Laboratory of Economics and Management
Sant'Anna School of Advanced Studies

Via Carducci, 40 - I-56127 PISA (Italy)
Tel. +39-050-883-341 Fax +39-050-883-344
Email: lem@sssup.it Web Page: http://lem.sssup.it

LEM

Working Paper Series

How Provincial is your Region?

Effects on Labour Productivity and Employment in Europe

Alfonso Gambardella[*]
Myriam Mariani [†]
and
Salvatore Torrisi [o]

[*] *Sant'Anna School of Advanced Studies, Pisa*
[†] *University of Camerino, Camerino, Italy and MERIT, University of Maastricht, Maastricht*
[o] *University of Camerino, Camerino, Italy*

2001/03 **February 2001**

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

Alfonso Gambardella
Sant'Anna School of Advanced Studies, Pisa
agambardella@info-net.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
storrisi@unicam.camserv.it

February 2001

We thank Stefano Breschi for helpful comments and suggestions on an earlier draft of this paper. We also acknowledge comments from Marco Darin, Alessandro Sembenelli, and the other participants in the MURST Project "Information Spillovers and Technological Diffusion in Italy and Europe". We had useful discussions with Davide Castellani, Mario Pianta, Roberto Schiattarella, David Soskice, Davide Ticchi, and Antonello Zanfei. 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)

Abstract

This paper estimates the determinants of labour productivity and employment in European NUTS2 regions. We focus on technological capabilities (proxied by regional patents), agglomeration economies (employment density), and openness, proxied by the number of airplane passengers embarked and disembarked in the region. We employ 1989-1996 data drawn from the Eurostat REGIO data base. By using instrumental variables, we confirm existing results in the literature that patents and employment density affect labour productivity. Our novel finding is that openness affects labour productivity as well. This suggests that regional advantages also stem from the ability of the regions to connect to the world that is outside them, and not just on internal factors like local infrastructures, local networks, etc.. In addition, we find that technological capabilities affect employment, while the effect of agglomeration economies and openness on the latter is less marked. Thus, technology seems to be the crucial variable for a thorough regional development. Agglomeration economies and openness benefit mostly those who are already employed, as it implies increases in their incomes with limited increases in employment.

1. INTRODUCTION

The economics of regions has drawn increasing attention in recent years. While the topic is rooted in the pioneering 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 the US States. This calls for a better understanding of the determinants of such 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. Saxenian, 1994; Porter, 1998; Swann, Prevetzer, and Stout, 1998). Krugman (1991) noted that these advantages may stem from initial conditions that are reinforced over time because of the increasing returns associated with the formation of a critical mass of economic activities. (See also Arthur, 1990.) Ciccone and Hall (1996), and Ciccone (2000) estimated the extent of the agglomeration economies. They find 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 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 regions can be assimilated to small countries, the extent to which they are linked to the world that is outside them can play an important role in explaining their performance. Many of the fast-growing regions of the world today exhibit significant international openness – e.g. Taiwan, Singapore, South Korea, Ireland, Israel, or the software industry in Bangalore or other Indian regions. (See for instance Arora and Jasundi, 1999; Saxenian, 2001; Arora, Gambardella, and Torrisi, 2001.)

This paper explores empirically the extent to which apart from technogical capabilities and agglomeration economies, the openness of the regions affects their economic performance. The analysis employs an unbalanced sample of NUTS2 European regions between 1989-1996 drawn from the official Eurostat data base REGIO. We first estimate an equation of regional labour productivity. Since the latter is equivalent to labour income in equilibrium, in this paper we shall use the two terms interchangeably. Apart from controls, we regress labour productivity on three variables. First, following Ciccone and Hall (1996), we derive a parameter to be estimated which accounts for the effects of the employment density of the regions, and hence for agglomeration economies. Second, our equation includes the stock of patents applied for by the inventors located in the regions. 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.

This finding raises the question about the meaning of openness and the mechanisms that induce greater labour productivity. First, more open regions can take greater advantage of international spillovers (Coe and Helpman, 1995), which 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. Second, these regions may experience a larger presence of multinational enterprises. To the extent that multinational firms replace less productive investments by local companies, this increases regional labour productivity (e.g. Rodriguez-Claire, 1996). Third, they may face a larger (international) demand. In addition, as the classical theory of international trade would suggest, open regions are more specialised, as they are part of a division of labour with other countries or regions. Specialisation would then have beneficial effects on their productivity.

Unfortunately, our data set is not fine enough to distinguish amongst these effects. In particular, we will not be able to distinguish between supply-side effects (international spillovers, or the productivity advantages associated with specialisation), and demand-side effects (e.g., greater "foreign" demand). However, the factors that we mentioned above are typically correlated with one another. 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; they are open to multinational enterprises; they show a pronounced specialisation of their international industries, and they are part of an extensive division of labour particularly with the US. (See Saxenian, 2001; Arora, Gambardella, and Torrisi, 2001.) This suggests that there could be more basic factors that reduce the costs of international openness. A related perspective is in terms of the present debate about the effects of globalisation. In this view, the scope of this paper would be to assess whether, apart from technology, agglomeration economies and other factors, there are net advantages in terms of labour income and employment accruing to more global regions as compared to less cosmopolitan ones.

We also estimated a more elaborate system composed of a labour income and a labour supply equation. This enables us to assess the effects of agglomeration economies, patents and our measure of openness on employment as well. We estimated two models. In the first model, we assume that labour markets clear, and hence the number of people actually employed is our labour supply. We label this the neoclassical model. In the second model, we assume that the labour supply is the active population of our regions (labour force). A simple inspection of our data revealed the well known fact about Europe that the active population is normally greater than the people employed.

We label this the keynesian model. This also implies that while the neoclassical model is composed of two equations (a labour income and a labour supply equation), the keynesian model is composed of three equations – a labour income equation, a labour supply equation, and an equation for the regional employment rate.

We find that agglomeration economies (employment density), patents and passengers are still important determinants of labour income both in the neoclassical and in the keynesian model. We also find that patents increase employment in both models, while employment density and openness raise employment only in the neoclassical model. This suggests that technology is the crucial factor for a thorough regional development. We also interpret our results as suggesting that with flexible labour markets – as implied by the neoclassical assumptions – openness and agglomeration economies have more pronounced effects on employment.

The paper is organised as follows. The next section presents a simple model as a guide to interpret our empirical results. Section 3 presents the data used in the analysis, the econometric specification, and the estimates of our labour productivity equation. Section 4 presents our models with employment, and the empirical results. Section 5 concludes.

