MERIT-Infonomics Research Memorandum series

How 'Provincial' is your Region? Effects on Labour Productivity in Europe

2002-002

Alfonso Gambardella, Myriam Mariani and Salvatore Torrisi

MERIT – Maastricht Economic Research Institute on Innovation and Technology
PO Box 616
6200 MD Maastricht
The Netherlands
T: +31 43 3883875
F: +31 43 3884905

http://meritbbs.unimaas.nl
e-mail:secr-merit@merit.unimaas.nl

International Institute of Infonomics

PO Box 2606
6401 DC Heerlen
The Netherlands
T: +31 45 5707690
F: +31 45 5706262

http://www.infonomics.nl
e-mail: secr@infonomics.nl

How 'Provincial' is your Region? Effects on Labour Productivity in Europe

Alfonso Gambardella, Myriam Mariani and Salvatore Torrisi

December 2001

Abstract

This paper estimates the determinants of labour productivity in European NUTS regions during 1989-1996. Unlike previous studies, which have focussed either on local technological capabilities or on agglomeration economies, we compare three potential explanations of regional advantages: Technological capabilities (proxied by regional patents), agglomeration economies (employment density), and openness. To study the latter we use a new measure, the number of airplane passengers embarked and disembarked in the region, and found that in spite of some limitations, this is a meaningful index for the openness of the regions and possibly of other locations (e.g. cities). By using instrumental variables, we confirm existing results that patents and employment density affect labour productivity. Our novel finding is that openness affects labour productivity as well.

Acknowledgements

We thank Zoltan Acs, Ashish Arora, Suma Athreye, Stefano Breschi, Simona Iammarino, Hugo Hollanders, Steve Klepper, Franco Malerba, Alessandro Nuvolari, David Soskice, Davide Ticchi, for helpful comments and suggestions. We acknowledge financial support from the Italian Ministry of University and Research (MURST), and the EU-DG XII projects AITEG and GRIT (Contract # HPSE-CT-1999-00043 and SOE2-CT98-3073). Myriam Mariani also acknowledges support by the European Commission TMR "Marie Curie Fellowship" (Grant # HPMF-CT-2000-00694).

Alfonso Gambardella
Sant'Anna School of Advanced Studies, Pisa
gambardella@sssup.it

Myriam Mariani
University of Camerino, Camerino, Italy and
MERIT, University of Maastricht, Maastricht
m.mariani@merit.unimaas.nl

Salvatore Torrisi
University of Camerino, Camerino, Italy
svtorrisi@libero.it

1. INTRODUCTION

The determinants of regional productivity have drawn increasing attention in recent years. While the topic is rooted in the work of Marshall (1920), Perroux (1950), Myrdal (1957), and Hirschman (1958), its growing popularity owes a great deal to the fortunes of some regions of the world. For example, the story of Silicon Valley prompted Saxenian (1994) to dig into the determinants of "regional advantages". At the same time, regional inequalities have raised a good deal of attention, especially in Europe. As noted for instance by Puga (1999), there are larger income disparities across European regions than US States. This calls for a better understanding of these differences.

Agglomeration economies have been a typical explanation of regional advantages. Several authors have emphasised the importance of local infrastructures and the local milieu for innovation and growth (e.g. Porter, 1998; Swann, Prevetzer, and Stout, 1998). Krugman (1991) and Krugman and Venables (1995) have highlighted the importance of increasing returns associated with the formation of a critical mass of economic activities in a given location. (See also Arthur, 1990.) Ciccone and Hall (1996), and Ciccone (2000) estimated the extent of the agglomeration economies. They found that increases in the density of employment both in the US and in Europe have a positive and significant impact on the labour productivity of a given area. Another typical explanation of regional advantages is technology. Audretsch and Feldman (1996) showed that in the US technological activities tend to cluster. Verspagen (1997), Caniels (1999), and Breschi (1999) obtained similar results for Europe. Paci and Usai (2000) found that in Europe regional patents per capita are positively correlated with labour productivity.

A common feature of these studies is that they look for explanations of regional advantages that are internal to the localities – e.g. local infrastructures or institutions; localised spillovers; local networks. (See also Jaffe, Trajtenberg, and Henderson, 1993.) While these are important factors, this paper argues that another relevant explanation is the "openness" of the regions, and in particular their international openness. As a matter of fact, a notable characteristic of many fast-growing regions of the world today is that they exhibit significant international openness – e.g. Taiwan, Singapore, South Korea, Ireland, Israel, or the software industry in Bangalore or other Indian regions. Moreover, their international connections, and in particular their connections to larger markets or economies, have been crucial for their success. (See Bresnahan, Gambardella, and Saxenian, 2002. See also Acs, 2000.)

The goal of this paper is to explore empirically the extent to which apart from technogical capabilities and agglomeration economies, the openness of the regions affects their economic performance. We employ data from the Eurostat data base REGIO to estimate the determinants of labour productivity in European NUTS (Nomenclature des Unités Territoriales Statistiques) regions during 1989-1996. Apart from controls, we regress labour productivity on three variables. First, following Ciccone and Hall (1996), we use the employment density of the regions as a measure of their agglomeration economies. Second, we use the stock of patents applied for by the inventors located in the regions as a proxy for their technological capabilities. Third, we measure regional openness by a sort of "airport capacity" variable given by the number of airplane passengers embarked and disembarked in the region, and we discuss the advantages and the limitations of this measure. Since patents, passengers and employment density are potentially endogenous, we use instrumental variables. We confirm the existing results in the literature that employment density and patents are correlated with higher productivity. Our novel finding is that airplane mobility is also significantly correlated with labour productivity.

Our analysis raises the question about the meaning of openness, its determinants, and the mechanisms by which it affects output per worker. Openness can take many forms. For example, more open regions can take greater advantage of international spillovers (Coe and Helpman, 1995). This may stem from the mobility of their human capital (e.g. international mobility of students; employees of multinational enterprises), or from the fact that they are better informed about new opportunities (technological or else) that take place elsewhere. More open regions are also more likely to host multinational enterprises. To the extent that multinational firms displace less productive investments by local companies, this increases regional labour productivity (e.g. Rodriguez-Claire, 1996). Fluency with an international language, communities of immigrants from or to the region, weather or other conditions that attract tourism, are other factors that may reduce trade costs, enhance the international mobility of human capital, and increase the potential for knowledge spillovers or technology transfer, or the exchange of final goods or inputs. As we shall see in our model section below, our approach is similar to the one suggested by Redding and Venables (2001). They define "market access" to be the set of factors that ease the market reach of the goods produced by a given location, and "supplier access" the factors that reduce the costs of acquiring inputs from other locations. One way to think about this paper is that we are trying to estimate the effects of the factors that raise the market or the supplier access of a region.

Unfortunately, our problem is that it is difficult to single out all the factors that may enhance the openness (whether market or supplier access) of a location. Moreover, in empirical work like ours, even if one could list a set of potential factors, it is hard to obtain data about them at the country level, let alone at the regional level. The positive twist however, is that most of these factors are highly correlated. For instance, the Asian Tigers, Ireland, Israel, or the Indian software industry show high levels of exports; they benefit from international spillovers because of their international linkages (immigrants, the use of the english language, etc.); they are open to multinational enterprises; they are part of an extensive division of labour particularly with the US. (See Saxenian, 2001; Arora *et al.* 2001.) This suggests that we may capture most of the relevant effects of openness by relying on fewer indicators, and this is what we do by using airplane passengers as a synthetic indicator for the extent to which the regions are linked to the world outside them. Our empirical results are encouraging, which suggests that this variable may indeed capture some underlying features of the openness of a location.

The next section discusses in greater detail our concept of openness, along with some relevant examples. Section 3 develops a basic model to derive the labour productivity equation to be estimated. Section 4 presents our data, the econometric specification, and the empirical estimates. It also discusses the use of our measure of openness and the way we address the endogeneity of employment density, patents and passengers. Section 5 concludes. The Appendix lists the NUTS regions which are part of our sample.

2. BENEFITS OF OPENNESS

There is a fairly long literature on openness especially in the context of trade. Frankel and Romer (1999) is one of the most recent and thorough work in this area. They study the effects of trade (exports plus imports over GDP), as a measure of openness, on income per capita for a sample of countries worldwide.[1] Two features of their study are worth mentioning here. First, unlike previous studies, they use instrumental variables to control for the endogeneity of trade. Second, they recognise that their measure of trade openness may be a proxy for the many ways in which interactions between countries affect income – e.g. specialisation, increasing returns associated with larger markets, mobility of people, wider knowledge spillovers.

