Papers in Evolutionary Economic Geography

10.12

Related variety and regional growth in Spain

R. Boschma and A. Minondo and M. Navarro

Utrecht University
Urban & Regional research centre Utrecht

http://econ.geo.uu.nl/peeg/peeg.html

Related variety and regional growth in Spain

R. Boschma*, A. Minondo** and M. Navarro***

* Urban and Regional Research Center Utrecht, Department of Economic Geography, Utrecht University, Utrecht (Netherlands)
** Deusto Business School, San Sebastian (Spain)
*** Orkestra-Basque Institute of Competitiveness and Deusto Business School, San Sebastian (Spain)

version 24 October 2010

Abstract

This paper investigates whether related variety, among other types of spatial externalities, affected regional growth in Spain at the NUTS 3 level during the period 1995-2007. We found evidence that related variety matters for growth across regions, especially when measured with the assistance of the Porter's cluster classification and the proximity index proposed by Hidalgo et al.. That is, Spanish provinces with a range of industries that are technologically related tend to show higher economic growth rates, controlling for the usual suspects. We did not find, however, any evidence of regional growth effects that come from technologically related sectors imports.

JEL Codes: D62, O18, R11
Keywords: related variety, Porter's cluster classification, product-proximity, regional growth, Spain

1. Introduction

There is a growing body of literature stating that variety is beneficial for economic growth (Saviotti, 1996). Regional scientists have taken up this point and incorporated the virtues of variety in their regional growth models (Glaeser et al., 1992). Inspired by the seminal work of Jane Jacobs (1969), they argue that not necessarily cities *per se*, but cities with a diversified set of industries will be characterized by high economic growth, because local diversity will spark creativity, new ideas and innovations. Having said that, some have argued that this concept of Jacobs' externalities needs to be refined and specified more precisely (Porter; 2003; Frenken et al., 2007). In particular, they claim that an urban structure that consists of a wide set of technologically related industries is more conducive for urban and regional growth. This in line with an expanding literature that suggests that technological relatedness is a major asset not only for economic growth in regions in general, but also for the process of regional diversification (Neffke et al., 2009; Boschma and Frenken, 2010).

This paper has three objectives. The first objective is to replicate studies on the regional effects of related variety conducted in countries like the Netherlands (Frenken et al., 2007), Italy (Boschma and Iammarino, 2009) and Great Britain (Bishop and Gripaios, 2010). We test empirically whether related variety, among other factors, matters for economic growth in 50 Spanish regions covering the period 1995-2007. The second objective is to test the effect of related variety by using three different criteria to establish relatedness across varieties. The first criteria, proposed by Frenken et al. (2007), and which has become conventional in subsequent studies, is to establish relatedness based on standard classifications of industries or products, such as the Standard Industrial Classification (SIC) or the Harmonized System (HS). The second criteria follows Porter's cluster classification and defines related industries on the basis of the geographical correlation of employment across traded industries (Porter, 2003). The third criteria rests on the products' proximity index developed by Hidalgo et al. (2007), which is based on the probability that a country develops comparative advantage in two products. The third objective is to assess the effects of extra-regional knowledge flows on economic growth of Spanish regions based on trade data, by making use of new methodology that has been developed by Boschma and Iammarino (2009). Among others, we tested whether related import flows (that is, imports that originate from sectors that are related to export sectors in the region) had an additional effect on regional growth in Spain.

The structure of the paper looks as follows. In Section 2, we explain the main theoretical ideas behind related variety, and discuss a number of existing empirical studies on this topic. In Section 3, we introduce the methodology used. Section 4 presents the main findings, followed by some concluding remarks in the final section.

2. Spatial externalities and related variety

In the spatial externalities and regional growth literature, a key question is whether firms in cities learn principally from other local firms in the same sector, or from other local firms in a range of other sectors (Glaeser et al., 1992; Feldman and Audretsch, 1999). The former form of spatial externalities is known as localization economies, and dates back to the work of Alfred Marshall in the late nineteenth century (Marshall, 1890; Asheim and Gertler, 2005; Potter and Watts, 2010). Firms in specialized regions would benefit from local externalities due to the presence of specialized input suppliers, a local pool of specialized labor skills, and specialized knowledge concerning the secrets of the respective trade. The latter form of spatial externalities has been associated with Jacobs' externalities, and builds on the seminal work of Jane Jacobs developed in the 1960s (Jacobs, 1969; Lambooy, 1984; Becattini et al., 1996; Van Oort, 2004). A diversified economy would bring benefits to local firms because it would trigger and generate new thinking, new ideas and innovations.

Since Glaeser et al. (1992), many regional scientists have embarked on this type of research. Despite all their efforts, this literature has led to inconclusive results so far with respect to the question whether localization economies and Jacobs' externalities matter for urban and regional growth, as concluded in a recently held meta-analysis (De Groot et al., 2009; see also Beaudry and Schiffauerova, 2009). Among other reasons, one plausible reason for this inconclusive finding is the potential misspecification of the notion of Jacobs' externalities. For example, one can seriously question whether knowledge will spill over across industries in diversified regions just because industries are each other's neighbors.

In the 1980s and 1990s, there was a lot of focus on the degree of relatedness between technologies used in sectors because this might affect the scope of knowledge spillovers (e.g. Rosenberg and Frischtak, 1983; Carlsson and Stankiewicz, 1991; Breshnahan and Trajtenberg 1995). As Nooteboom (2000) put it, knowledge is more likely to spill over across two industries when their cognitive distance is not too large, nor too small. That is, some degree of cognitive proximity between the two sectors ensures effective communication and common understanding, and some degree of cognitive distance is needed to avoid cognitive lock-in.

When applying this concept of industry relatedness to the spatial externalities literature, one may expect that the extent to which the variety of technologies present in a region is related will positively affect the scope for knowledge spillovers and learning, as local firms in different but related activities can profit more from mutual spillovers than local firms in unrelated industries. Porter (2003) made the claim that the distinction between localization economies and Jacobs' externalities is therefore too simple, because it focuses too much on the industry itself. Instead, there is a need to emphasize the importance of externalities among related industries, which Porter linked to his concept of clusters as geographic concentrations of linked industries. As Porter (2003) put it, "clusters are important because of the externalities that connect the constituent industries, such as common technologies, skills, knowledge and purchased inputs" (p. 562). According to Porter, specialization in clusters of related industries, not in industries *per se*, should lead to better regional performance. And next to having the benefits of clusters of related industries in a region, he argued that a range of overlapping clusters (caused by related industries that belong to more than one cluster) may be more beneficial for regional growth than having a diversity of clusters that are unrelated.

Frenken et al. (2007) incorporated this industry relatedness effect more explicitly in the spatial externalities and regional growth literature. They stated that the notion of Jacobs' externalities grasp two variety effects (i.e. related and unrelated variety) at the same time, and should therefore be disentangled. The related variety effect includes externalities that may come from a diversity of related industries in a region. The notion of regional related variety tries to capture a delicate balance between cognitive proximity and distance across sectors in a region that is needed for knowledge to spill over effectively between sectors. Thus, the more variety across related sectors in a region, the higher the number of technologically related sectors, the more learning opportunities there are for local industries, the more inter-sectoral knowledge spillovers are likely to take place, and the higher regional performance. This stands in contrast to localization economies in which regional specialization produces too much cognitive proximity between local firms (lock-in), while Jacobs' externalities *per se* may involve too much cognitive distance between local firms active in different industries.

The unrelated variety effect captures a portfolio-effect, which functions as a regional shock absorber (Essletzbichler, 2007). That is, when a region has a large number of unrelated industries, it may not be too vulnerable to sector-specific shocks. For instance, when an industry is affected by a sharp fall in demand, the workers that become redundant may find easily jobs in other local sectors that are unrelated and therefore will not be seriously damaged by this shock. Although both the related and unrelated variety effects are potential blessings for diversified regions, this latter effect is quite different from the related variety effect, and therefore both effects normally associated with Jacobs' externalities should be empirically separated from each other.

Empirical studies have investigated the significance of related variety for regional growth in the Netherlands (Frenken et al., 2007), Great Britain (Bishop and Gripaios, 2010) and Italy (Boschma and Iammarino, 2009). These studies found quite strong empirical evidence for the importance of related variety for regional growth. This was, however, less true for the unrelated variety effect. Nevertheless, Frenken et al. (2007) found that Dutch regions with a high degree of unrelated variety performed better in terms of unemployment rates. As expected, unrelated variety dampened regional unemployment growth.

These growth studies have not looked into possible externality effects that come from technologically related sectors in other regions. These studies predefine regions at a particular spatial scale, and they do not allow inter-industry externalities to spill over to other regions. Boschma and Iammarino (2009) have made an attempt to estimate externalities effects between related industries across regions on regional performance in Italy. Based on regional trade data, they found a positive correlation between employment growth of Italian regions from 1995 till 2003 on the one hand, and the degree of relatedness between export and import sectors in Italian regions on the other hand. This outcome might indicate that a region benefits from extra-regional knowledge when it originates from (import) sectors that are related or close, but not quite similar to existing export sectors in the region. This study also found evidence that trade similarity (i.e. regions with a great deal of overlap between their export and import sectors) was negatively correlated with regional growth in Italy. However, this issue of economic effects of related flows within and across regions on regional growth is still relatively unexplored. Due to limited data availability at the NUTS-3 regional level, the Italian study could not control for other factors that are normally accounted for in a conventional regional growth model, but in the present study on Spain, we can.

