



OPEN

Urbanization and economic complexity

Riccardo Di Clemente^{1,2}✉, Emanuele Strano³ & Michael Batty²

Urbanization plays a crucial role in the economic development of every country. The mutual relationship between the urbanization of any country and its economic productive structure is far from being understood. We analyzed the historical evolution of product exports for all countries using the World Trade Web with respect to patterns of urbanization from 1995 to 2010. Using the evolving framework of economic complexity, we reveal that a country's economic development in terms of its production and export of goods, is interwoven with the urbanization process during the early stages of its economic development and growth. Meanwhile in urbanized countries, the reciprocal relation between economic growth and urbanization fades away with respect to its later stages, becoming negligible for countries highly dependent on the export of resources where urbanization is not linked to any structural economic transformation.

It is an established fact that urbanization in developed countries is accompanied by economic growth and industrialization which mutually self-reinforce one another¹. This historic pattern generates an expectation of a virtuous circle between economic growth and urbanization regardless of local conditions^{2,3}. From classic urban economic theories^{4,5} to the more recent scaling approach to cities^{6,7}, the growth of urban population has routinely been used as a proxy for economic growth. This pattern has also been observed in rapidly developing countries such as China and India but it cannot be considered a universal blueprint⁸ for deviations from this norm have not yet been fully explained.

In fact, as pointed out in several studies^{9–12}, the increasing urbanization rate in persistently poor and non-industrialized countries poses an important dilemma for urban economic theory. Why, given the same rate of urbanization, does Asia contain a number of explosive economies while sub-Saharan Africa has seen very little growth? Moreover, in developed and advanced industrialized economies, is there appears to be a competitive advantage in continuing this urbanization process indefinitely?

There are several theories aimed at explaining urbanization processes. Some argue that rural poverty moves people to cities as was clearly the case in nineteenth century Europe and America¹³, driving the transformation from an agricultural to an industrial-service based economy^{14,15}. Others argue that in the last decades there has been urban-biased public policy that has led to over-urbanization¹².

The most intriguing approach however is rooted in the mutual indirect effects of the World Trade Web (WTW) on global urbanization^{16,17}. The dominant idea is that in open economies, domestic communities (such as cities) can trade easily with other communities, boosting their exports in substituting industrialization for urbanization policy¹⁸. In simple terms, the commodities can flow more freely using urban agglomerates as nodes in the trading networks between countries, generating the ever present virtuous circle between economic growth and urbanization.

Starting from this theory, we analyze the WTW to explore the mutual relationship between the urbanization of the countries and their economic production structure using the Economic Complexity (EC) framework. Economic Complexity^{19–27}, is a new and expanding field in the economic analysis, which proposes “Fitness” and “Complexity” metrics to quantify the fitness or competitiveness of countries and the complexity of products from a country's basket of exports. The main focus of EC is based on a bipartite representation of the World Trade Web where the nodes represent the set of world-countries and the set of exported products defined as different entities. Countries and products are connected to one another by imposing a threshold based on their Revealed Comparative Advantage (RCA)²⁸ which defines the criterion for the existence of relations.

The Fitness and Complexity algorithm is a kind of PageRank method applied to WTW, where Fitness F_c is the quantity for country c , and Complexity Q_p is the quantity for products p . The idea at the basis of the algorithm is that the countries with the highest fitness are those which are able to export the highest number of the most exclusive products i.e. those with the highest complexity. On the other hand this complexity is non linearly related

¹Department of Computer Sciences, University of Exeter, Exeter EX4 4QF, UK. ²Centre for Advanced Spatial Analysis, University College London, London W1T 4TJ, UK. ³MindEarth, 2502 Biel/Bienne, Switzerland. ✉email: r.di-clemente@exeter.ac.uk

to the fitness of its exporters so that products exported by low fitness countries have a low level of complexity and high complexity products are exported by high fitness countries only.

The Fitness metric is valuable in quantifying a country's productive structure and structural transformations which enable one to predict its future economic growth²³. It correlates with the extent of economic equality²⁹ and it has been used to analyze the country's growth path to industrialization³⁰.

In this work, we used a data driven approach borrowing tools recently introduced by statistical physics and network science to improve our understanding of the complex dynamics of human societies, with the aim of finding innovative insight³¹ to link urbanization process with the evolution of the international trade.

We couple the WTW data with the urbanization level of more than 144 countries worldwide, and analyze this between 1995–2010 thus capturing the fingerprint of urbanization on countries' productive systems through the lens of their exports. We notice that in rural economies, the increase in urban population fosters structural changes in industrial exports. It boosts the country's diversification improving the country fitness, and allowing the export of more complex products. These economic transformations fade away in countries that already have a high level of urban population (more than 60%) where there is no relation between the urbanization process and the country's fitness.

Within the sub-Saharan countries, we capture those where the virtuous circle between economic growth and urbanization is fostering structural changes in those countries' productive systems. On the other hand within countries with economies based on raw materials, we assess the implementation of policy leading to urbanization that does not support any structural transformations of their basket of exports.

Results

Economic complexity and urbanization. We represent the WTW as a bipartite network, i.e. by considering the set of world-countries and the set of products as different entities and linking a given country to a given product if (and only if) the former exports to the latter above a certain threshold (the so-called Revealed Comparative Advantage—RCA)²⁸. RCA is a general criterion adopted in order to understand whether a country can be considered, or not, a producer of a particular product. It quantifies how much the export of a given product p is relevant for the economy of a country c in relation to the global export of p for all countries (See “Methods” Section).

The country's fitness and product's complexity are the result of a non-linear iterative map applied to the WTW matrix M ^{19,20,32} (See “Methods”).

Through the algorithm's iterations, products exported by low fitness countries have a low level of complexity while high complexity products are exported by high fitness countries only. The countries' composition of their export basket depends on their fitness. Fitness and Complexity are thus non-monetary indicators of the economy's development: the fitness represents a measure of tangible and intangible assets and capabilities, which drive the country's development, such as political organization, its history, geography, technology, services, and infrastructures²¹. Meanwhile complexity measures the necessary capabilities which must be owned by a country in order that it can produce and then export the resulting product.

Within this framework, we include the dimension of a country's degree of urbanization defined as the percentage of the total population living in urban areas. Our aim is to quantify the link between a country's urbanization process and their exports as a proxy for their industrial economic system. To disentangle the relation between country productivity systems and their urbanization, we have divided the set of countries in terms of their degree of urbanization, defined by the Urban Range, which is expressed in four quantiles [Q1, Q2, Q3, Q4] (see urban range division in Fig. 1B top).

