XXV Conferenza Italiana di Scienze Regionali

Novara, 06-08 ottobre 2004

Conoscenza, innovazione e sviluppo territoriale

Labour Market Dynamics and Industrial Clusters: an Ecological Based Approach

Massimiliano R. Riggi* M. A. Maggioni †

June 2004

JEL Codes: J24, J61, R23

Keywords: Industrial Clustering, Labour Market Dynamics, Ecological Approach.

Abstract

The paper investigates the role of labour flows in determining the growth of industrial clusters (defined as a combination of industries and spatial areas). Within an ecological approach the role of inter and intra industry - as well as of inter and intra cluster - interactions is theoretically modelled and empirically tested. This approach is enriched with other economic perspectives in order to take into account the effect of workers' migration flows and its skill

*CERTeT and IEP, Università Bocconi, piazza Sraffa, 13, 20136 Milano, Italy. E-mail: massimiliano.riggi@unibocconi.it.

†DISEIS and Faculty of Political Science, Università Cattolica del Sacro Cuore, via Necchi, 5, 20126 Milano, Italy. E-mail: mario.maggioni@unicatt.it.

composition. The empirical evidence shows that industry specific behaviours and territorial peculiarities are crucial in determining the employment dynamics. In particular, empirical evidence from 50 US states and 80 3-digits NAICS industries are able to show that employment flows play different role on local wages according to industrial characteristics (within services the main distinction arise between services to production and to persons); within manufacturing industries between labour intensive and capital intensive activities). A final section drawing some policy suggestion from the above analysis concludes the paper.

1 Introduction

The production structure and dynamics within the US show both geographic and sector peculiarities. This stems employment flows and wages strongly depending on these specificities. How do wages and employment behave over time? At some industrial detail, there seems to be a relation in the way employment and wage covariate. This suggests that there exists a sector specific relationship between the two variables and that industrial and geographical ties matter to this extent. This feature spurs us to think of State – Industry couplets as different but non independent systems of production, later denoted as 'clusters'.

The issue of regional labour mobility has become crucial in the recent economic debate. In particular, it is often referred to as a main driver for laggard regions to catch up. Emphasis has been given alternatively to two factors: either the role of flexibility of wage adjustment in matching labour demand and supply at macro level (Blanchard et al., 1992), or the importance of wage determinants at a micro level (like in Bentivogli and Pagano (1999)), nested in personal and firm features such as the relevant industry and skills (like in Venturini (2001)).

The paper is devoted to investigate the structure and evolution of occupational flows with a particular reference to its skill composition.

In this way it is possible to underline the interplay of agglomeration economies and diseconomies in the growth process of a cluster (as defined in section 2), to distinguish among different kinds of interactions existing between different industries, within the same area (inter industry relations), and between different areas within the same industry (intra industry relations).

A complementary toolkit to this respect is constituted by the literature on skilled migration flows and the impact on receiving and sending regions. Why should migration flows involve more qualified workers at most? The relevant literature provides the following explanations:

shocks. In case of a regional shock and consequent drop outs from firms bankrupting, the opportunity cost of being unemployed is higher for skilled workers (Wood and Ridao-Cano, 2002) and they are willing to move away to compensate that cost;

transferability. Higher skills may be transferred more easily from one cluster to another than low skills, the range of readaptation to a different productive system being wider (Borjas, 1994);

mobility costs. Assuming wages being set proportionally to skills, highly skilled labour force is more likely to overcome mobility costs, if the latter are equal across different skills levels;

complementarities. The sending region being less productive, there may be a lack of factors (such as managerial talent) complementary to skills such that skilled workers receive an incentive to move (Commander et al., 2003);

risk taking behaviour. More skilled workers are generally more educated and this reduces their risk when moving (Shields and Shields, 1989).

The theoretical benchmark for the current paper is given by the Ecological Approach, so that interesting insights to model the evolution of clustering behaviour over time is taken into account. The inherent limitations of such an approach, namely the lack of micro-foundation of firms and workers decisions, is overcome through the explicit analysis of the role of wages as leading forces for workers to move and for firms to settle in some specific location. These two parts of the story are then matched considering the skill composition of regional migration flows. More specifically, the quality of movers is considered and their skills accounted as a part of the regional capital endowment. In order to understand what this implies in terms of regional disparities, two effects are identified following migration flows:

Convergent regional disparities (Supply Effect) The increase in labour supply in the West pushes wages downward narrowing the interregional wage rate gap (Bencivenga and Smith, 1997).

Cumulative regional disparities (Demand Effect) The effect of selective migration on cumulative growth can be understood as altering the relative labour demand in favour of the expanding region (Ghatak et al., 1996, Dolado et al., 1994) in detriment of the laggard one.

Combining agglomeration effects from labour mobility and the skill composition of labour flows, the Ecological Approach is given economic meaningfulness by the high skilled workers flows: in ecological models, the dynamics is usually driven by interactions between species and the availability of limited resources. The second point is tackled to endogeneise the adjustment process to the system carrying capacity, which can be affected by skilled workers through their contribution to the regional total capital endowment.

The Ecological Approach offers an interesting insight to model the evolution of clustering behaviour over time; the starting point being the basic ecological model where a cluster grows at a constant rate up to a carrying capacity (K_c). The economic foundation implemented here consists of giving an economic meaning to K_c, which will depend upon: the finite quantity of geographical benefits (which is related to the limited availability of local resources such as: labour, capital, land, and infrastructures); the agglomeration benefits (which depend on the strategic interactions between firms: competition, congestion and lobbying of incumbents). K_c is therefore determined by the relationship between the amount of resources, and in particular of skilled labour forces.

Firms decide to settle in a cluster on the basis of the expected profitability of being located there. This profitability depends on net locational benefits - obtained as the difference between gross locational benefits (opportunity to exploit synergies and infrastructure networks for firms, higher wages and employability opportunities for workers) and costs (moving costs and opportunity costs from leaving a location such as costs associated to the former location relative to the newer). Net locational benefits stem a probability to move to an alternative location where revenues are expected to be higher. A first distinction can be made between substitution effects (related to the adjustment mechanism of wages on employment), and cooperative effects (when a cumulative mechanism rises in response to labour flows), with a positive effect on wages. In figure [1], we show whether or not in-migration flows can boost or not cumulative effects on labour demand. More generally, according to the elasticity of labour supply and labour demand functions, two different outcomes can arise. Figure [1] (left side) illustrates the case in which the increase in labour demand is not able to compensate the migration inflows, because skills brought by movers are not enough to boost growth and by that way to shift significantly upright labour demand (substitution effects prevailing). The final equilibrium exhibits an increase in employment (to E_1), at lower wages (from w_0 to w_1). On the right side, instead, in-migration is considered to be able to boost growth within the host cluster and the final equilibrium is at higher employment and wages levels (complementary effects prevailing).

The way these substitution and complementary effects are tested, will be illustrated in section 4; in particular, two effects are defined:

- Intra-industry dependences, to proxy sector interactions; a positive (negative) coefficient on this variable shows cooperative (competitive) effects operating within the cluster, with positive (negative implications) on the employment dynamics.

- Inter-industry dependences, to proxy regional interactions; a positive (negative) coefficient on this variable shows cooperative (competitive) effects working within the cluster, with positive (negative implications) on the employment dynamics.

Under an empirical point of view, the objective of the paper will be to test the determinants of cluster dynamics taking into account the effects of the skill biased composition of the workforce (the migration of skilled workers) and compare the results with those stemming from the inference from a standard ecological-population approach, to investigate the role that sector inter and intra dependences play in regard to clusters' dynamics.

In conclusion, the effect of selective migration can be understood as altering the relative labour demand in favour of the expanding region. For a

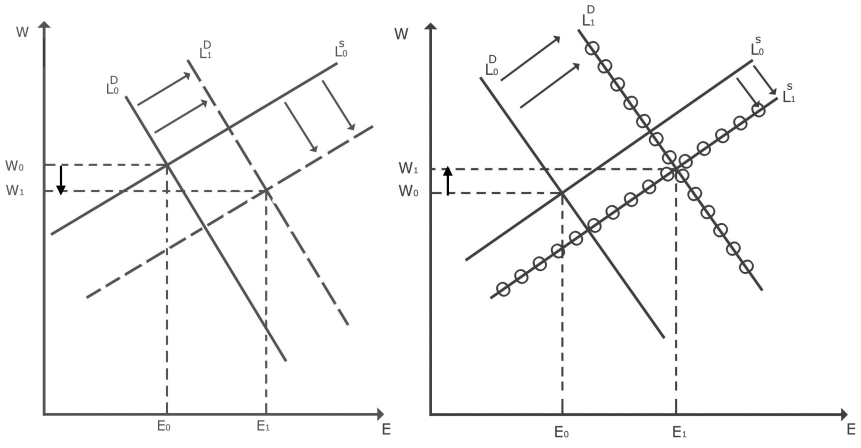

Figure 1: Labour migration: complementary and substitution effects.

reasonable combination of mobility costs and returns to skills, the predictions that only highly skilled workers move is consistent with the assumption by Krugman (1991) that workers in the agricultural sector are immobile. In graphical terms, the outward shift in the highly skilled labour supply in the *West* tends to imply a cumulative effect on labour demand; this effect is likely to overcome the traditional mechanism adjusting labour markets (Blanchard et al., 1992).

Consistently with these conclusions, Ghatak et al. (1996) discuss the long term impact of migration, arguing that it is positive through all forms of capital deepening. This turns to the role of human capital in growth; relevant contributions in this sense come from the endogenous growth models. In particular, Lucas (1988) is considered as the inspiring framework, in which human capital affects capital endowments, providing a mechanics suitable for studying economic development. Dolado et al. (1994) use a Solow (1956) growth model, augmented by migration, and assess the contribution of human capital – brought through in-movers' skills – to economic growth. They found empirical evidence at country level showing a positive impact, bigger, the greater is the human capital borne in movers. An important role in the cumulative process, is played by the technology in the most favoured region (the West), where the acquisition of high skilled workers can be easier due to skill specialisation and localised externalities. The introduction of a new technology may either switch around or reinforce the existing development pattern: region's attractiveness in fact depends on both wage levels and the productivity of the existing skills applied to the new technology (Desmet, 2000). On the one hand this gives rise to the two alternative development scenarios just sketched, on the other hand this argument can suggest a way to counteract regional disparities by setting up proper skills and stimulating

the localisation of new technology processes in lagging regions. Nevertheless these policies must be tightly integrated, to avoid pure subsidies to education that, within the migration model here proposed, can be further harmful to laggards (Suedekum, 2002).

Under this perspective, in section 7 it will be shown that both labour demand and supply policies are desirable for lagging regions; the first ones provide high skill workforce, the second one are aimed at reinventing a new technological trajectory (Dosi, 1982) so setting up a vocationally sustainable development path in the less favoured regions. It is reasonable to think that such policies being conceived to produce significant effects in the long run, adjustment interventions to sustain income have to be set up; nevertheless, the temptation to act up to bias the laggard's production opportunities has to be avoided. At the same time, infrastructures play a crucial role as an enabler for growth processes to be triggered. These interventions as a whole must be conceived as a compounded toolkit for regional policy, according to an integrating framework encompassing several aspects at the same time.

2 Defining and Gauging Clusters

The present work is built upon the definition of *clusters*. This label typically encompasses several meanings and it is not easy to set a cut point and draw a throughout definition.

