Mind your step: The heterogeneous effect of relatedness and diversification process in EU and ENP countries

SEARCH Working Paper, 2/09.

RON BOSCHMA[1] & GIANLUCA CAPONE[2]

[1] CIRCLE, Lund University
[2] URU, Utrecht University

Abstract

In this working paper we analyze the degree of relatedness between export products based on co-occurrence analysis (Hausmann & Klinger, 2007). We use the BACI Dataset (an elaboration from UN Comtrade Data), which guarantees ample coverage of ENP countries over time, at the 4-digit classification level. Our results show that in both EU and ENP countries the evolution of the export mix is strongly path-dependent: countries tend to keep a comparative advantage in products that are strongly related to their current productive structure, and they also diversify in nearby products. However, this effect is much stronger for ENP countries, signalling their lower capabilities to diversify in products that are not very near to their productive structure. Moreover, we also show that the future exports structure of countries is affected by imports: both EU and ENP countries keep a comparative advantage in products that are strongly related to their imports, but only EU countries show also evidence of learning to diversify in new products from related import sectors.

KEY-WORDS: diversification; relatedness; European Neighborhood Policy; trade

JEL: F19; O14; O33.

Acknowledgments
This research was supported by the Project "Sharing KnowledgE Assets: InteRregionally Cohesive NeigHborhoods" (SEARCH) within the 7th European Community Framework Program (FP7-SSH-2010.2.2-1), Grant Agreement 266834.

This paper is an updated version of SEARCH WP 2/04

1. Introduction

The evolution of the productive structure of countries is a relevant topic from both the academic and the policymaker point of a view. The wealth of countries depends on the richness and complexity of their products, but there is a strong heterogeneity in what they produce. The traditional paradigm of trade theory suggests that this heterogeneity should reflect underlying characteristics of countries, i.e. factor endowments or productivity. However, recent developments in the international trade literature question this paradigm by showing that the current productive structure of a country is affected by its own past productive structure, through a path-dependent process governed by the relatedness between products (Hausmann and Klinger, 2007). Relatedness matters because the product space is very heterogenous and it has a core-periphery structure (Hidalgo et al., 2007): jumping into new products is far from easy and straightforward, and it requires specific capabilities. Capabilities are not easily transferred across countries: if a country does not have most of the capabilities needed to produce a new good, it can be very difficult to start producing it (Hidalgo and Hausmann, 2009).

In this working paper, we propose two contributions to the relatedness and diversification literature. First, we claim that the path-dependent process of product diversification is driven not only by each country past productive structure, but also by its relationships with other countries: countries are not isolated monads, but are embedded in several networks through different channels, such as political relations and international trade. Jumping into new products because of these relationships might be more convenient, because of stronger economic incentives, or just easier, because of learning opportunities.

Second, we claim that the constraints of path-dependence are not equally binding for all countries. Capabilities may refer to very different domains: they include tangible inputs, such as infrastructure, or intangible ones, such as knowledge and institutions. However, while some capabilities are important only for specific products or groups of products (e.g. specific technological knowledge), there are also general-purpose capabilities that are relevant for all products, and are also country specific (e.g. institutions favoring or hindering entrepreneurship). Countries characterized by stronger general-purpose capabilities could find it easier to jump to any new product: therefore, the importance of product relatedness would be much lower in this case. On the contrary, countries characterized by weaker general-purpose capabilities would rely much more strongly on the links between products in order to jump into new industries. Similar considerations can also apply to learning: while some learning is certainly product specific, a more general absorptive capacity might differ across countries. Therefore, countries with higher learning capabilities might be better able to exploit their network relations to diversify into new products.

In the next section, we present the theoretical framework behind our analysis. Then we illustrate the methodology and the data. In the results section, we present both a descriptive and an econometric analysis. We conclude by discussing the policy and theoretical implications of our work.

2. The structure of the product space, diversification and relatedness

A striking characteristic of modern economies is the rich variety of the products they are able to provide. Some of the founding fathers of the economic science (Smith, 1776; Young, 1928) recognized that not all products are equal and they may have very different effects on the growth of countries. Only recently, though, economists have been able to provide a more precise and formal account of these ideas. By using international trade data, Hidalgo et al. (2007) drew a map of the product space and showed that some products are in a dense part of this map - that is, they are related to many other products - whereas other products are in the periphery of the map. Moreover, they also showed that countries specialized in products in the dense part of the product space have higher growth rates than countries specialized in periphery products.

Two causal mechanisms have been put forth to explain these empirical results. Hausmann and Klinger (2007) develop a model of product-specific human capital, with heterogeneous degree of substitutability across products and overlapping generations. In each period there is a young untrained worker who does not produce, but is trained by the old worker in the specific product the old worker is specialized on. Moreover, the old worker can choose whether to stick to the product in which she was trained in the previous period or to jump to an alternative good for which her product-specific training is an imperfect substitute. Given the marginal increase in revenue by unit of distance between products *f* and the following quadratic cost function:

$$C(\delta) = \frac{c \cdot \delta^2}{2}, \qquad (1)$$

it is possible to determine the optimal distance to jump, which is:

$$\delta^* = \frac{f}{c} \qquad (2).$$

If the product space is not continuous, but discrete, three immediate implications follow from this simple specification. First, at the optimal distance there could be no product and the workers might be forced to adopt non optimal solutions. Second, since the profits from jumping first increase and then decrease with distance, there could be no product such that jumping is convenient for the worker: in this case she keeps producing the old one. Third, generalizing the model, workers (countries) located in areas where products are more dense, will have higher opportunity to jump to new products, and therefore to grow.

An important assumption of this model is that workers might differ only in their specialization: both the revenues and the costs are the same for all workers. However, suppose there is heterogeneity among workers with respect to one of these dimensions (e.g. costs). Let worker *i* have higher costs than worker *j* (that is, $c_i > c_j$). Then the optimal distance to jump will be higher for worker *j* (that is, $\delta_j^* > \delta_i^*$). Therefore, the two workers will probably end up in specializing in different products, and the impact of distance on the diversification and branching process will be much stronger for the worker (country) that has the higher costs.

The second causal mechanism refers to the concept of capabilities. Capabilities might refer to different levels of analysis. At the firm level they identify complex routines or collection of routines that give an organization a set of options for producing specific outputs (Winter, 2003). They are an important source of firms competitive advantage because they cannot be easily imitated (Dosi, Nelson and Winter, 2000). At the country level, they refer to specific infrastructure, skills, knowledge, institutions or norms, and determine advantage for countries because they are not internationally tradable (Hausmann and Hidalgo, 2010). If a country is specialized in a certain product, it clearly has the capabilities to produce it, as well as most of the capabilities necessary to produce similar products. Therefore, countries specialized in products located in the dense part of the product space will have more opportunities to redeploy their capabilities in new products and will have higher growth rates. Recent empirical evidence indirectly confirms the role played by capabilities in the diversification process. Industry case studies show that the most successful firms in new industries are founded by entrepreneurs with experience in related industries: many carriage-makers were able to redeploy their experience in complex assembly in the newly born automobile industry (Klepper, 2002), and the dominant firms in the radio industry were also able to dominate the television receiver industry (Klepper and Simons, 2000). Boschma et al. (2012) show that the diversification process is much stronger at the regional rather than country level, which is compatible with the concept of non-tradable and localized capabilities.

An important difference between the capability perspective and the previous model is related to the heterogeneity between countries. Countries differ in the capabilities they have and these differences will affect the diversification process as much as the distances between products. Still, this heterogeneity between countries does not imply also a different effect of distance across countries, because the effect of distance is driven exactly by the heterogeneity in the capabilities, which refer to the specific characteristics of products. What if there are some capabilities that are not related to specific products, but to all products? General-purpose capabilities can be country specific rather than product specific. The diversification process in countries with strong general-purpose capabilities will be less affected by the existence of product specific capabilities. However, in countries with weak general-purpose capabilities the only way to move to new industries would require to exploit the existing product specific capabilities. An example might be useful to clarify this concept. Consider two countries A and B, where A has stronger general-purpose capabilities. Let F be a biotechnology company specialized in cancer diseases. If the company discovers a new drug related to a different type of cancer, then it will be profitable to exploit it both in country A and in country B. However, if the new drug refers to a quite different domain (say, organ transplant rejection), it would be much easier to exploit it if F is located in country A, where differentiation in distant markets is less costly. Moreover, even if F does not exploit the innovation directly, if in country A the conditions for the creation of spinoffs are easier, we will observe that country A diversifies in the new product, whereas this would not happen in country B.

So, our first claim in this paper is that the path-dependent process driving the diversification into new industries might offer different degrees of freedom in different countries. In countries where general costs are lower and general purpose capabilities are stronger, we will observe a lower role for distance or relatedness between products; in countries with higher costs and weaker capabilities we will observe the opposite pattern.

A natural follow-up question is to ask whether there is any other condition that might reduce the strength of path-dependence in the diversification process. The productive structure of a country is affected by its own past history, but is also embedded in a network of relationships. It is certainly possible that these relationships can exert some influence on the direction and the intensity of the diversification process. Here we will consider in detail trade relationships, although other type of links might be also relevant. For example, Bahar et al. (2012) consider geographical proximity between countries, while Boschma et al. (2013) consider the role of other regions within the same country.