Indeed, the perception that there are some basic factors that reduce trade costs, raise the openness of a country, and through that affect incomes in various ways, has become diffused. Amongst others, one factor that has drawn attention is knowledge of the English language. A recent issue of the European Edition of *Business Week* (2001) has argued that in European business there is an increasing divide between those who know English and those who do not. The article reports that in Continental Europe only 41% of the people speak English, and only 29% speak it well enough to carry on a conversation. It also provides examples of the fortunes of individuals who have learned English when they were children, and who have taken great advantage of this skill in the labour market today. Moreover, the number of English schools have increased in many European regions. The article also reports that, in several jobs, from factory floors to offices, workers who speak English command 25% to 35% higher salaries. What is the source of this premium? Individuals who know English could help tap larger international markets for the products of the firms, or they could help acquire

[1] They also mention several earlier studies that have used cross-country regressions to estimate the effects of trade on income per capita. These studies typically find a moderate positive relationship.

international inputs (financial, knowledge, or tangible inputs) more efficiently, or both.

A good example of the set of factors that may reduce the cost of international openness, and give rise to corresponding benefits in terms of growth and incomes, is provided by a recent study in which two of us were involved (Arora, Gambardella, and Torrisi, 2001). In the 1990s the international demand for information technology (IT) services has boomed in the developed countries, and particularly in the US. This has produced an IT skill shortage in these countries, to which IT producing and using firms have responded by outsourcing some of their activities to new locations. Several emerging economies were in the position to offer their services, but only a few have been able to catch the new market opportunities, particularly India, Ireland, and Israel. A critical reason was their natural exposure to international linkages in the form of language and connections with skilled emigrants, along with access to the services of foreign institutions such as multinational corporations and venture capital firms.

During the 1990s, many skilled people, particularly from Ireland and India, emigrated to the US and the UK to fill the gap of IT skills in those countries. These expatriates provided valuable links with foreign markets, helped Indian and Irish firms to absorb technical and managerial practices and establish contacts with foreign customers. For instance, some Indians who had emigrated to work in US firms in the 1980s have helped US buyers to find suppliers in India. On some occasions, the initial stimulus for outsourcing to India came from employees of Indian origin. In some cases successful Indian entrepreneurs have also provided money to Indian firms as "business angels", and they have helped establish links with US based venture capital firms. More recently, India is becoming a major location for international call centres and Indian trainers who have lived in the US teach local firms' employees phonetics and the

American jargon. (See also Arora *et al.*, 2001.)

Moreover, in India and in Ireland, returning emigrants have brought with them a background of working experience in advanced technological and business environments and personal linkages with the international business community. There are several examples of people with working experience abroad. Until recently Ireland, along with the UK, was the single EU largest source of emigration to the US, with a large share of emigrants being directed to California. Over time, however, emigration flows have changed. In the late 1990s net immigration to Ireland has increased (from –400 in 1993 to 22.8 thousands in 1998). Over 53% of immigrants to Ireland in this period were Irish returning emigrants, and over 25% of male emigrants with college education have returned to Ireland during the 1990s. Most of these have been in the 25-44 years age group.

India has been a net exporter of human capital for several decades. Although we lack systematic data, it is believed that Indians account for a very large fraction, perhaps over half of the H1-B work permit visas issued by the US. Similarly, recent reports suggest that many students graduating in computer science at the Indian Institute of Technology in Madras, an elite engineering institute in India, move overseas upon graduation. Thus, despite the growth of the software industry, India continues to be a net exporter of human capital. However, some of these exports are now tied to the growth of the Indian software industry, with many software developers and programmers working overseas for Indian software firms.

International linkages have also produced access to capital and spillovers from multinational corporations and venture capitalists. For instance, Sun Microsystem purchased a minority stake in Iona, a leading Irish software firm, while another large

Irish firm, Kindle Banking Systems, has been acquired by Mysis Group of the UK. Another example is Euristix, a medium-sized Irish firm (150 employees) specialised in telecommunication software, which was acquired by Marconi (UK) in 1999. Some Irish software firms have also benefited from spillovers, mostly in the form job mobility. For example, DLG, a small Irish firm specialised in localisation software development and testing (62 employees) has greatly benefited from its managing director's previous working experience with Lotus, which helped consolidate links between the two firms. The manager has helped the DLG's staff to absorb organisational and management best practices from Lotus. These practices include project management (clear tasks definition, use of milestones, rigorous assessment criteria) and relational and marketing capabilities (ability to conduct a business negotiate, sales skills and formal presentation skills). Moreover, many successful Irish software firms have started as programming houses (subcontractors) for the local subsidiaries of multinational corporations and have then exploited the network and reputation of these customers to gain access to foreign markets. In Israel too, expatriates from and to the US and factors similar to those discussed for Ireland and India have played an important role in the growth of its IT and related industries during the 1990s. (See De Fontenay and Carmel, 2001.)

To further appreciate the potential advantages of openness, compare these countries to Russia. The latter also had a large supply of technically skilled people (e.g. mathematicians) who could tap the IT skills that were demanded in the advanced countries. But no software industry has grown in Russia, at least to the extent that we have observed in India, Israel, or Ireland. Russia lacks a similar long-term historical connection with the leading US market. The political differences between the US and the former Soviet Union have been an obvious impediment. The importance of these connections is even more striking if one considers that as noted by De Fontenay and Carmel (2001), the flow of Russian immigrants to Israel during the 1990s has been an important source of skills for tapping the growing international demand of IT products in the latter country. Technically trained Russians could make a difference only when they became part of an environment that was internationally connected.

3. THE BASIC MODEL

To derive the labour productivity equation to be estimated, we employ a standard new trade theory model (see Redding and Venables, 2001; Overman, Redding, and Venables, 2001; Midelfart-Knarvik, Overman, and Venables, 2001), which we extend to take into account agglomeration economies and other factors. We assume that all the firms that operate in a region produce the same homogenous good competitively, and that this good is different from that produced in the other regions.[2] The good produced by each regions is used both in consumption and as an input in production. The production function for the output Q of region i is

$$Q = G(\frac{Q}{A}, M) \cdot X^{\alpha} \cdot L^{1-\alpha} \qquad (1)$$

where L is the quantity of labour employed; $X \equiv \left(\int_R q_{ji}^{\frac{\mu-1}{\mu}} \, dj \right)^{\frac{\mu}{\mu-1}}$ is the composite input made up of the quantities q_{ji} produced by region j and used in region i; $\mu > 1$ is the

[2] The assumption that the firms in each region produce a homogeneous good is not crucial. Standard new trade theory models equally assume that they produce differentiated goods. Also, the constant returns to scale assumption in (1) below implies that we can safely deal with the representative profit-maximisation problem of the region as a whole, rather than the problems of the individual firms.

elasticity of substitution; R is the relevant number of regions; α is the elasticity parameter; $G(\cdot)$ is a function of other factors that affect the productivity of the firms.[3] Apart from a vector of variables M, we follow Ciccone and Hall (1996) and assume that productivity is affected by the density of output in the region. The latter is measured by the ratio between the aggregate output Q and the area A of the region.[4]

In each region the firms maximise profits $p \cdot Q - w \cdot L - \int_R p_{ji} q_{ji} dj$ by taking the price of their good p as given. They also take as given the price of labour w, and the prices of the other inputs p_{ji}. A key feature of our analysis is that the latter are affected by trade costs. That is, while p_j is the price obtained by the suppliers of the *jth* good, the price actually paid by the users of the good in region i, p_{ji}, is affected by characteristics of the producer region j and of the user region i. The first order condition of this problem with respect to the generic input from region k is

$$p_{ki} = \frac{\alpha \cdot p \cdot Q \cdot q_{ki}^{-\frac{1}{\mu}}}{\int_R q_{ji}^{\frac{\mu-1}{\mu}} dj} \qquad (2)$$

which implies that the ratio of the inputs from two generic regions k and l is

$\frac{q_{ki}}{q_{li}} = \left(\frac{p_{ki}}{p_{li}} \right)^{-\mu}$. This expression can be used to replace both q_{ki} and q_{ji} as functions of p_{ki} and the specific p_{li} and q_{li} in (2). This produces the demand for the generic input l by region i, viz.