More urbanized countries [Q3, Q4] in the early 2000s, export a wide range of complex products such as: textiles, heavy manufacturing industries, and IT while rural countries [Q1, Q2] export products that require a low level of sophistication such as raw materials and agricultural products (Fig. 1A).

Highly urbanized countries maintain a similar distribution across the analysis years, with a long tail of low complexity products and a consistent increase in the number of high complexity products. On the other hand starting from 2005, we have noticed that rural countries change their export basket towards higher complexity products. This shift is shown by the cumulative distribution functions of the different Urban ranges that decrease their distance from one another over time (see Fig. 1A inset) together with their median and peak distance.

We notice that countries within the higher quantile of the Urban Range, Fig. 1B, are the ones with higher fitness and higher diversification, whilst low urbanized countries have a low diversification and fitness. Notable exceptions are countries with exports based on raw materials (i.e. Qatar, Kuwait, Gabon, Iraq, Libya). These countries reached higher levels of urbanization as result of policy decisions³³ meanwhile their exports are limited to a few products with low complexity.

The representation of the WTW in Fig. 1C shows country exports in 2010 rearranged by ranked fitness and complexity. The country exports' diversification is related to the urbanization level. Low urbanized countries are at the bottom of the matrix with low fitness and lower degree of diversification, whilst the urbanized countries, with the most advanced economies lie at the top, with a high degree of product diversification with different levels of complexity and high fitness.

Exports diversification and urbanization. It is known that low fitness countries have similar economies with low degrees of diversification and high similarity with respect to their export baskets^{20,34,35} i.e. they produce and export few of the same low technology products. We captured a shift in the distribution of the exported products within the rural countries (Fig. 1A). In particular, we noticed that rural countries start to produce and export more sophisticated products. This productive systems transformation in the EC literature is related to the development of new capabilities^{22,36,37}.

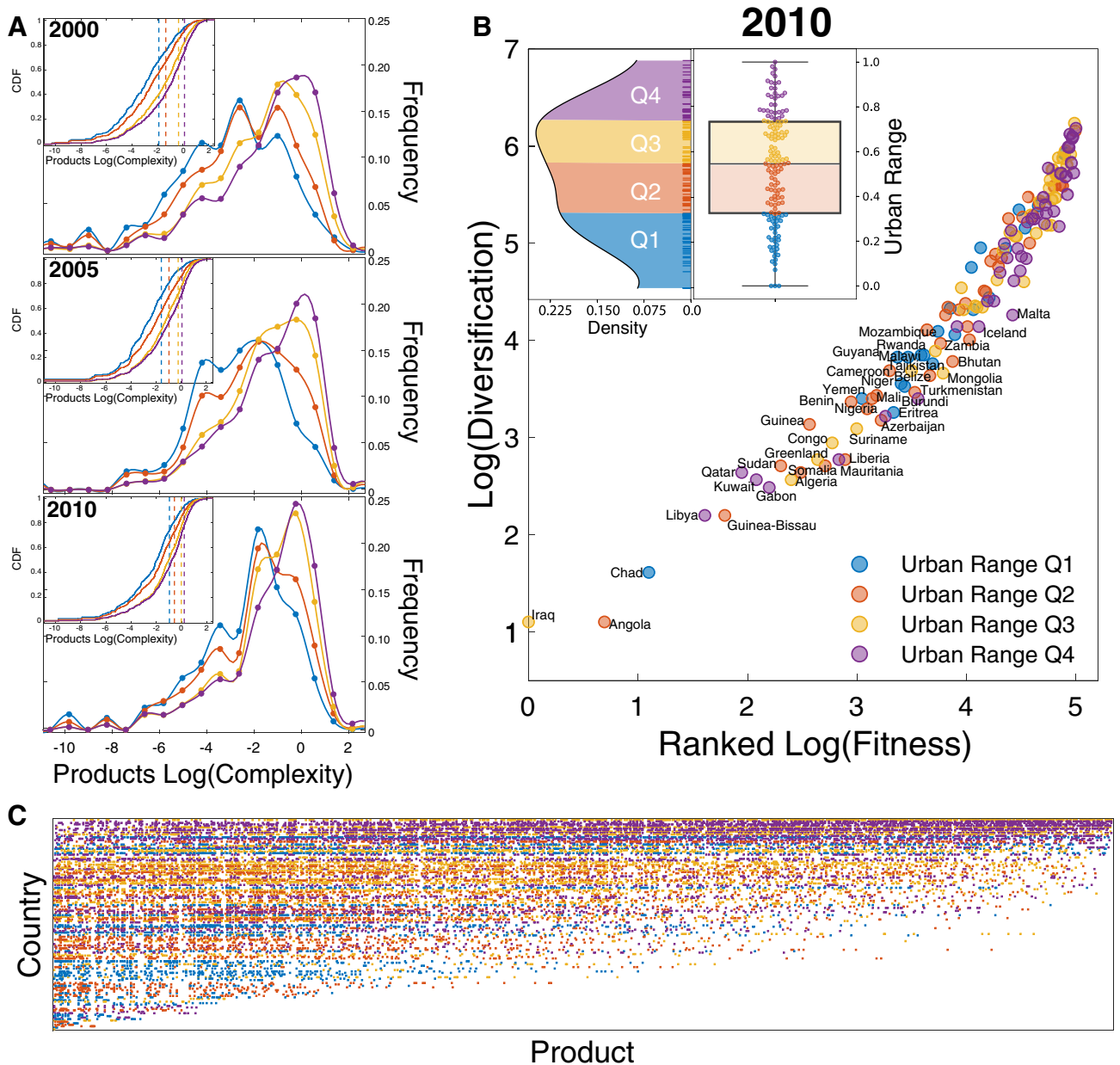


Figure 1. (A) Distribution of exported products complexity by different urbanization levels through the 2000–2010. There is a shift of lower urbanized countries towards the export of more complex products (B) Distribution of the Urban Range (percentage of the total population living in urban areas) of the 144 countries analyzed. (B) Ranked country fitness versus products export diversification, the highly diversified countries are one’s with more fitness and high urbanization, meanwhile low urbanized country are in the center bottom of the scatter plot, with some exceptions such as those with links to the oil countries. (C) Matrix of the country exports in 2010, reordering the countries and products by fitness and complexity; the color dots represent an exported product under the RCA threshold Eq. (2), the color gradient follows the urban range definition.

