Understanding European integration with bipartite networks of comparative advantage

Riccardo Di Clemente [ID][a,b,*], Balázs Lengyel [ID][c,d,*], Lars F. Andersson[e] and Rikard Eriksson [ID][f]

[a]Department of Computer Science, Exeter University, Exeter EX4 4PY, United Kingdom
[b]The Alan Turing Institute, London NW12DB, United Kingdom
[c]Agglomeration and Social Networks Lendület Research Group, Centre for Economic and Regional Studies, Eötvös Loránd Research Network, Budapest 1097, Hungary
[d]Laboratory for Networks, Technology and Innovation, Corvinus Institute for Advanced Studies, Corvinus University of Budapest, Budapest 1093, Hungary
[e]Department of Economic history, Umeå University, Umeå SE-90187, Sweden
[f]Department of Geography, Umeå University, Umeå SE-90187, Sweden
*To whom correspondence should be addressed: Emails: r.di-clemente@exeter.ac.uk
Edited By: Karen E. Nelson

Abstract

Core objectives of European common market integration are convergence and economic growth, but these are hampered by redundancy, and value chain asymmetries. The challenge is how to harmonize labor division to reach global competitiveness, meanwhile bridging productivity differences across the EU. We develop a bipartite network approach to trace pairwise co-specialization by applying the revealed comparative advantage method within and between the EU15 and Central and Eastern European (CEE). This approach assesses redundancies and the division of labor in the EU at the level of industries and countries. We find significant co-specialization among CEE countries but a diverging specialization between EU15 and CEE. Productivity increases in those CEE industries that have co-specialized with other CEE countries after EU accession, while co-specialization across CEE and EU15 countries is less related to productivity growth. These results show that a division of sectoral specialization can lead to productivity convergence between EU15 and CEE countries.

Keywords: labor, bipartite network, specialization, convergence, EU

Significance statement:

We build on bipartite network methodology to develop new data-driven tools to characterize integration trends of the European common market, across country groups, and on the level of countries and industries. We apply the widely used revealed comparative advantage method and trace pairwise co-specialization within and between the EU15 and Central and Eastern European member states from 2000. This new co-specialization approach can be used to assess redundancies and division in the system, and at the level of industries and countries as well. This latter feature enables us to investigate how co-specialization across countries impacts economic growth. These results inform European policy that a division of sectoral specialization can lead to productivity convergence between EU15 and Central and Eastern European (CEE) member states.

Introduction

The European integration has changed the production landscape of the continent and the development paths of its member states. The removal of European borders following the fall of the Iron Curtain and subsequent EU enlargement opened a free flow of goods and factor endowments between countries. In theory, such integration facilitates the convergence of countries that are at different development stages (1). In turn, others argue that the most and less developed countries and sectors might not equally tackle increasing competition within the common market (2). Some even claim that the emerging new division of labor creates asymmetric value-chain relations in the integrated macro-region, hamper-

ing economic and societal progress in less developed or accession countries (3,4).

A growing literature has therefore focused on the above duality of the EU integration process. Motivated by neoclassic growth theory, economists have investigated whether the unified market fosters economic growth of entrants and if so, whether convergence in income and productivity levels happen on the short or on the long run as a consequence of technological and knowledge spillovers toward less developed countries (5–12). Recent studies focusing on the Eastern enlargement find compelling support for the convergence hypothesis (13–15). Case studies on specific industries have looked into how relocation of production to

accession countries and integration of their industries to European value chains influence development (16,17). Some have even emphasised convergence of economic structures as well by measuring the structural similarity of trade and production between EU member countries (5,6,18,19).

However, labor division across nations in the global value chain is a natural process and a potential area of policy intervention to foster convergence in income by taking advantage of diverging specializations. Starting from the simple Ricardian model, comparative advantage has been considered a major source of national specialization into certain products or industries as international trade intensifies (20,21). New economic geography (NEG) models on European integration suggest diverging structures of specialization as agglomeration forces of industry production grow stronger at lower barriers of trade (22). This process facilitates economic progress due to the accumulation of knowledge and skills through specialization in line with endogenous growth theory (23–25). Dynamic models considering different stages of development suggest that countries tend to be more specialized the higher the levels of their income are (26), thereby questioning whether a deepened labor division actually can favor less developed countries when they integrate with more developed ones.

More than a matter of specialization, also the type of specializations should make a difference for future development. Previous studies have identified that whether nations, for example, co-specialize into identical industries or, on the contrary, specialize into divergent structures is crucial for their development (27,28). This is because specialization into identical products can signal integration into value chains that can help income and productivity convergence (17) but, at the same time, competition across countries as well (29). The latter would suggest that less developed countries can benefit less from co-specializing with more developed countries (30–32). Yet, the structural properties of specialization have remained under-researched, especially in the context of convergence within the EU.

In this paper, we offer a new methodological framework that provides tools to understand both structure and dynamics, and also the consequences of co-specialization across countries. Our case is the European integration and we define a bipartite network, in which EU member states are linked to industries if they have revealed comparative advantage (RCA) in that particular industry (20) in terms of workforce. The unipartite representations of RCA networks have been proven useful to understand the development of countries (27). Here, instead of projecting the bipartite network into countries or sectors (33), we keep the bipartite representation and investigate the abundance of co-specialization motifs (34,35). This approach has already been used to understand systems in ecology (36), trade networks (37), finance, (38) and scientific competition of countries (39) and offers new tools to better understand the dynamics of co-specialization at multiple levels.

We define the co-specialization motif directly on the bipartite network as a subgraph that captures if two countries are specialized in the same industry. By measuring the statistical significance of this motif we can assess whether and where production structures are overlapping or diverging in the EU (35,40,41). Here, we aim to disentangle co-specialization within and across the groups of EU15 and Central and Eastern European (CEE) member states. We can also quantify these metrics on the industry-country level to create a metric that can predict their economic growth. Together these exercises allow us to discuss how labor division and co-specialization across countries influence convergence in productivity.

Our measurement reveals no significant overlaps of comparative advantages across the EU15 countries after 2000, which is probably due to their gradual integration. However, we find that CEE countries tend to be specialized in similar, if not identical, sectors. The number of co-specialization network motifs including both EU15 and CEE countries decrease after enlargement signalling a deeper division of production between EU15 and CEE. We do however find that productivity increases in those CEE industries that had no significant overlap in specialization before EU accession but experience increasing co-specialization with other CEE countries after entering the EU. In the meantime, co-specialization across EU15 and CEE countries contribute slightly positive, if at all significant, to productivity growth after accession. The findings refer to the role of labor division between EU15 and CEE that foster co-specialization of CEE industries and facilitate their integration and convergence in the common market.