---

[3] We have dropped the subscript i for notational convenience. We employ it when it is relevant to do so, e.g. in q_{ji} above or p_{ji} below.

[4] We can safely assume that M includes the prices of other inputs which we do not focus upon in our analysis.

$$q_{li} = \alpha \cdot p \cdot Q \cdot p_{li}^{-\mu} \cdot \Pi^{\mu-1} \tag{3}$$

where $\Pi \equiv \left(\int_R p_{ji}^{1-\mu} dj\right)^{\frac{1}{1-\mu}}$. Moreover, by replacing (3) in the expression for *X*, one obtains $X = \frac{\alpha \cdot p \cdot Q}{\Pi}$, which can be replaced in the production function to obtain the following expression for the productivity of labour

$$\frac{Q}{L} = \alpha^{\frac{\alpha}{1-\alpha}} \cdot G(\cdot)^{\frac{1}{1-\alpha}} \cdot p^{\frac{\alpha}{1-\alpha}} \cdot \Pi^{-\frac{\alpha}{1-\alpha}} \tag{4}$$

We now introduce explicitly in (4) the supply- and demand-side factors that are in $G(\cdot)$ and in the inverse demand function $p(\cdot)$ to obtain the equation that we estimate. We start with the demand-side factors. Total demand is the sum of the demand for good *i* by all the regions when the good is used as an intermediate input, and its demand by all the regions when it is used as a consumption good. To derive the latter, we make the standard assumption in these models that the utility of the consumers in any given region *l* is equal to the CES expression for *X* defined earlier (with the index *l* replacing *i*). The budget constraint is $\int_R p_{jl} q_{jl} dj = wL$, where *wL* is the total labour income in region *l*, which is used to buy the consumption goods. The first order condition relative to the labour input in the profit-maximisation problem of the firms implies that $wL = (1-\alpha) \cdot p \cdot Q$. By solving the consumer problem, the demand of region *l* for the consumption good produced in region *i* is equivalent to (3) pre-multiplied by $(1-\alpha)$ rather than α. The total demand of region *l* for the good produced in *i* is then $q_{il} = p \cdot Q \cdot p_{il}^{-\mu} \cdot \Pi^{\mu-1}$.[5] The total demand Q_i for the good *i* is obtained by aggregating

[5] Note that here the indices are reversed with respect to (3) because we are now considering the demand of *l* for the good *i*, rather than the other way around as we were doing when deriving (3).

over the individual demands of all the regions, viz. $Q_i = \int_R q_{ij} dj$. By using the expression above for q_{ij} (i.e. with index j in lieu of l), one obtains

$Q_i = \int_R p_j \cdot Q_j \cdot p_{ij}^{-\mu} \cdot \Pi_j^{\mu-1} dj$.

To obtain the inverse demand function, we assume that the trade costs enter as a mark-up of the price obtained by the producers, p_i, and that this mark-up is separable into two parts, one that depends on characteristics of the exporting region, which we denote by v^{-1}, and the other that depends on characteristics of the importing region, which we denote by z^{-1}. That is, $p_{ij} = p_i \cdot (v_i \cdot z_j)^{-1}$. When we substitute this expression for p_{ij} in the aggregate demand above, the factors that depend on i can be factored out of the integral sign. The inverse demand function for i becomes $p_i = \left(\frac{v_i^{\mu} \cdot \Psi_i}{Q_i} \right)^{\frac{1}{\mu}}$, where $\Psi_i \equiv \int_R p_j \cdot Q_j \cdot z_j^{\mu} \cdot \Pi_j^{\mu-1} dj$ collects all the terms that have remained inside the integral. The latter include in particular the expenditure capacity of the other regions, $p_j Q_j$, and the price index Π_j that they face.

The inverse demand for i is positively correlated with the factors v_i that account for the ability of the *ith* region to reduce the trade costs for its good. These are the factors that Redding and Venables (2001) denote as the *market access* of region i. Since reductions in the trade costs of selling the own goods to other locations can significantly enhance the potential market faced by a region, we take this to be an important dimension of openness in the sense that we are trying to assess here. Another way to think about v_i is that it denotes how responsive is a certain region to variations in the expenditures of the other regions.

We also introduce explicitly the effects of the trade costs on the inputs purchased by region i from the other regions. The price index Π in (4) can be rewritten as $\left(\int_R p_j^{1-\mu} v_j^{\mu-1} z_i^{\mu-1} dj\right)^{\frac{1}{1-\mu}}$. The component of the trade costs that depends on the ability of region i to reduce its trade costs for the goods purchased from the other regions, viz. z_i, can be factored out from the integral sign. Redding and Venables (2001) label these factors as the *supplier access* of the region. This is because an increase in these factors reduces the full price of the inputs purchased from elsewhere, which in turn increases labour productivity. These factors then represent another dimension of the openness of the regions.

To complete the derivation of the productivity equation to be estimated, we assume that $G(\cdot) = \left(\frac{Q}{A}\right)^{\beta} \cdot M^{\gamma}$. Replace this expression and the one for p in the productivity equation (4). Solve for Q. Eventually one obtains

$$\frac{Q}{L} = \alpha^{\frac{\alpha\mu}{\delta}} \cdot \left(\frac{L}{A}\right)^{\frac{\beta\mu}{\delta}} \cdot M^{\frac{\gamma\mu}{\delta}} \cdot (v \cdot z)^{\frac{\alpha\mu}{\delta}} \cdot L^{-\frac{\alpha}{\delta}} \cdot \Omega^{\frac{\alpha\mu}{\delta}} \cdot \Psi^{\frac{\alpha\mu}{\delta}} \qquad (5)$$

where $\Omega \equiv \left(\int_R p_j^{1-\mu} v_j^{\mu-1} dj\right)^{\frac{1}{1-\mu}}$, $\delta \equiv (1-\alpha-\beta)\cdot\mu + \alpha$, and we assume that $\delta > 0$. Equation (5) says that regional labour productivity depends on agglomeration economies (measured by employment density), other factors M that increase productivity (amongst which we include the technological capabilities of the regions proxied by patents), and $(v \cdot z)$. The latter combines market and supplier access. We will not be able to estimate these two factors separately. However, we can estimate their total effect through variables that account for the openness of the region.

Finally, labour productivity is affected by Ω and Ψ. Note that Ω and Ψ are aggregations of variables over all the other regions. As a result, they are roughly constant across observations *i*. In pratice not all the variability across *i*'s is likely to be eliminated. For example, the trade costs may not be perfectly separable, and some variation across *i*'s may still be captured by Ω and Ψ. Another possibility is that the regions face a different number of regions *R* with which they collaborate. The controls that we employ in our empirical estimation will in part capture these variation. We assume that any remaining variation is captured by the error term.

4. DETERMINANTS OF LABOUR PRODUCTIVITY IN EUROPEAN REGIONS

4.1 Sample, data, and variables

To estimate (5) we employ an unbalanced sample of NUTS European regions during 1989-1996. We obtained our data from the Eurostat data base REGIO. We were forced to use an unbalanced sample because REGIO contains several missing values. Also, we wanted to exploit the richness of controls and instruments available in this data base. This prevented us from performing our estimations at the disaggregated NUTS3 level since most of the potential controls and instruments are reported only for NUTS regions at a higher level of aggregation (NUTS2 or NUTS1). We constructed fairly homogeneous regions. We employed NUTS2 regions for Italy, Spain, and France (e.g Lombardy, Cataluña, Bretagne). We employed NUTS1 regions for Germany, Belgium, the Netherlands, and Portugal (e.g. Baden-Wüttenbger, Bayern, Region of Bruxelles). This is because their overall magnitude and administrative role within the country are akin to the NUTS2 regions of Italy, Spain or France.[6] We also employed NUTS1 regions for the UK. In this case we were forced to use NUTS1 rather than NUTS2 regions (e.g Eastern Regions instead of East Anglia or Essex) because there are too many missing values for the NUTS2 UK regions. In the end however, we only had few observations for UK regions because of several missing values.[7]

Our final sample is composed of 622 observations. This includes regions from the 8 countries mentioned above – Belgium, France, Germany, Italy, the Netherlands, Portugal, Spain, and the UK. The bulk of our sample, however, is composed of the Italian, French, Spanish and German regions. (See the Appendix.) Data for practically all the NUTS2 regions from the former three countries and all the NUTS1 German regions are available systematically for the entire period 1989-1996. For the other countries, the available data cover only some of the regions in some of the years. Since we use country and time dummies in our estimation, we included these regions in our sample because they represent genuine observations, and we have no reasons for discarding them. Table 1 lists the variables employed in our analysis, along with their definition. Table 2 reports descriptive statistics.