In the theoretical literature on international trade is not uncommon to find references to the possibility of learning from exports and imports (Redding, 2010). At the empirical level, though, there is quite a lot of variation. Macro level studies refer to the effects of trade on productivity and growth of countries (Singh, 2010) or regions (Boschma and Iammarino, 2009), and then they infer that some learning is probably occurring to explain the observed outcomes.

More details on the mechanisms can be found in micro level studies (Wagner, 2007, 2012). A well documented stylized fact is that firms engaging in exports have higher productivity: empirical evidence shows that this is at least partially due to self-selection of better firms into exporting behavior (Lileeva and Trefler, 2010). There is much less consensus on the existence of ex-post effects, and existing studies tend to suggest economies of scale explanations rather than learning (Silva et al., 2012).

However, the situation is reverted in the case of imports: there is a limited evidence for self-selection in importing activities, because of fixed costs (Altomonte and Békés, 2010), while much more convincing is the case for indirect forms of learning (Castellani et al. 2010). In fact, importing firms may exploit the availability of more variety in inputs and also the possibility that imported products embody higher quality. These processes are feasible even when absorptive capacity by firms is missing (Cohen and Levinthal, 1990), because characteristics embedded in products do not require actual learning by firms. However, they might not be profitable in all countries: the availability of fitter and higher quality inputs can be exploited only when there is a demand (internal or foreign) that is sensitive to these issues; otherwise, imports will have no impact on the probability of moving into new products.

Imports might also operate on diversification through channels other than learning. Strong imports in a sector signal relevance of this sector for the productive structure of a country, and therefore provide a strong incentive for firms in the country to move into it. Still, if production was not occurring before, the required capabilities should be developed and this could be more difficult when suffering from strong international competition.

More generally, all trade relationships might also signal other type of links between countries, such as mobility of workers or cultural similarities, that could also favor learning opportunities or provide pressure to remain active in sectors that were already well developed in the past.

So, our second claim in this working paper is that the evolution of the productive structure of countries is affected by its international trade relationships. In this respect, we will focus our attention on imports, but we will also look at the impact of the productive structure of countries embedded in the trade network, that is of origin and destination countries.

3. Methodology and Data

In order to represent the product space we follow quite closely the approach outlined by Hidalgo et al. (2007) and Hidalgo and Hausmann (2009). Our starting point is the concept of comparative advantage developed by Balassa (1965). A country has a comparative advantage in a product *i* when the share of this product in its exports is larger than the share of the product in the world exports. We can easily compute the number of products in which each country has a comparative advantage; formally, we can write this number as:

$$\pi_{c,t} = \sum_k x_{k,c,t} \qquad (3)$$

where x takes the value of 1 if country c has a comparative advantage in product k at time t, and zero otherwise. Analogously, we can also define the number of countries that have a comparative advantage in exporting a certain product:

$$\nu_{i,t} = \sum_n x_{i,n,t} \qquad (4).$$

We can use these two measures to define an indicator providing more information on the characteristics of the products, and in particular of their complexity. Consider a country c, which has $\pi_{c,t}$ products. Then

$$\xi_{c,t} = \frac{\sum_k x_{k,c,t} \cdot \nu_{k,t}}{\pi_{c,t}} \qquad (5)$$

is the average number of countries that have a comparative advantage in the products in which also country c has a comparative advantage. A high value of $\xi_{c,t}$ means that many countries are able to produce the same products of country c, which signals that the products in which this country is specialized are not very complex and do not require specific capabilities. On the contrary, a low value of $\xi_{c,t}$ is an indicator of high complexity of the productive structure of country c. In the network jargon, if we consider international trade as a bipartite network where countries and products are nodes and comparative advantage is a link, $\xi_{c,t}$ is the average nearest neighbor degree.

The next step in building a product space requires a measure of the proximity between industries. The proximity (φ) between two products (i and j) in a given year t can be formally expressed as:

$$\varphi_{ijt} = min\{P(x_{i,t}|x_{j,t}), P(x_{j,t}|x_{i,t})\} \qquad (6)$$

that is, the proximity between product i and j in year t is the minimum between the conditional probability of having a comparative advantage in product i given a comparative advantage in product j, and the conditional probability of having a comparative advantage in product j given a comparative advantage in product i. The rationale behind the proximity measure is that if two products are related because they require similar institutions, infrastructure, productive inputs, organizational routines and capabilities, and technology, then they will be probably produced together. Conditional probabilities rather than joint probabilities must be used, so that the measure is not affected by the relevance of the products in the world trade. The minimum between conditional probabilities is used in order to ensure a symmetric and conservative measure.

Proximity is a property referring to the link between two products. In order to analyze countries we need to place them in this space. This can be done by using a density indicator, that measures how close a product is to the current productive structure of a country or a region. Formally, density can be expressed as follows:

$$d_{i,c,t} = \frac{\sum_k x_{k,c,t} \cdot \varphi_{i,k,t}}{\sum_k \varphi_{i,k,t}} \qquad (7)$$

where φ represents proximity (between product *i* and product *k*) and *x* takes the value of 1 if country *c* has a comparative advantage in product *k* at time *t*, and zero otherwise. So, density around a product will be high if a country has a comparative advantage in most of the products related to the focal one; at the very extreme, it will be equal to 1, if a country has a comparative advantage in all products with a non-zero proximity to the focal product. Conversely, density around a product will be low (zero) if a country does not have a comparative advantage in most (any) of the products related to the focal one.

To measure the role of imports, we adapt the density indicator and develop an import density indicator. The formal definition is analogous to the definition of density:

$$id_{i,c,t} = \frac{\sum_k m_{k,c,t} \cdot \varphi_{i,k,t}}{\sum_k \varphi_{i,k,t}} \qquad (8).$$

However, here *m* takes the value of 1 if product *k* has a higher share in the imports of country *c* than in the world imports (we could say if country *c* has a comparative "advantage" in importing product *k*), and zero otherwise. So, import density around a product will be high if a country has a comparative "advantage" in importing most of the products related to the focal one; conversely, import density around a product will be low if a country does not have a comparative "advantage" in importing most of the products related to the focal one.

Finally, to take into account the role of trade relationships with other countries, we employ two different strategies. First, we consider the most preferred countries in the trade relationships (Pinna, 2012), since the evolution of their productive structure might have important effects on the trade partners. In particular, we consider the (export) density of the most preferred origin and the import density of the most preferred destination: the former (most preferred origin) refers to the country (within the EU and ENP) that has the highest share in the focal country imports; the latter (most preferred destination) refers to the country (within the EU and ENP) that has the highest share in the focal country exports.

Second, we specify a more detailed measure, by considering all trade relationships: rather than considering only one country, we use a weighted sum of the densities (import densities) of all the partner countries; the weight is given by the share of the partner country in the imports (exports) of the focal country. Formally, the weighted density *wd* can be expressed as:

$$wd_{i,c,t} = \sum_b \frac{M^c_{i,b,t}}{M^c_{i,t}} \cdot d_{i,b,t} \qquad (9)$$

where $d_{i,b,t}$ is the density of country *b* around product *i* at time *t*, $M^c_{i,b,t}$ is the value of country *c* imports of product *i* from country *b* at time *t*, and $M^c_{i,t}$ is the total value of country *b* imports from the EU and ENP sample. In the case of destinations, import density is used instead of density, and exports values instead of imports values. The weighted density measure has a clear advantage with respect to the previous one (most preferred countries) because it is more detailed and it exploits all the available information; however, there is also an important limitation, as the weighted density is not an actual density (and it can take a zero value, if a country has no trade in a specific product category), and this might introduce some problems in the interpretation of the results.

In order to calculate all indicators we described so far, we use country-level world trade data from the BACI database for the period 1995-2010 (Gaulier and Zignago, 2010). This database is developed from UN Comtrade data using a procedure that reconciles the declarations of the exporter and the importer, allowing to extend considerably the number of countries with available trade data, including many countries in the European Neighborhood Policy group. Moreover, data are available at an high level of product disaggregation (6-digit Harmozized System), although for the current analysis we use a 4-digit level, which includes 1241 different products[1]. Descriptive statistics and correlations for all indicators described in this section and the variables used in the econometric analysis are reported in Table A1, distinguishing between two cases: 5-years interval data, that include 3 time points (1995, 2000, 2005); 3-years interval data, that include 5 time points (1995, 1998, 2001, 2004, 2007).

4. Results: Descriptive Analysis

Given the indicators described so far, we can represent countries in the space defined by both the number of products in which they specialize (the level of diversification, eq. 2) and the average number of countries exporting the products (the level of complexity, eq. 4). Figure 1 shows the position of all EU and ENP countries in this space. Our results confirm a positive relationship between the complexity and diversification (Hidalgo and Hausmann, 2009): countries with a high level of diversification (have a comparative advantage in a high number of products) produce also more complex products (products that require rarer capabilities and therefore are produced by a low number of countries). However, for our purposes is more relevant to notice that most of the ENP countries (in red) are located in the upper left corner (low complexity and diversification) with the notable exception of Israel and Russia, while EU countries (in blue) are in the center (mostly countries from recent enlargements) and in the downer right corner (mostly old EU countries). Moreover, it also worth to notice that in 15 years there are no big changes in the location of countries on this map.