2. THE BASIC MODEL

Our basic model combines demand and supply conditions to derive an expression for the determinants of the labour productivity of the regions. We start from the supply side, and assume that all the firms in a region produce the same good, which is different from the good produced in the other regions. The firms in region *i* take the price p_i as given, and produce their individual output q according to the following production function

$$q = G(\frac{Q}{A}, X) \cdot k^{a} \cdot l^{1-a} \tag{1}$$

where *k* and *l* are the quantity of the capital stock and labour employed, a is the capital share, and *G(·)* is a function of other factors that affect the productivity of the firms. Apart from a vector of variables *X*, 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. Denote *r* and *w* to be the prices of the capital stock and labour. These prices are common to all the firms in the regions, which take them as given.

The first order conditions of the profit-maximising firms with respect to *k* and *l* are

$$a \cdot \frac{q}{k} = \frac{r}{p} \tag{2a}$$

$$(1-a) \cdot \frac{q}{l} = \frac{w}{p} \tag{2b}$$

Replace the expression for *q* given by (1) in (2a), and solve for *k*. One obtains $k = \left(\frac{a \cdot p \cdot G}{r} \right)^{\frac{1}{1-a}} \cdot l$. Replace this expression in the production function (1), and obtain $q = G^{\frac{1}{1-a}} \left(\frac{a \cdot p}{r} \right)^{\frac{a}{1-a}} \cdot l$. In this expression only *q* and *l* are specific to the individual firm. Hence, by summing up both sides over all the firms in the region, the aggregate supply function is

$$Q = G^{\frac{1}{1-a}} \left(\frac{a \cdot p}{r} \right)^{\frac{a}{1-a}} \cdot L \tag{3}$$

where *Q* and *L* are the aggregate output and the aggregate level of employment.

To account for demand effects, we take *p* to be the inverse demand function of the good produced by region *i*. A natural way to characterise openness is to posit that more open regions are those that are more responsive to the expenditures of the consumers located in other regions. We assume that the utility of the representative consumer in each region depends on her consumption of the good produced in her own regions, and on the consumption of the goods produced in the other regions, viz. $u_j = \prod_{i=1}^{N} c_{ij}^{b_i}$, where *j* is an index for the regions, c_{ij} denotes the consumption by the *j*th region of the good produced in the *i*th region, and b_i measures the preference of a consumer located in the *j*th region for the good produced in the *i*th region.[1] Since the optimisers of a utility function are not influenced by its monotone transformations, we can safely assume that $\sum_{i=1}^{N} b_i = 1$. The budget constraint of our representative consumer located in region *j* is $\sum_{i=1}^{N} p_{ij} \cdot c_{ij} = y_j$, where p_{ij} is the price paid by the consumer located in region *j* for the good produced in region *i*, and y_j is the income of the representative consumer in region *j*.

This form of the utility function yields demand functions $c_{ij} = b_i \frac{y_j}{p_{ij}}$. The aggregate demand for the good produced by the *i*th region is then

$$\sum_{j=1}^{N} c_{ij} = b_i \cdot \frac{y_i}{p_{ii}} + b_i \cdot \sum_{j \neq i} \frac{y_j}{p_{ij}} \tag{4}$$

where the first term of this expression is the domestic demand, while the second term is the non-domestic demand. We make the assumption that the consumers face higher costs of buying the goods produced in the other regions, and that these additional costs

[1] We assume that the preference parameters *b* depend on *i* but not on the region in which the good is sold, *j*. We are then assuming that preferences depend on some "objective" characteristics of the goods that do not change according to the region in which the consumers are located. As we shall see below, we differentiate the costs of buying the goods coming from the various regions through some sort of transportation cost that affects the price paid by the consumers. It is not difficult to see that we could have alternatively assumed that differences in the demand for the non-domestic goods was a matter of preferences (by setting that *b* depended also on *j*) without changing the implications of the model that we are interested in.

are proportional to the producer price, viz. $p_{ij} = p_i \cdot t_{ij}$, where p_i is the price obtained by the producer and $t_{ij} > 1$ for $i \neq j$, and $t_{ij} = 1$ for $i = j$.

In equilibrium aggregate demand is equal to the total quantity produced of good i, $\sum_{j=1}^{N} c_{ij} \equiv Q_i$. Domestic income y_i is equal to the total sales of the region, viz. $y_i = p_i \cdot Q_i$, or $y_i/p_i = Q_i$. One can then re-write (4) as $Q_i = b_i \cdot Q_i + b_i \cdot \frac{\sum_{j \neq i} t_{ij}^{-1} y_j}{p_i}$. Solve for Q_i, and obtain the inverse demand function $p_i = b_i \cdot \frac{\sum_{j \neq i} t_{ij}^{-1} y_j}{Q_i}$, where b_i has been re-parameterised in an obvious way.

In the expression for the inverse demand function, the variations across regions of the term inside the summation sign are largely related to the "transportation" cost t_{ij}. This is because in a system of interconnected regions, like the European ones, the y_j's are in large part the same for many of the i's. Even if the individual regions dealt with different commercial partners (and hence the set of incomes y_j for different regions i is different), the term inside the summation sign that is unrelated to i can be thought of as being a stochastic term. That is, one can think of t_{ij} as being made up of two components, one that is specific to the exporting region i, and the other to the importing region j. The component related to i is systematic, as it applies repeatedly to all the terms of the summation, while the one related to j is idiosyncratic to each of these terms. In short, we can assume that $t^{-1}_{ij} = v_i \cdot u_j$, where v_i is the component that is systematic to all the terms in the summation, and u_j is the importing country stochastic shock. The inverse demand function becomes

$$p_i = \frac{b_i \cdot n_i}{Q_i} \Psi_i \qquad (5)$$

where $\Psi_i \equiv \sum_{j \neq i} m_j y_j$, and we assume that this term is stochastic across regions. Thus, apart from preferences b_i and the quantity Q_i, the inverse demand for the good produced in region i is affected by factors n_i that are specific to the exporting region, and that account for the extent to which such regions can reduce the costs of buying their goods by the consumers located in other markets. Put differently, n_i denotes how responsive is a certain region to the expenditures of other regions, and we take it to be a measure of the ability of the region to promote its goods outside its territory.[2]

To complete the model, we assume that the expression for $G(\cdot)$ in (1) takes the form $G = \left(\frac{Q}{A}\right)^g \cdot X^h$. Replace this expression and the one for p in the aggregate supply function of the region (3). Solve for Q. Eventually one obtains

$$\frac{Q}{L} = \left(\frac{a \cdot H}{r}\right)^{\frac{a}{1-g}} \cdot \left(\frac{L}{A}\right)^{\frac{g}{1-g}} \cdot X^{\frac{h}{1-g}} \cdot L^{-\frac{a}{1-g}} \tag{6}$$

where $H \equiv b \cdot n \cdot \Psi$ summarises the effects of preferences and openness arising from the demand side, and it includes the stochastic factor Ψ that depends on the characteristics of the set of regions in which the given region is exporting its products. Expression (6) is the labour productivity equation that we estimate in this paper. As we shall see, our econometric specification will employ variables that account for n and b, and it will assume that Ψ is part of the error term.

