correlation to enhance efficiency, as well as an optimal weighting matrix to account for serial correlation and heteroskedasticity.

Finally, there exists a branch of literature that can be considered a bridge connecting micro and macro productivity computation, since it focuses on aggregated productivity obtained from micro longitudinal data. Belong to this branch the empirical analysis related to the reallocation of factors and productivity dynamics provided by Baily et al. (1992), Olley and Pakes (1996), Bartelsman and Dhrymes (1998), Haltiwanger (1997) using plant-level manufacturing data from US; by Aw et al. (1997) for Taiwan; Tybout (1996) for Colombia, Chile and Morocco, and Griliches and Regev (1995) for Israel. All these studies have focused on the productivity effects caused by factors reallocation, noting that the reallocation is common within instead of between sectors.

These authors usually proceed to build an index of industry productivity through a weighted sum of firms' productivity, where weights are the firm's share. This kind of index, even if it allows to take into consideration firms' weights, it does not allow to distinguish between entrant and exiting firms. Further, selection and learning effects could lead to a biased measure. Baily et al. (1992) have proposed an alternative index that combines "within" effects, obtained as a sum of changes in firms' productivity weighted by the share of firm at the previous period, and "between" effects, obtained through the sum of changes in firms' share weighted by the productivity gap of firms with respect to the industry aggregate productivity, and firms' turnover.

Later, Griliches and Regev (1995) have proposed an alternative decomposition of changes in productivity that, as well as Baily et al. (1992), considers firms' turnover, within and between effects. In particular "within" effects are measured as the weighed sum of productivity, where weights are the average firms' share over time. "Between" effects are obtained summing changes in firms' share weighted by firms productivity gap (with respect to the industry productivity). Turnover effects imply that the entrant firms will contribute positively to the industry productivity if the average productivity of entrant firms is greater than the average productivity of the industry. The other way round is true for the retiring firms, which can improve the industry productivity if the average productivity of the firm is less than the average industry productivity. Another possible approach has been proposed by Olley and Pakes (1996), who, in addition to considering firm production function, provide a method of aggregation to investigate industry productivity. Their decomposition method does not explicitly mention the productivity of new entrant or exiting firms, since market dynamics have been considered in the production function estimation. Their approach will be explained in greater detail in the methodology section.

1.3 Empirical Strategy

1.3.1 Firm level models

Several estimators and methodologies have been proposed to derive consistent measures of inputs' elasticity and thus consistent productivity. In this work, we share this goal since TFP represents the main variables in the next analysis.

We have implemented both parametric and semi parametric approaches, in order to compare results obtained implementing different methods; in particular, we start by applying traditional estimators such as OLS and FE estimators, to then apply more efficient and unbiased estimators such as the Generalized Method of moments (GMM) and the LP estimator.

We depart from OP approach, because we have chosen to work with a balanced panel. Our data source does not permit us to distinguish between firm exit and entry decision. In order to obtain the equation to be estimated, we proceed by defining technology. In particular, we adopt the Cobb-Douglas production function:

$$Y_{jt} = A_{jt} K_{jt}^{\beta_k} L_{jt}^{\beta_l} M_{jt}^{\beta_m}$$

where Y_{jt} represents physical output of firm j at time t, K_{jt}, L_{jt}, and M_{jt} point to inputs of capital, labour, and materials, and A_{jt} is the Hicksian neutral efficiency level of firm j in period , which is unobservable.

The next step is to log-linearise the production function, in order to obtain an equation to be estimated, such as:

$$y_{jt} = \beta_0 + \beta_k k_{jt} + \beta_l l_{jt} + \beta_m m_{jt} + \varepsilon_{jt} \tag{1.1}$$

where $ln(A_{jt}) = \beta_0 + \varepsilon_{jt}$.

β_0 measures the average efficiency (both across firms and over time), ε_{jt} is the deviation from average productivity. The error term can be decomposed into an unobservable (ω_{jt}) that represents firm-level productivity and a *i.i.d.* shock (u_{jt}), $\varepsilon_{jt} = \omega_{jt} + u_{jt}$. The production function can be rewritten as follow:

$$y_{jt} = \beta_0 + \beta_k k_{jt} + \beta_l l_{jt} + \beta_m m_{jt} + \omega_{jt} + u_{jt} \tag{1.2}$$

After estimating inputs elasticity, we can determine TFP as a residual:

$$\omega_{jt} \widehat{+} u_{jt} = y_{jt} - \widehat{\beta}_k k_{jt} + \widehat{\beta}_l l_{jt} + \widehat{\beta}_m m_{jt}. \tag{1.3}$$

We start our analysis performing an OLS estimation, even if we know that it does not allow to consider firms' heterogeneity, so that productivity results correlated with the input factors and the error term is not orthogonal, resulting in biased estimation. In order to overcome this issue, the econometric literature suggests the implementation of the Fixed Effect estimator that is unbiased assuming time-invariant productivity, that is $\omega_{it} = \omega_i$. We can rewrite equation (1.4), as follows:

$$y_{jt} = \beta_0 + \beta_k k_{jt} + \beta_l l_{jt} + \omega_j + u_{jt} \tag{1.4}$$

where intermediate inputs are not explicitly introduced among the regressors, since we adopt as proxy of firm output the deflated firms' values added.

The Fixed Effects estimator implies transformation of the model by subtracting the average of each variable over time, so that ω_i disappears from the main equation.

$$(y_{it} - y_{i.}) = \alpha_o + \beta_1(l_{it} - l_{i.}) + \beta_2(k_{it} - k_{i.}) + (\omega_i - \omega_i) + (\varepsilon_{it} - \varepsilon_{i.}) \tag{1.5}$$

$$\widetilde{y}_{it} = \alpha_o + \beta_1 \widetilde{l}_{it} + \beta_2 \widetilde{k}_{it} + \widetilde{\varepsilon}_{it},$$

where $\tilde{y}$ represents the within transformation of value added, $\widetilde{l}$ and $\widetilde{k}$ represent the within transformation of labour and capital and $\widetilde{\varepsilon}$ represent the within transformation of the error term. The production function, so transformed, is characterized by an orthogonal error term and it can be estimated via OLS (GLS) estimator.

Recently, the literature has criticised the assumption of fixed productivity, focusing, instead, on estimators, which relax this assumption, such as the Generalized Method of Moments (GMM) and the Olley and Pakes/Levisohn and Petrin estimator that introduces the assumption that productivity could be varying over time, but firms could improve/worsen their performance. On the GMM estimator performed to estimate firms production function, an exhaustive guide is provided by Bond (2002) Instead, the seminal paper of Olley and Pakes (1996) marks the introduction of semi-parametric approaches to estimate firms' production function.

Previous equations represent the common step of each methodology, that consists of estimating inputs elasticity and then through the parameters obtaining firms' productivity.

Now, we briefly outline these methodologies and related econometric tools, starting from Bond approach. Recalling equation (1.4), we proceed introducing some assumptions drawn from Bond (2002). Firstly, the dependent variable is firm value added instead of firm output. This assumption implies that we remove intermediate inputs from the list of regressors, since value added is net of intermediate inputs cost. Furthermore, the unobservable firms specific component ω_{jt} is composed by fixed and time varying specific, that is $\omega_{jt} = \eta_j + a_{jt}$ and further time varying specific elements exhibits an auto-regressive pattern.

$$y_{it} = \beta_l l_{it} + \beta_k k_{it} + \gamma_t + (\eta_i + a_{it} + \varepsilon_{it})$$

$$a_{it} = \alpha a_{it-1} + e_{it},$$

$$e_{it},\ \varepsilon_{it} \sim MA(0),$$

where y_{it} indicates firm i output (value added) at time t, l_{it} and k_{it}represent input factors, γ_t indicates time fixed effects, η_i indicates firms fixed effects, a_{it} is the time varying firm specific effect that exhibits a regressive mean-reverting pattern ($\alpha < |1|$).

Then, it has been obtained a dynamic production function by substituting the auto-regressive process a_{it} in production function estimation. Redefining parameters, the production function can be rewritten as[6]:

[6]

$$y_{it} = \beta_l l_{it} + \beta_k k_{it} + \gamma_t + (\eta_i + \alpha a_{it-1} + e_{it} + \varepsilon_{it})$$

$$a_{it} = y_{it} - \beta_l l_{it} - \beta_k k_{it} - \gamma_t - (\eta_i + \varepsilon_{it})$$

$$a_{it-1} = y_{it-1} - \beta_l l_{it-1} - \beta_k k_{it-1} - \gamma_{t-1} - (\eta_i + \varepsilon_{it-1})$$

$$y_{it} = \beta_l l_{it} + \beta_k k_{it} + \gamma_t + (\eta_i + \alpha\,(y_{it-1} - \beta_l l_{it-1} - \beta_k k_{it-1} - \gamma_{t-1} - \eta_i - \varepsilon_{it-1})) + e_{it} + \varepsilon_{it})$$

$$y_{it} = \beta_l l_{it} + \beta_k k_{it} + \gamma_t + \eta_i + \alpha y_{it-1} - \alpha\beta_l l_{it-1} - \alpha\beta_k k_{it-1} - \alpha\gamma_{t-1} - \alpha\eta_i - \alpha\varepsilon_{it-1} + e_{it} + \varepsilon_{it})$$

$$y_{it} = \pi_1 l_{it} + \pi_2 l_{it-1} + \pi_3 k_{it} + \pi_4 k_{it-1} + \gamma_t^* + u_{it}$$

where $u_{it} = (\eta_i^* + \omega_{it})$

In line with Bond (2002), we perform both Difference and System GMM estimators, the latter based on two equations: level and first differences, and two sets of instruments, respectively, first differences and lags of independent variables.

Both estimators are characterized by some orthogonality conditions necessary to define appropriate instrumental variables, represented by lags of dependent and independent variables included in the model, instruments have to satisfy orthogonality conditions. Instruments suitables for difference GMM equation have to satisfy the following conditions: $E\left[y_i, \Delta\varepsilon_i\right] = 0 \qquad for \quad i = 1, ..., N$ where $\Delta\varepsilon_i = (\Delta\varepsilon_{i3}, \Delta\varepsilon_{i4}, ..., \Delta\varepsilon_{iT})$.

Orthogonality conditions that instruments need to satisfy for level equation are: $E(\triangle y_{is}, \varepsilon_{iT}) = 0 \quad for \ s = 2, 3, ..., T-1$ and $E\left(\varepsilon_{it} \triangle y_{it-1}\right) = 0 \quad t = 4, .., T$. These conditions allow to face endogeneity caused by the time lag of dependent variable, introducing dynamics in the model, but production function equation is also affected by endogeneity due to productivity and inputs demand relationship. In order to take into account this source of endogeneity we need account for further orthogonality condition, such as:$E(\triangle y_{is}, u_{iT}) = 0 for \quad s = 2, 3, ..., T-1$ and $E\left(u_{it}, \triangle y_{it-1}\right) = 0 \quad t = 4, .., T$.

Olley and Pakes (1996) approach represents a breakthrough in the estimation of production function. They have proposed an algorithm for estimating production function parameters, trying to overcome two issues neglected by the previous literature, such as selection bias, due to the relation between productivity and exit decision, and simultaneity, due to the relation between productivity and input demand. Endogeneity issue is faced by controlling the unobservable productivity ω_{it} through investment, instead, selection bias introduces productivity expectation which depends on age, capital and productivity at time $t-1$; indeed a certain firm will continue to produce if its productivity is greater than the critical value $\omega_t \geq \underline{\omega}_t(a_{t,} k_t)$.

The important founding that allow to define the estimator, is that the investment function is increasing in ω, this result allows inverting the investment function and writing the productivity as:

$$\omega_t = h_t(i_t, a_t, k_t)$$

Substituting the previous expression in the production function:

$$y_{it} = \beta_0 + \beta_a a_{it} + \beta_k k_{it} + \beta_l l_{it} + h_t(i_{it}, a_{it}, k_{it}) + \eta_{it}$$

assuming that $\beta_0 + \beta_a a_{it} + \beta_k k_{it} + h_t(i_{it}, a_{it}, k_{it}) = \phi_t(i_{it}, a_{it}, k_{it})$, we can rewrite:

$$y_{it} = \beta_l l_{it} + \phi_t(i_{it}, a_{it}, k_{it}) + \eta_{it} \tag{1.6}$$

This model does allow to obtain only the coefficient of labour factor, instead is not possible to distinguish the effect of capital and age. In order to estimate the parameters related to capital and age would be defined the probability of surviving of each firm:

$$Pr\left\{\chi_{t+1} = 1 | \underline{\omega}_{t+1}(k_{t+1}, a_{t+1}), J_t\right\} =$$

$$= Pr\left\{\omega_{t+1} \geq \underline{\omega}_{t+1}(k_{t+1}, a_{t+1}) | \underline{\omega}_{t+1}(k_{t+1}, a_{t+1}), \omega_t\right\} =$$

$$= \rho\left\{\underline{\omega}_{t+1}(k_{t+1}, a_{t+1}), \omega_t\right\} =$$

$$\rho(i_t, a_t, k_t) \equiv P_t$$

Obtained the coefficient of labour, we can proceed to the second step consisting in estimating the expected value of $(y_{t+1} - \widehat{\beta}_l l_{it})$ given the information at time t about firms surviving.

$$E\left[y_{t+1} - \widehat{\beta}_l l_{it} | a_{t+1}, k_{t+1}, \chi_{t+1} = 1\right] =$$

$$= \beta_0 + \beta_a a_{it} + \beta_k k_{it} + E\left[\omega_{t+1} | \omega_t, \chi_{t+1} = 1\right] =$$

$$= \beta_a a_{it} + \beta_k k_{it} + g(\underline{\omega}_{t+1}, \omega_t)$$

where function $g(.)$ represents the biases gap between current productivity and the critical value of productivity. The conditional density of $\omega_{t+1}$to ω_t is positive in a certain range of values of $\underline{\omega}_{t+1}$, and so can be substitute in g , as P_t and ω_t terms:

$$y_{t+1} - \beta_l l_{it} = \beta_a a_{it} + \beta_k k_{it} + g(P_t, \omega_t)$$

furthermore ω_t can be obtained through equation (1.6)

$$\phi_t = \beta_0 + \beta_a a_t + \beta_k k_t + \omega_t$$

$$\omega_t = \phi_t - \beta_0 - \beta_a a_t - \beta_k k_t$$

$$y_{t+1} - \beta_l l_{it} = \beta_a a_{it} + \beta_k k_{it} + g(P_t, \phi_t - \beta_a a_t - \beta_k k_t) + \xi_{t+1} + \eta_{t+1}$$

where $\xi_{t+1} = \omega_{t+1} - E\left[\omega_{t+1} | \omega_t, \chi_{t+1} = 1\right]$.

The fundamental system of equations is:

$$y_{it} = \beta_l l_{it} + \phi_t(i_{it}, a_{it}, k_{it}) + \eta_{it}$$

$$Pr\left\{\chi_{t+1} = 1|\underline{\omega}_{t+1}(k_{t+1}, a_{t+1}), J_t\right\} = \rho_t(i_t, a_t, k_t) \equiv P_t$$

$$y_{t+1} - \beta_l l_{it} = \beta_a a_{it} + \beta_k k_{it} + g(P_t, \phi_t - \beta_a a_t - \beta_k k_t) + \xi_{t+1} + \eta_{t+1}$$

The first equation of the system can be estimated through a polynomial approximation and a Kernel estimator. The probability of survival of firms is estimated through a probit model where the RHS of the model is a polynomial of (i_t, a_t, k_t).

Obtained estimates of $\hat{\beta}_l \hat{\phi}_t$ and $\hat{P}_t$ are substituted in the third equation of the system. The parameters (β_a, β_k) are determined by the minimization of the residual sum of squares.

The introduction of the probability of exit (predicted probability), allows Olley and Pakes (1996) to address the downward bias of capital coefficient due to the negative correlation between ε_{jt} e k_{jt}, correlation that arises because firms know their productivity level, that determines the choice of inputs factors. Implementing this procedure, we do not need to assume that inputs are chosen independently from firms productivity level.

Following OP's approach, LP have provided an extension of this approach. LP have criticized recourse to investment as a control variable, implying the truncation of firms with zero or negative investment. They have proposed as proxy variable intermediate inputs instead of investments. The demand function of the intermediate inputs can be expressed as $m_{jt} = m_t(k_{jt}, \omega_{jt})$, intermediate inputs is a function of capital and productivity at the firm level, intermediate inputs are assumed increasing together to productivity ω_t. This assumption allows inverting the function and writing productivity as a function of intermediate inputs and capital, that is the unobservable variable as a function of observable ones, $\omega_{jt} = s_t(k_{jt,} m_{jt})$ where $s_t(.) = m_t^{-1}(.)$. Substituting in the production function:

$$y_{jt} = \beta_0 + \beta_k k_{jt} + \beta_l l_{jt} + \beta_m m_{jt} + s_t(k_{jt,} m_{jt}) + u^q_{jt}$$

LP follow all the steps of the OP procedure, except for the adoption of intermediate inputs as a proxy. They sustain that intermediate inputs is a better proxy because, as investment, can address the simultaneity issue, and further, there exist some benefits arise from intermediate inputs data, that is, investment can be used as proxy only for firms with a positive investment, firms with negative or null investments are truncated. Using intermediate inputs as proxy avoids the loss of observations, since firms generally report positive use of intermediate inputs. Another advantages of using intermediate input as proxy is the lower adjustment cost afterwards a productivity shock.

While acknowledging the importance of productivity reallocation, we have to deal with shortage of data, about firms' turnover. As mentioned above, observing missing values in our data cannot be interpreted as firm exit from the market, in this works we have retain that the presence of missing values in the dataset it could be due to a choice of the firms to not sent balances sheet to the company collecting data.

Furthermore, our prior interest is not market dynamics, but productivity dynamics. This has led us to draw a dataset where firms stay in the market during all periods analysed.

1.3.2 Productivity Aggregation

In this section, we would like to provide a synthesis of Olley and Pakes (1996) methodology to compute aggregate productivity. Although we aim to analyse firm level productivity even in the next analysis, considering aggregated productivity permits to compare our results with other works and also to compare potential different results obtained by firms and aggregated models.

Following Olley and Pakes (1996) approach, we have proceeded to aggregate firm level productivity at province/industry level, taking into consideration the change in share of firms in order to not incurring in aggregation bias.

It is worth stressing that we have aggregated productivity, and we have estimated province/industry production function for sake of completeness. In this section, we can only explain the applied methods, that gives us productivity, which represents the dependent variable of next analysis steps.

We describe this method, that we adopts only partially, since we agree in applying the weighted sum of individual firm productivity, thinking that each firm has a different role and weight in the market, but we sustain that we do not have enough information to analyse turnover effect, since we are not able to distinguish firms'exit from missing values.

An excellent literature survey about aggregate productivity growth has been proposed by Foster et al. (1998) focusing on studies where aggregated productivity is obtained from micro longitudinal data. They are interested in how factors reallocation affects productivity , noting that the reallocation is common within instead of between sectors.

Olley and Pakes (1996), who investigate productivity trend of telecommunication industry have proposed to compute firms'productivity and then provide its industry aggregation. To this aim they have suggested an alternative decomposition, that do not explicitly mention the productivity of new entrant or exit firms, since market dynamic has been considered in the estimation procedure. The industry productivity is

$$p_t = \sum_{f=1}^{N_t} s_{ft} p_{ft}$$

Productivity can be decomposed as follow:

$$p_t = \sum_{i=1}^{N_t} (\bar{s}_t + \Delta s_{it})(\bar{p}_t + \triangle p_{it}) =$$

expanding the product:

$$p_t = N \overline{s}_t \overline{p}_t + \sum_{i=1}^{N_t} \triangle s_{it} \triangle p_{it} + \overline{s}_t \sum_{i=i}^{N_t} \triangle p_{it} + \overline{p}_t \sum_{i=1}^{N_t} \triangle s_{it}$$

$$p_t = N \overline{s}_t \overline{p}_t + \sum_{i=1}^{N_t} \triangle s_{it} \triangle p_{it} + \overline{s}_t \sum_{i=i}^{N_t} (p_{it} - \overline{p}_t) + \overline{p}_t \sum_{i=1}^{N_t} (s_{it} - \overline{s}_t)$$

$$p_t = N \overline{s}_t \overline{p}_t + \sum_{i=1}^{N_t} \triangle s_{it} \triangle p_{it} + \overline{s}_t \sum_{i=i}^{N_t} p_{it} - N \overline{s}_t \overline{p}_t + \overline{p}_t \sum_{i=1}^{N_t} s_{it} - N \overline{s}_t \overline{p}_t$$

$$p_t = \sum_{i=1}^{N_t} \triangle s_{it} \triangle p_{it} + \overline{s}_t \sum_{i=i}^{N_t} p_{it} - N \overline{s}_t \overline{p}_t + \overline{p}_t \sum_{i=1}^{N_t} s_{it}$$

$$p_t = \sum_{i=1}^{N_t} \triangle s_{it} \triangle p_{it} + \frac{1}{N} \sum_{i=i}^{N_t} p_{it} - N \bar{s}_t \bar{p}_t + \bar{p}_t \sum_{i=1}^{N_t} s_{it}$$

$$p_t = \sum_{i=1}^{N_t} \triangle s_{it} \triangle p_{it} + \bar{p}_t - N \bar{s}_t \bar{p}_t + \bar{p}_t \sum_{i=1}^{N_t} s_{it}$$

since $\overline{s_t}$ and $\overline{p_t}$ are constant across firm their sum can be expresses as the number of firms times $\overline{s_t p_t}$, further $\overline{s_t} = 1/N_t$, instead $\sum_{i=1}^{N_t} s_{it} = 1$,

$$\begin{aligned} p_t &= \sum_{i=1}^{N_t} \triangle s_{it} \triangle p_{it} + \bar{p}_t - N \bar{s}_t \bar{p}_t + \bar{p}_t \\ &= \sum_{i=1}^{N_t} \triangle s_{it} \triangle p_{it} + \bar{p}_t - N \frac{1}{N} \bar{p}_t + \bar{p}_t \end{aligned}$$

$$p_t = \bar{p}_t + \sum_{i=1}^{N_t} \triangle s_{it} \triangle p_{it}$$

where $\triangle s_{it} = s_{it} - \bar{s}_t$, $\triangle p_{it} = p_{it} - \bar{p}_t$.

Results found in literature related to the manufacturing sector have shown that the reallocation plays an important role in the productivity growth at industry level and that the productivity growth is sensitive to the decomposition method used. The choice of the decomposition methods is important, time series decompositions show a large contribution of the turnover of firms, (exit of the less productive and entry of the high productive), instead, using cross-section decompositions, the reallocation of inputs and output increases productivity.

Although our main goal is not investigating aggregated productivity, we wish to compare results obtained using firm level productivity and results obtained using province/industry level of aggregation, we rearrange OP decomposition to aggregate firms' productivity at industry and geographical level, obtaining the following equation:obtain

$$p^i_{rt} = \bar{p}^i_{rt} + \sum_{j=1}^{R} \sum_{i=1}^{I} \triangle s_{ft} \triangle p_{ft}$$

where p^i_{rt} indicates productivity of a given province r and sector i, at time t, $\bar{p}^i_{rt}$ is the relative average productivity , R indicate the total number of provinces where firms are localized, Δs_{ft} indicates changes in firms'share and $\triangle p_{ft}$ indicates changes in firms' productivity. The aggregated productivity will be used as alternative to firm level productivity in the succeeding analysis.

1.4 Empirical Application

1.4.1 Data description

We deal with panel data of Italian firms, sourced by the Italian section (AIDA) of the Bureau van Dijk database. We limit our analysis to Small and Medium Enterprises. This dataset mainly provides balance sheet data, but it also provides some industry and geographical information. For each firm we have collected the following information:

- CCIAA , which represents the registration number of the firm. This information has been used as univocal index
- Denomination
- Date of establishment
- Value Added. This items is obtained revenues and costs related to the main activity of the firm.
- Assets including property and equipments, used as proxy of capital. The values provided in the balance sheet is net of the depreciation of capital.
- Number of workers
- Cost of intermediate inputs (from the Profits and Losses table)
- Cost of labour (from the Profits and Losses table)

- ATECO code, which indicates the economic sectors of the firm and it corresponding to SIC-code.
- Location of firms: region, province and zip-code.

The variables used to estimate the production function are Value Added (Y), Capital (K), number of workers (L) and the cost of intermediate inputs (M). Before estimating the production function, we deflate these variables by a specific deflator for each of these items provided by the OECD - STAN-IT dataset. The deflators are provided per industry, in particular using a two digits SIC-code.

Our sample contains 65009 observations, with 9287 firms observed over 7 years from 1999 to 2005. Firms belong to different economic sectors and, in particular, the four main economic sectors represented are manufacture, services, construction and commerce.

Further processing on data are reported in appendix A.

1.4.2 Results

Table 1.1 reports results related to production function estimation obtained by applying at firm level the alternative estimators that have been described in the methodology section. Results appear to be consistent with expected ones, and rather homogeneous. In particular, firstly, as suggested by literature, firms are labour intensive, this result is shared by all methodologies, with the only exception, represented by first difference GMM (Arellano-Bond estimator). We can observe that OLS estimators underestimates capital with respect to LP estimator.

Introducing fixed effects, firstly province FE and later firms fixed effect, we can note that the magnitude of labour elasticity decreases and the magnitude of capital elasticity increases. These results can be interpreted as follows: labour elasticity is overestimated because it includes firms specifics characteristics; introducing fixed effects, we depurate the labour coefficient from these effects. Firms specific characteristics could be related to their location or workers skills or any capability of the firm that allows it to improve output.

We can observe that even the LP estimator, consistent for endogeneity, returns a higher value of capital elasticity with respect to the OLS estimator. Then, estimating the dynamic production function, we obtain that the magnitude of the labour input is lower due to the entry of the lagged dependent variable among the regressors; instead, the lagged capital and labour inputs do not seem to be relevant, given the small magnitude.

Looking at the dynamic specification results, these are characterized by a greater dimension of regressors' matrix due to the introduction of the dynamic of productivity, as suggested by Bond (2002).

GMM estimations, which apply the principal of Instrumental Variables, need to be followed by diagnostic tests, in particular Arellano-Bond auto-correlation test and Hansen/Sargan test for instruments are applied.

Performing these tests after difference and System GMM estimators, we found, in the first case, for the difference GMM estimation Arellano-Bond test rejects the null hypothesis of absence of auto-correlation of first but it does not reject the hypothesis of absence of second order auto-correlation; Sargan test rejects the null hypothesis, so again we have weak instruments. Applying tests to System GMM we obtain that: Arellano-Bond test rejects the null hypothesis of absence of auto-correlation of both first and second order; Sargan test rejects the null hypothesis of goodness of instruments. Test results suggest that our instruments are weak and our model is affected by auto-correlation, so GMM estimator does not seem to be suitable for our data.

Estimating the production function returns us inputs elasticity, so in order to determine firm productivity, we proceed with a further step, that according to the main literature, it is determined as a residual from the production function equation.

For the sake of completeness, we have provided results related to a battery of estimators; in the light of the econometric and economic theory, considerations about endogeneity and dynamic productivity, suggests us that the estimator that better than others allows overcoming simultaneity and considering changes in productivity, is LP.[7], so we have chosen to use in the next steps, as proxy of productivity, the residuals obtained from LP estimations.

Determined firms productivity, we next use this variable in the core analysis to investigate how spatial externalities can affect it, we also proceed to aggregate productivity at the province-industry level, applying OP procedure and analysing the relation between province/industry economies and province/industry productivity. We have proceeded analysing in parallel firms and provinces productivity. The latter can be obtained following two different methods:

[7]it could exist a learning mechanism, that leads firms to improve their efficiency and so their productivity

- aggregate firm's data (summing by province and industry main variables needed to estimate production function) in order to obtain province/industry capital, labour , intermediate inputs and value added.

- maintaining the estimation of the production function, and obtained firms' productivity, we can aggregate at the province/industry level following OP procedure.

Results obtained estimating Production Function at provincial level are reported in table 1.2 Also in this case, as well as firms' one, we found that provinces are labour intensive, and moreover that capital plays little relevance, how small magnitude of its elasticity shows, only results obtained through FE estimator return a higher elasticity of capital, instead we can observe that the other estimators, that assume dynamic productivity return a no significant coefficient. It seems reasonable conclude that the provincial value added is, mainly, explained by labour inputs. Although labour intensive characteristic, introducing province FE, magnitude of labour elasticity decreases, suggesting the importance of geographical effects, initially, absorbed by labour elasticity.

1.5 Conclusions

We have analysed the evolution of the concept of productivity over the last decade, and the importance attributed to this measure. Indeed, productivity is considered the engine of economic growth. Even if the earlier studies have investigated macro level productivity, mainly country-level, awareness of the limits of aggregated production function estimation, together with the increased availability of longitudinal micro-level data, and to the development of micro-foundation of economic theory, focused on the behavior of economic agent, their utility function and their preferences, have increased the interest in productivity measured at micro-level. The advantages of analysing the micro-level data is that they allow to understand the reallocation of input and output factors, to take into account firms' heterogeneity and productivity dispersion.

Indeed, productivity can be thought as the combination of several firm's skills, that combine cause an improvement in their efficiency; among possible causes that can explain firms' productivity we have focused on heterogeneity coming from localization choice and geographical characteristics , that is external elements that can affect firms' productivity of small and medium firms.

We have found productivity through two step analysis, in the first one we have proceeded to

Table 1.1: Production Function Estimation - Firm level

	OLS	OLS	OLS	FE	Difference_GMM	System_GMM	LP
l_{it}	0.7194***	0.69983***	0.6662***	0.35328***	0.15969***	0.3577***	0.7253***
	(0.0051)	(0.0052)	(0.0055)	(0.00514)	(0.0291)	(0.0189)	(0.00466)
k_{it}	0.1083***	0.12134***	0.1208***	0.197***	0.4773***	0.1279***	0.14197***
	(0.00408)	(0.0043)	(0.004677)	(0.0025)	(0.0536)	(0.01707)	(0.02186)
y_{it-1}					0.0832***	0.35578***	
					(0.0212)	(0.03048)	
l_{it-1}					0.0306***	0.0567***	
					(0.0098)	(0.0100)	
k_{it-1}					-0.0238*	-0.0856***	
					(0.01343)	(0.0132)	
$constant$	10.419***	9.69***	10.341***	10.538***	-	7.320***	-
	(0.04453)	(0.207)	(0.0575)	(0.03549)		(0.3071)	
time - FE	yes	yes	yes	yes	yes	yes	yes
firm- FE	no	no	no	yes	yes	yes	yes
province-FE	no	yes	yes	no	no	no	no
industry - FE	no	no	yes	no	no	no	no
R^2	0.6468	0.6591	0.6733	0.5813	-	-	-

Table 1.2: Aggregated Production Function - Provincial level (average input/output industry- province)

	OLS	OLS	OLS	FE	Difference_GMM	System_GMM	LP
l_{it}	0.7412*** (0.0132)	0.714*** (0.0143)	0.6832*** (0.01531)	0.36932*** (0.0125)	0.32977*** (0.11515)	0.5635*** (0.0532)	0.7322*** (0.0291)
k_{it}	0.1094*** (0.01018)	0.1281*** (0.011)	0.112*** (0.0131)	0.16745*** (0.0068)	0.4934*** (0.1578)	0.3014*** (0.0397)	0.1443*** (0.02439)
y_{it-1}					0.08180*** (0.0235)	0.2319*** (0.01842)	
l_{it-1}					-0.05368 (0.0606)	-0.04984 (0.03651)	
k_{it-1}					-.0647159 (0.06323)	-.1583*** (0.0246)	
$constant$	10.281*** (0.1089)	10.124*** (0.1164)	10.465*** (0.144)	10.867*** (0.0983)		7.296*** (0.4208)	
time - FE	yes	yes	yes	yes	yes	yes	yes
province-FE	no	yes	no	yes	yes	yes	yes
industry-dummy	no	no	yes	yes	no	no	no
R^2	0.6757	0.6913	0.7029	0.6275	-	-	-

estimate production function performing a battery of estimators, such as OLS, FE, GMM (both difference and system) and LP. Then, we have chosen among these estimators LP residual to determine our measure of productivity; this choice has been driven by the economic assumption underlying the estimator and econometric suggestions and diagnostic test applied in particular to GMM estimator, believing that LP estimator, can, better than other ones, overcome endogeneity issue and consider firms'heterogeneity.

Results have shown that firms are generally labour intensive.Further we have observed that introducing province FE causes a decrease of labour elasticity, probably geographical effect had been included in this parameter. According to OP and LP, we can further observe as OLS estimator underestimates capital parameter, with respect estimators that take into account firms heterogeneity.

Chapter 2

Agglomeration Effects

2.1 Introduction

In the previous chapter, we have discussed extensively the issue of productivity measurement and have proceeded to estimate productivity at the micro and macro level. Here, we move to the estimation of the role of the environment, and agglomeration economies in particular, for productivity. According to the theory, firms choose to localise close to other firms belonging to the same industry or close to firms belonging to other industries in order to exploit agglomeration externalities.

The importance of space and localisation decisions of firms have been studied since the works of Von Thünen (1826) and Marshall (1890). Recently the interest of researchers has been directed toward the so called "second nature advantages", represented by the positive effects due to human relationships. These are more likely at smaller geographical scales, because smaller areas allow easier exchange of information among agents.