3. Methodology

To measure the effect of various types of agglomeration economies, we make use of export data (see also Boschma and Iammarino, 2009). Obviously, not all industries are export sectors, so the export profile of a region may not fully reflect the industrial composition of a region. In our measurements, there will be some bias toward manufacturing activities, due to, among others, the relatively low tradability of most service industries. Having said that, in manufacturing industries, knowledge complementarities between sectors can be approximated by export structures of regions, since industries that are most open to international competition are also those that contribute most to new knowledge, innovation and economic growth (e.g., Dosi, 1988; Fagerberg, 1988). As exporting occurs in almost all manufacturing industries, and export industries are among the strongest in a region, we expect the effects of related and unrelated variety to matter most among export sectors. In addition, we make use of import data at the regional level, in order to account for the effects of extra-regional linkages in terms of (export) outflows and (import) inflows by industry.

Three ways to define relatedness between industries

To estimate the effects of related variety within and across regions, we first set out how these have been measured in our study. A major challenge is to determine the degree of (technological) relatedness between industries. As said in the introduction, we will assess the effect of related variety in three different ways.

First, we follow Frenken et al. (2007) by taking a standard product classification as a starting point. These authors have used the Standard Industrial Classification (SIC), and have defined 5-digit industries to be technologically related when these share the same 2-digit class. On the one hand, these industries are perceived to show some degree of cognitive proximity, because these 5-digit sectors (e.g. sub-branches in chemicals) will share some technology and product characteristics in the same 2-digit class (e.g. chemicals). On the other hand, these industries are considered to show some degree of cognitive distance, because these sectors differ at the 5-digit level. Then, the more variety there is at the 5-digit level within each 2-digit industry in a region, the more related variety and thus real learning opportunities are available in a region, the more a region might benefit from externalities from such a wide set of different but related industries. The Frenken et al. study measured unrelated variety as the degree of variety of industries at the 1-digit level in a region. This is because at the 1-digit level, industries are unlikely to have much in common with respect to technology and product characteristics. Consequently, this indicator grasps the portfolio effect of variety explained earlier.

In contrast to Frenken et al. (2007), our database uses the Harmonized System 6-digit classification which obliges us to introduce some minor changes in the definition of related and unrelated varieties. In particular, we assume that 6-digit export sectors that share the same 2-digit class have some but not too much cognitive proximity and, hence, consider them related varieties. We also assume that 1-digit export sectors are not close in technology or other product characteristics and consider them as unrelated varieties.

Second, we follow Porter (1998) and use a cluster classification to determine relatedness across products. Porter defines clusters as geographic concentrations of linked industries that encompass producers, suppliers and providers of specialized services that generate (knowledge) externalities to local firms. In his US study, he used the local correlation of employment across traded industries at the US state level to define clusters of related industries. As Porter (2003) put it, "if computer hardware employment is nearly always associated geographically with software employment, this provides a strong indication of locational linkages" (p. 562). After applying this basic rule, Porter did basically two things to eliminate cases of spurious correlation: (1) he left out those cases in which no 'logical' externality was to be expected between two industries; (2) he excluded those cases that did not have any substantial input-output flows. Following this procedure, Porter identified 41 different clusters in the US, with an average of 29 industries each. Recall that these clusters are not spatial entities (here is where the confusion about Porter's cluster concept often comes in), but are defined as a set of related industries based on their frequent co-occurrence at the US state level. This means that while some clusters are strongly concentrated in a few regions, other clusters are quite dispersed across regions.

This industry relatedness indicator, as defined by clusters of related industries, provides an advantage, as compared to the previous related variety measure. In Porter's definition, a cluster may consist of any set of industries, and may include both manufacturing and service industries, a possibility which is ruled out by definition in our first indicator that is based on standard classifications. As Porter (2003) describes himself, "clusters, then, represent a different way of dividing the economy that is embodied in conventional industrial classification systems that are based primarily on product type and similarities in production" (p. 563).

In order to construct our second related variety measure for Spanish regions, we made use of Porter's cluster classification, as outlined in his 2003 study in Table 3 on page 563. Based on the correspondence table provided by the Institute for Strategy and Competitiveness of Harvard University, we established a link between our HS 6-digit industry classification and 36 different Porter clusters (Harvard University).[1] In our Spanish case, each cluster has, on average, 345 different HS-6 digit products.

The third relatedness measure is based on the proximity indicator developed by Hidalgo et al. (2007). These authors argue that several dimensions may influence the degree of relatedness between two products: similarities in the combination of productive factors, the characteristics of the technology used in production, the use of a specific component, the features of the final customers or the use of specific distribution channels. Due to the myriad of factors that may determine relatedness between products, they use an outcome measure to calculate the degree of proximity between products. They argue that two products will be close to each other if countries tend to have revealed comparative advantage in both products. Based on this idea they calculate proximity (φ) between product *i* and product *j* at year *t* as:

$$\varphi_{ij\,t} = \min\left\{P(x_{i,t} \mid x_{j,t}), P(x_{j,t} \mid x_{i,t})\right\} \qquad (1)$$

where $P(x_{i,t} \mid x_{j,t})$ is the conditional probability of having revealed comparative advantage in product *i* given that the country has revealed comparative advantage in product *j*.

Based on this index and using network displaying techniques, Hidalgo et al. (2007) are able to draw a product space map. This map shows that products are not evenly distributed: there are sections of the map with a high density of products, whereas other sections of the map are sparsely populated. Our argument is that these discontinuities in the product map are very important to determine learning opportunities. If a country specializes in products that are close to other products, knowledge opportunities will be larger. In contrast, if a country specializes in products that are far from each other, learning opportunities will be scant.

[1] Table A1 in the appendix presents the list of clusters included in our analysis.

We calculated the proximity measures across products using a sample of 102 countries for the years 2004 and 2005 from the UN Comtrade database. We calculated the proximity for the 1,244 products that compose the HS 4-digit 2002 Classification.[2] Figure 1 presents the histogram of the proximity measure calculated with our sample.[3] As shown in the figure, the proximity index follows a bi-modal distribution: there are high frequencies at zero and 0.15. The difficulty with this index is to determine what level of proximity is needed to consider two products as related. We have taken a conservative position and have considered that two products are related if their proximity is equal or above 0.25. There are 107,275 product-pairs (14% of all product-pair combinations) that meet this criteria.

In addition to these indices, some authors have proposed other measures to establish relatedness across products or industries. For example, Neffke et al. (2009), following a similar idea to Hidalgo et al. (2007), developed a relatedness measure based on the frequency that two products are produced jointly at a plant level. Neffke and Henning (2009) calculated a relatedness index based on the intensity of labor flows between industries. In our Spanish study, we were not able to calculate those indices due to a lack of required data for Spain.

Variety indexes

We made use of entropy measures to calculate the different variety indexes. To calculate the related variety index, firstly, we grouped 6-digit HS products into related variety sets: S_r. As explained above, in the first conventional measure, a related variety set is composed of those 6-digit HS products that belong to the same 2-digit HS products' class. In the Porter measure, a related variety set is composed of those 6-digit HS products that belong to the same cluster. Finally, in the proximity indicator, we define a related variety set for each product, which is composed of the rest of HS 6-digit products that have, at least, a proximity equal to 0.25. We calculated the share of the 6-digit HS i product in total regional exports (p_i) and the share of each related variety set in total regional exports (P_r). With these shares we calculate the entropy within the related variety set (H_r) as follows:

[2] The Comtrade database allows the calculation of proximity indexes at the 6-digit HS level. However, there have been some changes in the HS classification at this disaggregation level. As we have to match these data with the Spanish regional data, we opt to calculate the proximity measures at the 4-digit level to minimize the loss of the data due to changes in classification.

[3] The proximity matrix encompasses 773,143 indexes: (1,244 products x 1,243 products)/2.

$$H_r = \sum_{i \in S_g} \frac{p_i}{P_r} \log_2 \left(\frac{1}{p_i / P_r} \right). \quad (2)$$

Related variety is calculated as the exports-weighted entropy in each related variety set.

$$RELATEDVARIETY = \sum_{r=1}^{R} P_r H_r \quad (3)$$

Unrelated variety is calculated in a similar way. Now, we have to define the unrelated variety sets for each measure. For the conventional and the Porter measure, there is only one unrelated variety set. In the first case, the unrelated variety set is composed of each 1-digit HS industry, while in the second case, it is composed of each cluster. For these two measures, the unrelated variety index is calculated as follows

$$UNRELATEDVARIETY = \sum_{j=1}^{N} P_j \log_2 \left(\frac{1}{P_j} \right) \quad (4)$$

where P_j denotes the share of each 1-digit sector or cluster in total exports.

For the proximity measure, there is one unrelated variety set for each product, which is composed of the rest of products whose proximity to the analyzed product is below 0.25. As in the related variety set, we calculate entropy using equation (2), and then calculate unrelated variety as the weighted sum of entropy at each unrelated variety set using equation (3).