In a very broad way, clusters refer to geographic agglomerations of firms in an industry or related industries. A short review of the definitions of clusters (as hereby meant) is now addressed. The original contribution can be attributed to Marshall (1921), who remarked that the importance for an industry to locate in a certain point in space depends on the presence of external economies that firms can benefit from by staying close to each others in terms of input provision or skilled labour force. Later, Porter (1990) observes that successful firms usually cluster in particular cities or States within a Nation, such that they take advantage from that location through a virtual circle between firms' growth, innovative success and industrial clustering. Porter (1998) defines an industrial cluster as a set of industries related through buyer-supplier (and vice-versa) relationship extending the definition to common technologies, the existence of joint distribution channels and common labour pools. Similarly, Swann et al. (1998) define the benefits of locating in a cluster as those related to the availability of skilled labour and intermediate goods suppliers, and also to the easy transmission and discussion of new ideas. Quite general, these definitions (especially Porters') are centred on firms business relations that can be transferred from the national level to the firm level focussing on industrial specialisation and competition. The idea of competitive advantage is introduced, based on a dynamic concept of competition that firms acquire though quality and innovation. The bulk

of innovation is the so called Porter's diamond; by exploiting factors conditions, demand conditions, inter-dependences and strategy, firm's growth is achieved.

In territorial empirical analyses the definition of the basic unit is not as straightforward as it is in other economic fields. Regions are irregular in terms of population, economic size, geographical dimension and their choice may lead to different policy implications. This argument is strengthened whether the focus of the analysis is centred on disparities, because spatial inequalities are sensitive to the definition of regions (Brülhart and Traeger, 2002). This problem is referred to as the 'Modifiable Areal Unit Problem' (MAUP), and concerns the arbitrariness of the geographical partition used; it implies that results from statistical data analysis for geographical dimensions can be different as long as geographical boundaries are varied. The MAUP can introduce a bias in any statistical measure, which will depend on the unit chosen whenever (and this is the case in territorial analyses) the unit is modifiable. The problem is twofold. On the one hand, the best aggregation scale has to be chosen (scale problem), on the other hand, units have to be correctly assigned to the right area (aggregation problem).

County Business Patterns provides the type of data described in section 8 for US counties and some thousand sub-sectors. Data are thereby collected at an administrative level, but at a such detailed one that it does not hinder us to set up functional regions to account for economic ties and flows cutting through administrative boundaries, but a strong limitation derives from the confidentiality issue discussed in section 8. The aggregation level for counties is not independent from the industry detail. This last choice is quite problematic: on the one hand industries should be as detailed as possible to allow peculiarities to rise, but too much detail implies that substantial economic ties are hidden, whereas too much aggregated sectors loose the meaningfulness of industry specificities. Hence, it is worth noting that the aggregation and scale problem here are twofold, involving both sectors and counties. The basic level of industrial data availability being a very much disaggregated digit-6 NAICS classification, data is managed at digit-3 NAICS level, which allows: a good level of detail (81 sectors at 3-digits NAICS level), self containment of the economic meaningfulness of the sector and the possibility to best bridge industries from SIC to NAICS [1].

The geographical dimension has been more complex to tackle and to simplify the analysis, following Swann et al. (1998) and Maggioni (2002), 'States' within the US are considered as the basic units; in this way, complexities (and confidentiality issues) that would arise from subdividing States at county level (i.e. counties), are overcome. Looking for economic flows

[1]The bridge to NAICS implies a new classification of sectors which is the more relevant the more detailed NAICS are. To avoid sensible breaks due to this switch, data from the SIC more detailed classification are aggregated to NAICS 3 level and hence the sector-switch bias becomes negligible.

between and within counties to aggregate them has not been considered worthwhile, because:

- too many clusters to deal with would arise;
- the industrial ties would be loose by the sector detail.

Besides, considering States as one of the dimensions to cut for basic units (Clusters), also allows us to encompass institutional State specificities.
On the basis of the above considerations, 'Clusters' (State/industry entities) are defined as basic units of analysis, combining one digit-3 NAICS sector and one State. Eventually, the work is implemented on 4233 clusters across the time span $1988 - 2001$.
The definition of industry-State couples is consistent with the purpose of the current analysis, oriented to investigate sector specific patterns of occupational flows. Another important choice which has driven the choice of the territorial units, is given by the confidentiality issue; at county level, wages are not reported whenever the establishment can be identified. This would bias the analysis whenever the sector structure implies concentration and firms can be identified.

Clusters, as defined here, include industrial districts of SMEs, concentrations of high technology firms related through the development and use of common technologies, and production systems that contain large hub firms and their local suppliers and spinoffs. A relatively broad definition of regional clusters is adopted, such that several aspects of the same phenomenon can be encompassed. This is not to say that differences among different types of agglomerations are ignored, but rather that each type provides with an insight which may reveal useful for the purposes hereby pursued. The most important common features relate the characteristics which shape basic units (e.g. firms) on the basis of inter-industry and inter-industry linkages, industrial *filière* peculiarities, competitive or synergetic use of raw material or basic services, systematic behaviours within the business cycles, pooled labour skills requirement, agreements to better fit the market. Besides, the context in which this phenomena occur is intertwined with the phenomena itself and affects the way relationships take place.
After the definition of the conceptual framework, a further step has to be undertaken to decide which degree of resolution is more suitable for the definition of cluster adopted to be economically meaningful.

3 Measurement Issues

In this section, two issues are addressed. The first one concerns the relation between employment dynamics and wages, whose intuitions rely on the skill selective migration model illustrated in section 4.3, which illustrates in detail the choices to be undertaken for this relation to be consistently and efficiently

estimated.
The second issue concerns the relational behaviour among clusters within the ecological framework. Maggioni and Riggi (2002) show that different clusters dynamics patterns are possible with a simulation approach, but here we address the possible interaction effects at intra and inter industry level. To this regard, the Panel Data technique is applied (see section 4).

3.1 Specification

The framework tackled concerns linear models. This can seem deceptive with regard to the ecological theoretical background; in particular, it could be raised the question whether or not such a modelling approach can give us the tools to grasp non linearities in the relation suspected in the evolution of the 'population' of employees within and between clusters. Such non linearities can be captured through a non linear specification with several methodologies (logit, probit, count data models, non parametric estimates), but the cost from the correct specification must be balanced with other costs arising from the increased complexity[2]. Here we address the question of non linearities within the linear model; one important point is that the use of the term *linear* can be misleading because a linear model is able to tackle non linearities in the explanatory variables, preserving the property of consistency. In practice, it is quite common to find some conditioning variables with squares and cross products values when non linearity is suspected in the relation with the dependent variable. What is important, is that the regression function is linear in its parameters. In section 4, we show the results of a linear panel data and a pooled regression. In order to comply with this issue, the final specification will also take into account for squared values of some conditioning variables.

3.2 Linda Indexes and Cluster Interdependences

In this section, some 'arenas' are defined, to identify the most important States in terms of establishment share by sector by using a modified version of Linda's index as defined by Maggioni (2002). Linda (1976) set up an original method to identify firms holding a considerable market share such that they are told to constitute an oligopolistic arena for a given sector. Maggioni (2002) applied this method to territorial analysis to gauge regional interactions within sectors; on this basis, we can define interactions between clusters. For each sector, it is possible to grasp whether or not there exist regions playing a crucial role in the structure of some given sectors and to which extend sectors are concentrated or dispersed.

The original formulation of the Linda index is:

[2]For example non linear models can harshly deal with individual fixed effects, which can play here an important role.

$$L_i = \frac{1}{n^* (n^* - 1)} \sum_{i=1}^{n^*-1} \frac{n^* - i}{i} \frac{CRt_i}{1 - CRt_i} \ . \tag{1}$$

Where L_i is the index for the $i-th$ unit, n^* is the unit under analysis and CRt_i represents the cumulative share of the characteristic examined (employment, establishment), ordered from the unit holding the greater share on. The most important feature of the series hinges in its non monotonicity; the larger n^*, the smaller L_i, but for growing n^*, CRt_i are progressively smaller, contributing to decrease L_i.

Following Linda (1976), the competitive segment of the market (firms holding a marginal share) is excluded. Though this implies a loss of information, this is consistent with the purposes here pursued to identify some sort of oligopolistic arena within an industry. This is carried out by excluding States holding marginal units of the characteristic observed (less than 0,8% within the main industry). As a consequence of this computational precaution, the oligopolistic arena, whenever it exists, arises in a more sharp way from the Linda's system.
Establishments and employment figures are averaged by sector, across time and expressed in sectoral shares, and then ranked in decreasing order of importance (those shares being defined CRi, where i denotes the generic sector). The single shares are then cumulated to get the share of the market held by respectively the first n^* clusters, where n^* assumes value two for the two biggest clusters, three for the three biggest ones and so on (those cumulative shares are indicated by $CRti$). Linda's index ranges between $\frac{1}{n^*}$ (lower limit) and ∞ (upper limit). Every index of the series is an arithmetic mean of the $(n^* - 1)$ oligopolistic equilibria ratios ($\frac{CRti}{1-CRti}/n^*$), previously divided by n^*, which represents the share of the market that firms within n^* engage compared to the residual market share.

When Linda's indices are ranked upon decreasing CRi and their graphical representation (the Linda's structural curve) shows at the beginning a decreasing monotonic path; the first discontinuity corresponds to the individuation of the competitive arena within the upper part of the series.

As a general finding, apart from some traditional sectors and natural resources, the number of States within each sector and each oligopolistic arena, witnesses that the most important sectoral shares of economic activities are not heavily concentrated (results omitted). In particular, sectors defining inter and intra industry dependences (that reveal to be significant in the inferential analysis carried out on 80 sectors on the 'ecological model', in section 4.2) show pretty dispersed sector shares.

Macrosectors in the County Business Patterns

NAICS	Classification
11	Agriculture, Forestry, Fishing and Hunting
21	Mining
22	Utilities
23	Construction
31	Manufacturing
420	Wholesale Trade
44	Retail Trade
48	Transportation and Warehousing
51	Information
52	Finance and Insurance
53	Real Estate and Rental and Leasing
54	Professional, Scientific, and Technical Services
55	Management of Companies and Enterprises
56	Administrative and Support and Waste Mgt and Remediation Services
61	Educational Services
62	Health Care and Social Assistance
71	Arts, Entertainment, and Recreation
72	Accommodation and Food Services
81	Other Services (except Public Administration)

Table 1: Industrial classification for macrosectors in the CBP.

4 Clusters Dynamics: Reduced Forms

We hereby address the original question about the determinants of clusters dynamics in a twofold way: on the one hand the focus is on the original ideas of purely ecological models (see subsection 4.2) and on the other hand, some crucial variables are included, such as wages, human capital and immigration rates (see section 4.3), which relate to the economic foundation of the ecological model discussed in the theoretical framework.

4.1 Variables Investigated

The CBP dataset has both longitudinal and temporal dimensions. This allows us to think of a panel data model to deal with heterogeneity across US clusters over time. Estimations are implemented by macro-sector to account for wider inter-sectoral differences; relations hereby analysed are suspected to show sector peculiarities that can be blurry on the whole industries. For each of the 17 macrosectors, employment and wage variables are expressed in deviation from the relevant industry mean, in accordance with Frisch and Waugh (1933); this allows us to tackle the heterogeneity embodied in the cross-section dimension analysis in a way that prevents to harm estimates efficiency within the OLS hypotheses.

The theoretical framework has been devoted to the investigation of the determinants' of clusters' growth. Particular attention has been paid to the skill composition of workers and its role in affecting the evolution of clusters

over time.

In this section the attention will be switched to the empirical analysis of the determinants of clusters' growth, bridging the theoretical issues and data availability/technicality issues. In particular, we analyse the occupational mass variation of clusters and its possible explanations. Bearing in mind the ecological perspective (clusters grow up to a carrying capacity, which can be overcome) combined with the economic proposed (qualified workforce as a way to foster growth and development), the growth of employment has to be justified in economic terms; besides, the potential carrying capacity of each cluster has to be controlled for in terms of sector specific features, affected by the different industry related skills. The skill composition of the employment workforce being not available from the CBP dataset, the hypothesis that sector specific skills are at work is made; in this way the skill component is controlled by distinguishing among different sectors. This hypothesis is very strong if considering that several skills are at work within each sector, but the argument still holds on average: each sector has a particular skill content which it is intensive in. The drawbacks are besides mitigated by the industrial classification which is adopted. The 'North Atlantic Industrial Classification System' groups activities instead of products, which makes industries functional to describe the skill content necessary within a process. Consistently with the theoretical model, the dependent variable to be investigated is defined as the temporal variation of the employment mass in each cluster (or its growth rate).