Some questions from this analysis emerge: do the rural countries evolve their productive systems in the same way? and do they continue to produce and export the same products? Is the pattern of economic development entangled with urbanization in same fashion for each rural country?

We can measure indirectly the transformation of the productive systems by analyzing the evolution of WTW topology³⁸. In particular we can assess the changes of countries’ similarities in their exports studying the abundance evolution of network motifs³⁹. A network motif is a particular pattern of interconnections occurring between the nodes of the network (i.e. between the countries and their products). In our case we are interested in the abundance of the similarity motif μ_{sim} in Fig. 2B (motifs 6⁴⁰, or X motif³⁵): it quantifies the co-occurrence of any two countries as producers of the same couple of products as Eq. (7) (and, viceversa, the co-occurrence of any two products in the basket of the same couple of countries). This represents the simplest motifs⁴⁰ that can quantify the similarities in the export countries’ diversification which maintains a pairwise correlation within

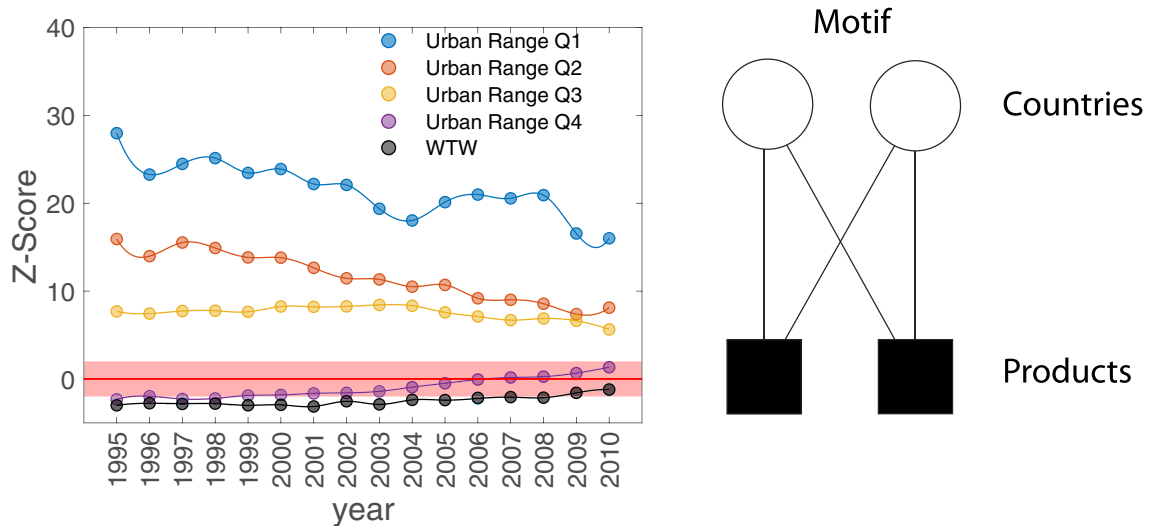


Figure 2. (A) Z-score of the export similarity motif by country groups with different Urban Ranges and the Z-score of the whole WTW in black. (B) Similarity motif as the co-occurrence of any two countries as producers of the same couple of products^{35,40}.

the products exported. Two economies with a fixed number of products exported are diversifying if the values of μ_{sim} is decreasing while their production similarity increases with high values of μ_{sim} .

To provide a benchmark and assess the μ_{sim} statistical significance of the WTW we use the Bipartite Configuration Model (BiCM)³⁴ as a null model. This framework is valuable in the analysis of the abundance of the bipartite motifs⁴⁰, enabling us to detect financial crisis effects on a country's export basket³⁵ as well as export similarities between countries with same level of economic development⁴¹.

We generated 1000 matrices using the BiCM³⁴ (see "Methods" Section) and we compare the observed abundances of the similarity motif (Eq. 7) in the real network with the corresponding expected values in the null ensemble using the Z-score.

The whole WTW manifests a progressive increase of the abundance of similarity motifs with respect to the null case³⁵ (Black line Fig. 2A). Highly urbanized countries show a similar trend of increasing similarity in their products exports. This measure implies that rural economies are very similar with a higher abundance of the similarity motif with respect to the random case having a high value Z-score. Interestingly, low urban range countries diversifying between each other manifest an opposite trend. The exports diversification trends of the low urbanized countries coupled with the increasing complexity of the product exported imply a nontrivial connection between urbanization and production capabilities. This measure outlines how rural economies follow different development patterns based on their production systems. The urbanization phenomenon coupled with the capabilities already presented in the country enable the production of different sophisticated products depending on their environment.

Urbanization growth and country fitness. The economic transformation of a rural country has an impact on its overall fitness value, and the competitiveness of its productive system. In this respect, the urbanization process is key element in a country's development and its economic growth^{33,42}. To assess the relation between the country's fitness and the urbanization process we analyzed the Urban Range growth rate in relation to the growth rate of country fitness ranking between 1995 and 2010, as we show in Fig. 3. The country fitness ranking is the country's ordered position with respect to the country's fitness value in a given year. The growth rate of the country's fitness ranking is an easily understood tool to compare the transformations of a country's productive systems with respect to its competitors. It has been proven a reliable tool in quantifying the country's relative degree of competitiveness across different years providing a more stable measurement than the raw fitness value⁴³.

For each of the four Urban Range quantiles we find a linear relation between the urbanization rate and the Fitness ranking growth rate in Fig. 3b. Increasing urbanization within lowly urbanized countries is interwoven with increasing Fitness. Meanwhile, the effects are minimal in highly urbanized countries (Urban Range Q3,Q4). We validate the urbanization/fitness relation analyzing a 25% quantile sliding window on the whole urbanization distribution, which we show in black in Fig. 3B.

We notice that in many rural economies, the urbanization process affects or has been affected by structural changes in its economic production. (An example are countries such as Uganda, Nepal, Somalia.) On other hand, there are many countries (such as IvoryCoast, Paraguay, Chad) where the urbanization process does not provide improvement in the fitness ranking⁴⁴.