TABLE 1 AND 2 ABOUT HERE

As one can see from Table 1, for some of the variables we employed averages across years rather than the full panel variable. This is because of missing values for some of the years. In some cases, we computed the averages over a pre-sample period to avoid

[6] For instance, the regional governments of the NUTS1 German Länders perform functions that are similar to those of, say, the Italian regional governments, and in both cases they are the administrative sub-divisions of the country right below the national government.

[7] Caniels (1999) defined NUTS regions in a way similar to ours to obtain regions that were more comparable with one another.

potential endogeneity of the contemporaneous variables. Also, REGIO only provides the 1997 value of *NIGHT*. As we shall see in section 4.3 we use this variable to account for the touristic attractiveness of the region. Since touristic attractiveness does not change dramatically over time, the problem is not severe. For *HOUSPOP* we employed the average for the number of households during 1992-1994 because no other years were available.

4.2 The econometric specification

We estimated the following log-linear specification

$$log\left(\frac{Q_{it}}{L_{it}}\right) = constant + country\ dummies + time\ dummies +$$

$$\eta_1 \cdot log(KPAT_{it}) + \eta_2 \cdot log\left(\frac{L_{it}}{A_{it}}\right) + \eta_3 \cdot log(PASS_{it}) +$$

$$\eta_4 \cdot log(AGR_i) + \eta_5 \cdot log(ARABLE_i) + \eta_6 \cdot log(MTW_{it}) +$$

$$\eta_7 \cdot log(SUI_i) + \eta_8 \cdot log(L_{it}) + \varepsilon_{it} \qquad (6)$$

The terms *log(KPAT)*, *log(L/A)*, and *log(PASS)* are our variables of interest; *log(L)* is the log of total employment in the region; ε_{it} is an error term; and the η's are parameters to be estimated.[8] We will treat *log(KPAT)*, *log(L/A)*, *log(PASS)*, and *log(L)* as endogenous. All the other variables are exogenous controls.

We will discuss the variable *PASS* in greater detail in the next section. The discussion of the other variables is more straight forward. We follow Ciccone and Hall (1996) and

[8] Because there are a few regions with no airports, and hence *PASS=0*, we employed *log(1+PASS)*.

Ciccone (2000), and use employment density *(L/A)* to account for the extent of the agglomeration economies.[9] The stock of patents *KPAT* accounts for the technological intensity of the regions. While patents have well known limitations (e.g. they measure only the most important innovations, they do not take into account differences in the values of the innovations themselves), they are the most commonly used measure in cases like ours.[10] We also found that the use of R&D as an alternative measure was impratical because REGIO's series on regional R&D expenditures contained several missing values.[11] Moreover, *KPAT* is likely to capture other factors that we may want to control in our regression. For example, it is associated with differences in the educational levels of the regions, human capital, and similar characteristics. While we are unable to make these finer distinctions, we are content with the fact that *KPAT* enables us to control for some important effects that are correlated with the technological capabilities and other technology-related differences among our regions.

As far as the other controls are concerned, we use dummies to control for time- and country-specific effects. The use of country dummies is particularly important, as it implies that our results are likely to stem from genuine variations across regions, even within the same countries, and they will not depend mainly on variations across countries. The variables *AGR* and *ARABLE* control for the composition of the regional output, and particularly for the importance of agricultural activities. The motorways

[9] Because we do not use NUTS3 regions, our problem in using employment density is similar to the one faced by Ciccone and Hall (1996). In estimating the agglomeration economies in the US, they develop a measure of the employment density for the US States which is composed of the aggregate employment density at the State level and a correction factor that takes into account the differences in employment density in the individual counties within the State. To simplify our analysis, we assume that this correction factor is part of the error term of our regressions.

[10] See for instance Paci and Usai (2000) and the other regional studies on technology cited in the introductory section.

[11] We also employed the patent annual flows rather than the stocks with no significant changes in the results.

variable *MTW* proxies for infrastructures. This avoids that infrastructures be in fact captured by the employment density of the regions. The rationale for including the number of suicides, *SUI*, is that it is correlated with the general education of the region. Suicides are relatively more common in more advanced societies vis-à-vis poorer ones, and they are more common amongst more educated people.[12]

4.3 Airplane passengers as a measure of openness

The variable *PASS* is our measure of openness.[13] The mobility of people is correlated with several factors that we discussed earlier. For example, knowledge spillovers and other information flows are more pronounced if more people travel in and out of a location. Likewise, the presence of multinational enterprises commands airplane mobility. Airplane flights imply longer travels than the mere movement of people across the regional borders. They are then likely to be correlated, especially in Europe, with international travels and therefore with international openness.

Figure 1 offers a visual inspection of this variable by showing the distribution of *PASS* across the NUTS regions in our sample, as well as the distribution of the ratio between *PASS* and the population of the region. As the Figure shows, the distribution of *PASS* is skewed. However, almost all our regions had airports and airplane passengers. This suggests that this variable is an adequate measure even for smaller regions, and not just

[12] In the XIX century, the French sociologist Emile Durkheim wrote a famous essay entitled "The Suicide". He argued that suicides were more common "in industry than in agriculture", "amongst foremen rather than simple workers", and "in economically more developed countries". We also correlated our data on the regional labour income *(Q/L)* with the ratio between suicides and population, and obtained a correlation coefficient of 0.40.

[13] As indicated in Table 1, we defined *PASS* to be the average number of passengers over three years. This is to reduce the impact of yearly shocks. We also used a five year average for *PASS*, as well as simply annual passengers in our regressions with no major change in the results.

for the major locations.[14] REGIO provided potential alternative measures, notably the annual number of maritime passengers in the region, the annual maritime freight of goods, and the annual freight of goods embarked and disembarked by planes. Unlike airports, there are maritime passengers only in the relatively few regions with large harbours and that border with the sea. In addition, only few people travel by sea nowadays. The maritime freight of goods had much of the same problems, and only very special kinds of goods are moved by plane.

FIGURE 1 ABOUT HERE

To check the extent to which we can effectively use *PASS* as a proxy for openness, we run a number of experiments. First and foremost, we correlated it with some classical measures of openness in the trade literature. Since the latter variables are only available at the national level, from REGIO we aggregated *PASS* for all the 15 countries in the European Union during 1980-1996.[15] We obtained data on export, import and GDP for the same years and the same 15 countries from the World Bank's *World Development Indicators 1999*. The correlation coefficient between *PASS* over the population of the country, and trade openness (exports + imports over GDP) for the full sample period is 0.15. However, when we distinguished between 1989-1996, which corresponds to the sample period of our analysis in this paper, and 1980-1988, we obtained that the

[14] To avoid that it is actually biased towards the major hubs, we run all our regressions in this paper by using a dummy for the regions with major European airport hubs in our sample. We wanted to control for the fact that in these airports there may be many passengers in transit, or that they serve passengers who live in other, near-by regions. All the results in this paper are largely unaffected by the inclusion of this dummy.

[15] Since *PASS* is the average over the current and the previous two years, we used the annual number of airplane passengers that are available from REGIO since 1978.

correlation is 0.24 in the former case, and 0.11 in the latter.[16]

This is confirmed when we add controls. In the first part of Table 3, the OLS regression with trade openness as the dependent variable shows that for 1989-1996 the coefficient of *PASS* over the population of the country is sizable and statistically significant, after controlling for time dummies and the country's GDP per capita. The magnitude of this coefficient drops and becomes statistically insignificant for the same OLS regression during 1980-1988. It is also worth noting that the estimated impact of the GDP per capita does not change in the two sample periods, which reinforces the conjecture that it is something about *PASS* over population that is peculiar between the two periods. The increased economic "globalisation" during the 1990s, which is affecting classical trade as well, may be increasingly associated with spillovers, services, knowledge transfers, and other "soft" factors that are more strongly correlated with the mobility of people than in the past.