Results
Bipartite network approach to european integration

The core of our empirical approach is a bipartite, undirected, and dynamic network between 21 European countries $C = [1, ..., c, ..., 21]$, 31 industry sectors $S = [1, ..., s, ..., 31]$, and over 14 years $t = [2000, ..., t, ..., 2014]$ where the countries and the industries are the two types of nodes in each year (Fig. 1A). The data for the analysis have been retrieved from the OECD webpage and allow us to assess changes before the 2004 enlargement as well as evaluating the postenlargement changes (see data description in the "Methods" section).

We start from a rectangular matrix $\tilde{M}_t$ with dimensions C and S where each entry $\tilde{M}_{cs,t}$ is the number of employees of country c in industry s and year t. The links between industries and countries are established by applying the RCA index as developed by Balassa (20) over the entries $\tilde{M}_{cs,t}$

$$\text{RCA}_{cs,t} = \frac{\frac{\tilde{M}_{cs,t}}{\sum_{s' \in S} \tilde{M}_{cs',t}}}{\frac{\sum_{c' \in C} \tilde{M}_{c's,t}}{\sum_{c' \in C} \sum_{s' \in S} \tilde{M}_{c's',t}}}. \tag{1}$$

The entries of the adjacency matrix that represents the binary undirected bipartite industry-country network in year t are defined by the rule

$$\begin{cases} M_{cs,t} = 1 & \text{when } \text{RCA}_{cs,t} \geq 1 \\ M_{cs,t} = 0 & \text{otherwise} \end{cases}. \tag{2}$$

Eq. (2) indicates that the industry-country link is established if $\text{RCA}_{cs,t}$ is over the threshold =1. In this case, country c is considered to have comparative advantage in sector s given the specialization in that sector in terms of employment (17). According to theory, such specialization leads to economic and technological progress (42) and thus international trade and economic development literatures have recently focused on the dynamics of RCA (28,43).

Our bipartite network approach allows us to further this research in two major methodological domains:

- dynamics of $\text{RCA}_{cs,t}$ in relation to sectors in other countries;
- the decomposition of $\text{RCA}_{cs,t}$ to groups of countries that are either of similar or of different stages of economic development.
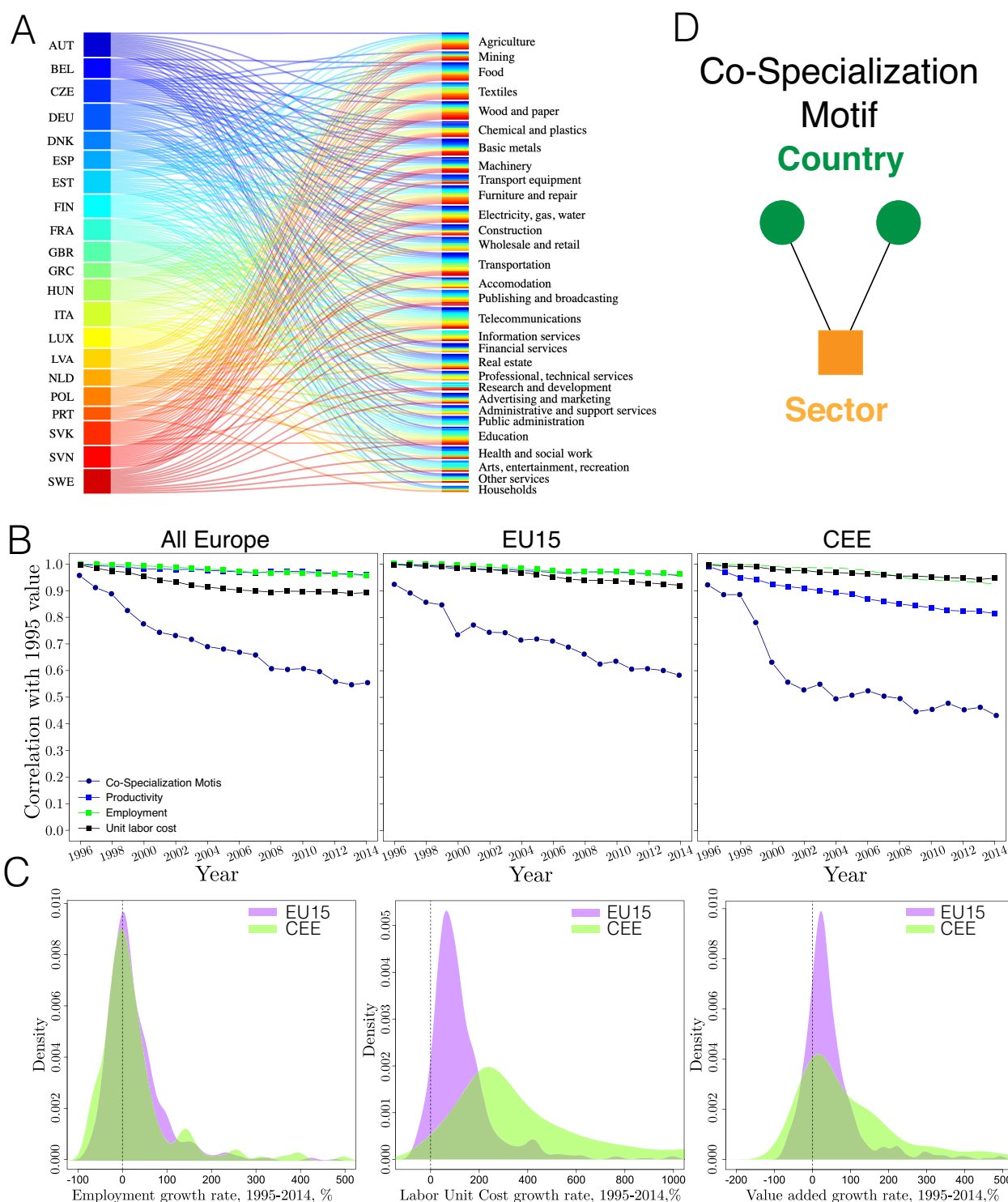
Fig. 1. Variables and their dynamics. (A) Bipartite networks of European countries and industry sectors of 2005. (B) Correlation of country-industry variables with their 1995 value. The drop of coefficients of the co-specialization motif signal dynamic change whereas other measures suggest proportional growth across sectors. (C) Employment is stable in most EU15 and CEE industries and the number of workers grows extensively only in a limited number of sectors. Unit labor cost (average cost of employment per unit of output produced in million Euros) have grown more in CEE sectors than in EU15 sectors. Value added per capita in million Euros has grown slowly in both country groups, but the tail of the distribution is fatter in the CEE. (D) Co-specialization motif across countries that have RCA in identical sectors.

We propose that such methodological progress enables us to characterize the dynamics of co-specialization at multiple levels and thus infer the trends of labor division during European integration and also the role of different types of co-specialization in value-added dynamics (30).