Marshall distinguished between internal and external economies of scale. The former refers to economies arising within the firm because of a decreasing average cost function of a firm, which, further, require a monopolistic power and generally firms of a large size.

The expression "external economies of scale" indicates the positive effects accruing to a firm, due to the decisions and actions of other firms. For example, we can have external economies when the production cost of one firm decreases as the output of another firm increases. External economies increase firms' productivity, allowing firms to pay higher wages and cover the higher commuting costs, caused by the concentration of production in given geographical area. This kind of external ties can be internal to the industry and/or geographical area. When internal to the industry these externalities are referred to as localising economies and when internal to a geographical area they are referred to as urbanisation economies.

Localisation economies have firstly been proposed by Marshall, who coined the term "industrial atmosphere" and then referred as Marshallian externalities. Among these we can distinguish between technological and pecuniary externalities. Technological externalities refer to improvements in production capacity, on how inputs factors are combined, due to the innovation adopted by other firms in the same sector. Pecuniary externalities, instead, operate through prices and generally require imperfect competition. It is worth noting that sectoral agglomeration may imply advantages and disadvantages. The main advantages have been well described by Marshall (1890) and are related to the sharing of input suppliers, labour pooling and information (knowledge /technology spillovers). The main disadvantages are related to congestion costs, such as pollution and higher competition. After Marshall, Arrow (1962) and Romer (1990) contributed to this theory, so that this kind of externalities takes the name of Marshall Arrow Romer (MAR) externalities.

Economies external to the firm but internal to a geographical area have been studied firstly by Jacobs (1969), who focused on externalities arising in cities and due to the greater concentration of demand. These externalities are usually called *urbanisation economies,* and focus on local diversity, differently from MAR externalities which focus on specialisation.

Urban economic theories have emphasised the importance to concentrate the production in few sites, stating that cities have an important role in the economy since in a market based economy, concentrating economic activities in a city can facilitate innovation, since creativity can be increased by personal communication and sharing common interests. Face to face communication allows the transmission of idea and so the development of new product or process, so we can say that cities provides opportunity for knowledge spillovers. Moreover, human capital, which is depend on knowledge and skills of workers depend on schooling, working experience and social interactions, an environment like a city facilitates the contact among workers performing similar tasks and so the learning process and finally their productivity. Indeed, production concentration can increase learning opportunities for workers. Further, cities serve also as production centre since the production of some goods is more efficient in high-density urban environments, this characteristic implies that agglomeration can facilitate trade and commerce, and the production of a large variety of goods and services.

Producing in few places, such as cities allows to benefit of economies of scale. Cities have gradually become centres of production. When has been recognised the advantages that production concentration can arise, productivity process has started to move from the home production to the firm production, . Indeed, producing in a single place enhances returns to scale for two main reasons: factor specialisation and lower average costs. Firms production allows to every worker to specialise in a special task, specialisation implies an increase of labour productivity and time saving. Instead, home production does not allow to achieve an high level of specialisation and all tasks need to be performed by each worker to produce the final output. Moreover, concentrating the production in a factory allow to scaled indivisible inputs, while, when machines are used by only one person they cannot be scaled.

Another factor that encourages clusters of production in cities is represented by the higher wages, generally associated to cities. Wage rates represent the variable that affects workers decisions and can make workers indifferent between working at home or in a factory. So, the wage rate should be equal to the amount that allows workers to satisfy their consumption bundle and cover the opportunity cost to move from home to factory.

Workers of factories will live close to production places, in order to minimise the commuting costs, so around the factory will arise an urban city with a high population density.

Before the contribution of Porter (1990), competition have been seen with a centrifugal force with respect the agglomeration. Porter argues, in line with MAR, that knowledge spillovers favour the industrial concentration, but contrary to MAR point of view, Porter has sustained that competition stimulate firms to innovate.

This work aims to investigate how firm productivity may be affected by spatial externalities. In particular, we have proceeded by analysing the relationship between firm level productivity and province spatial externalities measured through localisation, urbanisation and competition indices.

This rest of the chapter is organised as follows: section 2 contains an overview of the main studies on agglomeration economies, section 3 describes the methodology adopted, summarising the most common indices used in literature and the specification chosen. Finally, section 4 reports on the empirical results and section 5 concludes.

2.2 Literature review

Within the literature on endogenous growth theory, several studies have attempted to investigate the role of technological spillovers in generating growth (see Lucas (1988) and Romer (1986b, 1990)). In regional and urban economics, it is assumed that geographical concentration can favour the spreading of knowledge, intended as the sedimentation of communication, that is easier across firms located close each other.

A large literature analyses the effects of agglomeration economies on productivity. In particular, these works have adopted different measures of productivity, such as employment growth and total factor productivity, while, few studies have used, as proxy of productivity, output produced estimating an augmented production function.

Glaeser et al. (1992), Cainelli and Leoncini (1999), Paci and Usai (2006), belong to the mainstream that analyse the effect of agglomeration externalities on employment growth.

This is the methodology followed by the seminal paper by Glaeser et al. (1992) who analyse the growth of large industries in 170 US cities[1] between 1956 and 1987 and find that employment growth is favoured by local competition and urban variety, in agreement with Jacobian and Porterian externalities, but not by regional specialisation, as expected if MAR-type externalities are at work. They argue that MAR and urbanisation externalities are probably more relevant to the firm's location choice, rather than its growth.

Cainelli and Leoncini (1999) analyse the same relation for Italian provinces, using data from 1961 to 1991. They find strong spatial effects of specialisation and variety on knowledge spillovers. They also observe that patterns of externalities depend on the differences in market structure, endowment and factor accumulation.

Moreover, Paci and Usai (2006) use data on the 784 Italian Local Labor Systems and 34 sectors at the 2 digit ATECO 91-ISIC 3 level 1 (21 manufacturing and 13 service sectors) from 1991 to 2001 in order to analyse employment dynamics in the services and manufacturing sectors of local labour systems, distinguishing among MAR and Jacobian externalities. They find a negative effect of specialisation. Local employment seems to benefit from the presence of a diversified and competitive environment of small firms. Employment growth seems to benefit also from the performance of the surrounding areas.

[1]For example, New York apparel and textiles, Philadelphia apparel and textiles, Philadelphia electrical equipment, etc...

Recent works have departed from the previous mainstream by introducing as proxy of economic performance the Total Factor Productivity, thanks to the increasing availability of data on capital. These procedures have been adopted by Dekle (2002), De Lucio et al. (2002) and Cingano and Schivardi (2004), who compare the effects of agglomeration on TFP and employment growth.

Dekle (2002) uses direct measures of total factor productivity (TFP) growth at the regional level for Japanese prefectures analysing four main sectors: manufacturing, services, finance, and wholesale and retail trade. He finds that significant MAR externalities exist for the finance sector. Evidence of strong MAR, some Porter and no Jacobs externalities are identified for the services sector and the wholesale and retail trade industries. No externalities for the manufacturing sector are detected.

De Lucio et al. (2002) has analysed Spanish Manufacturing firms, using data available in the Spanish Industry Survey from 1978 to 1992 for 26 manufacturing branches to discuss the role of externalities in promoting regional industrial growth. They find evidence of dynamic effects due to specialisation (MAR). However, these effects depend on the level of this variable. At first, specialisation affects productivity growth negatively, but once it reaches a certain level, its effect on growth becomes positive due to knowledge sharing. They do not find clear evidence on diversity and competition externalities.

Cingano and Schivardi (2004) estimate the production function, using the approach of Olley and Pakes (1996) at the firm level and then aggregate the obtained TFP estimates by sector and by Local Labor Systems. They regress both TFP and employment growth on four main indicators - specialisation, variety, city size and competition - from different sources[2]. Their regressions yield different results for TFP and employment growth. TFP growth is enhanced by specialisation and city size but not by urban diversity, the opposite result is reached for employment growth. This contrasting evidence could depend on the fact that employment-based equations might not be able to disentangle the determinants of local industry growth from the sources of productivity growth.

However, agglomeration effects are not only important to examine the economic performance of firms or geographical area considered as economic unit, but external economies also affect the firm decision of localisation. The birth of new firms is more likely in areas where many firms are localised and where it is easier to benefit of positive externalities.

This aspect has been analysed by Rosenthal and Strange (2004),who argue that agglomeration

[2]Their sources are the 1991 Census data on manufacturing firms, data from INPS (the National Institute for Social Provision), and Centrale dei Bilanci (Balance Sheet Collection Centre).

effects should lead new firms to locate close to other firms and in this respect both the size of urban space and its organisation are important. They use the employment rate as a measure of productivity and new firms establishments, bypassing the issues related to the availability and measurement of capital, endogeneity (the establishment of a new firm is decided at t-1 and is uncorrelated with the errors), and the use of aggregate data (disaggregate data are available). In particular, they employ the Dun & Bradstreet Marketplace database, providing data for over 12 million establishments, and use data from the fourth quarter of 1997 to extract two possible dependent variables, i.e. new establishments and their employment. New establishments are those younger than one year as of the fourth quarter of 1997. The new establishments are assumed to have made their location decision on the grounds of the previous year (fourth quarter of 1996) level of employment. On these data, they implement a fixed effects Tobit approach using as dependent variables the number of new establishments per square mile and the new-establishments' employment per square mile. They reject the hypothesis that the industry fixed effects are jointly insignificant, confirming the presence of differences in local attributes across units. Competition within the same industry increases new births and the employment of new firms, but competition in other industries affects both negatively. Diversity of employment increases births and leads to greater employment. Moreover, they aggregate employment into four concentric rings (within 1 mile, between 1 mile and 5 miles, between 5 miles and 10 miles, and between 10 miles and 15 miles from the centroid) and estimate the degree of dissipation of localisation economies. For most industries (five out of six), the effects of localisation economies are more rapidly dispersed after the first few miles and less farther away.

Other variables commonly used as proxy of economic performance and/or productivity are value added or the total sales. These variables lead to an alternative specification implemented by Henderson (2003) and by Martin et al. (2008), which involve the estimation of an augmented production function introducing agglomeration indices as regressors. In particular, Handerson estimates plant level production functions for machinery and high-tech industries. He allows for localisation and diversity externalities and considers these effects for both affiliate (within a corporation) and not affiliate plants. Data on plants' productivity are drawn from the LRD of the Census Bureau from 1972 to 1992. The dataset involved potentially cover 742 counties in 317 metropolitan areas. They find that high-tech industries experience significant localisation economies, while machinery industries do not. Externalities are strongly localised within the own county, and it seems that there are no spillovers from plants in other counties. Moreover, non-affiliate plants seem to generate greater externalities than corporate plants. Evidence of static Jacobs-diversity economies of any type, however, is not found for all industries. Evidence of static urbanisation-scale economies appears for corporate machinery plants.

An additional departure from the standard literature in Martin et al. (2008) is the estimation of an augmented production function at the firm level, combining as regressors firm level factors of inputs and agglomeration indices related to the local area where firms are localised.

Martin et al. (2008)analyse the effect of spatial agglomeration of activities on the productivity of a large panel of individual firms with French data from 1996 to 2004. Their results suggest that French firms benefit from localisation economies, but not from urbanisation economies or from competition effects. However, these benefits, although positive and significant, have a modest size and are geographically limited.

Cainelli and Lupi (2009) use 24,089 Italian manufacturing firms for the period 1998-2001 from AIDA.[3] Variables are expressed in term of changes. They measure localisation economies using the number of firms belonging to the same industry and located within a certain distance and urbanisation economies using the Shannon's entropy index. Agglomeration economies over space are measured using the actual distance between each pair of firms in the sample. Their results suggest that localisation effects are positive, but decrease in distance. On the contrary, diversity effects are negative for shorter distances and become positive for greater distances.

In a recent review of the agglomeration literature, Beaudry and Schiffauerova (2009) present a summary of the indexes used to capture localisation (MAR) and/or urbanisation (Jacobs) externalities. They argue that empirical studies' conclusions are not unanimous in indicating which kind of externalities prevail, but they underline how results are driven by many factors, such as the choice dependent and independent variables, the index used to measure externalities (size, share e diversity), the industry level considered (different digits can lead to different results), on the technology intensity of the sector, on the chosen geographical unit. In agreement with economic theory, there is greater evidence of MAR externalities for larger geographical units (such as countries), and greater evidence of Jacobian externalities when the geographical units considered are smaller.

Table 2.1 summarises main papers on agglomeration economics, their approach and results.

Table 2.1: Agglomeration Literature summary

Papers	Aggregation level	Variables	Results
GlaeserGlaser et al. (1992) "What When Space Matters Little for Firm Productivity? A multilevel analysis of localized knowledge externalities,"	the largest industries in 170 US cities	dependent variable: employement growth regressors: relative spacialization index , competitivity (number of firm per worker), diversity (a measure of a variety of industries in the city outside the industry in question. represented by the fraction of the city's employment the largest five industries other than the industry in question)	Employment growth is favored by local competition and urban variety, in agreement with Jacobian and Porterian externalities, but not by regional specialization, as expected is MAR-type externalites are at work.
Cianelli and Leoncini (1999) "Externalities and long term local industrial development. Some empirical evidence from Italy,"	Italian provinces	dependent variable: employement growth regressors: spacialization index, competitivity (number of firms per worker), diversity (Herfindahl index)	strong spatial effects of specialization and variety on knowledge spillovers. They also observe that patterns of externalities depend on the differences in market structure, endowment and factor accumulation.
De Lucio et al (2002) "The effects of externalities on productivity growth in Spanish industry,"	Spanish province 1978 to 1992 for 26 manufacturing branches	to discuss the role of externalities on productivity. External externalitities are measure by industrial specialization, region specialization, competion and diversity	They find evidence of dynamic effects due to specialization (MAR). Specialization affects productivity growth negatively, but once it reaches a certain level, its effect on growth becomes positive due to knowledge sharing. They do not find clear evidence on diversity and competition externalities.
Deckle (2002) "Industrial Concentration and Regional Growth: Evidence from the Prefectures,"	Japanese prefectures 3 sectors:manufacturing, services, trade (wholesale and retail)	dependent variable:TFP regressors: relative spacialization index, competitivity (number of firm per worker), diversity (Herfindahl index)	significant MAR externalities exist for the finance sector. Evidence of strong MAR, some Porter and no Jacobs externalities are identifies for the services sector and the wholesale and retail trade industries. No externalities for the manufacturing sector are detected.
Handerson (2003) "Marshall's scale economies,"	Plant level for machinery and high-tech industries. The dataset involved 742 counties in 317 metropolitan areas from 1972 to 1992.	estimation of production function allowing for localization and specialization externalities and considers these effects for both affiliate and not affiliate plants.	high-tech industries experience significant localization economies, while machinery industries do not. Externalities are strongly localized within the own county, and it seems that there are no spillovers from plants in other counties. Moreover, non-affiliate plants seem to generate greater externalities than corporate plants. Evidence of static urbanization-scale economies appears for corporate machinery plants.

Papers	Aggregation level	Variables	Results
Cingano e Schivardi (2004) "Identifying the Sources of Local Productivity Growth,"	Italian Local Labor Systems	Two step procusdere. In the first step they determine the value of TFP, then they proceed to aggregate productivity, which they regress in the second step together with employment growth on four main indicators - specialization, variety, city size and competition	TFP growth is enhanced by specialization and city size but not by urban diversity, the opposite result is reached for employment growth.
Rosenthal and Strange (2004) "Evidence on the Nature and Sources of Agglomeration Economies"	American establishment level data	They use the employment rate as a measure of productivity and new firms establishment. Tobit using as dependent variables the number of new establishments per square mile and the new-establishments' employment per square mile. Moreover, they aggregate employment into four concentric rings (distinguished by mile distance) and estimate the degree of dissipation of localization economies.	They reject the hypothesis that the industry fixed effects are jointly insignificant, confirming the presence of differences in local attributes across units. Competition within the same industry increases new births and the employment of new firms, but competition in other industries affects both negatively. Diversity of employment increases births and leads to greater employment. For most industries (five out of six), the effects of localization economies are more rapidly dispersed after the first few miles and less farther away.
Paci et al. (2006) "Agglomeration economies and growth-The case of Italian local labour systems, 1991-2001,"	784 Italian Local Labor Systems and 34 sectors at the (manufacturing and service sectors) from 1991 to 2001.	Employment dynamics in the service and manufacturing sectors of local labor systems among MAR (specialization index) and Jacobian externalities (inverse of Herfindhal index computed on sectoral employment), competition (Herfindhal index computed on employees distribution over plants).	They find a negative effect of specialization. Local employment seems to benefit from the presence of a diversified and competitive environment of small firms. Employment growth seems to benefit also from the performance of the surrounding areas.
Cainelli and Lupi (2009) "Does Spatial Proximity Matter? Micro-evidence from Italy"	24,089 Italian manufacturing firms for the period 1998-2001	They measure localization economies using the number of firms belonging to the same industry and located within a certain distance and urbanization economies Shannon's entropy index.	Their results suggest that localization effects are positive, but decreasing in distance. On the contrary, the diversity effects are negative for shorter distances and become positive for greater distances.
Martin et al (2008) "Spatial Concentration and Firm-Level Productivity in France,"	French firm level data from 1996 to 2004	The empirical approach of Martin et al. (2008)consists in introducing the spatial agglomeration effects as further explanatory variables of the production function.	Their results suggest that French firms benefit from localization economies, but not from urbanization economies or from competition effects. However, these benefits, although positive and significant, have a modest size and are geographically limited.

2.3 Methodology

We plan to empirically verify how dynamic externalities, represented by localisation and urbanisation economies, may affect firms' productivity. Dynamic externalities are associated to the improvement of firm performance thanks to the innovation adopted by another firm, through the knowledge spillover. In order to verify our idea, we are going to measure agglomeration and urbanisation economies, in particular we have focused on three kinds of externalities, proposed by: Marshall, Jacobs, and Porter. Localisation economies, and measured, respectively, through localisation economies, urbanisation economies and competition effects. We are going to proceed to briefly described the most popular indices that permit to capture these externalities.

2.3.1 Computing Spatial indices

Since we have extensively spoken about productivity that has been computed in the previous section, and that, now, it becomes the phenomenon we want to investigate the predictors of productivity.

Concepts introduced in order to capture spatial externalities are specialisation, localisation, urbanisation and diversity/agglomeration indeces. For each of these macro-categories there exist a large variety of indices.

These concepts are quite similar, and their different effects are not always easy to identify. For example, the term specialisation indicates the trend of economic activity, belonging to the same industry to localise in the same region. Localisation and agglomeration are closely related and both indicate that a region presents a high density of economic activities, regardless the economic sectors. Indices which measure specialisation, agglomeration and localisation, are built through the number of workers. These indices of spatial externalities should allow us to capture the advantages (in terms of productivity) that firms can enjoy localising close to other firms of the same industry (specialisation) or close firms belonging to the other industries (diversity). Working close each others encourage firms to establish long run relationships. Foe example, firms within the same sector could specialise in different stages of the productivity process and benefiting of these

[3] Firms in Sicily and Sardinia are excluded

relationships, which represent linkages between suppliers and customers. The greater opportunity to generate backward and forward linkages represent another reason that pushes firms to cluster.

Instead, cross-fertilisation, due to a large varieties of sectors, together with advantages arising from the demand side constitute the so called urbanisation economies. We have measured this second kind of economies by an index of population density, that captures the potential demand faced by firms of the region.

Brakman et al. (2009) distinguishes between agglomeration, concentration and specialisation as follows. Specialisation indicates the presence of only one sector in a given region. Concentration is used as a relative measure of region with respect to country.[4] Finally, agglomeration occurs when firms belonging to many sectors are localised in the same region. So Van Marrewijk definition of agglomeration resembles that of diversity.

In this work, we refer to the traditional definitions of external economies, in agreement with Marshall (1890), Jacobs (1969)and Porter (1990). The large variety adopted in agglomeration literature has led us to compare these indices. Below, we list the main indices used in the empirical literature. In particular, we are going to distinguish between indices that capture the so called Marshallian externalities (specialisation) from the indices that capture agglomeration/ diversity and, finally, indices capturing Porter's externalities (competition).

MAR externalities

The basic variables used to built these indices are the number of workers and the number of firms in a given area (region, county, province). The latter has been used only in few cases, for example in Cainelli and Lupi (2009) , they measure Marshallian externalities, associate to the concept of localisation, as $L_{is}^d = n_s^d$, where n_s^d denotes the number of firms belonging to the same industry s and localised inside a given distance d.

More often, MAR externalities are measured by specialisation indices , both relative and absolute. Among the others, Henderson (2003), Cainelli and Leoncini (1999) have adopted the following specialisation index:

$$S_j(t) = \sum_i \left[\frac{E_{ij}(t)}{E_j(t)} - \frac{E_i(t)}{E(t)} \right]^2$$

[4] In particular, we have concentration if economic activities are localised in one o few regions.

where $E_{ij}(t)$ represents the number of workers in the industry i and location (region/city) j , $E_j(t) = \sum_i E_{ij}(t)$ represents the total number of workers in location j, $E_i(t)$ is the country's total workers in the industry i, moreover $E(t) = \sum_i E_i(t)$ is the country's total workers.

$S_j(t)$ denotes the squares of the deviations of the share of workers in location j in the industry i with respect the share of workers in the industry i in the country. This relative measure indicates the degree of specialisation of a local area comparing the presence of a given sector i in the area j.

A similar index has been used by De Vor and De Groot (2010), who have measured specialisation through the localisation quotient, LQ:

$$LQ_{s,i} = \frac{E_{si} / \sum_{i=1}^{n} E_{s,i}}{\sum_{s=1}^{m} E_{s,i} / \sum_s \sum_i E_{si}}$$

where E_{si} indicates the number of employees in the industry s localised in i. This index differs from the one used by Henderson (2003)because the former takes ratios and the latter takes differences.

Cingano and Schivardi (2004) have , instead, computed an absolute geographical specialisation index, comparing workers in a city and industry with the total workers in the city; that is, this index measures the presence of industry s in the city,

$$spec_{c,s} = L_{c,s}/L_c.$$

Moreover, Duranton and Puga (2004) measure MAR externalities building an absolute and a relative specialisation index, based on the share of a given industry j in city i. Respectively, the absolute and the relative indices are:

$$ZI_i = \max_j(s_{ij}),$$

and

$$RZI_i = \max_j (s_{ij}/s_j)$$

these indices are obtained by dividing the share of each sector in local employment by its share in national employment. The related relative index allows taking into account that in a country there could exist one or few industries holding a larger share of national employment than others.

Furthermore, using firm level data, MAR externalities have been captured by **?**, through two alternative measures of localisation and specialisation (the first ones is an absolute and the other ones is a relative measure). Respectively;

$$ind\_loc_{it}^{sz} = ln(N_t^{sz} - N_{it}^{sz} + 1)$$

where N_t^{sz} represents the number of workers in the location z and industry s, N_{it}^{sz} denotes the number of workers in firm i.

From the point of view of the firm i, this index denotes the number of workers of the same industry and in the same location of firm i, excluding itself. The specialisation index is built as follow:

$$specialization_{it}^{sz} = ln\frac{(N_t^{sz} - N_{it}^{sz} + 1)}{(N_t^{z} - N_{it}^{sz} + 1)} = ln(N_t^{sz} - N_{it}^{sz} + 1) - ln(N_t^{z} - N_{it}^{sz} + 1)$$

where N_t^z is the number of workers in the location z. As a localisation index, for firm i, this index captures the number of other workers localised in z and working in the sector s compared with the total number of workers in z.

Jacobians'externalities

Jacobs has proposed a theory of cross fertilisation sustaining that firms benefit from localising in regions characterised by a large variety of sectors. Generally, this kind of externalities and the related diversity/variety indices are obtained through the Herfindahl index. Seldom, some entropy indices have been used . We briefly proceed to outline some of indices adopted in literature.

The Shannon's entropy index, used by Cainelli and Lupi (2009), formulate as:

$$H = -\sum_{s=1}^{S} p_s ln(p_s),$$

where $p_s = n_s/N$, $N = \sum_{s=1}^{S} n_s$ and n_s represent the number of firms in each industry $s \in [1, ..., S]$. This index assumes the maximum values when all sectors have the same number of firms.

Cainelli and Leoncini (1999)have adopted a Variety index, based on the Herfindahl diversity index, such as:

$$V_i^k = \sum_{j \neq i} s_{jk}^2$$

where s_{jk} represents the share of workers of every sector other than industry i.

An increase of the index V_i^k can be interpreted as an increase of the industry i concentration outside the local system and consequent decrease of the varieties in the local system.

Another variety index rather similar has been used by Cingano and Schivardi (2004), who adopt the following formulation:

$$variety_{c,s} = \sum_{j \neq s}^{J_c} \left(\frac{L_{c,s}}{L_c - L_{c,s}} \right)^2.$$

This index measures the variety that a certain sector s faces in the city c. The variety of sector s is defined as the square of the total share of workers in the other sectors (different from s), where the number of total workers in the city is obtained without considering the workers of sector s in city c. This index will assume values close to 1 if sector s is characterised by a small number of firms in the city.

The inverse of the variety index adopted by Cingano and Schivardi (2004) have been used by Martin et al. (2008), in order to capture Jacobians 'externalities. Their formulation is the following

$$ind\_div_t^{sz} = ln \left(\frac{1}{H_t^{sz}} \right)$$

where $H_t^{sz} = \sum_{s \neq j} \left(\frac{N_{jt}^z}{N_t^z - N_t^{sz}} \right)^2$, represents the Herfindahl diversity index, N_{jt}^z indicates workers of industry j in location z, N_t^z denotes workers in location z. This formulation allows an easier interpretation of the results, since an increase of this index is associated an increase in diversity.

De Vor and De Groot (2010) have measured the diversity using a Relative Diversity Index (RDI),

$$RDI_i = \frac{1}{\sum_s \mid \frac{E_{si}}{\sum_{i=1}^{n} E_{s,i}} - \frac{\sum_{s=1}^{m} E_{s,i}}{\sum_s \sum_i E_{si}} \mid}.$$

This index, contrary to the previous indices of diversity, varies only across local areas (cities), and it can be associated to the specialisation index used by Henderson. In particular, it looks like the inverse of the specialisation index.

Another diversity index, seldom used is the so called Dixit -Stiglitz index adopted by Wheeler (2003)

$$DS_i = \left[\sum_s \left(\frac{E_{si}}{E_i} \right)^{1/2} \right]^2$$

where E_{si} is the number of employees of sector s in location i. This index will assume greater values, the greater the variety in the location i.

? have adopted a common measure of diversity represented by the inverse of a Hirshman-Herfindahl index, summing for each city over all sectors the square of each sector's share in local employment. Formally, this index can be written as follows:

$$DI_i = 1/\sum_j s_{ij}^2$$

This index assumes a value of one when economic activities in city i are concentrated in a sector. The value assumed by the index increases as activities in this city become more varied. Duranton and Puga, as for the specialisation index, suggest that an absolute index needs to be correct for differences in sectoral employment shares at the national level, and they have proposed a relative measure obtained summing for each city, the absolute value of the difference between local employment share of a given sector and the relative share in national employment. Formally:

$$RDI_i = 1/\sum_j \mid s_{ij} - s_j \mid$$

Fu and Junjie (2010) have used as measure of urban diversity, represented by the complement to the inverse of the Hefindahl index, so that the greater the value of the index, the greater the local diversity:

$$Urban\_diversity_k = 1 - \sum_{m=1}^{M} \left(\frac{E_{mk}}{\sum_m E_{mk}} \right)^2$$

where, now, the subscripts used are k for the geographical area and m for the industry.

Urbanisation indices

In some papers, Jacobs externalities are measured through urbanisation indices. In others cases urbanisation is considered as a phenomenon associated to the demand, instead of cross-fertilisation, considering workers as consumers.

Among the urbanisation indices used in previous works, we can mention(?) the index adopted by Martin et al. (2008):

$$ind\_urb_t^{sz} = ln(N_t^z - N_t^{sz} + 1)$$

and urbanisation index that considers potential consumers demand and can be represented by the density index used by Cainelli and Leoncini (1999) and computed dividing provincial (or regional) population by the area expressed in squares kilometres :

$$Density_{it} = \frac{Population_{it}}{Area_i}$$

Competition indices

Another important kind of externality, that often is neglected, focuses on competition, that, as stressed by Porter, competition pushes firms to innovate, in order to improve their performance and so their productivity.

Measures adopted to capture competition effects are an index à -la- Herfindahl indicating the concentration of a sector in a given geographical area, such as the one used byCingano and Schivardi (2004):

$$comp_{c,s} = \sum_{j \in c,s} \left(L_{c,s,j}/L_{c,s}\right)^2$$

where j denotes the firm. This index indicates the number of workers in the same sector of firm j locate in the city c. This index measures the distribution of the employment share at firm level within the city-industry set. Small values of this index mean that there is a uneven distribution of workers across existing firms. The inverse of this index has been used to measure of competitiveness by Martin et al. (2008),where Porterian externalities have been computed through the following index:

$$ind\_comp_t^{sz} = ln\left(\frac{1}{Herf_t^{sz}}\right),$$

where $Herf_t^{sz} = \sum_{i \in S_t^z} \left(\frac{N_{it}^{sz}}{N_t^{sz}}\right)^2$ represents a traditional Herfindahl index that has been inverted in order to interpret the increase of the index as an increase in competitiveness.

An alternative index is built by comparing per-capite number of firms in a city-industry with respect the per-capite number of firms in the geographical area i.

$$COM_{si} = \frac{F_{si}/E_{si}}{\sum_s F_{si}/\sum_s E_{si}}.$$

This index, used by **?**and Glaeser et al. (1992), returns values greater than 1 when one industry has more plants/firms in that city relative to its size in the country; values grater than one indicates that the industry in that city is locally more competitive.

Agglomeration indices

Some works focus on the role of agglomeration in a certain area, without pay attention to the kind of externalities generated, but focusing only on how economic activities are distributed across space. Looking at the specialisation and diversity indices computed for a given geographical area, we can observe that these two phenomena are one the inverse of the other. Thus we can analyse one effect and obtain relevant also for the other.

If agglomeration is considered as opposite of specialisation, it can be measured like urbanisation.

Among the agglomeration indices, we can recall the *entropy index* and the *Krugman index* employed by Brulharta and Sbergami (2009). Formally we can write the indices as follows:

$$Entropy_i = \sum_j \left[E_{ij} / \sum_j E_{ij} * ln \left(E_{ij} / \sum_j E_{ij} \right) \right] / ln(k)$$

where k denotes the number of provinces (or, more generally, the geographical area where firms are localised), i indicates the industry and j indicates the geographical area. This index assumes values in the interval $(-1, 0)$ and it will be equal to zero when there is a perfect concentration or close to $-1$with dispersion. Krugman index of concentration is,instead :

$$Krugman_i = \sum_j \mid E_{ij} / \sum_j E_{ij} - \sum_{z \neq i} E_{ij} / \sum_j \sum_{z \neq i} E_{ij} \mid .$$

This index takes values between 0 and 2, with values close to zero indicating that a certain sector is homogeneously distributed across regions and values close to 2 when there is perfect concentration.

An innovative index of agglomeration has been proposed by Ellison and Glaeser (1997). This is an index of geographic concentration and of co-agglomeration allowing for the comparison of geographical concentration among areas of different size, by controlling for differences in the size distribution of plants and for differences in the size of the geographic areas for which data are available. Starting from the definition of the Gini geographical concentration index $G = \sum_i (s_i - x_i)^2$, where x_i indicates the share of employment in the industry in the area i, and s_i is the share of employment in the area i , the index of Industry concentration proposed by Ellison and Glaeser is the following:

$$\gamma = \frac{G - (1 - \sum_i x_i^2)H}{(1 - \sum_i x_i^2)(1 - H)} = \frac{\sum_i^M (s_i - x_i)^2 - (1 - \sum_i^M x_i^2) \sum_j^N z_j^2}{(1 - \sum_i x_i^2)(1 - \sum_j^N z_j^2)},$$

where $H = \sum_j^N z_j$ is the Herfindahl index of the industry plant size distribution. Further, using the original notation, we can indicate by M the total number of geographical areas and by N the total number of economic sectors. This can take values belonging to the range $(-1\ 1)$. When the index is zero, there is complete lack of agglomeration forces. [5]

2.3.2 Empirical specification

Choosing among the large set of indices used in literature, to capture each kind of spatial externalities, without overlapping their effects. represents a rather tricky task

The choice of the indices to introduce in our specification has been driven by the following criteria: first we introduce one index for each kind of externality (MAR, Jacobs, Porter); second, we compare indices through a correlation matrix and we introduce those indices that are characterised by lowest correlation, in order to avoid, or at least minimise, the overlapping of effects.

Another important step is represented by the decision of the geographical area to analyse. The geographical units that we choose is represented specifically provinces area. This choice is characterised by the data availability and the size of the area, small enough to allow to analyse human relationships that can arise inside them and large enough to consider this geographical area as a natural market for firms.

Following this principle, and partially departing from the standard procedure, we have built indices that vary both across sectors and across provinces. Then, we have computed a correlation matrix of indices and compared alternative specification through a collinearity test, to avoid entering redundant information.

Defined and built the indices, we next choose one index for each kind of externality, in particular, we think that a suitable criterion could be to choose indices characterised by low correlations, where low correlation is represented by a level of correlation below the cut off of 0.75.

[5] For example, values of this index computed using our dataset yields a smaller range , $(-0.23, 0.26)$, indicating that, industries are dispersed across provinces and that only few provinces exhibits high agglomeration (concentration). Except Prato, the other provinces with a high value of γ are provinces characterised by a low number of firms and few sectors.