A certain degree of stability characterizes also density and import density. In order to study their evolution over time, we divide our countries in four groups: old EU countries (EU 15), new accession EU countries (EU new), eastern ENP countries, and southern ENP countries. Figure 2a shows that there is a clear difference in the average level of density for the four groups of countries: old EU countries have the highest level, the other EU countries have an intermediate level, while both eastern and southern ENP countries have the lowest levels. Moreover, the dynamics of density over time is also very interesting: in old EU countries it consistently increased over time; in the other EU countries, it first declined and then increased again starting from 2003. Density shows also a decline in the eastern ENP countries and an increase in the southern ENP countries. So, this analysis confirms the insights we got from the previous graph: strong differences between groups and slow processes of change over time. Here we can add the following element: since density signal that countries are strongly embedded in the network of products, we see that EU countries are in a much denser part of the product space than ENP countries. Figure 2b shows the evolution of import density over

[1] We use 4-digit data rather than 6-digit because the computation of conditional probabilities is highly demanding for memory. However, even as it is, our analysis is more fine-grained than what can be found in previous studies: Hausmann and Klinger (2007) use a specification with 1006 products, while Boschma et al. (2012) have 775 products.

time. The most striking characteristic of this graph is that the rank is exactly the same as in the case of density: EU countries have higher import density than ENP countries, signaling that they are much more embedded in international trade. While the level of import density of EU countries remains substantially constant over time, in both groups of ENP countries it tends to increase. However, the increase in import density is much steeper for the eastern countries, that experienced a decrease in density: this could be a signal of substitution between internal production and imports.

Moving to the issue of diversification, we replicate in our sample the following stylized fact: countries have a higher probability to develop a comparative advantage in products characterized by higher density. Figure 3 presents the probability of developing a comparative advantage in a new product (five years later) for different density ranges in the EU (a) and ENP (b) case, respectively. In both cases, to higher levels of density correspond higher probabilities to develop a comparative advantage in new products. However, there are two important differences between the two groups: first, among the ENP countries there are no products with density higher than 0.4, although from this level on EU countries have the highest probability to develop a comparative advantage in new products; second, for almost all levels of density, ENP countries have a higher probability to develop a comparative advantage in new products.

Interesting details emerge from a more disaggregated representation. In Figure 4, we plot average density in products with no comparative advantage against the number of new products where a comparative advantage has been developed five years later, for three different points in time (a: 1995; b: 2000; c: 2005). A positive relationship is evident in all cases; however, the plots also show a stronger relationship, lower average densities, and a higher variation over time in the number of new products for ENP countries (in red).

So, both Figure 3 and Figure 4 hint at the existence of a difference in the effect of density between EU and ENP countries. What about import density? Remember that the density indicator measures the closeness of a good to the productive structure of a country; density drives the evolution of countries productive structure, because it is easier for them to move to nearby goods rather than jump far away. The meaning and impact of import density are less clear-cut. A product with high import density is close to goods that are strongly needed in the country: this could be an important incentive for the country to produce it locally and could also provide an opportunity for learning from international trade. However, import density could also signal the lack of significant production capabilities: in this case, high import density would not be a driver for the evolution of the country productive structure, and it could even be associated to a lower probability of developing a comparative advantage in new products. Finally, a low import density might favor within-country production, because it could provide a space sheltered from the international competition that could lead to the emergence of a strong national sector.

Analogously to what we have done for density, in Figure 5 we present the probability of developing a comparative advantage in a new product (five years later) for different import density ranges in the EU (a) and ENP (b) case, respectively. The results are quite striking: in the EU case, the probability of developing a comparative advantage in a new product increases strongly with import density; however, in the ENP case, the pattern is far less clear, and the highest probability peak is at very low levels of import density. This suggests that the mechanisms we have mentioned before might have different strength in different groups of countries.

5. Results: Econometric Analysis

All the results presented in the previous section are descriptive in nature: we have to perform more formal tests to validate them. So we estimate the following econometric equation:

$$x_{i,c,t+k} = \alpha + \gamma x_{i,c,t} + \beta_d d_{i,c,t} + \beta_{id} id_{i,c,t} + \delta_d ENP \cdot d_{i,c,t} + \delta_{id} ENP \cdot id_{i,c,t} + \pi X + \varepsilon_{i,c,t} \quad (10)$$

where the dependent variable takes value 1 if country *c* has a comparative advantage in product *i* at time *t* + *k* (*k* = 3, 5) and zero otherwise, $d_{i,c,t}$ denotes the density around product i in country c at time t, $id_{i,c,t}$denotes import density around product i in country c at time t, ENP is a dummy variable that takes value 1 if the country belongs to the European Neighborhood Policy group, and zero if it belong to the European Union, and *X* is a vector of country-year and product-year dummy variables, which control for any time-varying country or product characteristic. The coefficient $\delta_d$and δ_{id}capture any eventual difference in the impact of density between EU and ENP countries. Both density and import density are normalized by subtracting the mean and dividing by the standard deviation.

The results obtained from OLS estimation with standard errors clustered at the country level are presented in Table 1. Model 1 presents the results with a 5-years interval between the dependent and the independent variables. Both density and import density have a positive and significant effect. In particular, the positive and significant coefficient of import density suggests that countries tend to diversify also in new industries that are close to their import needs. In model 2, we add interaction effects, to check whether density and import density work differently in EU and ENP countries. In both cases, we can observe significant differences: ENP countries are characterized by a stronger impact of density and no role for import density. The last result is particularly interesting, since it suggests that the mechanisms behind the role of import density might work effectively only in some countries. In models 3 and 4 we check the robustness of our results with a different time specification (a 3-years interval): both density indicators and interactions are weaker in these models, whereas the lagged dependent variable has a stronger effect. This strengthens the idea that density and import density have a stronger impact over longer periods of time.

These results should be carefully interpreted. In particular, the model specified before does not distinguish between the effects of density and import density in keeping a comparative advantage in a certain product and developing a comparative advantage in a new product. Following Hausmann and Klinger (2007) and Boschma et al. (2012) we perform this refined analysis by estimating the following equation:

$$\begin{aligned} x_{i,c,t+5} = {} & \alpha + \gamma x_{i,c,t} + \beta_d^o (x_{i,c,t}) \cdot d_{i,c,t} + \beta_d^n (1 - x_{i,c,t}) \cdot d_{i,c,t} + \beta_{id}^o (x_{i,c,t}) \cdot id_{i,c,t} \\ & + \beta_{id}^n (1 - x_{i,c,t}) \cdot id_{i,c,t} + \delta_d^o (x_{i,c,t}) \cdot ENP \cdot d_{i,c,t} + \delta_d^n (1 - x_{i,c,t}) \cdot ENP \cdot d_{i,c,t} \\ & + \delta_{id}^o (x_{i,c,t}) \cdot ENP \cdot id_{i,c,t} + \delta_{id}^n (x_{i,c,t}) \cdot ENP \cdot id_{i,c,t} + \pi X + \varepsilon_{i,c,t} \qquad (11) \end{aligned}$$

where β_d^o (β_{id}^o) captures the impact of density (import density) in keeping a comparative advantage in product i, β_d^n (β_{id}^n) captures the impact of density (import density) in developing a comparative advantage in a new product, and the δ coefficients capture the existence of any difference in the impact of density and import density across EU and

ENP countries. As in the previous case, both density and import density are standardized. The model is estimated using OLS with country-clustered standard errors. Results are presented in Table 2. Model 1 shows that both density and import density have a positive effect on both keeping a comparative advantage in a current product and developing a comparative advantage in a new product. However, consistently with previous findings, the former effect is stronger than the latter: both density and import density play a larger role in keeping a comparative advantage than in developing a new one. In model 3 we introduce the interaction effects: the results confirm that density has a stronger impact in the ENP country, and this holds in the case of both current products and new ones. However, the impact of import density on keeping current comparative advantage is similar across the two groups of countries (the interaction effect is not significantly different from zero). Finally, while import density plays an important role in the development of a comparative advantage in new products in EU countries, it has no importance in the case of the ENP group, suggesting that the difference in the mechanisms behind import density refers specifically to the creation of a comparative advantage in new products rather than to the retention of old products. The 3-years interval robustness checks (Model 3 and 4) confirm these outcomes.

6. Results: The Role of Trade Partners

The results obtained in the previous section suggest that the evolution of the productive structure of countries might depend not only on their past history but also on the trade relationships with other countries. In this section we explore more in detail this possibility.

Whatever the measure we use, there are two reasons why density and import density of trade partners might affect the evolution of the productive structure of a country. First, high density and import density signal the centrality of a product in the productive structure of a country: this provides the opportunity and the incentive for trade partners to keep active and even strengthen the trade relationships around the product. Second, density and import density also signal the existence and the widespread diffusion of the capabilities required for getting an advantage in a given product: trade relationships with partners where these capabilities are high enough might provide important learning opportunities.