TABLES 3 ABOUT HERE

We also studied the correlation between *PASS* at the national level and other measures of openness. From the *European Science & Technology Indicator 1997* we obtained data for the same European countries on high-tech imports and exports during 1990-1995, and on inward and outward FDIs during 1990-1994. OLS regressions in the second part of Table 3 show that *PASS* is highly correlated with both high-tech imports and exports, and *PASS* over population is highly correlated with high-tech openness (high-tech imports + high-tech exports over GDP). The correlation with FDIs is less marked. However, *PASS* is positively correlated with inward FDIs, while it is not

[16] In computing these correlations and in the regressions below we are relating a variable, *PASS*, which is an average over three years, with yearly flows for the other variables. However, the results do not change when we consider yearly flows of passengers, or three-year averages for the other variables.

correlated with outward FDIs. All in all, Table 3 suggests that *PASS* is correlated with some classical measures of openness.

As a further check, we compared our regional measure of *PASS* with a measure of the sectoral specialisation of the regions themselves, in line with the classical view that openness and trade encourage specialisation. We computed the Herfindhal index of the shares of the regional value added in six sectors.[17] We found that the correlation coefficient between *PASS* over the population of the region and this index was 0.69.[18] Moreover, we run our regression (6) by using both *PASS* and this index, and by treating both variables as endogenous. When both variables were included, different estimation techniques (OLS, Two-Stage-Least Square, Generalised Method of Moments) implied either that both had a smaller and less significant effect, or that one of the two was significant and the other was not. This suggested to us that the two variables are measuring similar effects.

Finally, we clarify the assumptions that are implied by our use of *PASS* as a proxy for openness in (6). Rather than *PASS*, in (6) one would like to list all the variables that influence either v (i.e. the demand for the good of region i) or z (i.e. the trade costs of purchasing external inputs by region i). Let us summarise these variables by a vector $\boldsymbol{W}$, and take $\boldsymbol{W'}\cdot\boldsymbol{\chi}$ to be the linear combination of these factors where $\boldsymbol{\chi}$ is the vector of impacts of the factors in $\boldsymbol{W}$ on productivity. For simplicity call this index *OPEN*, viz. $OPEN \equiv \boldsymbol{W'}\cdot\boldsymbol{\chi}$. Now, if we observed all the variables in $\boldsymbol{W}$, we could simply list them as regressors in (6). Our problem is that it is difficult to collect data on all the factors $\boldsymbol{W}$

---

[17] These are Agriculture, Forestry & Fishery; Fuel & Power Products; Manufactured Products; Building & Construction; Market Services; Bank Services; Non-market Services.

[18] We also computed this correlation coefficient for the year 1992 to avoid that it be driven largely by the relatively similar values of these variables for the same region over the years. The correlation coefficient in this case was 0.67.

that may affect *v* or *z*. Even from a theoretical point of view, it is not clear what the full list of these variables should be (e.g. language skills, linkages through immigrants or other related factors, communications or other infrastructures).

Our specification (6) can then be interpreted as follows. The variable *log(PASS)* is affected by the same linear combination $\boldsymbol{W'}\cdot\chi$ that affects productivity plus an error term ξ – that is, $log(PASS) = \lambda \cdot OPEN + \xi$, where λ is the impact of *OPEN* on *log(PASS)*. In short, *log(PASS)* is the "true" index *OPEN* measured with error. But this implies that *log(PASS)* is negatively correlated with the error term of the productivity equation.[19] Other things being equal, OLS estimates will then underestimate the impact of *PASS*. Note also that the restriction that is implied by our assumption is that the set of unobserved factors that affect *v* or *z*, and hence productivity, is such that the ratio of the impacts of any two variables in ***W*** on log-productivity is equal to the ratio of the impacts of the same two variables on *log(PASS)*.[20]

4.4 Addressing endogeneity

In estimating (6) we face a classical endogeneity problem. The natural way of thinking about it is that we cannot be sure whether the potential correlation between *KPAT*, *(L/A)* and *PASS* on the one hand, and labour productivity *(Q/L)*, on the other, arises because *KPAT*, *(L/A)* and *PASS* affect *(Q/L)*, or the other way around. There are reasons for

[19] To see this, take (6) with *OPEN* rather than *log(PASS)* as a regressor. Since $OPEN = (1/\lambda)\cdot log(PASS) - (\xi/\lambda)$, and log(PASS) is postively correlated with ξ, then when we replace *OPEN* with *log(PASS)*, the latter is negatively correlated with the error term.

[20] In fact, the restriction is slightly weaker. Because we are using other controls in the regression, we could equally assume that $log(PASS) = \lambda\cdot OPEN + \boldsymbol{Z'}\cdot\zeta + \xi$, where ***Z*** is a vector of variables which also enter directly as regressors in the productivity equation, and ζ is a vector of impacts. If ***Z*** includes some of the variables in ***W***, then for each of those common variables and any other variable in ***W***, it no longer applies that there is a constant ratio between their impacts on *log(PASS)* and log- productivity.

both directions of causation. In the case of patents, while *KPAT* may augment labour productivity, higher labour productivity may provide more resources that encourage new investments in research and technology. Similarly, employment density may be higher because regions with higher incomes attract people. We already noted that *PASS* may be econometrically endogenous because of measurement errors. An additional source of endogeneity is that while the international openness of a region may induce higher productivity, the latter may encourage more intensive business activities, which leads to greater international mobility of people from and to the region. These problems entail that we have to resort to instrumental variable estimation. In order to be able to estimate the effects of our variables on regional productivity, rather than vice versa, we then need to find factors that account for differences in innovation, openness, or employment density independently of the regional incomes.

As far as employment density is concerned, Ciccone (2000) argued that the total land area of the region, *A*, is a powerful instrument to identify the effect of *(L/A)*. His argument is that the total area of a region is uncorrelated with changes in regional productivity. This is because the borders of the European regions were defined several years ago, and in most cases even more than one century ago. Yet, as he finds, and as we confirmed with our own data, the area of the regions is negatively correlated with their employment density. Therefore, while *A* is not affected by today's regional productivities, it is nonetheless correlated with the variable we are interested in, *(L/A)*.

The variables *NIGHT*, *MOTO*, and *SEA* are good instruments to identify the effect of *PASS*. *NIGHT* is the number of nights spent by non-residents in the region over the number of non-residents that visited the region. It is therefore a measure of the average number of nights spent by the visitors to the region. This is correlated with its touristic attractiveness. When people visit for business, they spend fewer days on average. By contrast, one is likely to stay longer in touristic areas. The correlation between *NIGHT* and the touristic attractiveness of the region is apparent from Table 4, which lists the top 20 regions in the REGIO data base ranked by *NIGHT*. A simple inspection of Table 4 reveals that these are all highly touristic regions.[21]

TABLE 4 ABOUT HERE

However, a region visited because of touristic attractions is likely to imply greater openness and exchange for other purposes as well. For example, tourism may induce the construction of larger airports or it implies a higher number of flights per day, which can also be used for business. Similarly, with tourism, people are likely to speak more languages, which encourages international openness. At the same time, tourism covers a small share of regional economic activities. For example, a recent official report of the Italian Ministry of Industry on the economic perspectives of tourism in Italy indicated that direct and indirect activities linked to the tourist sector account for 5% of the Italian GDP on average, and for 8% of the GDP of the most touristic regions. (See *Ministero dell'Industria*, 2000) Since these figures include activities that are very indirectly associated with tourism (e.g. the food industry), the effective share of relevant touristic activities is probably quite smaller. This means that at the aggregate level, the direct effect of tourism on productivity is negligible for most of our regions. This may then be a reasonable exclusion restriction for our purposes.[22]

[21] Table 4 is constructed using all the regions in the REGIO data base for which *NIGHT* was available. It also includes regions that are not in our final sample.

[22] In this respect, the fact that we deal with relatively large territorial areas, like the NUTS1 or NUTS2 regions, helps. If we used NUTS3 or even smaller areas, the share of tourism, and hence the potential direct effect of *NIGHT* on productivity, would have probably been more important.