Our approach builds on network motifs, which are patterns that typically consist few nodes of the network and the links between them (34,40). In bipartite networks, motifs carry valuable information that is otherwise hidden in classic network analysis of unipartite networks projected from bipartite networks

(38,44,45). Our focus is on the most simple co-specialization motif in this bipartite country-industry network defined by Eq. (3) and illustrated in Fig. 1D.

The "Co-Specialization Motif" $\mu^{co-spec}$ defines the pairs of countries in a given year that are both specialized in a given industry as $M_{cs,t} = M_{c's,t} = 1$. $\mu^{co-spec}$ captures similarity between countries in having comparative advantage in certain industries (35). To quantify the co-specialization in one or two sectors in a country with the same sectors in other countries, we count the number of $\mu_t^{co-spec}$ for every country-industry observation in each year

$$\mu_t^{co-spec} = \frac{1}{2}\sum_{s=1}^{S} u_{s,t}(u_{s,t} - 1), \tag{3}$$

where, $u_{s,t} = \sum_{c \in C} M_{cs,t}$ is a vector of length C, giving the degree of each columns, it is representing how an industry sector is ubiquitous. Certainly, more complex motifs (35,41) can be quantified to unfold various types of RCA relations across countries and industries offering tools to analyse the intensity of inter-relatedness of specialization.

The indicator introduced in Eq. 3 enables us to capture the dynamics of RCA relations between sectors and across countries and provide new information on economic restructuring during European integration. To illustrate this in Fig. 1B, we correlate $\mu_t^{co-spec}$ with its 1995 values and do the same exercise of employment, unit labor costs, and value added in sectors for the entire EU and decomposed to EU15 and CEE countries. Stable correlation coefficients of employment and value added in EU15 industries denote a rather stable structure of proportional growth. Together with the observation in Fig. 1C we find that most EU15 industries have not or have only slightly grown over 1995 to 2014. There is however a small deviation in terms of unit labor costs starting from 2004, the entry year of the accession countries denoting a restructuring shock of accession to EU15 wages. In the CEE countries, the growth of employment and unit labor costs are proportional despite the immense growth rate of the latter. The correlation coefficient of value added in CEE decreases monotonically but remains high as well, denoting that ranks of CEE sectors in terms of productivity have slightly changed over the period.

In contrast, the correlation coefficient of $\mu_t^{co-spec}$ is decreasing monotonically and sharply suggesting that RCA relations are changing much more over time than other indicators frequently used to understand EU integration (see Fig. 1C). We find that co-specialization of the CEE countries were experiencing great internal turbulences before the entry, but less so after the entry in 2004. See the Materials and Methods for further explanation of our bipartite network approach.

Co-specialization and labor division revealed by motif significance

To assess whether network motifs carry significant information about co-specialization of countries into industries, we compare the occurrence of observed motifs to a null model for each year using the Bipartite Configuration Model (40). This is done by calculating the z-score of $\mu_{cs,t}^{co-spec}$ by:

$$z_{cs,t}^{\mu} = \frac{\mu_{cs,t}^{co-spec} - \langle\mu\rangle}{\sigma_\mu} \tag{4}$$

where $\langle\mu\rangle$ is the average of $\mu_{cs,t}^{co-spec}$ in a set of 10,000 random matrices that persevere the average degree sequence of nodes in $M_{cs,t}$ and σ_μ is the SD of these randomized distributions. Values of $z_{cs,t}^{\mu}$ around zero mean that the observed motif distribution does not

differ significantly from the null model. Positive high values capture significant co-specialization; while negative low values denote dissimilar RCA structures across countries. For a detailed description of the bipartite configuration model, see the "Methods" section.

Tracing $z_{cs,t}^{\mu}$ over time allows us to infer on co-specialization trends during the EU enlargement process. We label countries in the bipartite network CEE (accession countries) and EU15. Fig. 2A demonstrates considerable dynamics between 2000 and 2014, in which CEE industries have a growing share of RCA links (for example in Basic Metals, Transport Equipment, or Telecommunications). To quantify co-specialization dynamics within and across these country groups, we distinguish three cases of co-specialization motif that are illustrated in Fig. 2B. Internal EU15 $\mu_{c \in EU15, t}$ denotes motifs with two EU15 countries specialized in the same sector while Internal CEE $\mu_{c \in CEE, t}$ captures specialization of two CEE countries in the same sector. External $\mu_{EXT, t}$ stands for motifs that capture specialization of one EU15 country and one CEE country in the same industry. We can decompose the abundance of the co-specialization motif in

$$\mu_{EXT,t} = \mu_t - (\mu_{c \in EU15,t} + \mu_{c \in CEE,t})$$
$$= \sum_s \sum_{c' \in EU15} \sum_{c'' \in CEE} M_{c's,t} M_{c''s,t}. \tag{5}$$

Fig. 2C reports on $z_{cs,t}^{\mu}$ trends over EU enlargement. The z-score measured from the full matrix containing all countries is stable around -0.5. Restricting the countries to EU15 produces similar z-scores that even converge to zero. These findings mean that co-specialization across all EU countries and across EU15 countries is not significantly different from the null model.

However, both internal CEE co-specialization and external co-specialization are significantly different from the null model and reveal interesting trends. Internal CEE has high positive values that decrease over time, especially after 2008. There are two potential reasons for the drop of co-specialization after the financial crisis. First, the unequal impact of the crisis across member states, and second, the different paths of recovery during its aftermath (46–48). External co-specialization z-scores have low negative values meaning that EU15 and CEE countries tend to have comparative advantages in distinct industries. The z-score becomes significant in 2006 signalling that EU accession has been followed by a stronger division of production across old and new member states. Supplementary Material 1 (Fig. S1.1) contains the z-score trends decomposed to sectors, which reveal different dynamics. For example, there is a significant co-specialization in primary production of the CEE countries, but this sector has a nonoverlapping RCA structure in EU15 countries. We find that EU15 and CEE countries tend to not specialize in the same service sectors.

Impact on productivity

In this section, we give an example of the industry-country level application of our framework. By rewriting Eq. 3, it is possible to study the co-specialization of a sector in a country. The significance level of the co-specialization motif of industry s in country c changes in case the of number of industries in other countries that the given sector in a given country is co-specialized with increases or decreases. This enables us to analyse the relation between the dynamics of co-specialization motifs and productivity of industry-country pairs.