To analyze whether disentangled variety measures better identify learning opportunities than conventional variety measures, we also calculated a Jacobs' externalities or variety measure. The variety index is calculated as follows:

$$VARIETY = \sum_{i=1}^{N} p_i \log_2 \left(\frac{1}{p_i} \right) \quad (5)$$

where p_i stands for the share of 6-digit HS product i in total regional exports

As set out in Section 2, we also aimed to estimate the effects of possible externalities that spill over from related industries in other regions. Following Boschma and Iammarino (2009), our intention was not to account for pure spatial autocorrelation, but we wanted to estimate real industry flows between regions. A drawback of using trade data is that inter-regional spillovers within Spain cannot be accounted for, since these data do not exist at such a detailed regional level. We used the same trade variety indicators as developed by Boschma and Iammarino (2009).

As mentioned earlier, we expect a region to benefit particularly from extra-regional flows when they originate from sectors that are related, but not identical to those present in the region. Boschma and Iammarino (2009) developed an indicator, denominated as related trade variety, that proxies the possible benefits that export sectors in a region can derive from learning opportunities in related import sectors. To calculate this indicator, first, for each 6-digit HS product in a region, we measured the entropy of related imports. Related imports are identified using the same criteria explained before. Once we calculate entropy in related imports, we multiply it by the relative size of the 6-digit export industry. We repeat this calculation for all 6-digit export industries in the region and weight them by their share in total exports. The logic behind this indicator is that the higher the variety in related (but not similar) imports, the more learning opportunities, and the larger the relative size of the respective export sectors in the regional economy, the more these learning opportunities may be transformed into regional growth.

Analytically related trade variety is calculated as follows

$$RELATED\,TRADE\,VARIETY = \sum_i OE_i^M * p_i \qquad (6)$$

where OE_i^M is entropy in the import sectors related to the 6-digit HS product *i*.[4]

[4] As before, in the conventional measure related imports are those 6-digit HS imports that belong to the same 2-digit class as product i (excluding product *i* imports); in the Porter measure those 6-digit HS imports that belong to product *i*'s cluster (excluding product *i* imports); finally, in the proximity measure those HS 6-digit imports with a proximity equal or higher than 0.25 to product *i*.

We also calculated an indicator that accounts for extra-regional inflows that come from import sectors in which the region exports. In that case, the region (or, better, the export sector) can absorb the external knowledge, but the new knowledge will not add substantially to the existing knowledge base of the region. As a result, we expect no additional effect on regional growth. We constructed a trade similarity variable by means of the sum of the products of the absolute sizes of the 6-digit HS exports and imports in each region. This indicator gets its maximum value when a region is specialized in only one and the same sector both in imports and exports, while the value gets lower the more diversified a region is (in both imports and exports), and the less similar the import and export profiles of the region look like. Analytically:

$$TRADE\ SIMILARITY = \log\left(\sum_i X_i * M_i\right) \tag{7}$$

Finally, we estimated the effects of extra-regional linkages that bring a high degree of variety into a region through a diversified set of import industries. We assumed that the wider the spectrum of import industries, the more diversified the knowledge flows that enter a region through its trade linkages. We measured the degree of import variety in each region by means of an entropy measure at the 6-digit level, where now p_i stands for the share of the 6-digit HS product i in total regional imports:

$$IMPORT\ VARIETY = \sum_{i=1}^{N} p_i \log_2\left(\frac{1}{p_i}\right) \tag{8}$$

The geographic unit used in our analysis are provinces, classified as NUTS-3 in Eurostat's regional classification. Spain is divided in 52 provinces. Due to their special circumstances we have excluded the two Spanish provinces that are located in Africa (Ceuta and Melilla).[5] Data on Spanish provinces exports and imports at the Harmonized System 6-digit level were obtained from the Spanish Dirección General de Aduanas - Agencia Tributaria database.

[5] Seven of the fifty provinces are uni-provincial autonomous communities. It is worth mentioning that, even though Spain is not formally a federal country, in fact it is very decentralized. The decentralization process has benefited the autonomous community level, but not the provincial and municipality levels.

4. Empirical findings

Following Boschma and Iammarino (2009), we assess the impact of related variety, among other factors, on three dependent variables: employment growth, value-added growth and labor-productivity growth. The period of analysis is 1995-2007.[6] We divide the period of analysis in four-year intervals.[7] Growth is measured as the average annual growth in the 4-year interval. All independent variables are measured at the initial year of the interval.

We present a table with the results of the empirical analyses for each dependent variable. The results are presented in four models. The first model includes the urbanization economies (population density) and the Jacobs' externalities (variety) coefficients. In the second model, the Jacobs' externality coefficient is divided into a related variety coefficient and an unrelated variety coefficient. The third model includes the import variety, while the fourth model introduces a related trade variety coefficient and a trade similarity coefficient.

Within each of these four models, we estimate three regressions. In the first regression, we use the pool of observations and simple OLS to analyze the impact of externalities on the dependent variable. In the second regression, we control for common variables, such as human capital, labor productivity, employment and capital-labor ratio that may influence regional performance.[8] As our data spans a long time-period (12 years), we can divide this period in intervals and estimate a fixed-effects model in the third regression. This model allows us to control for time invariant characteristics at the province level that may also influence regional growth. Tables 1a-1c present the outcomes in which the related and unrelated variety indicators have been obtained using the conventional measures, while Tables 2a-2c present the findings based on the Porter's cluster classification measures, and Tables 3a-3c on the proximity measures.[9] Table A2 in the appendix presents a correlation matrix for the independent variables used in the empirical analyses.

[6] Although the Spanish Statistical Institute (INE) provides data before 1995, these data are less reliable.
[7] The use of alternative intervals, such as 3-year or 2-year intervals, did not alter the results of the econometric analyses.
[8] In addition to these variables, we also introduced an additional variable developed in Minondo (2010) to control for the sophistication level of provinces' production. This variable did not alter the results.
[9] Data on employment, valued added and population of Spanish provinces come from the Spanish Statistical Institute's (INE) Regional Economic Accounts database. Human capital data are obtained from the Instituto Valenciano de Investigaciones Económicas (Ivie) database. Human capital is proxied by the percentage of occupied population that has upper-secondary or tertiary studies. Finally, data on provincial capital stock are obtained from Fundación BBVA. Capital is proxied by the stock of machinery and other products.

Table 1a presents the results with regional employment growth as the dependent variable and conventional related and unrelated variety measures. Urbanization economies, as proxied by population density, has a positive and statistically significant impact on regional employment growth, once we control for province-level time-invariant fixed effects. All variety variables are not statistically significant. These results are not in line with those obtained by Frenken et al. (2007) for the Netherlands, where related variety has a positive and statistically significant effect in all regressions. The same is true in case of the Italian study, which obtained a positive and statistically significant effect in Model 3-1 (Boschma and Iammarino, 2009). With respect to regional trade variety variables, we can conclude that the coefficient signs are not robust to changes in the model specification and they are not statistically significant. In relation to the control variables, we find that provinces with an initial high level of labor productivity show higher employment growth rates, while the opposite is true for initial levels of human capital. The coefficient on initial employment levels is negative and statistically significant when we control for province-level effects. Furthermore, we can observe that coefficients are not biased due to spatial autocorrelation (Moran I's is never statistically significant), and that there are strong province-specific effects.

Table 1b presents the results when value-added growth is selected as dependent variable. In this case, province-level effects seem to play a smaller role: the F-test statistic for the null hypothesis that all province effects are zero is very low. The coefficient of urbanization economies is not robust anymore, as it shifts from positive to negative when adding province-level effects. However, Jacobs' externalities coefficient has now a positive and statistically significant effect, which is highly persistent in all model specifications. When we divide variety into related and unrelated variety, we observe that it is related variety which has a positive and a statistically significant effect on regional value-added growth. These results are in line with those obtained by Boschma and Iammarino (2009). The sign of unrelated variety is not robust, and never statistically significant. With respect to the regional trade variety effects, our findings show that related trade variety has a positive coefficient (as expected), but not statistically significant. The coefficient of trade similarity is persistently negative, as expected, but not statistically significant. Finally, the control variables do quite poorly in these models. Only initial human capital levels in regions are negatively correlated with value-added growth, but only when province-specific effects are not controlled for.

Table 1c presents the findings with regional labor productivity growth as the dependent variable. As in the previous table, the coefficient for population density is not robust, and becomes negative and statistically significant when one controls for fixed effects. All variety variables, including related variety, are not statistically significant, and in most model specifications, the coefficient has a negative sign. These results tend to be in line with those obtained by Frenken et al. (2007). With respect to the regional trade variety variables, again, we did not find robust results. In relation to the control variables, the initial labor productivity level in a region has always a negative and statistically significant effect. This result points out that Spanish provinces are converging in terms of labor productivity. The sign of the capital-labor ratio and human capital coefficients are not robust to the model specifications. In both cases, the signs of the coefficients change from positive to negative once we control for province-level fixed effects.