[2] While in our introduction we argued that openness may have supply-side and demand-side effects, here we model it as a demand-side effect. In part, this is because we believe that demand-side effects (i.e. exports, or the international demand induced by the location of multinational enterprises) are in the end more important than supply-side factors, like international spillovers or the productivity benefits arising from specialisation. Note also that we are characterising openness in terms of a factor n that accounts for the extent to which the region responds to the expenditures of other regions. Such a factor is typically correlated with some general manifestations of regional openness – e.g. mobility of people, airplane passengers, etc., as also argued in the introduction.

3. THE DETERMINANTS OF LABOUR PRODUCTIVITY IN EUROPEAN REGIONS

3.1 Sample, data, and variables

Our estimation employs an unbalanced sample of NUTS2 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 quite a few 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 more disaggregated NUTS3 level since most of these potential controls and instruments are reported only for the NUTS2 regions. We tried to construct fairly homogeneous regions. For example, the NUTS2 regions in Germany (e.g. Stuttgart or Tübingen) compare more naturally to the NUTS3 regions in other countries (e.g. they are similar to the Italian Provinces). We then employed NUTS1 regions for Germany (e.g. Baden-Wüttenbger, Bayern), whose overall magnitude and administrative role within the country are similar to the NUTS2 regions of Italy, Spain or France.[3] Similarly, the NUTS2 regions for Belgium, the Netherlands, and Portugal are quite small, and we employed NUTS1 regions for these countries as well (e.g. Region of Bruxelles). We also employed NUTS1 regions for the UK. The correspondence between NUTS1 UK regions (e.g. Eastern Regions, rather than the NUTS2 East Anglia or Essex) and the NUTS2 regions for France, Italy or Spain is harder to justify. However, we were forced to use NUTS1 regions for the UK because there are too many missing values for the NUTS2 UK regions. Our final sample however contains only few UK regions because of several missing values.[4]

Our final sample is composed of 622 observations. It includes regions from 9 European countries – Belgium, France, Germany, Italy, the Netherlands, Portugal, Spain, Sweden, and the UK. The bulk of our sample, however, is composed of the Italian, French,

[3] For instance, the regional governments of the NUTS1 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 that come at the level right below the national government. Moreover, some regions are classified as NUTS1 regions (e.g. Berlin, Brandeburg, or Sicily) and they are subdivided into a number of smaller NUTS3 regions without having a NUTS2 classification.

[4] Caniels (1999) also rearranged her definition of regions from the NUTS classification in a way similar to ours to obtain regions that were more comparable with one another.

Spanish and German regions. 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 do 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 potential endogeneity of the contemporaneous variables. Also, REGIO only provides the 1997 value of *NIGHT*. As we shall see in section 3.3 we use this variable to account for the touristic attractiveness of the region. Since this does not change dramatically over time, the problem may not be that severe. For *HOUSPOP* we employed the average for the number of households during 1992-1994 because no other years were available.

3.2 The econometric specification

To estimate (6) we employ the following log-linear specification

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

$$m_1 \cdot log(KPAT_{it}) + m_2 \cdot log\left(\frac{L_{it}}{A_{it}}\right) + m_3 \cdot log(PASS_{it}) +$$

$$m_4 \cdot log(AGR_i) + m_5 \cdot log(ARABLE_i) + m_6 \cdot log(MTW_{it}) +$$

$$m_7 \cdot log(SUI_i) + m_8 \cdot log(L_{it}) + e_{it} \qquad (7)$$

where *log(KPAT), log(L/A),* and *log(PASS)* are our variables of interest; *log(L)* is the log of total employment in the region; e_{it} is an error term; and the *m*'s are parameters to be estimated.[5] We will treat *log(KPAT), log(L/A), log(PASS),* and *log(L)* as endogenous. All the other variables are exogenous controls.

As extensively discussed so far, and following Ciccone and Hall (1996), and Ciccone (2000), the employment density *(L/A)* accounts for the effects of the agglomeration economies.[6] 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.[7] 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.[8] The variable *KPAT* is also 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.

Note also that in terms of our labour productivity equation (6), *KPAT* can be either part of the vector *X*, or it can affect the preferences b_i for the good produced by the *i*th region. In the latter case, the effect of patents is via the demand *p(·)*, and we assume

---

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

[6] Because we do not use NUTS3 regions, our problem in using the NUTS2 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.

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

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

that other things being equal technologically more intensive goods have a higher unit value. We are again unable to distinguish whether patents account for demand or supply-side factors. However, we like to think that they also capture demand side factors like the preference for technologically more intensive goods, rather than looking at them – as typically done by the literature – only as another factor of production, spillovers, etc..

The variable *PASS* is our measure of international openness. To justify its use, recall that we are looking for variables that denote the propensity of our regions to capture a larger share of the demand of the other regions, i.e. our term n in section 2. We argue that this variable is correlated with the openness of the regions in the sense that we are trying to denote here. In the first place, the mobility of people is correlated with exports and imports. These activities typically imply movement of people inside and outside the region. The presence of multinational enterprises also commands airplane mobility of passengers. More generally, airplane mobility of people and international trade flows are correlated. Second, 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. Third, we realised from our data that almost all our regions had airports and airplane passengers. Hence, this variable is not biased towards the major hubs. Relatedly, REGIO provided the information on the number of airplane passengers excluding the passengers in transit. Again, this limits the bias towards the international hubs as the passengers are only recorded in their airport of initial departure and final destination.[9]

We also considered alternative measures available from REGIO, and particularly 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. But unlike airports, there are maritime passengers only in the relatively few regions that border with the sea. This variable is then biased towards the regions with larger

[9] Clearly, some bias is likely to persist, since people may reach close-by regions from a major hub with means other than the airplane.

harbours. In other words, much more than for airports this variable is likely to overstate the role of the hubs. In addition, only few people travel by sea nowadays. The maritime freight of goods had much of the same problems. As far as the freight by airplane are concerned, only special kinds of goods are moved by plane. This variable could then capture factors like the composition of demand or supply in the region.

Finally, we compared *PASS* with a measure of the sectoral specialisation of the regions, on the ground that specialisation is correlated with the openness to international trade. We computed the Herfindhal index of the shares of the regional value added in six sectors.[10] We found that the correlation coefficient between *PASS* over the population of the region and this index was 0.69.[11] Moreover, we run our regression (7) using both *PASS* and this index, and by treating both variables as endogenous. Whenever 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 these two variables were measuring very similar effects. In our analyses, we used *PASS* rather than the sectoral specialisation index for two reasons. First, compared to passengers, this is at any rate only an indirect measure of openness. Second, *PASS* was available for a larger number of observations, and these were better spread across countries which provided a greater variability of the other variables in our equations.[12]

All the other regressors in (7) are controls. We included *AGR* and *ARABLE* to control for the composition of the regional output, and particularly for the importance of agricultural activities. The motorways variable *MTW* proxies for infrastructures. This

[10] These are Agriculture, Forestry & Fishery; Fuel & Power Products; Manufactured Products; Building & Construction; Market Services; Bank Services; Non-market Services. Finer sectoral distinction were not possible because of missing data in REGIO.