The self-reinforced mechanism between urbanization and fitness reaches a plateau within the urbanized countries (Q3,Q4), where the urbanization does not affect or has not been affected by changes in fitness ranking. In this respect, the resource exports countries manifest a shift toward a negative relation between urbanization and fitness. In fact in countries that are heavily dependent on resource exports, urbanization appears to be

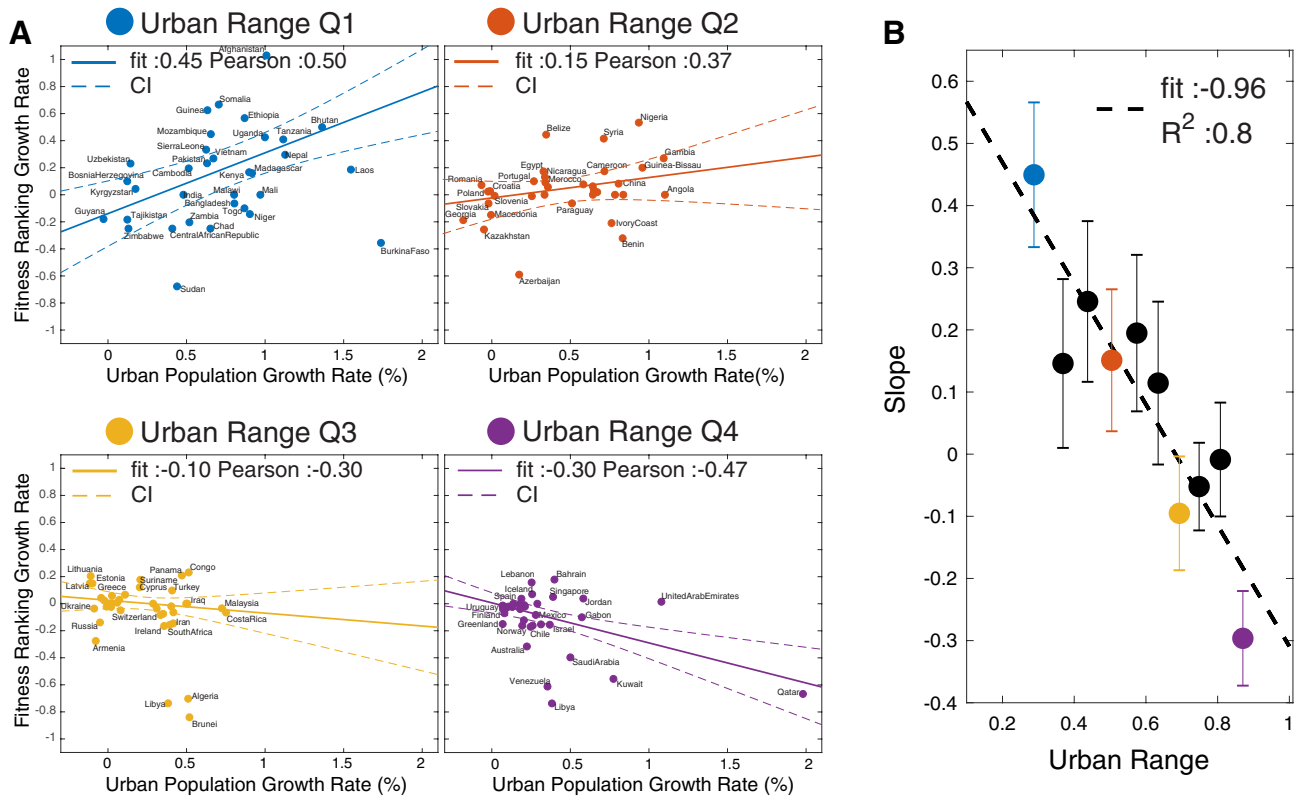


Figure 3. (A) The Fitness Ranking Growth Rate versus Urbanization Growth Rate. The effect of urbanization growth on the transformation of the economic systems (or vice-versa) is more relevant in low urbanize countries. The dashed lines represent the 95% Confidence Interval (CI) of the linear regression. (B) Slope coefficient of a sliding window across 25% of the countries (corresponding to 36 countries) of its fitness ranking growth rate versus urban population growth rate. The error bar corresponds to the fit's 95% confidence interval. The colors follow the Urban Range Scheme.

concentrated in the cities where the economies consist primarily of non-tradeable services⁴⁵. To support our result we provide the same analysis using instead of the Fitness Ranking metric, the Fitness, Gross Domestic Product (GDP) and GDP Ranking respectively (see “Methods” section: Urban Range vs Fitness and GDP). We do not find any evidence of relation between the other three metrics and the urbanization rate.

Urban fitness trends. The process of urbanization is often entangled with a country’s industrialization¹¹. As countries develop, people move out of rural areas and agricultural activities into urban centers, where they engage in manufacturing products⁴⁶ which are more sophisticated with higher complexity. This transformation is outlined by the increasing level of fitness of low urbanized countries that are involved in the urbanization process. To leverage this information and capture its trends, we define the country Urban Fitness $F_c^{urb}(t) = F_c(t) * U_c(t)$; this is the value of country fitness F_c weighted by the percentage of urban population U_c .

We cluster the countries Urban Fitness trends using the Louvain algorithm⁴⁷ which is based on their correlation matrix shown in Fig. 4B. Three clusters emerge with high correlations disentangling the non-trivial geographical relations we show in Fig. 4A–C.

In Fig. 4A countries with a clear urbanization trend (in orange) are ones with a stable increase in fitness ranking. Meanwhile the blue cluster contains developed countries, where the urbanization does not provide any new input to the economic development and resource dependent countries, where the urbanization is not only lead by deep structural economic change. These results are in agreement with the Urban Range study in Fig. 3 that show a poor effect of the urbanization on the country fitness, implying that over a given value of urbanization, other factors have a more important role in economic development and growth. Finally, the third cluster (in red) are the countries without any clear trend and are thus uncategorized.