The industry-country level co-specialization is defined by

$$\mu_{cs,t}^{co-spec} = (u_{s,t} - 1)M_{cs,t}. \tag{6}$$
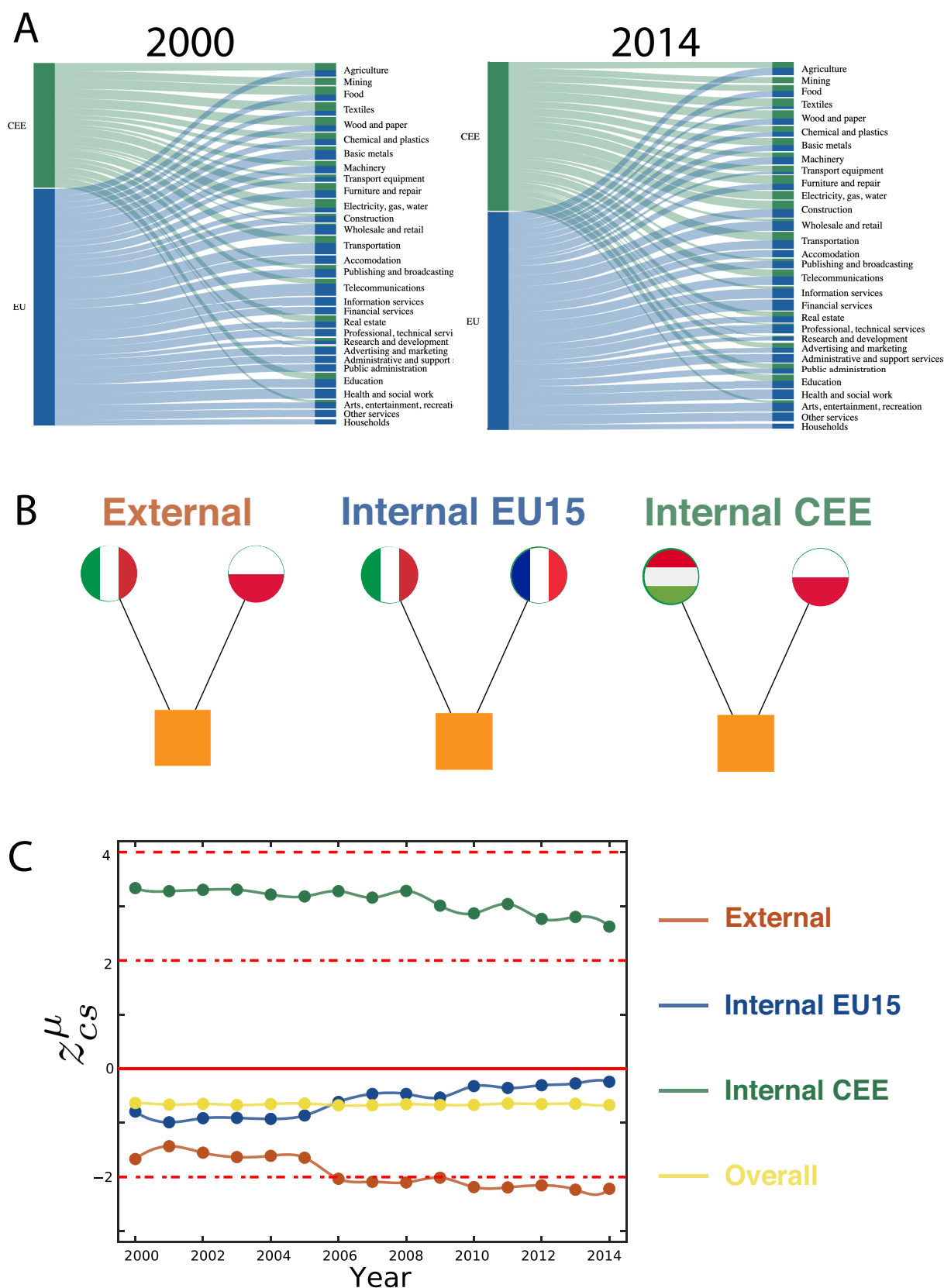
Fig. 2. Co-specialization within and between country groups. (A) The share of CEE countries in the bipartite RCA network has grown from 2000 to 2014. (B) Illustration of internal and external co-specialization motifs. The exemplar external Motif is illustrated with a link between Italy (EU15) and Poland (CEE); the Internal EU15 is represented by a link between Italy and France and the internal CEE with a link between Hungary and Poland. (C) Trends of co-specialization motifs compared to a null model of co-specialization, quantified by the z-score, illustrate that CEE countries co-specialize in identical industries but EU15 countries do not. The negative z-score of External co-specialization motifs denote an intensifying labor-division between EU15 and CEE that becomes significantly different from the null model after 2006.
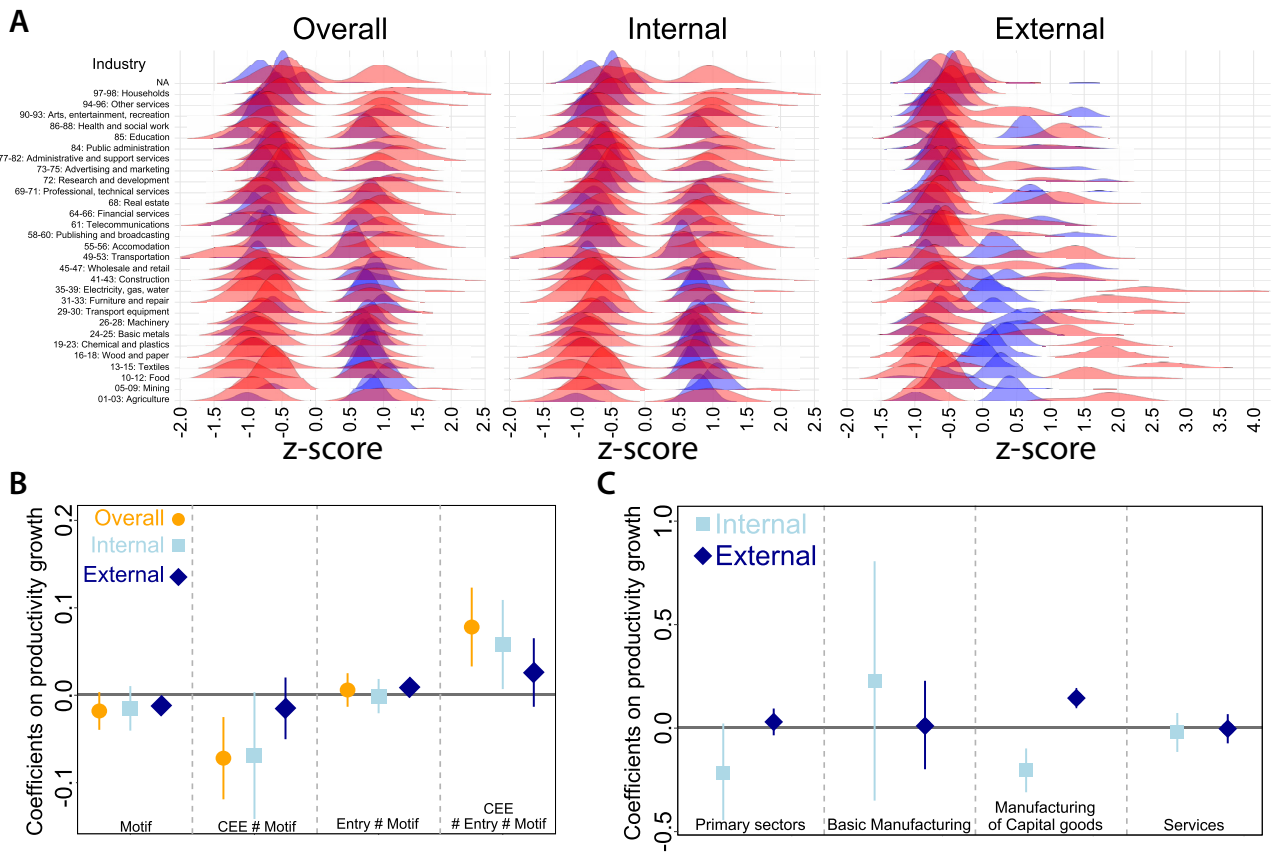
Fig. 3. Co-specialization motifs and value-added growth. (A) z-scores of co-specialization are plotted over the 2000 to 2014 period. Blue depicts z-scores in CEE industries and red is z-scores in EU15 industries. (B) The role of co-specialization motifs and their effect in CEE countries on value-added growth after entering the EU. (C) The role of External and Internal motifs for value-added growth by industries during EU integration.