The skill effect component can be included exploiting the large variety in sector classification available in CBP and is embedded in the variable wage, computed as a difference between cluster's wage and average industry wage. Wage also expresses the quality of labour within a sector, and can also be related to the total capital component increase, stemming from skill selective migration flows, as described in the model for regional migration; this is consistent with Marshall's idea of industry specific skills (Marshall, 1890) as one of the reasons for firms to cluster. Eventually, though migration flows data are not available, an important information can be inferred from the variation of wage over time for each cluster. An increase in wages (at constant prices) over time within the same unit shows that capital has increased. Recall the definition of capital provided in the theoretical framework; it is both physical and human capital to be part of the cluster's capital endowment. When observing an increase of wages over time, it can be reconducted to either labour skill content or an increase in capital intensity or a shift in the production. The importance of skills within sectors goes back to the distinction between codified and tacit knowledge (Nonaka and Umemoto, 1996) and the importance of knowledge within production. When the effects of human capital formation and movement are also taken into account, as this paper is intended to do, clustering behaviours can become crucial; in fact, conditions can be provided under which human capital is better created

and more efficiently used within industrial clusters containing similar firms Almazan et al. (2003).

Recalling the nature of relevant regressors and the dependent variable, the issue of endogeneity arises. In particular, observe the following equation:

$$\frac{dN_t^{is}}{dt} = \alpha + \beta w_t^{is} + \gamma N_t^{is} + \delta N_t^{is} N_t^{ik} + \lambda N_t^{is} N_t^{js} + u_{it} \tag{2}$$

where N_t^{is} is the employment mass of State s, industry i at time t, w is the relevant wage in cluster is and u_{it} is an unobservable random variable with $E\left[u_{it} \mid w\right] = 0$. Equivalently equation [2] can be expressed as:

$$E\left[\frac{dN_t^{is}}{dt} \mid w, N_t^{is}\right] = \alpha + \beta w_t^{is} + \gamma N_t^{is} + \delta N_t^{is} N_t^{ik} + \lambda N_t^{is} N_t^{js} \tag{3}$$

This specification encounters a big trouble due to the simultaneity of w and N_t^{is}, which are far from being independent. Nevertheless, alternative data is used to instrument N_t^{is} with its underlying generation process:

$$N_t^{is} \quad = \quad \gamma_0 + \gamma_1 est_t^{is} + \delta_2 empsize_t^{is} + v_t^{is} \tag{4}$$

where $estsize_t^{is}$ represents establishments size and $empsize_t^{is}$ the employment size of firms. This equation opens a route to identification of eq. [2], through the additional information born by the establishment size and the employment size of firms, which are at the same time correlated to N_t^{is} and uncorrelated with w_t. In reality $cov\left(w_t, (estsize_t, empsize_t)\right)$ is not zero strictly speaking, but the correlation being considerably reduced, this is considered a way to operationalise the estimation process Greene (2003), page 381-2. Eventually, the reduced form that is estimated is obtained putting eq. [4] in eq. [2], where the endogeneity issue is solved:

$$\begin{aligned}\frac{dN_t^{is}}{dt} = \kappa + \beta w_t^{is} + \psi estsize_t^{is} + \tau (estsize_t^{is})^2 + \phi empsize_t^{is} \\ + \pi (empsize_t^{is})^2 + \delta N_t^{is} N_t^{ik} + \lambda N_t^{is} N_t^{js} + \zeta_{it}\end{aligned} \tag{5}$$

where $\kappa = \alpha + \gamma_0$, $\psi = \gamma\gamma_1$, $\phi = \gamma\gamma_2$ and ζ is a white noise error.

With respect to this equation, to operationalise the estimation, control variables are included. In particular, the focus of the analysis being the skill selective effect of wage on employment, other effects, influencing this relation, have to be controlled for. In this sense, the nationwide industrial rate of growth (g_t^{i-}), the mean educational attainment (HK^{-s}) and the rate of immigration on employment ($Imm - rate^{-s}$) control for other sources of wage variability are added. It is reasonable to assume that, with the model formulation in eq. [5], estimated by macro-sectors and with the other regressors just outlined, with the last attempt to infer variation in employment that can be explained with the skill composition of workers inflows.

4.2 Model I

The reduced form to be estimated is:

$$\frac{dN_t^{is}}{dt} = \alpha_0 + \alpha_1 w_t^* + \alpha_2 estsize_t^{is} + \alpha_3 empsize_t^{is} + \alpha_4 Intra + \alpha_5 Inter + \alpha_6 g^{i-} + \epsilon_{it} \quad (6)$$

Equation [6] implicitly considers the 'two clusters' case. In reality, the sample encompasses some 4000 clusters; such a wide variety would complexify the empirical analysis and retain too many degrees of freedom if all possible combinations are taken into account. Sector peculiarities are tackled by applying the Frisch-Waugh-Lovell theorem within macrosectors and implementing estimates for each of them (listed in table 1). This trick will allow us to both take into account the strong inter sector heterogeneity existing, and to gain many degrees of freedom. In order to account for clusters inter dependences, the following hypotheses are adopted to set up variable **Intra** and **Inter**: two kinds of clusters are considered: the one examined and all the others. The rest of the world is supposed to affect one's development path, if it belongs to the competitive arena defined through the Linda series (see section 3.2); in this way, only the most important clusters [3] are considered to be able to affect the 'rest of the world' development path. After this first selection, a distinction is made between clusters belonging to the same State where the one under analysis is, but operating in a different sector (inter-industry dependences) and clusters belonging to the same sector where the one in exam is, but operating in different States (intra-industry dependences). *Intra* and *Inter* variables, so defined, provide me with cooperation and competition effects of regional and sector interactions discussed in section 3. The main objective hereby pursued is to test whether positive or negative dependences work to explain the evolution of clusters.

4.3 Model II

The reduced form to be estimated is:

$$\frac{dN_t^{is}}{dt} = \beta_0 + \beta_1 w_t^{is} + \beta_2 estsize_t^{is} + \beta_3 empsize_t^{is} + \beta_4 Intra + \beta_5 Inter + \beta_6 g_t^{i-} + \beta_7 HK_t^{-s} + \beta_8 Imm - rate_t^{-s} + \zeta_{it} \quad (7)$$

Equation [7] implicitly considers the 'two clusters' case. In reality, the sample encompasses some 4000 clusters; such a wide variety would complexify the empirical analysis and retain too many degrees of freedom if all possible combinations are taken into account. Sector peculiarities are tackled by

[3] In terms of market share held.

applying the Frisch-Waugh-Lovell theorem within macrosectors and implementing estimates for each of them (listed in table 1). This trick will allow me to both take into account the strong inter sector heterogeneity existing, and to gain many degrees of freedom. The distinction is carried out between a purely ecological (see section 4.2) and a model that considers the migration of skilled labour (defined Skill biased model). A further advancement is decided on the basis of the previous results, and an analysis is implemented at NAICS-3 level aggregation, which implies analysis for some 80 sectors.

5 Interpretation

Table 2 report a short explanation of the labels adopted in the estimation results. For the full explanation of variables refer to section 4.1.

The main advantage of the NAICS classification in comparison to the SIC based one, is that activities are not classified by product, but by functions necessary to accomplish a given task; this is particularly useful in this context where the skills within a given activity are the hinge of the discussion. Estimations are implemented by macro-sector to account for wider inter-sectoral differences; relations hereby analysed are suspected to show sector peculiarities that can be blurry on the whole industries. For each of the 19 macrosectors, employment and wages are expressed in deviation from the relevant industry mean, in accordance with Frisch and Waugh (1933); this allows me to tackle the heterogeneity embodied in the cross-section analysis in a way that prevents to harm estimates efficiency within the OLS hypotheses.

5.1 Results and Comments: Model I

At macrosector level (see table [3]), composite results arise, showing in general that the level of aggregation is not the most suitable. On this basis, the ecological model is implemented on the 80 3-digits NAICS sectors. Table 4 summarises these results on the ecological setting. Interestingly, inter and intra industry dependences never show both positive or negative significant coefficients: competitive (synergic) intra industry dependences are generally associated to synergic (competitive) inter industry dependences. In many cases the coefficients of either intra or inter industry dependences (sometimes both) are positive or negative and significant at least at 10% of the significance interval, but the focus is here on the case in which both are significant. Two groups of sectors can be thus identified. Within the first one (Agriculture support activities, Heavy construction, Food manufacturing, Fabricated metal product manufacturing, Miscellaneous manufacturing, Publishing industries, Real estate, Nursing care facilities, Recreation industries, Food and drinking services) clusters grow together with the ones situated within

Variable	Explanation
dempw	Variation in employment over time
empsize	Occupational Class size (see table [10])
dwageD	log variation of wages
empwg	National Industrial Trend
hk	Educational Attainment
Imm-rate	Immigrants rate to employment
inter	Industrial Inter-dependences
intra	Industrial Intra-dependences
estD	Establishments, deviations from mean
estD2	Establishments squared, deviations from mean
year	Temporal control variable
$*$	10% sign. level (t-stats reported)
$**$	5% sign. level (t-stats reported)
$***$	1% sign. level (t-stats reported)

Table 2: Labels in estimations

the same State. Territorial contiguities boost growth whereas similar industries in different States are negatively related to clusters' growth. Apart from some special cases, competition is typically working among similar and close clusters. Within the second group (Motor vehicle dealers, Building material and garden dealers, Health and personal care stores, Gasoline stations, Clothing and accessories stores, Sporting and hobby stores, General merchandise stores, Credit and related Activities, Ambulance health care services, Hospitals, Personal and laundry services, Religious grantmaking and like organizations) there are sectors with a strong territorial vocation in serving local markets, while on the other hand, they can benefit from the positive cycles that other sectors are experiencing on the same territory (follow-the-wave sectors). That is why their growth is positively associated with the size of other industries in the same State and negatively with the size of the same industry in other States (spatial competition).