However, this index captures the level of agglomeration, but it could be insufficient to evaluate the degree of specialisation and urbanisation in a province. Indeed they could both work and any improvement of productivity could be due to the combination of externalities.

We present a correlation matrix, and we compare three groups of indices. The first that captures Marshallian Externalities, the second group captures Jacobian externalities and finally, the smaller one captures Porter externalities.

Indices belonging to the same group should be more correlated. Instead, indices belonging to different groups should be less correlated in order to capture different aspects.

We can observe that indices that capture MAR extenalities have low correlation to the indices that capture Jacobian externalities and/or Porterian exnalitites.

The only exceptions are represented by *ind_com* (the competitivity index adopetd by Martin et al. (2008)), that is highly correlated with *ind_loc*, and *ind_urb* (respectively indicating localization and urbanization indices adopted by Martin et al. (2008)) and *log_density*, even if only the correlation with *ind_loc* is too high and it should be investigated for collinearity.

We can see that the entropy and Krugman index are, as expected, highly and inversely correlated with the diversity index, since both these indices measure the degree of agglomeration and can be considered one the inverse of the diversity indices. It is worth noting that often, in literature, researchers, who have analysed the Jacobian externalities through diversity indices, have considered the diversity of a region as the opposite of its degree of specialisation. This reciprocal relationship arises when indices are computed only accounting for the geographical level. Both Entropy and Krugman indices are examples of this relationship, as we can observe in table 2.4.Their correlation is highly negative and close to one.

Observing table 2.4 and so considering the correlation among indecs, we can define two alternative specifications: one based on absolute indices and one on relative indices. The two alternative specifications are, respectively:

$$TFP_{it} = f\left(spec\_abs,\, log\_diversity,\, ind\_com,\, density\right)_{i \in p,t}$$

$$TFP_{it} = f\left(spec\_rel,\, RDI,\, COMP,\, density\right)_{i \in p,t}$$

After choosing the indicators for our specification, we perform a further test of collinearity:

Table 2.4: indices correlation matrix

	spec_abs	spec_rel	LQ	ind_loc	ind_urb	density	diversity	RDI	DS	Urban_var	DI	Entropy	Krugman	COMP	ind_com
spec_abs	1														
spec_rel	0.6427	1													
LQ	0.6796	0.5099	1												
ind_loc	0.4920	0.2115	0.2031	1											
ind_urb	-0.2418	-0.2729	-0.2954	0.6381	1										
density	-0.0727	-0.0934	-0.0850	0.5163	0.7320	1									
diversity	0.0091	-0.0251	-0.0528	0.3317	0.3809	0.1450	1								
RDI	-0.1800	-0.2963	-0.2851	0.2478	0.4612	0.1347	0.5004	1							
DS	-0.2608	-0.3005	-0.3200	0.5125	0.8344	0.4763	0.6766	0.5136	1						
Urban_variety	-0.2462	-0.2845	-0.2871	0.1961	0.4695	0.1674	0.6281	0.4978	0.7562	1					
DI	-0.2639	-0.3105	-0.3020	0.3109	0.6079	0.2912	0.7780	0.6031	0.8687	0.8962	1				
Entropy	0.2669	0.3096	0.3175	-0.4494	-0.7726	-0.4297	-0.7357	-0.5874	-0.9683	-0.8049	-0.9494	1			
Krugman	0.1590	0.2833	0.2625	-0.1581	-0.3277	-0.0052	-0.4587	-0.9700	-0.4178	-0.4651	-0.5435	0.4985	1		
COMP	-0.2511	-0.1420	-0.3867	-0.1062	-0.0599	-0.0435	-0.0281	-0.0366	-0.0608	-0.0205	-0.0358	0.0532	0.0319	1	
ind_com	0.4733	0.2304	0.1460	0.8916	0.6586	0.6013	0.3207	0.2229	0.4845	0.1627	0.2805	-0.4260	-0.1204	0.0958	1

Table 2.5: VIF specification 1

Variable	VIF	SQRT VIF	Tolerance	R-Squared
ln_spec_abs	1.84	1.36	0.5438	0.4562
log_diversity	1.19	1.09	0.8438	0.1562
log_density	2.18	1.48	0.4591	0.5409
ind_com	3.14	1.77	0.3187	0.6813

Table 2.6: VIF specification 2

Variable	VIF	SQRT VIF	Tolerance	R-Squared
ln_spec_rel	1.13	1.06	0.8853	0.1147
log_RDI	1.12	1.06	0.8953	0.1047
log_density	1.02	1.01	0.9764	0.0236
log_COMP	1.03	1.01	0.9708	0.0292

the Variance Inflation Factor (VIF)[6] test, which measures the impact of collinearity among the variables in the regression. The VIF assumes values greater or equal to 1, there is no an official value for determine collinearity, but it is common practise to consider values greater than 10 as a signal of multicollinearity. Testing our specifications we find here results reported in tables 2.5 and 2.6, respectively for the first and the second specification. There results suggest that the specification chosen are not affected by collinearity problem, since the value of VIF statistic is always less then the critical value of 10.

2.4 Results

We start showing the results obtained running a plain OLS estimator and then we introduce some control variables such as geographical fixed effects and industry fixed effects. Further, we account for efficiency of OLS estimator assuming that may exist provincial clusters or industry clusters. So we assume that the variance matrix is characterised by a special pattern that could affect its structure.[7]. We present in table 2.7 results related to the relationship between firm level productivity and province/industry spatial externalities indices. We found that specialisation

[6]VIF indicates a measure of tolerance for multicollinearity. Formally $tolernce = 1 - R_i^2$ and $VIF = \frac{1}{tolerance}$, where R indicates how the regressor i explains in the regression.

[7]Variance-Covariance matrix, could be$V = \begin{matrix} \sigma_{1p} & 0 & 0 & 0 & 0 \\ 0 & \sigma_{2p} & 0 & 0 & 0 \\ 0 & 0 & \ddots & 0 & 0 \\ 0 & 0 & 0 & \sigma_{n-1,p} & 0 \\ 0 & 0 & 0 & 0 & \sigma_{np} \end{matrix}$

does not have a positive effect on firms productivity, and positive evidence for MAR economies is obtained controlling for first nature advantages, captured through province fixed effects. Our results do not support Marshallian theories, since the specialisation index has generally a negative sign and loses its significance when we assume an industry clusters structure in standard errors. The relationship between productivity and diversity is positive and significant when provincial fixed effects are introduced, but the diversity coefficient loses its significance assuming clusters in the variance covariance matrix. These results imply that there is not strong evidence for Jacobian externalities. Mixed results are obtained even for competitivity effects on firms productivity; these are negative when provincial fixed effects are included, positive otherwise. Finally, the relation between density and productivity is significant and positive, indicating that the greater the number of consumers in the province, the higher the productivity, because the higher demand. Density is not significant when provincial fixed effects are introduced; this effect could be associated to the fact the both variables vary across provinces, so they could capture similar effects.

Table 2.8 shows the results obtained estimating the "relative" specifications. We found that the relative specialisation is positive and significant when no clusters are assumed to estimate the variance-covariance matrix. We obtain positive and significant signs for relative diversity index and density index, supporting Jacobian theories that urbanisation and diversity causes cross - fertilisation and improvements in productivity. The coefficient of competitivity is negative and significant; the only exception is represented by the case when industry fixed effects are introduced in the model specification. In this case, the coefficient becomes positive, revealing the important role of differences among sectors for understanding competitivity effects.

Our results have shown two main insights; first, theories proposed by Marshall, Jacobs and Porter hold also considering the micro firm-level of productivity. Second, there is not unanimous evidence in favour of these effects. Indeed, using different indices to measure the spatial externalities and introducing different assumption for cluster and/or fixed effects, lead to different conclusions.

Finally, we want to implement the standard approach followed in the agglomeration literature, based on estimating the relation between aggregated productivity and spatial externalities. This further step in our analysis aims to compare results obtained following these two different approaches. Our approach that analyses the relation between variables with a different level of aggregation and the approach that relates the same variables at the same level of aggregation (province/industry level of aggregation). We employ provide two types of aggregate productivity:

one is a weighted productivity, obtained from aggregating firm-level productivity, obtained using LP approach. The other is an unweighted productivity, obtained as a residual from aggregate production function estimation.

Using province/industry productivity we may lead to simultaneity issue because of the uncertain direction of causality between productivity and spatial externalities (computed to the same level of aggregation). Indeed, productivity could increase because of externalities, and at the same time, an increase in productivity could enhance spatial externalities. This issue is overcome at micro level analysis since it is assumed that firms are small economic agents, that their decisions cannot affect provincial market, but the external forces can affect firms' performances. In the aggregate production function estimation, endogeneity has been accounted for through Instrumental Variables; in particular, since we do not have proper external instruments (excluded instruments), we resort to internal instruments, represented by the time lag of regressors (included endogenous variables).

Table 2.9 shows the results obtained by running the regressions of aggregate productivity (adopting the OP procedure) on absolute indices. We find that specialisation indices has a negative sign and the associated coefficient is statistically significant, except when province fixed effects are introduced. There is no evidence for Jacobian externalities, since the diversity index is always insignificant, so that province-level productivity does not seem to be affected by the diversity of the area. We find robust results for competitiveness, supporting Porter's idea that competition stimulates innovation and improvements in productivity. Furthermore, there is no evidence for a density effect, since its coefficient is not statistically significant except when industry level fixed effects are introduced, when density becomes statistically significant, but negative.

Table 2.10 reports results obtained estimating the second specification with relative indices. We find that relative specialisation is positive and statistically significant when we depurate the model from province fixed effects (in column c, g, h), in the other cases it is not significant. Moreover this specification does not show evidence in favour of competitiveness effect, but we find strong evidence in favour of Jacobi an effects, as the diversity index is positively signed and statistically significant.

Results related to the density index are ambiguous, the relative coefficients are significant but they take both positive and negative sign, in particular we find a negative sign when province fixed effects are introduced (columns c and g), and positive otherwise (columns a, b, d, e, f and h).

Comparing estimates obtained for the weighted productivity with estimates obtained by unweighted productivity, we can observe that results are rather different, leading us to stress the importance of the choice of aggregation method and the relative weights.

Results in tables 2.11 and 2.12 suggest that unweighted province/industry productivity is affected by externalities to a greater extent compared to the weighted productivity, and coefficients are generally greater in magnitude and more statistically significant.

Concentrating on the first specification, with absolute indices, we can see that, when we consider weighted productivity, specialisation results negative and significant when province fixed effect are introduced. The opposite is true for unweighted productivity, so that there is not a clear effect of MAR externalities on province/industry productivity. Looking at the competition coefficient we find a positive and significant coefficient in both cases, but using unweighted productivity we have greater magnitudes of the coefficients. These results gives evidence in support of Porterian theories.

Running a regression of weighted productivity on the density index, we find a negative and statistically significant coefficient when we account for first nature advantages. When we perform the regression for unweighted productivity, excluding province fixed effects, the coefficient is positive and significant.

Finally, the diversity index, is never significant for weighted province/industry productivity, while it is negative and significant considered the unweighted productivity, and only if we include first nature advantages.

Comparing results obtained estimating aggregate productivity on relative indices of spatial externalities, we find that specialisation is always positive and significant for unweighted productivity. Instead, considering weighted productivity specification results positive and significant only when geographical area effects are considered by dummy variables.

Competitivity has a poor effect on weighted productivity. Instead, it results positive and significant for unweighted productivity, except for the case when industry fixed effects are introduced in the model.

The same results are obtained for density, positive effects are found except when first nature advantages are introduced in the model.

Finally, relative diversity results positive and significant in both cases, but coefficients magnitudes are greater for the unweighted province/industry productivity. We can, further, observe that estimating through instrumental variables gives better data fitting.

2.5 Conclusions

In this chapter, we have analysed how space, through spatial externalities, affects firms productivity. We have investigated how firms' performance is affected by local environment where firms are localised. Indeed, the choice of firms to locate close to other firms of the same industry (specialisation) or other industries (diversity /urbanisation) gives rise to advantages, such as facilitating innovation, increasing learning opportunities, facilitating trade.

Concentrating production in few places allows, especially small firms to benefit from external scale economies, since the small size does not permit to achieve internal ones.

We have proceeded estimating firm level productivity and computing a large set of indices able to capture these mentioned advantages. In particular, we have referred to MAR, Jacobs and Porter externalities. These indices have been compared in order to choice a specification without collinearity.

We have implemented OLS and IVs estimators to two main equations and to the sub-specifications represented by the introduction of information about geography and/or economic sector. Also, we compare both results obtained using robust methods. The main results found are synthesised in the tables 2.13, 2.14, and 2.15. As we can observe, the results obtained are heterogeneous and depend on the particular specifications adopted. For example, including province fixed effects, or, considering variance-covariance correction method implies different. results.

Results show that MAR externalities affects positively firms' productivity only when province fixed effects are introduced in the specification, when we adopt absolute indices. Instead, adopting relative indices we observe a positive effect on productivity, when no adjustment for variance covariance matrix is made, indicating that an increase in specialisation of a given province with respect the national economy improves the productivity of firms localised in that area. We can observe a marked evidence of positive effects of diversification externalities in the second specification. Thus, increasing the variety of goods produced in a province with respect to the national diversity, determines an increase in firm-level productivity.

Considering absolute indices, we find a greater evidence for Porter externalities, which seem almost absent in the second specification. Finally, density, when significant, has a positive effect on productivity. Its significance can be linked to the introduction of province-level fixed effects.

At the province level, we observe the opposite results between the two specifications, but it is worth noting that when aggregated province productivity is used as dependent variable and a relative specification is adopted, we obtain a stronger evidence for MAR, Jacobs, urbanisation externalities, and only mixed evidence of competition effects.

Table 2.7: Firm level productivity - first specification

	TFP_{it}^{LP}							
	(a)	(b)	(c)	(d)	(e)	(f)	(g)	(h)
$SPEC\_ABS_{pt}^{s}$	-0.0218***	-0.0218*	0.0367***	0.0366***	-0.034***	-0.022	-0.034***	-0.027**
	(0.00305)	(0.0116)	(0.00422)	(0.01343)	(0.0031)	(0.02122)	(0.0084)	(0.01131)
$IND\_COM_{pt}^{s}$	0.0344***	0.0344***	-0.0408***	-0.0408***	0.0549***	0.0344	0.0549***	.0356***
	(0.00297)	(0.0120)	(0.00458)	(0.0122)	(0.0031)	(0.0210)	(0.009)	(0.01324)
$DIVERSITY_{pt}^{s}$	0.00946***	0.00946	0.0230***	0.023	-0.0066*	0.0095	-0.0066	0.0169
	(0.0034)	(0.00981)	(0.00653)	(0.0171)	(0.0034)	(0.0179)	(0.0086)	(0.019)
$DENSITY_{pt}^{s}$	0.0430***	0.0430***	0.2341**	0.2341	0.0112***	0.043**	0.0112	0.1451
	(0.0030)	(0.0155)	(0.10594)	(0.2857)	(0.0031)	(0.0185)	(0.0083)	(0.2797)
$constant$	9.5007***	9.501***	8.648***	8.1828***	9.566***	9.5***	9.566***	8.556***
	(0.01934)	(0.0740)	(0.6247)	(1.426)	(0.0206)	(0.0941)	(0.0445)	(1.412)
$province\ FE$	no	no	yes	yes	no	no	no	yes
$cluster\_province$	no	yes	no	yes	no	no	no	yes
$ateco\_FE$	no	no	no	no	yes	no	yes	yes
$cluster\_ateco$	no	no	no	no	no	yes	yes	no
R^2	0.0231	0.0231	0.0393	0.0393	0.0654	0.0231	0.0659	0.0738

* significant at 10%; ** significant at 5%; *** significant at 1%

Table 2.8: Firm level productivity - second specification

	TFP_{it}^{LP}							
	(a)	(b)	(c)	(d)	(e)	(f)	(g)	(h)
$SPEC\_REL_{pt}^{s}$	0.0048***	0.0048	0.00589***	0.0059	0.00772***	0.0048339	0.00772	0.00899
	(0.00171)	(0.00976)	(0.00168)	(0.00959)	(0.0017)	(0.0093)	(0.00508)	(0.0057)
$COMP\_REL_{pt}^{s}$	-0.03220***	-0.03220***	-0.02523***	-0.0252**	0.0508***	-0.0322	0.0508***	0.06326***
	(0.00478)	(0.00984)	(0.00424)	(0.01023)	(0.0076)	(0.0486)	(0.0162)	(0.01516)
RDI_{pt}^{s}	0.09428***	0.09428**	-1.411	0.6229***	0.0628***	0.0943***	0.0628***	0.5174***
	(0.00941)	(0.03976)	(1.15e+09)	(0.081)	(0.0096)	(0.0271)	(0.01923)	(0.0728)
$DENSITY_{pt}^{s}$	0.0681***	0.0681***	0.1755*	0.1755	0.0536***	0.0681***	0.0536***	0.229
	(0.00207)	(0.0124)	(0.10586)	(0.2931)	(0.0021)	(0.0097)	(0.0082)	(0.2903)
$constant$	9.6835***	9.6835***	7.0251	9.5763	9.602*** (0.0231)	9.683***	9.6016***	9.101***
	(0.02097)	(0.07584)	(1.55e+09)	(1.320)		(0.08594)	(0.0394)	(1.3377)
$province\ FE$	no	no	yes	yes	no	no	no	yes
$cluster\_province$	no	yes	no	yes	no	no	no	yes
$ateco\ FE$	no	no	no	no	yes	no	yes	yes
$cluster\_ateco$	no	no	no	no	no	yes	yes	no
R^2	0.0225	0.0225	0.0387	0.0659	0.0625	0.0225	0.0625	0.0747
* significant at 10%; ** significant at 5%; *** significant at 1%								

Table 2.9: Weighted province/industry level productivity - specification 1

	TFP_{pst}^{LP}							
	(a)	(b)	(c)	(d)	(e)	(f)	(g)	(h)
$SPEC\_ABS_{pt}^{s}$	-0.0470***	-0.0311***	0.02115	-0.0471***	-0.0311**	-0.0261***	-0.000984	-0.00098
	(0.00887)	(0.0090)	(0.01401)	(0.0107)	(0.01164)	(0.00995)	(0.01551)	(0.0158)
$IND\_COM_{pt}^{s}$	0.0942***	0.08493***	0.0176	0.0942***	0.0849***	0.11895***	0.08055***	0.0806***
	(0.01037)	(0.01062)	(0.0158)	(0.01825)	(0.0182)	(0.00984)	(0.0205)	(0.02178)
$DIVERSITY_{pt}^{s}$	0.00088	0.0052	-0.0291	0.0009	0.0052	-0.0135	-0.05282*	-0.0528*
	(0.0120)	(0.01267)	(0.0263)	(0.0180)	(0.0181)	(0.01123)	(0.02792)	(0.02786)
$DENSITY_{pt}^{s}$	0.0131	0.0203	-0.6146***	0.0131	0.02026	0.00211	-0.64083***	-0.64083**
	(0.0128)	(0.0134)	(0.22471)	(0.01347)	(0.01375)	(0.00954)	(0.2272)	(0.3074)
constant	9.665***	9.672***	12.406***	9.665***	9.672***	9.749***	12.559***	12.559***
	(0.0798)	(0.0832)	(1.127)	(0.1145)	(0.11565)	(0.0643)	(1.138)	(1.4883)
estimator	OLS	IV	IV	OLS	IV	IV	IV	IV
$cluster\_province$	yes	yes	yes	no	no	no	yes	no
$province\_FE$	no	no	yes	no	no	no	yes	yes
$cluster\_ateco$	no	no	no	yes	yes	yes	no	yes
$ateco\_FE$	no	no	no	no	no	yes	yes	yes
R^2	0.0563	0.0544	0.1093	0.0563	0.0544	0.1482	0.1879	0.1879
* significant at 10%; ** significant at 5%; *** significant at 1%								

Table 2.10: Weighted province/industry level productivity - specification 2

	TFP_{pst}^{LP}							
	(a)	(b)	(c)	(d)	(e)	(f)	(g)	(h)
$SPEC\_REL_{pt}^{s}$	-0.00693	-0.00375	0.0133**	-0.00693	-0.0037	0.00235	0.0176***	0.0176***
	(0.00549)	(0.0056)	(.00542)	(0.0068)	(0.007921)	(0.0053)	(0.00609)	(0.0053)
$COMP\_REL_{pt}^{s}$	0.02441*	-0.013	-0.00123**	0.0244	-0.0131	0.0090	0.011	0.011
	(0.01424)	(0.01502)	(0.0149)	(0.0264)	(0.0266)	(0.0194)	(0.0172)	(0.017)
RDI_{pt}^{s}	0.1177**	0.1093**	0.7817***	0.1177***	0.10933***	0.1049***	0.7438***	0.7438**
	(0.0533)	(0.053017)	(0.0427)	(0.0396)	(0.04023)	(0.036)	(0.0419)	(0.354)
$DENSITY_{pt}^{s}$	0.0611***	0.0596***	-0.585**	0.0611***	0.0596***	0.05834***	-0.5723*	-0.5723*
	(0.0189)	(0.0187)	(0.227)	(0.00994)	(0.01013)	(0.00962)	(0.2267)	(0.3091)
$constant$	9.7919***	9.805***	13.661***	9.792***	9.805***	9.773***	13.536***	13.536***
	(0.1281)	(0.1265)	(1.072)	(0.0757)	(0.08056)	(0.0702)	(1.078)	(1.622)
$estimator$	OLS	IV	IV	OLS	IV	IV	IV	IV
$cluster\_province$	yes	yes	yes	no	no	no	yes	no
$province\_FE$	no	no	yes	no	no	no	yes	yes
$cluster\_ateco$	no	no	no	yes	yes	yes	no	yes
$ateco\_FE$	no	no	no	no	no	yes	yes	yes
R^2	0.0251	0.02020	0.1059	0.0251	0.0202	0.1018	0.1762	0.1762

* significant at 10%; ** significant at 5%; *** significant at 1%

Table 2.11: Aggregated province/industry level productivity - specification 1

	TFP^{LP}_{pst}							
	(a)	(b)	(c)	(d)	(e)	(f)	(g)	(h)
$SPEC\_ABS^s_{pt}$	-0.0062	0.00277	0.1025***	-0.0062	0.0028	0.0118	0.1203***	0.1203***
	(0.0125)	(0.0122)	(0.0192)	(0.01229)	(0.0118)	(0.0109)	(0.0219)	(0.0191)
$IND\_COM^s_{pt}$	1.156***	1.152***	1.027***	1.156***	1.152***	1.1786***	1.019***	1.019***
	(0.01442)	(0.0144)	(0.0217)	(0.0232)	(0.0232)	(0.01515)	(0.0300)	(0.02836)
$DIVERSITY^s_{pt}$	0.0294	0.0191	-0.0959***	0.02935	0.0191	0.0061	-0.0893**	-0.0893**
	(0.0184)	(0.0176)	(0.0306)	(0.0244)	(0.02146)	(0.0149)	(0.03556)	(0.03404)
$DENSITY^s_{pt}$	0.04239**	0.0465**	-1.861***	0.0424**	0.0465***	0.03312**	-1.872***	-1.872***
	(0.0207)	(0.02007)	(0.3803)	(0.0158)	(0.0160)	(0.0142)	(0.3803)	(0.4717)
$constant$	9.5823***	9.6332***	18.71***	9.5823*** (0.1285)	9.633***	9.7097***	18.767***	18.767***
	(0.124)	(0.1169)	(1.906)		(0.1221)	(0.08657)	(1.902)	(2.306)
$estimator$	OLS	IV	IV	OLS	IV	IV	IV	IV
$cluster\_province$	yes	yes	yes	no	no	no	yes	no
$province\_FE$	no	no	yes	no	no	no	yes	yes
$cluster\_ateco$	no	no	no	yes	yes	yes	no	yes
$ateco\_FE$	no	no	no	no	no	yes	yes	yes
R^2	0.7884	0.8021	0.8145	0.7884	0.8020	0.8125	0.8229	0.8229
* significant at 10%; ** significant at 5%; *** significant at 1%								

Table 2.12: Aggregated province/industry level productivity - specification 2

	TFP_{pst}^{LP}							
	(a)	(b)	(c)	(d)	(e)	(f)	(g)	(h)
$SPEC\_REL_{pt}^{s}$	0.2103***	0.2417***	0.36719***	0.2103***	0.2417***	0.11658***	0.2458***	0.2458***
	(0.0227)	(0.02557)	(0.0216)	(0.02824)	(0.0314)	(0.02765)	(0.01743)	(0.0273)
$COMP\_REL_{pt}^{s}$	0.2041***	0.19244***	0.28553***	0.2041***	0.1924*	-0.08695*	-0.06464	-0.06464
	(0.0425)	(0.0453)	(0.04189)	(0.0932)	(0.10253)	(0.0502)	(0.0432)	(0.0536)
RDI_{pt}^{s}	1.656***	1.677***	3.229***	1.656***	1.67701***	1.6036***	3.697***	3.697***
	(0.2923)	(0.2944)	(0.12232)	(0.200)	(0.20244)	(0.17789)	(0.12422)	(0.3407)
$DENSITY_{pt}^{s}$	0.5155***	0.52381***	-0.611	0.5155***	0.5238***	0.5580***	-0.95136***	-0.95136***
	(0.1134)	(0.113636)	(0.3766)	(0.0344)	(0.034669)	(0.0357)	(0.35621)	(0.5155)
$constant$	11.412***	11.527***	18.497***	11.412***	11.5266***	11.11***	20.397***	20.397***
	(0.6934)	(0.6959)	(1.740)	(0.3782)	(0.38366)	(0.2736)	(1.6552)	(1.6552)
$estimator$	OLS	IV	IV	OLS	IV	IV	IV	IV
$cluster\_province$	yes	yes	yes	no	no	no	yes	no
$province\_FE$	no	no	yes	no	no	no	yes	yes
$cluster\_ateco$	no	no	no	yes	yes	yes	no	yes
$ateco\_FE$	no	no	no	no	no	yes	yes	yes
R^2	0.1911	0.1930	0.3891	0.1911	0.1930	0.3199	0.54	0.54

* significant at 10%; ** significant at 5%; *** significant at 1%

Table 2.13: firm level results

specification1	MAR	Jacobs	Porter	density	specification 2	MAR	Jacobs	Porter	density
(a)	-	+	+	+	(a)	+	+	-	+
(b)	ns	ns	+	+	(b)	ns	+	-	+
(c)	+	+	-	+	(c)	+	ns	-	ns
(d)	+	ns	-	ns	(d)	ns	+	-	ns
(e)	-	ns	+	+	(e)	+	+	+	+
(f)	ns	ns	ns	+	(f)	ns	+	ns	+
(g)	-	ns	+	ns	(g)	ns	+	+	+
(h)	-	ns	+	ns	(h)	ns	+	+	ns

Table 2.14: weighted province results

specification 1	MAR	Jacobs	Porter	density	specification 2	MAR	Jacobs	Porter	density
(a)	-	ns	+	ns	(a)	ns	+	ns	+
(b)	-	ns	+	ns	(b)	ns	+	ns	+
(c)	ns	ns	ns	-	(c)	+	+	-	-
(d)	-	ns	+	ns	(d)	ns	+	ns	+
(e)	-	ns	+	ns	(e)	ns	+	ns	+
(f)	-	ns	+	ns	(f)	ns	+	ns	+
(g)	ns	ns	+	-	(g)	+	+	ns	ns
(h)	ns	ns	+	-	(h)	+	+	ns	ns

Table 2.15: unweighted province results

specification 1	MAR	Jacobs	Porter	density	specifition 2	MAR	Jacobs	Porter	density
(a)	ns	ns	+	+	(a)	+	+	+	+
(b)	ns	ns	+	+	(b)	+	+	+	+
(c)	+	-	+	-	(c)	+	+	+	ns
(d)	ns	ns	+	+	(d)	+	+	+	+
(e)	ns	ns	+	+	(e)	+	+	-	+
(f)	ns	ns	+	+	(f)	+	+	-	+
(g)	+	-	+	-	(g)	+	+	-	-
(h)	+	-	+	-	(h)	+	+	-	-

Part II

Does distance matter?

Chapter 3

Spatial econometric application to Province-Level Productivity

3.1 Introduction

As we have seen in chapter 2 agglomeration economies can be considered important to reduce production costs and improving firms' efficiency. Studies on agglomeration have distinguished between two types of externalities: static and dynamic. The latter kind has been investigated in the previous chapter, where we have defined knowledge spillovers as the introduction of an innovation by one firm improves the performance of other firms, that have not to pay any kind of compensation. A consequence of these externalities is the accumulation of knowledge within the geographical area where firms are localized. Furthermore, knowledge spillovers and, more generally, physical proximity favors long-term relationships and knowledge accumulation, which represent a free resource for local firms, through a mechanism of cumulative causation which finally improves firms' performance.

The decreasing costs due to the developments in transportation and communication technologies, have led to many commenters to believe in the loss of importance of distance, the so called "death of distance", and agglomeration effects, but this idea is not unanimously shared. Indeed some economic geographers, such as Scott (1988), Storper (1997), and more recently McCann (2005), have stressed the importance of geography for knowledge spillovers, since certain types of knowledge need human interactions that can be more easily established when people are located near each other.

The literature empirical on knowledge spillovers has found favourable evidence for distance, stating that location matters, and have encouraged spatial analysis exploiting spatial econometrics tools, considered the best technique to analyze the presence of a spatial structure of innovations, to investigate the presence of clusters and to estimate model which explicitly include geography in their specification.

In light of new findings within Geographical Economics, the traditional approach followed to analyse spatial externalities has been criticized because of the assumption that geographical units are independent.

Guillain and Le Gallo (2006)along these lines, have stressed how indeces computed to assess the role of space and geographic concentration, wrongly treating geographical units as identical cause biases in spatial agglomeration effects estimation.

We aim, in this chapter, to focus on distance as a variable that allows to analysing linkages among economic agents, combining the traditional approach, based on agglomeration indices, with recent tools provided by spatial econometrics, allowing to explicitly account for spatial autocorrelation.

We follow this method in order to verify the presence of spatial patterns of TFP, investigating whether provincial Total Factor Productivity may be affected by neighboring provinces productivity.

This chapter is organized as follows: section 2 places this work within the relative economic framework, section 3 provides a description of spatial econometric characteristics, section 4 outline the methodology adopted and the results obtained, section 5 concludes.

3.2 Literature review

The topics analysed in this chapter arise from two different strands of literature, one represented by the agglomeration economies framework, discussed in the previous chapter, and the second represented by the spatial convergence of TFP.

Indeed, we aim to investigate the presence of spatial patterns for TFP focusing on how geography can affect knowledge spillovers, introducing spatial interactions into the traditional framework.

We have extensively discussed of agglomeration economies in the previous chapter, where we have considered how the traditional approaches have neglected to consider the problem of operating on different spatial scales, as also stressed by Rosenthal and Strange (2004), who criticize empirical studies on agglomeration externalities since they usually refer to fixed geographical units, such as region or city. Moreover, they have noted that there is not enough information about the scale sensitivity and spatial analysis often neglects the possible presence of agglomeration externalities at different spatial scales. Spatial composition and aggregation effects are known as Modifiable Areal Unit Problem (MAUP). A possible approach that tries to overcome the MAUP[1] introducing, consists of introducing distance to account for spatial patterns.

Along this lines, Guillain and Le Gallo (2006) have suggested a methodology to measure the degree of spatial agglomeration and proposed a new approach to measure of agglomeration and of identification of economic sector pattern. Their approach combines a locational Gini index with the tools of Exploratory Spatial Data Analysis (ESDA). They apply this methodology to identify the location of employment at municipality patterns for 26 manufacturing and services sectors in Paris, in order to overcome some of the weaknesses of the previous approaches.

A good review of MAUP is provided by Burger et al. (2007), who analyse the effect of agglomeration externalities on sectoral concentration growth, from 1996 to 2004, in six sectors considered at the following levels of aggregation: 483 municipalities, 129 economic geographic areas, and 40 regional level. These three spatial levels1 have been analysed through spatial autoregression model in order to control for spatial spillover effects in growth.

Burger, van Oort and van der Knaap state that accounting for spatial autocorrelation does not guarantee overcoming MAUP, but that it could be addressed moving from from a meso-economic approach to a micro-economic approach, for example using continuous space as suggested by Arbia (2001)or multilevel analysis as suggested by Hox (2002). Further they suggest the use of firm level analysis and after model identification at the micro spatial level, draw policy implications at more aggregated level.

In spite of a large literature on convergence of TFP, only few studies have analysed that analyse country TFP applying spatial econometrics tools. In particular, most of these works try to extend the traditional framework proposed by Nelson and Phelps (1966)including cross-sectional

[1] Openshaw (1984) has provided a detailed description of MAUP. He has argued that "the areal units (zonal objects) used in many geographical studies are arbitrary, modifiable, and subject to the whims and fancies of whoever is doing, or did, the aggregating". MAUP and the consequent scale effects may cause variation in statistical results between different levels of aggregation. These variation depend on the size of spatial unit considered.

interactions. We can notice this approach in Abreu et al. (2004),Mishra et al. (2008), Di Liberto and Usai (2010).

Abreu et al. (2004) investigate the presence of a spatial dimension of technology flows across countries, implementing explanatory spatial data analysis and spatial econometric techniques.