The model we estimate is analogous to the one in equation (11), where we disentangle the effect of density and import density on keeping a comparative advantage in old products and developing a comparative advantage in new products. We report again the results from this model in Column 1 of Table 3, in order to allow comparisons more easily. To this model, we add the density of the most preferred origin and the import density of the most preferred destination (Model 2) and the corresponding ENP interactions (Model 3). In Model 4, instead, we use weighted densities of origin and weighted import densities of destinations; finally, we include the ENP interactions in Model 5.

The first result that can be extracted from Table 3 is that neither the magnitude nor the significance of the main effects and ENP interactions emerging from the previous specifications are affected once we introduce the densities of trade partners: this was not obvious, as they could capture effects similar to those already revealed by the import density of the focal country. Second, in both the most preferred country and the weighted density specification, the import density of destination countries and the density of origin countries have a positive and significant effect on the probability of keeping a comparative advantage[2]. Third, there is no significant impact of trade partners densities on the development of new products in the most preferred countries specification and a significant, but very low effect in the weighted densities specification. Fourth, there are no relevant differences in the effects of trade partners across EU and ENP countries. Finally, all results hold also when considering a 3-years interval, with the usual caveat that the effect of densities is smaller and is compensated by a stronger role of the lagged dependent variable. Taken together, these results strongly suggest that opportunities and incentives mechanisms are at work here: they are more powerful in keeping old products than in developing new ones, and they are also less affected by the peculiar characteristics of the countries.

7. Conclusions

In this article, we investigated the process of diversification of EU and ENP countries, by using the proximity approach developed by Hausmann, Hidalgo and Klinger. Our results confirm the path-dependence in the diversification process: all analyzed countries tend to jump into new industries that are related to their current productive structure, because they can exploit the existing capabilities. However, the effect of density is much stronger in the case of ENP countries, signaling the existence of different types of capabilities: EU countries are also able to diversify into less related industries because of general-purpose capabilities, while ENP have to rely much more on the relatedness between products and the specific capabilities necessary to produce them.

Moreover, we also show that imports may have an impact on the trajectory of the productive structure of countries, provided that absorptive capabilities exist: in our sample, only EU countries are able to diversify into sectors related to their imports. The productive structure of trade partners, instead, does not have any impact on the diversification process, but it provides economic incentives to both EU and ENP countries to keep producing in old sectors that are related to what their partners do.

Both these results contribute to the literature on country diversification by showing that, although path dependence matters, there is still the possibility that the network of relations in which countries are embedded might change the direction and the intensity of the process. Further research should look more generally at the links between countries, considering also geographical and cultural types of relationships. More specifically, what is really missing in our analysis is a detailed account of why differences between countries in terms of costs or general-purpose capabilities emerge and persist over time. Considering our specific sample, European and Neighboring countries differ under many respects: education, culture, physical infrastructure. All these elements could be effectively summarized under the label of institutions (North, 1990), and there is actually a strong consensus on the role that institutions play in determining innovation and competitiveness of countries (Cantwell, 2005). Institutions refer to how societies are organized at all levels, including laws, customs, habits and traditions, and how this has an impact on the incentives, frameworks, ideas and behaviors of individuals and social entities. Some institutions directly favor or hinder the

[2] The magnitude difference of the density of origin countries in the two specifications should be probably explained by the strong difference in the standard deviations between the two variables (see Table A1).

emergence of innovations (Lipsey, 2009; von Tunzelmann, 2003); moreover, they interact at different levels generating complex structures such as national (Lundvall, 1992), regional (Cooke et al., 1997) or sectoral (Malerba, 2002) systems of innovation.

Moreover, institutions can also be actively shaped by government intervention (Rodrik, 1996; Glaeser et al., 2004). This is strictly linked with the issue of drawing policy implications, which is a very delicate exercise with respect to the analysis conducted so far. The most striking characteristic emerging from the study of the product space and the diversification process of countries is the strong path-dependence: the productive structure of the past keep exerting its influence many years later, and the position of countries in the product space, in terms of both diversification and complexity, is very stable over the whole period under analysis (15 years). Moreover, also the econometric analysis show that the effects of density and import density are stronger when considering a 5-years interval. Therefore, policy interventions related to the following comments should take into account that effects might display only over a long time period.

Our first results shows that in the case of ENP countries product relatedness measured through density has a stronger effect both in keeping a comparative advantage in old products and in diversifying into new products. Policy aimed at improving and speeding up the diversification process should consider that in these countries this could be obtained mostly by favoring the development of nearby sectors. Directly favoring the creation of very distant industries might result in severe failures, since the lack of necessary supporting infrastructure and institutions may doom these initiatives before positive diffusion effects may occur. However, together with these interventions focused on nearby industries, policy makers might also consider actions aimed at improving the quality of the supporting institutions: creating an environment where firms can emerge and grow more easily, or returns from innovation can be appropriated also by local companies and new innovators, might provide stronger incentives and opportunities for diversification even in very far products and therefore boost the future growth of countries.

Our second result on the role of import provides similar insights. While imports provide powerful incentives for both EU and ENP countries to keep producing in old sectors where they already have comparative advantage, learning effects are circumscribed only to EU countries. In the ENP countries, the availability of a wider variety of inputs or of higher quality products does not produce positive effects, because of the lack of institutions, capabilities and probably demand. Therefore, policy aimed at improving institutions might be very useful also in this respect. However, more specific policies might also consider trade flows as a whole in these countries: sectors opening to international imports should also be opened very soon to opportunities in exports, so to have the possibility to grow and support the diversification process of countries.

Finally, our results on the role of trade partners do not support any role for general trade policies even in developed countries. Neither EU nor ENP countries benefit from the existing productive structures and capabilities of their trade partners: although trade openness might have beneficial effects on countries, it would not improve the diversification process by itself.

References

Altomonte, C., and Békés, G., 2010. Trade complexity and productivity. *CeFig Working Papers,* 12. Center for Firms in the Global Economy.

Bahar, D., Hausmann, R., and Hidalgo, C.A., 2012. International knowledge diffusion and the comparative advantage of nations. *CID Working Paper*, No. 235. Cambridge, MA: Center for International Development, Harvard University.

Balassa, B., 1965. Trade liberalization and revealed comparative advantage. *Manchester School of Economics and Social Studies,* 33: 99-123.

Boschma, R., and Iammarino, S., 2009. Related variety, trade linkages and regional growth. *Economic Geography,* 85: 289-311.

Boschma, R., Minondo, A., and Navarro, M., 2013. The emergence of new industries at the regional level in Spain: A proximity approach based on product relatedness. *Economic Geography,* 89: 29-51.

Cantwell, J., 2005. Innovation and competitiveness. In Fagerberg, J., Mowery, D.C., and Nelson, R.R. (Eds), *The Oxford Handbook of Innovation.* Oxford, Oxford University Press: 243-267.

Castellani, D., Serti, F., and Tomasi, C., 2010. Firms in international trade: Importers' and exporters' heterogeneity in Italian manufacturing industry. *The World Economy,* 33: 424-457.

Cohen, W.M., and Levinthal, D.A., 1990. Absorptive capacity: A new perspective on learning and innovation. *Administrative Science Quarterly*, 35: 128-152.

Cooke, P., Gomez Uranga, M., and Etxeberria, G., 1997. Regional innovation systems: Institutional and organizational dimensions. *Research Policy,* 26: 475-491.

Dosi, G., Nelson, R.R., and Winter, S.G. (Eds), 2000. *The Nature and Dynamics of Organizational Capabilities.* Oxford: Oxford University Press.

Gaulier, G., and Zignago, S., 2010. BACI: International trade database at the product level. The 1994-2007 version. *CEPII Working Paper*, No. 2010-23. Paris: Centre d'Etudes Prospectives et d'Informations Internationales.

Glaeser, E.L., La Porta, R., Lopez-de-Silanes, F., and Shleifer, A., 2004. Do institutions cause growth? *Journal of Economic Growth,* 9: 271-303.

Hausmann, R., and Hidalgo, C.A., 2010. Country diversification, product ubiquity, and economic divergence. *CID Working Paper*, No. 201. Cambridge, MA: Center for International Development, Harvard University.

Hausmann, R., and Klinger, B., 2007. The structure of the product space and the evolution of comparative advantage. *CID Working Paper*, No. 146. Cambridge, MA: Center for International Development, Harvard University.

Hidalgo, C.A., and Hausmann, R., 2009. The building blocks of economic complexity. *Proceedings of the National Academy of Sciences*, 106: 10570-10575.

Hidalgo, C.A., Klinger, B., Barabàsi, A.L., and Hausmann, R., 2007. The product space conditions the development of nations. *Science*, 317: 482-487.

Klepper, S., 2002. The capabilities of new firms and the evolution of the US automobile industry. *Industrial and Corporate Change,* 11: 645-666.

Klepper, S., and Simons, K.L., 2000. Dominance by birthright: Entry of prior radio producers and competitive ramifications in the U.S. television receiver industry. *Strategic Management Journal,* 21: 997-1016.

Lileeva, A., and Trefler, D., 2010. Improved access to foreign markets raises plant-level productivity... for some plants. *The Quarterly Journal of Economics,* 125: 1051-1099.