The second set of tables (Table 2a-2c) presents the results of the analyses when related and unrelated variety, and related and unrelated trade variety are computed using the Porter's cluster classification. We exclude from the tables Models 1-1 to 1-3 as they do not change with respect to the first set of tables. Table 2a presents the results with regional employment growth as the dependent variable. In the fixed-effects models, urbanization economies has a positive and significant effect on regional employment growth, as in Table 1a. Our findings also show that related variety now has a positive and statistically significant effect when we include additional control variables but leave out province-level fixed effects. In the fixed-effects models, however, the coefficient of related variety turns negative, though not statistically significant. The coefficient sign of unrelated variety is not robust to the model specifications: it shifts from a negative to a positive sign and statistically significant coefficient when province-level fixed effects are controlled for. As previously, the regional trade variety variables do not show any significant effect on regional employment growth.

Table 2b presents the results when regional value-added growth is selected as dependent variable. The effect of urbanization economies is again not robust, as in Table 1b. The coefficient of related variety is always positive and statistically significant, except when province-level fixed effects are included. In this latter case, however, the F-test shows this is of less relevance, because it does not reject the null hypothesis that all province-level effects are zero. The coefficient for unrelated variety is, in most cases, negative. As expected, related trade variety has a positive coefficient and trade similarity has a negative coefficient, but both coefficients are not statistically significant.

Table 2c presents the results with regional labor productivity growth as the dependent variable. As opposite to the previous findings in Table 1c, where related variety had in most regressions a negative coefficient, we now obtain a positive coefficient in the majority of regressions, although not statistically significant. Unrelated variety shifts from positive to negative coefficients when we control for province-level fixed effects. Neither related trade variety and trade similarity have robust coefficients.

Finally, Tables 3a-3c present the results when we use the proximity index to build related and unrelated variety measures. Table 3a presents the results with regional employment growth as the dependent variable. As was the case in the previous tables, in the fixed-effects models, urbanization economies has a positive and significant effect on regional employment growth. In all specifications, we find a positive coefficient for related variety; moreover, in the majority of cases, the coefficient is statistically significant. As was the case with the cluster classification, the sign of unrelated variety is not robust to the model specifications. Trade variety coefficients are not either statistically significant. Table 3b presents the results when regional value-added growth is selected as dependent variable. The effect of urbanization economies is not robust to changes in the econometric model. The coefficient of related variety is always positive and in most cases statistically significant. The sign of the coefficient for unrelated variety and trade varieties are not robust to changes in model specifications. Finally, Table 3c presents the results with regional labor productivity growth as the dependent variable. The sign of all variety coefficients are not robust to changes in model specifications.

To sum up, the use of alternative measures to establish relatedness between industries enhance the number of cases in which we observe a positive and statistically significant relationship between related variety and regional economic growth. These results suggest that clusters and product proximity better identify the relatedness across industries than the conventional measures based on standard industry or product classification.

5. Conclusions

This study has investigated the importance of related variety for regional growth in Spain. One objective was to replicate recent studies on this same topic in other countries like the

Netherlands and Italy. Doing so, we tried to be as close as possible to the definitions used in other studies. Broadly speaking, our findings tend to confirm the positive effect of related variety on regional value-added growth found in the Italian study. However, the insignificant effect of related variety on regional employment growth we found for Spain (though the coefficient was often found positive) contradicts findings in both the Netherlands and Italy. The insignificant effect of related variety on regional labor productivity growth tends to confirm the results of the Dutch study, but is opposite to those obtained in the Italian study.

Of course, such a cross-country comparison is not unproblematic because of different definitions (like the use of more disaggregated data to calculate related variety measures), different periods that are covered, among other factors. Therefore, to make more precise inter-country comparisons, it might be useful to investigate the impact of related variety on the performance of some industries at the regional level, not on regional performance in general. This could also improve our estimations, because some have suggested that related variety may be important only for some industries, but not for others (Bishop and Gripaios, 2010). Since we expect that the growth of an industry depends on the local presence of industries that are technologically related to that industry, such an exercise would allow one to measure more directly the effect of related variety at the regional level.

Another objective was to measure the effect of related variety using alternative indices. In the conventional manner, one makes use of standard industrial or product classifications, and defines industries as related when they share the same digit class at a more aggregated level. However, this measure of relatedness between industries is not unproblematic. One reason is that the relatedness measure based on these classifications does not fully capture the degree of technological relatedness between industries (Breschi et al., 2003). In our study, first, we used an alternative measure that was proposed by Porter, in which industries are classified by means of their geographic correlation. Our findings show that the use of Porter's industrial classification tends to increase the number of cases in which we observed a positive relationship between related variety and regional growth. Second, we also calculated an additional related variety measure based on the product proximity concept developed recently by Hidalgo et al. (2007). With this measure, we obtain the largest number of cases in which there is a positive and statistically significant relationship between related variety and regional growth.

A final objective of our study was to account for the fact that new and related variety may also be brought from other regions through inter-sectoral linkages. None of our hypotheses could be confirmed however, regardless of the model specifications. In this study, we made use of regional trade data to assess possible spillover effects from technologically related sectors located outside the region. Of course, trade data only account for spillover effects from other countries, but leave out the effects of inter-regional flows within Spain. This can be gauged by labor flow data though, if available at the regional and industry level. Another advantage of inter-regional labor flows is that they better capture knowledge flows between industries and, therefore, may more accurately assess the spillover effects of technologically related industries between regions on regional growth (Boschma et al., 2009).

References

Asheim, B. and Gertler, M. (2005) The geography of innovation. Regional innovation systems, in J. Fagerberg, D. Mowery, and R. Nelson (eds.), *The Oxford Handbook of Innovation*, Oxford: Oxford University Press, pp. 291-317.

Beaudry, C. and A. Schiffauerova (2009). Who's right, Marshall or Jacobs? The localization versus urbanization debate, *Research Policy* 38 (2), 318-337.

Becattini, G., M. Bellandi, G. dei Ottati and F. Sforzi (2003), *From industrial districts to local development. An itinerary of research*, Cheltenham: Edward Elgar.

Bishop, P., Gripaios, P. (2010) Spatial Externalities, Relatedness and Sector Employment Growth in Great Britain. *Regional Studies*, 44,4: 443-454.

Boschma, R. and K. Frenken (2010), Technological relatedness and regional branching, in H. Bathelt, M.P. Feldman and D.F. Kogler (eds.), *Dynamic Geographies of Knowledge Creation and Innovation*, Routledge, Taylor and Francis, forthcoming.

Boschma, R. and Iammarino, S. (2009). "Related Variety, Trade Linkages, and Regional Growth in Italy", *Economic Geography*, **85**, 3, 289-311.

Boschma, R., R. Eriksson and U. Lindgren (2009), How does labor mobility affect the performance of plants? The importance of relatedness and geographical proximity. *Journal of Economic Geography* 9 (2): 169-190

Breschi, S., F. Lissoni and F. Malerba (2003) Knowledge-relatedness in firm technological diversification, *Research Policy* 32: 69-87.

Breshnahan, T.F. and M. Trajtenberg (1995) General purpose technologies. Engines of growth? *Journal of Econometrics* 65 (1): 83-109.

Carlsson, B. and R. Stankiewicz (1991) On the nature, function and composition of technological systems, *Journal of Evolutionary Economics* 1, 93-118.

De Groot, H.L.F., Poot, J. and Smit, M.J. (2009), Agglomeration externalities, innovation and regional growth: theoretical perspectives and meta-analysis. In: R. Capello and P.Nijkamp (eds) Handbook of Regional Growth and Development Theories. Northampton MA: Edward Elgar, pp. 256-281.

Dosi, G. (1988), Technical change and international trade. In *Technical change and economic theory,* ed. G. Dosi, C. Freeman, R. Nelson, G. Silverberg, and L. Soete, 401-431. London: Pinter.

Essletzbichler, J. (2007) Diversity, Stability and Regional Growth in the United States, 1975-2002. In Frenken, K., ed. Applied Evolutionary Economics and Economic Geography , 203-229. Edward Edgar, Cheltenham.

Fagerberg, J. (1988) International competitiveness. *Economic Journal* 98:355–74.

Feldman, M.P., D.B. Audretsch (1999), Innovation in cities. Science-based diversity, specialization and localized competition, *European Economic Review*, 43, 409-429.

Frenken, K., van Oort, F. G. and Verburg, T. (2007) Related variety, unrelated variety and regional economic growth", *Regional Studies*, **41**, 5, 685–97

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

Hidalgo, C.A., Klinger, B., Barabási, A.L. Hausmann, R. (2007) The Product Space

Conditions the Development of Nations, *Science*, **317**, 5837, 482-487.

Jacobs, J. (1969), *The Economy of Cities*. New York: Vintage.

Lambooy, J.G. (1984). The regional ecology of technological change. In: Lambooy, J.G. (Ed.). *New Spatial Dynamics and Economic Crisis*. Tampere: Finnpublishers, 63-77.

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

Minondo, A. (2010) Exports' productivity and growth across Spanish regions, *Regional Studies*, 44, 5, 568-577.