Their measure of TFP is obtained by a constant returns to scale Cobb-Douglas production function and setting the capital share of income to 1/3 and the labour share to 2/3. The capital stock series is constructed using the perpetual inventory method. Then, they add the spatial lag of TFP growth to the Nelson-Phelps model, and employ spatial econometric estimators such as SAR and SEM. Further, they provide some diagnostic tests with respect to the choice of the appropriate specification.

Looking at 73 countries over the 1960-200 period, their exploratory analysis shows that TFP growth rates and levels are positively autocorrelated over space; they find statistically significant clusters of low TFP values in Sub-Saharan Africa, Central and South America, and clusters of high values in North America and Europe. These results may suggest the presence of barriers to technology adoption, and the possibility that technology levels are converging locally, with countries forming spatial convergence clubs. A possible implication is that countries would benefit from closer ties to the local technology leader.

Mishra et al. (2008)have proposed a model for analysing the spatial interdependence of total factor productivity (TFP). They examine whether TFP growth is correlated across spatial locations and explore how human capital accumulation dynamics can explain interdependence in TFP growth among countries.

Their analysis is applied to a sample of 15 Asian countries for the period 1970-2000, in a model with feedback effects in a semi-parametric spatial Vector Autoregressive (VAR) setting in order to estimate dynamic spatial externalities among countries' TFP growth processes and their error terms with respect to human capital differences. The distance between countries is perceived as due to the human capital differences.

Their idea is to shed light on a new explanation of cross-country TFP complementaries by utilizing the dynamics of age-structured human capital accumulation across countries. The finding of spatial correlation in TFP and/or growth is significant in that differences in human capital across countries could explain the non-linear spatial correlation. Further they argue that conditional convergence in TFP could occur, and that human capital plays a crucial role in determining the dynamic path of TFP.

Di Liberto and Usai (2010) have investigated TFP convergence across European regions accounting for TFP heterogeneity and dynamics by means of a fixed-effects panel estimator in a standard convergence equation framework. TFP levels are estimated by means of growth regressions, applying, firstly, spatial panel tools,

Traditional approach to study productivity convergence has been merged with the spatial dimension of TFP across regions, combining, at the same time, traditional panel data estimators with spatial econometric tools.

They have not found evidence for TFP convergence but, on the contrary, of TFP dispersion. In spite of the absence of convergence, they find a dynamic pattern across regions, revealed by exploratory spatial data techniques. In particular, they find that between 1985 and 2006, there have been numerous changes in clusters patterns observing with miracles and some disasters in terms of TFP performance, causing intra-distribution movements with significant changes in regional rankings and cluster composition. Their results confirm that cross-regional gaps in TFP levels are significant and persistent.

3.3 Spatial Econometrics Methods

Spatial analysis involves both statistics and econometrics, however spatial econometrics and spatial statistics, differ with respect to the object of analysis and the approach followed. Spatial econometrics follows more often a deductive approach where tools are applied to regional and urban analysis and considered as instruments to verify a previously defined theory; instead, spatial statistics mostly follows a descriptive and inductive approach, where tools are is mainly employed to analyse natural phenomena and often without a reference model.

Hence, we aim to apply spatial econometric to analyse the presence of spatial effects in TFP across provinces. Before proceeding with our analysis, we outline the main characteristics and aims of this branch of econometrics, which includes all the techniques dealing with space characteristics. Two are the main characteristics of spatial econometrics: spatial dependence and spatial heterogeneity.

Spatial dependence refers to data where observations associated to a location i depend on observations associated to another location $j \neq i$, formally:

$$y_i = f(y_j)\ , i = 1, ..., n \quad j \neq i.$$

Two main reasons may lead to spatial dependence: measurement errors and spatial interactions among economic agents, as suggested by regional studies. Measurement errors can occur when dealing with data associated to spatial units, such as regions, provinces, cities, etc, theoretical boundaries of the geographical unit analysed are not a good representation of the true data generating process. Spatial interactions, are the most important factors that implies spatial dependence, since regional science has stressed the importance of distance and location in human geography and market activity.

Spatial heterogeneity, instead, relates to the differences in relations across space. Formally, it can be expressed through a linear relation, such as:

$$y_i = X_i \beta_i + \varepsilon_i$$

where i indicates points in space, X_i is a matrix of explanatory variables varying across space, associated to a vector of coefficients β. Given a sample of N observations, we cannot estimate N relations (and so N parameters), since we do not have enough information and there exists a problem of degrees of freedom. This implies the use of a more parsimonious analysis.

The main goal of spatial analysis is to introduce variables related to neighboring locations. This aim is achieved by defining contiguity among locations. Several methods exists to define contiguity, but the starting points of each method is represented by the geographical coordinates, latitude and longitude, or alternatively geographical borders. Contiguity is introduced through a matrix, generally indicated by W, that contains connectiveness relations. Two main kinds of matrices can be build, binary and distance matrix, where binary matrix is characterized by two elements: 1 or 0, respectively indicating contiguity or not. Among the alternative methods to define contiguity relations, we can mention linear contiguity, rook contiguity, bishop contiguity, double linear contiguity, double rook contiguity and queen contiguity. Instead, distance is generally, computed using the *great circle distance method.* The W matrix is symmetric and has zeros on main diagonal, further it is common practice to row standardise W , allowing to deal with the mean of observations in contiguous regions.

Formally, we can rewrite the spatial dependence relation as:

$$y = \rho W y + \varepsilon$$

where W represents the row standardised contiguity matrix, ρ captures the spatial dependence in our sample measuring the average influence of neighboring observations. The previous equation is the most basic, but explanatory variables can be added, obtaining the more complete specification:

$$y = \rho W y + X\beta + \varepsilon.$$

In order to face spatial heterogeneity, two are the main approaches: the spatial expansion proposed by Casetti (1972)and the geographically weighted regression as proposed by McMillen (1996). The first is based on the idea that location information is function of latitude and longitude. Estimation is based on two steps:

$$y = X\beta + \varepsilon \tag{3.1}$$

$$\beta = ZJ\beta_0 \tag{3.2}$$

where Z includes the latitude and longitude information and J represents a structure of identity matrices. The first step consists in estimating equation (3.2) and then substitute the fitted β in equation (3.1).

The second approach is based on the following relation:

$$W_i y = W_i X\beta_i + \varepsilon_i$$

that consists in a transformation of the standard model, $y = X\beta_i + \varepsilon_i$, obtained premultiplying the weights matrix. This transformation allows to achieve the orthogonality condition for error terms and applies the OLS estimator to the transformed model. Parameters are obtained as:

$$\widehat{\beta}_{\mathbf{i}} = (X'W_i^2X)^{-1}(X'W_i^2y).$$

Empirically, the increased availability of geo-referenced data and the interest in understanding the role of space for economic activity has led a deepen interest for the analysis of spatial autocorrelation.

According toAnselin and Bera (1998) spatial autocorrelation can be defined as follow: *"Spatial autocorrelation can be loosely defined as the coincidence of value similarity with locational similarity. In other words, high or low values for a random variable tend to cluster in space (positive spatial autocorrelation) or locations tend to be surrounded by neighbors with very dissimilar values (negative spatial autocorrelation). Of the two types of spatial autocorrelation, positive autocorrelation is by far the more intuitive. Negative spatial autocorrelation implies a checkerboard pattern of values and does not always have a meaningful substantive interpretation*".

Spatial autocorrelation analysis can be analysed in both univariate or multivariate frameworks.

The univariate analysis of spatial autocorrelation is based on global indicators, such as the Moran and Geary Indices, and the Local Index of Spatial Autocorrelation, such as the local version of Moran *I.*

Moran's I (Moran (1950)) is a weighted correlation coefficient used to investigate spatial patterns, such as clusters. Positive spatial autocorrelation can that nearby areas exhibit similar rates, negative values of Moran's I are associated dissimilar rates.

We can formally write the Moran Index as follows:

$$I = \frac{N}{S}\frac{\sum_{ij}^{N} W_{ij}(x_i - \bar{x})(x_j - \bar{x})}{\sum_{i}^{N}(x_i - \bar{x})^2}$$

where N indicates the number of spatial points, S represents the sum of all weights w_{ij}. As N becomes larger, Moran's index I will approximate to a standard normal $Z(I)$. This statistic is based on the null hypothesis of absence of spatial autocorrelation; this hypothesis is rejected for significant values positive (or negative) indicate positive (or negative) autocorrelation.

It is easy to note the similarity between Moran Index and Durbin Watson (DW) test, with the former representing a DW suitable for geo-referenced data.

In empirical analysis the Moran's Index as global indicator of autocorrelation can be combined with a local indicators that allows analysing the correlation across each region. The Local Indicator of Spatial Autocorrelation (LISA) is the local version of Moran statistic, formally:

$$I_i = \frac{z_i}{\sum_i z_j^2 / N} \sum_{j \in J_i} w_{ij} z_j$$

z_i represents the normalized value of Moran Index associated to the i-th location. Local Moran Index is also based on the null hypothesis of no spatial autocorrelation and it can be interpreted in the same way. Positive and significant values of this statistic denote the existence of clusters among geographical units, characterized by similar characteristics. Instead, negative and significant values indicate that neighboring regions (provinces) have different characteristics.

Multivariate spatial dependence relate to the variance-covariance matrix, where the off diagonal element are different from zero. Formally:

$$Cov[x_i, x_j] = E[x_i, x_j] - E[x_i]E[x_j] \neq 0$$

where i and j represent spatial units.

In order to estimate the variance-covariance matrix, a definition of spatial stochastic process is required. Spatial econometrics helps us to define the stochastic process and in particular a linear model with spatial lags of dependent variable is proposed. This model is characterized by simultaneity issues due to the endogeneity of the spatial lag of dependent variable, implying the biaseness of Ordinary Least Squares (OLS) estimator. OLS estimator, however, is still suitable, returning unbiased coefficients, if spatial dependence concerns only the residuals. In this case, spatial autocorrelation can be faced as temporal autocorrelation, that is, through an efficient estimator and OLS could be still applied by adjusting the residuals for spatial dependence. Before explaining alternative regressors, we briefly outline a general spatial model:

$$y = \rho W_1 y + X\beta + \varepsilon$$

$$\varepsilon = \lambda W_2 \varepsilon + u$$

where $u \sim N(0, \Omega)$, W_1, and W_2represent the row standardised spatial weights matrices, associated respectively with a spatial autoregressive process in the dependent variable and in the disturbance term. The parameters to be estimated are ρ, β, λ, and σ^2 but setting some of these parameters to zero, we can obtain simpler specifications. For example, assuming $\rho = 0$ and $\lambda = 0$, we obtain the classical linear regression; for $\lambda = 0$, we obtain a mixed regressive-spatial autoregressive model; for $\rho = 0$, we obtain the linear regression model with spatial autoregressive disturbance.

These basic models could be integrated introducing the time dimension. This determines an increase in complexity since aspects related to both dimensions, geographical and temporal, need to be considered, such as spatial and time autocorrelation and heterogeneity.

While the linear regression model with spatial autoregressive disturbance can be estimated through classical econometric tools, adjusting the variance covariance matrix to account for the spatial dependence in the error terms, the spatial autoregressive model, instead, requires suitable spatial econometric tools to overcome the endogeneity problem generated by the spatial lag of the dependent variable.

When the OLS estimator is biased, alternative estimators include Maximun Likelihood Estimator (hereafter,MLE), Instrumental Variables (IVs, 2SLS) and robust OLS.

The MLE is unbiased, efficient and asyntotically normal, these characteristics still hold when it is applied to spatial models. Models that include spatial interactions requires the existence of a likelihood function for estimated parameters, that this function is differentiable and that the number of coefficients is fixed.

The main drawback of the the ML estimator is that it is based on numerical methods of non linear optimization, which make this estimator unsuitable when the number of observations is large.

In a mixed regressive-spatial autoregressive model, the MLE assumes that the disturbance terms are jointly normally distributed. The estimator applied to the spatial model differ with respect to the standard one, because of the endogeneity caused by the spatial lag of the dependent variable. This issue can be overcome applying the following steps:

1. perform the OLS estimator for the model , $y = X\beta_0 + \varepsilon_0$;
2. perform the OLS estimator for the model $Wy = X\beta_L + \varepsilon_L$;
3. determine the residuals from the previous two regressions, e_0 and e_L;
4. find the value of ρ that maximizes the concentrate likelihood function:$L_C = C - \left(\frac{n}{2}\right) ln \left(\frac{1}{n}\right) (e_0 - \rho e_L)' (e_0 - \rho e_L) + ln|I - \rho W|$,
5. then determine the parameter $\hat{\beta} = \hat{\beta}_0 - \rho\hat{\beta}_L$.

One important feature is that we have to impose to the parameter ρ to take feasible values, parameter must assume values in this range:$\frac{1}{\lambda_{min}} < \rho < \frac{1}{\lambda_{max}}$ where λ_{min} is the minimum eigenvalue of the standardized spatial contiguity matrix and λ_{max} represents the maximum.

The endogeneity of the spatial lag of the dependent variable can be addressed applying an alternative approach represented by instrumental variables.

The Instrumental Variables estimator has the same asyntotic characteristics of the ML estimator, but it is easier to compute. This method allows overcoming endogeneity by instrumenting the spatial lag of the dependent variable. Instruments have to respect the properties set by traditional econometrics. The Instrumental Variable approach has been proposed by Kelejian and Prucha (1998) and Anselin (2002). According to Kelejian and Prucha (1998)Kelejan and Prucha (1998), the instrumental variables employed to instrument the spatial lag of the dependent variables are respectively the regressors' matrix X, the spatial lag of the regressors' matrix, WX and the squares of spatial lag of independent variables, W^2X. The inclusion of higher lags may improve the precision of the estimation, but at the same time can lead to a larger bias.

Anselin (2002) has proposed as instruments the spatial lag of the predicted values of a regression, where the dependent variable is regressed on no spatial regressors, or where the regressors are spatial lag of exogenous variables, although we can incur in multicollinearity issue.

The last approach, called "robust", has been widely used through bootstrap or jacknife resampling methods. This estimator aims to check for biaseness, comparing efficient results, obtained implementing this method, with variables original data.

3.3.1 Spatial panel

Earlier works in spatial econometrics have focused on cross- section dimension. Only recently, thanks to the developments of spatial panels by Balatgi et al. (2003),Balatgi et al. (2007),Balatgi and Liu (2008), Anselin and Le Gallo (2008)and Elhorst (2009), these techniques have started to spread. Spatial panel tools allow exploiting the traditional advantages of spatial data and so allow considering individuals heterogeneity and, at the same time, to take into account spatial dependence.

According to Hsiao (2003) , working with panel data yields many benefits. Indeed, panel data allow controlling for individual heterogeneity, overcoming a weakness shared by time series and cross sectional analysis, that cannot take into consideration heterogeneity presents across firms, states, countries and individuals. Further, panel data are characterized by less collinearity and more variability, providing more informative data. These data are better able to study the dynamics of adjustment. Indeed, many economic phenomena change over time, such as spells of unemployment, job turnover, residential and income mobility, unemployment and poverty, and holding a panels long enough, allow to investigate for example, speed of adjustments to economic policy changes. Moreover, micro panel data may be more accurately measured than similar variables measured at the macro level, allowing to overcome aggregation bias issue.

However, panel data do not cause only benefits, but they are characterized by some limitations. Among others, there are: difficulties related to data survey and data collection, distortions due to measurement errors, that can be caused by unclear questions, memory errors, inappropriate informants. Furthermore, micro panels are often characterized by a short time-series dimension, implying that asymptotic arguments rely on the number of individuals tending to infinity; Finally, another drawback that panel data may be affected is the cross-section dependence. Usually, this issue has been analysed the time autocorrelation, but macro panels on countries or regions they could be affected by cross-country dependence, neglecting it may lead to misleading inference.

Spatial panel tools permit to overcome the drawbacks related to this last point, allowing to take into account cross-sectional dependence and individuals heterogeneity.

Starting from a linear panel model with spatial specific effects, but without spatial interactions, such as:

$$y_{it} = x_{it}\beta + \mu_i + \varepsilon_{it},$$

where i indicates the spatial units, t time, y_{it} denotes the dependent variable and x_{it} regressors' matrix. In matrix form:

$$Y = X\beta + \mu + \varepsilon$$

The reasoning is similar to that of standard panel data models, that is individual specific effects are time-invariant variables whose omission could bias the estimates. In particular, in this case, the individual fixed effects are represented by the spatial specific effects.

This model can be completed by introducing spatial interactions through a spatially lagged dependent variable or a spatial autoregressive process in the error term, respectively we would have a spatial lag and the spatial error model.

The previous equation can be rewritten as follow: for a spatial lag model:

$$Y = \rho WY + X\beta + \mu + \varepsilon$$

and

$$Y = X\beta + \mu + \lambda W\varepsilon + u$$

for a spatial error model.

The matrix W describes the spatial pattern of the units in the sample. This matrix is of order N, that is the same number of individuals. The spatial lag model is considered as the formal specification for the equilibrium outcome of a spatial or social interaction process, where the dependent variable referred to an individual is determined by neighboring individuals characteristics. Instead, the spatial error model assumes that the dependent variable depends on a set of observed local characteristics and that the error terms are correlated across space.

The W matrix, that as we have said above, can assume two kinds of structure, binary or distance matrix, becomes a weights matrix through the standardization procedure. Generally, the weights matrix is row standardized. This procedure allows keeping the mutual proportions between elements of W. This property is very important when we deal with an inverse of the distance matrix, because the weights sum to one losing the originary economic interpretation ((Anselin, 1988)). Two main approaches have been suggested in the literature to estimate models, which include spatial interaction effects. One is based on the maximum likelihood (ML) principle and the other on instrumental variables or generalized method of moments (IV/GMM) techniques.

Standard fixed effects are treated through a within transformation or the introduction of $N-1$ dummy variables losing degrees of freedom. The within transformation allows eliminating individual fixed effects μ_i, and it is applied as follows:

$$\tilde{y}_{it} = y_{it} - \sum_{t=1}^{T} y_{it}$$

$$\tilde{x}_{it} = x_{it} - \sum_{t=1}^{T} x_{it}$$

When spatial interactions are introduced, the transformation will assume the following form:

$$y_{it}^{*} = y_{it} - \sum_{j=1}^{N} w_{ij} y_{jt} = y_{it} - \frac{1}{N} \sum_{j=1}^{N} y_{jt} \tag{3.3}$$

$$x_{it}^{*} = x_{it} - \sum_{j=1}^{N} w_{ij} x_{jt} = x_{it} - \frac{1}{N} \sum_{j=1}^{N} x_{jt} \tag{3.4}$$

This transformation differs from the traditional within transformation that considers fixed effects in time.

In empirical work, it has been notes that estimation with or without accounting for spatial fixed effects leads to different results. These differences are due to the different kinds of variation considered, because models with controls for spatial fixed effects utilize the time-series component of the data, whereas models without controls for spatial fixed effects utilize the cross-sectional component of the data.

According to Anselin et al. (2006), extending the Fixed Effects Estimator to spatial lag model implies dealing with th endogeneity of the spatial lag of the dependent variable that does not respect the orthogonality condition. In order to face this issue it has been extended the Maximum Likelihood estimator adopted in cross-sectional spatial models to the panel dimension. The log likelihood function when spatial specific effects are considered as fixed is[2]:

$$LogL = -\frac{NT}{2}log(2\pi\sigma^2) + T \cdot log\left|I_N - \rho W\right| + -\frac{1}{2\sigma^2}\sum_{i=1}^{N}\sum_{t=1}^{T}\left\{y_{it}^{*} - \rho\left[\sum_{j=1}^{N} w_{ij}y_{jt}\right]^{*} - x_{it}^{*}\beta\right\}^2$$

.

Assuming a spatial error model leads to the following log-Likelihood function:

$$LogL = -\frac{NT}{2}log(2\pi\sigma^2) + T \cdot log\left|I_N - \rho W\right| +$$

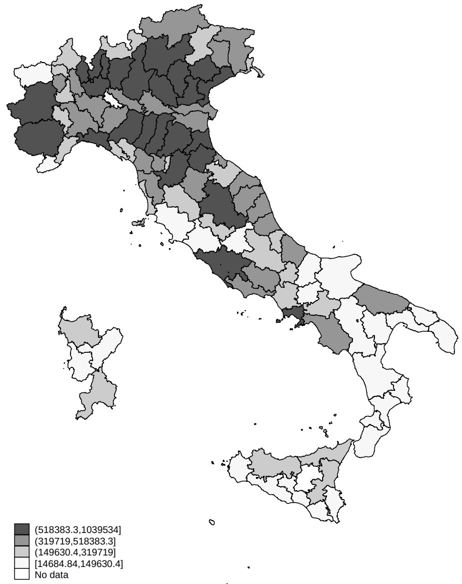
$$-\frac{1}{2\sigma^2}\sum_{i=1}^{N}\sum_{t=1}^{T}\left\{\left(y_{it}^{*} - \rho\left[\sum_{j=1}^{N} w_{ij}y_{jt}\right]^{*}\right)^2 - \left(x_{it}^{*} - \rho\left[\sum_{j=1}^{N} w_{ij}x_{jt}\right]^{*}\right)\beta\right\}^2.$$

3.4 Methodology and Results

Our empirical analysis focuses on the spatial effects of province productivity. In particular we want to test if productivity, besides being affected by spatial agglomeration, receives some benefits from neighboring provinces productivity. We want to investigate whether or not, there exists knowledge spillovers across provinces closely localized. We deal with an aggregate dataset, but province TFP has been built adopting the Olley and Pakes procedure, starting from firm level productivity, allowing us to extend our results to firms productivity. The presence of a spatial pattern of province TFP implies that even firms'productivity is affected by neighboring productivity and so that distance matters.

[2]For further explanations on the spatial panel Fixed effects estimator, see Spatial Panel Data Models, in Fischer M.M., Get is A. (eds.) Handbook of Applied Spatial Analysis, pp. 377-407. Springer, Berlin.

Figure 3.1: TFP_overtime



We first proceed through the univariate analysis and later to the multivariate analysis, employing an autoregressive model SAR, which includes only the spatial lag of dependent variable as regressor and the alternative SEM model, assuming that TFP presents a spatial structure in the error term.

Before proceeding with the SAR and SEM models, we begin, as suggested in Dettori et al. (2011) from the graphical inspection of spatial dependence. Figure 1 shows the spatial structure of TFP across Italian provinces. We can observe that northern provinces are characterized by higher productivity, and lower productivity can be observed in southern provinces. Moreover, this figure shows the presence of clusters of high (low) productivity. Since our measure of productivity is obtained aggregating (through OP procedure) firm level productivity, this seems to suggest that high productivity firms tend to cluster.

Then, we move to test for spatial dependence through global Moran I and Geary C tests, and local Moran I, whose results are reported, respectively, in table 3.1 and 3.2.

Global spatial dependence tests lead us to reject the null hypothesis of no spatial autocorrelation, supporting our idea that TFP presents a spatial pattern, bringing us to introduce spatial

Table 3.1: Global Moran I - Geary C test on TFP

TFP	*I*	*E(I)*	*sd(I)*	*z*	*p-value**
Moran's I	0.100	-0.001	0.034	2.975	0.001
Geary's C	0.904	1.000	0.046	-2.101	0.018

Table 3.2: Local Moran I test on TFP

Province	1999	2000	2001	2002	2003	2004	2005
Agrigento	R	R				R	
Aosta					R		
Avellino					R		
Benevento					R		
Brindisi		R	R			R	
Caltanissetta	R	R				R	
Campobasso	R					R	R
Caserta						R	
Catania						R	R
Catanzaro		R					
Cosenza		R					
Crotone		R				R	
Foggia	R	R				R	R
Isernia						R	R
Lecce		R	R			R	
Nuoro						R	R
Oristano	R	R				R	R
Palermo	R						
Potenza	R					R	R
Ragusa	R	R	R		R	R	R
Rieti	R						
Roma	R	R					R
Sassari						R	R
Siracusa	R	R			R	R	
Taranto						R	
Terni	R	R					
Varese							R
Vibo Valentia	R	R				R	R

R indicates that we reject the null hypothesis of no spatial correlation.

We do not report provinces when the null hypothesis of no Spatial Agglomeration is not rejected interactions in our specification. Local Moran Index allows testing for the presence of spatial correlation referring to every spatial unit considered. Results show that only few provinces, with respect the 102 that we are examining, present spatial correlation, in particular we can observe that these provinces are mainly localized in the south of the country.[3]

Moreover, we have tested the presence of spatial dependence through the Cross-Section Dependence (CD) test proposed by Pesaran (2004), which, as proved by Pesaran (2004), is applicable to a large variety of panel data models, including stationary and non-stationary dynamic heterogeneous panel with short T and large N, as it is the case for the panel of data used in this study. The test, which is based on the average of the pair-wise correlation coefficients, is calculated as follows:

$$CD = \sqrt{\frac{2T}{N(N-1)}\left(\sum_{i=1}^{N-1}\sum_{j=i+1}^{N}\hat{\rho}_{ij}\right)}$$

Performing Pesaran's test, assuming that the null hypothesis is of cross sectional independence, we obtain a value of 12.372 for CD test and a p-value of 0, that leads us to reject the null hypothesis, hence, we can proceed our analysis including spatial interaction terms.

The SAR model, that we are proceeding to estimate, can be formally be written as follow:

$$TFP_{pt} = \rho W \cdot TFP_{pt} + province\ dummy + time\ dummy + \varepsilon_{pt}$$

where $p = 1, ..., N = 102$, $t = 1, ...7$, ρ captures the spatial dependence, ε represents the random error term and W is the row standardized weights matrix, whose elements are :

$$w_{ij} = \frac{1}{d_{ij}^2}.$$

Departing from the standard practice, our decay matrix does have not zeros in the main diagonal but when $i = j$ we compute the internal distance as:

$$d_{ii} = \frac{2}{3}\sqrt{\frac{area_i}{\pi}}$$

[3]Graphical support of LISA analysis are available on request.

The SEM model can be written as :

$$TFP_{pt} = province\, dummy + time\, dummy + \lambda W \varepsilon_{pt} + u_{pt}$$

We choose to perform our estimation through Maximum Likelihood estimator (MLE), because we are dealing with a rather small sample that does not affect the MLE characteristics[4]. Further, when applied to spatial dataset, MLE allows to account for endogeneity of spatial lags of the dependent variable, as explained in the previous section.

We present results obtained applying the MLE, firstly to a pooled panel and then introducing some control variables. In particular in columns (1) and (5) we report results related to the pooled panel. In the other columns, we have introduced dummy variables in order to control for individual and time fixed effects, since our sample has a panel structure. We further report results related to two specifications, respectively an autoregressive model and a mixed regressive, autoregressive model. In the former, only the spatial lag of TFP is introduced as regressor; instead, in the latter, we introduce as regressor also the agglomeration indices, constructed in chapter 2. We follow the same approach for the SAR and SEM models.

Results are not unanimous. In particular when we consider the SAR model, we find positive evidence in favor of a spatial pattern of productivity. Results change introducing the time fixed effects. In particular, dealing with pooled panel data and introducing individual fixed effects leads to a significant and positive coefficient of spatial interaction, represented by the spatial lag of the dependent variable, suggesting that provincial productivity is affected positively by neighboring province productivity. This result supports the idea that productivity can improve through knowledge spillovers and in particular closer locations can have greater benefits since closeness favors the spreading of knowledge.

Then, we have proceeded to estimate a mixed regressive autoregressive model. In this specification, we have introduced as regressors the agglomeration indices adopted in the previous chapter, and departing from the theoretical model, we have not added the spatial lag of the regressors

[4] The MLE is consistent, efficient and under certain regularity conditions, maximum likelihood estimator has an asymptotically normal distribution.

matrix since these variables are highly correlated and could introduce multicollinearity problems.

Entering the agglomeration indices as regressors, we find that only the specification in column (6) reports a significant and positive coefficient for the spatial lag of TFP, while in the other cases this parameter results insignificant.

We can observe in table 3.4, where the results related to the spatial error model are reported, that this model leads to the same results of the SAR model. In particular, the coefficient associated to spatial interactions is positive and significant when MLE is applied to the pooled panel and when province fixed effects are introduced.

Again, the same results have been obtained performing the mixed regressive auto-regresissive model; indeed, we have found support for a spatial pattern of the error component only considering the the specification (6) in table 3.4, that includes province fixed effects.

Introducing fixed effects in this model is...... to apply a standard within transformation that could be not suitable for spatial panel data. In this case the appropriate transformation is represented by equations (3.3) and (3.4). Estimators adopting this transformation are also performed, and the results obtained are reported in table 3.5 and table 3.6 respectively for the SAR and the SEM specifications.

We can observe that the spatial interactions coefficients are always positive and significant, even if their magnitude is very modest. Regarding the agglomeration indices, we have found poor evidence, and in particular, the density index is positive and significant only when province and time fixed effects are introduced in the SAR and SEM specifications. In the same specifications competition coefficient is negative and significant. A positive sign for competition is found when the specification includes only time fixed effects. The other indices are never found significant.

3.5 Conclusions

In this chapter we have analysed how province productivity is affected by neighboring productivity. Approach implemented have allowed to investigate the presence of knowledge spillovers and the importance of geography, since we sustain that human interactions favour the spreading of information and knowledge and human interactions are favored by closeness of workers. The

Table 3.3: SAR province TFP

TFP_{pt}^{LP}								
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
ρ	0.1085***	0.125***	0.0835*	0.0635	0.063	0.124***	0.0338	0.0570
	(0.457)	(0.0385)	(0.0467)	(0.0393)	(0.0424)	(0.0384)	(0.0438)	(0.0385)
log_spec					-0.1326	-1.512	-0.1239	-0.8980
					(0.2097)	(1.214)	(0.2023)	(1.190)
$log\_diversity$					0.002	-0.0677	0.0033	-0.0341
					(0.0187)	(0.0583)	(0.0183)	(0.0571)
$log\_density$					0.001	0.3467	0.0009	1.131***
					(0.0119)	(0.269)	(0.0116)	(0.3714)
log_comp					0.04898***	-0.1396	0.0491***	-0.317**
					(0.00996)	(0.1086)	(0.0098)	(0.127)
$constant$	8.912***	8.39***	9.146***	8.983***	9.3199***	7.277***	9.583***	3.741***
	(0.0455)	(0.38293)	(0.468)	(0.3894)	(0.438)	(1.427)	(0.4464)	(1.906)
$province\ FE$	no	yes	no	yes	no	yes	no	yes
$time\_FE$	no	no	yes	yes	no	no	yes	yes
* significant at 10%; ** significant at 5%; *** significant at 1%, values in (.) represent standard errors								

Table 3.4: SEM spatial regression- province TFP

	TFP_{pt}^{LP}							
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
λ	0.108***	0.125***	0.0835	0.0635	0.0269	0.1277***	-0.0058	0.05716
	(0.0455)	(0.0385)	(0.0467)	(0.0392)	(0.0454)	(0.0405)	(0.04695)	(0.0404)
log_spec					-0.125	-1.196	-0.1183	-0.7867
					(0.2109)	(1.252)	(0.2026)	(1.219)
$log\_diversity$					0.00313	-0.0699	0.0047	-0.0366
					(0.019)	(0.0586)	(0.0184)	(0.0576)
$log\_density$					4.53e-06	0.4165	0.00078	1.163***
					(0.012)	(0.2919)	(0.012)	(0.3805)
log_comp					0.0498***	-0.16	0.0499***	-0.3101***
					(0.01012)	(0.111)	(0.0099)	(0.129)
$constant$	9.996***	9.624***	9.979***	9.60633***	9.94***	8.019***	9.915***	4.093**
	(0.00872)	(0.0623)	(0.0245)	(0.057)	(0.1087)	(1.5064)	(0.1102)	(1.948)
$province\ FE$	no	yes	no	yes	no	yes	no	yes
$time\_FE$	no	no	yes	yes	no	no	yes	yes

* significant at 10%; ** significant at 5%; *** significant at 1%, values in (.) represent standard errors

Table 3.5: SAR spatial panel regression

	TFP_{pt}^{LP}			
	(1)	(2)	(3)	(4)
ρ	0.007***	0.007***	0.006***	0.006***
	[4.234]	[4.2328]	[3.62995]	[26.0921]
log_spec	-1.1601	-1.977	-0.7625	-0.766
	[-1.0045]	[-1.6947]	[-0.76288]	[-0.76656]
$log\_diversity$	-0.0362	-0.0725	-0.0767	-0.0771
	[-0.5640]	[-1.125]	[-0.7688]	[-0.772]
$log\_density$	1.233***	0.42603	0.3861***	0.3855
	[2.689]	[1.2409]	[4.837]	[4.7528]
log_comp	-0.2954***	-0.10799	-0.135*	-0.1338
	[-2.207]	[-0.861]	[-1.526]	[-1.5128]
$constant$	3.8527***	8.0465***	8.199	8.198
	[3.4943]	[6.489]	[0.0303]	[13.6686]
$province\ FE$	yes	yes	no	no
$time\_FE$	yes	no	yes	no
* significant at 10%; ** significant at 5%; *** significant at 1% , values in [.] represents t statistic.				