The dummy *SEA* is another good instrument for *PASS*. Historically, the sea has been a major factor in enhancing communication and openness, and from there economic growth. This was very clear to the French historian Fernand Braudel who glorified the role of the Mediterranean sea for enhancing the Italian economic development during the XV and XVI centuries: *"La mer attire elle les trois continents qui l'enserrent, elle les transperce. C'est l'eau, le trafic de la mer Intérieure qui créent la gloire vénitienne. La mer est plus que jamais le centre du monde."* (Braudel, 1996, p.451.) Such a role of the sea has continued unabated since our very own days, as suggested by the recent geography literature. Overman, Redding, and Venables (2001) report that countries that are landlocked have 50% higher transportation costs and 60% lower trade volumes. Since regions that border with the sea are associated with more intensive transportation activities and related infrastructures, this may have a direct effect on productivity, which may cast some doubt about the exclusion of *SEA* among the regressors of (6). But one of the reasons for employing *MTW* in (6) was to control for such transportation activities and the associated infrastructures. Moreover, as we shall see in the next Section, we check the robustness of our empirical results by running our productivity regression under different exclusion restrictions for our instruments, with no appreciable changes in the results. The number of motorcycles, *MOTO*, is also likely to be correlated with the pleasantness of the regional weather. In addition, we found that *MOTO* over population is positively correlated with *NIGHT*, which confirms that that the former may be associated with places where life is more pleasant. Since there is no special reason why *MOTO* should directly affect productivity, we employed it as another instrument.

In (6) *log(L)* is also endogenous. We then included *HOUSPOP* and the average population in working age in the region (*POP25-65*) in our set of instruments These are both factors that affect the labour supply, and hence *L*. For instance, the number of households per inhabitant may reflect sociological characteristics of the family structures. Thus, regions where people marry earlier, or simply where young people leave their parents' house earlier, are more likely to have a larger labour supply, which would in turn affect *L* independently of the regional productivity.[23] Similarly, the population in working age reflects whether a region is composed of a relatively young or old population, which would also affect the labour supply, and hence *L*.[24] Finally, there are enough instruments and exclusion restrictions so far to be able to identify the effect of *KPAT* as well.

4.5 Empirical results and robustness check

Table 5 reports our empirical results obtained by using OLS, Two-Stage-Least Squares (2SLS), and the Generalised Method of Moment (GMM). The results are fairly robust across the three types of estimation techniques. The elasticities with respect to the stock of patents, *KPAT*, and the employment density, *(L/A)*, are around 10-11%, and they are statistically significant. This confirms the importance of innovation and the technological capabilities of the regions in raising labour productivity. The estimated effect of the agglomeration economies is higher than the one estimated by Ciccone

---

[23] Unfortunately the number of household in the pre-sample period was not available from REGIO. The 1992-1994 average may be affected by changes in the population of the region during our sample period, which could create some potential endogeneity problem for this variable. We can only argue that the sociological characteristics of the family structures are unlikely to change in the short-run.

[24] If there is inter-regional migration, the working age population of a region may also be endogenous. We use however a pre-sample average for *POP25-65* rather than its yearly measure during the sample period, which mitigates the problem.

(2000). However, Ciccone's estimates are based on a sample of NUTS3 regions, while we use a wider spatial aggregation. Ciccone found that there are sizable spillovers across neighbouring regions. Our estimation has internalised these spillovers.

TABLE 5 ABOUT HERE

The estimated elasticity with respect to *PASS* is about 3.5% in the 2SLS and GMM estimations, and 1.2% in the OLS estimation. In all three cases the effect is statistically significant. We therefore find that there is an independent effect of openness on regional productivity in Europe. This effect occurs in addition to, say, the potential increase in demand due to other desirable characteristics of the goods produced by the region, like its innovation and technological content.

The higher value of the two instrumental variable estimates suggests that the effect of the measurement error in *PASS* dominates that of other sources of endogeneity. Interestingly, this is the same finding as in Frankel and Romer (1999), who also obtained higher instrumental variable estimates of the effect of their measure of trade openness on income than OLS. They also appealed to measurement error problems, and argued largely in the same fashion as we have done in this paper. They maintain that their measure of trade openness is an imperfect measure of the actual set of interactions with other countries that represent the real determinants of higher incomes. The implied measurement error introduces a negative correlation between the error of the estimated equation and the trade openness regressor, which creates a downward bias in the OLS estimate. Like trade openness, we maintain that *PASS* is an imperfect proxy for the set of factors that give rise to a higher market or supplier access of countries or regions. Moreover, to the extent that the mobility of people has become more tightly linked to the production of services and more generally to the production or use of intangible inputs and outputs, and these economic activities have become more important than in the past, *PASS* is an interesting proxy for the relevant factors that account for the economically valuable openness of a location.

Finally, we performed some robustness check to evaluate whether our estimates are affected by the exclusion restrictions that we have imposed in Table 5. Table 6 presents the GMM estimates of (6) after including any one of *NIGHT*, *MOTO*, and *SEA* as regressors in the equation, as well as any two of them. The estimated parameters, and particularly those of *KPAT*, *(L/A)*, and *PASS*, are largely unaffected by the set of instruments used. Also, whenever *NIGHT* appears as a direct regressor, its impact is small and statistically insignificant. This is consistent with the remark made earlier that tourism has a small weight on the economy of our regions. Hence, the direct effect of a variable like *NIGHT*, which is correlated with the touristic attractiveness of the region, appears to be unimportant. This suggests that the exclusion of *NIGHT* in the regression, and its use to identify the effect of *PASS*, is a reasonable restriction.

TABLE 6 ABOUT HERE

5. CONCLUSIONS

The determinants of labour productivity is one of the most widely asked questions in economics, and increasingly the literature is paying attention to this question at the level of regions rather than countries. However, previous studies have focussed either on regional technological capabilities or on agglomeration economies. This paper is one of the first attempts to compare three major potential explanations of regional advantages: Technological capabilities, agglomeration economies, and openness. By looking at the

NUTS European regions during 1989-1996, we find that both technology and agglomeration economies have a significant and sizable impact. We also find that the openness of the regions has an additional independent effect.

This suggests that policies aimed at encouraging regional development should not focus only on factors that are internal to the localities, like local infrastructures, local networks, etc.. Actions aimed at making the regions more "cosmopolitan" are also important. In the paper, we were unable to distinguish whether the effects of openness depended on the ability of the regions to access larger potential markets for their goods, or on other factors, like spillovers due to mobile and internationalised human capital, the presence of multinational corporations, or else. These can be important topics for future and more focussed research. However, the experience of some of the fast growing regions of the world today (e.g. the Asian Tigers, or countries like Ireland and Israel) indicate that these factors are probably very correlated with one another. In short, there may be underlying factors that account for the extent to which some regions are more open than others, and we found that these underlying factors matter.

Appendix

List of the regions used in the empirical analysis

Région Bruxelles-capitale/ Brussels hoofdstad gewest	be1
Vlaams Gewest	be2
Région Wallonne	be3
Baden-Wurttemberg	de1
Bayern	de2
Berlin	de3
Bremen	de5
Hamburg	de6

Hessen	de7
Niedersachsen	de9
Nordrhein-Westfalen	dea
Rheinland-Pfalz	deb
Saarland	dec
Schleswig-Holstein	def
Galicia	es11
Principado de Asturias	es12
Cantabria	es13
Pais Vasco	es21
Comunidad Foral de Navarra	es22
La Rioja	es23
Aragón	es24
Comunidad de Madrid	es3
Castilla y León	es41
Castilla-la Mancha	es42
Extremadura	es43
Cataluña	es51
Comunidad Valenciana	es52
Baleares	es53
Andalucia	es61
Murcia	es62
Canarias (ES)	es7
Île de France	fr1
Champagne-Ardenne	fr21
Picardie	fr22
Haute-Normandie	fr23
Centre	fr24
Basse-Normandie	fr25
Bourgogne	fr26
Nord - Pas-de-Calais	fr3
Lorraine	fr41
Alsace	fr42
Franche-Comté	fr43
Pays de la Loire	fr51
Bretagne	fr52
Poitou-Charentes	fr53
Aquitaine	fr61
Midi-Pyrénées	fr62
Limousin	fr63
Rhône-Alpes	fr71
Auvergne	fr72

Languedoc-Roussillon	fr81
Provence-Alpes-Côte d'Azur	fr82
Piemonte	it11
Valle d'Aosta	it12
Liguria	it13
Lombardia	it2
Trentino-Alto Adige	it31
Veneto	it32
Friuli-Venezia Giulia	it33
Emilia-Romagna	it4
Toscana	it51
Umbria	it52
Marche	it53
Lazio	it6
Abruzzo	it71
Molise	it72
Campania	it8
Puglia	it91
Basilicata	it92
Calabria	it93
Sicilia	ita
Sardegna	itb
Noord-Nederland	nl1
Oost-Nederland	nl2
West-Nederland	nl3
Zuid-Nederland	nl4
Portugal (Continent)	pt1
Yorkshire and The Humber	uke
East Midlands	ukf
West Midlands	ukg
South West	Ukk
Wales	Ukl

References

Acs, Z. (ed.) (2000) *Regional Innovation, Knowledge, and Globalization*, Pinter Publishers, London.