Neffke, F., Henning, M. and Boschma, R. (2009) How do regions diversify over time? Industry relatedness and the development of new growth paths in regions, Papers in Evolutionary Economic Geography, no. 9.16, Utrecht University.

Neffke, F., Henning, M. (2009) Skill relatedness and firm diversification, Papers on Economics and Evolution, no. 09.06, Jena, Max Planck Institute of Economics.

Nooteboom, B. (2000), *Learning and innovation in organizations and economies*, Oxford, Oxford University Press.

Porter, M.E. (1998), Clusters and the new economics of competition, *Harvard Business Review*, **76**, 6, 77-90.

Porter, M.E. (2003). “The Economic Performance of Regions”, *Regional Studies*, **37**, 6&7, 549-578.

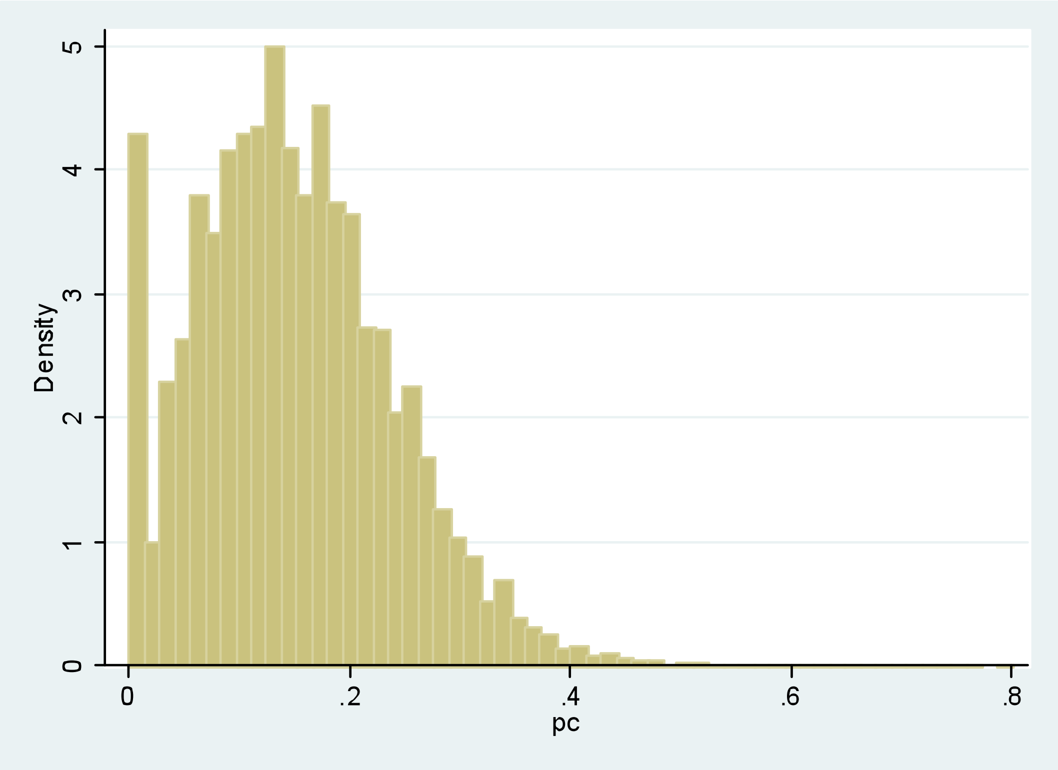
Potter, A. and D.G. Watts (2010) Evolutionary agglomeration theory. Increasing returns, diminishing returns, and the industry life cycle, *Journal of Economic Geography*, doi: 10.1093/jeg/lbq004

Rosenberg, N. and R. Frischtak (1983), Long waves and economic growth. A critical appraisal, *American Economic Review*, vol. 73, no. 2, pp. 146-151.

Saviotti, P. P. 1996. *Technological evolution, variety and the economy*. Cheltenham, U.K., and Brookfield, Vt.: Edward Elgar.

Van Oort, F. G. (2004), *Urban Growth and Innovation. Spatially Bounded Externalities in the Netherlands*. Aldershort: Ashgate.

Figure 1. Product proximity. Histogram



Source: authors' calculations based on Comtrade database.

Table 1a. Dependent variable employment growth, 1995-2007 (4-year intervals). Conventional relatedness measures

	Model 1-1	Model 1-2	Model 1-3	Model 2-1	Model 2-2	Model 2-3	Model 3-1	Model 3-2	Model 3-3	Model 4-1	Model 4-2	Model 4-3
Pop. density (Log)	0.00373**	-0.000909	0.165***	0.00321	-0.00129	0.172***	0.00348	-0.00129	0.179***	0.00229	-0.00149	0.180***
	(0.00174)	(0.00351)	(0.0344)	(0.00219)	(0.00351)	(0.0358)	(0.00218)	(0.00350)	(0.0374)	(0.00257)	(0.00353)	(0.0373)
Variety	0.00128	0.00168	0.00133									
	(0.00138)	(0.00130)	(0.00291)									
Related variety				0.00280	0.00337	-0.00173	0.00334	0.00343	-0.00224	0.000234	0.00453	-0.00258
				(0.00287)	(0.00326)	(0.00509)	(0.00343)	(0.00333)	(0.00464)	(0.00473)	(0.00480)	(0.00521)
Unrelated variety				-0.000869	0.000440	0.00313	-0.000369	0.000521	0.00170	-0.000308	0.000211	0.00222
				(0.00293)	(0.00360)	(0.00693)	(0.00338)	(0.00388)	(0.00747)	(0.00379)	(0.00398)	(0.00727)
Import variety							-0.000877	-0.000140	0.00186			
							(0.00204)	(0.00160)	(0.00214)			
Rel. trade variety										0.00226	-0.00151	0.00178
										(0.00423)	(0.00391)	(0.00295)
Trade similarity										0.000574	-0.000295	-0.00112
										(0.00106)	(0.00103)	(0.00139)
Lab. prod. (Log)		0.132***	0.0274		0.133***	0.0261		0.133***	0.0275		0.136***	0.0324
		(0.0313)	(0.0586)		(0.0317)	(0.0571)		(0.0317)	(0.0579)		(0.0343)	(0.0605)
Emp. (Log)		0.00292	-0.252***		0.00281	-0.255***		0.00287	-0.255***		0.00379	-0.252***
		(0.00414)	(0.0511)		(0.00436)	(0.0502)		(0.00442)	(0.0508)		(0.00467)	(0.0533)
Human cap. (Log)		-0.0409***	-0.000287		-0.0428***	0.00334		-0.0425***	0.00305		-0.0407***	0.00326
		(0.0122)	(0.0201)		(0.0148)	(0.0206)		(0.0151)	(0.0209)		(0.0132)	(0.0208)
Constant	0.0118	-1.216***	0.384	0.0174	-1.209***	0.407	0.0198*	-1.208***	0.379	-0.00662	-1.236***	0.365
	(0.00880)	(0.307)	(0.805)	(0.0106)	(0.309)	(0.782)	(0.0117)	(0.309)	(0.806)	(0.0459)	(0.327)	(0.823)
Observations	150	150	150	150	150	150	150	150	150	150	150	150
R-squared	0.074	0.303	0.800	0.076	0.304	0.801	0.078	0.304	0.803	0.082	0.305	0.803
Model	Pooled OLS	Pooled OLS	Fixed-effects	Pooled OLS	Pooled OLS	Fixed-effects	Pooled OLS	Pooled OLS	Fixed-effects	Pooled OLS	Pooled OLS	Fixed-effects
Moran's I p-value	0.411	0.410	0.464	0.400	0.402	0.712	0.394	0.398	0.529	0.391	0.409	0.532
F test for province-level effects			8.79			8.70			8.72			8.59

Note: Robust standard errors in parentheses. All regressions include period dummies. *** p<0.01, ** p<0.05, * p<0.1

Table 1b. Dependent variable value-added growth, 1995-2007 (4-year intervals). Conventional relatedness measures