5.2 Results and Comments: Model II

At macrosector level, the NAICS classification applied to the model with highly skill migration (skill biased flows) shows the following results (table [7]): Utilities present a positive relation between variation in wages and variations in employment; this is mainly attributed to the specialised workforce employed within this sector. Retail Trade and Construction, for the low skill content embedded in workers, exhibit a negative relation between wages and employment variations. Manufacturing, Information, Arts and Entertainment show the expected positive relation. Management of Com-

NAICS classification

Dep. Var.: dempw	empsize	empwg	intra-est	inter-est	estD	estD2	year	Constant	Obs.
Agric.	-0.28	(33.05)***	(2.47)**	-1.16	(-2.17)**	-0.32	-0.79	-0.78	1939
Mining	(-1.67)*	(29.32)***	-0.51	-0.6	-1.15	-0.18	(2.15)**	(-2.16)**	1926
Utilities	-0.17	(-13.80)***	-1.53	-1.06	-1.07	-0.18	-0.74	-0.75	663
Constr.	-0.54	(60.85)***	-1.56	-0.41	(-4.30)***	(5.22)***	(2.40)**	(-2.41)**	1989
Manuf.	-1.16	(45.44)***	(2.73)***	-0.51	(-1.93)*	-0.48	-0.34	-0.34	13834
Retail Tr.	-0.91	(90.19)***	(7.34)***	-0.77	(-9.27)***	(17.89)***	-0.24	-0.18	7956
Transp.	(-3.42)***	(63.48)***	-0.19	-0.49	(-3.12)***	-0.48	-0.05	-0.06	5155
Inf.	(2.11)**	(28.27)***	-0.84	-0.22	(3.90)***	-1.32	-0.09	-0.07	2652
Finance	-0.96	(35.39)***	(2.98)***	-0.62	(-3.11)***	(4.89)***	(1.74)*	(1.72)*	2647
Real Est.	-0.14	(22.91)***	-0.55	-1.41	(-4.96)***	(5.02)***	-0.57	-0.58	1975
Prof. Serv.	-1.42	(89.59)***	-1.39	-0.23	(2.84)***	(-2.95)***	(14.36)***	(-14.38)***	663
Mgmt of Companies	(2.27)**	(43.89)***	-0.17	-0.58	-0.57	-0.41	-0.96	-0.96	663
Admin. Serv.	-0.86	(61.37)***	(1.67)*	-0.86	-1.1	(2.11)**	(1.84)*	(-1.85)*	1326
Educat. Serv.	-0.52	(28.68)***	(2.52)**	(-1.97)**	(-3.79)***	-1.05	-0.67	-0.63	663
Health Care	-0.91	(194.59)***	-1.44	-0.81	(3.52)***	(-3.30)***	(27.32)***	(-27.35)***	2652
Arts	(1.67)*	(15.20)***	(2.43)**	-1.5	(-4.38)***	(-2.36)**	(-3.97)***	(3.95)***	1989
Accom.	-0.05	(129.13)***	-0.28	-0.85	(2.02)**	(-2.06)**	(19.30)***	(19.33)***	1326
Other Serv.	-0.54	(138.69)***	(-3.92)***	-1.17	(6.34)***	(-8.04)***	(20.47)***	(-20.48)***	1989
Wholesale Tr.	-0.6	(74.98)***	-0.2	-0.59	(3.09)***	(-2.41)**	(8.14)***	(-8.15)***	663

t-stats reported
∗ significant at 10%
∗∗ significant at 5%
∗ ∗ ∗ significant at 1%

Table 3: Ecological model (model I).

Panel Data: 80 **sectors**

Model I (Ecological)	empsize	empwg	imm-rate	intra-est	inter-est	estD	estD2
Forestry and logging	-1.19	34.86***	-5.28***	-1.69*	-0.91	-1.25	-0.02
Fishing and hunting	-2.59***	40.67***	-6.96***	-0.59	-0.2	-1.92*	-0.5
Agr. and forestry supp act.	-1.41	21.31***	-6.11***	-2.62***	-1.66	-0.69	-1.14
Oil and gas extr.	-0.49	35.9***	-6.39***	-0.33	-0.43	-0.82	-0.17
Mining (exc. oil and gas)	-1.73*	20.41***	-11.11***	-1.65*	-1.22	-0.3	-0.96
Mining supp. act.	-1.53	35.35***	-3.75***	-0.14	-0.84	-1.71*	-0.75
Utilities	-0.02	15.13***	-12.21***	-1.26	-0.84	-0.63	-0.38
Building contracting	-1.71*	44.15***	-9.9***	-0.05	-0.73	-0.04	-3.8***
Heavy construction	-0.81	27.73***	-12.86***	3.35***	-2.27**	-3.25***	-0.79
Special trade contract.	-3.55***	69.94***	-4.61***	-1.12	-0.09	-3.04***	-2.99***
Food mfg	-1.77*	31.22***	-5.43***	2.1**	-1.67*	-3.46***	-1.92*
Beverage, tobacco prod. mfg	-1.12	48.75***	-4.88***	-0.48	-0.39	-1.58	-1.2
Textile mills	-0.25	16.31***	-3.43***	-0.55	-0.29	-0.85	-0.24
Textile prod. mills	-0.14	30.58***	-4.67***	-1.31	-0.96	-1.73*	-0.27
Apparel mfg	-3.63***	38.71***	-5.56***	-0.98	-0.07	-1.35	-0.14
Leather prod. mfg	-1.14	20.85***	-5.3***	-0.26	-0.65	-0.3	-0.97
Wood product mfg	-0.88	24.46***	-2.02**	-2.03**	-0.61	-1.07	-0.18
Paper mfg	-0.15	33.13***	-1.19	-1.79*	-0.89	-1.14	-0.08
Printing and supp. activities	-0.96	25.3***	-12.12***	-0.44	-0.04	-3.08***	-1.15
Petroleum and coal prod. mfg	-2.25**	21.56***	-5.9***	-0.05	-0.08	-0.55	-0.29
Chemical mfg	-1.98**	28.15***	-8.18***	-1.13	-1.18	-1.08	-0.12
Plastics prod. mfg	-0.37	27.26***	-4.61***	-1.1	-0.93	-1.67*	-0.54
Nonmet. mineral prod. mfg	-1.92*	60.49***	-6.04***	-0.81	-0.35	-0.99	-0.17
Primary metal mfg	-1.28	11.66***	-3.55***	-0.5	-0.48	-0.92	-1.06
Fabricated metal prod. mfg	-1.55	23.51***	-4.87***	2.14**	-2.06**	-2.38**	-1.1
Machinery mfg	-0.13	28.65***	-2.49**	-2.06**	-0.82	-1.56	-0.78
Computer and el. prod. mfg	-2.69***	16.96***	-6.72***	-0.92	-0.71	-0.16	-0.71
Electr. and appl. Mfg	-0.96	20.53***	-4.54***	-1.56	-0.97	-0.29	-1.13
Transp. equip. mfg	-0.38	13.04***	-7.78***	-0.31	-1.05	-0.31	-0.47
Furnit. and rel. prod. mfg	-1.14	40.09***	-7.08***	-2.31**	-1.27	-0.19	-0.86
Miscellaneous mfg	-2.24**	25.75***	-5.18***	2.89***	-1.89*	-0.53	-1.1
Wholesale trade	-0.83	70.79***	-5.6***	-1.33	-1.48	-2.06**	-0.22
Motor vehicle dealers	-0.01	75.9***	-0.41	-7.7***	3.81***	-4.56***	-12.2***
Furniture stores	-1.11	75.23***	-3.92***	-2.02**	-1.12	-1.66*	-3.59***
Electr. equip, appl. stores	-1.9*	77.06***	-0.18	-0.67	-0.94	-3.14***	-0.21
Bldg mat. and garden dealers	-0.47	80.46***	-1.39	-4.46***	3.1***	-6.06***	-10.55***
Food and beverage stores	-0.07	47.06***	-7.5***	-0.35	-0.53	-2.29**	-5.45***
Health and pers. care stores	-0.79	72.74***	-1.76*	-6.03***	3.98***	-7.21***	-9.36***
Gasoline stations	-6.1***	74.97***	-0.87	-4.21***	3.72***	-2.34**	-6.46***
Clothing and access. stores	-1.09	75.15***	-4.12***	-6.39***	4.28***	-4.66***	-7.17***
Sporting and hobby stores	-4.32***	66.47***	-5.51***	-6.09***	4.61***	-1.16	-1.93*
General merch. stores	-2.07**	73.85***	-0.38	-7.17***	4.43***	-0.72	-8.72***
Miscellaneous stores	-1.14	70.99***	-6.95***	-2.76***	-0.57	-0.57	-7.04***
Nonstore retailers	-1.99**	32.4***	-9.39***	-0.45	-0.19	-4.95***	-1.67*
Water transp.	-1.11	20.61***	-5.89***	-0.13	-0.3	-0.01	-0.39
Truck transp.	-1.01	77.61***	-1.46	-0.36	-0.08	-2.28**	-3.94***
Transit passenger transp.	-1.58	19.01***	-11.76***	-0.93	-0.55	-2.32**	-0.54
Pipeline transp.	-1.47	28.36***	-3.76***	-0.26	-0.82	-1.26	-0.2
Scenic transp.	-0.78	19.71***	-8.06***	-1.43	-0.95	-1.62	-0.4
Transp. supp. act.	-0.62	27.36***	-6.58***	-0.43	-0.5	-1.34	-0.27
Couriers and mess.	-0.1	51.41***	-5.1***	-0.32	-0.3	-0.33	-0.82
Wareh. and storage	-3.18***	23.24***	-5.53***	-0.29	-0.14	-0.17	-1.13
Publish. ind.	-2.54**	25.32***	-24.04***	-2.9***	1.99**	-1.78*	-1.63
Motion pict. and sound rec. ind.	-0.26	14.64***	-8.54***	-1.23	-0.75	-0.14	-0.87
Broadcast. and telecom.	-1.66*	20.6***	-9.8***	-1.54	-1.34	-2.51**	-1.21
Inform. and data process. serv.	-1.12	32.64***	-5.75***	-0.19	-0.31	-1.03	-0.66
Credit and rel. Act.	-2.77***	81.06***	-4.85***	-2.31**	-1.04	-4.66***	-3.87***
Security and like act.	-2.11**	21.77***	-11.11***	-0.57	-0.17	-1.48	-2.24**
Ins. carriers and rel. Act.	-0.49	56.16***	-5.85***	-1.21	-0.61	-3.04***	-1.66*
Funds and other fin. Veh.	-0.14	18.88***	-11.01***	-0.19	-0.16	-0.54	-1.28
Real estate	-2.07**	76.73***	-9.05***	1.74*	-2.1**	-6.46***	-4.12***
Rental and leasing serv.	-1.52	49.11***	-9.56***	-0.33	-0.54	-4.22***	-2.11**
Lessors of oth. intang. asset	-1.19	14.45***	-9.43***	-1.34	-0.94	-0.21	-0.57
Prof., scient. and techn. serv.	-0.89	76.93***	-7.55***	-0.11	-0.19	-2.78***	-0.68
Mgmnt of comp. and enterpr.	-1.89*	44.85***	-7.94***	-0.02	-0.75	-0.35	-2.13**
Admin. and supp. serv.	-0.26	32.4***	-9.99***	-0.6	-0.45	-0.59	-2.47**
Waste mgmnt and remed. serv.	-1.13	69.92***	-7.3***	-0.36	-0.35	-1.56	-0.28
Educational serv.	-2.29*	29.49***	-9.17***	-1.8*	-1.64	-3.69***	-0.2
Amb. health care serv.	-0.28	81.47***	-8.08***	-3.02***	2.41**	-4.71***	-1.14
Hospitals	-0.92	87.42***	-5.36***	-1.6	-1.35	-0.89	-0.13
Nursing care facilities	-0.62	83.04***	-5.05***	2.49**	-2.51**	-1.71*	-0.9
Social assistance	-2.44**	67.95***	-6.51***	-1.27	-0.61	-0.31	-3.4***
Performing arts, sports etc.	-1.54	25.63***	-15.31***	-0.4	-0.16	-0.05	-1.86*
Museums and like instit.	-0.95	25.6***	-5.73***	-1.23	-0.69	-2.77***	-1.21
Recreation industries	-1.8*	22.35***	-18.3***	3.33***	-3.12***	-5.52***	-2.85***
Accommodation	-0.65	71.01***	-3.28***	-0.38	-0.96	-0.67	-0.59
Food and drinking serv.	-5.55***	79.56***	-7.58***	1.88*	-2.49**	-4.47***	-1.57
Repair and maint	-2.76***	74.74***	-4.19***	-1.71*	-0.28	-3***	-2.91***
Pers. and laundry serv.	-0.82	70.63***	-5.83***	-4.85***	3.13***	-6.22***	-4.71***
Religious, grantm. like organiz.	-1.64	61.12***	-3.89***	-2.49**	1.67*	-3.05***	-2.81***

Table 4: Ecological model (model I): estimation by 3-digits NAICS.

panies has a significant negative coefficient; this will result less surprising after considering the wide range of economic activities encompassed in such category (in-house activities with a loose both sectoral and geographical connotation). Agriculture and Mining present puzzling results. A negative relation was expected between variation in wages and employment, a positive relation is found. This model is probably not appropriate to represent the evolution dynamics of those sectors across States. In particular in Agriculture sectoral characteristics are predominant, while in Mining locational features are crucial. In order to shed further light on the interplay between sectoral and territorial peculiarities on the one hand and wage dynamics on the other, the industrial classification is applied to 80 sub-sectors, and to the skill selective migration model. A first distinction is made between manufacturing-related activities and services. After controlling for a certain number of variables, attention is focused on the relation between variations in wages and variations in employment, and discuss the coefficients significant at 10% level of the confidence interval or more. Both manufacturing and services show interesting results: within some sectors there is a positive relation between variations in wage and variations in employment (table [6]), within some others an increase in wages is associated to a decrease in employment (table [6]). The explanation of these results resides in the different features that sectors exhibit. As to manufacturing industries, capital intensive sectors present a positive relation, whereas in labour intensive sectors the relation works the other way round. This is considered as an evidence of the adjustment mechanism in the labour market when the sector is labour intensive (an increase in the number of workers is absorbed through a decrease in wages) and of cumulative effects in capital intensive sectors, where skills are more likely to matter within production (this may be due to strong complementarities between labour and capital). As to services, the boundary line between positive and negative influence of wage variations and employment variation is the difference between services to production and to persons. In the first case employment shrink as wage goes up, working as a classical adjustment mechanism; in the second case (mainly social services and entertainment), higher wages and higher employment are positively related. These results are summarised in table [5], to highlight the switch of sign in the relation between variations in wages and variations in employment on the one hand, and the content embedded in the two groups in the different cases (services to production vs person, capital vs labour intensive manufacturing).