Arora, A., Gambardella, A. and Torrisi, S. (2001) "In the Footsteps of the Silicon Valley? Indian and Irish Software in the International Division of Labour", Working Paper N.41, Stanford Institute for Economic Policy Research (SIEPR), Stanford University, Stanford CA.

Arora, A., Arunachalam, V.S., Asundi, J., Fernandes, R. (2001), "The Indian Software Services Industry", *Research Policy*, Vol. 30 (8), 1267-1287.

Arthur B. W. (1990) "Silicon valley Locational Clusters: When Do Increasing Returns Imply Monopoly?", *Mathematical Social Sciences*, 19, 235-251.

Audretsch, D.B. and M. P. Feldman (1996), "Knowledge Spillovers and the Geography of Innovation and Production", *American Economic Review*, 86 (3), 630-640.

Braudel, F. (1996), *Autour de la Méditerranée*, Editions de Fallois, Paris.

Breschi S. (1999) *Spatial Patterns of Innovation: Evidence from Patent Data*, in A. Gambardella, F. Malerba (eds.), *The Organisation of Innovative Activity in Europe*, Cambridge university Press, Cambridge (MA)

Bresnahan, T., Gambardella, A., and Saxenian, A. (2002), "'Old' Economy inputs for 'New Economy' Outcomes: Cluster Formation in the new Silicon Valleys", *Industrial and Corporate Change*, forthcoming.

Business Week (2001) "The Great English Divide", August 13, 36-40, European Edition.

Caniels M. (1999) *Regional Growth Differentials. The Impact of Locally Bounded Knowledge Spillovers*, Ph.D dissertation, University of Maastricht.

Ciccone A. and R. E. Hall (1996) *Productivity and the Density of Economic Activity*, American Economic Review, 86 (1), 54-70.

Ciccone A. (2000) "Agglomeration Effects in Europe", mimeo, Universitat Pompeu Fabra, Barcelona, Spain.

Coe D. and E. Helpman (1995) "International R&D Spillovers" *European Economic Review*, 39, 859-887.

De Fontenay, C. and Carmel, E. (2001) "Israeli's Silicon Wadi: The Forces Behind Cluster Formation", Working Paper N.40, Stanford Institute for Economic Policy Research (SIEPR), Stanford University, Stanford CA.

Frankel, J. and Romer, P. (1999), "Does Trade Cause Growth?", *American Economic Review*, 89 (3), 379-99.

Hirschman A. O. (1958) *The Strategy of Economic Development*, Yale Studies in Economics, n. 10, New Haven.

Jaffe, A., Trajtenberg, M., and R. Henderson (1993) "Geographic Localization of Knowledge Spillovers as Evidenced by Patent Citations", *Quarterly Journal of Economics*, 63 (3), 577-598.

Krugman P. (1991), *Geography and Trade*, MIT Press, Cambridge, MA.

Krugman, P and Venables, A.J. (1995), "Globalisation and the Inequality of Nations", *Quarterly Journal of Economic*s, 110(4), 857-80.

Marshall, A. (1920) *Principles of Economics*, 8th Ed. London: MacMillan.

Midelfart-Knarvik, K.H., Overman, H., and Venables, A. (2000), "Comparative Advantage and Economic Geography", CEPR Discussion Paper, 2618.

Ministero dell'Industria (2000), *Nono rapporto sul Turismo Italiano 2000*, Ministero dell'Industria, Rome.

Myrdal G. (1957) *Rich Lands and Poor*, World Perpectives Series, Vol. 16. Harper Brothers, New York.

Overman, H., Redding, S., and Venables, A.J. (2001) "The Economic Geography of Trade, Production and Income: A Survey of Empirics", paper prepared for the *World Bank ABCDE*, April.

Paci R. and S. Usai (2000) "Technological Enclaves and Industrial Districts: An Analysis of the Regional Distribution of Innovative Activity in Europe", *Regional Studies*, 34 (2), 97-114.

Perroux F.(1950) "Economic Space: Theory and Applications", *Quarterly Journal of Economics*, 64, 89-104.

Porter M. E. (1998) "Clusters and the New Economics of Competition", *Harvard Business Review*, Nov-Dec, 77-90.

Puga, D. (1999) "The Rise and Fall of Regional Inequalities", *European Economic Review* 43, 303-334.

Redding, S. and Venables, A. (2001) "Economic Geography and International Inequality", CEPR Discussion Paper, 2568.

Rodriguez-Claire A., (1996) "Multinationals, Linkages, and Economic Development", *American Economic Review*, 86 (4), 852-873.

Saxenian A. (1994) *Regional Advantage. Culture and Competition in Silicon Valley and Route 128*, Harvard University Press, Cambridge (MA).

Saxenian A. (2001) "Taiwan's Hsinchu Region: Imitator and Partner for Silicon Valley", Working Paper N.44, Stanford Institute for Economic Policy Research (SIEPR), Stanford University, Stanford CA.

Swann P., Prevezer M. and D. Stout (eds.) (1998) *The Dynamics of Industrial Clustering. International Comparisons in Computing and Biotechnology*, Oxford University Press, Oxford UK.

Verspagen B. (1997) "European Regional Clubs: Do They Exist and Where Are They Heading? On Economic and Technological Differences between European Regions", in Adams J. and Pigliaru F. (eds.) *Economic Growth and Change: National and Regional Patterns of Convergence and Divergence*, Edward Elgar, Cheltenham.

Table 1: List of variables

Q_{it} / L_{it}	Regional GDP (in PPP and corrected for inflation) over number of people employed in the region, 1989-1996 [in 000 euros].
L_{it}	Number of people employed in the region, 1989-1996 [in 000].
$KPAT_{it}$	Stock of 1989-1996 European patent applications in the region, computed from the number of annual patent applications using a 0.25 depreciation rate. Initial value of stock for 1989 (first year of available patent applications in REGIO) obtained as the ratio between the 1989 number of patent applications in the region and the depreciation rate, 0.25.
$PASS_{it}$	Average annual number of airplane passengers embarked and disembarked in the region during the past three years, 1989-1996 [in 000].
AGR_i	Utilised agricultural area, average for 1984-1988 [in Km^2].
$ARABLE_i$	Arable land, average for 1984-1988 [in Km^2].
$NIGHT_i$	Number of nights spent in the region per non-resident arrived in the region (data for 1997).
MTW_{it}	Motorways in the region in 1989-1996 [in Km].
SEA_i	Dummy equal to 1 if region borders with the sea.
A_i	Area of the region [in Km^2]
$MOTO_{it}$	Number of motorcycles over 50 cm^3 owned by residents in the region, 1989-1996 [in 000].
SUI_i	Number of suicides in the region, average for 1985-1988.
$HOUSPOP_i$	Family structure, 1992-1994 average number of households in the region over 1989-1996 average population in the region.
$POP25\text{-}65_i$	Population of age between 25 and 65, average for 1985-1988 [in 000].