	Model 1-1	Model 1-2	Model 1-3	Model 2-1	Model 2-2	Model 2-3	Model 3-1	Model 3-2	Model 3-3	Model 4-1	Model 4-2	Model 4-3
Pop. density (Log)	0.00204**	0.00192*	-0.0874***	0.00152	0.00125	-0.101***	0.00171	0.00133	-0.0947***	0.00202*	0.00161	-0.0799**
	(0.000824)	(0.000956)	(0.0303)	(0.00104)	(0.00110)	(0.0284)	(0.00107)	(0.00114)	(0.0316)	(0.00120)	(0.00124)	(0.0370)
Variety	0.00127*	0.00160**	0.00681**									
	(0.000683)	(0.000674)	(0.00291)									
Related variety				0.00271*	0.00376**	0.0117*	0.00308*	0.00387**	0.0113*	0.00293	0.00355	0.0104*
				(0.00147)	(0.00185)	(0.00620)	(0.00169)	(0.00192)	(0.00599)	(0.00261)	(0.00260)	(0.00585)
Unrelated variety				0.000136	9.80e-06	0.00138	0.000479	0.000152	0.000287	-0.000263	-0.000287	-0.000618
				(0.00229)	(0.00233)	(0.00524)	(0.00234)	(0.00244)	(0.00554)	(0.00258)	(0.00253)	(0.00585)
Import variety							-0.000602	-0.000234	0.00142			
							(0.000955)	(0.000961)	(0.00193)			
Rel. trade variety										0.000427	0.000767	0.00192
										(0.00198)	(0.00209)	(0.00384)
Trade similarity										-0.000422	-0.000348	-0.00237
										(0.000604)	(0.000588)	(0.00203)
Lab. prod. (Log)		0.0203	0.0265		0.0211	0.0255		0.0207	0.0263		0.0214	0.0309
		(0.0133)	(0.0245)		(0.0135)	(0.0249)		(0.0138)	(0.0252)		(0.0144)	(0.0251)
Human cap. (Log)		-0.0131**	-0.00180		-0.0159*	-0.00617		-0.0154*	-0.00640		-0.0154*	-0.00599
		(0.00650)	(0.0172)		(0.00808)	(0.0177)		(0.00839)	(0.0178)		(0.00805)	(0.0176)
Constant	0.0160***	-0.150	0.0920	0.0196**	-0.144	0.186	0.0213**	-0.141	0.140	0.0358	-0.136	0.154
	(0.00474)	(0.128)	(0.269)	(0.00740)	(0.129)	(0.283)	(0.00799)	(0.131)	(0.291)	(0.0277)	(0.137)	(0.286)
Observations	150	150	150	150	150	150	150	150	150	150	150	150
R-squared	0.105	0.126	0.190	0.110	0.186	0.188	0.113	0.137	0.191	0.115	0.140	0.209
Model	Pooled OLS	Pooled OLS	Fixed-effects	Pooled OLS	Pooled OLS	Fixed-effects	Pooled OLS	Pooled OLS	Fixed-effects	Pooled OLS	Pooled OLS	Fixed-effects
Moran's I p-value	0.597	0.532	0.783	0.582	0.527	0.770	0.579	0.523	0.769	0.567	0.524	0.782
F test for all province-level effects			1.45			1.39			1.39			1.40

Note: Robust standard errors in parentheses. All regressions include period dummies. *** p<0.01, ** p<0.05, * p<0.1

Table 1c. Dependent variable productivity-productivity growth, 1995-2007 (4-year intervals). Conventional relatedness measures

	Model 1-1	Model 1-2	Model 1-3	Model 2-1	Model 2-2	Model 2-3	Model 3-1	Model 3-2	Model 3-3	Model 4-1	Model 4-2	Model 4-3
Pop. density (Log)	-0.00175	0.00280*	-0.0661**	-0.00170	0.00271*	-0.0733**	-0.00180	0.00316*	-0.0717**	-0.000223	0.00304*	-0.0768**
	(0.00133)	(0.00150)	(0.0321)	(0.00155)	(0.00155)	(0.0325)	(0.00162)	(0.00162)	(0.0333)	(0.00167)	(0.00160)	(0.0368)
Variety	-3.10e-05	-0.000756	-0.000258									
	(0.00101)	(0.000792)	(0.00207)									
Related variety				-0.000186	-0.000822	0.00372	-0.000373	-0.000404	0.00360	0.00271	-0.000974	0.00390
				(0.00196)	(0.00149)	(0.00398)	(0.00219)	(0.00149)	(0.00409)	(0.00277)	(0.00256)	(0.00420)
Unrelated variety				0.00103	-0.000843	-0.00289	0.000857	-0.000259	-0.00311	2.67e-05	-0.00113	-0.00261
				(0.00131)	(0.00116)	(0.00563)	(0.00169)	(0.00130)	(0.00608)	(0.00168)	(0.00133)	(0.00588)
Import variety							0.000305	-0.000985	0.000331			
							(0.00136)	(0.000877)	(0.00152)			
Rel. trade variety										-0.00189	0.000708	-0.000227
										(0.00293)	(0.00247)	(0.00325)
Trade similarity										-0.00104*	-0.000346	0.000388
										(0.000622)	(0.000484)	(0.00108)
Lab. prod. (Log)		-0.122***	-0.242***		-0.122***	-0.245***		-0.124***	-0.245***		-0.121***	-0.247***
		(0.0225)	(0.0228)		(0.0228)	(0.0239)		(0.0230)	(0.0240)		(0.0233)	(0.0243)
Human cap. (Log)		0.0217***	-0.00474		0.0213***	-0.00909		0.0228***	-0.00912		0.0219***	-0.00915
		(0.00777)	(0.0161)		(0.00787)	(0.0165)		(0.00772)	(0.0164)		(0.00788)	(0.0166)
Cap./labor (Log)		0.0135***	-0.0207***		0.0134***	-0.0189**		0.0144***	-0.0192**		0.0132***	-0.0186**
		(0.00373)	(0.00703)		(0.00377)	(0.00807)		(0.00395)	(0.00828)		(0.00381)	(0.00813)
Constant	0.00524	1.051***	2.994***	0.00302	1.051***	3.068***	0.00218	1.066***	3.058***	0.0451*	1.058***	3.073***
	(0.00608)	(0.202)	(0.276)	(0.00533)	(0.205)	(0.285)	(0.00578)	(0.206)	(0.288)	(0.0246)	(0.209)	(0.286)
Observations	150	150	150	150	150	150	150	150	150	150	150	150
R-squared	0.059	0.444	0.756	0.060	0.443	0.760	0.061	0.447	0.760	0.086	0.445	0.761
Model	Pooled OLS	Pooled OLS	Fixed-effects	Pooled OLS	Pooled OLS	Fixed-effects	Pooled OLS	Pooled OLS	Fixed-effects	Pooled OLS	Pooled OLS	Fixed-effects
Moran's I p-value	0.383	0.400	0.625	0.372	0.397	0.615	0.368	0.402	0.607	0.374	0.392	0.633
F test for all province-level effects			4.30			4.33			4.24			4.22

Note: Robust standard errors in parentheses. All regressions include period dummies. *** p<0.01, ** p<0.05, * p<0.1

Table 2a. Dependent variable employment growth, 1995-2007 (4-year intervals). Porter's classification

	Model 2-1	Model 2-2	Model 2-3	Model 3-1	Model 3-2	Model 3-3	Model 4-1	Model 4-2	Model 4-3
Population density (Log)	0.00464**	-0.000556	0.148***	0.00473**	-0.000553	0.154***	0.00327	-0.000891	0.159***
	(0.00221)	(0.00332)	(0.0298)	(0.00224)	(0.00334)	(0.0314)	(0.00253)	(0.00338)	(0.0322)
Related variety	0.00362	0.00931***	-0.00443	0.00386	0.00926***	-0.00471	0.00237	0.00855**	-0.00495
	(0.00318)	(0.00337)	(0.00395)	(0.00337)	(0.00336)	(0.00388)	(0.00343)	(0.00380)	(0.00387)
Unrelated variety	-1.81e-05	-5e-05**	0.00013***	-1.55e-05	-5e-05**	0.00013***	-3.37e-05	-6.0e-05***	0.00014***
	(1.87e-05)	(1.94e-05)	(3.84e-05)	(1.89e-05)	(2.12e-05)	(3.97e-05)	(2.11e-05)	(1.92e-05)	(4.19e-05)
Import variety				-0.000701	0.000236	0.00125			
				(0.00192)	(0.00150)	(0.00166)			
Related trade variety							0.0113	0.00673	-0.00791
							(0.00804)	(0.00881)	(0.00817)
Trade similarity							0.000336	-0.000177	-0.00159
							(0.000939)	(0.000933)	(0.00115)
Labor productivity (Log)		0.155***	0.00982		0.155***	0.0112		0.153***	0.0150
		(0.0321)	(0.0558)		(0.0323)	(0.0567)		(0.0331)	(0.0570)
Employment (Log)		0.00425	-0.260***		0.00420	-0.260***		0.00453	-0.258***
		(0.00419)	(0.0454)		(0.00421)	(0.0461)		(0.00459)	(0.0473)
Human capital (Log)		-0.0508***	0.00711		-0.0512***	0.00730		-0.0522***	0.0119
		(0.0145)	(0.0206)		(0.0147)	(0.0207)		(0.0138)	(0.0205)
Constant	0.00891	-1.427***	0.653	0.0120	-1.430***	0.606	-0.000870	-1.397***	0.596
	(0.0104)	(0.304)	(0.740)	(0.0130)	(0.307)	(0.763)	(0.0383)	(0.311)	(0.758)
Observations	150	150	150	150	150	150	150	150	150
R-squared	0.079	0.359	0.815	0.080	0.359	0.816	0.100	0.363	0.820
Model	Pooled OLS	Pooled OLS	Fixed-effects	Pooled OLS	Pooled OLS	Fixed-effects	Pooled OLS	Pooled OLS	Fixed-effects
Moran's I p-value	0.380	0.419	0.445	0.375	0.416	0.476	0.363	0.408	0.436
F test for all province-level effects			8.64			8.60			8.63

Note: Robust standard errors in parentheses. All regressions include period dummies. *** p<0.01, ** p<0.05, * p<0.1

Table 2b. Dependent variable value-added growth, 1995-2007 (4-year intervals). Porter's classification