Discussion

It is well-known that urbanization provides several advantages to the economics of scale and division of labour, boosting productivity and competition. It helps in accessing the labor force and inputting materials to the production process as well as decreasing the geographical distance between firms, reducing transaction costs, and fostering competition⁴⁸. These urbanization advantages⁴⁹ together with the appropriate bureaucratic environment³³, investment in infrastructures⁵⁰ and companies market structure⁵¹, are some of intangible attributes, the capabilities, that a country needs to drive economic growth and innovation³⁶. We noticed that the country Fitness,

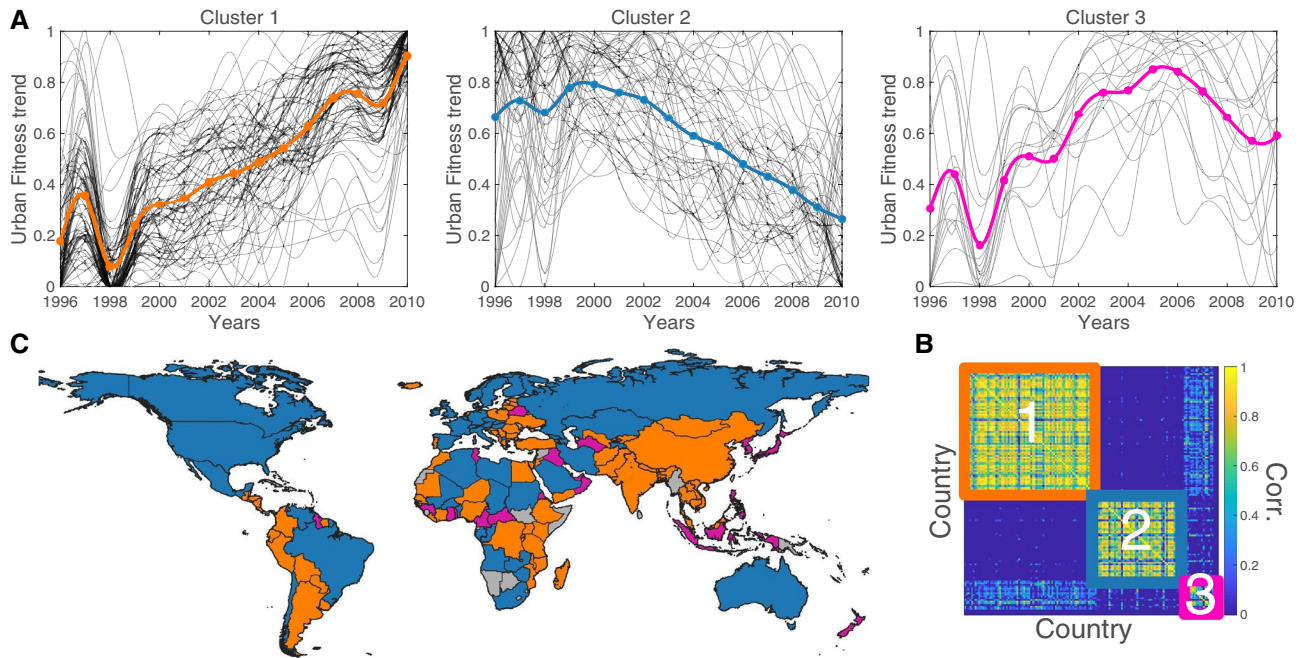


Figure 4. (A) Clusters of normalized Urban Fitness Trends. (B) Correlation Matrix of the countries urban fitness trends clustered with the Louvain algorithm. (C) Geographical cluster distribution. The map in this figure was created using the software QGIS.

the production and export of goods, is interwoven within the urbanization process during the early stages of country’s economic development and growth. We show that the information carried by WTW can provide a different perspective on analyzing the complex process of urbanization, enlightening the relation between a country’s exports, economic development and its urban growth.

Methods

Data. *World trade web.* The dataset used in this work is the BACI (Base pour l’Analyse du Commerce International) World Trade Database (Gaulier, S. Baci: International trade database at the product-level <http://www.cepii.fr/CEPII/fr/publications/wp/abstract.asp?NoDoc=2726> Date of access: 18/01/2021). The data contains information on the trade of 200 different countries for more than 5000 different products, categorized according to the 6-digit code of the Harmonised System 2007 (<http://www.wcoomd.org/> Date of access: 18/01/2021). The products’ sectors follows the UN categorization (<http://unstats.un.org/unsd/ct/registry/regcst.asp?Cl=8> Date of access: 18/01/2021). We create a map between the two systems converting the HS2007 in to the ISIC revision 2 code at 2-digit (<http://www.maclester.edu/research/economics/PAGE/HAVEMAN/Trade.Resources/TradeConcordances.html#FromISIC> Date of access: 18/01/2021). We represent the trade relation between the 144 countries $c \in [1, C]$ and the 1131 products $p \in [1, P]$ between the years [1995, 2010] through the bipartite matrix \tilde{M} with dimension $(C \times P)$ where each entry $\tilde{m}_{c,p}$ measures the export in US dollars. The framework of the Economic-Complexity^{19–22} based on the interaction between countries and products is expressed by the application of the Revealed Comparative Advantage (RCA)²⁸ threshold over the entries $\tilde{m}_{c,p}$:

$$RCA_{c,p} = \frac{\frac{\tilde{m}_{c,p}}{\sum_{i,p} \tilde{m}_{c,p'}}}{\frac{\sum_{i,c} \tilde{m}_{c,p}}{\sum_{i,p} \sum_{i,c} \tilde{m}_{c,p'}}} \tag{1}$$

Finally, we define the entries of the biadjacency matrix M of the undirected bipartite network analyzed in this work as:

$$\begin{cases} m_{cp} = 1 & \text{when } RCA_{cp} \geq 1 \\ m_{cp} = 0 & \text{otherwise} \end{cases} \tag{2}$$

This indicates that the connection (country-product link) is established if and only if the relative RCA is relevant (over the threshold), otherwise it can be ignored. Each row of M represents the export basket of a given country (or its diversification k_c), while each column represents the subset of producers of a given product (or its ubiquity k_p)⁵².

$$k_c = \sum_p m_{cp} \quad k_p = \sum_c m_{cp} \tag{3}$$

Urbanization. The data of the urban population from 1995 to 2010 are available at the World Bank database (<https://data.worldbank.org/> Date of access: 18/01/2021).

Fitness and complexity. Fitness and Complexity are a metric for countries and products applied to bipartite binary matrix M of the WTW^{19–22,24}. The basic idea of EC is to define a non-linear map through an iterative process which couples the Fitness of countries to the Complexity of products. At every step of the iteration, the Fitness F_c of a given country c is proportional to the sum of the exported products, weighted by their complexity parameter Q_p . In particular, the Fitness F_c for the generic country c and Quality Q_p for the generic product p at the n -th step of iteration, are defined as:

$$\begin{cases} \tilde{F}_c^{(n)} = \sum_p m_{cp} Q_p^{(n-1)} \\ \tilde{Q}_p^{(n)} = \frac{1}{\sum_c m_{cp} \frac{1}{F_c^{(n-1)}}} \end{cases} \rightarrow \begin{cases} F_c^{(n)} = \frac{\tilde{F}_c^{(n)}}{\langle \tilde{F}_c^{(n)} \rangle} \\ Q_p^{(n)} = \frac{\tilde{Q}_p^{(n)}}{\langle \tilde{Q}_p^{(n)} \rangle} \end{cases}, \quad (4)$$

where the symbols $\langle \cdot \rangle$ indicate the average taken over the proper set. The initial condition are taken as $F_c^0 = Q_p^0 = 1 \forall c \in N_c, \forall p \in N_p$, where N_c and N_p are the number respectively of countries and products (the convergence of the algorithm described by Eq. (4) depends on the shape of the matrix M , as it has been discussed in⁴³).