Lipsey, R.G., 2009. Economic growth related to mutually interdependent institutions and technology. *Journal of Institutional Economics,* 5: 259-288.

Lundvall, B.A., 1992. *National Systems of Innovation: Towards a theory of Innovation and Interactive Learning.* London: Pinter.

Malerba, F., 2002. Sectoral systems of innovation and production. *Research Policy,* 31: 247-264.

North, D.C., 1990. *Institutions, Institutional Change and Economic Performance*. New York: Cambridge University Press.

Redding, S.J., 2010. Theories of heterogeneous firms and trade. *NBER Working Paper,* 16562. Cambridge, MA: National Bureau of Economic Research.

Rodrik, D., 1996. Understanding economic policy reform. *Journal of Economic Literature,* 34: 9-41.

Silva, A., Afonso, O., and Africano, A.P., 2012. Learning-by-Exporting: What we know and what we would like to know. *The International Trade Journal,* 26: 255-288.

Singh, T., 2010. Does international trade cause economic growth? A survey. *The World Economy,* 33: 1517-1564.

Smith, A., 1776. *An Inquiry into the Nature and Causes of the Wealth of Nations.* London: Strahan and Cadell.

Wagner, J., 2007. Exports and productivity: A survey of the evidence from firm-level data. *The World Economy,* 30: 60-82.

Wagner, J., 2012. International trade and firm performance: A survey of empirical studies since 2006. *Review of World Economics,* 148: 235-267.

Winter, S.G., 2003. Understanding dynamic capabilities. *Strategic Management Journal,* 24: 991-995.

von Tunzelmann, N., 2003. Historical coevolution of governance and technology in the industrial revolutions. *Structural Change and Economic Dynamics,* 14: 365-384.

Young, A.,1928. Increasing returns and economic progress. *The Economic Journal,* 38: 527-542.

Table 1

Determinants of Having a Comparative Advantage in the Future

	(1)	(2)	(3)	(4)
Model	LPM	LPM	LPM	LPM
D.V.	CA_{t+5}	CA_{t+5}	CA_{t+3}	CA_{t+3}
CA_t	0.556**	0.556**	0.62**	0.621**
	(0.0169)	(0.017)	(0.0161)	(0.0161)
Density	0.199**	0.187**	0.172**	0.167**
	(0.0157)	(0.0153)	(0.0138)	(0.0143)
Import Density	0.031**	0.044**	0.015*	0.022**
	(0.012)	(0.0127)	(0.0075)	(0.0084)
ENP * Density		0.04*		0.011
		(0.0192)		(0.0162)
ENP * Import Density		-0.042**		-0.024*
		(0.015)		(0.01)
Observations	156,366	156,366	260,568	260,568
R-squared	0.5189	0.5191	0.5752	0.5752

Country-clustered standard errors in parentheses.
All models include country-year and product-year dummy variables.
**,*,+ statistically significant at .01, .05 and .10 percent respectively.

Table 2

Determinants of Having a Comparative Advantage in the Future:
Keeping a Current Product and Developing a New Product

	(1)	(2)	(3)	(4)
Model	LPM	LPM	LPM	LPM
D.V.	CA_{t+5}	CA_{t+5}	CA_{t+3}	CA_{t+3}
CA_t	0.531**	0.527**	0.593**	0.597**
	(0.0157)	(0.0182)	(0.0148)	(0.0177)
Density on Current	0.202**	0.177**	0.176**	0.153**
	(0.0181)	(0.0172)	(0.0158)	(0.0147)
Density on New	0.168**	0.129**	0.139**	0.108**
	(0.0148)	(0.0118)	(0.0125)	(0.01)
Import Density on Current	0.064**	0.082**	0.05**	0.058**
	(0.0158)	(0.0145)	(0.0119)	(0.0116)
Import Density on New	0.031*	0.051**	0.014+	0.029**
	(0.013)	(0.0114)	(0.0083)	(0.0072)
ENP * Density on Current		0.118**		0.116**
		(0.0404)		(0.0403)
ENP * Density on New		0.123**		0.094**
		(0.0199)		(0.0174)
ENP * Imp. Density on Current		-0.041		-0.011
		(0.0347)		(0.0292)
ENP * Imp. Density on New		-0.057**		-0.038**
		(0.0158)		(0.011)
Observations	156,366	156,366	260,568	260,568
Adjusted R-squared	0.5214	0.5223	0.5781	0.5788

Country-clustered standard errors in parentheses.
All models include country-year and product-year dummy variables.
**,*,+ statistically significant at .01, .05 and .10 percent respectively.

Table 3

The effects of most preferred countries and weighted densities.
5-years Interval

	(1)	(2)	(3)	(4)	(5)
Model	LPM	LPM	LPM	LPM	LPM
D.V.	CA_{t+5}	CA_{t+5}	CA_{t+5}	CA_{t+5}	CA_{t+5}
		Most Preferred Densities		Weighted Densities	
CA_t	0.527**	0.534**	0.535**	0.471**	0.461**
	(0.0182)	(0.0134)	(0.0147)	(0.0129)	(0.0155)
Density on Current	0.177**	0.186**	0.162**	0.192**	0.168**
	(0.0172)	(0.0172)	(0.016)	(0.0162)	(0.0157)
Density on New	0.129**	0.167**	0.133**	0.166**	0.129**
	(0.0118)	(0.0143)	(0.011)	(0.0138)	(0.011)
Import Density on Current	0.082**	0.053**	0.068**	0.045**	0.061**
	(0.0145)	(0.0153)	(0.0132)	(0.0148)	(0.0134)
Import Density on New	0.051**	0.033**	0.053**	0.03*	0.049**
	(0.0114)	(0.0127)	(0.012)	(0.013)	(0.0114)
Imp. Density of Destination		0.037**	0.026**	0.019**	0.026**
on Current		(0.007)	(0.0095)	(0.0055)	(0.007)
Imp. Density of Destination		0.004	0.001	-0.002+	-0.005*
on New		(0.0052)	(0.0073)	(0.0013)	(0.0021)
Density of Origin		0.028**	0.031**	0.136**	0.141**
on Current		(0.0097)	(0.009)	(0.0109)	(0.0162)
Density of Origin		0.003	-0.002	0.009**	0.009**
on New		(0.0069)	(0.0064)	(0.0019)	(0.0023)
ENP * Density on Current	0.118**		0.137**		0.115**
	(0.0404)		(0.0393)		(0.0346)
ENP * Density on New	0.123**		0.113**		0.127**
	(0.0199)		(0.0203)		(0.0202)
ENP * Imp. Density on Current	-0.041		-0.038		-0.041
	(0.0347)		(0.035)		(0.0293)
ENP * Imp. Density on New	-0.057**		-0.054**		-0.053**
	(0.0158)		(0.0166)		(0.0156)
ENP * Imp. Density			0.016		-0.014
of Destination on Current			(0.0143)		(0.0105)
ENP * Imp. Density			0.003		0.006*
of Destination on New			(0.0079)		(0.0027)
ENP * Density of Origin			-0.016		-0.008
on Current			(0.0125)		(0.0257)
ENP * Density of Origin			0.004		-0.002
on New			(0.007)		(0.0029)
Observations	156,366	156,366	156,366	156,366	156,366
Adjusted R-squared	0.5223	0.5226	0.5236	0.5244	0.5254

Country-clustered standard errors in parentheses.
All models include country-year and product-year dummy variables.
**.*,+ statistically significant at .01, .05 and .10 percent respectively.

Table 4

The effects of most preferred countries and weighted densities.
3-years Interval

	(1)	(2)	(3)	(4)	(5)
Model	LPM	LPM	LPM	LPM	LPM
D.V.	CA_{t+3}	CA_{t+3}	CA_{t+3}	CA_{t+3}	CA_{t+3}
		Most Preferred Densities		Weighted Densities	
CA_t	0.597**	0.596**	0.6**	0.53**	0.526**
	(0.0177)	(0.0129)	(0.0151)	(0.0105)	(0.0139)
Density on Current	0.153**	0.163**	0.142**	0.167**	0.144**
	(0.0147)	(0.015)	(0.014)	(0.0139)	(0.0134)
Density on New	0.108**	0.139**	0.113**	0.138**	0.109**
	(0.01)	(0.0118)	(0.0096)	(0.0114)	(0.009)
Import Density on Current	0.058**	0.039**	0.051**	0.03**	0.039**
	(0.0116)	(0.0107)	(0.0102)	(0.0109)	(0.0114)
Import Density on New	0.029**	0.016*	0.034**	0.014^{+}	0.027**
	(0.0072)	(0.0081)	(0.0077)	(0.0081)	(0.0067)
Imp. Density of Destination		0.028**	0.02**	0.013**	0.021**
on Current		(0.0056)	(0.0064)	(0.0048)	(0.0062)
Imp. Density of Destination		-0.002	-0.008	-0.002*	-0.006**
on New		(0.004)	(0.0061)	(0.0011)	(0.0017)
Density of Origin		0.023**	0.022**	0.145**	0.151**
on Current		(0.0087)	(0.0077)	(0.0106)	(0.0152)
Density of Origin		0.004	-0.003	0.008**	0.007**
on New		(0.0049)	(0.0045)	(0.0016)	(0.0021)
ENP * Density on Current	0.116**		0.131**		0.112**
	(0.0403)		(0.0396)		(0.0325)
ENP * Density on New	0.094**		0.09**		0.098**
	(0.0174)		(0.0165)		(0.0159)
ENP * Imp. Density on Current	-0.011		-0.021		-0.019
	(0.0292)		(0.0284)		(0.0252)
ENP * Imp. Density on New	-0.038**		-0.044**		-0.035**
	(0.011)		(0.0122)		(0.0105)
ENP * Imp. Density			0.012		-0.016
of Destination on Current			(0.0109)		(0.0095)
ENP * Imp. Density			0.01		0.007**
of Destination on New			(0.0064)		(0.0021)
ENP * Density of Origin			-0.002		-0.014
on Current			(0.012)		(0.0255)
ENP * Density of Origin			0.011*		-0.0002
on New			(0.0056)		(0.0025)
Observations	260,568	260,568	260,568	260,568	260,568
Adjusted R-squared	0.5788	0.579	0.5797	0.5813	0.582

Country-clustered standard errors in parentheses.
All models include country-year and product-year dummy variables.
**.*,+ statistically significant at .01, .05 and .10 percent respectively.