[11] 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.

[12] Since the openness of a region is unlikely to undergo dramatic changes from year to year, we took *PASS* to be the average number of passengers over three years. (See Table 1.) This reduces the potential effect of yearly shocks to the number of passengers which are unrelated to changes in the openness of the regions. We also used a five year average for *PASS*, as well as simply annual passengers with no major changes in the results.

avoids that the employment density of the region would in fact proxy for such infrastructures. The rationale for including the number of suicides, *SUI*, is that this variable 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.[13]

We employ dummies to control for time- and country-specific effect. This also follows Ciccone (2000) who employs dummies for larger regions to account for the unobserved price of capital in the individual NUTS3 regions used in his estimation. The justification is that *r* is similar across NUTS3 regions belonging to the same larger regional areas, and hence dummies for such larger regions are likely to capture their effects (possibly up to a stochastic error). Unlike Ciccone, here we employ NUTS2 or even NUTS1 regions. However, *r* is likely to vary largely because of factors that are common within the same country (e.g. monetary policy), and to the extent that they change over time they are captured by our time dummies. In addition, we assume that further differences in *r* across regions of the same country, are correlated with the other controls that we use in the regression, and they are partly contained in the error term. The error term in (7) also accounts for other unobserved characteristics coming either from the demand or the supply side, including the stochastic shock Ψ which captures the characteristics of the other regions with which our region trades.

3.3 Addressing the endogeneity issue

In estimating (7) we face a classical endogeneity problem. The natural way of thinking of our problem 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 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

[13] One of the best known essay on the matter is the one written in the XIX century by the famous French sociologist Emile Durkheim, entitled “The Suicide”. Durkheim 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 over population, and obtained a correlation coefficient of 0.40.

density may be higher because regions with higher incomes attract people. Finally, while the international openess of the 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. We then need 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 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, we follow Ciccone (2000) and use the total land area of the region, A_i, as an instrument for *(L/A)*. Ciccone's argument is that the total area of a region is uncorrelated with changes in regional productivity. This is because the areas of the European regions were defined several years ago, and in most cases even more than one century ago. For example, the Italian and the German regions still reflect the borders of the States that composed their territories before their unification in 1861 and in 1870. Similarly, the French regions as well as the regions of the other European countries reflect historical conditions which originated very long ago. Yet, as noted by Ciccone (2000), and as also confirmed by our data, the area of the regions is negatively correlated with their employment density. Thus, while A_i is not affected by today's regional productivities, it is nonetheless correlated with the variable we are interested in, *(L/A)*.

As far as *KPAT* is concerned, it is known that research activities tend to be located in areas that are more enjoyable to live. We then used *NIGHT*, *MOTO*, and *SEA* as instruments that proxy for the "quality of life" of the region. As indicated in Table 1, *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. The rationale is that 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 3, which lists the top 20 regions in the REGIO data base ranked by *NIGHT*.

A simple inspection of Table 3 reveals that these are all highly touristic regions.[14] Since touristic regions are more pleasant to live, other things being equal they attract research. Similarly, the number of motorcycles, *MOTO*, is correlated with the pleasantness of the regional weather.[15] The sea also increases the attractiveness of a location. Hence, other things being equal, *SEA* is correlated with research activities.

TABLE 3 ABOUT HERE

The same variables can be employed as instruments for *PASS*. The variable *NIGHT* plays again a key role. 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. Thus, an entrepreneur in a region with a large airport because of tourism may find it easier to move abroad than if she was in another region. 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 quite indirectly associated with tourism (e.g. the food industry), the effective share of relevant touristic activities is even 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.

The variables *MOTO* and *SEA* are good instruments for *PASS* as well. As the former accounts for the pleasantness of the weather, it is likely to imply greater attraction of people from outside the region. Historically, the sea has been a major factor in

[14] Table 3 is constructed using all the regions in the REGIO data base for which *NIGHT* was available. As a result, it also includes regions that are not in our final sample.

[15] *MOTO* over population is positively correlated with *NIGHT*, which suggests that the former is also associated with places where life is more pleasant.

enhancing communication and openness. Since, as noted earlier, the borders of our regions have been established long ago, *SEA* is exogenous. In fact, 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. This may cast some doubt on the exclusion of *SEA* among the regressors of (7). But one of the reasons for employing *MTW* in (7) was to control for such transportation activities and the associated infrastructures. In other words, to the extent that the sea implies the presence of harbours, and therefore of related communications and transportation of goods, these factors also imply the development of motorways to move both goods and passengers to other inland destinations. Thus, once we include *MTW* directly in (7), the problems potentially associated with the exclusion of *SEA* may be less severe.[16]

Finally, in (7) *log(L)* is also endogenous. We then use *HOUSPOP* and the average population in working age in the region (*POP25-65*) as additional instruments in our estimation. The rationale for these instruments is that they 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.[17] Similarly, the population in working age

[16] As the attentive reader has certainly noted, we are identifying our effects of interest through exclusion restrictions. As discussed in the text, we have tried to justify why *NIGHT* and *SEA* may not be crucial exclusions in (7). One might argue that the pleasantness of the weather could also have direct effects on productivity. In this case, *MOTO* or even *SEA* would not be proper exclusion restrictions. However, the direct effect of the weather on productivity is likely to be negligible. Moreover, there could be arguments going both ways – viz., good weather may prompt greater productivity by workers, and people more generally, because of the vigor associated with places in which there is more light, sun, and an enjoyable life; by contrast, it may make people more willing to undertake leasurely activities.

[17] 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. Compare for instance what is commonly held about the propensity of the young people in the Northern European regions to leave their parents' house vis-à-vis the young Italians or Spanish. Even within countries there can be substantial differences across regions, like between Northern and Southern Italian regions.

reflects whether a region is composed of a relatively young or old population, which would also affect the labour supply, and hence *L*.[18]

3.4 Empirical results

Table 4 reports our empirical results obtained 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. Particularly, the elasticities with respect to the stock of patents, *KPAT*, and the employment density, *(L/A)*, are both around 10-11%, and they are statistically significant. The elasticity with respect to *PASS* is around 3.5% (slightly smaller in the OLS estimation), and it is statistically significant.

TABLE 4 ABOUT HERE

This confirms the importance of innovation and the technological capabilities of the regions in raising labour productivity. Our estimated effect of the agglomeration economies is higher than the one estimated by Ciccone (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 neighboring regions. This suggests that our estimation has internalised these spillovers. Finally, we found an independent effect of openness. This occurs in addition to the potential increase in demand due to other desirable characteristics of the goods produced by the region, like its innovation and technological content. In short, this suggests that being internationally linked matters. Our controls also have the expected sign. Particularly, note that *SUI* is positive and statistically significant, which is consistent with our conjecture that it proxies for the educational level of the regions.