6 General Remarks

From the NAICS classification (model II), estimation results across macrosectors are composite. Four kinds of behaviours can be identified: industries

	positive relation	**negative relation**
Services	Services to persons	Services to production
Manufacturing	K-intensive	L-intensive

Table 5: Wage and employment variations: a taxonomy.

Negative relation	**Positive relation**
Manufacturing	
Building contracting	Fishing and hunting
Beverage, tobacco prod. mfg	Mining (exc. oil and gas)
Leather prod. mfg	Utilities
Nonmet. mineral prod. mfg	Textile mills
Electr. equip, appl. stores	Textile prod. mills
Bldg mat. and garden dealers	Apparel mfg
Food and beverage stores	Petroleum and coal prod. mfg
Pipeline transp.	Chemical mfg
Publish. ind.	Primary metal mfg
	Fabricated metal prod. mfg
	Transp. equip. mfg
	Miscellaneous mfg
Services	
Inform. and data process. serv.	Motor vehicle dealers
Security and like act.	Furniture stores
Real estate	Agr. and forestry supp act.
Mgmnt of comp. and enterpr.	Health and pers. care stores
Amb. health care serv.	Miscellaneous stores
Hospitals	Motion pict. and sound rec. ind.
	Broadcast. and telecom.
	Admin. and supp. serv.
	Nursing care facilities
	Social assistance
	Museums and like instit.
	Recreation industries
	Accommodation
	Food and drinking serv.

Table 6: Wage and employment variations by sectors.

Sector classification												
Dep. Var.: dempw	empsize	dwageD	empwg	hk	imm-rate	intra-est	inter-est	estD	estD2	year	Constant	Obs.
Agr.	(2.82)***	(2.28)**	(44.70)***	-0.19	(-6.62)***	(2.75)***	(-2.04)**	(-3.13)***	-0.01	-0.1	-0.09	1022
Mining	-1.02	(2.72)***	(28.47)***	-0.04	(-3.24)***	-0.28	-0.4	-1.62	-0.06	-0.46	-0.47	1134
Utilities	-1	(2.01)**	(13.32)***	-0.27	(-11.95)***	-1.37	-0.82	-0.55	-0.47	-0.04	-0.04	636
Constr.	(-1.79)*	(-5.27)***	(57.51)***	-0.12	(-12.78)***	-0.28	-1.29	(-5.18)***	(5.71)***	-0.23	-0.23	1828
Manufact.	(5.08)***	(3.82)***	(35.43)***	(-5.35)***	(-15.55)***	(3.30)***	-0.82	(-3.13)***	-0.47	-0.89	-0.85	10805
Retail Tr.	-1.08	(-2.40)**	(90.24)***	-0.34	(-16.43)***	(4.57)***	-1.28	(-10.37)***	(19.48)***	-0.79	-0.76	7598
Transp.	(2.27)**	-0.21	(49.12)***	(-3.08)***	(-9.89)***	-0.06	-0.43	(-3.40)***	-0.24	-0.04	-0.03	3380
Inf.	(2.25)**	(2.44)**	(25.92)***	-0.68	(-15.90)***	-1.3	-0.17	(-4.19)***	-1.24	-1.38	-1.38	2134
Finance	(2.52)**	-1.26	(39.24)***	-1.53	(-14.57)***	(4.35)***	(-1.91)*	(-5.40)***	(8.30)***	0	-0.01	2015
Real Est.	(2.46)**	-1.48	(84.54)***	(-2.49)**	(-14.19)***	-0.42	-0.42	-0.79	-1.03	(10.55)***	(-10.62)***	1476
Prof. Serv.	-0.9	-0.33	(85.89)***	-0.78	(-5.07)***	(-1.99)**	-0.84	(3.85)***	(-2.36)**	(11.56)***	(-11.60)***	660
Mgmt of Comp.	-0.82	(-5.00)***	(42.10)***	-0.05	(-7.42)***	-0.45	-0.81	-1.1	-1.49	-0.5	-0.51	639
Adm. Serv.	-0.28	-0.36	(84.90)***	-1.59	(-4.77)***	(1.86)*	-0.82	(-1.85)*	(2.65)***	(5.08)***	(-5.09)***	1095
Ed. Serv.	(2.19)**	-0.87	(27.83)***	-0.19	(-9.36)***	(-2.02)**	(1.83)*	(-4.08)***	-0.63	(-2.07)**	(2.04)**	652
Health Care	-0.33	-1.49	(186.25)***	-1.14	(-5.69)***	(-1.97)**	-1.14	(3.80)***	(-3.16)***	(23.31)***	(-23.41)***	2482
Arts	-1.57	(5.84)***	(52.72)***	-1.46	(-9.98)***	-0.93	-0.55	(-4.84)***	(-2.41)**	-1.05	-1.01	1807
Accom.	-0.9	(4.96)***	(128.54)***	-0.72	(-3.72)***	-0.55	-0.28	(2.51)**	(-2.52)**	(18.26)***	(-18.34)***	1315
Other Serv.	-0.18	(2.38)**	(136.97)***	-1.6	(-5.51)***	(-4.49)***	-1.63	(7.24)***	(-7.38)***	(18.00)***	(-18.07)***	1978
Wholesale Tr.	-0.55	-0.69	(73.29)***	-0.54	(-4.86)***	-0.31	-0.17	(3.79)***	(-2.31)**	(6.95)***	(-6.98)***	658

t-stats reported
∗ significant at 10%
∗∗ significant at 5%
∗ ∗ ∗ significant at 1%

Table 7: Skill selective model (model II).