Table A1

Summary Statistics and Correlations

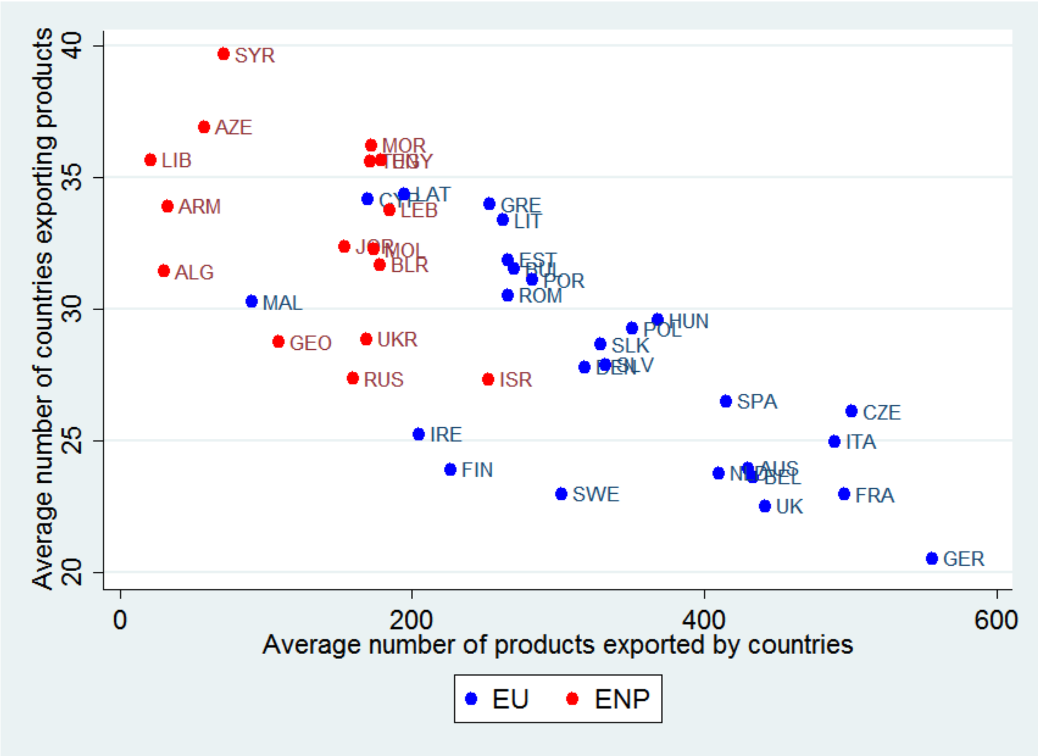
5-years Interval

Variable	Mean	S.D.	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
CA_t (1)	0.216	0.411	1							
CA_{t+5} (2)	0.222	0.416	0.687	1						
Density (3)	0.232	0.133	0.449	0.403	1					
Import Density (4)	0.363	0.09	0.279	0.262	0.722	1				
Import Density of Destination - MPC (5)	0.426	0.055	0.189	0.185	0.423	0.382	1			
Density of Origin - MPC (6)	0.418	0.151	0.116	0.123	0.29	0.325	0.266	1		
Import Density of Destination - WD (7)	0.373	0.141	0.106	0.112	0.289	0.366	0.263	0.39	1	
Density of Origin - WD (8)	0.319	0.174	0.296	0.268	0.521	0.519	0.325	0.183	0.382	1

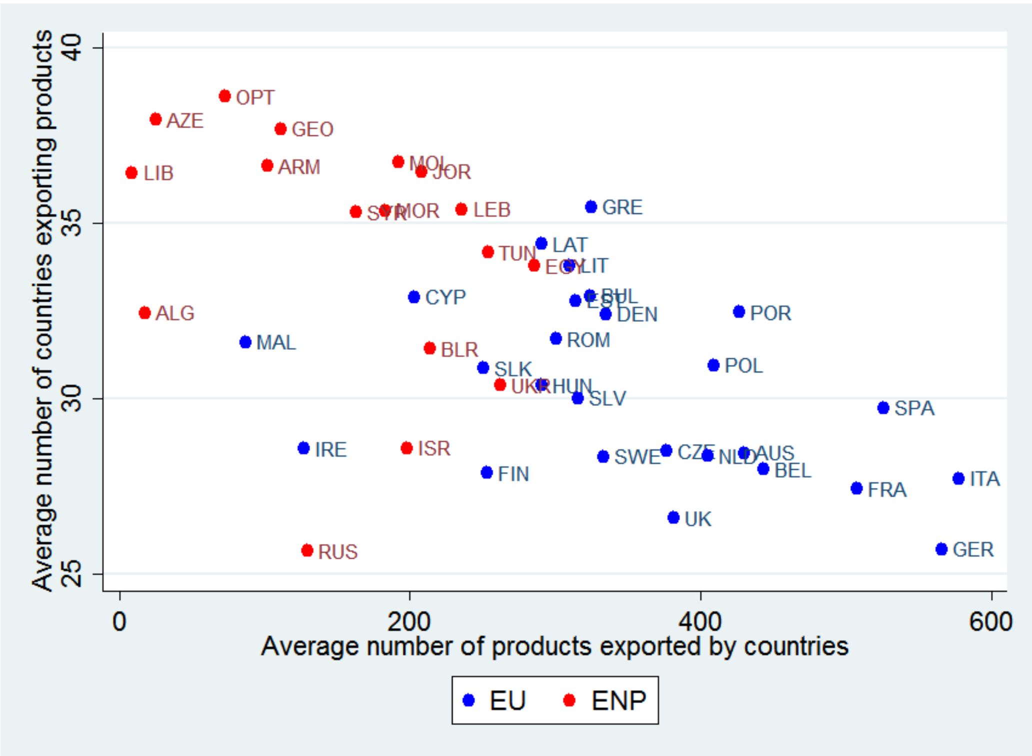
3-years Interval

Variable	Mean	S.D.	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
CA_t (1)	0.216	0.412	1							
CA_{t+3} (2)	0.22	0.414	0.735	1						
Density (3)	0.233	0.13	0.448	0.412	1					
Import Density (4)	0.361	0.087	0.283	0.27	0.724	1				
Import Density of Destination - MPC (5)	0.42	0.054	0.192	0.192	0.439	0.42	1			
Density of Origin - MPC (6)	0.411	0.149	0.117	0.123	0.295	0.351	0.292	1		
Import Density of Destination - WD (7)	0.369	0.136	0.108	0.111	0.281	0.349	0.267	0.419	1	
Density of Origin - WD (8)	0.318	0.171	0.293	0.271	0.52	0.519	0.328	0.205	0.386	1

Figure 1. Relation between the level of diversification of countries and the level of complexity of the products in which they have a comparative advantage, with t = 1995 (a) and t = 2010 (b)