To give some sense of the extent to which our variables affect the observed differences in labour productivity across European regions, we computed the effects of a one standard deviation of *KPAT*, *(L/A)*, and *PASS* from their sample average. By using the GMM estimates in Table 4, the elasticities are respectively 11.1%, 10.5%, and 3.6%.

[18] If there is interregional 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.

From Table 2, the sample average for the labour productivity *(Q/L)* is 36.5 thousand euros. The sample averages of *KPAT*, *(L/A)*, and *PASS* are respectively 1.1 thousands, 0.130, and 2,108 thousands. Then, if *KPAT* increased by one standard deviation (2.5 in Table 2), the estimated percentage increase in labour productivity for a region with *KPAT* equal to its sample average would be 25.2%, viz. (11.1*2.5/1.1). If a region had labour productivity equal to the sample average, this would produce an increase in *(Q/L)* from 36.5 to 45.7.

Similar calculations indicate that if *(L/A)* increased by one standard deviation (0.345 in Table 2), the percentage increase in labour productivity for a region with *(L/A)* equal to the sample average would be 27.9%, viz. (10.5*0.345/0.130). At the sample average, labour productivity would increase from 36.5 to 46.7. Finally, if *PASS* increased by one standard deviation (3,945 in Table 2), the percentage increase in labour productivity for a region with *PASS* equal to the sample average would be 6.7% (3.6*3,945/2,108). The effect on labour productivity at the sample average would be an increase from 36.5 to 38.9. While smaller than the other two, the effect of openness is nonetheless sizable, and it contributes to enhance the observed differences in regional labour productivity in Europe.

4. LABOUR PRODUCTIVITY AND EMPLOYMENT IN EUROPEAN REGIONS

4.1 The neoclassical model

■ ***Structure of the model.*** In this section we estimate the effects of agglomeration economies, technology and openness on labour income and employment. To do so, we have to deal with two issues. First, we need to estimate a labour supply function. The covariates in our labour productivity equation (6) can be interpreted as factors affecting the demand for labour. As we shall see below, we assume that labour supply depends on time and country dummies, the variables *HOUSPOP* and *POP25-65*, and labour income proxied by *(Q/L)*. By jointly estimating labour demand and supply, not only we will estimate the effects of *KPAT*, *(L/A)*, and *PASS* on labour income, as we did in the previous section, but also on employment.

The second issue is the measurement of the labour supply. Here we have to clarify what is our prior belief about what is measured by the series on the regional labour force (or active population) available from REGIO. Does it measure the number of people who are willing to work at the prevailing wage but who cannot find a job at that wage, or those who express a desire to work but who would actively seek a job only if the wage was higher? In the latter case, we need to estimate a typical neoclassical model, where labour markets clear. The proper measure of labour supply is then the number of people employed, and not the active population. In the former case, the proper measure of labour supply is the active population. But we then have to add an employment adjustment equation explaining the level of employment and its potential gap from the active population. In this subsection we estimate the neoclassical model. In the next subsection we estimate the keynesian model.

The neoclassical model is based on two equations. We write them in their general form as

$$(Q/L) = f[KPAT; (L/A); PASS; Z_1; L; e_1] \qquad (8)$$

$$L = g[Z_2, (Q/L), e_2] \qquad (9)$$

where *(Q/L)*, *KPAT*, *(L/A)*, and *PASS* are our usual variables, and *L* is the number of people employed. Equation (8) is the labour demand equation, which is the same as (7). Thus, Z_1 is a vector of controls composed of a constant term, country dummies, time dummies, *AGR*, *ARABLE*, *MTW*, and *SUI*, while e_1 is the error term. Equation (9) is the labour supply equation. As noted earlier, we assume that labour income *(Q/L)* affects labour supply. The covariates in Z_2 are: constant, time dummies, country dummies, *HOUSPOP* and *POP25-65*. We estimated (8) and (9) jointly by GMM. We employ the same instruments used in the previous section – viz., Z_1 and Z_2, along with *log(MOTO)*, *log(NIGHT)*, and *SEA* The empirical specification of (8) and (9) is log-log.

■ ***Empirical results.*** The empirical results are in Table 5. The estimated elasticities of *(Q/L)* with respect to *KPAT* and *(L/A)* are similar to those estimated in the previous section. The elasticity with respect to *PASS* is slightly higher (4.2%). The estimated elasticity of labour supply with respect to income is 44.2%.

TABLE 5 ABOUT HERE

Like in the previous section, we can perform some simple simulations about the changes in labour income and employment produced by interregional differences in *KPAT*, *(L/A)*, and *PASS*. Unlike the previous section, however, we have to take into account that (8)-(9) is a system of equations in its structural form. This means for example that there is both a direct effect of *KPAT*, *(L/A)*, and *PASS* on *(Q/L)* in equation (8), as well as an indirect effect coming through *L*. Since *(Q/L)* also affects the labour supply (9), one has to take this effect into account as well. In other words, in the previous section, we looked for the effects of *KPAT*, *(L/A)*, and *PASS* on *(Q/L)* holding the labour supply constant, and therefore holding *L* constant in the labour productivity equation. Now, to take the indirect effects into account, we first have to transform (8)-(9) in its reduced form. Since (8)-(9) are linear in logs, the standard procedure is to solve the system for *log(Q/L)* and *log(L).* The elasticities with respect to *KPAT*, *(L/A)* and *PASS* computed from the reduced form will be a function of the estimated structural parameters. These elasticities are reported in Table 6.

TABLE 6 ABOUT HERE

As in section 3.4, we can look for the effects of one standard deviation changes in *KPAT*, *(L/A)*, and *PASS* from their sample averages. By using the estimated elasticities in Table 6, labour income changes respectively by 20.9%, 23.6%, and 6.7%.[19] In turn, this implies changes in labour productivity from its sample average of 36.5 thousands euros to 44.1, 45.1, and 38.9. As far as changes in employment are concerned, one standard deviation increases in *KPAT*, *(L/A)* and *PASS* from their sample averages

[19] These are the result of the following calculations 9.2*2.5/1.1; 8.9*0.345/0.130; and 3.6*3,945/2,108.

imply changes in employment of, respectively, 9.3%, 10.4%, and 3.0%.[20] Since the sample average of *L* is 1.321 million people (see Table 2), the one standard deviation increase in *KPAT* from its sample average implies an estimated increase in the employment of the "average" region from 1.321 millions of about 123 thousand people. The one standard deviation increase in *(L/A)* implies an estimated increase in employment of 137 thousands, and the one standard deviation increase in *PASS* increases employment by 40 thousand units.

4.2 The keynesian model

■ ***Structure of the model.*** The keynesian model makes the alternative assumption that the active population denotes the number of people who are willing to work at the prevailing wage. Hence, it includes people who are "involuntarily" unemployed. Our two models are really two polar cases – the case in which all unemployment recorded by the official statistics is involuntary, and the case in which it is entirely voluntary. By estimating both models, we can evaluate to what extent our estimates are sensitive to different assumptions about the structures of the regional labour markets in Europe.