Panel Data: 80 **NAICS sectors**

Model II (Skills)	empsize	dwageD	empwg	hk	imm-rate	intra-est	inter-est	estD	estD2
Forestry and logging	-1.05	0.39	26.87**	-0.14	-4.89**	1.71***	1.12	1.37	0.2
Fishing and hunting	-0.84	2.66**	18.78**	-0.46	-3.91**	0.38	0.57	0.85	0.68
Agr. and forestry supp act.	3.38**	1.66	30.50**	-0.19	-3.24**	2.01*	1.28	2.62**	0.39
Oil and gas extr.	0.29	1.33	29.95**	0.61	0.63	0.8	1.22	0.21	0.4
Mining (exc. oil and gas)	1.21	4.83**	20.59**	1.44	-10.30**	1.67***	1.26	2.23*	0.46
Mining supp. act.	0.25	-0.89	47.73**	2.68**	3.84**	0.82	0.17	1.73***	1.64
Utilities	1.01	2.02*	13.34**	-0.28	-12.01**	1.37	0.82	0.55	0.47
Building contracting	0.21	-5.15**	40.58**	0.32	-9.07**	0.29	0.94	0.42	3.91**
Heavy construction	0.75	-1.27	26.69**	1.65***	-12.93**	3.54**	2.39*	3.11**	0.66
Special trade contract.	2.74**	-0.44	69.30**	6.03**	-6.10**	2.26*	1.17	3.01**	2.81**
Food mfg	2.13*	-1.14	26.29**	0.41	-6.13**	2.04*	1.79***	4.39**	2.67**
Beverage, tobacco prod. mfg	1.56	-2.28*	54.12**	2.65**	-6.78**	1.06	0.79	1.28	0.79
Textile mills	2.48*	4.59**	29.11**	0.77	-2.72**	0.13	0.47	1.90***	0.8
Textile prod. mills	1.52	2.01*	25.77**	-1.15	-4.02**	1.25	0.87	0.97	0.45
Apparel mfg	3.61**	5.30**	38.65**	-1.16	-4.33**	0.72	0.03	1.85***	0.6
Leather prod. mfg	0.77	-2.05*	23.56**	-0.2	-7.19**	0.61	0.16	1.4	0.26
Wood product mfg	2.38*	-0.13	22.33**	0.11	-5.22**	1.85***	0.59	2.09*	0.73
Paper mfg	0.1	0.65	30.86**	1.03	-3.89**	1.94***	0.81	0.14	1.03
Printing and supp. activities	1.42	0.77	36.29**	2.05*	-6.03**	1.03	1.21	3.11**	1.76***
Petroleum and coal prod. mfg	0.7	4.16**	18.23**	2.16*	-4.93**	0.27	0.02	0.59	0.26
Chemical mfg	1.32	3.27**	36.47**	-0.92	-12.00**	0.74	0.8	1.43	0.13
Plastics prod. mfg	1.25	0.3	22.97**	-0.39	-4.23**	0.78	0.73	1.80***	0.42
Nonmet. mineral prod. mfg	0.9	-1.66***	56.99**	3.03**	-5.89**	0.67	0.41	1.51	0.09
Primary metal mfg	0.18	2.85**	14.11**	-0.32	-4.55**	1.2	1.02	0.26	0.35
Fabricated metal prod. mfg	1.02	1.89***	30.32**	-0.89	-1.37	2.77**	2.38*	2.37*	0.96
Machinery mfg	0.13	-0.08	25.77**	-0.62	-2.20*	2.08*	0.82	2.07*	1.14
Computer and el. prod. mfg	2.77**	-0.33	15.20**	-2.93**	-6.58**	1.33	0.92	1.38	2.03*
Electr. and appl. Mfg	1.57	-1.39	20.47**	-0.59	-12.14**	2.01*	1.33	0.27	1.41
Transp. equip. mfg	0.57	7.17**	13.89**	-0.63	-1.83***	0.13	0.91	1.18	0.5
Furnit. and rel. prod. mfg	1.57	-1.61	33.32**	-0.84	-7.09**	2.18*	1.16	0.22	0.93
Miscellaneous mfg	1.79***	3.05**	22.84**	0.08	-2.75**	2.93**	1.91***	0.43	1.25
Wholesale trade	0.67	-0.13	70.63**	3.37**	-6.62**	0.62	0.87	2.48*	0.68
Motor vehicle dealers	0.81	0.55	73.26**	6.93**	-3.39**	6.71**	3.00**	3.28**	12.34**
Furniture stores	1.01	0.07	77.20**	6.00**	-4.80**	2.63**	1.43	0.96	2.98**
Electr. equip, appl. stores	4.11**	-1.70***	71.19**	5.71**	-3.47**	0.67	0.08	2.06*	1.07
Bldg mat. and garden dealers	1.16	-2.07*	74.00**	3.90**	-1.03	4.01**	2.62**	5.49**	9.43**
Food and beverage stores	0.08	-3.64**	43.89**	1.80***	-7.50**	0.25	0.65	2.38*	5.13**
Health and pers. care stores	0.88	2.57*	74.72**	5.67**	-3.59**	5.84**	3.68**	5.39**	7.91**
Gasoline stations	5.66**	1.44	77.33**	7.13**	-3.46**	4.69**	3.94**	1.62	7.10**
Clothing and access. stores	0.88	-0.89	76.93**	6.39**	-6.42**	6.19**	4.14**	5.73**	6.42**
Sporting and hobby stores	4.30**	0.59	67.63**	6.34**	-7.28**	6.67**	5.02**	1.72***	2.42*
General merch. stores	0.93	0.16	75.09**	2.02*	0	6.64**	4.07**	1.26	8.65**
Miscellaneous stores	0.72	2.62**	74.51**	7.83**	-9.73**	3.03**	0.64	2.18*	7.09**
Nonstore retailers	2.53*	0.26	31.46**	2.61**	-9.85**	1.07	0.65	5.25**	1.85***
Water transp.	1.12	-0.06	15.75**	1.54	-6.02**	0.05	0.24	0.21	0.39
Truck transp.	-1.58	1.05	70.49**	5.13**	-2.97**	0.01	0.45	1.26	3.39**
Transit passenger transp.	0.72	-0.25	34.86**	0.42	-13.95**	0.61	0.57	3.50**	1.46
Pipeline transp.	0.95	-1.84***	31.53**	-1.09	9.64**	0.87	1.38	0.6	1.16
Scenic transp.	0.23	1.41	31.42**	0.12	-3.61**	0.18	0.04	1.37	0.05
Transp. supp. act.	1.73***	0.78	22.02**	-0.31	-5.75**	0.58	0.55	1.25	0.35
Couriers and mess.	2.07*	1.63	53.19**	1.56	-6.53**	0.33	0.13	1.52	0.03
Wareh. and storage	0.43	0.1	18.67**	-0.09	-5.63**	0.86	0.35	0.02	1.83***
Publish. ind.	0.75	-1.67***	70.80**	4.44**	-4.84**	1.12	0.9	1.83***	1.75***
Motion pict. and sound rec. ind.	2.22*	7.68**	15.52**	-0.94	-11.08**	0.99	0.52	1	0.84
Broadcast. and telecom.	0.55	5.59**	19.63**	0.7	-10.81**	1.94***	1.62	2.59**	1.23
Inform. and data process. serv.	0.11	-3.32**	28.37**	0.02	-5.06**	0.24	0.02	0.04	0.39
Credit and rel. Act.	2.61**	0.82	76.68**	5.84**	-5.18**	2.85**	1.33	5.03**	4.40**
Security and like act.	2.05*	-4.10**	23.60**	1.83***	-20.08**	0.44	0.58	2.48*	2.99**
Ins. carriers and rel. Act.	0.52	0.11	57.31**	3.24**	-6.20**	2.00*	1.01	3.47**	2.63**
Funds and other fin. Veh.	0.52	0.16	16.57**	-1.90***	-8.60**	0.06	0.07	0.07	1.05
Real estate	2.65**	-2.09*	73.55**	4.79**	-9.65**	0.6	1.08	5.75**	3.78**
Rental and leasing serv.	0.15	0.33	64.47**	6.84**	-6.95**	3.73**	2.85**	0.43	0.38
Lessors of oth. intang. asset	0.04	1.41	29.86**	-0.19	-9.60**	0.66	0.84	0.72	0.29
Prof., scient. and techn. serv.	1.42	0.42	78.23**	4.88**	-8.20**	1.12	0.59	2.55*	0.52
Mgmnt of comp. and enterpr.	-1.07	-5.03**	42.17**	-0.07	-7.48**	0.4	0.77	1.1	1.49
Admin. and supp. serv.	0.41	1.84	81.78**	5.41**	-12.45**	0.73	0.54	4.80**	1.92***
Waste mgmnt and remed. serv.	-2.27*	0.81	54.82**	-0.12	-6.54**	0.2	0.36	2.30*	0.21
Educational serv.	2.20*	0.8	27.90**	-0.62	-9.13**	1.88***	1.69***	3.64**	0.12
Amb. health care serv.	0.5	-1.88***	84.48**	6.68**	-9.56**	4.06**	3.10**	5.19**	1.22
Hospitals	1.08	-1.69***	78.12**	4.34**	-6.03**	1.17	1.08	0.48	0.47
Nursing care facilities	-3.98**	3.97**	85.54**	3.95**	-5.19**	0.72	1.26	1.25	0.57
Social assistance	-0.96	2.93**	67.87**	5.88**	-7.26**	2.42*	1.57	1.1	4.07**
Performing arts, sports etc.	1.14	0.27	40.07**	2.46*	-6.48**	0.2	0.15	1.26	0.42
Museums and like instit.	0.94	7.78**	25.76**	-0.47	-5.04**	1.33	0.76	2.81**	1.26
Recreation industries	0.84	2.85**	61.46**	3.09**	-8.71**	1.26	1.45	4.69**	2.46*
Accommodation	0.05	2.32*	70.44**	5.04**	-4.54**	0.1	1.22	0.79	0.58
Food and drinking serv.	-4.61**	7.92**	86.04**	5.92**	-9.27**	1.28	2.17*	5.08**	1.70***
Repair and maint	2.53*	0.05	76.46**	6.16**	-5.93**	2.54*	0.85	3.57**	3.47**
Pers. and laundry serv.	0.8	1.14	73.06**	6.64**	-7.61**	5.38**	3.49**	6.52**	4.46**
Religious, grantm. like organiz.	1.15	1.54	60.47**	3.25**	-4.50**	2.49*	1.6	3.43**	2.84**

Table 8: Skill selective model (model II): estimation by 3-digits NAICS.

in which it is verified the hypothesis that higher wages are associated to positive employment dynamics for the skills embedded in sector peculiarities (Utilities, Manufacturing, Information, Health Care, Arts, Accomodation Services, Other Services); industries in which the poor potentialities of skills in boosting growth reflect in wages as a pure adjustment mechanism — the higher wages, the lower employment — (this is the case in Construction and Retail); industries with non significant relationship and industries with counterintuitive relation (this is the case of Agriculture and Mining, whose sectoral and geographical peculiarities respectively may even question the idea of considering them as industrial clusters as here intended).

When the analysis is applied for each of the 80 3-digits NAICS sectors, the findings become clearer and they allow me to draw the following conclusions. The pure ecological model (model I) identifies, among the others, two sets of clusters: (1) a group of clusters (mainly raw services) show both negative intra-industry dependences and positive inter-industry dependences; (2) a group with positive intra-industry dependences and negative inter-industry dependences. The first group identifies clusters for which sectoral competition seems to play a crucial role as well as territorial synergies. In these cases, clusters grow with diversity and compete with the same sectors in different States. The second group corresponds to clusters that can, to some extent, benefit from vertical integration or competing on region specific demand or supply factors with other clusters in the same State. In these cases, specialisation seems to boost growth. Recalling Glaeser et al. (1992), in the first case the engine for growth hinges in territorially specialised industries, as predicted by Romer (1986), Arrow (1962), Marshall (1890) and Porter (1990). In the other case, Jacob's argument (Jacobs, 1969) that variety and diversity of geographically proximate industries boost clusters growth is verified. The skill selective migration model (model II) has allowed me to identify a taxonomy between services and manufacturing activities. Within services, higher wages identify higher employment for services to persons, whereas the opposite result holds in the case of services to production; within manufacturing, the positive relation concerns capital intensive activities, whereas labour intensive ones experience lower employment for higher wages.

7 Conclusions

Empirically, two different outcomes can arise; when the increase in labour demand effect is not able to compensate the migration inflows, probably because, within the human capital setting hereby discussed, skills brought by movers are not enough to boost growth and by that way to shift significantly upright labour demand, the ecological setting forecasts a prevailing competition effect, positively affecting the size of one cluster in detriment of the other one (in a two-cluster setting). When in-migration is considered

to be able to boost growth within the host cluster and the final equilibrium is at higher employment and wage levels, the ecological framework exhibits complementarities, providing advantages for both interacting clusters. As a main conclusion, the role of wage in exploring employment is basically sector specific and this is considered as an evidence of relevant skills to affect such relationship. From an empirical point of view, several results have been described. Among the main ones, from the sector classification, estimation results across 20 macrosectors highlight three kinds of behaviours can be identified with reference to the relation between the variation in wages and the variation in employment: industries in which it is verified the hypothesis that higher wages are associated to positive employment dynamics for the skills embedded in sector peculiarities (Utilities, Manufacturing, Information, Health Care, Arts, Accomodation Services, Other Services); industries in which the poor potentialities of skills in boosting growth reflect in wages as a pure adjustment mechanism — the higher wages, the lower employment — (this is the case in Construction and Retail); industries in which the relationship between wages and employment is counterintuitive (this is the case of Agriculture and Mining, whose sectoral and geographical peculiarities respectively may even question the idea of considering them as industrial clusters as here intended).

When the analysis is applied for each of the 80 3-digits NAICS sectors, the findings become clearer. The pure ecological model (model I) identifies, among the others, two sets of clusters: (1) a group of clusters (mainly raw services) show both negative intra-industry dependences and positive inter-industry dependences; (2) a group with positive intra-industry dependences and negative inter-industry dependences. The first group identifies clusters for which sectoral competition seems to play a crucial role as well as territorial synergies. In these cases, clusters grow with diversity and compete with the same sectors in different States. The second group corresponds to clusters that can, to some extent, benefit from vertical integration or competing on region specific demand or supply factors with other clusters in the same State. In these cases, specialisation seems to boost growth. Recalling Glaeser et al. (1992), in the first case the engine for growth hinges in territorially specialised industries, as predicted by Romer (1986), Arrow (1962), Marshall (1890) and Porter (1990); in the other case, the argument by Jacobs (1969) that variety and diversity of geographically proximate industries boost clusters growth is verified. The skill selective migration model (model II) identifies a taxonomy between services and manufacturing activities. Within services the higher wages identifies higher employment for services to persons, whereas the opposite result holds in the case of services to production; within manufacturing, the positive relation concerns capital intensive activities, whereas labour intensive ones experience lower employment for higher wages.

As a general implication for policy making, regional disparities are more

likely to persist and widen rather than narrow for the existence of automatic adjustment mechanisms based on congestion costs. Counteracting this tendency can imply a trade off between efficiency and equity and optimal policies can be based on education and training provision as escapes to cumulative effects acting in favour to more advanced regions. These policies can be successful only within a systemic framework encompassing on the one hand targeted skill building policies and on the other hand, the promotion of the attraction of new technologies to close the interregional productivity gap. At the same time, a necessary condition for the growth process to be triggered is to set up targeted infrastructure policies. Scholars usually support (Baldwin et al., 2003) the idea that an important market failure would come from some factors being immobile. As a consequence, it is not agglomeration *per se* to be harmful, but the fact that the welfare of immobile factors is not taken into account. If this were the plain story, the reduction of transaction costs to ease mobility would be the response. Thus on the basis of the empirical results obtained, it would be desirable to displace workers from clusters with negative inter/intra dependences to clusters with positive values. Such an intervention, however, would imply social costs for people moving and require targeted training policies. Moreover, this risks to desert peripheral areas, which can only rely on rents and wages in the advanced region to speed up congestion costs dynamics (after rents and wages are high enough in the richer region, firms start to settle in the laggard region). In this sense, integrated policies would address the question about which opportunities the laggard regions can be given, in particular pursuing proper technological oriented policy would offer the chance to leapfrog (Brezis and Krugman, 1993, Desmet, 2000). When technological change takes place within a given trajectory, cumulative effects are likely to occur. When a new breakthrough technology is invented, two outcomes are possible: either persistence or leapfrogging. If technological progress seems inferior to the older and is sharply different in skill requirements, the economic leadership can become the source of its own downfall. Brezis et al. (1993) identifies the following conditions for leapfrogging, thus resulting in catch up opportunities for laggard regions with respect to the richer:

- high wages in the leading region relatively to the laggard's;
- appearance to the old technology owner for the new technology to seem inferior;
- low readaptation of skills from the old to the newer technology;
- productivity improvement of the new technology over the old one.