	Model 2-1	Model 2-2	Model 2-3	Model 3-1	Model 3-2	Model 3-3	Model 4-1	Model 4-2	Model 4-3
Population density (Log)	0.00253**	0.00224**	-0.101***	0.00260**	0.00225**	-0.0949**	0.00264**	0.00223*	-0.0871**
	(0.00101)	(0.00109)	(0.0327)	(0.00104)	(0.00111)	(0.0372)	(0.00126)	(0.00127)	(0.0383)
Related trade variety	0.00423**	0.00590***	0.00532	0.00440**	0.00592***	0.00503	0.00367*	0.00536**	0.00432
	(0.00173)	(0.00183)	(0.00327)	(0.00176)	(0.00183)	(0.00334)	(0.00199)	(0.00213)	(0.00372)
Unrelated variety	-1.71e-05	-2.22e-05*	5.71e-05	-1.53e-05	-2.18e-05	5.24e-05	-2.10e-05	-2.63e-05**	5.67e-05
	(1.32e-05)	(1.28e-05)	(5.61e-05)	(1.34e-05)	(1.35e-05)	(5.92e-05)	(1.32e-05)	(1.29e-05)	(5.89e-05)
Import variety				-0.000506	-0.000102	0.00118			
				(0.000957)	(0.000905)	(0.00214)			
Related trade variety							0.00499	0.00478	0.0132
							(0.00567)	(0.00561)	(0.0116)
Trade similarity							-0.000367	-0.000226	-0.00243
							(0.000524)	(0.000536)	(0.00206)
Labor productivity (Log)		0.0326**	0.0263		0.0325**	0.0270		0.0314**	0.0305
		(0.0137)	(0.0264)		(0.0139)	(0.0267)		(0.0144)	(0.0244)
Human capital (Log)		-0.0191**	-0.00113		-0.0189**	-0.000943		-0.0196**	-0.00425
		(0.00797)	(0.0181)		(0.00814)	(0.0182)		(0.00803)	(0.0191)
Constant	0.0133***	-0.260**	0.154	0.0155**	-0.258**	0.116	0.0285	-0.235*	0.174
	(0.00454)	(0.126)	(0.300)	(0.00626)	(0.128)	(0.319)	(0.0202)	(0.135)	(0.291)
Observations	150	150	150	150	150	150	150	150	150
R-squared	0.133	0.176	0.182	0.135	0.176	0.185	0.140	0.182	0.217
Model	Pooled OLS	Pooled OLS	Fixed-effects	Pooled OLS	Pooled OLS	Fixed-effects	Pooled OLS	Pooled OLS	Fixed-effects
Moran's I p-value	0.575	0.528	0.720	0.572	0.525	0.726	0.543	0.514	0.709
F test for all province-level effects			1.22			1.21			1.30

Note: Robust standard errors in parentheses. All regressions include period dummies. *** p<0.01, ** p<0.05, * p<0.1

Table 2c. Dependent variable productivity-productivity growth, 1995-2007 (4-year intervals). Porter's classification

	Model 2-1	Model 2-2	Model 2-3	Model 3-1	Model 3-2	Model 3-3	Model 4-1	Model 4-2	Model 4-3
Population density (Log)	-0.00211	0.00255	-0.0570	-0.00214	0.00279	-0.0535	-0.000563	0.00297	-0.0640*
	(0.00166)	(0.00182)	(0.0346)	(0.00170)	(0.00180)	(0.0355)	(0.00166)	(0.00189)	(0.0379)
Related variety	0.000485	-0.00316	0.00113	0.000397	-0.00290	0.000980	0.00120	-0.00302	0.00133
	(0.00214)	(0.00192)	(0.00289)	(0.00225)	(0.00193)	(0.00313)	(0.00203)	(0.00201)	(0.00307)
Unrelated variety	8.58e-07	1.20e-05	-4.52e-05	-6.97e-08	1.49e-05	-4.80e-05	1.32e-05	1.55e-05	-5.33e-05
	(1.01e-05)	(1.06e-05)	(3.54e-05)	(9.89e-06)	(1.13e-05)	(3.55e-05)	(1.40e-05)	(1.11e-05)	(3.75e-05)
Import variety				0.000253	-0.000936	0.000646			
				(0.00124)	(0.000856)	(0.00145)			
Related trade variety							-0.00653	-0.00179	0.00870
							(0.00469)	(0.00494)	(0.00603)
Trade similarity							-0.000741	-0.000291	0.000597
							(0.000602)	(0.000492)	(0.000961)
Labor productivity (Log)		-0.128***	-0.238***		-0.130***	-0.237***		-0.126***	-0.239***
		(0.0228)	(0.0227)		(0.0230)	(0.0227)		(0.0231)	(0.0226)
Human capital (Log)		0.0256***	-0.00695		0.0268***	-0.00691		0.0275***	-0.0111
		(0.00839)	(0.0168)		(0.00827)	(0.0167)		(0.00810)	(0.0172)
Capital/labor (Log)		0.0124***	-0.0210***		0.0132***	-0.0217***		0.0123***	-0.0202***
		(0.00384)	(0.00710)		(0.00394)	(0.00748)		(0.00389)	(0.00749)
Constant	0.00533	1.109***	2.922***	0.00421	1.125***	2.900***	0.0316	1.102***	2.962***
	(0.00778)	(0.210)	(0.277)	(0.00853)	(0.211)	(0.277)	(0.0251)	(0.210)	(0.276)
Observations	150	150	150	150	150	150	150	150	150
R-squared	0.059	0.455	0.761	0.060	0.459	0.762	0.092	0.458	0.765
Model	Pooled OLS	Pooled OLS	Fixed-effects	Pooled OLS	Pooled OLS	Fixed-effects	Pooled OLS	Pooled OLS	Fixed-effects
Moran's I p-value	0.356	0.382	0.610	0.352	0.386	0.604	0.350	0.372	0.594
F test for all province-level effects			4.22			4.14			4.19

Note: Robust standard errors in parentheses. All regressions include period dummies. *** p<0.01, ** p<0.05, * p<0.1

Table 3a. Dependent variable employment growth, 1995-2007 (4-year intervals). Relatedness based on proximity

	Model 2-1	Model 2-2	Model 2-3	Model 3-1	Model 3-2	Model 3-3	Model 4-1	Model 4-2	Model 4-3
Population density (Log)	0.00339**	-0.00185	0.146***	0.00382**	-0.00178	0.148***	0.000465	-0.00173	0.141***
	(0.00156)	(0.00392)	(0.0354)	(0.00166)	(0.00391)	(0.0369)	(0.00252)	(0.00380)	(0.0367)
Related variety	0.00854**	0.00723**	0.00834	0.00896**	0.00744**	0.00789	0.00758	0.00475	0.0110*
	(0.00425)	(0.00356)	(0.00548)	(0.00437)	(0.00360)	(0.00561)	(0.00541)	(0.00512)	(0.00653)
Unrelated variety	-0.00349	-0.00247	0.00315	-0.00283	-0.00224	0.00311	-0.00461*	-0.00233	0.00341
	(0.00245)	(0.00259)	(0.00346)	(0.00249)	(0.00263)	(0.00350)	(0.00256)	(0.00264)	(0.00382)
Import variety				-0.00130	-0.000587	0.000443			
				(0.00181)	(0.00149)	(0.00196)			
Related trade variety							0.00176	0.00225	-0.00327
							(0.00221)	(0.00228)	(0.00315)
Trade similarity							0.00135	0.000520	-0.000850
							(0.000863)	(0.000869)	(0.00137)
Employment (Log)		0.00404	-0.244***		0.00421	-0.244***		0.00194	-0.240***
		(0.00455)	(0.0508)		(0.00447)	(0.0510)		(0.00455)	(0.0520)
Human cap. (Log)		-0.0372***	-0.00433		-0.0365***	-0.00409		-0.0397***	-0.000379
		(0.0134)	(0.0210)		(0.0135)	(0.0210)		(0.0137)	(0.0212)
Constant	0.0190**	-1.189***	0.441	0.0216**	-1.184***	0.426	-0.0267	-1.160***	0.464
	(0.00928)	(0.290)	(0.777)	(0.0100)	(0.289)	(0.788)	(0.0296)	(0.295)	(0.802)
Observations	150	150	150	150	150	150	150	150	150
R-squared	0.111	0.326	0.810	0.115	0.327	0.810	0.132	0.331	0.813
Model	Pooled OLS	Pooled OLS	Fixed-effects	Pooled OLS	Pooled OLS	Fixed-effects	Pooled OLS	Pooled OLS	Fixed-effects
Moran's I p-value	0.420	0.435	0.521	0.414	0.431	0.538	0.419	0.421	0.520
F test for all province-level effects			8.87			8.77			8.76

Note: Robust standard errors in parentheses. All regressions include period dummies. *** p<0.01, ** p<0.05, * p<0.1

Table 3b. Dependent variable value-added growth, 1995-2007 (4-year intervals). Relatedness based on proximity