The keynesian model implies that we now have three endogenous variables – labour productivity *(Q/L)*; labour supply, which is now the active population *S*; and the people employed, *L*. The labour productivity equation is the same equation estimated in the previous sections. The labour supply equation is the same as the labour supply equation estimated in section 4.1, with *S* instead of *L* as the dependent variable. We can write the two equations in their generic form as

$$(Q/L) = f[KPAT; (L/A); PASS, Z_1; L; e_1) \qquad (10)$$

$$S = g[Z_2, (Q/L), e_2] \qquad (11)$$

where the arguments of these functions have been discussed in the previous section. By assuming that the active population of the region can be different from the people actually employed we are assuming that there is some form of disequilibrum in the

[20] Viz. 4.1*2.5/1.1; 3.9*0.345/0.130; and 1.6*3,945/2,108.

labour market. In other words, the labour income can be persistently different from the equilibrium wage because the labour markets do not clear. The level of employment would then be affected by the extent to which the labour markets flexibly adjust the wage to the level that would cover the existing unemployment. This also means that we have to specify a third equation for the employment rate *L/S*, which we write in its generic form as

$$L/S = h(Z_3, X, S, e_3) \qquad (12)$$

Apart from the error term e_3, we assume that the employment rate is a function of time dummies and country dummies, which are the variables included in the vector Z_3. Time dummies account for cyclical effects, while the country dummies account for national factors, like national legislations and the like, that may affect the institutional conditions of the labour markets. We include a few other variables in (12). First, we introduce a set of variables *X*. In *X* we include the labour income proxied by *(Q/L)*, as well as *KPAT*, *(L/A)*, *PASS*, and the other controls in the labour demand equation (10). The goal is to estimate whether increases in income or in the variables that affect labour demand affect the employment rate. One hypothesis could be that in regions with lower incomes, or with little openness or research, there are greater social interventions which keep the labour market from flexibly reducing the wages. Another way to see this is that regions with higher incomes or with higher levels of *KPAT*, *(L/A)*, or *PASS* imply a greater number of activities wherein people move flexibly across jobs according to individual conditions as well as to changes in wages or salaries. This is typically the case of jobs associated with greater skills like research, or similarly of jobs associated with the presence of multinational enterprises, more globalised firms and the like, as implied by the openness of the region. If employment and wages adjust, the gap between people employed and the active population will reduce. In (12) we include *S* amongst the covariates, which accounts for the fact that differences in labour supply may affect the employment rate as well.

Thus, the model that we estimate is composed of equations (10), (11), and (12), which we assume to take a log-linear form. The endogenous variables are *(Q/L)*, *S* and *L/S*. In the estimation we also treat *KPAT*, *(L/A)*, *PASS*, and *L* as econometrically endogenous. We employ the usual set of instruments, which include all the *Z* variables in the system, along with *log(MOTO)*, *log(NIGHT)* and the *SEA* dummy. We show the results obtained by GMM.

■ ***Empirical results.*** The results of the estimation are in Table 7. The first column of Table 7 estimates a version of our system with no *X* variables among the regressors in (12). This assumes that the degree of flexibility of the labour market is captured entirely by the time and country dummies. The second column estimates a version that includes *(Q/L)*, but no other variables in *X*. This assumes that regions with higher incomes are associated with more flexible labour markets, which reduces the distance between people employed and those seeking jobs. In the third column, *X* includes *(Q/L)*, as well as all the other regressors in Z_1 – that is, *KPAT*, *(L/A)*, *PASS*, *AGR*, *ARABLE*, *MTW*, and *SUI*. This is to estimate the different potential effects of labour income, as proxied by *(Q/L)*, and of the determinants of labour income, on the employment rate. For example, this enables us to assess whether variables like *KPAT*, *(L/A)*, or *PASS* imply more flexible labour markets (by affecting the employment rate *L/S*), apart from their direct effects on labour demand, and hence on the wage, as implied by equation (10). The parameters of interest in Table 7 are the elasticities of *(Q/L)* with respect to *KPAT*, *(L/A)* and *PASS*; the elasticity of the labour supply *S* with respect to *(Q/L)*; and the impacts of *(Q/L)*, *KPAT*, *(L/A)*, *PASS* and the other determinants of labour income on the employment rate *L/S*.

TABLE 7 ABOUT HERE

The results of the estimation can be summarised as follows. First, the elasticities of *(Q/L)* with respect to *KPAT*, *(L/A)* and *PASS* are roughly similar to those of the neoclassical model. Second, the elasticity of the labour supply with respect to *(Q/L)* is far smaller than in the neoclassical model. For example, in the fully specified model (third column of Table 7), the elasticity of the labour supply with respect to income is quite small, 5.7% (as compared to 44.2% in the neoclassical model). The active population then appears to be far less elastic with respect to changes in income than the people actually employed. This suggests that interregional mobility in Europe is not very pronounced.

As far as the determinants of the employment rate are concerned, the second column of Table 7 shows that regions with higher income exhibit a higher employment rate. In the third column of Table 7, the statistical significance of the covariates in the employment rate equation is not impressive. If we focus on the point estimates, *KPAT* has the most sizable positive impact on the employment rate, while the impact of *(Q/L)* on the employment rate becomes negative. The impact of *(L/A)* is also positive, although smaller than *KPAT*. The impact of *PASS* is positive, but quite small. Interestingly enough, *SUI* has a positive impact on *(L/S)*, with a non-negligible statistical significance. If we keep with our interpretation that *SUI* measures the level of education of the region, this suggests that more educated regions entail lower unemployment rate.   The negative impact of *(Q/L)* in the third column of Table 7 suggests that two effects may be at work here. On the one hand, *KPAT* and the other factors that affect the demand for labour, imply higher employment rate. On the other hand, as they induce higher incomes, they have an indirect negative effect on employment because of the potential increase in labour costs.

By using the estimated parameters in the third column of Table 7, we can perform our usual simulation. As we did in the previous section, we first have to transform system (10)-(12) in its reduced form. Given that all three equations are linear in logs, we solve the system for *log(Q/L)*, *log(S)*, and *log(L/S)* as a function of all the other variables. We then obtain the elasticities of each of these variables with respect to *KPAT*, *(L/A)*, and *PASS*. In the reduced form, these will be functions of the estimated structural parameters. These elasticities are reported in Table 8.