The policy maker has to be aware of this possibility and act consistently.

References

A. Almazan, A. de Motta, and S. Titman. Firm Location and the Creation and Utilization of Human Capital. *NBER Working Paper Series*, (WP N. 10106), November 2003.

K. J. Arrow. The Economic Implications of Learning by Doing. *Review of Economic Studies*, 29(3):155–173, June 1962.

R. E. Baldwin, R. Forslid, P. Martin, G. Ottaviano, and F. R. Nicoud. *Economic Geography and Public Policy*. Princeton University Press, 2003.

V. Bencivenga and B. Smith. Unemployment, Migration, and Growth. *The Journal of Political Economy*, 105(3):582–608, June 1997.

C. Bentivogli and P. Pagano. Regional Disparities and Labour Mobility: the Euro-11 versus the USA. *Labour*, 13(3):737–760, 1999.

O. J. Blanchard, L. F. Katz, R. E. Hall, and B. Eichengreen. Regional evolutions. *Brookings Papers on Economic Activity*, 1992(1):1–75, 1992.

G. J. Borjas. The Economics of Immigration. *Journal of Economic Literature*, 32(4):1667–1717, December 1994.

E. S. Brezis and P. R. Krugman. Technology and the life cycle of cities. *NBER Working Papers*, 4561, December 1993.

E. S. Brezis, P. R. Krugman, and D. Tsiddon. Leapfrogging in international competition: A theory of cycles in national technological leadership. *The Americal Economic Review*, 83(5):1211–1219, December 1993.

M. Brülhart and R. Traeger. An Account of Geographic Concentration Patterns in Europe. *Mimeo, University of Lausanne*, 2002.

S. Commander, M. Kangasniemi, and L. A. Winters. The Brain Drain: Curse or Boon? *IZA Discussion Papers*, (809):1–34, June 2003.

K. Desmet. A perfect foresight model of regional development and skill specialization. *Regional Science and Urban Economics*, (30):221–242, 2000.

J. J. Dolado, A. Goria, and A. Ichino. Immigration, Human Capital and Growth in the Host Country: Evidence from Pooled Country Data. *Journal of Population Economics*, 2(7):193–215, 1994.

G. Dosi. Technological paradigms and technological trajectories. *Research Policy*, (11):147–162, 1982.

R. Frisch and F. Waugh. Partial Time Regressions as Compared with Individual Trends. *Econometrica*, (1):387–401, 1933.

S. Ghatak, P. Levin, and S. Wheatley Price. Migration Theories and Evidence: An Assessment. *Journal of Economic Surveys*, 10(2):159–198, June 1996.

E. L. Glaeser, H. D. Kallal, J. A. Scheinkman, and A. Shleifer. Growth in cities. *Journal of Political Economy*, 100(6):1126–52, 1992.

W. H. Greene. *Econometric Analysis. Fifth Edition*. Prentice Hall, 2003.

J. Jacobs. *The Economy of Cities*. New York: Random House, 1969.

P. R. Krugman. Increasing Returns and Economic Geography. *The Journal of Political Economy*, 99(3):483–499, 1991.

R. Linda. Méthodologie de l'analyse de la concentration appliquée à l'étude des secteurs et des marchés. *European Commission*, 1976.

R. E. Lucas. On the mechanics of economic development. *Journal of Monetary Economics*, 22(1):3–42, 1988.

M. A. Maggioni. *Clustering Dynamics and the Location of High-Tech Firms*. Heidelberg and New York: Springer-Verlag, 2002.

M. A. Maggioni and M. R. Riggi. Forme alternative di collective learning: un approccio simulativo. In R. Camagni and R. Capello, editors, *Collective learning e competitività regionale*. Franco Angeli, 2002.

J. Mairesse. Lectures on applied panel data. Bertinoro. Summer School in Econometrics, CIDE, 2003.

A. Marshall. *Principles of Economics*. London: Macmillan, 1890.

A. Marshall. *Industry and Trade*. London: Macmillan, 1921.

I. Nonaka and H. Umemoto. A theory of organizational knowledge creation. *International Journal of Technology Management*, 11:833–845, 1996.

M. E. Porter. *The Competitive Advantage of Nations*. New York: The Free Press, 1990.

M. E. Porter. Clusters and competition: new agendas for companies, governments, and institutions. In M. E. Porter, editor, *On Competition*, pages 197–287. Boston: Harvard Business School Press, 1998.

P. Romer. Increasing Returns and Long-Run Growth. *Journal of Political Economy*, (94):1002–1037, October 1986.

G. M. Shields and M. P. Shields. The Emergence of Migration Theory and a Suggested New Direction. *Journal of Economic Surveys*, 3(4):277–304, December 1989.

R. M. Solow. A Contribution to the Theory of Economic Growth. *Quarterly Journal of Economics*, 70(1):65–94, February 1956.

J. Suedekum. Subsidizing Education in the Economic Periphery: Another Pitfall of Regional Policies? *University of Goettingen, Mimeo*, November 2002.

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

A. Venturini. *Le migrazioni e i Paesi sudeuropei. Un'analisi economica.* Torino: UTET libreria, 2001.

A. Wood and C. Ridao-Cano. Skill, Trade and International Inequality. *IDS*, (47), 2002.

8 Appendix: Data Issues

The County Business Patterns, by the US Census Bureau, provides data on establishments, employment and annual payroll by detailed industry for all counties in the United States across the time spans 1988 to 2001. Data are converted to the later NAICS classification, which offers a good level of detail, a satisfactory degree of economic meaningfulness and the possibility to bridge earlier data (1988 to 1997), from the SIC (Standard Industry Classification) to the NAICS (North American Industrial Classification System) classification. This bridging is non fully possible, but at the degree of detail chosen, it has been proven to be save from significant time series breaks due to errors in measurement. The degree of reliability is deemed satisfactory, even once the disclosure issue has been taken into account; for confidentiality reasons, information on employment and payroll are withheld, but in these cases further information on employment class sizes by establishment, allows the implementation of the analysis.

Other data are used in the empirical analysis, from various sources, mainly derived from the Census Bureau. In particular, from the U.S. Citizenship and Immigration Services (USCIS) data have been made available about Immigration Statistics. From the U.S. Department of Education, National Center for Education, the following proxy for human capital has been used: Population age 25 and over, percent of population age 25 and over with bachelor's degree are provided by "School Enrollment-Social and Economic Characteristics of Students," (P-20 Current Population Reports).

8.1 Overview

The data base which the current analysis is based on, and all information concerning data, is provided by the US Census Bureau.
County Business Patterns provides data by detailed industry for all counties in the United States. The choice of the empirical unit of analysis is mainly driven by the long tradition in data availability by the U.S. Census Bureau, allowing for a significant industrial[4] and territorial level of analysis[5]. The North American Industry Classification System (NAICS) replaces the U.S. Standard Industrial Classification (SIC) system. NAICS identifies hundreds of new, emerging, and advanced technology industries, and reorganises industries into more meaningful sectors–especially in the service-producing segments of the economy. NAICS provides for comparable statistics among the three NAFTA trading partners (USA, Canada and Mexico). NAICS defined industries according to a consistent principle: businesses that use similar production processes are grouped together, giving a more adaptability of industrial classification to the real economic world. While many of the individual SIC industries correspond directly to industries as defined under the NAICS system, most of the higher level groupings do not. Particular care should be taken in comparing data for retail trade, wholesale trade, and manufacturing, which are sector titles used in both NAICS and SIC, but cover somewhat different groups of industries.
The aggregation level for counties and sectors is dealt with in detail in section 2; eventually, 4233 clusters (State – industry couplets) across the time span $1988-2001$ are considered.

[4]The classification adopted is 4-digit Standard Industrial Classification (SIC) for the time span 1993 to 1997 and North American Industry Classification System (NAICS) starting from 1998. A conversion, later discussed, has been carried out to express the whole data set accordingly the NAICS system.

[5]Business Patterns presents data on the total number of establishments, employment and payroll for more than $40,000$ ZIP Code areas nation-wide.

The definition of industry – State couplets is consistent with the target here pursued to investigate sector specific patterns of occupational flows, and widely discussed in section 2.
The time span ranges between 1988 and 2001. Earlier data are not considered, thought it would have been possible to go back to 1986, because of another structural change in industrial classification occurred in 1987, when the recording system switched to SIC-1987. Table 9 reports the highest level of industry detail adopted. The first two figures of the relevant NAICS code indicate that two different sectors belong to the same group in the 2-digits NAICS classification, so defining groups of sectors defined as macrosectors and illustrated in table [1][6].

From the data set, we have computed the following list of variables used in the estimations:

Variation in Employment (dempw) Both theoretically and empirically crucial the variation in employment over time is assumed as the dependent variable to be explained, since representing the evolution of clusters over time. It is defined as the time differences of the employment mass by cluster in logarithms (ln $empw_t$ - ln $empw_{t-1}$).

Number of Establishments and squared term (estD and estD2) An establishment is a single physical location at which business is conducted and/or services are provided. It is not necessarily identical with a company or enterprise, which may consist of one establishment or more. Sector heterogeneity is taken into account by adjusting variables by their within-industry mean. It mainly represents scale effects within clusters and the squared term allows to control for possible non-linearities.

Establishments size and Employment Size (estsize, empsize) By the theoretical insights, employment is the main independent variable, but it cannot be included as a regressor for the endogeneity issues illustrated in section 4, where the solution proposed consists in instrumenting it, through establishments size and firms's employment size. Variables have a twofold meaning. On the one hand they play the role of instrumenting N_t^{is}, the level of employment which cannot be included as in the original ecological derived specification (remember that as a statistical consequence, the consistency and the efficiency of the estimation procedure could lead to misleading interpretations of the results). On the other hand, both variables allow me to control for some dimension-related variable, preventing the results to be affected by different behaviours across different clusters size.

Interdependences (Intra and Inter) Glaeser et al. (1992) test knowledge spill-overs in territorially specialised vs diversified industries. Spill-overs may occur within or between industries. In the first case, empirical evidence is found for the predictions by Romer (1986), Arrow (1962) and Marshall (1890), which, to different extents[7] stress the relevance of industrial specialisation for externalities between firms: through spying, imitation, and skilled labour poaching, ideas rapidly disseminate in an industry. In the second case, the argument by Jacobs (1969) is found to be consistent: a variety and diversity of

[6]The only exception being the Wholesale trade, which is both a 2-digits and a 3-digits sector. It has been artificially created by grouping the durable and non durable goods distinction, because the bridging from SIC to NAICS classification was not possible.

[7]In particular, Romer (1986), Arrow (1962) and Marshall (1890) focus on the role of monopoly for growth, restricting the flows of ideas and allowing for externalities to be internalised by the innovator, whilst Porter (1990) argues that knowledge spill-overs are more relevant to growth in a competitive environment, and firms' dimension is considered as a proxy of monopolistic power.