Following the logic of Eq. 5, we decompose $\mu_{cs,t}^{co-spec}$ to $\mu_{cs,t}^{EXT}$, $\mu_{cs,t}^{CEE}$, and $\mu_{cs,t}^{EU15}$. For each country groups CEE and EU15, we define the INTERNAL indicator that includes co-specialization within CEE and EU15 groups of countries ($\mu_{cs,t}^{CEE}$, $\mu_{cs,t}^{EU15}$). The EXTERNAL indicator includes co-specialization across CEE and EU15 groups ($\mu_{cs,t}^{EXT}$). From now on, we refer to $\mu_{cs,t}^{co-spec}$ as the OVERALL Motif of co-specialization.

Fig. 3A illustrates the distributions of OVERALL, INTERNAL, and EXTERNAL Motifs by sectors. The figure shows that CEE countries (in blue) tend to co-specialize in manufacturing sectors while EU15 countries (in red) tend to co-specialize in service sectors. Decomposing OVERALL similarity to EXTERNAL and INTERNAL, we find that CEE countries co-specialize with other CEE countries in manufacturing sectors to a high degree and EU15 countries co-specialize with other EU15 countries in services but in a slightly lower degree. Co-specialization across EU15 and CEE countries mostly appears for CEE industries and mostly in manufacturing industries. However, in few cases, this is highly significant for EU15 industries as well.

To estimate the impact of the above described changes in motifs on the performance (value added per full-time equivalents) of each industry in every country, we have employed a difference-in-differences (49) approach in a fixed effect panel model for the period 2000 to 2014 that is specified by

$$y_{i,t} = \beta' X_{i,t-1} + \gamma' Z_{i,t-1} + \lambda_i + \epsilon_{i,t}, \qquad (7)$$

where y denotes value added per capita, t denotes 1-year intervals from 2000 to 2014, i denotes the pair sector-country $i = (c, s)$,

X represents a vector of co-specialization variables, λ_i denotes the industry-country fixed effects, Z stands for a set of control variables, and ϵ is the case- and time-specific error term. The definition, descriptive statistics and correlation values of the indicators can be found in Supplementary Material 2 (Tables S2.1, S2.2, and S2.3) while Supplementary Material 3 explains the motivation for this specification in detail.

This regression technique allows us to assess how changes in motifs influence value added in relation to (i) the entry of CEE in the EU, and (ii) whether such changes occur in CEE or in EU15. This is important to distinguish because some capital adjustment is likely to occur as new sites of production and new markets become available within the common market. Hence, a dummy variable (Entry) that equals one for the period 2004 and onwards was created to capture the role of the enlargement as such. Moreover, another dummy variable (CEE) that equals 1 if a given part of the CEE countries joining the EU in 2004 (i.e., Czech Republic, Estonia, Hungary, Latvia, Poland, Slovakia, or Slovenia). Finally, two controllers capturing the size of the industry (EMP) and gross capital formation (GFC) are added since value added might be driven by labor saving technologies. All continuous variables are standardized to ease interpretation.

The results from the regressions on the full models that include interaction terms are displayed in Fig. 3B. The first results on the OVERALL Motif indicate that the main effect of co-specialization does not significantly influence value-added growth. In fact, there is a negative overall influence of increasing EU-wide co-specialization in the accession countries before the entry, but that turns positive after the entry. In other words,

increasing postentry co-specialization seems to primarily benefit the CEE countries. This would indicate that, if anything, an increasing convergence in overall EU co-specialization is beneficial for growth.

However, when decomposing this co-specialization into the INTERNAL and EXTERNAL terms (CEE and EU15, respectively), our results indicate a somewhat more complex pattern. First, while the main effect of increasing internal co-specialization is insignificant, the corresponding estimate on increasing external co-specialization is significantly negative. Due to the corresponding interaction effects this implies that this negative association mainly is identified in the EU15 before the entry and then slightly moderated after the entry. Concerning the CEE countries, our findings on internal co-specialization suggest that before the entry increasing co-specialization within the CEE hampers growth, while after the entry increasing internal co-specialization in the CEE explain the overall positive association with growth in the CEE. Hence, it is not convergence within the EU that contribute to growth, but rather convergence within the accession countries after the entry. As is displayed in Fig. 3C, where we decompose INTERNAL and EXTERNAL terms in different groups of sectors, the impact on productivity growth is positively associated with internal co-specialization in basic manufacturing and external co-specialization in manufacturing in capital goods. While the division of labour between CEE and EU15 in basic manufacturing largely impact on productivity growth, capital goods are not affected by the entry as such. Regression tables and robustness checks are presented in Supplementary Material 4 (Table S4.4).

Discussion

Altogether, our findings imply a deeper division of labor across diversified EU15 and specialized CEE countries. This division follows from the principle of comparative advantages, where EU15 and CEE countries specialize in producing goods and services at a lower opportunity cost to trade. While NEG models assume that as the EU become more integrated, economic activity become more regionally concentrated and specialized (22), recent empirical studies fail to find consistent support for this argument (50). Our findings on the structural changes of specialization indicate that the convergence effects of this evolution depends on the pattern of specialization. Thus, with whom's structure the accession countries converge toward. The regression results confirm that productivity increases in those CEE industries that converge in specialization with other CEE countries, not with other EU members, while co-specialization across CEE and EU15 countries is less related to productivity. If anything, co-specialization across EU-15 and CEE is more beneficial for the EU15 which suggest difficulties for CEE countries to specialize in similar activities as EU15 countries after the accession.