Table 2: Descriptive Statistics

Variable	Mean	Standard Deviation	Minimum	Maximum
Q_{it}/ L_{it}	36.5	5.7	18.0	57.5
L_{it}	1320.9	1331.5	50.8	7544.4
L_{it}/A_i	0.130	0.345	0.006	3.748
$KPAT_{it}$	1131.3	2491.8	2.0	12680.0
$PASS_{it}$	2107.6	3944.8	0.0	27998.3
AGR_i	1241.0	1130.2	0.7	5688.2
$ARABLE_i$	699.6	743.5	0.5	3999.1
$NIGHT_i$	2.8	1.7	1.3	10.0
MTW_{it}	427.4	396.0	0.0	2192.0
SEA_i	0.5	0.5	0.0	1.0
A_i	22806.5	19722.6	161.4	94193.9
$MOTO_{it}$	86.7	102.4	4.8	547.0
SUI_i	457.0	477.7	19.0	2609.8
$HOUSPOP_i$	0.376	0.056	0.273	0.566
$POP25\text{-}65_i$	1656.0	1534.4	62.1	9069.7

N. of observations = 622.

Table 3: Correlations between PASS and some "classical" measures of openness, OLS regressions for EU-15 countries

Trade openness

1990s

Dependent Variable	Const.	PASS	POP	GDP	PASS/POP	GDP/POP
Imports (R^2 = 0.94)	27351.6 (8493.65)	0.229 (0.206)	-0.969 (0.524)	0.248 (0.020)	--	--
Exports (R^2 = 0.94)	25312.7 (8845.98)	0.040 (0.168)	-1.171 (0.476)	0.277 (0.019)	--	--
Open (R^2 = 0.29)	21.85 (17.44)	--	--	--	14.571 (6.79)	3.574 (0.663)

Sample period 1989-1996. N. obs. = 106. Heteroskedastic consistent standard erros. All equations include time dummies. Open = (Imports + Exports)/GDP.

1980s

Dependent Variable	Const.	PASS	POP	GDP	PASS/POP	GDP/POP
Imports (R^2 = 0.89)	15993.1 (4714.1)	-0.068 (0.198)	-0.312 (0.225)	0.243 (0.023)	--	--
Exports (R^2 = 0.87)	14970.5 (4890.96)	-0.127 (0.174)	-0.349 (0.197)	0.251 (0.022)	--	--
Open (R^2 = 0.09)	50.15 (18.42)	--	--	--	4.520 (8.096)	3.427 (1.101)

Sample period 1980-1988. N. obs. = 105. Heteroskedastic consistent standard erros. All equations include time dummies. Open = (Imports + Exports)/GDP.

High-tech trade

Dependent Variable	Const.	PASS	POP	GDP	PASS/POP	GDP/POP
Imports (R^2 = 0.90)	2159.57 (1035.86)	0.152 (0.041)	-0.375 (0.089)	0.034 (0.004)	--	--
Exports (R^2 = 0.85)	2051.43 (1394.31)	0.233 (0.053)	-0.689 (0.120)	0.046 (0.006)	--	--
Open (R^2 = 0.05)	0.020 (0.025)	--	--	--	0.023 (0.012)	0.001 (0.001)

Sample period 1990-1995. N. obs. = 75. Heteroskedastic consistent standard erros. All equations include time dummies. Open = (High-tech Imports + High-Tech Exports)/GDP.

Foreign Direct Investments

Dependent Variable	Const.	PASS	POP	GDP	PASS/POP	GDP/POP
Inward	3289.88	0.116	-0.051	-0.001	--	--
(R^2=0.43)	(1203.94)	(0.031)	(0.069)	(0.003)		
Outward	2943.30	-0.011	-0.155	0.018	--	--
(R^2=0.62)	(1235.77)	(0.035)	(0.117)	(0.005)		
Open	0.052	--	--	--	-0.004	-0.001
(R^2=0.03)	(0.015)				(0.007)	(0.001)

Sample period 1990-1994. N. obs. = 62. Heteroskedastic consistent standard erros. All equations include time dummies. Open = (Inward + Outward)/GDP.

Table 4: Top 20 regions ranked by NIGHT

	NIGHT
Canarias (ES)	10.0
Baleares (ES)	9.9
Notio Aigaio (GR)	9.1
Ionia Nisia (GR)	9.0
Voreio Aigaio (GR)	8.8
Kriti (GR)	8.3
Madeira (PT)	8.1
Comunidad Valenciana (ES)	6.6
Anatoliki Makedonia, Thraki (GR)	5.7
Trentino-Alto Adige (IT)	5.7
Marche (IT)	5.5
Kentriki Makedonia (GR)	5.5
London (UK)	5.3
South West (UK)	5.2
Scotland (UK)	5.2
Sardegna (IT)	5.2
North East (UK)	5.0
Northern Ireland (UK)	5.0
Eastern (UK)	5.0
Abruzzo (IT)	4.9

Table 5: Determinants of labour productivity – OLS, 2SLS, and GMM estimation

Dependent variable log(Q_{it}/L_{it})

	OLS	2SLS	GMM
Const.	4.275	4.538	4.560
	(0.087)	(0.096)	(0.094)
log($KPAT_{it}$)	0.087	0.114	0.111
	(0.005)	(0.007)	(0.007)
log(L_{it}/A_{it})	0.120	0.103	0.105
	(0.020)	(0.019)	(0.019)
log($PASS_{it}$)	0.012	0.034	0.036
	(0.002)	(0.005)	(0.005)
log(AGR_i)	0.084	0.086	0.089
	(0.019)	(0.017)	(0.016)
log($ARABLE_i$)	- 0.017	- 0.007	-0.007
	(0.006)	(0.007)	(0.007)
log(MTW_{it})	0.013	0.017	0.019
	(0.002)	(0.003)	(0.003)
log(SUI_i)	0.027	0.032	0.043
	(0.014)	(0.018)	(0.017)
log(L_{it})	-0.239	-0.340	-0.350
	(0.027)	(0.031)	(0.029)
N. of obs.	622	622	622
Adjusted R^2	0.76	0.72	0.71

Heteroscedastic consistent Standard Errors in parenthesis. All equations include time and country dummies. 2SLS and GMM employ the following instruments: constant, time dummies, country dummies, *log(AGR)*, *log(ARABLE)*, *log(NIGHT)*, *log(MTW)*, *SEA*, *log(A)*, *log(SUI)*, *log(MOTO)*, *log(HOUSPOP)*, *log(POPM25-65)*.

Table 6: Determinants of labour productivity, alternative specifications – GMM

Dependent variable $log(Q_{it}/L_{it})$

Const.	4.613 (0.113)	4.559 (0.113)	4.546 (0.100)	4.744 (0.201)	4.598 (0.118)	4.402 (0.169)
$log(KPAT_{it})$	0.105 (0.009)	0.111 (0.007)	0.124 (0.015)	0.097 (0.014)	0.117 (0.017)	0.144 (0.024)
$log(L_{it}/A_{it})$	0.105 (0.020)	0.105 (0.019)	0.096 (0.023)	0.106 (0.022)	0.095 (0.024)	0.081 (0.028)
$log(PASS_{it})$	0.046 (0.011)	0.036 (0.007)	0.042 (0.009)	0.061 (0.022)	0.053 (0.015)	0.043 (0.009)
$log(AGR_i)$	0.094 (0.018)	0.089 (0.016)	0.090 (0.017)	0.099 (0.020)	0.094 (0.019)	0.087 (0.018)
$log(ARABLE_i)$	-0.006 (0.007)	-0.007 (0.007)	-0.005 (0.008)	-0.002 (0.010)	- 0.004 (0.008)	-0.007 (0.007)
$log(MTW_{it})$	0.023 (0.005)	0.019 (0.004)	0.021 (0.004)	0.030 (0.009)	0.026 (0.006)	0.020 (0.004)
$log(SUI_i)$	0.053 (0.014)	0.043 (0.018)	0.039 (0.019)	0.069 (0.029)	0.050 (0.021)	0.023 (0.025)
SEA_i	0.022 (0.021)	--	--	0.040 (0.031)	0.023 (0.022)	--
$log(NIGHT_i)$	--	0.000 (0.014)	--	-0.018 (0.021)	--	0.024 (0.022)
$log(MOTO_{it})$	--	--	-0.023 (0.024)	--	-0.023 (0.025)	-0.054 (0.039)
$log(L_{it})$	-0.383 (0.041)	-0.354 (0.034)	-0.359 (0.031)	-0.431 (0.074)	-0.388 (0.044)	-0.329 (0.042)
N. of obs.	622	622	622	622	622	622
Adjusted R^2	0.68	0.72	0.68	0.61	0.64	0.67

Heteroscedastic consistent Standard Errors in parenthesis. All equations include time and country dummies. The instruments are listed in Table 5.

Figure 1: Distribution of *PASS* across Regions (data for 1992)