Bipartite configuration model (BICM). The Bipartite Configuration Model (BICM), as defined by^{34,35}, is a null model of general applicability that is able to generate a grandcanonical ensemble of bipartite, undirected, binary networks in which the two layers Country and Products have respectively C and P nodes. The ensemble generated by the BICM constrained the number of connections for each node, on both layers (in our case d_c and u_p) to match, on average, the observed one. Each network M in such ensemble is assigned a probability coefficient:

$$P(M|\mathbf{x}, \mathbf{y}) = \prod_c x_c^{d_c(M)} \prod_p y_p^{u_p(M)} \prod_{c,p} (1 + x_c y_p)^{-1}, \quad (5)$$

x_c and y_p are the Lagrange multipliers associated to the constrained degrees.

Constraining the ensemble average values of countries and products degree induces the probability that a link exists between country c and industry sector p independently of the other links:

$$p_{cp} = \frac{x_c y_p}{1 + x_c y_p}. \quad (6)$$

The numerical values of the unknown parameters \mathbf{x} and \mathbf{y} have to be determined by solving the following system of $C + P$ equations, which constrains the ensemble average values of countries diversification and products ubiquities to match the real values, $\langle d_c \rangle = d_c^*$, $c = 1 \dots C$ and $\langle u_p \rangle = u_p^*$, $p = 1 \dots P$.

Where $\{d_c^*\}_{c=1}^C$ and $\{u_p^*\}_{p=1}^P$ 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. (6)—as $\langle d_c \rangle = \sum_s p_{cp}$ and $\langle u_s \rangle = \sum_c p_{cp}$. Indicated with an asterisk, “*” are the parameters that satisfy the systems.

Similarity motifs. In the present work we have sampled the grand canonical ensemble of binary, undirected, bipartite networks induced by the BiCM, according to the probability coefficients $P(M|\mathbf{x}^*, \mathbf{y}^*)$ and calculated the average and variance of the motif μ_{sim} , define as b-motif6 in⁴⁰.

The Similarity Motif represents the symmetric and complete connections between two countries c, c' and two industry sectors p, p' . The number of similarity motifs is:

$$\mu_{sim} = \frac{1}{4} \sum_{c=1}^C \sum_{c'=1}^C \mathcal{Z}_{cc'} (\mathcal{Z}_{cc'} - 1) - \frac{1}{4} \sum_{c=1}^C d_c (d_c - 1) \quad (7)$$

with \mathcal{Z} is the matrix of dimension (C, C) , that represents the projection of M over the countries. Each entry $\mathcal{Z}_{cc'}$ counts the number of industry sectors in common between the countries c and c' , it is defined as: $\mathcal{Z}_{cc'} = \sum_{s=1}^S M_{cs} M_{c's} = MM^T$

This motif represents the co-occurrence of two products in two countries' export basket within the bipartite matrix of the country exports. The accuracy of the BiCM prediction in reproducing the value of quantity μ_{sim} please follows³⁴.

Urban range versus fitness and GDP. To validate our analysis of the relation between the country's fitness and the urbanization process we analyzed the urbanization growth rate in relation to the growth rate of three different metrics: the country Fitness (Fig. 5A), country GDP (Fig. 5B), and country GDP ranking (Fig. 5C) between 1995 and 2010.

We study the variation of the slope coefficient of a sliding window across 25% of the countries urban range and the three metrics above. Both the metrics extracted from the GDP do not have statistical significant results.

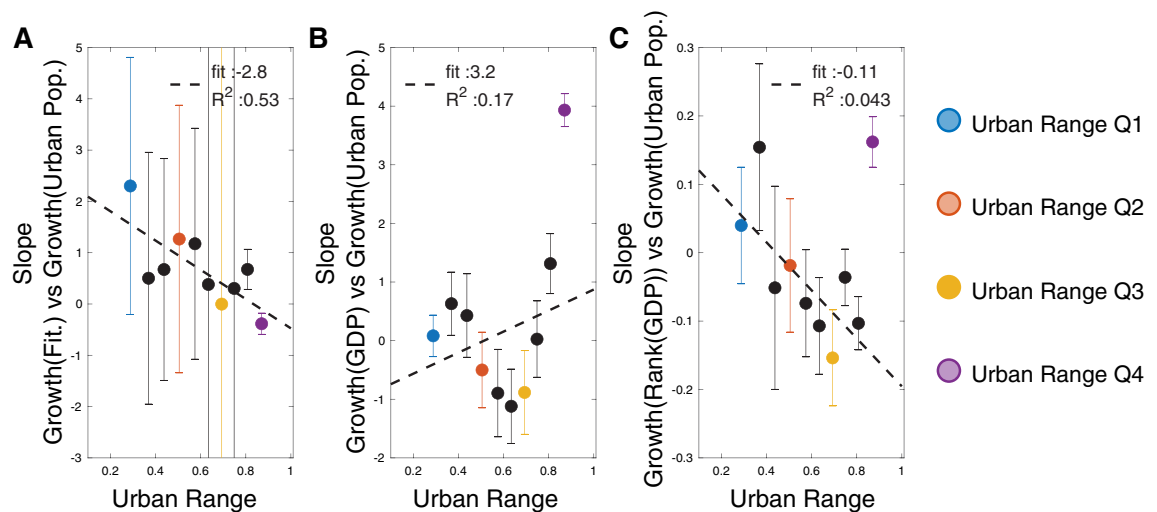


Figure 5. Slope coefficient of a sliding window across 25% of the countries (corresponding to 36 countries) of respectively its Fitness Growth Rate (A)— GDP Growth Rate (B)— GDP Ranking Growth Rate (C) versus Urban Population Growth Rate. The error bar corresponds to the fit's 95% confidence interval. The colors follow the Urban Range scheme.

Although the growth rate of fitness in relation with the urbanization growth rate manifests a linear relation (Fig. 5A) with an $R^2 = 0.53$, as Fig. 3B, we notice that the fitness ranking is a more reliable tool than the raw fitness value⁴³. The fitness ranking provides a more stable metric across each sliding window.