Since these two country groups are known to be at different stages of development (51) this indicates two different processes of EU integration. On the one hand, integration has made offshoring of certain stages of the value chain possible, which has increased internal CEE co-specialization and growth. This may also reflect the fact that countries that have similar production endowments tend to trade with similar but related products to utilize returns to scale by specializing in different tasks (52). On the other hand, in contrast to the intersectoral convergence previously identified for Western European countries as lagging-behind countries has shifted from industrialized to service economies (5), our findings suggest that the differences between the EU15 and the CEE countries instead have increased over time. This, in turn,

suggests greater differences of sector specializations and thus a persistent and even intensifying labor division between the EU15 and the CEE countries. Based on the decomposed estimates, this could be explained by the fact that basic manufacturing is expanding across all CEE countries after the accession, which increases both co-specialization within CEE and productivity. Given the negative estimate for external co-specialization this pattern is partly driven by relocation of such activities from the EU15. Finally, our findings suggest that the remaining co-specialization mainly benefit the EU15 that already have a productive edge as compared to the accession states.

Conclusions

While great emphasis has been put on tackling territorial disparities within the EU over time, persistent differences in employment and output still remains. Since the bulk of present Cohesion Policy funding is concentrated on less developed European countries and regions to help them catching-up, the main objective of this paper was to assess how the structures of employment have changed in the CEE countries when entering the EU. While this is far from a new empirical assessment, the contribution of this paper was the use of a novel bipartite network approach to trace the significance of co-specialization in employment over the European integration 2000 to 2014. We present a new methodology to revealed comparative advantage interactions using the bipartite network approach that enables a more direct understanding of the countries' interaction. Tehereby overcoming the limitations of monopartite approaches.

The methodological contributions offered in this paper regard the decomposition of significant network motifs. This, in turn, enables a quantification of the dynamics of co-specialization at the level of the whole EU, groups of countries, and industry-country pairs compared to more general specialization indices typically applied in the empirical literature. Since this method reveals distinct co-specialization trends across EU member countries, more detailed patterns of convergence and divergence within the EU can be revealed. For example, while the co-specialization among EU15 member states is not significant, the pairwise specialization of CEE countries is significantly higher, and the co-specialization across EU15 and CEE countries is significantly lower than random distributions would suggest.

The bulk of previous studies on convergence either advocates neoclassical growth theory or more endogenous mechanisms. The division of labour at different levels of development, and its relation to productivity growth identified in this paper, however, underlines the interdependence between growth and trade suggested in the NEG models on "new trade theory" (53). The increasing returns to scale associated with co-specialization furthermore underlines the ability to exploit economies of scale and market size suggested by endogenous growth theory (54).

The division of labor between EU15 and CEE is persistent and is seemingly growing, and therefore requires special policy focus. Our findings highlight that the current EU policy on smart specialization (55–57) need to be applied in new member states differently from old member states to reflect labor division that can support convergence in productivity. One of the risk factors this paper touches upon is the rapidly increasing unit labor costs in manufacturing over the past two decades in CEE economies. Although, such a development entails social benefits in terms of higher incomes and potential for tax revenues, such a development may in the longer-run compromise competitiveness and reduce market share for manufacturing output.

Table 1. Country Groups.

Group name	Countries ISO code
CEE	CZE, EST, HUN, LVA, POL, SVK, and SVN
EU15	AUT, BEL, DEU, DNK, ESP, FIN, FRA, GBR, GRC, ITA, LUX, NDL, PRT, and SWE

Table 2. Industries Sectors.

Sectors name	Codes STANI4 2016
Primary production	D01T03, D05T09, D10T12, D13T15, and D16T18
Basic manufacturing	D19T23, D24T25
Manufacturing of capital goods	D26T28, D29T30, and D31T33
Infrastructure	D35T39, D41T43
Retail	D45T47, D49T53
Services	DD55T56, D58, T60, D61, D62T63, D64T66, D68, D69T71, D72, D73T75, and D77T82
Personal services	D84, D85, D86T88, D90T93, D94T96, D97T98, and D99

While the advantage of the proposed method is its' ability to detect signals of co-specialization dynamics from simple data sources, the empirical approach provided here is not without limitations. We focused only on employment structures and its association with productivity, while the exact mechanisms of economic integration can be better captured with more detailed data. Future studies should therefore complement this approach with trade data to assert the potential evolving interdependence between specialization and trade.

Methods
Data
Country grouping is presented in Table 1 and sectors classes are presented in Table 2.

Mathematical appendix
Notation and Bipartite Configuration Model
In this section, we will follow the definitions provided by (35,37,41). All the formulas are defined for the analysis of given year t. The rows of $M_{cs,t}$ correspond to Countries $C = [1, ..., c, ..., 21]$, the columns correspond to Industry Sectors $S = [1, ..., s, ..., 31]$. $u_{s,t} = \sum_c^C M_{cs,t}$ is a vector of length C, giving the degree of each columns, it represents the ubiquity of an industry sector, whereas $d_{c,t} = \sum_s^C M_{cs,t}$ is a vector of length S, giving the degree of each row and quantifies how diversified a country is.

The bipartite configuration model (BiCM) is a null model for bipartite, undirected, binary networks that is able to generate a grandcanonical ensemble of networks as defined by (37,40). BiCM generates a network ensemble where each matrices has constrained the number of links for each node, on both layers (i.e., $d_{c,t}$ and $u_{s,t}$) to match, on average, the observed one. Each network $\mathbf{M}$ in such ensemble is assigned a probability coefficient

$$P(\mathbf{M}|\vec{x}, \vec{y}) = \prod_c x_c^{d_c(\mathbf{M})} \prod_s y_s^{u_s(\mathbf{M})} \prod_{c,s} (1 + x_c y_s)^{-1}, \tag{8}$$

with x_c and y_s the lagrange multipliers associated to the constrained degrees, of countries and sectors, respectively.

The constrains of the ensemble average values of countries' and sectors' degree allow us to calculate the probability that a link exists between country c and industry sector s independently

of the other links

$$p_{cs} = \frac{x_c y_s}{1 + x_c y_s}. \tag{9}$$

By solving the system of $C + S$ equations it is possible to determine the numerical values of the unknown parameters $\vec{x}$ and $\vec{y}$, which constrains the ensemble average values of countries' diversification and sectors' ubiquity to match the real values, $\langle d_c \rangle = d_c^*$, $c = 1 \ldots C$ and $\langle u_s \rangle = u_s^*$, $s = 1 \ldots s$.