Table 3.6: SEM spatial panel regression

	TFP_{pt}^{LP}			
	(1)	(2)	(3)	(4)
ρ	0.006***	0.007***	0.007***	0.007***
	[3.623]	[4.2379]	[4.225]	[4.2378]
log_spec	-1.11531	-1.9896*	-0.0716	-0.0758
	[-0.964]	[-1.7063]	[-0.4788]	[-0.5011]
$log\_diversity$	-0.0365	-0.0758	0.01126	0.01083
	[-0.5755]	[-1.2008]	[0.7537]	[0.7165]
$log\_density$	1.26755***	0.46798	0.000483	9.445e-05
	[2.7501]	[1.3388]	[0.0404]	[0.0078]
log_comp	-0.2922***	-0.1125	0.04836***	0.049
	[-2.2053]	[-0.901]	[3.6525]	[3.653]
$constant$	4.0901***	8.2734***	9.9022	9.906***
	[3.707]	[6.6735]	[0.244]	[110.603]
$province\ FE$	yes	yes	no	no
$time\_FE$	yes	no	yes	no
* significant at 10%; ** significant at 5%; *** significant at 1% , values in [.] represents t statistic.				

literature on knowledge spillovers has found evidence in favour of the role of for distance, arguing that location matters, and spatial econometrics tools have been considered the best technique to analyze the presence of a spatial structure in innovations, and more generally to estimate models explicitly including geography in their specification. Adopting new tools of spatial analysis permits to overcome the main drawbacks of the traditional approach, where every location has been treated identically.

In the light of these results, combining traditional literature on agglomeration externalities together with the recent literature that studies the spatial pattern of countries TFP, we have investigated the spatial pattern of province productivity performing a spatial panel estimators on our model, characterized by TFP as a dependent variable and including geography explicitly in our specification, through the spatial lag of TFP, together with agglomeration indices as regressors.

Our results are mixed. We have not obtained evidence in favour of spatial spillovers, because, even if spatial interaction coefficients are positive and significant their magnitude is rather small.

The analysis involving the spatial structure of TFP cannot be considered ended, we have proceeded to implement the basic approach, but other hypothesis can be tested and more tools can be applied. We aim to deep further this topic, in particular, we would like to introduce dynamic in this model exploiting the panel dimensions, further, in line with the previous chapters we are going to apply this model at more disaggregated level, such as firm level.

Chapter 4

Spatial economies within a General Equilibrium framework: New Economic Geography[1]

4.1 Introduction

The last two decades have witnessed the surge of "New Economic Geography" (NEG) models. This term refers to a branch of economics, which analyses the localization decisions of workers and economic activities, overcoming the traditional relation between localization and natural endowment. The book written by Fujita et al. (1999), can be considered the first formalization of the NEG theory. New Economic Geography models develop from the International trade theory proposed by Krugman (1979). Gradually, some of the original assumptions and scope of these models have been modified to analyse the disparities among regions within a country. In particular, the NEG underlines the importance of transport costs, recovering the significance of spatial phenomena which cause agglomeration. Indeed, firms do not appear uniformly distributed over a territory and their distribution is the result of exogenous and endogenous forces. Traditionally, the former were identified as advantages of "first nature" (Krugman (1993)) and explain firms' localisation through the unequal endowment of resources, climate, proximity to ports and rivers, scant factors mobility.

[1]This chapter represents an extension of my MSc dissertation, work supervised by Prof. Fingleton

The latter refers to externalities generated by workers, firms, and policy makers, and are therefore, mainly related to people. More recently, NEG attributes the agglomeration/concentration phenomena to the interplay between centrifugal and centripetal forces. Centrifugal forces that cause dispersion, are represented by competition effects, a larger market with many firms producing the same goods induces more competitiveness, leading firms to migrate towards markets with lower costs from competition. Conversely, centripetal forces promote concentration and are referred to as agglomeration externalities. Externalities are produced by firms' clusters and arise from forward and backward linkages, which together enhance the pecuniary externalities of firms, keeping the transport costs low and enjoying large goods markets and a specialized local labour markets. Current theories, which NEG theory belongs to, explain economic concentration/localisation through the concept of increasing returns to scale and more realistic market structures, such as monopolist competition, overcoming the traditional assumptions of constant returns to scale and competitive markets. Geographers and regional economists have stressed that many of the assumptions of the NEG are not completely "new", but they were proposed in traditional models of regional and urban economics. As stressed by Ottaviano and Thiesse (2005), the main merit of NEG models is the capability of merging existing theories of economic geography and localisation in a general equilibrium framework.

There are five are the main features of NEG: general equilibrium approach, increasing returns to scale, imperfect competition, transport costs and localization decisions. Initially, researchers focused on the theoretical development of NEG, while the empirical validation is relatively recent ((Hanson, 2005)). In empirical models, the main equation employed links nominal wages with the market potential (MP), analysing the effect of market potential on the prices of factors and, in particular, the cost of labour. Empirical works concentrate on testing inequality of wages among regions and the presence of spatial wage structure.

In line with the empirical literature, in this chapter we aim to test the spatial structure of wages in Italy, and verify whether or not the NEG theory holds. In particular, this work departs from the standard strand of literature for the methodology employed, because the analysis is conduced at micro-level and different levels of aggregation have been adopted for dependent and independent variables. We have proceed by providing first an overview of the main literature (section 2) and the theoretical model (section 3), and then we present the empirical investigation of wage equation (sections 4 and 5). Section 6 concludes

4.2 Literature Review

In the last decade, there has increased the interest in New Economic Geography (hereafter, NEG), in part due to the admission of the improvements and the innovation introduced by this theory in the study of agglomeration and spatial economics. We think that the existing literature on New Economic Geography can be divided into two main strands: firstly, the theoretical models and the simulation approaches to explain the equilibrium; secondly, the empirical verification concentrated on wage disparities. The first strand includes the seminal works, such as Krugman (1980, 1991) and the book of Fujita, Krugman and Venables (FKV, 1999), which represent the core of the New Economic Geography Theory. These works have enlarged the international trade theory incorporating labour mobility across nations. This feature makes the NEG able to overcome the barrier of international economics and engage also issues typical of regional economics. Also, this process stresses one again the importance of geography and space in mainstream economic theory.[2]The NEG model combines the Core-Periphery and the Dixit-Stiglitz Model, in order to overcome some weaknesses of the Neoclassical framework, by introducing the possibility of increasing returns to scale and monopolistic competition, transport costs and factor mobility. These features can lead to agglomeration phenomena and generate a core–periphery structure endogenously. An analogous approach is followed by Ottaviano and Puga (1998), who examine how the combination of trade costs and house' prices produce agglomeration and dispersion forces and how the pecuniary externalities encourage workers and firms to cluster.

The richness of its assumptions is such that NEG models present some difficulties in realising analytical closed form solution. Hence, it is often necessary to introduce some simplifying assumptions, which allow to obtain closed (solvable) forms. In order to solve these models, researchers often recur to simulation and calibration. Works that adopt the simulation-based approach have been developed, among the others, by Forslid (1999), Baldwin and Forslid (2000),Baldwin (2001).

The original NEG model, as presented by Krugman, has been later improved by successive works which aim to make the model analytically solvable, such as the work of Forslid (1999), or introduce new assumptions in order to make the model closer to reality, such as assuming a number

[2] NEG theory have attributed to the space a more relevant role, it is not only considered as physical space, but it has assumed an economic meaning, the geographical area represent the space where the economic relation are realized. It is overcome the administrative border limit and the space has started to be considered as continuous.

of regions, R, greater than two.

The departure from the standard model regards the assuming that fixed costs in the manufacturing sector consists of a divisible factor, which can migrate between regions. This factor can be human capital and/or physical capital.

Forslid found different agglomeration effects for human capital and physical capital; in particular, in the former case, manufacturing tends to agglomerate for low trade costs as in the standard core-periphery model. Instead, in the physical capital case, economic integration does not cause agglomeration.

Baldwin and Forslid (2000)have combined the traditional core-periphery model with the endogenous growth theory model; in order to contribute to the debate on incentivating growth vs incentivating dispersion. Indeed, these models can contribute to debate on economic integration analysing the role of transport costs and knowledge sharing. The model proposed by Baldwin and Forslid (2000)introduces a migration equation based on utility (not only the real wage gap), captured by the shadow value of migration. They find two steady states; one core-periphery outcome and one dispersion outcome with equal real wages and equal growth. The long run equilibrium, however, can be unstable because of trading costs.

They also stress the importance of the cost of sharing knowledge and that policies aimed at reducing this kind of costs can have a stabilizing effect, since they can reduce the agglomeration of the growing sector. Finally, they show that agglomeration enhances welfare in growth regions restoring partially the periphery welfare.

Baldwin (2001) studies the role of forward-looking expectations, finding that the migrants' behavior does not change and that the assumption of forward-looking expectations does not qualitatively affect equilibrium.

Further, the original NEG model has been modified to relax some of the restrictive original assumptions, in order to make the model closer to the real world. Many of the new assumptions have been proposed by FKV, who, for example, have formulated the model with three symmetric regions instead of two and have studied the effect of introducing transport costs in the agricultural sector. This assumption affects nominal wages and the price index, based on the prices of both sectors.[3]

[3] In addition, FKV have investigated the case where the agricultural sector exhibits decreasing returns to scale. This assumption is also related to the strand of core-periphery models with manufacturing goods as intermediates and with workers' mobility only between sectors and not between regions.

In spite of a well developed literature, New Economic Geography has not been fully explored from empirical standpoint. Fujita and Krugman (2004), in a paper/interview, have suggested three possible directions for future research: 1) expand the theory, 2) sustain the theory with the empirical analysis and 3) contribute to the policy implications. In particular, the first suggestion of Fujita and Krugman refers to the consideration of a broader class of general equilibrium models with imperfectly competitive markets and spatial factors, and a greater focus on labour market structures, embodying the heterogeneity of workers, an aspect still only analysed in an aspatial context. Recently, improvements in this direction have been made by Ottaviano et al. (2002), who propose a monopolistic competition model using a quasi-linear utility function and quadratic sub-utility and linear transport costs. Workers' heterogeneity has been addressed by Tabuchi and Thiesse (2002), Murata (2003), Mori and Turrini (2005), Amiti and Pissarides (2005). Their analysis has concentrated on the spatial sorting of workers and the matching externalities.

Yet, the second suggestion of Fujita and Krugman has only been partially undertaken by empirical researchers and the empirical support for the NEG framework is rather recent. Among the others, we can mention the works of Redding and Venables (2004), Hanson (2005), Brakman et al. (2009), and Redding (2009).

There are five main features of the NEG framework which are empirically testable:

- the home market effect;
- the relationship between market potential and factors inflows;
- the relationship between market potential and local factors price;
- the effect of the change in economic environment on the equilibrium spatial distribution of economic activity;
- the effect of the decline in trade costs for agglomeration.

In a series of papers Davis and Weinstein (1996)Davis and Weinstein (1999)and Davis and Weinstein (2003) and Head and Mayer (2005) have tested the 'home market effect' hypothesis. They found that a larger demand for products of a given industry in any given region causes an increase in the regional production of that industry. The hypothesis that a large market potential increases factor prices is investigated by Brakman et al. (2002), who found evidence of a spatial wage structure in Europe. Their analysis also explores the role of trade costs and, in particular, how the decline in the magnitude of these costs may have affected economic agglomeration. Commonly, the most extensive empirical literature focuses on the relation between wage and market potential, considering labour as the only input factor. Empirical works generally adopt the theoretical model introduced by Helpman (1998) and first applied by Hanson (1998), referred to as the Helpman and Hanson model, instead of the traditional framework proposed by Krugman (1991), since the latter is unable to explain some stylized facts. Also, this model is based on the wage equation and on the assumption that wages are higher in regions belonging to or located close to a large market. The most relevant innovation of this model is the introduction of the housing services as non-tradable goods. The housing sector is a perfectly competitive sector that works as a centrifugal force, since houses are more expensive in the centre. In order to make the model suitable for empirical analysis, two more assumptions are introduced: there is no labour mobility across regions and labour supply comes only from the sector with increasing returns to scale. The Helpman and Hanson model was first applied by Hanson (1998), who analysed the wage structure in the USA. This work has been revisited in 2005 by Hanson (2005), who compares the standard model with an augmented version. The 2005 version introduces changes in the wage equation over time, in order to remove the mean deviation of the average wage. This version emphasizes the importance of workers' skills in determining the differences in regional wages. Implementing the non linear least squares (NLLS) and the GMM estimators, he finds that regional wage disparities are explained by the proximity to a large market, and that the coefficients of the augmented model are consistent with the theory.

Only later, the analysis was extended to the European level, where, however, there are different labour market structures and wages are, generally, the result of central bargaining and are not the result of supply and demand forces. The delay in the application of NEG theory to European countries is probably also the consequence of the scarce mobility of workers, as stressed by Mion (2004) for the Italian case. The poor mobility of workers prevents the price system from clearing labour market. This characteristic can be slightly neglected when the analysis is managed at a more disaggregate level, such as, for example, when we consider cities or provinces, and even more when the wage equation refers to the firm-level.

In the literature analysing testing the relation between wage and market potential for European countries, we can mention Roos (2001), Brakman et al. (2002), De Bruyne (2002), Mion (2004), and Piser (2006). These papers adopt the same empirical approach, based on the Helpman and Hanson's model, and analyse respectively, Germany, Belgium, Italy and Spain. The common goal of these papers is to verify the presence of a spatial wage structure and analyse the wage disparities among regions within the same country.

As mentioned earlier, the model applied is the Helpman-Hanson's model, where the wage of region r depends on ... ..., as follows:

$$log(W_r) = k_0 + \epsilon^{-1} log \left(\sum_s Y_s^{\epsilon+(1-\epsilon)/\delta} H_s^{(1-\delta)(\epsilon-1)/\delta} W_s^{(\epsilon-1)/\delta} T^{(1-\epsilon)D_{rs}} \right) + err_r,$$

which implies the application of non linear estimators. Estimating this equation yields three structural parameters : the share of income spent in manufacturing goods,δ, the elasticity of substitution, ϵ, and transport costs, τ, implicit in T. These papers share, beyond the model, also the econometric techniques implemented. The list of estimators employed includes non linear least squares and, depending of data availability, the non linear instrumental variables and the Generalized Method of Moments. Results support the NEG theory and the importance of final demand linkages, which influence the distribution of wages, enforcing the role of spatial externalities.

Europe and the USA are not the only countries analysed in the literature. In particular, we can recall the works of Hering and Poncet (2006)for China and Kiso (2005) for Japan. In particular, Hering and Poncet have analysed the effects of market potential on individual wages. They have implemented a simple equation where the dependent variable is the wage and the independent variable is market potential; further, they have controlled for gender skills and whether the firm operates in the private or public sector. The empirical evidence supports the presence of a relationship between wages and MP only in the case of highly skilled workers. Also, and interestingly,only the private sector reacts to changes in MP.

Similarly to the theory, the recent empirical literature focuses on the effects of workers heterogeneity to explain the spatial wage structure. Empirical evidence in support this theory for Italy is provided by Mion (2004), for France by Combes et al. (2008) and for Germany by Redding and Sturm (2008).

A clear gap between the theoretical and empirical analyses of NEG models emerges from the review of the literature, as these models concentrate only on the equilibrium wage equation, neglecting other important aspects. Furthermore, some criticisms were raised about the validity and the relevance of these models by Head and Ries (2001) and Davis and Weinstein (2003).

Very recently, Fallah et al. (2010) have proposed an empirical verification of the NEG model, focusing on the distinction between skilled and unskilled workers. This paper departs from the standard literature that analyses the relationship between wage and market potential, since it investigates how market access affects the wage distribution in U.S. metropolitan areas, highlighting the importance of pecuniary externalities due to backward and forward linkages. Their interest moves from the asymmetric effect of market access on wage inequality. They compare skilled to unskilled wages looking at the 90^{th},$50^{th}$and 10^{th} percentiles of the wage distribution. Contrary to the relation between average wage and market access, the authors find that the greater the MP, the greater the wage gap between the wages of the most and least skilled workers and between the wages of the medium and the least skilled workers. In their empirical model. they include in the LHS the difference of log wages in the percentiles considered ($90^{th} - 10^{th}$ and $50^{th} - 10^{th}$). The wage gap is explained by market potential, agglomeration, human capital, industry composition and amenities. The measure of Market Potential used is the original measure proposed by Harris (1954), since they assume that transport costs are negligible and price indices across metropolitan areas are constant. They have shown, in agreement with the predictions that the average wage is higher in metropolitan areas, but also that wage differentials are unequally distributed and higher MP is associated with lower outcomes at the bottom of the wage distribution.

The NEG theory is not free from criticisms. In particular, doubts have been raised about its degree of novelty and its ability to explain the real world better than the traditional model from which it borrows some assumptions. This has induced Fingleton (2005)and Fingleton (2007) to test the validity of the NEG model specification, comparing it with traditional models, such as the neoclassical model and the urban economy model, in order to understand which specification better fits the stylized facts about wage disparities and which models can better predict and explain regional convergence.

Fingleton (2005) compares the neoclassical with the NEG model in their ability to explain the globalization and income disparity.The NEG models are represented by the wage equation in which the explanatory variables are MP and the two dummy variables that capture the education level and workers experience. The specification of the neoclassical model is the same except that the MP is replaced with the growth rate of technology and population and the depreciation rate.

Fingleton (2007), further, compares the urban economic model with the NEG model; the two specifications differ from each other in that the former includes employment density whereas the latter includes MP. In both these works the econometric approach used involves the application of a bootstrap J-test. Since the two models are not nested , these two specifications are be compared by performing an Artificial Nested Model. The bootstrap J-tests suggests, in the first case, that only one of the two compared non-nested hypotheses should be rejected, the NEG model rejects the neoclassical model, but the neoclassical model fails to reject the NEG model. This leads to the conclusion that the NEG model specification is superior. Results indicate that the NEG model fits the data as well as the neoclassical model, but it yields the advantage of introducing the assumption of increasing returns to scale, so relevant in geographical economics and regional science. In the other paper, instead, the bootstrapping J-test can not discharge one specification completely.

Other criticisms come from Martin (1999) and Neary (2001). Martin (1999), who is a geographer, states that NEG is, first of all, an economic model and so linked to the social and political aspects of the economy. This could compromise the mathematical formalism claimed by the NEG . Martin, also, states that, while geographers follow the traditional approach incorporating the natural landscape of the countries and regions analysed, economists try to introduce a conceptual model where the space has more economic characteristics than geographical ones. In Martin's critique of the NEG models, he mainly refers to the theoretical works, while he accepts the works that include the quantitative investigation of the spatial dimension. Another critique has been originated by Neary (2001), who stresses that NEG models are not suitable to explain the dynamics of industrial location, since they do not consider the possibility of strategic interactions among firms, vertical integration or horizontal mergers. Considering the benefits of localization only due to the market-externalities and not to the technological externalities makes NEG models less representative of the real economy. Furthermore, he has argued that the most suitable market structure is the oligopoly, because it allows the firms to build barriers to entry and earn positive profits.

4.3 The Theoretical model

Although the NEG theory was firstly formalised only in Fujita et al. (1999), we can consider Fujita (1988), Krugman (1991) and Venables (1996) as the pioneering papers of NEG. Initially, they have analysed different aspects of the theory; in particular, Krugman approached this theory from the international trade point of view. He considered the effects of increasing European integration for factor mobility and realized that his work was departing from the traditional international trade assumptions. Indeed, while he kept the assumption of perfectly mobile factors, he assumed the existence of transport costs, an assumption more closely related to the classical location theory than to international trade theory. Fujita, instead, followed Von Thünen's work. He tried to explain the structure of modern cities through endogenous forces, internal forces that are not generated by the competitive markets, but from market structures such as oligopoly. Firms enjoy some advantages by locating close to households and vice versa, but agglomeration can occur because of the price interaction mechanism and may not be driven by the externalities. Venables has centred his studies on the investigation of forward and backward linkages between firms linked by an input-output structure, showing that agglomeration phenomena are possible without labour mobility. As mentioned in the introduction, some of the assumptions underlying the model are borrowed from the international trade theory and from traditional regional and urban economics. Traditionally, international trade theories focus on the concepts of comparative advantages and specialization of production. These theories are based on the so-called "first nature" advantages determined by the endowment of natural resources, the proximity to channels of communication (such as ports and canals). The NEG theory, on the contrary, focuses the attention on the notion of "second nature" advantages, grounding on human relations and circular causation that affects the localization choice. Firms prefer to localize close to large markets, and these generally are created by the clustering of economic agents (firms and households). Agglomeration produces externalities, through the interaction of transport costs, increasing returns to scale and factors' mobility. Despite the criticisms of the scarce originality of the assumptions of NEG models, it has to be stressed that the NEG theory provides a consistent framework combining centrifugal and centripetal forces, within a general equilibrium framework with micro-foundations. These forces are generated by the proximity of firms to other firms or workers; firms close to other firms can enjoy a higher demand from the other firms and the availability of cheaper intermediate goods, but the proximity can generate also a strong competition for cheap inputs and labour. The proximity of firms to workers determines high demand from consumers and weak competition for labour. Workers take advantages from the higher demand for labour and the availability of cheap consumer goods, but they face a high competition for vacancies. The equilibrium is generated by the balance of these forces working in opposite directions, but the implementation of a general equilibrium model is possible by sacrificing some generality of the Traditional Arrow-Debreau model and adopting the model proposed by Dixit and Stiglitz (1977). In agreement with Dixit and Stiglitz, NEG models assume monopolistic competition, increasing returns to scale and utility function characterized by constant elasticity of substitution.

4.3.1 Dixit-Stiglitz Model of Monopolistic Competition

The monopolistic competition model proposed by Dixit and Stiglitz relax the assumptions of perfect competition and constant returns to scale, fundamental in the neoclassical framework, allowing to explain the uneven distribution of the economic activities and population.

The Dixit-Stiglitz Model within the economic geography framework is based on the following assumptions. There exist two sectors: Agriculture (A) exhibiting constant returns to scale and Manufacturing (M) characterized by increasing returns to scale. This model, according to the micro-founded general equilibrium model, define the behaviour of the representative agents. It is assumed that consumer exhibits a Cobb-Douglas utility function

$$U = M^{\mu} A^{1-\mu}$$

where M is a composite consumption index of the manufacturing goods, defined through an utility function on the manufacturing variety of goods, M. M is defined by a CES utility function:

$$M = \left[\int_0^n m(i)^{\rho} di \right]^{1/\rho}$$

where ρ indicates the intensity of preference for variety in manufactured goods, in particular, when $\rho \to 1$ differentiated goods are closely perfect substitutes. Instead, when $\rho \to 0$, differentiated goods are not perfect substitutes and there is a preference for greater variety. The elasticity of substitution is related to the values assumed by the parameter ρ, $\sigma = 1/(1-\rho)$, when $\rho \to 1$, the elasticity of substitution assumes values greater than one and it will assume values close to one when $\rho \to 0$, meaning that goods are not perfect substitutes, and each location (area) will specialize in a single type of good, this implies that varieties produced by the manufacturing sector are differentiated. It is further assumed that production of a particular variety requires a fixed input cost. This implies that it is not possible for firms in a region to produce all the varieties and that each firm will enjoy economies of scale and increasing returns to scale in the production of that variety. Both imply monopolistic power over that variety. Assuming that the other factors are constant, in particular, keeping constant transport costs, this result encourages firms agglomeration.

Indicating income by Y, the price in the agriculture sector by p^A and the price of manufacturing goods by $p(i)$, the main goal of consumers is to maximize the utility function subject to the budget constraint:

$$p^A A + \int_0^n m(i)p(i)di = Y$$

Please refer to appendix A for the mathematical solution of the maximization problem. This can be solved in two steps; the first involves the minimisation of a cost function of manufacturing goods, M, and allows obtaining the optimal demand for each variety of manufacturing goods, the second step involves the maximization of the consumer utility obtaining the optimal share of income spent respectively in agriculturing and manufacturing goods. The first step leads to obtain the optimal demand for variety j among manufacturing goods:

$$m(j) = \left(\frac{p(j)}{G}\right)^{1/(\rho-1)}.$$

Recalling that $\frac{1}{1-\rho} = \sigma$, the previous can be rewritten as:

$$M = \left(\frac{p(j)}{G}\right)^{-\sigma} M,$$

where the manufacturing price index is $G = \left[\int_0^n [p(i)]^{1-\sigma} di\right]^{1/(1-\sigma)}$. The shares of income spent in manufacturing and agriculture sector are:

$$M = \frac{\mu Y}{G} \qquad A = \frac{(1-\mu)Y}{p^A}$$

Combining the previous results, we can obtain the demand for A and M as:

- $A = \frac{(1-\mu)Y}{p^A}$

- $m(j) = \mu Y \left(\frac{p(j)^{-\sigma}}{G^{-(\sigma-1)}} \right)$

Another important assumption underlying the Dixit- Stigltz model is that each variety is produced in a different location and that all the varieties are characterized by same technology and price, i.e. varieties are symmetric.

Let's assume that the number of locations is R, and the total number of varieties produced in a given location r will be indexed by n_r. A further assumption is related to the transport costs. Both sectors, Agriculture and Manufacturing, face "iceberg" transport costs (this assumption is removed for agriculture sector in the Krugman model, where the transport of agriculture goods is costless)

This type of transport costs are based on the idea that if a variety is transferred from location r to location s, only a fraction $1/T_{rs}^A$, $1/T_{rs}^M$ of the initial quantity will be delivered and a portion will "melt " along the way.

In order to satisfy the demand, the quantity of goods shipped should be greater than the amount requested, in order to take into account the percentage of goods melt away, T_{rs}^A e T_{rs}^M represent the quantity of good sent for a unit of good received.

The costly trade of varieties produced in a region and sold in the others has effects on the price of these goods. For example, a variety produced in region r and sold at price p_r will have a price $p_{rs} > p_r$ when that variety is delivered to region s.

$$p_{rs} = p_r T_{rs}$$

The assumption of "iceberg transport costs" together with the assumption that all the varieties produced in the same region have the same price, implies that the price index $G = \left[\int_0^n [p(i)]^{1-\sigma} di\right]^{1/(1-\sigma)}$ can be rewritten as:

$$G_s = \left[\sum_{r=1}^{R} n_r (p_r T_{rs})^{1-\sigma} di\right]^{1/(1-\sigma)} \qquad s = 1, 2, ..., R.$$

The demand of region s for a good of region r is

$$m(j) = \mu Y_s (p_r T_{rs})^{-\sigma} G_s^{(\sigma-1)}.$$

The total sales of variety r can be obtained aggregating the amount sold in all the regions

$$q_r = \mu Y_s (p_r T_{rs})^{-\sigma} G_s^{(\sigma-1)} T_{rs} \tag{4.1}$$

As well as the representative consumer, also the representative producer aims to maximize its objective function, i.e. the profit function.

The assumptions related to the production of manufacturing goods imply that they are produced by employing the same technology in all regions. This is characterized by marginal inputs c^M e fixed costs F. Since the only input factor is labour, we indicate with l^M the amount of input required to produce the quantity q^M, i.e.:

$$l^M = F + c^M q^M. \tag{4.2}$$

Each variety is produced in a single region and by only one firm that will specialize; the effect is that the number of firms and varieties will be the same.

In this model, a firm produces a specific variety in a given region r bearing a cost for the input factor, wage w_r^M, the production price is p_r^M, then the profit function will be:

$$\pi_r = p_r^M q_r^M - w_r^M l^M$$

and substituting equation (4.2), we obtain the following profit equation:

$$\pi_r = p_r^M q_r^M - w_r^M (F + c^M q^M) \tag{4.3}$$

Each producer takes the price index G_s of manufacturing goods as given, but given monopolistic behavior, he can choose the price (related to the specific variety produced) that maximizes profits. Profit maximization leads to the following mark-up pricing rule:

$$p_r^M \left(1 - \frac{1}{\sigma}\right) = c^M w_r^M,$$

or

$$p_r^M = \frac{c^M w_r^M}{\rho}.$$

Given the pricing rule, the profits of firms in region r are:

$$\pi_r = \frac{c^M w_r^M}{\rho} \cdot q_r^M - w_r^M (F + c^M q^M),$$

or

$$\pi_r = w_r^M \left[\frac{q_r^M c^M}{\sigma - 1} - F\right].$$

Equilibrium is obtained assuming the zero profit condition, due to firm entry and exit in the long run, which implies the following equations:

$$q^* = F,$$

$$w_r^M \left[\frac{q_r^M c^M}{\sigma - 1} - F\right] = 0,$$

$$\left[q_r^M\right] = w_r^M F(\sigma - 1)\frac{1}{w_r^M c^M},$$

$$q_r^M = F(\sigma - 1)\frac{1}{c^M} = q^*,$$

$$l^* = F + \frac{c^M F(\sigma - 1)}{c^M} = F(1 + \sigma - 1) = F\sigma.$$

l^* and q^* are constant for each firm. The number of manufacturing firms in a given region can be expressed as the ratio of the number of workers in the manufacturing sector in a given region r , L_r^M over the equilibrium workers and/or the fixed cost times the elasticity of substitution.

$$n_r = L_r^M / l^* = L_r^M / F\sigma$$

This result suggests that the market size will not affect neither the markup or the scale economies, but the scale effects operated through the change in variety of produced goods. This is a consequence of the assumptions of constant elasticity of the demand function and that firms are price takers. In particular, the latter assumption implies that firms do not have strategic behaviour. Otherwise, firms realize that their decisions have effects on the price index and that the market size (power) leads to a decrease in the output produced and an increase in prices.

The zero profit condition leads to the equilibrium quantity produced by each firm q^*. This amount has to be equal to the aggregate demand, as expressed in equation (4.1) :

$$q^* = \mu \sum_{s=1}^{R} Y_s (p_r^M)^{-\sigma} (T_{rs}^M)^{1-\sigma} G_s^{\sigma-1}.$$

Rewriting the inverse demand function:

$$(p_r^M)^{\sigma} = \frac{\mu}{q^*} \sum_{s=1}^{R} Y_s (T_{rs}^M)^{1-\sigma} G_s^{\sigma-1},$$

using the mark-up pricing rule:

$$w_r^M c^M = p_r^M \left[1 - \frac{1}{\sigma} \right],$$

$$w_r^M c^M [\frac{\sigma}{\sigma-1}] = p_r^M,$$

and substituting for p_r^M, it is possible to write the following wage equation:

$$w_r^M = \left(\frac{\sigma-1}{\sigma c^M}\right)\left[\frac{\mu}{q^*}\sum_{s=1}^{R} Y_s (T_{rs}^M)^{1-\sigma} G_s^{\sigma-1}\right]^{1/\sigma}.$$

This expression represents the break-even wage for each firm in the manufacturing sector. The higher are wages, the higher incomes (profits) of firms and the better the market access (due to less competition).

The real wage can be obtained deflating nominal wages by the cost of living index, that is determined by accounting for the price level in both the economic sectors present in the geographical unit considered, that is manufacturing and agriculture, $G_r^{\mu}(p_r^A)^{1-\mu}$,

$$\omega_r^M = w_r^M G_r^{-\mu}(p_r^A)^{-(1-\mu)}$$

The price index and the wage equation can be simplified by a normalization process. In particular FKV proceed to normalize assuming that:

$$c^M = \frac{\sigma-1}{\sigma} = \rho$$

which implies that the price rule becomes:

$$p_r^M = w_r^M c^M [\frac{\sigma}{\sigma-1}],$$

$$p_r^M = w_r^M,$$

and so:

$$l^* = q^*.$$

Some simplifying assumptions have been introduced, such that fixed inputs satisfy the relation:

$$F = \frac{\mu}{\sigma},$$

and that the number of firms belongs to a numerical range in the real interval [0, n], and that number of firms, for each location, is related to the size of labour force, and given by the ratio between manufacturing workers and elasticity of substitution:

$$n_r = L_r^M / \mu$$

The output for each firm will be

$$q^* = l^* = \mu.$$

Using the previous assumption, price and wage equations may be, respectively, rewritten as:

$$G_r = \left[\sum_{s=1}^{R} n_r (p_r^M T_{rs}^M)^{1-\sigma} \right]^{1/(1-\sigma)} = \left[\frac{1}{\mu} \sum_{s=1}^{R} L_s^M (p_r^M T_{rs}^M)^{1-\sigma} \right]^{1/(1-\sigma)}$$

$$w_r^M = \left(\frac{\sigma - 1}{\sigma c^M} \right) \left[\frac{\mu}{q^*} \sum_{s=1}^{R} Y_s (T_{rs}^M)^{1-\sigma} G_s^{\sigma-1} \right]^{1/\sigma}$$

$$w_r^M = \left[\sum_{s=1}^{R} Y_s (T_{rs}^M)^{1-\sigma} G_s^{\sigma-1} \right]^{1/\sigma} \tag{4.4}$$

Under the Core-Periphery model, the equilibrium conditions are affected by migration phenomena and so by the change of manufacturing workers share. It is worth to remember these effects that are respectively: Price Index Effect and Home Market Effect , the former is related to the effect that change in labour (workers) inputs can cause on price index, the latter, instead, measures the effects of change in workers on wage. The mathematical investigation of these effects, as reported in Appendix B2, leads us to conclude that change in the employment in the manufacturing sector $\frac{dL}{L}$ has a negative "price index effect". This result is obtained assuming that labour supply is perfectly elastic, implying that $dw = 0$ and remembering that $\sigma > 1$ and $1 - \sigma < 0$ e $T > 1$. Further, regions with a larger manufacturing sector have a lower price index because a smaller share of consumption will require transport costs.