Received: 5 November 2020; Accepted: 24 January 2021

Published online: 17 February 2021

References

- Martin, P. & Ottaviano, G. I. P. Growth and agglomeration. *Int. Econ. Rev.* **42**, 947–968. <https://doi.org/10.1111/1468-2354.00141> (2001).
- Buckley, R. M., Annez, P. C. & Spence, M. *Urbanization and Growth* (The World Bank, 2008).
- Duranton, G. Growing through cities in developing countries. *The World Bank Research Observer* **30**, 39–73. <https://doi.org/10.1093/wbro/lku006> (2014).
- Jacobs, J. *The Economy of Cities* (Vintage, New York, 2016).
- Glaeser, E. L., Kallal, H. D., Scheinkman, J. A. & Shleifer, A. Growth in cities. *J. Polit. Econ.* **100**, 1126–1152. <https://doi.org/10.1086/261856> (1992).
- Bettencourt, L., Lobo, J., Helbing, D., Kühnert, C. & West, G. Growth, innovation, scaling, and the pace of life in cities. *PNAS Proc. Natl. Acad. Sci.* **104**, 7301–7306. <https://doi.org/10.1073/pnas.0611072104> (2007).
- Youn, H. *et al.* Scaling and universality in urban economic diversification. *J. R. Soc. Interface* **13**, 20150937. <https://doi.org/10.1098/rsif.2015.0937> (2016).
- Ravallion, M., Chen, S. & Sangraula, P. New evidence on the urbanization of global poverty. *Popul. Dev. Rev.* **33**, 667–701. <https://doi.org/10.1111/j.1728-4457.2007.00193.x> (2007).
- Montgomery, M. R. The urban transformation of the developing world. *Science* **319**, 761. <https://doi.org/10.1126/science.1153012> (2008).
- Petrakos, G. Urbanization and international trade in developing countries. *World Dev.* **17**, 1269–1277. [https://doi.org/10.1016/0305-750X\(89\)90237-4](https://doi.org/10.1016/0305-750X(89)90237-4) (1989).
- Gollin, D., Jedwab, R. & Vollrath, D. Urbanization with and without industrialization. *J. Econ. Growth* **21**, 35–70. <https://doi.org/10.1007/s10887-015-9121-4> (2015).
- Glaeser, E. L. A world of cities: the causes and consequences of urbanization in poorer countries. *J. Eur. Econ. Assoc.* **12**, 1154–1199. <https://doi.org/10.1111/jeea.12100> (2014).
- Ades, A. F. & Glaeser, E. L. Trade and circuses: explaining urban giants. *Q. J. Econ.* **110**, 195–227. <https://doi.org/10.2307/2118515> (1995).
- Davis, J. C. & Henderson, J. Evidence on the political economy of the urbanization process. *J. Urban Econ.* **53**, 98–125. [https://doi.org/10.1016/S0094-1190\(02\)00504-1](https://doi.org/10.1016/S0094-1190(02)00504-1) (2003).
- Henderson, V. The urbanization process and economic growth: the so-what question*. *J. Econ. Growth* **8**, 47–71. <https://doi.org/10.1023/a:1022868000744> (2003).
- Fujita, M. & Thisse, J.-F. Does geographical agglomeration foster economic growth? and who gains and loses from it?. *Jpn. Econ. Rev.* **54**, 121–145. <https://doi.org/10.1111/1468-5876.00250> (2003).
- Krugman, P. R. *Geography and Trade* (MIT Press, Cambridge, 1993).
- Krugman, P. & Elizondo, R. L. Trade policy and the third world metropolis. *J. Dev. Econ.* **49**, 137–150. [https://doi.org/10.1016/0304-3878\(95\)00055-0](https://doi.org/10.1016/0304-3878(95)00055-0) (1996).
- Hidalgo, C. A. & Hausmann, R. The building blocks of economic complexity. *Proc. Natl. Acad. Sci.* **106**, 10570–10575. <https://doi.org/10.1073/pnas.0900943106> (2009).
- Tacchella, A., Cristelli, M., Caldarelli, G., Gabrielli, A. & Pietronero, L. A new metrics for countries' fitness and products' complexity. *Sci. Rep.* **2**, 723. <https://doi.org/10.1038/srep00723> (2012).
- Cristelli, M., Gabrielli, A., Tacchella, A., Caldarelli, G. & Pietronero, L. Measuring the intangibles: a metrics for the economic complexity of countries and products. *PLoS One* **8**, e70726. <https://doi.org/10.1371/journal.pone.0070726> (2013).