Where $\{d_c^*\}_{c=1}^C$ and $\{u_s^*\}_{s=1}^S$ are the real degree sequence of countries, and industry sectors, respectively, and $\langle \cdot \rangle$ represents the ensemble average of a given quantity, over the ensemble measure defined by Eq. (9)—as $\langle d_c \rangle = \sum_s p_{cs}$ and $\langle u_s \rangle = \sum_c p_{cs}$. Indicated with an asterisk, "$*$" are the parameters that satisfy the systems.

$\mathcal{Z}_t$ is the matrix of dimension (CxC), that represents the projection of $M_{cs,t}$. Each entry $\mathcal{Z}_{cc',t}$ counts the number of industry sectors in common between the countries c and c'. It is defined as

$$\mathcal{Z}_{cc',t} = \sum_{s=1}^S M_{cs,t} M_{c's,t}. \tag{10}$$

Co-specialization motif
The degree of an industry sector in the biadjacency matrix represents the number of countries that the industry in the focal country co-specialize with. To count the abundance of the co-specialization motif we need to count all the possible couple that can be generated within a set of u_s countries. Hence, we have $\binom{u_s}{2} = \frac{1}{2} u_s(u_s - 1)$ possibilities to pick two countries that together with s will form the motif. Following Eq. (3), we can obtain the overall abundance of the co-specialization motif for the entire matrix M.

Research context of BiCM
The BiCM it is an extension of the binary configuration model to bipartite networks (58), which both are null model that uses the framework of exponential random graphs (ERG) (59). The BiCM is of general applicability, and it has been proven valuable in several multidisciplinary studies to quantify the structural modification on the bipartite network over time. For example, in the bipartite network representation of the World Trade Web (i.e., network layers Country / Exported Products) (37,40), BiCM was able to define benchmark to highlight meaningful correlations between

countries and products. The method detected early-warning signals of the changing WTW topology by analyzing the motifs of products' similarities in emerging economies. BiCM makes the validation of the one-mode projection of the bipartite WTW possible. This permits to recognize country pair-wise similarities in industrial systems and to detect statistically significant signals of export specialization dynamics developing from basic to more sophisticated products (44,60). In ecology, BiCM was used to assess the interaction between the species in mutual networks and to show the nested shape of this interaction (61). Finally, BiCM applied on the bipartite representation of stocks ownership by financial institutions showed that increasing portfolio similarities between financial institutions before the 2008 financial crisis enhanced the systemic risk from fire sales liquidation and could be used to forecast market crashes and bubbles (38).

Acknowledgment

Supplementary Material

Supplementary material is available at PNAS Nexus online.

Funding

Authors' Contributions

R.D.C. developed and performed the mathematical framework of the bipartite network analysis. B.L. analyzed the industry dynamics in the EU. L.F.A., and R.E. defined and performed the econometric model. R.D.C., B.L., L.F.A., and R.E. designed the study and performed the research. All the figures, plots, and visualization have been created by the authors. All the authors wrote the paper and gave the final approval to the publication.

Preprints

A preprint of this article is published at: DOI:10.48550/arXiv.2202.01080.

Data Availability

Data have been retrieved from the OECD webpage [https://stats.oecd.org/Index.aspx?DataSetCode=STANI4_2016 (last checked October/2022)].

Code Availability

Bipartite configuration model code can be found here: https://bicm.readthedocs.io/en/latest/ (last checked October/2022)

Motifs and position motifs code can be found here: https://github.com/SimmonsBI/bmotif (last checked October/2022).

References

1. Solow RM. 1956. A contribution to the theory of economic growth. Quart J Econ. 70(1):65.
2. Knox AD, Myrdal G. 1960. Economic theory and under-developed regions. Economica. 27(107):280.
3. Rodríguez-Pose A. 2018. The revenge of the places that don't matter (and what to do about it). Camb J Reg Econ Soc. 11(1):189–209.
4. Smith A, et al. 2002. Networks of value, commodities and regions: reworking divisions of labour in macro-regional economies. Prog Human Geograp. 26(1):41–63.
5. Palan N, Schmiedeberg C. 2010. Structural convergence of European countries. Struc Change Econ Dyn. 21(2):85–100.
6. Midelfart KH, Overman HG, Venables AJ. 2003. Monetary Union and the Economic Geography of Europe. JCMS: J Common Mark Stud. 41(5):847–868.
7. Barro RJ, Martin XSI, Blanchard OJ, Hall RE. 1991. Convergence across states and regions. Brookings Pap Eco Ac. 1991(1):107–182..
8. Borsi MT, Metiu N., 2015. The evolution of economic convergence in the European Union. Emp Econ. 48(2): 657–681.
9. Kutan AM, Yigit TM. 2007. European integration, productivity growth and real convergence. Eur Econ Rev. 51(6): 1370–1395.
10. Kutan AM, Yigit TM. 2009. European integration, productivity growth and real convergence: evidence from the new member states. Econ Syst. 33(2):127–137.
11. Petrakos G, Rodríguez-Pose A, Rovolis A. 2005. Growth, integration, and regional disparities in the European Union. Environ Plan A: Econ Space. 37(10):1837–1855.
12. Henrekson M, Torstensson J, Torstensson R. 1997. Growth effects of European integration. Eur Econ Rev. 41(8):1537–1557.
13. Cabral R. Castellanos-Sosa FA. 2019. Europe's income convergence and the latest global financial crisis. Res Econ. 73(1): 23–34.
14. Benczes I, Szent-Ivanyi B. 2015. The European Economy in 2014: fragile recovery and convergence. JCMS: J Common Mark Stud. 53(S1):162–180.
15. Nagy SG, Šiljak D. 2022. Is the European Union still a convergence machine? Acta Oeconomica. 72(1):47–63.
16. Crestanello P, Tattara G. 2011. Industrial Clusters and the Governance of the Global value chain: the Romania–Veneto network in footwear and clothing. Reg Stud. 45(2):187–203.
17. Frigant V, Layan JB. 2009. Modular production and the new division of labour within europe. Eur Urban Reg Stud. 16(1):11–25.
18. Crespo N, Simões N. 2012. On the measurement of a multidimensional concept of structural similarity. Econ Lett. 116(1):115–117.
19. Kallioras D, Petrakos G. 2010. Industrial growth, economic integration and structural change: evidence from the EU new member-states regions. Annals Reg Sci. 45(3):667–680.
20. Balassa B. 1965. Trade liberalisation and "Revealed" comparative advantage. Manchester School. 33(2):99–123.
21. Steuer MD, Leamer EE. 1987. Sources of international comparative advantage, theory and evidence. volume 54(214). Cambridge, MA: MIT Press.
22. Krugman P, Venables AJ. 1996. Integration, specialization, and adjustment. Eur Econ Rev. 40(3–5):959–967.
23. Romer P. 1987. Growth based on increasing returns due to specialization. Am Econ Rev. 77(2):56–62.
24. Marshall A. 1890. Principles of Economics (Eighth Edition). Cosimo, Inc..