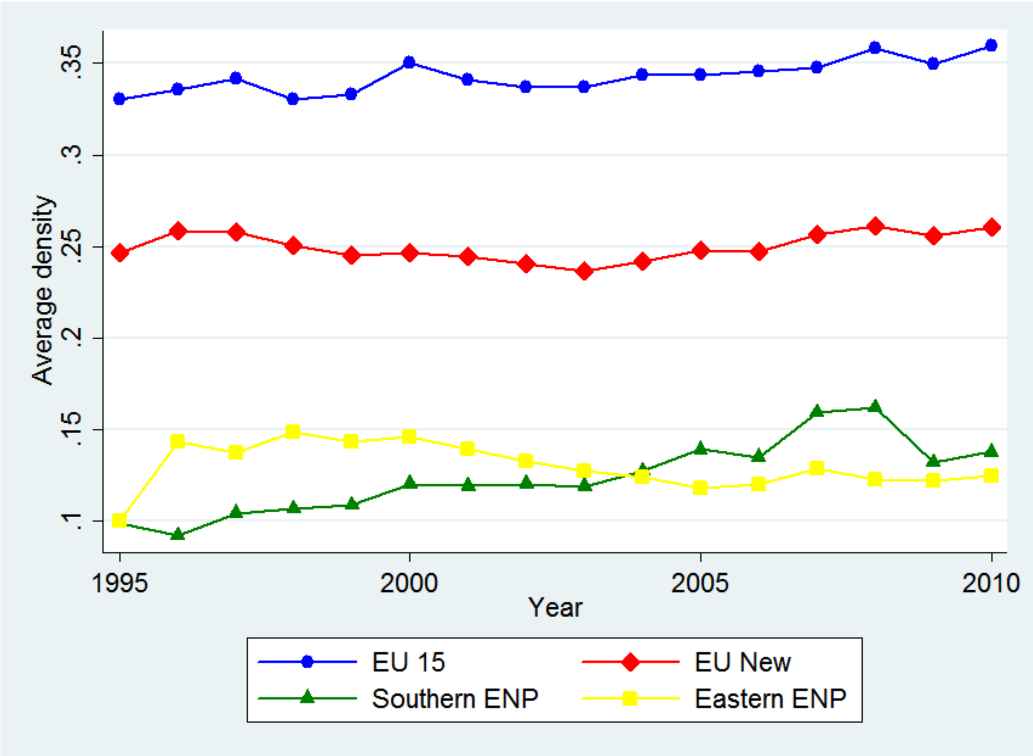


(a)

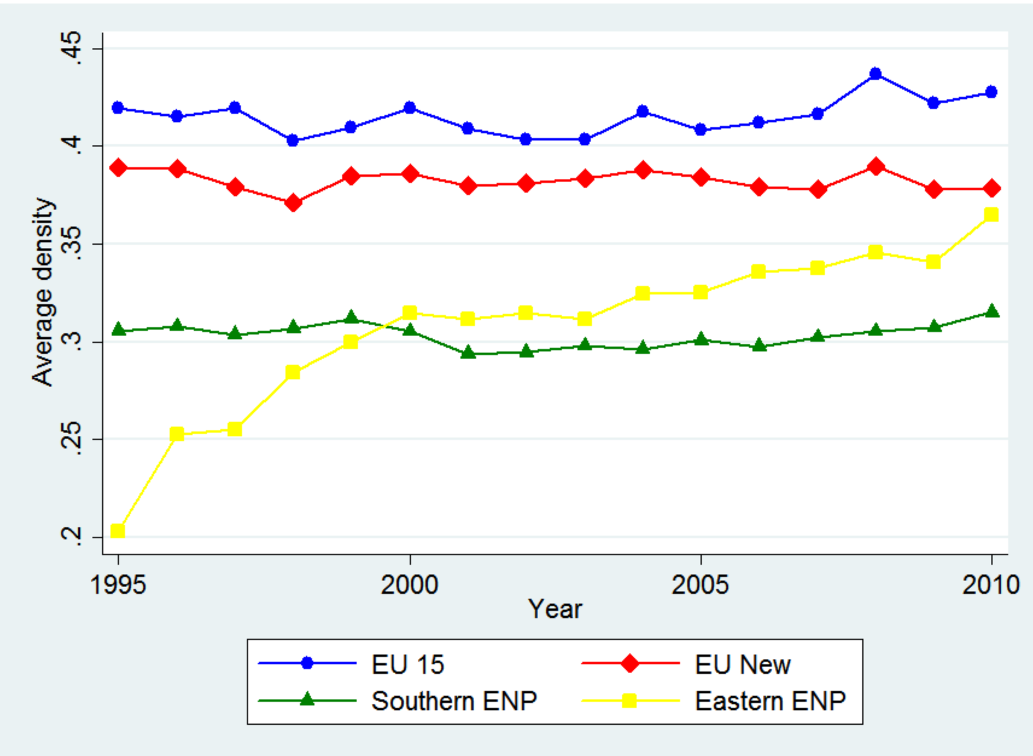


(b)

Figure 2. Evolution of density (a) and import density (b) for EU 15, new EU countries, southern ENP countries and eastern ENP countries.

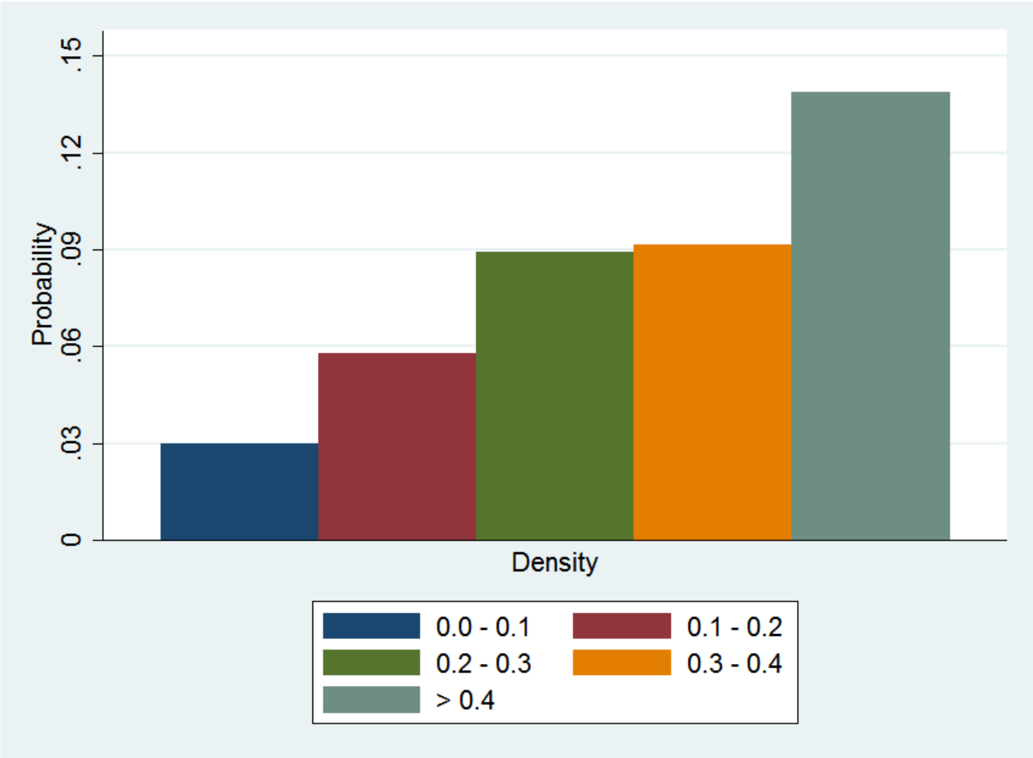


(a)

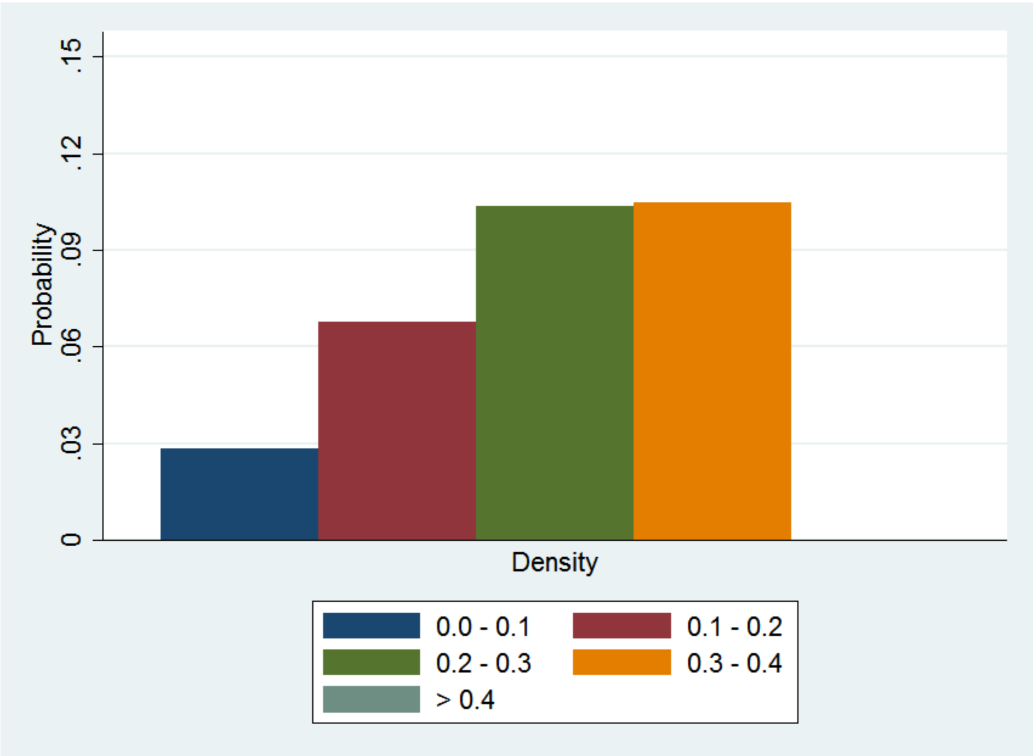


(b)

Figure 3. Probability of developing a comparative advantage in a new product (five years later) for different density ranges in the EU (a) and ENP (b) countries.