In order to investigate the "Home market effect" we need to decompose the change in demand as the change in price index, wage and workers. Home market effect allows to investigate how firms localization and the change in workers can be affected by the relative demand. This relation can be expressed as follow:

$$\left[\frac{\sigma}{Z} + Z(1-\sigma)\right]\frac{dw}{w} + Z\frac{dL}{L} = \frac{dY}{Y}$$

where $Z = \frac{1-T^{1-\sigma}}{1+T^{1-\sigma}}$ represents a transport cost index, its range of values varies between 0, when trading is costless, and 1, when trading is not workable. As already done early, we assume perfectly elastic labour supply and so $dw = 0$, this assumption reduces the "home market effect" to a relation between change in demand and change in manufacturing workers, an increase of 1% of the manufacturing demand dY/Y causes an increase equal to $1/Z$ (>1) of the employment in the manufacturing sector. This relation suggests that the locations with a greater home market, they will have a larger manufacturing sector and (due the the assumption $dw = 0$) the increase in the demand will be traduced in an increase in export.

Relaxing the assumption of perfectly elastic labour supply , $dw = 0$ and assuming that the labour supply is a positively sloped curve some advantages of the home market will reflect on wages instead of export.

An increase in employment L causes a decrease in the price index, the location with a larger manufacturing demand will bid higher real wage, since price index will relatively lower.

Turning the Price Index Effect, we assume that there exist two symmetric regions (region 1 and region 2), remembering, further, that normalization process lead to the equality between price and marginal cost. The price index of each region can be written as:

$$G_1^{(1-\sigma)} = \frac{1}{\mu}\left[L_1 w_1^{1-\sigma} + L_2\left(w_2 T\right)^{1-\sigma}\right]$$

$$G_2^{(1-\sigma)} = \frac{1}{\mu}\left[L_1 (w_1 T)^{1-\sigma} + L_2 w_2^{1-\sigma}\right]$$

and the wage equations become:

$$w_1^{\sigma} = Y_1 G_1^{\sigma-1} + Y_2 G_2^{\sigma-1} T^{1-\sigma}$$

$$w_2^{\sigma} = Y_1 G_1^{\sigma-1} T^{1-\sigma} + Y_2 G_2^{\sigma-1}$$

The relation between price index and wage can be investigated through a linearisation around the symmetric equilibrium. We need to analyse these equations separately, but symmetry allows us to consider only one region. We start from the index price in the region 1 (which is the same of region 2):

$$G_1^{(1-\sigma)} = \frac{1}{\mu}\left[L_1 w_1^{1-\sigma} + L_2\left(w_2 T\right)^{1-\sigma}\right]$$

totally differentiating the previous equation:

$$(1-\sigma) G_1^{-\sigma} dG_1 = \frac{1}{\mu}\left[L_1 (1-\sigma) w_1^{-\sigma} dw_1 + L_2 T^{1-\sigma} (1-\sigma) w_2^{-\sigma} dw_2\right] + \frac{1}{\mu}\left[w_1^{1-\sigma} dL_1 + \left(w_2 T\right)^{1-\sigma} dL_2\right]$$

and keeping in mind that the symmetric equilibrium implies that each change of a variable in the region 1 coincide with an opposite change in the same variable in region 2, of the same magnitude, $dw_1 = -dw_2$ and $dL_1 = -dL_2$

$$(1-\sigma)G_1^{-\sigma}dG_1 = \frac{L}{\mu}\left[(1-\sigma)w_1^{-\sigma}dw_1 - T^{1-\sigma}(1-\sigma)w_2^{-\sigma}dw_1\right] + \frac{1}{\mu}\left[w_1^{1-\sigma}dL_1 - (w_2T)^{1-\sigma}dL_1\right]$$

substituting the symmetric equilibrium conditions: $w_1 = w_2 = w$, $G_1 = G_2 = G$, and $L_1 = L_2$, we obtain:

$$(1-\sigma)G^{-\sigma}dG = \frac{L}{\mu}\left[(1-\sigma)w^{-\sigma}dw - T^{1-\sigma}(1-\sigma)w^{-\sigma}dw\right] + \frac{1}{\mu}\left[w^{1-\sigma}dL - w^{1-\sigma}T^{1-\sigma}dL\right]$$

$$(1-\sigma)G^{-\sigma}dG = \frac{L}{\mu}(1-\sigma)w^{1-\sigma}\frac{dw}{w}\left[1-T^{1-\sigma}\right] + \frac{1}{\mu}w^{1-\sigma}\left[1-T^{1-\sigma}\right]dL$$

$$(1-\sigma)\frac{dG}{G} = \frac{L}{\mu}\frac{w^{1-\sigma}}{G^{1-\sigma}}\left[1-T^{1-\sigma}\right]\left((1-\sigma)\frac{dw}{w} + \frac{dL}{L}\right)$$

$$(1-\sigma)\frac{dG}{G} = \frac{L}{\mu}\left(\frac{G}{w}\right)^{\sigma-1}\left[1-T^{1-\sigma}\right]\left((1-\sigma)\frac{dw}{w} + \frac{dL}{L}\right) \tag{4.5}$$

$$\frac{dG}{G} = \frac{1}{(1-\sigma)}\frac{L}{\mu}\left(\frac{G}{w}\right)^{\sigma-1}\left[1-T^{1-\sigma}\right]\left((1-\sigma)\frac{dw}{w} + \frac{dL}{L}\right)$$

The same procedure has been followed for the nominal wage of region 1:

$$w_1^{\sigma} = Y_1G_1^{\sigma-1} + Y_2G_2^{\sigma-1}T^{1-\sigma}$$

totally differentiating:

$$\sigma w_1^{\sigma-1} dw_1 = (\sigma-1) Y_1 G_1^{\sigma-1} G_1^{-1} dG_1 + G_1^{\sigma-1} dY_1 + (\sigma-1) Y_2 G_2^{\sigma-1} G_2^{-1} T^{1-\sigma} dG_2 + G_2^{\sigma-1} T^{1-\sigma} dY_2$$

and recalling that $dY = dY_1 = -dY_2$ e $dG = dG_1 = -dG_2$

$$\sigma w_1^{\sigma-1} dw_1 = (\sigma-1) Y_1 G_1^{\sigma-1} G_1^{-1} dG_1 + G_1^{\sigma-1} dY_1 - (\sigma-1) Y_2 G_2^{\sigma-1} G_2^{-1} T^{1-\sigma} dG_1 - G_2^{\sigma-1} T^{1-\sigma} dY_1$$

further, the symmetric equilibrium, allows to rewrite a simpler form of the previous equation because of $w_1 = w_2 = w$, $G_1 = G_2 = G$ and$Y_1 = Y_2 = Y$

$$\sigma w^{\sigma-1} dw = (\sigma-1) Y G^{\sigma-1} G^{-1} dG + G^{\sigma-1} dY - (\sigma-1) Y G^{\sigma-1} G^{-1} T^{1-\sigma} dG - G^{\sigma-1} T^{1-\sigma} dY$$

$$\sigma w^{\sigma-1} dw = (\sigma-1) Y G^{\sigma-1} G^{-1} dG \left[1 - T^{1-\sigma}\right] + G^{\sigma-1} dY \left[1 - T^{1-\sigma}\right]$$

$$\sigma w^{\sigma-1} dw = Y G^{\sigma-1} \left[1 - T^{1-\sigma}\right] \left[(\sigma-1) \frac{dG}{G} + \frac{dY}{Y}\right]$$

$$\sigma dw = Y \cdot \frac{G^{\sigma-1}}{w^{\sigma-1}} \left[1 - T^{1-\sigma}\right] \left[(\sigma-1) \frac{dG}{G} + \frac{dY}{Y}\right]$$

$$\sigma \frac{dw}{w} = \frac{Y}{w} \left(\frac{G}{w}\right)^{\sigma-1} \left[1 - T^{1-\sigma}\right] \left[\frac{dY}{Y} + (\sigma-1) \frac{dG}{G}\right] \qquad (4.6)$$

$$\frac{dw}{w} = \frac{1}{\sigma}\frac{Y}{w}\left(\frac{G}{w}\right)^{\sigma-1}\left[1 - T^{1-\sigma}\right]\left[\frac{dY}{Y} + (\sigma - 1)\frac{dG}{G}\right]$$

Consider now, the behaviour of the real wage in the case of autarky, when $Z = 1$, associated to high trade costs that yield trade impossible to realize and assuming the expenditure is constant ($dY = 0$). The effects that changes in workers can cause on real wage is twofold, we can observe these effects by totally differentiating real wage equation, given by the combination of the nominal wage (4.6)and price index (4.5):

$$\frac{d\omega}{\omega} = \frac{dw}{w} - \mu\frac{dG}{G}$$

$$\frac{d\omega}{\omega} = (1 - \mu)\frac{dY}{Y} + \left[\frac{\mu - \rho}{\rho}\right]\frac{dL}{L}$$

Firstly, the increase in the number of workers causes a decrease in average real wage, this effect is due to the assumption of constant expenditure that can be interpreted as constant total wage (that represents a cost for firms). Secondly, the greater is the number of workers the larger will be the varieties of manufacturing goods implying a lower price's index that will lead to increase real wages.

If the second effect (increase of real wage due to an increase in the number of workers) prevail on the first effect (an increase of the number of workers lead to a decrease in the individual wage), the net effect is an increase in the real wages.

In order to avoid that economy collapse in a single point, the black hole location, caused by the positive causation arising by the increase in real wage that appeal workers in the location with high real wage, determining agglomeration, we need to assume that the "no-black holes" holds, that is there exist a negative relation among workers and real wage, given by:

$$\frac{\sigma - 1}{\sigma} = \rho > \mu,$$

so $(\mu - \rho)$ is negative and a positive change in L will imply a negative change in real wage.

4.3.2 Core - Periphery

The New Economic Geography framework arises from developments of traditional model of Regional and Urban Economics but in a framework of General Equilibrium Model. The General Equilibrium Model determined by the Dixit-Stiglitz model as outlined in the previous section. Now, we proceed to outline the Regional Economics Framework borrowed by the New Economic Geography framework, that is the Core-Periphery model.

The main assumptions that underlying this model are:

- there are two regions;
- the economy is constituted by two sectors: Manufacturing (M), characterized by monopolistic competition and Agriculture (A) characterized by perfect competition;
- every sector employs only one input factor, labour, factor supply is defined a priori and is fixed;
- there exist L^A farmers and each region disposes of a fixed portion of the total farmer in the economy ϕ_r;
- manufacturing labour force is characterized by a high mobility and the share of workers in a given region is indicated by λ_r;
- the total number of workers is normalized to one , such that $L^A = 1 - \mu$ e $L^M = \mu$
- manufacturing goods are traded bearing "iceberg transport cost" , that implies that when one units of a good produced in region r and sent to region s, only the quantity $1/T_{rs}$will be delivered.
- trade of agriculture goods is costless, the technology employed exhibits constant return to scale, this implies that wage is equal to the marginal product and that is equal for each farmer
- agriculture wage is assumed as works and it is equal to one ,$w_r^A = 1$

- manufacturing wage may be expressed as nominal or real wage respectively indicated as w_r^M e ω_r^M.
- manufacturing workers mobility is determined by the real wage, workers will move toward regions with higher real wage.
- migration dynamic is the following $\dot{\lambda} = \gamma(\omega_r - \bar{\omega})\lambda_r$.
- the size of the manufacturing sector across regions evolves at any time due to the migration.

In order to determine equilibrium solution it is assumed that there exist two regions, for each one four equations represented by:

- regions' income;
- price's index of manufacturing good;
- nominal wage rate;
- real wage.

Income equation is determined taking into consideration the assumptions that only manufacturing sector transport costs are positive and agriculture workers receive a wage equal to one. We have further assumed that $L^A = 1 - \mu$ and $L^M = \mu$. Symmetric assumption related to the region allows to write the income equation for a generic region r as:

$$Y_r = \mu\lambda_r w_r^M + (1-\mu)\phi_r w_r^A = \mu\lambda_r w_r^M + (1-\mu)\phi_r$$

where $w_r^A = 1$

The assumption that manufacturing workers in the location s is $L_s^M = \mu\lambda_s$, leads to rewrite the price index equation [4] as:

[4] $G_r = \left[\sum_{s=1}^{R} n_r (p_r^M T_{rs}^M)^{1-\sigma}\right]^{1/(1-\sigma)}$

$$G_r = \left[\sum_{s=1}^{R} \lambda_s (w_s T_{rs})^{1-\sigma} \right]^{1/(1-\sigma)}$$

the price index of region r is lower the higher is the number of varieties in the region. The wage rate that breaks-even the manufacturing sector in region r is

$$w_r = \left[\sum_{s=1}^{R} Y_s (T_{rs})^{1-\sigma} G_s^{\sigma-1} \right]^{1/\sigma}$$

Assuming that the price index across regions is the same, nominal wage will be higher the higher is regional income, and the lower the transport costs from region r. Regions will pay higher wage the higher is the market access. Equilibrium solution is affected by consumers/workers behaviour driven by the their purchasing power, and so by the real wage:

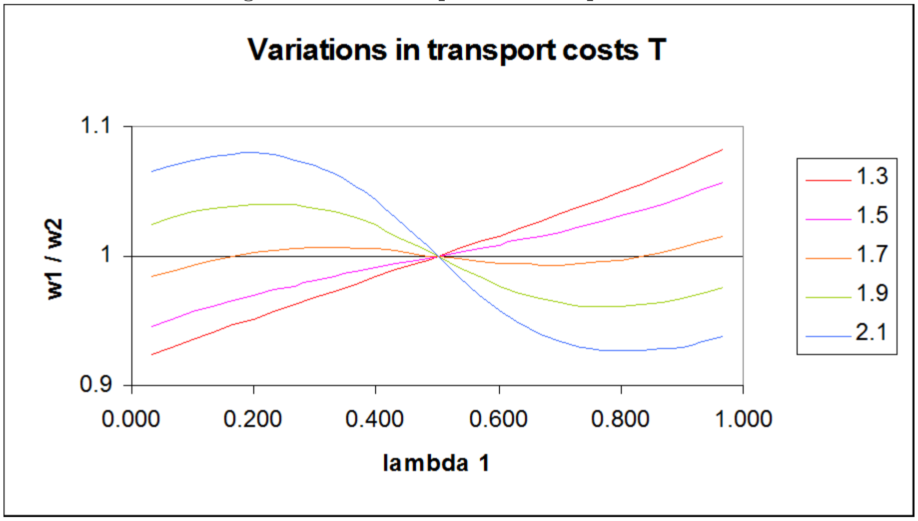
$$\omega_r = w_r G_r^{-\mu}$$

Equilibrium is obtained solving simultaneous equations related to income, price index , nominal and real wage.

This model is not analytically solvable without imposing specific assumptions, such as, limiting the economy to two regions and that industry exhibiting constant returns to scale is evenly distributed between regions; indeed three scenario are possible for the manufacturing sector: even the M sector is evenly distributed between regions, M sector is concentrated in the region 1 and A sector is localized in region 2, the latter two scenario coincide with core-periphery model, where manufacturing is localized in the core and agriculture in the periphery.

In order to show equilibrium solutions is often used numerical and simplified example, in particular following Fujita et al. (1999)we set the simultaneous equations of a simplified Core-Periphery model, assuming that there are two regions and two economic sectors, manufacturing and agriculture. Agriculture is evenly distributed between the two regions, so that the share of agriculture workers is 1/2. Further we indicate by λ the share of workers in the manufacturing sector in region 1 and $(1 - \lambda)$ the share of workers in region 2. Equilibrium is obtained solving the simultaneous

Figure 4.1: The impact of transport costs



Brakman, Garretsen and van Marrewijk (2009)

equations of income, price index, wage rate and real wage, related to the two regions.[5]

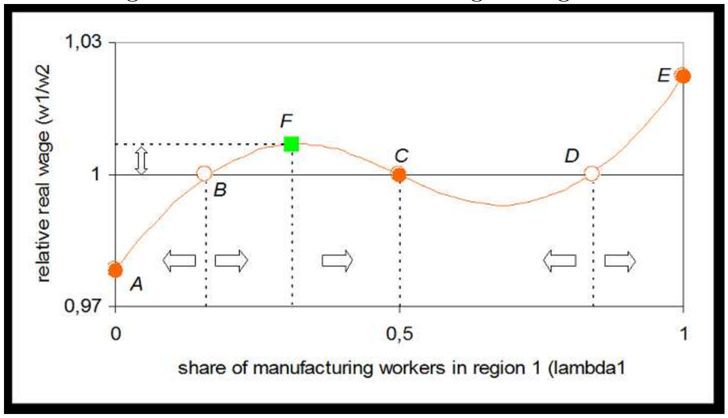
This model can be solved through simulations, assigning values to parameters. Simulation results are usually shown graphically to ease their interpretation. One of the most interesting graphs shows the relation between relative real wages and λ, share of manufacturing workers in region 1, we know that migration of manufacturing workers and so the change in the share of these workers in a region affect the real wage and so the equilibrium.

We report graphs from Brakman et al. (2009). The first one gives us an overall picture for possible values of T ($T = 2.1,\ 1.75,\ 1.5$)

If $T = 2.1$, the difference in real wages ($\omega_1 - \omega_2$) is positive when $\lambda < 0.5$ and negative when values are greater. The only possible equilibrium is represented by $\lambda = 0.5$ and the difference in

[5] $Y_1 = \mu\lambda w_1 + (1-\mu)\frac{1}{2}$
$Y_2 = \mu(1-\lambda)w_2 + (1-\mu)\frac{1}{2}$
$G_1 = \left[\lambda w_1^{1-\sigma} + (1-\lambda)(w_2 T)^{1-\sigma}\right]^{1/(1-\sigma)}$
$G_2 = \left[\lambda (w_1 T)^{1-\sigma} + (1-\lambda)w_2^{1-\sigma}\right]^{1/(1-\sigma)}$
$w_1 = \left[Y_1 G_1^{\sigma-1} + Y_2 G_2^{\sigma-1} T^{1-\sigma}\right]^{1/\sigma}$
$w_2 = \left[Y_1 (G_1 T)^{1-\sigma} + Y_2 G_2^{\sigma-1}\right]^{1/\sigma}$
$\omega_1 = w_1 G_1^{-\mu}$
$\omega_2 = w_2 G_2^{-\mu}$

Figure 4.2: The relative real wage in region 1



Brakman, Garretsen and van Marrewijk (2009)

wages is equal to zero, that is, wages are equal on both regions, i.e. there is a long run equilibrium.

For $T = 1.5$, the relationship is positive, i.e. the greater the share of manufacturing sector in a region, the more appealing the region. The symmetric share of manufacturing sectors between two regions $\lambda = 0.5$ is, also in this case, a possible, but unstable, equilibrium . Indeed, if a region is characterized by a larger manufacturing sector, the sector should become larger and larger until the sector is concentrated in only one region, leading to a pattern of economic activities of the core-periphery type.

Finally, the case of $T = 1.75$, represented by an intermediate level of transport costs, leads to a more complicated situation, because more than one equilibrium is possible. This case is shown in figure 4.2.

The symmetric equilibrium is again a possible and locally stable equilibrium, together with two unstable equilibria. In order to determine the equilibrium, it is important to know the initial condition of λ. Indeed, a relatively high or small initial level could yield different equilibrium, respectively symmetric and/ or core-periphery pattern.

Once the equilibria are identified, the next step consists in verifying their stability and sustainability, characteristics that allow to distinguish between short run and long run equilibrium. The latter is achieved when real wages are the same across regions, so that workers have no incentive to migrate.

In order to investigate the features the of equilibrium points we need to compare real wages in the two regions. If $\omega_1 \geq \omega_2$, workers in region 1 may have an incentive to migrate towards region 2.

Further, this relationship will have a different shape depending on the level of transport costs, affecting the the equilibrium points.

It could be useful to this goal proceeding with a numerical example such as in Fujita et al. (1999), it is assumed that nominal wage is equal to one, $w_1 = 1$, this assumption implies that both sectors pay the same wage rate, but the manufacturing sector produce a larger variety of goods, causing that price index will be lower in the region where the manufacturing sector is localized, and real wage will be greater, further the regional income will be higher than region 2.

Numerically, for$\lambda = 1$ and $w_1 = 1$, the model equations are the following:

$Y_1 = \mu\lambda w_1 + (1-\mu)\frac{1}{2} = \mu + (1-\mu)\frac{1}{2} = \frac{2\mu+1-\mu}{2} = \frac{\mu+1}{2}$

$Y_2 = \mu(1-\lambda)w_2 + (1-\mu)\frac{1}{2} = (1-\mu)\frac{1}{2}$

$G_1 = \left[\lambda w_1^{1-\sigma} + (1-\lambda)(w_2 T)^{1-\sigma}\right]^{1/(1-\sigma)} = 1$

$G_2 = \left[\lambda (w_1 T)^{1-\sigma} + (1-\lambda)w_2^{1-\sigma}\right]^{1/(1-\sigma)} = (T^{1-\sigma})^{1/(1-\sigma)} = T$

$w_1 = \left[Y_1 G_1^{\sigma-1} + Y_2 G_2^{\sigma-1} T^{1-\sigma}\right]^{1/\sigma}$.

Since $w_1 = 1$ by assumption and we obtain $G_1 = 1$, we can conclude that nominal and real wages coincide, so real wage $\omega_1 = 1$. Further, we need to determine the nominal and real wage of region 2, as the weighted average of income of both regions, where weights are represented by the transport costs. The nominal and real wage of region 2, respectively, are the following:

$$w_2 = \left[Y_1 (G_1 T)^{1-\sigma} + Y_2 G_2^{\sigma-1}\right]^{1/\sigma} = \left[\frac{\mu+1}{2} T^{1-\sigma} + (1-\mu)\frac{1}{2} T^{\sigma-1}\right]^{1/\sigma}$$

$$\omega_2 = T^{-\mu}\left[\frac{\mu+1}{2} T^{1-\sigma} + (1-\mu)\frac{1}{2} T^{\sigma-1}\right]^{1/\sigma}$$

If $T = 1$, so that there are no transport costs, the real wage of region 2 is equal to zero,$\omega_2 = 1$. Differentiating the real wage with respect to T and evaluating the derivative at point $T = 1$ and $\omega_2 = 1$, we obtain:

$$\frac{d\omega_2}{dT} = \frac{\mu(1-2\sigma)}{\sigma} < 0 \tag{4.7}$$

since $\sigma > 1$[6]

This result shows that for low transport costs, agglomeration equilibrium is sustainable because the real wage of region 1 is greater then region 2, $\omega_2 < 1 = \omega_1$.

Instead, for high transport costs (T high), we need to observe the behaviour of the two terms, which determine the real wage of region 2: $\frac{\mu+1}{2}T^{1-\sigma-\mu\sigma}$ and $\frac{(1-\mu)}{2}T^{\sigma-1-\mu\sigma}$ to define an equilibrium as stable and sustainable. Two scenarios are possible:

1. $\sigma - 1 - \mu\sigma < 0$ is very small, so that $T^{\sigma-1-\mu\sigma}$ is very small and the real wage goes to zero;
2. $\sigma - 1 - \mu\sigma > 0$ or $\frac{\sigma-1}{\sigma} < \mu$, so that agglomeration forces are stronger than the spreading ones, so the core-periphery will be an equilibrium.

Fujita et al. (1999) have shown that the symmetric equilibrium is a stable ones for high values of transport costs, but it is unstable when T is low, this is very intuitive, since, high transport costs make trade more expensive and producer may prefer to produce where goods will be sell, instead, low transport costs, will stimulate trade between regions, and so agglomeration forces may lead to more benefits.

They have determined the break-even point that coincides with the curve of difference in real wage $(\omega_1 - \omega_2)$ that is horizontal to the symmetric equilibrium. In order to determine the break-even point they proceed by totally differentiate the equilibrium system with respect to λ and find: $\frac{d(\omega_1-\omega_2)}{d\lambda}$.

[6] To obtain equation (4.7) these are the steps performed:

$\omega_2 = T^{-\mu}\left[\frac{\mu+1}{2}T^{1-\sigma} + (1-\mu)\frac{1}{2}T^{\sigma-1}\right]^{1/\sigma}$

$\frac{d\omega_2}{dT} = -\mu T^{-\mu-1}\left[\frac{\mu+1}{2}T^{1-\sigma} + (1-\mu)\frac{1}{2}T^{\sigma-1}\right]^{1/\sigma} +$

$+\frac{1}{\sigma}T^{-\mu}\left[\frac{\mu+1}{2}T^{1-\sigma} + (1-\mu)\frac{1}{2}T^{\sigma-1}\right]^{(1/\sigma)-1}\left[(1-\sigma)\frac{\mu+1}{2}T^{-\sigma} + (\sigma-1)(1-\mu)\frac{1}{2}T^{\sigma-2}\right] = 0$

Evaluating at the point $T = 1$ and $\omega_2 = 1$, obtaining:

$\frac{d\omega_2}{dT} = -\mu + \frac{1}{\sigma}\left[(1-\sigma)\frac{\mu+1}{2} + (\sigma-1)(1-\mu)\frac{1}{2}\right] = -\mu + \frac{1}{\sigma}\left[\frac{(1-\sigma)(\mu+1)+(\sigma-1)(1-\mu)}{2}\right] = \left[\frac{(1-\sigma)(\mu+1)+(\sigma-1)(1-\mu)-2\sigma\mu}{2\sigma}\right] =$

$\left[\frac{1-\sigma\mu-\sigma+\mu+\sigma-1+\mu-\sigma\mu-2\sigma\mu}{2\sigma}\right] = \left[\frac{2\mu-4\sigma\mu}{2\sigma}\right] = \left[\frac{\mu(1-2\sigma)}{\sigma}\right]$

$$\frac{d\omega_2}{dT} = \left[\frac{\mu(1-2\sigma)}{\sigma}\right] < 0$$

In order to differentiate real wage, ω, with respect to λ , it is necessary to proceed by totally differentiate wage rate, price index and income.

The total differential will be computed around the symmetric equilibrium, so the endogenous variables assume the following values:

where $w_1 = w_2 = 1$; $\lambda = 1/2$; $Y_1 = Y_2 = 1/2$; $G_1^{1-\sigma} = G_2^{1-\sigma} = \left[\frac{1+T^{1-\sigma}}{2}\right]$.[7]

The mathematical steps performed in order to obtain the relation between the real wage and the share of manufacturing workers are reported in detail in the appendix B3. The relation obtained is the following:

$$\frac{d\omega}{d\lambda} = 2ZG^{-\mu}\left(\frac{1-\rho}{\rho}\right)\left[\frac{\mu(1+\rho) - Z(\mu^2+\rho)}{1-\mu Z(1-\rho) - \rho Z^2}\right]$$

Recalling the no black hole condition which claims a negative relationship between real wage and the number of manufacturing workers, as consequence the symmetric equilibrium is stable if $\frac{d\omega}{d\lambda}$ is negative.

Since denominator is positive, the sign of the $\frac{d\omega}{d\lambda}$ depends on the numerator. When Z goes to zero, the numerator is positive and so the symmetric equilibrium is unstable, when Z is close to one, two are the possible scenario

- $\rho < \mu$, implying a positive numerator,
- $\rho > \mu$, implying negative numerator.

The symmetric equilibrium is a stable ones, and so the "no-black-hole" condition is satisfied, when $\rho > \mu$ and for sufficiently high transport costs.

When $T = 1$, that is, in the case of zero transport costs, inputs reallocation (labour) has no effects on the real wage , zero transport costs implies that regions are not distinguishable, intermediate values of T labour mobility causes an increase in the real wage of the region where

[7] $Y_1 = \mu\lambda w_1 + (1-\mu)\frac{1}{2} = \frac{1}{2}\mu + (1-\mu)\frac{1}{2} = \frac{1}{2}$
$Y_2 = \mu(1-\lambda)w_2 + (1-\mu)\frac{1}{2} = \mu(1-\frac{1}{2}) + (1-\mu)\frac{1}{2} = \frac{1}{2}$
$G_1 = \left[\lambda w_1^{1-\sigma} + (1-\lambda)(w_2 T)^{1-\sigma}\right]^{1/(1-\sigma)} = \left[\frac{1}{2} + \frac{1}{2}(T)^{1-\sigma}\right]^{1/(1-\sigma)} = \left[\frac{1+T^{1-\sigma}}{2}\right]^{1/(1-\sigma)}$
$G_2 = \left[\lambda(w_1 T)^{1-\sigma} + (1-\lambda)w_2^{1-\sigma}\right]^{1/(1-\sigma)} = \left[\frac{1}{2}(T)^{1-\sigma} + \frac{1}{2}\right]^{1/(1-\sigma)} = \left[\frac{1+T^{1-\sigma}}{2}\right]^{1/(1-\sigma)}$

labour move to, implying that $\frac{d\omega}{d\lambda} > 0$ and unstable symmetric equilibrium. Finally, high transport costs, with $T \to \infty$(that represents a condition of autarky), an increase on the number of workers in manufacturing sector decreases the real wage, indeed an increase in workers causes a increase in the supply of manufacturing varieties that now cannot be exported .

4.4 Empirical model

The empirical investigation of New Economic Geography model has a shorter tradition with respect the theoretical ones. Although researchers have analysed different aspects of the model, the equation that has received more attention is the wage equation, which represents the clearing condition of the labour market.

Wage equation relates average wage with market potential, Market potential or market access is a measure that allow to quantify the accessibility of a market; it is an index of the concentration of possible contracts within a market. It indicates the potential demand faced by firms located in a given region.

This concept has not been introduced by NEG theory, but it was proposed by Harries (1954). The empirical verification of NEG theory focuses on the MP; this implies that the transport costs and the measurement of the distance between regions become very important to calculate this variable. Later, in literature has been developed other definitions of MP, many empirical works have adopted the definition proposed by Helpman (1998). This alternative considers the land rent for urban housing as the centrifugal force, and differs from the previous definition which centers on local income. Another definition of MP is the definition adopted by FKV, that weights the transport costs by elasticity of substitution, this is the concept of MP that we use in our empirical verification.

We perform our analysis following the traditional NEG model as proposed by Fujita, Krugman and Venables (1999) that obtain the following wage equation:

$$W_i = \left[\sum_r Y_r (G^M)^{\sigma-1} (\bar{T}_{ir})^{1-\sigma} \right]^{1/\sigma} \tag{4.8}$$

where the RHS represents the market potential, σ denotes the elasticity of substitution, Y_r

indicates the regional income (regional expenditure), determined as weighted sum of manufacturing and agriculture wages in a given region; $(G^M)^{\sigma-1}$ indicates manufacturing price index and $(T_{ir})^{1-\sigma}$ are transport costs. In particular the variable Income is defined as follow:

$$Y_r = \mu\lambda_r w_r^M + (1-\mu)\phi_r w_r^A \tag{4.9}$$

where ϕ_r is the share of workers in the competitive sector (agriculture) in the region and μ and $(1-\mu)$ are respectively the share of Manufacturing and Agriculture workers normalizing the total number of worker to 1. Transport costs take the form of iceberg cost, in a variant of Samuelson (1952), proposed by Krugman (1995), and called "the delivered price of the good":

$$T_{ij} = e^{\tau \cdot ln(D_{ij})}$$

The basic assumption that underlying iceberg transport costs is that a certain quantity of goods is consumed during the shipment, while the remaining quantity is delivered. NEG models have adopted the Krugman's version of iceberg transport costs because the original version, and the concavity assumption, could imply unstable results. Krugman's version departs from the original version in three ways:

- The transport cost function is convex with respect to the distance;
- Transport cost is directly proportional to the original price of the good
- Transport rate per ton-kilometer is independent of the quantity shipped.

Although the convexity assumption allows overcoming the issue of unstable results, it is not supported by the empirical evidence that, instead, suggests a concave transport function. More precisely, Krugman has referred to the more general concept of trade costs, instead of the transport costs and it is worth noting that convexity is induced by the set of distance-transaction costs (like information costs, institutional barriers, cultural and linguistic differences), which are embodied in Krugman's iceberg costs.

4.5 Empirical application

4.5.1 Methodology

The empirical analysis concerns the estimation of the basic relation between wage and market potential. The equation to be estimated is the log of wage equation (4.10)

$$W_i = \left[\sum_r Y_r (G^M)^{\sigma-1} (\bar{T}_{ir})^{1-\sigma} \right]^{1/\sigma} \tag{4.10}$$

According to Fujita et al. (1999), the empirical literature has analysed the wage equation at regional level, at regional or more recently at province level. We aim to investigate this relation at micro level, in particular analysing the relationship between wage at firm level and MP at provincial level, departing from the common approach that investigate the relation between regional wage and regional MP.

We proceed to test our micro level model and then to substituting firms' wage with province wage, in other to compare results.

The choice of investing the effect of MP on firm level wage, allows to overcome the endogeneity problem faced by the traditional literature since wage enters in the computation of market potential, which at the same time constitutes the dependent variables, causing simultaneity.

According to Roos (2001), the endogeneity effect can be reduced through aggregating independent variables. If the dependent variable is measured at a finer geographic level than the independent variables, there are fewer states in which a shock to a lower level region seriously affects the aggregate of regions. We have dealt with a model that seems to reflect the suggestion of Roos, indeed we have used firm level data for the dependent variable, and we have combined the income and transport costs at province level to construct the independent variables.