	Model 2-1	Model 2-2	Model 2-3	Model 3-1	Model 3-2	Model 3-3	Model 4-1	Model 4-2	Model 4-3
Population density (Log)	0.00183**	0.00178*	-0.107***	0.00214**	0.00201*	-0.110***	0.00142	0.00124	-0.105**
	(0.000825)	(0.000931)	(0.0338)	(0.000917)	(0.00102)	(0.0351)	(0.00136)	(0.00133)	(0.0409)
Variety									
Related variety	0.00550**	0.00526**	0.0127**	0.00580**	0.00551**	0.0134**	0.00460	0.00431	0.0165**
	(0.00215)	(0.00219)	(0.00597)	(0.00217)	(0.00222)	(0.00593)	(0.00331)	(0.00334)	(0.00692)
Unrelated variety	-0.00150	-0.000923	0.00677	-0.00103	-0.000615	0.00683	-0.00156	-0.00101	0.00687
	(0.00155)	(0.00160)	(0.00613)	(0.00160)	(0.00162)	(0.00623)	(0.00156)	(0.00161)	(0.00657)
Import variety				-0.000920	-0.000709	-0.000656			
				(0.000958)	(0.000968)	(0.00191)			
Related trade variety							0.000777	0.000881	-0.00496
							(0.00146)	(0.00147)	(0.00348)
Trade similarity							0.000135	0.000225	-0.00180
							(0.000410)	(0.000435)	(0.00182)
Labor productivity (Log)		0.0176	0.0216		0.0165	0.0210		0.0171	0.0178
		(0.0126)	(0.0248)		(0.0128)	(0.0245)		(0.0129)	(0.0248)
Human capital (Log)		-0.0114*	-0.00648		-0.0105	-0.00685		-0.0124*	-7.96e-06
		(0.00666)	(0.0168)		(0.00689)	(0.0168)		(0.00660)	(0.0177)
Constant	0.0195***	-0.125	0.215	0.0213***	-0.115	0.242	0.0152	-0.124	0.317
	(0.00595)	(0.121)	(0.291)	(0.00617)	(0.124)	(0.287)	(0.0149)	(0.123)	(0.302)
Observations	150	150	150	150	150	150	150	150	150
R-squared	0.146	0.161	0.242	0.152	0.165	0.243	0.148	0.164	0.264
Model	Pooled OLS	Pooled OLS	Fixed-effects	Pooled OLS	Pooled OLS	Fixed-effects	Pooled OLS	Pooled OLS	Fixed-effects
Moran's I p-value	0.628	0.560	0.779	0.626	0.558	0.794	0.610	0.551	0.805
F test for all province-level effects			1.52			1.49			1.58

Note: Robust standard errors in parentheses. All regressions include period dummies. *** p<0.01, ** p<0.05, * p<0.1

Table 3c. Dependent variable productivity-productivity growth, 1995-2007 (4-year intervals). Relatedness based on proximity

	Model 2-1	Model 2-2	Model 2-3	Model 3-1	Model 3-2	Model 3-3	Model 4-1	Model 4-2	Model 4-3
Population density (Log)	-0.00159	0.00266*	-0.0633*	-0.00174	0.00309**	-0.0597*	0.000998	0.00361**	-0.0607
	(0.00121)	(0.00143)	(0.0323)	(0.00135)	(0.00152)	(0.0328)	(0.00165)	(0.00161)	(0.0377)
Related variety	-0.00295	-0.00135	0.000594	-0.00310	-0.000978	-4.42e-05	-0.00299	0.000284	-0.000945
	(0.00299)	(0.00183)	(0.00474)	(0.00309)	(0.00191)	(0.00509)	(0.00387)	(0.00314)	(0.00568)
Unrelated variety	0.00183	-4.70e-05	-0.00280	0.00160	0.000312	-0.00287	0.00293*	0.000127	-0.00295
	(0.00148)	(0.00138)	(0.00346)	(0.00141)	(0.00134)	(0.00354)	(0.00172)	(0.00146)	(0.00364)
Import variety				0.000462	-0.000959	0.000615			
				(0.00124)	(0.000868)	(0.00161)			
Related trade variety							-0.000943	-0.00153	0.00181
							(0.00182)	(0.00177)	(0.00317)
Trade similarity							-0.00126**	-0.000414	0.000354
							(0.000589)	(0.000464)	(0.00107)
Labor productivity (Log)		-0.121***	-0.239***		-0.123***	-0.238***		-0.120***	-0.237***
		(0.0220)	(0.0228)		(0.0221)	(0.0224)		(0.0220)	(0.0213)
Human capital (Log)		0.0212**	-0.00415		0.0222***	-0.00384		0.0231***	-0.00626
		(0.00826)	(0.0160)		(0.00804)	(0.0161)		(0.00855)	(0.0165)
Capital/labor (Log)		0.0132***	-0.0211***		0.0141***	-0.0219***		0.0131***	-0.0218***
		(0.00371)	(0.00690)		(0.00387)	(0.00719)		(0.00374)	(0.00698)
Constant	0.00182	1.043***	2.980***	0.000886	1.058***	2.957***	0.0446**	1.042***	2.938***
	(0.00609)	(0.198)	(0.277)	(0.00660)	(0.199)	(0.269)	(0.0198)	(0.198)	(0.270)
Observations	150	150	150	150	150	150	150	150	150
R-squared	0.070	0.444	0.759	0.071	0.448	0.760	0.104	0.450	0.761
Model	Pooled OLS	Pooled OLS	Fixed-effects	Pooled OLS	Pooled OLS	Fixed-effects	Pooled OLS	Pooled OLS	Fixed-effects
Moran's I p-value	0.378	0.400	0.625	0.375	0.404	0.627	0.383	0.393	0.621
F test for all province-level effects			4.29			4.21			4.17

Note: Robust standard errors in parentheses. All regressions include period dummies. *** p<0.01, ** p<0.05, * p<0.1

Table A1. Porter's clusters

Cluster code	Cluster name
1	Aerospace engines
2	Aerospace vehicles and defense
3	Agricultural products
4	Analytical instruments
5	Apparel
6	Automotive
7	Biopharmaceuticals
8	Building fixtures, equipment and services
9	Chemical products
10	Communications equipment
11	Construction materials
12	Reproduction equipment
13	Fishing and fishing products
14	Footwear
15	Forest products
16	Furniture
17	Heavy machinery
18	Information technology
19	Jewellery and precious metals
20	Leather products
21	Lighting and electrical equipment
22	Marine equipment
23	Medical devices
24	Metal manufacturing
25	Motor driven products
26	Oil and gas products and services
27	Plastics
28	Power generation and transmission
29	Prefabricated enclosures
30	Processed food
31	Production technology
32	Publishing and printing
33	Sporting, recreational and children’s goods
34	Textiles
35	Tobacco
36	Coal

Table A2. Correlation matrix for independent variables

	Pop. density	Variety	R.V. Conv.	R.V. Porter	R.V. Prox.	U.V. Conv.	U.V. Porter	U.V. Prox.	Imp. V.	R.T.V. Conv	R.T.V. Porter	R.T.V. Prox.	Trade Sim.	Employ.	Lab. prod.	Human cap.	Cap./Lab.
Pop. density	1																
Variety	0.5561	1															
R.V. Conv.	0.6667	0.84	1														
R.V. Porter	0.5785	0.8832	0.9145	1													
R.V. Prox.	0.4751	0.8181	0.6836	0.7686	1												
U.V. Conv.	0.0685	0.6435	0.1596	0.334	0.5099	1											
U.V. Porter	0.7068	0.8408	0.8271	0.7754	0.7233	0.3834	1										
U.V. Prox.	0.4402	0.8876	0.8329	0.8374	0.7351	0.4442	0.82	1									
Imp. V.	0.5107	0.6391	0.5891	0.6019	0.5915	0.3203	0.651	0.6417	1								
R.T.V. Conv	0.6231	0.7311	0.8549	0.7609	0.6954	0.1298	0.7884	0.7761	0.6698	1							
R.T.V. Porter	0.6682	0.746	0.8628	0.7236	0.5687	0.1536	0.7939	0.759	0.57	0.8156	1						
R.T.V. Prox.	0.5337	0.7762	0.6631	0.7371	0.911	0.4507	0.7038	0.6678	0.6471	0.7092	0.6009	1					
Trade Sim.	0.6688	0.4528	0.6456	0.4999	0.3039	-0.0745	0.6335	0.4374	0.3063	0.6057	0.7412	0.3815	1				
Employ.	0.8811	0.6413	0.7096	0.6687	0.5262	0.1592	0.7726	0.5519	0.5702	0.6513	0.6848	0.5886	0.7148	1			
Lab. prod.	0.3911	0.2602	0.3492	0.1844	0.2483	-0.0866	0.3443	0.2296	0.2242	0.4538	0.4302	0.2828	0.4721	0.2282	1		
Human cap.	0.3496	0.3973	0.4683	0.3916	0.3235	0.0011	0.4291	0.4062	0.4025	0.5346	0.5144	0.3874	0.5509	0.3383	0.5805	1	
Cap./Lab.	-0.1125	0.0902	-0.0216	-0.1036	0.1604	0.1347	0.1061	0.1121	0.0578	0.1343	0.0965	0.1163	-0.0315	-0.2002	0.5355	0.2635	1