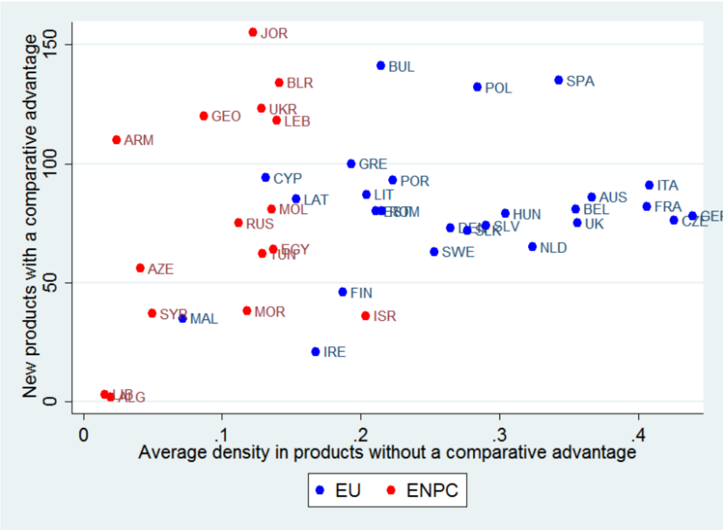


(a)

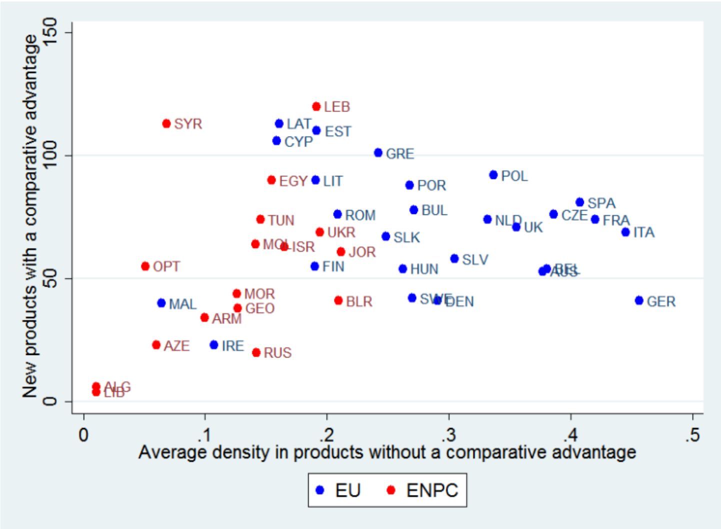


(b)

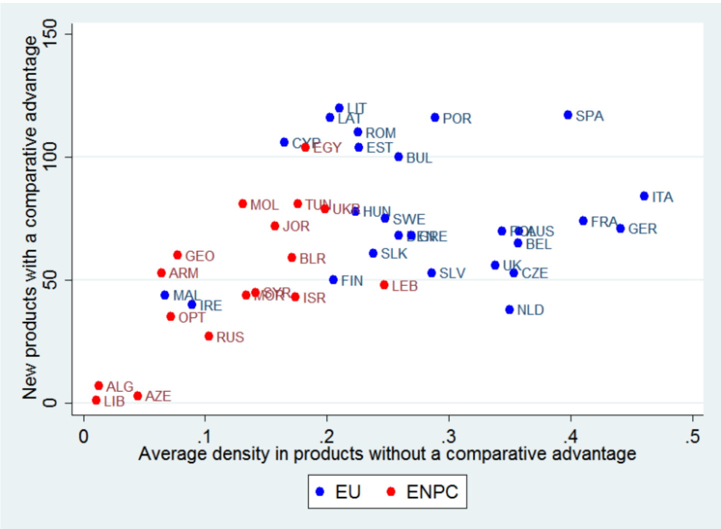
Figure 4. The relationship between density in products without a comparative advantage at time *t* and new products with a comparative advantage at time *t* + 5 in EU and ENP countries, with t = 1995 (a), t = 2000 (b), t = 2005 (c).



(a)

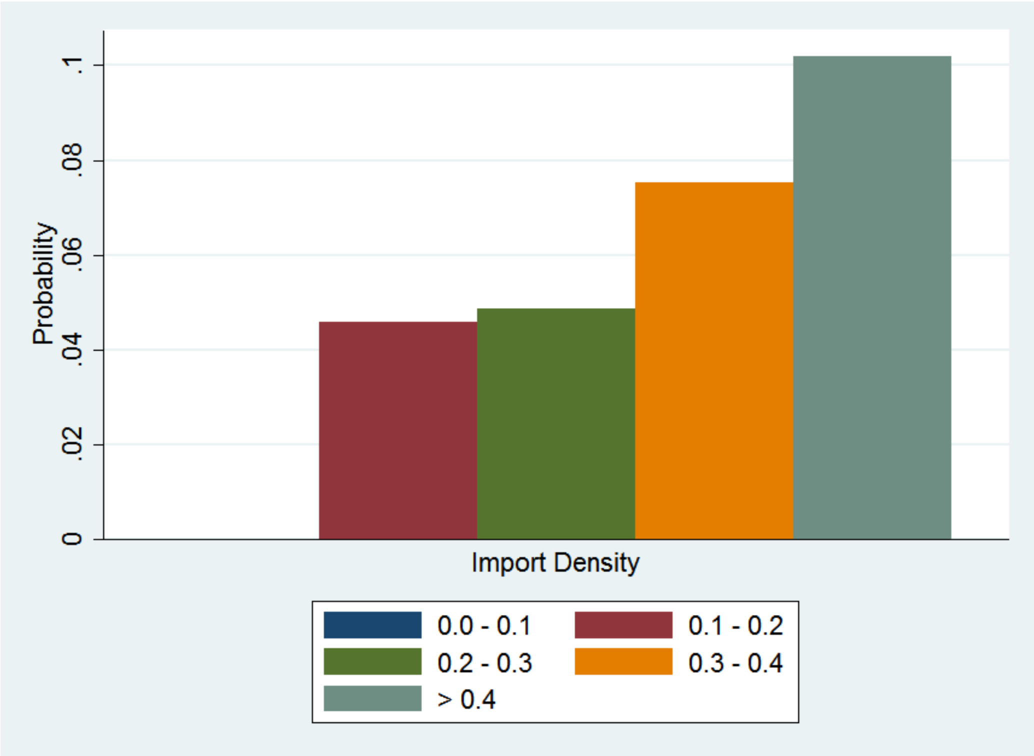


(b)

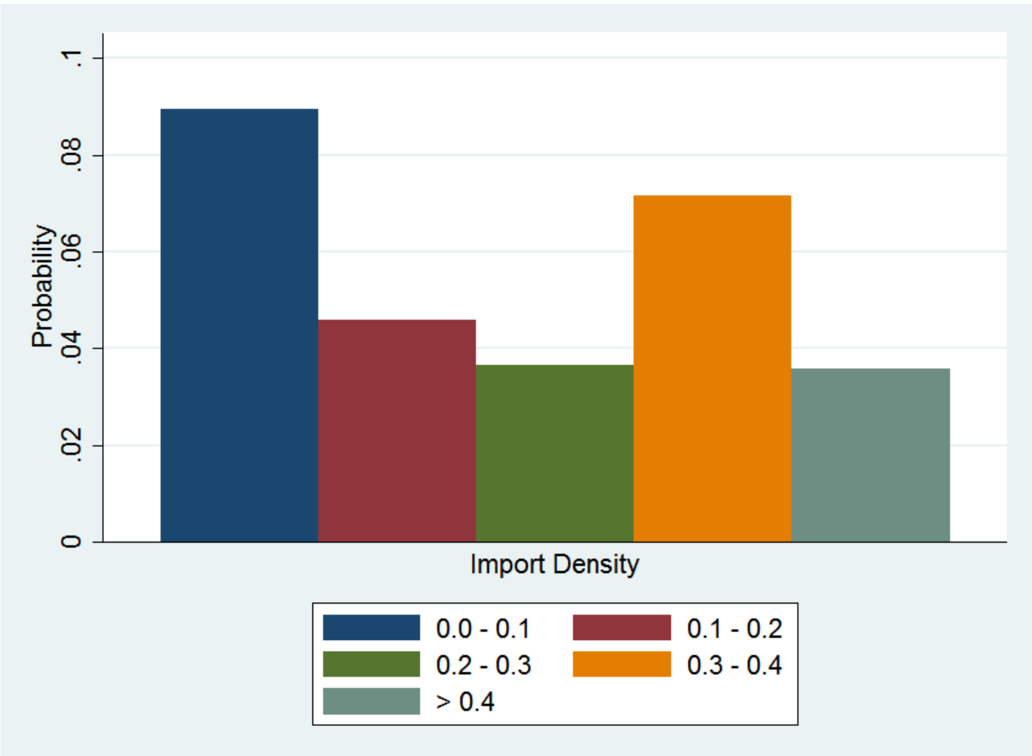


(c)

Figure 5. Probability of developing a comparative advantage in a new product (five years later) for different import density ranges in the EU (a) and ENP (b) countries.



(a)



(b)