25. Archibugi D, Pianta M. 1994 Aggregate convergence and sectoral specialization in innovation. J Evol Econ. 4(1):17–33.

26. Imbs J, Wacziarg R. 2003. Stages of diversification. Am Econ Rev. 93(1):63–86.

27. Hidalgo CA, Klinger B, Barabási AL, Hausmann R. 2007. The product space conditions the development of nations. Science. 317(5837):482–487.

28. Hidalgo CA, Hausmann R. 2009. The building blocks of economic complexity. Proc Natl Acad Sci, 106(26):10570–10575.

29. Fontagn L, Gaulier G, Zignago S. 2009. Specialization across Varieties and North–South competition. Econ Polic. 23(53):53–93.

30. Levchenko AA, Zhang J. 2012. Comparative advantage and the welfare impact of European integration. Econ Polic. 27(72):567–602.

31. Di Clemente R, Chiarotti GL, Cristelli M, Tacchella A, Pietronero L. 2014. Diversification versus specialization in complex ecosystems. PLoS One. 9(11):e112525.

32. Saracco F, Di Clemente R, Gabrielli A, Pietronero L. 2015. From innovation to diversification: a simple competitive model. PLoS One. 10(11):e0140420.

33. Zweig KA, Kaufmann M. 2011. A systematic approach to the one-mode projection of bipartite graphs. Soc Network Anal Min. 1(3):187–218.

34. Milo R. 2002. Network Motifs: simple building blocks of complex networks. Science. 298(5594):824–827.

35. Simmons BI, et al. 2019. bmotif: a package for motif analyses of bipartite networks. Meth Ecol Evol. 10(5):695–701.

36. Dormann CF, Frund J, Bluthgen N, Gruber B. 2009. Indices, graphs and null models: analyzing bipartite ecological networks. Open Ecol J. 2(1):7–24.

37. Saracco F, Di Clemente R, Gabrielli A, Squartini T. 2015. Randomizing bipartite networks: the case of the World Trade Web. Sci Rep. 5(1):10595.

38. Gualdi S, Cimini G, Primicerio K, Di Clemente R, Challet D. 2016. Statistically validated network of portfolio overlaps and systemic risk. Sci Rep. 6(1):39467.

39. Cimini G, Gabrielli A, Sylos Labini F. 2014. The scientific competitiveness of nations. PLoS One. 9(12):e113470.

40. Saracco F, Di Clemente R, Gabrielli A, Squartini T. 2016. Detecting early signs of the 2007–2008 crisis in the world trade. Sci Rep. 6(1):30286.

41. Di Clemente R, Strano E, Batty M. 2021. Urbanization and economic complexity. Sci Rep. 11(1):3952.

42. Metcalfe JS, Foster J, Ramlogan R. 2006. Adaptive economic growth. Camb J Econ. 30(1):7–32.

43. Tacchella A, Cristelli M, Caldarelli 'G, Gabrielli A, Pietronero L. 2012. A new metrics for countries' fitness and products' complexity. Sci Rep. 2(1):723.

44. Saracco F, et al. 2017. Inferring monopartite projections of bipartite networks: an entropy-based approach. New J Phys. 19(5):053022.

45. van Lidth de Jeude J, Di Clemente R, Caldarelli G, Saracco F, Squartini T. 2019. Saracco, and tiziano squartini. Reconstructing mesoscale network structures. Complexity. 2019: 1–13.

46. Kolev A. 2012. The impact of the recession in 2008–2009 on EU regional convergence. ECON Department, European Investment Bank. p.12. Economic Studies Division SG/ECON/ES/2012-522, http://www. eib. org/infocentre/publications/all/econ-note-2012-regional-convergence. htm,(Erişim, 02.12. 2012)

47. Hristov N, Hülsewig O, Wollmershäuser T., 2012. Loan supply shocks during the financial crisis: evidence for the Euro area. J Int Money Finance. 31(3):569–592.

48. Szekely I, van den Noord PJ. 2009. Economic crisis in Europe: cause, consequences, and responses. A report by the European Commission. European Commission, Directorate-General for Economic and Financial Affairs. Available at: http://dx.doi.org/10.2765/67727.

49. Pavlínek P. 2002. Transformation of the Central and East European passenger car industry: selective peripheral integration through foreign direct investment. Environ Plan A: Econ Space 34(9):1685–1709.

50. Cutrini E, Gardiner B, Martin R. 2022. EU integration and the geographies of economic activity: 1985–2019. Environ Plan A: Econ Space. Available at:https://doi.org/10.1177/0308518X221127028.

51. Caju PD, Kátay G, Lamo A, Nicolitsas D, Poelhekke S. 2010. Nicolitsas, and Steven Poelhekke. inter-industry wage differentials in EU countries: what do cross-country time varying data add to the picture? J Eur Econ Assoc. 8(2–3):478–486.

52. Grossman GM, Rossi-Hansberg E. 2012. Task trade between similar countries. Econometrica. 80(2):593–629.

53. Neary JP. 2009. Putting the "New" into new trade theory: paul krugman's nobel memorial prize in economics. Scand J Econ. 111(2):217–250.

54. Helpman E, Krugman P. 1987. Market structure and foreign trade: Increasing returns, imperfect competition, and the international economy. MIT Press.

55. McCann P, Ortega-Argilés R., 2015. Smart specialization, regional growth and applications to European Union Cohesion Policy. Reg Stud. 49(8):1291–1302.

56. Thissen M, van Oort F, Diodato D, Ruijs A. 2013. Regional competitiveness and smart specialization in Europe. Edward Elgar Publishing.

57. Capello R, Kroll H. 2016. From theory to practice in smart specialization strategy: emerging limits and possible future trajectories. Eur Plan Stud. 24(8):1393–1406.

58. Cimini G, Squartini T, Saracco F, Garlaschelli D. 2019. Andrea Gabrielli, and Guido Caldarelli. The statistical physics of real-world networks. Nat Rev Phys. 1(1):58–71.

59. Holland PW, Leinhardt S. 1981. An exponential family of probability distributions for directed graphs. J Am Stat Assoc, 76(373):33–50.

60. Straka MJ, Caldarelli G, Saracco F., 2017. Grand canonical validation of the bipartite international trade network. Phys Rev E. 96(2):022306.

61. Payrató-Borràs C, Hernández L, Moreno Y., 2019. Breaking the spell of nestedness: the entropic origin of nestedness in mutualistic systems. Phys Rev X. 9(3): 031024.