NAICS code	Sector
113	Forestry and logging
114	Fishing, hunting and trapping
115	Agriculture and forestry support activities
211	Oil and gas extraction
212	Mining (except oil and gas)
213	Mining support activities
221	Utilities
233	Building, developing and general contracting
234	Heavy construction
235	Special trade contractors
311	Food mfg
312	Beverage and tobacco product mfg
313	Textile mills
314	Textile product mills
315	Apparel manufacturing
316	Leather and allied product mfg
321	Wood product mfg
322	Paper mfg
323	Printing and related support activities
324	Petroleum and coal products mfg
325	Chemical mfg
326	Plastics and rubber products mfg
327	Nonmetallic mineral product mfg
331	Primary metal mfg
332	Fabricated metal product mfg
333	Machinery mfg
334	Computer and electronic product mfg
335	Electrical equip, appliance and component mfg
336	Transportation equipment mfg
337	Furniture and related product mfg
339	Miscellaneous mfg
420	Wholesale trade
441	Motor vehicle and parts dealers
442	Furniture and home furnishing stores
443	Electronics and appliance stores
444	Bldg material and garden equip and supp dealers
445	Food and beverage stores
446	Health and personal care stores
447	Gasoline stations
448	Clothing and clothing accessories stores
451	Sporting goods, hobby, book and music stores
452	General merchandise stores
453	Miscellaneous store retailers
454	Nonstore retailers
483	Water transportation
484	Truck transportation
485	Transit and ground passenger transportation
486	Pipeline transportation
487	Scenic and sightseeing transportation
488	Transportation support activities
492	Couriers and messengers
493	Warehousing and storage
511	Publishing industries
512	Motion picture and sound recording industries
513	Broadcasting and telecommunications
514	Information and data processing services
522	Credit intermediation and related activities
523	Security, commodity contracts and like activity
524	Insurance carriers and related activities
525	Funds, trusts, and other financial vehicles (part)
531	Real estate
532	Rental and leasing services
533	Lessors of other nonfinancial intangible asset
541	Professional, scientific and technical services
551	Management of companies and enterprises
561	Administrative and support services
562	Waste management and remediation services
611	Educational services
621	Ambulatory health care services
622	Hospitals
623	Nursing and residential care facilities
624	Social assistance
711	Performing arts, spectator sports, and related industries
712	Museums, historical sites and like institutions
713	Amusement, gambling and recreation industries
721	Accommodation
722	Food services and drinking places
811	Repair and maintenance
812	Personal and laundry services
813	Religious, grantmaking, civic, prof and like organizations

Table 9: Sector classification by 3-digits NAICS

geographically proximate industries foster knowledge transfers from outside the industry. Let me apply this to the current context.

The growth of the relevant industry at a nationwide level (g_t^{i-}) is included as a regressor, to control for industrial nationwide trend when investigating cluster's growth. Inter industry spill-overs are instead taken into account with the rate of growth of firms outside the core industry but inside the relevant macrosector[8].

To set up variables **Intra** and **Inter**, two kinds of clusters are considered: the one examined and all the others. The rest of the world (USA nationwide) is supposed to affect one's development path, if it belongs to the competitive arena defined through the Linda series (see section 3.2); in this way, only the most important clusters [9] are considered to be able to affect the 'rest of the world' development path. After this first selection, a distinction is made between clusters belonging to the same State where the one under analysis is, but operating in a different sector (inter-industry dependences) and clusters belonging to the same sector where the one in exam is, but operating in different States (intra-industry dependences). *Intra* and *Inter* variables, so defined, provide me with cooperation and competition effects of regional and sector interactions.

Wages (w^*) Exploiting the industrial variability available on wages, we address two different but related questions, in a twofold way with the same variable: which is the role of factor cost component embedded in wages coming from State peculiarities ($w_t^S = w_t^{is} - w_t^{i-}$) and the skill effect component specific to each industry ($w_t^I = w_t^{is} - w_t^{-s}$)[10].
The skill effect component can be included exploiting the large variety in sector classification available in CBP and is embedded in the variable wage, computed as a difference between cluster's wage and average industry wage. In this way wages are considered as differences between two close sectors, assuming intra sector labour mobility and to some extent determinants of labour poaching are considered (i.e. the effect of a relative pressure of labour demand on labour supply). Notice that the effect embedded in w_t^I is accounted for by the "National Industrial Trend" (g_t^{i-}).

Wages are included in the regression as deviations to grasp the performances of clusters relative to the others within the same State. How much does a cluster pay relative to other industries within the same State? Let w^{is} be the prevailing wage in the cluster identified by State s and industry i, and w_t^{-s} be the mean of wage across States. This allows the variable w^*=ln w^{is}-ln w^{-s}, to take into account for differences in living costs across States and the level of different purchasing power[11]. The analysis being aimed at assessing the contribution to the employment mass, w^* allows me to identify the specific contribution of wages to cluster's growth. This factor has to be taken into account in order to properly assess clusters' evolution, that could otherwise be blurry because of sector wide trends and structural differences across States. Besides, this variable represents the skill component of employment growth; to this extend, further explanations will be provided in section 4.3.

National Industrial Trend (g_t^{i-}) How much does a cluster grow relatively to the nationwide industry it belongs to? This effect is controlled through this

[8]Industries whose content is completely different may bias the analysis with spurious correlations to the variables under investigation.

[9]In terms of market share held.

[10]The superscript (the middle score -) represents average across the index it replaces.

[11]The cluster's wage is doubled for computational purposes only.

variable which allows in part to control for a source of unobserved heterogeneity which is here particularly relevant, the one stemming from the industry specific skilled content of employment which is theoretically relevant to the framework outlined but that cannot be tested because the CBP database does not contain such information. In this way, skills are proxied without knowing them and through the information available on sectors, represented in a detailed way, which have strong skill content specificities.

Educational Attainment(HK_t^{-s}) As a proxy of human capital, the hypothesis is made that the degree of education of the State population can control for State specific human capital endowment, which is able to explain one part of the variation of wage[12].

International Immigration ($Imm - rate_t^{-s}$) The Immigration from outside the US to each State is assumed to control for occupational effects on wages for reasons not strictly related to workforce skill composition[13].

8.2 Data Management

NAICS being introduced for a more detailed and reasonable classification, analyses earlier than 1998 bear a problem of non full comparability. In the reminder of this section it is discussed why it is deemed worthwhile to deep the question for a better assessment of the possible implications of this switch.
The creation of new sectors and the expansion of the industry code from four to six digits certainly suggest that this revision of the industry classification system is more profound than earlier SIC revisions. Yet neither of these changes necessarily affects the ability to link old data on an SIC basis to new data on a NAICS basis. Data for more than two-thirds of all 4-digit SICs will be derivable from the NAICS system, either because the industry is not being changed (other than in code), or because new industries are being defined as subdivisions of old ones. The total number of industry classifications is increasing with NAICS (1170 industries applicable in the U.S., up from 1004 SIC). The switch is less pervasive if data are bridged for the most detailed SIC availability to closest NAICS, and then data aggregated at NAICS 3 level. In this case most time series breaks collapse when data are aggregated, and vanish if further aggregations to macrosectors are implemented. Besides, given the purpose of the analysis, relation between variables at the same time are insensitive to such issue. Most of the numbers used take on greater meaning when compared to data for other time periods, allowing inference of development or change. This is the main reason why a huge work has been undertaken to translate data earlier than 1998 into NAICS classification[14]. The implementation of NAICS causes disruptions in the availability of time series data, not only for individual industries that are redefined but also for the broad sectors, like manufacturing and

[12]Educational Attainment of the Population 25 Years and Over, By State, Bachelor's degree or more. Percent of population Data for 1992 being not available, a computational trick is adopted. For 1992, data is available for percentage of high school attenders. In this case, Bachelor's degree or more is proxied by the high school attenders, corrected by the difference between Bachelor's degree or more and high school attenders for 1993, for which both data are available. Data for District of Columbia are missing. Data for 1990 are provided by the U.S. Department of Commerce, Bureau of the Census, Current Population Reports. Data for 1988 are missing.

[13]Immigrants Admitted by State of intended residence, fiscal years 1989-2001. For the empirical analysis, the relative ratio to employment is used.

[14]The alternative being to translate more recent data back to SIC classification. This was not possible because of the later part of data was not available for a suitable detailed level of aggregation and because it has been deemed worthwhile to draw a picture consistent with a more meaningful classification.

retailing. Ideally, given the need to start building the new time series, it would have been desirable to apply the new classification criteria to existing data. This is possible where SICs have direct NAICS, but most of the new classifications require information not available for earlier data. The reasons why the bridge to NAICS is not the cleanest, are mainly three.
First, aggregations of SIC classified subsectors to total NAICS 3-digits sector totals[15] do not always return the sector total, since some detailed data are withheld for data confidentiality. In fact, while nearly all industries have been defined in such a way that there are enough companies present to support tabulation at the national level, many of the SIC-by-NAICS combinations may be so small that they are dominated by individual companies.
Second, some SIC industries simply do not directly bridge to NAICS, because some SIC has been split further or simply because they are introduced *ex novo*.
Third, businesses have reacted to the new classification and took advantage from the new classification to stress the core of their activity (which the classification is based on).
Managing the conversion, all newly recorded industries and those no more existing are left aside[16].
Besides, the only 'pure' alternative is to drop all information back to 1998, the date NAICS were introduced; to this extent, whatever procedure adding more information will allow the user to get data at least as good as the narrower sample does. This corresponds to the principle that 'data never hurt' (Mairesse, 2003): getting an estimate by an extended data sample, the bias incurred is at most the same as if further data were not included in the sample.

8.3 Disclosure

In accordance with the U.S. Code, Title 13, Section 9, no data are shown that would disclose the operations of an individual employer. However, the number of establishments in an industry classification and the distribution of these establishments by employer-size class are not considered disclosures, and so this information may be released even though other information is withheld. In these cases, a computational trick is adopted. The problem arises whenever firms are easily identifiable. In those cases, the number of establishments are provided by 9 employment classes, for firms engaging on average 1 to more than 1000 employees (see table [10]); besides, the the number of employees by firm size class in substitution of the withheld exact value is reported (see table [11]). This implies a loss of information that has been overcome setting up a new variable on employment, such that data are reported whenever available and replaced, if missing, by a composition of the information provided for withheld data (data suppression flags and employment class-size firms). From the number of establishments for each employment class, firms are supposed to engage the median value of the employment class under consideration[17]; from the variable

[15]This is the way available data have been bridget to NAICS: starting from the most detailed available SIC sector to the NAICS 3-digits sector, which is the detail deemed relevant for the purposes of this paper.

[16]In total less than 10 sectors have been neglected, resulting in the following NAICS 3-digits: 92 (public administration), 95 (auxiliaries, except corporate, subsidiary and regional management), 111 (agriculture), 481 (air transportation) and 521 (monetary authorities central bank). Sectors 421 (wholesale trade, durable goods) and 422 (wholesale trade, nondurable goods) have been jointly classified in a joint sector 420 (wholesale trade). Overlapping SIC to NAICS sectors have been assigned to NAICS industries on a business-like basis whenever possible, and with trial and error to reduce time series breaks effects when no criteria was available.

[17]for example, for withheld data on employment, if it is known that in one sector the number of establishments N belong to an employment class size of 1 to 4 employees, it

Flag	Average employees
A	0-19
B	20-99
C	100-249
E	250-499
F	500-999
G	1,000-2,499
H	2,500-4,999
I	5,000-9,999
J	10,000-24,999
K	25,000-49,999
L	50,000-99,999
M	100,000 or More

Table 10: Establishments' average employee class size

Number of Establishments	Employee Size Class
N1_4	1-4
N5_9	5-9
N10_19	10-19
N20_49	20-49
N50_99	50-99
N100_249	100-249
N250_499	250-499
N500_999	500-999
N1000	1,000

Table 11: Number of establishments by employee class size

'Data Suppression Flag' (establishment class), we get the employment size class for data subject to confidentiality[18]. The average of both information is compounded in one variable. A proxy for employment is so set up whenever its actual value is withheld in accordance with the U.S. Code, Title 13, Section 9, on data confidentiality.
The Data Suppression Flag denotes employment size class for data withheld to avoid disclosure according to table [10], whereas the number of establishments for each employment class is shown in table [11].

is supposed that those firms engage on average 2.5 employees, and that the number of employees within that sector is $2.5 * N$.

[18]So, for example, it is known whether within a given sector there are 0 to 19 workers or between $50,000$ and $99,999$, and assume the median to be a good indicator from the size class.