TABLE 8 ABOUT HERE

As the Table shows, in the keynesian model the elasticities of labour income are higher than the corresponding elasticities in the neoclassical model, while the elasticities of employment are smaller. Particularly, note that the estimates of the keynesian model imply that only patents (i.e. technology) increases employment. Agglomeration economies and openness only increase labour income without increasing employment. The same applies if one looks at the employment rate rather than the employment level in Table 8.[21]

This result is intriguing. It says that technology is the crucial factor in regional development, as it increases both income and employment. By contrast, agglomeration economies and openness benefit primarily those who are already employed, by raising their incomes, without providing any particular advantage to the unemployed or the new entrants in the job market. It is also interesting to compare the results of the keynesian model with those of the neoclassical model. Broadly speaking, the estimates of the keynesian model denote higher elasticities of income and lower elasticities of employment. One way to think about these results is that the beliefs about what is measured by the active population is important. Relatedly, our results are consistent with the view that when labour markets are flexible, as postulated by the neoclassical model, economic opportunities, like technology, agglomeration economies, or openness, translate into higher incomes, and because of the flexibility of the labour markets, they increase employment. By contrast, if the labour markets are not flexible, as assumed by the keynesian model, all the advantages of the economic opportunities associated with agglomeration economies and openness translate into higher incomes, with no benefits on employment. As noted, only the effect of technology is strong enough to induce increases in both income and in employment.

To conclude, we also evaluated the effects of one standard deviation increases in *KPAT*, *(L/A)* and *PASS* with respect to their sample averages on labour income *(Q/L)* and emplyment *L*. Given the estimated elasticities in Table 8, labour income would increase respectively by 24.3%, 31.6%, and 9.9%, while the percentage increases in employment

[21] In Table 8, we show the elasticities of both the employment level *L*, and the employment rate *(L/S)*. To compute the latter we simply solved our system (10)-(12) for its reduced form. The former can be computed by the same reduced form by taking the elasticity of *L* with respect to *S* to be one plus the estimated elasticity of the employment rate *(L/S)* with respect to S. This is obtained by moving *logS* to the right hand side of (12).

will be respectively 6.6%, -0.3%, and –2.1%.[22] For a region with *(Q/L)* equal to the sample average, this would imply an increase from 36.5 to respectively 45.4, 48.0, and 40.1. For a region with employment equal to the sample average of 1.321 million people, we obtain an estimated increase in the number of people employed of 87 thousands in the case of *KPAT*, about zero for *(L/A)*, and even slightly negative (about 3 thousand people) for *PASS*. This compares with increases in labour productivity from the sample average up to 44.1, 45.1, and 38.9, and with increases in employment by 123, 137, and 40 thousand people in the neoclassical model, as computed in the previous section. (See Figure 1.)

FIGURE 1 ABOUT HERE

5. CONCLUSIONS

Regional advantages have been the subject of a good deal of discussion in recent years. The topic is especially hot in Europe, where regional inequalities are more pronounced than in the US States, and the European Commission is keen about the developement of the less advanced regions of the Union. The literature has focussed on two main explanations of regional advantages – agglomeration economies and technological capabilities. In this respect, there are two main novelties in our paper. First, we estimate the effects of regional openness on labour productivity and employment. Second, we take into account agglomeration economies and technology as well. We empirically assess the relative importance of all three factors in explaining European differences in regional labour productivity and employment.

We found that regional openness, as measured by the annual number of airplane passengers, contributes to explain differences in regional labour productivity, together with technology and agglomeration economies. This suggests that policies aimed at encouraging regional development should not focus only on factors that are "internal" to the localities, like local infrastructures, 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 depend on the ability of the regions to export their goods, or on

[22] As usual, these are obtained by multiplying the elasticities in Table 8 by the ratio between the standard

other factors, like spillovers due to mobile and internationalised human capital, the presence of multinational corporations, or else. However, the experience of some of the fast growing regions of the world today (e.g. the Asian Tigers, or the newcoming countries like Ireland and Israel) indicate that these factors are correlated with one another. In short, there may be underlying differences in the extent to which some regions are more open than others, and we found that this matters.

We also found that the effects of openness, and partly of the agglomeration economies, on employment are less pronounced. This may be part of the more general problem faced by Europe in creating new jobs. However, we find that the technological capabilities of the regions do raise employment. We therefore conclude that technology is a powerful factor in promoting both higher income and jobs. By contrast, openness and agglomeration economies seem to benefit mostly those who are already employed, without helping as much regional employment growth. We also found that there are differences between what we labelled the "neoclassical" *vs* the "keynesian" model, with the former model showing a more pronounced effect of our variables on employment as well. We interpret this finding as suggesting that the effects of openness and agglomeration economies on employment may be more marked when the labour markets are more flexible.

References

Arora, A., Gambardella, A. and S. Torrisi (2001) "Human Capital, International Linkages, and Growth: The Software Industry in India and Ireland", paper presented at the *American Economic Association Meeting*, New Orleans, January 4-8.

Arora, A. and J. Jasundi (1999) "Quality Certification and the Economics of Contract Software Development: A Study of the Indian Software Industry", NBER Working Paper n. 7260.

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.

---

deviation and the sample average of *KPAT, (L/A),* and *PASS* in Table 2.

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)

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.

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), *Economic Geography and Trade,* MIT Press, Cambridge, MA.

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

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.

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.

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”, mimeo, University of Berkley, Berkeley, 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].
S_{it}	Active population 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 (excluding passengers in transit), 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
S_{it}	1449.2	1416.4	50.8	8030.2
L_{it}/S_{it}	0.911	0.088	0.681	1.580
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: 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 4: Determinants of labour productivity - OLS, 2SLS, and GMM estimation

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

	OLS	2SLS	GMM
Const.	4.275 (0.087)	4.538 (0.096)	4.560 (0.094)
log($KPAT_{it}$)	0.087 (0.005)	0.114 (0.007)	0.111 (0.007)
log(L_{it}/A_{it})	0.120 (0.020)	0.103 (0.019)	0.105 (0.019)
log($PASS_{it}$)	0.012 (0.002)	0.034 (0.005)	0.036 (0.005)
log(AGR_i)	0.084 (0.019)	0.086 (0.017)	0.089 (0.016)
log($ARABLE_i$)	- 0.017 (0.006)	- 0.007 (0.007)	-0.007 (0.007)
log(MTW_{it})	0.013 (0.002)	0.017 (0.003)	0.019 (0.003)
log(SUI_i)	0.027 (0.014)	0.032 (0.018)	0.043 (0.017)
log(L_{it})	-0.239 (0.027)	-0.340 (0.031)	-0.350 (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 5: The "neoclassical" model - GMM estimation

Dependent variables $log(Q_{it}/L_{it})$ and $log(L_{it})$

	GMM
Labour Productivity Equation $log(Q_{it}/L_{it})$	
Const.	4.668 (0.093)
$log(KPAT_{it})$	0.107 (0.007)
$log(L_{it}/A_{it})$	0.104 (0.019)
$log(PASS_{it})$	0.042 (0.004)
$log(AGR_{i})$	0.083 (0.016)
$log(ARABLE_{i})$	0.002 (0.006)
$log(MTW_{it})$	0.020 (0.003)
$log(SUI_{i})$	0.048 (0.016)
$log(L_{it})$	-0.376 (0.029)
Labour Supply Equation $log(L_{it})$	
Const.	-1.529 (0.168)
$log(HOUSPOP_{i})$	0.192 (0.054)
$log(POP25\text{-}65_{i})$	0.990 (0.004)
$log(Q_{it}/L_{it})$	0.442 (0.044)
N. of obs.	622