The New Economic Geography wage equation can be rearranged to our model, with to level of data aggregation as follow:

$$W_i = f(MP_{p_i}^{1/\sigma}) = f\left(\left[\sum_p Y_p G_p^{\sigma-1} (\bar{T}_{p_i p,})^{1-\sigma} \right]^{\frac{1}{\sigma}} \right), \qquad p = 1, \ldots, p_i \ldots, P \tag{4.11}$$

where i indicated firms and p indicates province and p_i denotes the province where firm i is localized. The lack of data availability about price index at micro level led to assume that price index is the same across regions of a certain country, this assumption implies that the equation to be estimated is:

$$ln(W_i) = \frac{1}{\sigma} ln \left[\sum_p Y_p (\bar{T}_{p_i p,})^{1-\sigma} \right], \qquad p = 1, \ldots, p_i \ldots, P \tag{4.12}$$

$$ln(W_i) = \frac{1}{\sigma} ln \left[\sum_p Y_p e^{\tau(1-\sigma) d_{p_i,p}} \right], \qquad p = 1, \ldots, p_i \ldots, P \tag{4.13}$$

As we can observe the relationship between wage and MP is not linear, since parameters are at the exponent. The consequence is that the OLS estimator is not suitable for this model and we need to recur to non linear estimation. In particular, the literature suggests implementing the Non Linear Least Squares estimator.

Furthermore, since we proceed to estimating the NEG relation at the firm and province levels, in latter case we may incur in endogeneity issue, since wage, that represents our dependent variable, enters in the Market Potential formula, thus MP is a function of regional wage and, at the same time, wage rate depends on MP. The double directions of causality implies that MP could not respect the orthogonality condition, requests to obtain unbiased estimation when OLS is performed

To address this issue, the econometric literature suggests to use instrumental variables. The choice of instrumental variables is rather challenging and some difficulties can arise from the availability of the data and the technical characteristics that they must satisfy. An instrumental variable must have the following properties:

- it must be correlated with the instrumented variable;
- it must be not correlated to the disturbance terms.

Empirical NEG literature uses two main instruments, such as the level of education and working experience. The absence of this information at the firm level prevents us to make use of the traditional instruments for the wage equation. Furthermore, the dataset in use, created from balance sheet data, does not include any additional information that can be used to differentiate workers. An alternative approach, which allows to take into account the endogeneity effect without searching for out of sample data is to adopt group the instrumental variables approach [8] (see Fingleton (2006)).

This method involves the creation of a rank variable, where the rank refers to the income groups. In particular, sorting income in increasing order, and deciding the number of groups, each group represents a rank. For example, the highest income will be allocated in the highest rank. In particular, we decide to define 6 groups, so the instrumental variable assumes values from one to six, for the smallest and the highest income level, respectively.

Then, after constructing the instrumental variable, in order to continue using our STATA routine for Non Linear Least squares, we have proceeded to regress provincial incomes on the group instrumental variable and the relative fitted values have been introduced in place of provincial income in the wage model.

4.5.2 Data Description.

The source of the data employed is the Italian section of Bureau van Dijk dataset. This dataset includes balance sheet information for about 90% of Italian companies. From this dataset it has been drawn a sample of small and medium firms, selected through the definition of this type of firm[9], based on:

- total assets, which has to be less or equal to 43 Million;
- total revenues, which have to be less or equal to50Million;
- number of workers, which must not exceed 250 workers.

Other characteristics that we have considered are related to the economic sector, in particular, following Krugman model, we analyse manufacturing and agriculture firms, that respectively exhibit increasing and constant return to scale.

[8] For greater details, see Peter Kennedy, Introduction to Econometrics

[9] Definition issued by European Commision with the official recommendation 2003/361/CE

We focus our analysis on small and medium firms, considered the most representative of economic activity in Italy.

The dataset includes firms data referred to the period from 2000 to 2005, so that we deal with a panel dataset.

The data used for the analysis concerns:

- "wages and salaries" from the Profit and Loss Statement;
- The total number of employees from the exposure draft.

The entry "wages and salaries" summarizes the cost that firms bear for the labour input. The average wage of the firm has been obtained dividing the total costs of workers by their number.

The other information drawn from AIDA is related to the economic sector of firms operation, as defined by the Italian Statistical Institute (ISTAT) in 2002, and their localization, i.e. the region, province. The province is the geographical unit under study here.

In reference to the distance matrix, geographical coordinates (latitude and longitude) , as well as province surface, are freely available. Distance was defined by using the Great Circle Distance Formula that allows computing the distance between two points characterized by different latitude and longitude.

The distance of a point with itself is clearly zero, but we have dealt with localities, characterized by a surface, so it cannot be assumed that there is a null distance between two points in the same province, but firms and/or consumers face transportation costs even within a province,

In the literature it is usually applied the Internal Distance formula in order to determine the main diagonal of the distance matrix. The formula is the following:

$$D_{ii} = \frac{2}{3}\sqrt{\frac{area_i}{\pi}}$$

As mentioned before, the empirical model is based on two main variables, nominal wage and market potential (MP). Whilst the former is obtained from the balance sheet, the latter is a composite variable and it is constructed by using information about distances and province income.

Distance matrix will assume the dimension of the number of observations, we analyse $N = 5241$ firms over $T = 6$ years, so matrix assumes following dimensions: D_N when firms are analysed cross sectionally, and $D_N \otimes I_T$.

Market potential is defined for each province and firms localized in the same province take the same value of market potential.

Income has been defined for each province by equation (4.9). In order to compute the share of workers in each sector: and, population of each province as a proxy of the total number of employed for each geographical unit. It has been further determined the global share of worker in each sector, and calculated by normalizing to 1 the total number of workers, and dividing the number of workers in each sector by the total number of workers in province. Price index has been overlooked in the computation of the variable MP, since we do not have this information at province level. We try to overcome the problem by assuming that price index is equal across provinces.

4.5.3 Results

In this paragraph we report results obtained by performing Non Linear Least Squares Estimator respectively at firm level and at province level.

The model's structure is the same for each estimation, and the spatial aspects are taken into account indirectly by the computation of the market potential determined through income and transport costs, which is computed starting from distance matrix. and permits to consider not only the physical distance but even the economic distance. Two are the parameters obtained from estimation results, the elasticity of substitution, σ that as suggested by theory is greater than 1, and exponent related to the transport costs, that we call *beta* and that is composed as : $\tau \cdot (1 - \sigma)$ so we can indirectly obtain the parameter associated to transport costs.

Parameters found respect sign expected while they diverge in magnitude, we found a rather smaller value of elasticity of substitution, but a value that it seem stable over time and specification, instead we deal with a rather small transport cost coefficient in firm level specification and a rather high value at province level.

Table 4.1 illustrates results obtained by performing respectively on pooled panel, between

Table 4.1: Non Linear Estimation- firm level wage

	w_i	w_i	w_i	w_i	w_i	w_i	w_i	w_i	w_i
σ	1.423***	1.4186***	1.799***	1.431***	1.431***	1.439***	1.429***	1.3933***	1.383***
	(0.0005)	(0.0006)	(0.0033)	(0.0017)	(0.0006)	(0.0008)	(0.00089)	(0.0009)	(0.0013)
$beta$	0.0106***	0.0477***	0.0904***	0.0852***	0.0012	0.011**	0.001	0.0921***	0.1040***
	(0.0032)	(0.0037)	(0.0019)	(0.0115)	(0.0038)	(0.0048)	(0.00567)	(0.00647)	(0.008)
implied τ	0.026	0.114	0.113	0.198	-	.0251	-	0.2341	0.2723
year(s)	Pooled OLS	between estimator	pooled OLS	2000	2001	2002	2003	2004	2005
R-squares	0.99	0.99	0.99	0.99	0.99	0.99	0.99	0.99	0.99
time FE	no	no	yes	no	no	no	no	no	no

where $beta = \tau(1-\sigma)$, implying that $\tau = beta/(1-\sigma)$

Table 4.2: Non Linear Estimation- province level wage

	w_p	w_p	w_p	w_p	w_p	w_p	w_p	w_p	w_p
σ	1.6397***	1.625***	1.6655***	1.423**	1.858***	1.877***	1.602***	1.614***	1.618***
	(0.2301)	(0.5774)	(0.2332)	(0.577)	(0.5614)	(0.5417)	(0.713)	(0.544)	(0.5379)
$beta$	1.033***	1.0236***	1.046***	0.984***	1.124***	1.1438***	0.947***	1.031***	1.027***
	(0.1227)	(0.3076)	(0.1237)	(0.337)	(0.285)	(0.277)	(0.357)	(0.294)	(0.2894)
implied τ	1.615	1.63	1.57	2.32	1.31	1.30	1.58	1.68	1.6
year(s)	Pooled OLS	between estimator	pooled OLS	2000	2001	2002	2003	2004	2005
time FE	no	no	yes	no	no	no	no	no	no
R-squares	0.43	0.4299	0.43	0.4246	0.4203	0.4258	0.3910	0.4330	0.4407

Table 4.3: Non Linear Estimation- province level wage - IV

	w_p	w_p	w_p	w_p	w_p	w_p	w_p	w_p	w_p
σ	1.0486***	1.0127***	1.0355***	1.035***	1.057***	1.0564***	1.051***	1.0424***	1.0405***
	(0.0043)	(0.0097)	(0.0093)	(0.01102)	(0.0106)	(0.011)	(0.0106)	(0.011)	(0.011)
$beta$	0.0316***	0.5577***	0.0315***	0.0323***	0.0294***	0.0319***	0.0325***	0.0351***	0.0356***
	(0.0011)	(0.3525)	(0.00113))	(0.00241)	(0.0026)	(0.00301)	(0.0034)	(0.0037)	(0.00385)
implied τ	0.6502	43.91	0.8873	0.9116	0.5122	0.565	0.6371	0.8254	0.8728
year(s)	Pooled OLS	between estimator	pooled OLS	2000	2001	2002	2003	2004	2005
time FE	no	no	yes	no	no	no	no	no	no
R-squares	0.99	0.99	0.99	0.99	0.99	0.99	0.99	0.99	0.99

estimator, pooled panel with time fixed effect and finally results referred to each year in the dataset. Results suggest that there exists a positive and statistically significant relationship between firms' wage and provincial Market Potential.

Our results presents some differences, specially in the magnitude of coefficients, with respect results we are used to see in similar works. Indeed, we can observe that our results present a smaller values of elasticity of substitution, (generally σ assumes values close to 5 , even greater). The greater is the values of elasticity of substitution the lower is the monopolistic power of firms. We observe the opposite, the elasticity is rather low indicating high monopolistic power. The positive relation between wage rate and MP gives empirical evidence of the presence of a geographical pattern of wages. In particular we find that the higher is the market access of the province where firm is localized, the higher are firms wages.

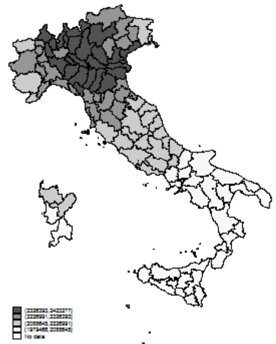
Comparing results obtained adopted our approach with results obtained at provincial level we can observe that without take into account endogeneity, running NLS returns higher values of both parameters, a higher elasticity and transport cost, instead introducing instrumental variables, elasticity of substitution becomes smaller than the firm level results, while transport cost parameter assume greater with respect firm level results and smaller values is compared with the provincial level results without instruments.

The statistical significance and the positive sign of MP suggest that it affects positively the wage level. This variable gives us information about the presence of a spatial structure of the wage, confirming the NEG theory.

Summarising, we have found evidence that wages are highly affected by MP. This result supports the NEG theory, and the idea that wages are affected by the potential demand of the local area where firms are localized, furthermore these results suggest that market intensity affects price of input factor. The magnitude of elasticity of substitution found does not appear consistent with previous empirical literature. The relatively low value of the elasticity indicates that consumers attach much value to goods' variety and that in equilibrium the number of large firms tends to be high.

The rather small magnitude of substitution elasticity implies a relevant relation between MP and wages, since this relation is quantified by the inverse of σ parameters, reinforcing the evidence of wage spatial pattern, that can be inferred graphically observing that northern provinces are characterized by a more intensive MP.

Figure 4.3: Italian Market Potential



4.6 Conclusion

Starting from the contribution of Fujita, Krugman and Venables (1999), more and more researchers have addressed their interest in economic geography, motivated by the introduction of general equilibrium model for explaining economic concentration/localisation and by the assumptions of increasing returns to scale and more realistic market structures, such as monopoly and oligopoly. This branch of economics is called New Economic Geography.

NEG theory employs the variable market potential to measure the effect of firms concentration; this variable was firstly defined by Harries (1954) as the demand for goods produced in a location, determined as the sum of purchasing power in other locations, weighted by transport costs. Most of the empirical analysis has focused on the impact that the variable MP has on the inputs' price, in particular on the cost of the input labour, verifying the relationship between wage and MP.

In line with the empirical literature, in this thesis, we have examined the presence of a spatial structure of wages in Italy and whether market potential is able to explain the wages disparities in Italy. The investigation has involved data related to 5241 Italian firms, belonging to two sectors: manufacturing and agriculture. Firms are localized in 99 provinces.

The study engages firm level data for the dependent variable, the wage, and province level data for independent variable, the market potential. The methodology adopted aims to capture the spatial effects, through transport costs, included in the variable MP. Results obtained suggest that there exists a positive relation between wages and market potential and also show that the geographical localization matter, and that there exists a wages disparity in Italy.

The larger is Market potential, the larger the demand for production factors determining a raise in their compensations rate; likely, this force will work together with an increase in the factors inflows, given that input factors are attracted by the market where rewards are relatively higher. Applying this concept to the labour factor, workers move towards regions where the wages are higher and where there are more varieties of goods, determining an increase in price, due to the increased demand, that lead to equalize the real wage.

Although the Italian labour market is rigid, our results show a strong relationship [10] between wage and Market Potential. This result could be driven by our sample since we are considering only capital companies, that are more prone to flexible markets. However, even if we found this relation, MP across provinces is not the same and the intensity of MP can be better observed in Figure 4.3.

[10] Represented by the inverse of the elasticity of substitution

Conclusions

In recent years, we have seen an increase of interest in the role of geography and space for economic activity, recognising the uneven distribution of people and firms across space. The traditional economic theory has attributed this phenomenon to the natural endowment of locations, i.e. the so called first nature advantages. Despite the wide literature on these topics, for a long time the role of space has been ignored in mainstream economics, mainly because of the inability of the traditional model to deal with space in a coherent general equilibrium framework. Instead, more recently, economic theories, and in particular the New Economic Geography framework, have proposed a general equilibrium model where the inequality of economic activity are ascribed to the so called second nature advantages, which arise through human interactions occurring between agents localised close to each other. The economies arising because of closeness favour the spreading of knowledge, and the transfer of tacit knowledge that requires personal contacts among agents. Indeed, sharing knowledge is not easy and distance represents an important barrier in this direction.

We have, further, observed how the greater availability of data (longitudinal dataset) and more powerful computershave contributed to the evolution of the methodologies adopted to analyse the economic activity distribution, leading to an increase in micro level studies.

These changes, observed mainly in the empirical literature, come from the awareness that individual agents represent the natural level in order to assess the validity of microfounded economic theories.

Along these lines, this thesis aims to investigate the role of space for economic activity, looking at the smallest possible level of analysis. i.e. the firm.

This work follows two main directions. It starts looking at the firms' productivity and the importance that of external environment for their performances, according to the main theories on the role of agglomeration. To the other side, it has been examined at micro level, one of the main relationship arising from equilibrium solution of New Economic Geography. This relationship links firm's wage rate with Market Potential of the geographical area where a firms are located.

With respect to the first part, since estimating firm level productivity represents a challenging topics, we have dedicated the first chapter to this issue. In particular, it has been stressed the importance of potential aggregation bias issues, especially when macro-aggregate data are used.

Therefore, we have focused on the micro level production function in order to obtain a consistent measure of firms' productivity for a large sample of Italian firms. Moreover, in the first chapter, we have proceeded to estimate aggregate productivity, in order to compare disaggregate and aggregate measures of productivity, obtained through a variety of estimators.

This analysis is carried out using alternative specifications and alternative estimators (OLS, FE, Difference GMM, System GMM, LP). The results seem to show that firms in our sample do not exhibit internal economies of scale and, as expected, firms are labour intensive. . Finally, we have observed how traditional approaches, which do not account for simultaneity and endogeneity issues, underestimate capital. These results corroborate our desire to investigate the effects of external economies.

Therefore, in the second chapter, we have focused on how productivity is affected by spatial externalities. We aim to verify the presence of second nature advantages for firms localised in the same geographical area, i.e. the province. To this goal we have built a battery of indices capturing externalities as defined by Marshall, Jacobs and Porter (specialisation, diversity, and competition indices) and we have analysed the effect of province/industry externalities on the firm-level productivity estimated in the first chapter. This approach allows to overcome the potential causality issue in these types of investigation without incurring in other potential problems embedded in the Instrumental Variables estimator. Further, following the same approach adopted in chapter one, we have also performed our analysis at the province level and compared results. Looking at the relationship between TFP and spatial externalities we find mixed results, highly affected by the specification adopted (first nature advantages and/or industry control variable). These results provide rather weak empirical evidence in favour of Marshallian, Jacobian and Porterian externalities. This result is not due to the level of aggregation but to the indices introduced in the specifications, since performing our analysis at province level, also leads to mixed evidence.

In the third chapter, we have been focused on the recent developments in spatial analysis in order to overcome some of the drawbacks affecting traditional approaches. In particular, we have analysed the spatial interactions among agents introducing explicitly geography in the model to be estimated and relaxing the assumption that locations are independent.

Implementing this methodology is possible thanks to the recent tools provided by spatial econometrics, which exploit the advantages arising from combining panel data and spatial analysis. Further, in order to apply this approach, we have relaxed the assumption of homogeneous space and taken into account the role of neighbouring economies.

We have implemented spatial econometric models, such as SAR and SEM models, both in their plain form and adding fixed effects. Results obtained suggest that spatial interactions are rather week, as they have a small magnitude, but the relationship between TFP at one location and the TFP of neighbouring locations is positive and statistically significant.

In the fourth chapter, we have empirically tested a relation coming from a General Equilibrium Model, in particular as proposed by NEG models. In chapters two and three, we have followed two different methodologies in order to account for space, but in both cases, we have looked to the firms point of view, analysing the supply side of the economy. In the last chapter we have taken into account a more complete framework, that cannot be solved analytically, so we have chosen a single equation of the model represented by the wage equation. Our model is non linear and to this reason we have applied non linear estimator (such as Non Linear Least Squares). The relationship between wage rate and Market Potential underlines that there exists a spatial pattern of Italian wage rates in manufacturing sector. Results suggest that there exists a spatial structure and uneven distribution of wage rates at the firm as well as the province level.

Our analysis has not shown unanimous results with respect to the role of spatial externalities. We can state that the overall these effects seems to be positive, but weak; this conclusion is reached from the balance of how single effects work in different directions.

This difference can be observed in the second chapter, where we have distinguished among three main kinds of externalities and we have obtained different sign for the same indices in different specifications.

In agreement with the NEG theory, the distribution of economic activities over space is driven by the interplay of two forces, centrifugal and centripetal, the former promotes geographical dispersion, the latter geographical concentration. In our analysis centripetal forces seem to prevail, and agglomeration generates some benefits for firms, even if these are not so strong.

This result highlights the difficulty to implement policies in order to encourage agglomeration, as the border between the advantages and disadvantages arising from agglomeration is very thin, and excessive clustering can cause centrifugal forces to dominate centripetal one.

On the other side, promoting agglomeration could lead to the following scenario: workers migrate toward the greater centres, characterised by higher market access and higher wages, population, workers and economic activities are then localised in few places, which, according to the

NEG theory, are characterised by a positive growth rate due to the cumulative causation. On the contrary, peripherical areas will be progressively abandoned.

This phenomenon has been observed in US, for example Glaeser (2007) has analysed the case of Buffalo, historically characterised by a high concentration of firms because of the richness of resources, while recently, it continues to lose people. Glaeser stresses how the policies adopted by the federal government to revitalise Buffalo have not obtained the desired results and the city continues to decay. Glaeser suggest to pursue policies that help Buffalo's citizens, not the city as a geographical place.

Policies that favour agglomeration, according to NEG theories lead the wage rate to increase. Greater wages, imply greater regional income and so higher demand. Of course, this scenario does not take into account the proportion of the population not represented by workers, such as the elderly and children.

If the main goal of the policymaker is the national growth, its policy will be oriented toward the promotion of agglomeration, regardless of the uneven distribution of wealth. Instead, if its main aim is to ensure a fair distribution of wealth, dispersion policies are more suitable. Policies which favour agglomeration could lead to face problems similar to Buffalo. A high national growth rate could be achieved at the expense of desertification or impoverishment in some areas.

Two the main kinds of policies may be adopted: incentives to the firm or to the territory. The former could not solve structural issues faced by firms implying low productivity, while the latter, improving the environment where firms operate could improve their long run performance. Probably, it is also important, to combine this kind of policy with a certain degree of geographical aggregation. Indeed, different policies could be implemented at different levels, although they need to be coordinated.

Clearly, the issues discussed in this thesis present a degree of complication that is extremely difficult to summarise in empirical work. We have done our best to do so. In particular, we have been able to limit the issue of aggregation bias and endogeneity implicit in these kind of investigations by the use of micro-level data. Also, using a large panel dataset, we have been able to exploit the greater information they contain and to account for firm-level and geographical heterogeneity in suitable econometric models. Future work could address some further issues, also important in order to conceive the role of space for economic activity. In particular, we can mention the so called MAUP in the analysis of agglomeration economies. Moreover, the spatial econometric analysis has also considered interactions at the province level. Although this could be computationally intensive, future work could extend the analysis to consider firm-level contiguity and distance.

Finally, with respect to the empirical validation of NEG model, the main literature on this topic, we have applied the Non Linear Least Squares estimator. Alternative estimators could be use in order to improve on numerical accuracy.

Overall, this analysis has confirmed space as an important determinant of economic activity and the results obtain encourage further investigation in the empirical economics of geographical economics at the firm level.

Bibliography

Abreu, M., Groot, H. L. de and Florax, R. J. (2004). Spatial patterns of technology diffusion: An empirical analysis using tfp. *Tinbergen Institute Discussion Paper* .

Ackerberg, D. A., Caves, K. and Frazer, G. (2006). Structural identification of production functions. *mimeo, UCLA Department of Economics* .

Alonso, W. (1964). *Location and Land Use*. Harvard University Press.

Amiti, M. and Pissarides, C. (2005). Trade and industrial location with heterogeneous labor. *Journal of International Economics* 67: 392–412.

Anselin, L. (1988). *Spatial econometrics: methods and models*. Dordrecht: Kluwer.

Anselin, L. (2002). Under the hood. issue in the specification and interpretation of spatial regression models. *Agriculture Economics* 27: 247–267.

Anselin, L. and Bera, A. (1998). *Handbook of Applied Economic Statistics*. Marcel Dekker, chap. Spatial dependence in linear regression models with an introduction to spatial econometrics. Ullah, a. and giles, d.e. ed., 237–289.

Anselin, L. and Le Gallo, J. (2008). *Spatial Panel Econometrics In The Econometrics of Panel Data, Fundamentals and Recent Developments in Theory and Practice*. Springer-Verlag.

Arbia, G. (2001). Modellling the geography of economic activities on a contiuous space. *Papers in Regional Science* 80: 411–424.

Arrow, K. (1962). The economic implications of learning by doing. *Review of Economic Studies* 29: 155–173.

Aw, B., Chen, X. and Roberts, M. (1997). Firm-level evidence on productivity differentials, turnover, and exports in taiwanese manufacturing, working Papers 6235, NBER.

Baily, N., Hulten, C. and Campbell, D. (1992). *Brookings Papers on Economic Activity: Microeconomics*. Washington, DC: Brookings Institution, chap. Productivity dynamics in manufacturing establishments.

Balatgi, B. and Liu, L. (2008). Testing for random effects and spatial lag dependence in panel data models. *Statistics and Probability Letters* 78: 3304–3306.

Balatgi, B., Song, S. H. and Koh, W. (2003). Test panel data regression models with spatial error correlation. *Journal of Econometrics* 117: 123–150.

Balatgi, B., Song, S. H. and Koh, W. (2007). Test for serial correlation, spatial autocorrelation and random effects using panel data. *Journal of Econometrics* 140: 5–51.

Baldwin, R. (2001). Core-periphery model with forward-looking expectations. *Regional Science and Urban Economics* 31: 21–49.

Baldwin, R. and Forslid, R. (2000). Core-periphery model and endogeneous growth:stabilizing and destabilizing integration. *Economica* 67: 307–324.

Bartelsman, E. and Dhrymes, P. (1998). Productivity dynamics: U.s. manufacturing plants. *Journal of Productivity Analysis* 9: 5–34.

Beaudry, C. and Schiffauerova, A. (2009). Who's right marshall or jacobs? the localization versus urbanization debate. *Research Policy* 38: 318–337.

Bond, S. (2002). Dynamic panel data models: a guide to micro data methods and practice. *Working paper CWP09/02, CEMMAP* .

Brakman, S., Garretsen, H. and Marrewijk, C. van (2009). *The New Introduction to Geographical Economics*. Cambridge University Press.

Brakman, S., Garretsen, H. and Schramm, M. (2002). New economic geography in germany: Testing the helpman-hanson model, working Paper, Hamburg Institute of International Economics.

Brulharta, M. and Sbergami, F. (2009). Agglomeration abd growth: Cross-country evidence. *Journal of Urban Economics* 65: 48– 63.

Burger, M., Oort, F. van and Knaap, B. van der (2007). A treatise on the scale-dependency of agglomeration externalities and the modifiable areal unit problem, paper presented at the Workshop on Agglomeration and Growth in Knowledge-Based Societies, Kiel, Germany.

Cainelli, G. and Leoncini, R. (1999). Externalities and long term local industrial development. some empirical evidence from italy. *Revue d'Economie Industrielle* 4: 25–39.

Cainelli, G. and Lupi, C. (2009). *Does Spatial Proximity Matter? Micro-evidence from Italy*. Routledge, chap. Internalization, Technological Change and the Theory of the Firm. De liso, n. and leoncini,r. ed.

Casetti, E. (1972). Generating models by the expansion method: Applications to geographical research. *Geographical Analysis* 4: 81–91.

Chamberlain, G. (1982). Multivariate regression models for panel data. *Journal of Econometrics* 18: 5–46.

Chamberlain, G. (1984). *Handbook of Econometrics*. Elsevier Science Publishers, *4*, chap. Panel Data. Z. griliches and m. d. intriligator ed.

Cingano, F. and Schivardi, F. (2004). Identifying the sources of local productivity growth. *Journal of the European Economic Association* 2: 720–742.

Combes, P.-P., Duranton, G. and Laurent, G. (2008). Spatial wage disparities: Sorting matters! *Journal of Urban Economics* 63: 723–742.

Davis, D. R. and Weinstein, D. E. (1996). Does economic geography matter for international specialization, working paper No. 5706,NBER.

Davis, D. R. and Weinstein, D. E. (1999). Increasing returns and international trade: An empirical confirmation. *European Economic Review* 43: 379–407.

Davis, D. R. and Weinstein, D. E. (2003). Market access, economic geography and comparative advantage: an empirical test. *Journal of International Economics* 59: 1–23.

De Bruyne, K. (2002). The location of economic activity. is there a spatial employment structure in belgium?, katholiek University of Leuven.

De Lucio, J. J., Herce, J. A. and Goicolea, A. (2002). The effects of externalities on productivity growth in spanish industry. *Regional Science and Urban Economics* 32: 241–258.

De Vor, F. and De Groot, H. (2010). Agglomeration externalities and localized employment growth: the performance of industrial sites in amsterdam. *The Annals of Regional Science* 44: 409–431.

Debreu, G. (1959). *Theory of value*. Yale University Press.

Dekle, R. (2002). Industrial concentration and regional growth: Evidence from the prefectures. *The Review of Economics and Statistics* 84: 310–315.

Dettori, B., Marrocu, M. and Paci, R. (2011). Total factor productivity, intangible assets and spatial dependence in the european regions. *Regional Studies* : Forthcoming.

Di Liberto, A. and Usai, S. (2010). Tfp convergence across european regions: a comparative spatial dynamics analysis, working Paper CRENoS 201030, Centre for North South Economic Research.

Dixit, A. K. and Stiglitz, J. E. (1977). Monopolistic competition and optimum product diversity. *American Economic Review* 67: 297–308.

Duranton, G. and Puga, D. (2004). Micro-foundations of urban agglomeration economies. *Handbook of Regional and Urban Economics* 4: 2063–2117.

Elhorst, P. (2009). *Handbook of Applied Spatial Analysis*. Springer: Berlin Heidelberg New York, chap. Spatial Panel Data Models. M.m. fischer and a. getis ed.

Ellison, G. and Glaeser, E. L. (1997). Geographic concentration in u.s. manufacturing industries: A dartboard approach. *Journal of Political Economy* 105: 889–927.

Fallah, B., Partridge, M. and Olfert, R. (2010). Urban sprawl and productivity: Evidence from us metropolitan areas. *Papers in Regional Science* .

Felipe, J. and Fisher, F. (2003). Aggregation in production functions: what applied economists should know. *Metroeconomica* 54: 208–262.

Fingleton, B. (2005). Beyond neoclassical orthodoxy: A view based on the new economic geography and uk regional wage data. *Regional Science* 84: 351–375.

Fingleton, B. (2006). The new economic geography versus urban economics: an evaluation using local wage rates in great britain. *Oxford Economic Papers* 58: 501–530.

Fingleton, B. (2007). *New Directions in Economic Geography*, chap. Testing the new economic geography: a comparative analysis based on EU regional data. Bernard fingleton ed.

Forslid, R. (1999). Agglomeration with human and physical capital: an analytically solvable case, cEPR Discussion Papers 2102.

Foster, L., Haltiwanger, J. and Krizan, C. (1998). Aggregate productivity growth: Lessons from microeconomic evidence, working papers 6803 ,NBER.

Fu, S. and Junjie, H. (2010). Testing urbanization economies in manufacturing industries: Urban diversity or urban size? *Economics and Econometrics Research Institute Working Paper* .

Fujita, M. (1988). A monopolistic competition model of spatial agglomeration : Differentiated product approach. *Regional Science and Urban Economics* 18: 87–124.

Fujita, M. and Krugman, P. (2004). The new economic geography: Past, present and the future. *Reginal Studies* 5: 139–164.

Fujita, M., Krugman, P. R. and Venables, A. J. (1999). *The Spatial Economy*. MIT Press.

Glaeser, E. L. (2007). Can buffalo ever come back?, city Journal Autumn 2007.

Glaeser, E. L., Kallal, H. D., Scheinkman, J. A. and Shleifer, A. (1992). Growth in cities. *Journal of Political Economy* 100: 1126–52.

Gorman, W. (1959). Separable ulitility and aggregation. *Econometrica* 27: 469–481.

Griliches, Z. (1979). Issues in assessing the contribution of research and development to productivity growth. *The Bell Journal of Economics* 10: 92–116.

Griliches, Z. and Regev, H. (1995). Firm productivity in israeli industry: 1979-1988. *Journal of Econometrics* 65: 175–203.

Grunfeld, Y. and Griliches, Z. (1960). Is aggregation necessarily bad? *The Review of Economic and Statistics* 17.

Guillain, R. and Le Gallo, J. (2006). Measuring agglomeration: an exploratory spatial analysis approach applied to the case of paris and its surroundings, rEAL Working Paper 06-T-10.

Hall, R. and Jones, C. (1996). The productivity of nations, working Paper No. 5812 ,NBER.

Haltiwanger, J. (1997). Measuring and analyzing aggregate fluctuations: The importance of building from microeconomic evidence. *Federal Reserve Bank St. Louis Review* 79: 55–77.

Hanson, G. (2005). Market potential, increasing returns, and geographic concentration. *Journal of International Economics* 67: 1–24.

Harris, C. (1954). The market as a factor in the localization of industry in the united states. *Annals of the Association of American Geographers* 64: 315 – 348.

Head, K. and Mayer, T. (2005). *Handbook of Regional and Urban Economics*, chap. The Empirics of Agglomeration and Trade. Henderson v. and thisse j.f. ed., 2609–2669.

Head, K. and Ries, J. (2001). Increasing returns versus national product differentiation as an explanation for the pattern of us-canada trade. *American Economic Review* 91: 858– 876.

Heckscher, E. (1919). The effects of foreign trade on the distribution of income. *Ekonomisk Tidskrift* 21: 497–512.

Helpman, E. (1998). *Topics in Public Economics*. Cambridge University Press, chap. The size of regions. Pines, d., sadka, e., zilcha, i ed., 33–54.

Henderson, V. (2003). Marshall's scale economies. *Journal of Urban Economics* 53: 1– 28.

Hering, L. and Poncet, S. (2006). Market access impact on individual wage: Evidence from china, working paper 2006-23 CEPII.

Hotelling, H. (1929). Stability in competition. *The Economic Journal* 39: 41–57.

Hox, J. (2002). *Multilevel Analysis: Techniques and Applications*. London: Lawrence Erlbaum.

Hsiao, C. (2003). *Analysis of Panel Data*. Cambridge University.

Islam, N. (1995). Growths empirics: a panel data approach. *Quarterly Journal of Economics* 110: 1127–1170.

Islam, N. (1999). Intenational comparison of total factor productivity: a review. *Review of Income and Wealth* 45: 493–518.

Jacobs, J. (1969). *The Economy of Cities*. Jonathan Cape.

Jorgenson, D. and Griliches, Z. (1967). The explanation of productivity change. *Review of Economic Studies* 34: 349–383.