22. Hidalgo, C. A., Klinger, B., Barabási, A.-L. & Hausmann, R. The product space conditions the development of nations. *Science* **317**, 482–487. <https://doi.org/10.1126/science.1144581> (2007).
23. Tacchella, A., Mazzilli, D. & Pietronero, L. A dynamical systems approach to gross domestic product forecasting. *Nat. Phys.* **14**, 861. <https://doi.org/10.1038/s41567-018-0204-y> (2018).
24. Tacchella, A., Cristelli, M., Caldarelli, G., Gabrielli, A. & Pietronero, L. Economic complexity: conceptual grounding of a new metrics for global competitiveness. *J. Econ. Dyn. Control* **37**, 1683–1691. <https://doi.org/10.1016/j.jedc.2013.04.006> (2013).
25. Albeaik, S., Kaltenberg, M., Alsaleh, M. & Hidalgo, C. A. Improving the economic complexity index. arXiv preprint arXiv:1707.05826. <http://arxiv.org/1707.05826> (2017).
26. Servedio, V. D. P., Buttà, P., Mazzilli, D., Tacchella, A. & Pietronero, L. A new and stable estimation method of country economic fitness and product complexity. *Entropy* **20**, 783. <https://doi.org/10.3390/e20100783> (2018).
27. Sciarra, C., Chiarotti, G., Ridolfi, L. & Laio, F. Reconciling contrasting views on economic complexity. *Nat. Commun.* **11**, 3352. <https://doi.org/10.1038/s41467-020-16992-1> (2020).
28. Balassa, B. Trade liberalisation and “revealed” comparative advantage. *The Manchester School* **33**, 99–123. <https://doi.org/10.1111/j.1467-9957.1965.tb00050.x> (1965).
29. Hartmann, D., Guevara, M. R., Jara-Figueroa, C., Aristrán, M. & Hidalgo, C. A. Linking economic complexity, institutions, and income inequality. *World Dev.* **93**, 75–93. <https://doi.org/10.1016/j.worlddev.2016.12.020> (2017).
30. Pugliese, E., Chiarotti, G. L., Zaccaria, A. & Pietronero, L. Complex economies have a lateral escape from the poverty trap. *PLoS One* **12**, e0168540. <https://doi.org/10.1371/journal.pone.0168540> (2017).
31. Perc, M. The social physics collective. *Sci. Rep.* **9**, 16549. <https://doi.org/10.1038/s41598-019-53300-4> (2019).
32. Caldarelli, G. *et al.* A network analysis of countries’ export flows: firm grounds for the building blocks of the economy. *PLoS One* **7**, e47278. <https://doi.org/10.1371/journal.pone.0047278> (2012).
33. Henderson, J. V., Shalizi, Z. & Venables, A. J. Geography and development. *J. Econ. Geogr.* **1**, 81–105. <https://doi.org/10.1093/jeg/1.1.81> (2001).
34. Saracco, F., Di Clemente, R., Gabrielli, A. & Squartini, T. Randomizing bipartite networks: the case of the world trade web. *Sci. Rep.* **5**, 10595. <https://doi.org/10.1038/srep10595> (2015).
35. Saracco, F., Di Clemente, R., Gabrielli, A. & Squartini, T. Detecting early signs of the 2007–2008 crisis in the world trade. *Sci. Rep.* **6**, 30286. <https://doi.org/10.1038/srep30286> (2016).
36. Saracco, F., Di Clemente, R., Gabrielli, A. & Pietronero, L. From innovation to diversification: a simple competitive model. *PLoS One* **10**, e0140420. <https://doi.org/10.1371/journal.pone.0140420> (2015).
37. Tacchella, A., Di Clemente, R., Gabrielli, A. & Pietronero, L. The build-up of diversity in complex ecosystems. arXiv preprint arXiv:1609.03617. <http://arxiv.org/1609.03617> (2016).
38. van Lidth de Jeude, J., Di Clemente, R., Caldarelli, G., Saracco, F. & Squartini, T. Reconstructing mesoscale network structures. *Complexity*. <https://doi.org/10.1155/2019/5120581> (2019).
39. Gualdi, S., Cimini, G., Primicerio, K., Di Clemente, R. & Challet, D. Statistically validated network of portfolio overlaps and systemic risk. *Sci. Rep.* **6**, 39467. <https://doi.org/10.1038/srep39467> (2016).
40. Simmons, B. I. *et al.* bmotif: a package for motif analyses of bipartite networks. *Methods Ecol. Evol.* **10**, 695–701. <https://doi.org/10.1111/2041-210X.13149> (2019).
41. Saracco, F. *et al.* Inferring monopartite projections of bipartite networks: an entropy-based approach. *New J. Phys.* **19**, 053022. <https://doi.org/10.1088/1367-2630/aa6b38> (2017).
42. Bertinelli, L. & Black, D. Urbanization and growth. *J. Urban Econ.* **56**, 80–96. <https://doi.org/10.1016/j.jue.2004.03.003> (2004).
43. Pugliese, E., Zaccaria, A. & Pietronero, L. On the convergence of the fitness-complexity algorithm. *Eur. Phys. J. Spec. Top.* **225**, 1893–1911. <https://doi.org/10.1140/epjst/e2015-50118-1> (2016).
44. Gindelsky, R. J. L. C. M. *Demography, Urbanization and Development: Rural Push, Urban Pull and... Urban Push?* (The World Bank, 2015). <https://doi.org/10.1596/1813-9450-7333>.
45. Gollin, D., Jedwab, R. & Vollrath, D. Urbanization with and without industrialization. *J. Econ. Growth* **21**, 35–70. <https://doi.org/10.1007/s10887-015-9121-4> (2016).
46. Michaels, G., Rauch, F. & Redding, S. J. Urbanization and structural transformation *. *Q. J. Econ.* **127**, 535–586. <https://doi.org/10.1093/qje/qjs003> (2012).
47. Blondel, V. D., Guillaume, J.-L., Lambiotte, R. & Lefebvre, E. Fast unfolding of communities in large networks. *J. Stat. Mech. Theory Exp.* **2008**, P10008. <https://doi.org/10.1088/1742-5468/2008/10/p10008> (2008).
48. Turok, I. & McGranahan, G. Urbanization and economic growth: the arguments and evidence for Africa and Asia. *Environ. Urban.* **25**, 465–482. <https://doi.org/10.1177/0956247813490908> (2013).
49. Verona, G. & Ravasi, D. Unbundling dynamic capabilities: an exploratory study of continuous product innovation. *Ind. Corporate Change* **12**, 577–606. <https://doi.org/10.1093/icc/12.3.577> (2003).
50. Esfahani, H. S. & Ramirez, M. T. Institutions, infrastructure, and economic growth. *J. Dev. Econ.* **70**, 443–477. [https://doi.org/10.1016/S0304-3878\(02\)00105-0](https://doi.org/10.1016/S0304-3878(02)00105-0) (2003).
51. Di Clemente, R., Chiarotti, G. L., Cristelli, M., Tacchella, A. & Pietronero, L. Diversification versus specialization in complex ecosystems. *PLoS One* **9**, e112525. <https://doi.org/10.1371/journal.pone.0112525> (2014).
52. Hausmann, R. & Hidalgo, C. Country diversification, product ubiquity, and economic divergence. *SSRN Electron. J.* <https://doi.org/10.2139/ssrn.1724722> (2010).

Acknowledgements

Riccardo Di Clemente as Newton International Fellow of the Royal Society acknowledges support from the Royal Society, the British Academy, and the Academy of Medical Sciences (Newton International Fellowship, NF170505). The authors would like to thank Fabio Saracco, Enrico Ubaldi, Bernardo Monechi, Andrea Zaccaria, Andrea Gabrielli, Luciano Pietronero and Marta C. González for the insightful discussions and comments.

Author contributions

R.D.C. and E.S. designed the study and performed the research. R.D.C. analyzed the data and generated the plots, E.S. created the Map. R.D.C., E.S. and M.B. wrote the paper. All the authors gave the final approval for the publication.

Competing interests

The authors declare no competing interest. The corresponding author is responsible for submitting a competing interests statement on behalf of all authors of the paper.

Additional information

Correspondence and requests for materials should be addressed to R.D.C.

Reprints and permissions information is available at www.nature.com/reprints.

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2021