Heteroscedastic consistent standard errors in parenthesis. All equations include time and country dummies. 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: Neoclassical model - "Full" elasticities of labour income and employment with respect to technology (*KPAT*), agglomeration economies *(L/A)*, and openness (*PASS*) (*)

	KPAT	*(L/A)*	PASS
Labour income *(Q/L)*	9.2%	8.9%	3.6%
Employment (*L*)	4.1%	3.9%	1.6%

(*) Elasticities computed from the reduced form of (8)-(9)

Table 7: The "keynesian" model - GMM estimation

Dependent variables *log(Q_{it}/L_{it}), log(S_{it}), log(L_{it}/S_{it})*

Labour Productivity Equation log(Q_{it}/L_{it})			
Const.	4.709 (0.096)	4.603 (0.088)	4.801 (0.101)
log($KPAT_{it}$)	0.125 (0.007)	0.112 (0.006)	0.121 (0.007)
log(L_{it}/A_{it})	0.098 (0.021)	0.116 (0.017)	0.119 (0.020)
log($PASS_{it}$)	0.046 (0.005)	0.034 (0.004)	0.048 (0.005)
log(AGR_i)	0.093 (0.017)	0.087 (0.014)	0.104 (0.017)
log($ARABLE_i$)	0.000 (0.006)	-0.003 (0.005)	0.001 (0.007)
log(MTW_{it})	0.023 (0.003)	0.015 (0.003)	0.021 (0.003)
log(SUI_i)	0.047 (0.019)	0.036 (0.015)	0.034 (0.018)
log(L_{it})	-0.414 (0.030)	-0.348 (0.027)	-0.415 (0.031)
Labour Supply Equation log(S_{it})			
Const.	-0.801 (0.132)	-0.467 (0.137)	-0.402 (0.135)
log($HOUSPOP_i$)	0.016 (0.048)	0.021 (0.048)	0.031 (0.047)
log($POP25\text{-}65_i$)	1.010 (0.003)	1.015 (0.003)	1.017 (0.003)
log(Q_{it}/L_{it})	0.177 (0.032)	0.077 (0.033)	0.057 (0.033)

Cont.

Table 7: Cont.

Employment Rate Equation log(L_{it}/S_{it})			
Const.	0.054 (0.028)	-1.276 (0.109)	1.796 (2.317)
log($KPAT_{it}$)	--	--	0.076 (0.050)
log(L_{it}/A_{it})	--	--	0.045 (0.051)
log($PASS_{it}$)	--	--	0.008 (0.020)
log(AGR_i)	--	--	-0.006 (0.050)
log($ARABLE_i$)	--	--	0.012 (0.006)
log(MTW_{it})	--	--	-0.003 (0.011)
log(SUI_i)	--	--	0.036 (0.019)
log(Q_{it}/L_{it})	--	0.388 (0.033)	-0.340 (0.520)
log(S_{it})	-0.017 (0.003)	-0.026 (0.003)	-0.184 (0.166)
N. of obs.	622	622	622

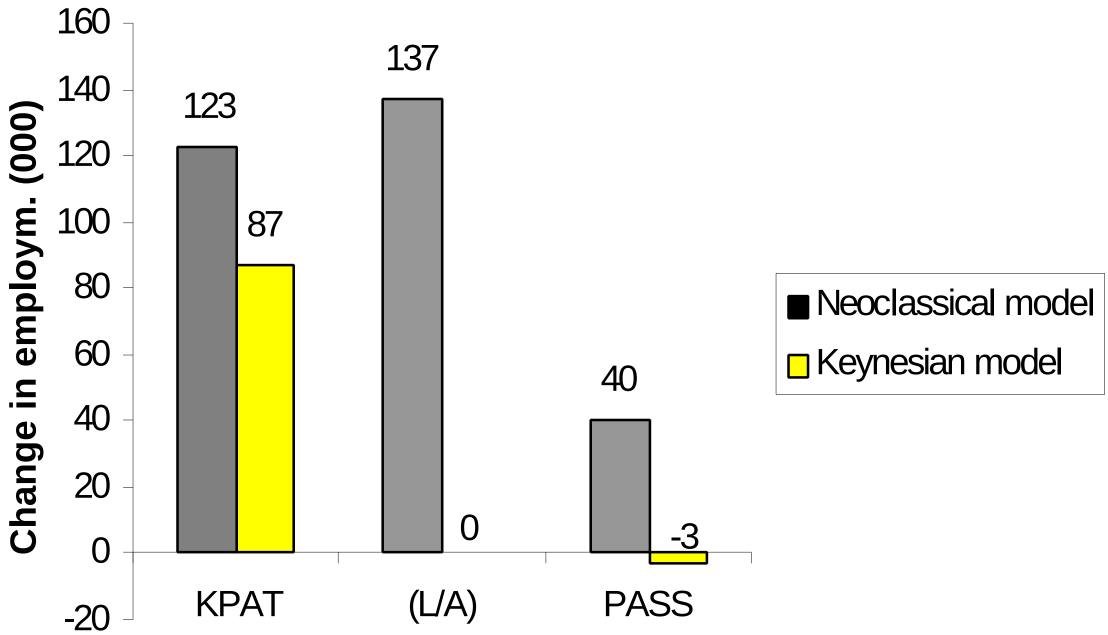
Heteroscedastic consistent standard errors in parenthesis. All equations include time and country dummies. Instruments: constant, time dummies, country dummies, *log(AGR)*, *log(ARABLE)*, *log(NIGHT)*, *log(MTW)*, *SEA*, *log(A)*, *log(SUI)*, *log(MOTO)*, *log(HOUSPOP)*, *log(POPM25-65)*.

Table 8: Keynesian model - "Full" elasticities of labour income, employment, and employment rate with respect to technology (*KPAT*), agglomeration economies *(L/A)*, and openness (*PASS*) (*)

	KPAT	*(L/A)*	PASS
Labour income *(Q/L)*	10.7%	11.9%	5.3%
Employment (*L*)	2.9%	-0.1%	-1.1%
Employment rate (*L/S*)	3.6%	0.6%	-0.7%

(*) Elasticities computed from the reduced form of (10)-(12)

Figure 1: Increases in employment caused by one standard deviation increases in technology (*KPAT*), agglomeration economies *(L/A)*, and openness (*PASS*) - Comparison between the neoclassical and the keynesian models



(*) All variables are evaluated at their sample average. Particularly, these are increase in employment from its sample average of 1.321 millions.