Kelejian, H. and Prucha, I. (1998). A generalized spatial two-stage least squares procedure for estimating a spatial autoregressive disturbances. *Journal of Real Estate Finance and Economics* 17: 99–121.

Kiso, T. (2005). Does new economic geography explain the spatial distribution of wages in japan?, mimeo University of Tokyo.

Klein, L. (1964). Remarks on the theory of aggregation. *Econometrica* 14: 303–312.

Krugman, P. (1979). Increasing returns, monopolistic competition and international trade. *Journal of International Economics* 9: 469–479.

Krugman, P. (1980). Scale economies product differentiation and the pattern of trade. *American Economic Review* 70: 950–959.

Krugman, P. (1991). Increasing returns and economic geography. *Journal of Political Economy* 99: 483–499.

Krugman, P. (1993). First nature, second nature, and metropolitan location. *Journal of Regional Science* 33: 129–144.

Leontief, W. (1947). Introduction to a theory of internal structure of functional relationships. *Econometrica* 15: 361–373.

Levinsohn, J. and Petrin, A. (2003). Estimating production functions using inputs to control for unobservables. *Review of Economic Studies* : 317–342.

Lintunen, J., Olli, R. and Yrjo, V. (2009). Micro meets macro via aggregation, helsinki Center of Economic Research Discussion Paper.

Losch, A. (1954). *The Economics of Location: A Pioneer Book in the Relations Between Economic Goods and Geography*. Yale University Press.

Lucas, R. E. (1988). On mechanism of economic development. *Journal of Monetary Economics* 22: 3–42.

Lutero, G. (2010). The aggregation problem in its hystorical perspective: a summary overview, working paper ISTAT.

Marshall, A. (1890). *Principles of Economics*. Macmillan.

Martin, P., Mayer, T. and Mayneris, F. (2008). Spatial concentration and firm-level productivity in france, cEPR Discussion Papers 6858.

Martin, R. (1999). The new geographical turn in economics: some critical reflections. *Cambridge Journal of Economics* 23: 65–91.

McCann, P. (2005). Transport costs and new economic geography. *Journal of Economic Geography* 5: 305–318.

McMillen, D. (1996). One hundred fifty years of land values in chicago: A nonparametric approach. *Journal of Urban Economics* 40: 100–124.

Mion, G. (2004). Spatial externalities and empirical analysis: the case of italy. *Journal of Urban Economics* 56: 97–118.

Mishra, T., Jumah, A. and Parhi, M. (2008). Age-structured human capital and spatial total factor productivity dynamics, economic series 226, Institute for Advanced Studies.

Moran, P. A. (1950). Notes on continuous stochastic phenomena. *Biometrika* 37: 17–23.

Mori, T. and Turrini, A. (2005). Skills, agglomeration, and segmentation. *European Economic Review* 49: 201–225.

Murata, Y. (2003). Product diversity, taste heterogeneity, and geographic distribution of economic activities: Market vs. non-market interactions. *Journal of Urban Economics* 53: 126 –144.

Nataf, A. (1976). Sur la possibilite de construction de certain macro-model. *Econometrica* 16: 232–244.

Neary, P. (2001). Of hype and hyperbolas: Introducing the new economic geography. *Journal of Economic Literature* 39: 536 – 561.

Nelson, R. and Phelps, E. (1966). Investment in humans, technological diffusion, and economic growth. *American Economic Review* 56: 65–75.

Ohlin, B. (1933). *Interregional and International Trade*. Harvard University Press, Cambridge.

Olley, G. S. and Pakes, A. (1996). The dynamics of productivity in the telecommunications equipment industry. *Econometrica* 64: 1263–1297.

Ottaviano, G. and Puga, D. (1998). Agglomeration in the global economy: A survey of the new economic geography. *World Economy* 21: 707–731.

Ottaviano, G., Tabuchi, T. and Thiesse, F. (2002). Agglomeration and trade revisited. *International Economic Review* 43: 409–436.

Ottaviano, G. and Thiesse, F. (2005). New economic geography: what about the n. *Environment and Planning A* 37: 1707–1725.

Paci, R. and Usai, S. (2006). Agglomeration economies and growth-the case of italian local labour systems, 1991-2001. *Working Paper CRENoS, Centre for North South Economic Research, University of Cagliari and Sassari* .

Pesaran, H. (2004). General diagnostic tests for cross section dependence in panels, working paper in Economics 435 and CESifo WP 1229.

Piser, A. J. (2006). Estimating krugman s economic geography model for the spanish regions. *Spanish Economic Review* 8: 83–112.

Porter, M. E. (1990). *The CompetitiveA dvantage of Nations*. New York: Free Press.

Redding, S. (2009). The empirics of new economic geography, cEP Discussion Paper No 925.

Redding, S. and Sturm, D. (2008). The costs of remoteness: evidence from german division and reunification. *American Economic Reviews* 98: 1766–1797.

Redding, S. and Venables, A. (2004). *Globalization: Analyzing the Economics*. National Bureau of Economic Research, Inc, chap. Geography and Export Performance: External Market Access and Internal Supply Capacity. De liso and leoncini ed., 95–130.

Ricardo, D. (1821). *Principles of Political Economy and Taxation*. Dover Pubblications.

Romer, P. (1990). Endogenous technological change. *Journal of Political Economy* 98.

Romer, P. M. (1986a). Increasing returns and long run growth. *Journal of Political Economy* 94: 1002–1037.

Romer, P. M. (1986b). Increasing returns and long run growth. *Journal of Political Economy* 94: 1002–1037.

Roos, M. (2001). Wage and market potential in germany. *Review of Regional research* 21: 171–195.

Rosenthal, S. S. and Strange, W. C. (2004). *Handbook of Urban and Regional Economics*. Routledge, *4*, chap. Evidence on the Nature and Sources of Agglomeration Economies. N. de liso and r. leoncini ed.

Scott, A. (1988). *New industrial spaces*, chap. Flexible production organisation and regional development in North America and Western Europe. Pion shah s. and m. ed.

Solow, R. (1994). Perspectives on growth theory. *The Journal of Economic Perspectives* 8: 45–54.

Storper, M. (1997). *The regional world*. New York: The Guidford Press, chap. Territorial development in a global economy.

Tabuchi, T. and Thiesse, F. (2002). Taste heterogeneity, labor mobility and economic geography. *Journal of Development Economics* 69: 155–177.

Theil, H. (1954). *Linear aggregation of economic relations*. Amsterdam, North Holland.

Tybout, J. (1996). *Industrial Evolution in Developing Countries: Micro Patterns of Turnover, Productivity, and Market Structure*. Roberts , M and Tybout, James.

Venables, A. (1996). Equilibrium locations of vertically linked industries. *International Economic Review* 37: 341–359.

Von Thünen, J. H. (1826). *The Isolated State*. Hamburg: Perthes, 1826. English Translation. Oxford: Pergamon, 1966.

Weber, A. (1909). *Theory of the Location of Industries*. The University of Chicago Press.

Wheeler, C. H. (2003). Evidence on agglomeration economies, diseconomies and growth. *Journal of applied econometric* 18: 79 –104.

Wooldridge, J. (2009). On estimating firm-level production functions using proxy variable to control for unobservables. *Economics Letters* 104: 112–114.

Zellner, A. (1962). An efficient method of estimating seemingly unrelated regressions and tests for aggregation bias. *Journal of the American Statistical Association* 57.

Appendix A

Statistical Appendix

Data are collected from the Italian section (AIDA) of the Bureau van Dijk database, limiting our analysis to Small and Medium Enterprises. Before to start our empirical analysis we have proceeded to analyse our dataset. Through descriptive statistics, we have noted that some outliers are present. Since we cannot verify balance sheet of firms directly, we cannot distinguish between an odd but true value from a type error. In order to avoid biases in our analysis, we aim to eliminate these outliers without lose information.

Focusing on data related to capital, value added, number of workers, intermediate inputs cost, we have proceeded testing our data to find them, then we have interpolated missing data generated because outliers. We have trying to keep the series as much as possible close to the original ones, intervening as little as possible.

Original dataset did not have missing values except for the number of workers, but it has outliers and unfeasible values,such as negative values for capital, labour cost and intermediate inputs cost.

Our attention has been turned to the number of workers, in order to fill missing data and check for down toward and up toward peaks. We have proceeded to divide our sample using three range of values that L can assume ($5 < L < 10$; $10 < L < 50$ and $L > 50$) considered as peaks , respectively:

- When L at time t is fivefold with respect to L at time $t-1$ and at $t+1$ halves with respect to the value at time t

- When L at time t triples with respect to L at time $t-1$ and at $t+1$ halves with respect to the value at time t

- When L at time t dual with respect to L at time $t-1$ and at $t+1$ halves with respect to the value at time t.

Some procedure has been implemented in order to find down toward outliers. Outliers have been substituted by linear interpolation. This procedure have led to substitute 389 observation and fill 1175 missing values. Referring to the other variables, we have interpolate and substitute 5 observation for capital, 5 observations for intermediate inputs cost , while 25 missing values have been kept for variable labour cost.

Appendix B

Mathematical steps of NEG framework

B.1 Dixit- Stiglitz Model

Cobb-Douglas Utility function:

$$U = M^{\mu} A^{1-\mu}$$

where M is an composite consumption index of the manufacturing goods, defined through an utility function on the manufacturing variety of goods. M, $m(i)$. M is defined by a CES utility function:

$$M = \left[\int_0^n m(i)^{\rho} di \right]^{1/\rho}$$

Representative consumer maximize the utility function subject to the following balance constrain:

$$p^A A + \int_0^n m(i) p(i) di = Y$$

The first step involves the minimization of cost function of manufacturing goods M

$$min \int_0^n m(i)p(i)di$$

$$s.to\ M = \left[\int_0^n m(i)^\rho di\right]^{1/\rho}$$

The minimization problem can be expressed through a Lagrangian:

$$min\mathcal{L} = \int_0^n m(i)p(i)di - \lambda\left[M - \left[\int_0^n m(i)^\rho di\right]^{1/\rho}\right]$$

$$\frac{\partial \mathcal{L}}{\partial m(i)} = p(i) - \lambda\frac{1}{\rho}\rho m(i)^{\rho-1} = 0$$

this represents a generic FOC, that can be rewritten identical for all the manufacturing good $i = 0...j....n$, considering the $j - th$ good:

$$\frac{\partial \mathcal{L}}{\partial m(j)} = p(j) - \lambda\frac{1}{\rho}\rho m(j)^{\rho-1} = 0$$

$$p(i) = \lambda\frac{1}{\rho}\rho m(i)^{\rho-1}$$

$$p(j) = \lambda\frac{1}{\rho}\rho m(j)^{\rho-1}$$

in order to solve the system of FOCs, we proceed to relate $p(i)/p(j)$

$$\frac{m(i)^{\rho-1}}{m(j)^{\rho-1}} = \frac{p(i)}{p(j)}$$

$$m(i) = \left[\frac{p(i)}{p(j)}\right]^{\frac{1}{\rho-1}} m(j)$$

$$\frac{\partial \mathcal{L}}{\partial \lambda} = M - [m(i)^{\rho}]^{1/\rho} = 0$$

$$M^{\rho} = m(i)^{\rho}$$

substituting $m(i)$ in the constraint:

$$\left[\int_0^n m(j)\left[\frac{p(i)}{p(j)}\right]^{\frac{\rho}{\rho-1}} di\right]^{1/\rho} = M$$

we can rewrite the previous equation taking into account that $m(j)$ and $p(j)$ are constant with respect to i:

$$m(j)p(j)^{1/(1-\rho)}\left[\int_0^n [p(i)]^{\frac{\rho}{\rho-1}} di\right]^{1/\rho} = M$$

$$m(j)p(j)^{1/(1-\rho)} = \frac{1}{\left[\int_0^n [p(i)]^{\frac{\rho}{\rho-1}} di\right]^{1/\rho}} M$$

$$m(j) = \frac{p(j)}{\left[\int_0^n [p(i)]^{\frac{\rho}{\rho-1}} di\right]^{(\rho-1)/\rho}} M$$

The total expenditure for the variety j, will be:

$$\int_0^n p(j)m(j)dj = \left[\int_0^n [p(i)]^{\frac{\rho}{\rho-1}} di\right]^{(\rho-1)/\rho} M$$

the term $\left[\int_0^n [p(i)]^{\frac{\rho}{\rho-1}} di\right]^{(\rho-1)/\rho} = G$ indicates the price index that we can rewrite substituting the parameter $\rho$with the elasticity of substitution σ

$$G = \left[\int_0^n [p(i)]^{1-\sigma} di\right]^{1/(1-\sigma)}$$

The demand for $m(i)$ can be rewritten in a more concise:

$$m(j) = \left(\frac{p(j)}{G}\right)^{1/(\rho-1)}$$

$$M = \left(\frac{p(j)}{G}\right)^{-\sigma} M.$$

The second step of the optimization problem involves the maximization of the consumers' utility through the optimal share of income between the agriculture and manufacturing goods.

$$max\ U = M^{\mu} A^{1-\mu}$$

$$s.to\ GM + p^A A = Y$$

$$M = \frac{\mu Y}{G} \qquad A = \frac{(1-\mu)Y}{p^A}$$

Combining the previous results:

- demand for A: $A = \frac{(1-\mu)Y}{p^A}$

- demand for M, $m(j) = \mu Y \left(\frac{p(j)^{-\sigma}}{G^{-(\sigma-1)}} \right)$

If G is constant across regions the demand of manufacturing goods M is constant and equal to σ.

Considering now the problem of the producer, that maximise his objective function How to attain this result staring from the profits function:

$$\pi_r = p_r^M q_r^M - w_r^M (F + c^M q^M)$$

where

$$q_r^M = \mu Y_s (p_r^M T_{rs}^M)^{-\sigma} G_s^{(\sigma-1)}$$

$$G_s = \left[\sum_{s=1}^{R} n_r (p_r^M T_{rs}^M)^{1-\sigma} \right]$$

since firms work in monopolistic competition, the control variable is the price, p_r^M. In order to determine the optimal quantity to produce, the profit function has to be maximised with respect the control variable, we proceed to rewrite the profit function highlighting that quantity is a function of price.

$$\pi_r = p_r^M q_r^M (p_r^M) - w_r^M (F + c^M q^M (p_r^M))$$

and differentiate with respect to price

$$\frac{\partial \pi_r}{\partial p_i} = q_r^M + (p_r^M - w_r^M c^M) \frac{\partial q_r^M}{\partial p_r} = 0$$

$$w_r^M c^M = q_r^M \frac{\partial p_r^M}{\partial q_r^M} + p_r^M$$

$$w_r^M c^M = p_r^M \left[1 + \frac{q_r^M}{p_r^M}\frac{\partial p_r^M}{\partial q_r^M}\right] = p_r^M \left[1 - \frac{1}{\sigma}\right]$$

recalling the pricing rule:

$$p_r^M = \frac{c^M w_r^M}{\rho}$$

$$\pi_r = \frac{c^M w_r^M}{\rho} q_r^M(p_r^M) - w_r^M(F + c^M q^M(p_r^M)) =$$

$$= w_r^M \left[\frac{q_r^M c^M}{\rho} - F\right] - c^M q_r^M = w_r^M \left[\frac{(1-\rho) q_r^M c^M}{\rho} - F\right]$$

and the relation between ρ and σ, $\frac{1-\rho}{\rho} = \frac{1}{\sigma - 1}$.

B.2 No Black hole condition

From totally differentiate real wage we obtain:

$$\frac{d\omega}{\omega} = \frac{dw}{w} - \mu \frac{dG}{G},$$

and given that:

$$(1-\sigma)\frac{\partial G}{G} = \frac{L}{\mu} \cdot \left(\frac{G}{w}\right)^{\sigma-1} \left[1 - T^{1-\sigma}\right] \left[\frac{dL}{L} + (1-\sigma)\frac{dw}{w}\right]$$

$$\sigma \frac{dw}{w} = \frac{Y}{w} \left(\frac{G}{w}\right)^{\sigma-1} \left[1 - T^{1-\sigma}\right] \left[\frac{dY}{Y} + (\sigma - 1)\frac{dG}{G}\right]$$

and that $1 + T^{1-\sigma} = \frac{\mu}{L}\left(\frac{G}{w}\right)^{1-\sigma} = \frac{w}{Y}\left(\frac{G}{w}\right)^{1-\sigma}$, substituting in the above equations, we take:

$$(1-\sigma)\frac{\partial G}{G} = \frac{\left[1 - T^{1-\sigma}\right]}{\left[1 + T^{1-\sigma}\right]}\left[\frac{dL}{L} + (1-\sigma)\frac{dw}{w}\right]$$

$$\sigma\frac{dw}{w} = \frac{\left[1 - T^{1-\sigma}\right]}{\left[1 + T^{1-\sigma}\right]}\left[\frac{dY}{Y} + (\sigma - 1)\frac{dG}{G}\right]$$

where $\frac{\left[1-T^{1-\sigma}\right]}{\left[1+T^{1-\sigma}\right]} = Z = 1$

$$(1-\sigma)\frac{\partial G}{G} = \left[\frac{dL}{L} + (1-\sigma)\frac{dw}{w}\right]$$

$$\sigma\frac{dw}{w} = \left[\frac{dY}{Y} + (\sigma - 1)\frac{dG}{G}\right]$$

solving we obtain :

$$(1-\sigma)\frac{\partial G}{G} = \left[\frac{dL}{L} + \frac{(1-\sigma)}{\sigma}\left[\frac{dY}{Y} + (\sigma - 1)\frac{dG}{G}\right]\right]$$

$$(1-\sigma)\frac{\partial G}{G} - \frac{(1-\sigma)}{\sigma}(\sigma - 1)\frac{dG}{G} = \left[\frac{dL}{L} + \frac{(1-\sigma)}{\sigma}\frac{dY}{Y}\right]$$

$$\frac{\partial G}{G}\left[\frac{(1-\sigma)(\sigma - (\sigma - 1)}{\sigma}\right] = \left[\frac{dL}{L} + \frac{(1-\sigma)}{\sigma}\frac{dY}{Y}\right]$$

$$\frac{\partial G}{G} = \left[\frac{dL}{L} + \frac{(1-\sigma)}{\sigma}\frac{dY}{Y}\right]\left[\frac{\sigma}{(1-\sigma)}\right]$$

$$\frac{\partial G}{G} = \left[\frac{\sigma}{(1-\sigma)} \frac{dL}{L} + \frac{dY}{Y} \right]$$

Substituting in the wage differential

$$\sigma \frac{dw}{w} = \left[\frac{dY}{Y} + (\sigma - 1) \left[\frac{\sigma}{(1-\sigma)} \frac{dL}{L} + \frac{dY}{Y} \right] \right]$$

$$\sigma \frac{dw}{w} = \frac{dY}{Y} (1 + \sigma - 1) - \sigma \frac{dL}{L}$$

$$\frac{dw}{w} = \frac{dY}{Y} - \frac{dL}{L}$$

and substituting in the real wage:

$$\frac{d\omega}{\omega} = \frac{dY}{Y} - \frac{dL}{L} - \mu \left[\frac{\sigma}{(1-\sigma)} \frac{dL}{L} + \frac{dY}{Y} \right] = (1-\mu) \frac{dY}{Y} + \left[\frac{\sigma \mu}{(\sigma - 1)} - 1 \right] \frac{dL}{L}$$

$$\frac{d\omega}{\omega} = (1-\mu) \frac{dY}{Y} + \left[\frac{\mu - \rho}{\rho} \right] \frac{dL}{L}$$

where the relation between ρ and σ is: $\frac{\sigma-1}{\sigma} = \rho$

B.3 Changes in Manufacturing workers share effects

In order to totally differentiate the real wage, it is important keep in mind that the symmetric equilibrium implies that the change in one region of the endogenous variables are equal to the mutual change in the other region, that is $dY = dY_1 = -dY_2$.

Fujita et al. (1999) start differentiating the income equations,

$Y_1 = \mu\lambda w_1 + (1-\mu)\frac{1}{2}$

$Y_2 = \mu(1-\lambda)w_2 + (1-\mu)\frac{1}{2}$

$$dY_1 = \mu w_1 d\lambda + \mu\lambda dw_1$$

$$dY_2 = -\mu w_2 d\lambda + \mu(1-\lambda)dw_2$$

the symmetric equilibrium implies $w_1 = w_2 = 1$ and $\lambda = \frac{1}{2}$

$$dY_1 = \mu d\lambda + \mu\frac{1}{2}dw_1$$

$$dY_2 = -\mu d\lambda + \mu(1-\frac{1}{2})dw_2$$

$$dY = \mu d\lambda + \frac{\mu}{2}dw_1$$

At the same way has been treated the price index:

$$G^{1-\sigma} = \left[\lambda w_1^{1-\sigma} + (1-\lambda)(w_2T)^{1-\sigma}\right] =$$

$$G^{1-\sigma}(1-\sigma)\frac{dG}{G} = \left[\lambda(1-\sigma)w_1^{1-\sigma}\frac{dw_1}{w_1} + (1-\sigma)(1-\lambda)w_2^{1-\sigma}T^{1-\sigma}\frac{dw_2}{w_2}\right] + w_1^{1-\sigma}d\lambda - w_2^{1-\sigma}T^{1-\sigma}d\lambda$$

$$G^{1-\sigma}(1-\sigma)\frac{dG}{G} = \left[\lambda(1-\sigma)w_1^{1-\sigma}\frac{dw_1}{w_1} - (1-\sigma)(1-\lambda)w_2^{1-\sigma}T^{1-\sigma}\frac{dw_1}{w_2}\right] + w_1^{1-\sigma}d\lambda - w_2^{1-\sigma}T^{1-\sigma}d\lambda$$

substituting $\lambda = \frac{1}{2}$

$$G^{1-\sigma}(1-\sigma)\frac{dG}{G} = \left[\frac{1}{2}(1-\sigma)w^{1-\sigma}\frac{dw}{w} - \frac{1}{2}(1-\sigma)w^{1-\sigma}T^{1-\sigma}\frac{dw}{w}\right] + w^{1-\sigma}d\lambda - w^{1-\sigma}T^{1-\sigma}d\lambda$$

$$(1-\sigma)\frac{dG}{G} = G^{1-\sigma}\left[\frac{1}{2}(1-\sigma)w^{1-\sigma}\frac{dw}{w}(1-T^{1-\sigma})\right] + G^{1-\sigma}w^{1-\sigma}(1-T^{1-\sigma})d\lambda$$

and evaluating the derivative at the point $w = 1$

$$(1-\sigma)\frac{dG}{G} = G^{1-\sigma}\left[\frac{1}{2}(1-\sigma)\frac{dw}{w}(1-T^{1-\sigma})\right] + G^{1-\sigma}(1-T^{1-\sigma})d\lambda$$

$$(1-\sigma)\frac{dG}{G} = G^{1-\sigma}(1-T^{1-\sigma})\left[(1-\sigma)\frac{1}{2}\frac{dw}{w} + d\lambda\right]$$

Recall the variable Z that has been introduced above $Z = \frac{1-T^{1-\sigma}}{1+T^{1-\sigma}} = \frac{1-T^{1-\sigma}}{2G^{1-\sigma}}$, which derives by the relation $G_1^{1-\sigma} = G_2^{1-\sigma} = \left[\frac{1+T^{1-\sigma}}{2}\right]$.

Substituting Z in the derivative of index price, we obtain:

$$(1-\sigma)\frac{dG}{G} = \left(\frac{2}{1+T^{1-\sigma}}\right)(1-T^{1-\sigma})\left[(1-\sigma)\frac{1}{2}\frac{dw}{w} + d\lambda\right] = Z(1-\sigma)\frac{dw}{w} + 2Zd\lambda$$

$$\frac{dG}{G} = Z\frac{dw}{w} + \frac{2Z}{(1-\sigma)}d\lambda$$

The same approach has been used for wage rate total differentiate.

$$w_1 = \left[Y_1 G_1^{\sigma-1} + Y_2 G_2^{\sigma-1} T^{1-\sigma}\right]^{1/\sigma}$$

$$w_1^{\sigma} = \left[Y_1 G_1^{\sigma-1} + Y_2 G_2^{\sigma-1} T^{1-\sigma}\right]$$

$$\sigma w_1^{\sigma-1} dw_1 = (\sigma-1) Y_1 G_1^{\sigma-1} \frac{dG_1}{G_1} + (\sigma-1) Y_2 G_2^{\sigma-1} T^{1-\sigma} \frac{dG_2}{G_2} + G_1^{\sigma-1} dY_1 + G_2^{\sigma-1} T^{1-\sigma} dY_2$$

symmetric equilibrium implies $dY = dY_1 = -dY_2$ and$dG = dG_1 = -dG_2$

$$\sigma w_1^{\sigma-1} dw_1 = (\sigma-1) Y_1 G_1^{\sigma-1} \frac{dG_1}{G_1} - (\sigma-1) Y_2 G_2^{\sigma-1} T^{1-\sigma} \frac{dG_1}{G_2} + G_1^{\sigma-1} dY_1 - G_2^{\sigma-1} T^{1-\sigma} dY_1$$

symmetric equilibrium leads to the following results: $Y_1 = Y_2$ and $G_1 = G_2$

$$\sigma w^{\sigma-1} dw = (\sigma-1) Y G^{\sigma-1} \frac{dG}{G} - (\sigma-1) Y G^{\sigma-1} T^{1-\sigma} \frac{dG}{G} + G^{\sigma-1} dY - G^{\sigma-1} T^{1-\sigma} dY$$

$$\sigma w^{\sigma-1} dw = (\sigma-1) Y G^{\sigma-1} \frac{dG}{G} \left[1 - T^{1-\sigma}\right] + G^{\sigma-1} dY \left[1 - T^{1-\sigma}\right]$$

$$\sigma dw = (\sigma-1) Y \frac{G^{\sigma-1}}{w^{\sigma-1}} \frac{dG}{G} \left[1 - T^{1-\sigma}\right] + \frac{G^{\sigma-1}}{w^{\sigma-1}} dY \left[1 - T^{1-\sigma}\right]$$

$$\sigma dw = Y\left(\frac{G}{w}\right)^{\sigma-1}\left[1-T^{1-\sigma}\right]\left[(\sigma-1)\frac{dG}{G}+\frac{dY}{Y}\right]$$

substituting the variable Z we obtain:

$$\sigma dw = (\sigma-1)ZY\frac{dG}{G}+2ZdY \tag{B.1}$$

$\omega = wG^{-\mu}$

$$d\omega = G^{-\mu}dw - \mu G^{-\mu}\frac{dG}{G}$$

$$G^{\mu}d\omega = dw - \mu\frac{dG}{G}$$

[1]

substituting dY in (B.1)

$$\sigma dw = (\sigma-1)ZY\frac{dG}{G}+2Z\left(\mu d\lambda+\frac{\mu}{2}dw\right)$$

$$\sigma dw - Z\mu dw = (\sigma-1)ZY\frac{dG}{G}+2Z\mu d\lambda$$

[1] Obtained recall the following relations:

$$dY = \mu d\lambda + \frac{\mu}{2}dw$$

$$\frac{dG}{G} = Z\frac{dw}{w}+\frac{2Z}{(1-\sigma)}d\lambda$$

$$\sigma dw = (\sigma-1)ZY\frac{dG}{G}+2ZdY$$

$$G^{\mu}d\omega = dw - \mu\frac{dG}{G}$$

$$\frac{(\sigma - Z\mu)}{(1-\sigma)}dw - Z\frac{dG}{G} = \frac{2Z\mu}{(1-\sigma)}d\lambda$$

that together with

$$\frac{dG}{G} - Z\frac{dw}{w} = \frac{2Z}{(1-\sigma)}d\lambda$$

we can write the problem as a system of equations and solve it by Cramer rule:

$$\begin{bmatrix} 1 & -Z \\ Z & \frac{(\sigma - Z\mu)}{(1-\sigma)} \end{bmatrix} \begin{bmatrix} \frac{dG}{G} \\ dw \end{bmatrix} = \begin{bmatrix} \frac{2Z}{(1-\sigma)}d\lambda \\ \frac{2Z\mu}{(1-\sigma)}d\lambda \end{bmatrix}$$

$$\frac{dG}{G} = \frac{d\lambda}{\Delta}\frac{2\sigma Z}{(1-\sigma)^2}\left[1 - \mu Z\right]$$

$$dw = \frac{d\lambda}{\Delta}\frac{2Z}{(1-\sigma)}\left[\mu - Z\right]$$

$$det\begin{bmatrix} 1 & -Z \\ Z & \frac{(\sigma - Z\mu)}{(1-\sigma)} \end{bmatrix} = \Delta = \frac{(\sigma - Z\mu)}{(1-\sigma)} + Z^2 = \frac{(\sigma - Z\mu) + Z^2(1-\sigma)}{(1-\sigma)}$$

$$det\begin{bmatrix} 1 & \frac{2Z}{(1-\sigma)}d\lambda \\ Z & \frac{2Z\mu}{(1-\sigma)}d\lambda \end{bmatrix} = \frac{2Z\mu}{(1-\sigma)}d\lambda - \frac{2Z^2}{(1-\sigma)}d\lambda = \frac{2Z}{1-\sigma}d\lambda(\mu - Z)$$

$$dw = \frac{\frac{2Z}{1-\sigma}d\lambda(\mu - Z)}{\frac{(\sigma-Z\mu)+Z^2(1-\sigma)}{(1-\sigma)}} = \frac{det\begin{bmatrix} 1 & \frac{2Z}{(1-\sigma)}d\lambda \\ Z & \frac{2Z\mu}{(1-\sigma)}d\lambda \end{bmatrix}}{det\begin{bmatrix} 1 & -Z \\ Z & \frac{(\sigma-Z\mu)}{(1-\sigma)} \end{bmatrix}} = \frac{d\lambda}{\Delta}\frac{2Z}{1-\sigma}(\mu - Z)$$

$$det\begin{bmatrix} \frac{2Z}{(1-\sigma)}d\lambda & -Z \\ \frac{2Z\mu}{(1-\sigma)}d\lambda & \frac{(\sigma-Z\mu)}{(1-\sigma)} \end{bmatrix} = \frac{2Z}{(1-\sigma)}d\lambda \cdot \frac{(\sigma - Z\mu)}{(1-\sigma)} + \frac{2Z^2\mu}{(1-\sigma)}d\lambda = d\lambda\frac{2Z}{(1-\sigma)}\left[\frac{(\sigma - Z\mu)}{(1-\sigma)} + Z\mu\right] =$$

$$= d\lambda\frac{2Z}{(1-\sigma)}\left[\frac{(\sigma - Z\mu) + Z\mu(1-\sigma)}{(1-\sigma)}\right] = d\lambda\frac{2Z}{(1-\sigma)}\left[\frac{\sigma(1 - Z\mu)}{(1-\sigma)}\right]$$

$$\frac{dG}{G} = \frac{det\begin{bmatrix} \frac{2Z}{(1-\sigma)}d\lambda & -Z \\ \frac{2Z\mu}{(1-\sigma)}d\lambda & \frac{(\sigma-Z\mu)}{(1-\sigma)} \end{bmatrix}}{det\begin{bmatrix} 1 & -Z \\ Z & \frac{(\sigma-Z\mu)}{(1-\sigma)} \end{bmatrix}} = \frac{d\lambda}{\Delta}\frac{2\sigma Z}{(1-\sigma)^2}\left[1 - Z\mu\right]$$

substituting the results of this system in (B.1)

$$G^{\mu}d\omega = dw - \mu\frac{dG}{G}$$

$$G^{\mu}d\omega = \frac{d\lambda}{\Delta}\frac{2Z}{1-\sigma}(\mu - Z) - \mu\frac{d\lambda}{\Delta}\frac{2\sigma Z}{(1-\sigma)^2}\left[1 - Z\mu\right]$$

dividing by $d\lambda$ and by G^{μ}

$$\frac{d\omega}{d\lambda} = \frac{G^{-\mu}}{\Delta}\frac{2Z}{1-\sigma}(\mu - Z) - \mu\frac{G^{-\mu}}{\Delta}\frac{2\sigma Z}{(1-\sigma)^2}\left[1 - Z\mu\right]$$

$$\frac{d\omega}{d\lambda} = \frac{2ZG^{-\mu}}{\Delta(1-\sigma)}\left[(\mu - Z) - \mu\frac{\sigma}{(1-\sigma)}\left[1 - Z\mu\right]\right] =$$

$$= \frac{2ZG^{-\mu}}{(\sigma - Z\mu) + Z^2(1-\sigma)}\left[\frac{(1-\sigma)(\mu - Z) - \mu\sigma\left[1 - Z\mu\right]}{(1-\sigma)}\right] =$$

$$\frac{d\omega}{d\lambda} = \frac{2ZG^{-\mu}}{(\sigma-1)}\left[\frac{-(1-\sigma)(\mu - Z) + \mu\sigma\left[1 - Z\mu\right]}{(\sigma - Z\mu) - Z^2(\sigma-1)}\right] = \frac{2ZG^{-\mu}}{(\sigma-1)}\left[\frac{\mu(2\sigma - 1) - Z(\sigma(1+\mu^2) - 1)}{(\sigma - Z\mu) - Z^2(\sigma-1)}\right]$$

remembering that $\sigma = \frac{1}{1-\rho}$

$$\frac{d\omega}{d\lambda} = 2ZG^{-\mu}\left(\frac{1-\rho}{\rho}\right)\left[\frac{\mu(1+\rho) - Z(\mu^2+\rho)}{1 - \mu Z(1-\rho) - \rho Z^2}\right].$$