

Papers in Evolutionary Economic Geography

#18.30

The Principle of Relatedness

César A. Hidalgo & Pierre-Alexandre Balland & Ron Boschma
& Mercedes Delgado & Maryann Feldman & Koen Frenken &
Edward Glaeser & Canfei He & Dieter F. Kogler & Andrea
Morrison & Frank Neffke & David Rigby & Scott Stern & Siqi
Zheng & Shengjun Zhu



Utrecht University

Urban & Regional research centre Utrecht

The Principle of Relatedness

César A. Hidalgo^{1*}, Pierre-Alexandre Balland^{2*}, Ron Boschma^{2,3}, Mercedes Delgado⁴, Maryann Feldman⁵, Koen Frenken⁶, Edward Glaeser^{7,8}, Canfei He⁹, Dieter F. Kogler¹⁰, Andrea Morrison², Frank Neffke¹¹, David Rigby¹², Scott Stern^{4,8}, Siqi Zheng¹³, and Shengjun Zhu⁹

¹ Collective Learning Group, The MIT Media Lab; Massachusetts Institute of Technology

² Department of Human Geography and Planning; Utrecht University

³ Business School, University of Stavanger

⁴ MIT Sloan School of Management

⁵ Department of Public Policy, University of North Carolina at Chapel Hill

⁶ Copernicus Institute of Sustainable Development, Utrecht University

⁷ Department of Economics, Harvard University

⁸ National Bureau of Economic Research

⁹ College of Urban and Environmental Sciences, Peking University

¹⁰ Spatial Dynamics Lab and School of Architecture, Planning & Environmental Policy; University College Dublin

¹¹ Center for International Development Harvard University

¹² Department of Geography & Department of Statistics; University of California, Los Angeles

¹³ Department of Urban Studies and Planning, MIT; School of Civil Engineering, & Tsinghua University

*contributed equally.

hidalgo@mit.edu

Abstract. The idea that skills, technology, and knowledge, are spatially concentrated, has a long academic tradition. Yet, only recently this hypothesis has been empirically formalized and corroborated at multiple spatial scales, for different economic activities, and for a diversity of institutional regimes. The new synthesis is an empirical principle describing the probability that a region enters—or exits—an economic activity as a function of the number of related activities present in that location. In this paper we summarize some of the recent empirical evidence that has generalized the principle of relatedness to a fact describing the entry and exit of products, industries, occupations, and technologies, at the national, regional, and metropolitan scales. We conclude by describing some of the policy implications and future avenues of research implied by this robust empirical principle.

Keywords: Economic Complexity, Relatedness, Economic Geography.

1 Introduction

Why are human creativity and productivity concentrated in space and time? Is the computer revolution in Silicon Valley similar to England's Industrial Revolution or to Florence's Renaissance?

An old idea in economic geography is that in dense industrial clusters the secrets of a trade are, "as it were in the air." [1] James Watt's steam engine required the metal works of John "Iron Mad" Wilkinson, just like Apple's Macintosh required the idea of the Graphical User Interface (GUI) developed at Xerox's PARC. The idea that skills, technology, and knowledge, are spatially concentrated, has a long academic tradition, dating back to Johan von Thünen and Alfred Marshall in the nineteenth century, and to Harold Hotelling, Walter Christaller, August Lösch, Waldo Tobler, Jane Jacobs, and Michael Porter in the twentieth. Yet, only recently this hypothesis has been empirically formalized and corroborated at multiple spatial scales, for different economic activities, and for a variety of institutional regimes. The new synthesis is an empirical principle describing the probability that a region enters (or exits) an economic activity as a function of the number of related activities present in that location. This synthesis is *the principle of relatedness*.

2 The Principle of Relatedness

In principle, we say that two activities, such as products, industries, or research areas are related when they require similar knowledge or inputs (Figure 1.A). For instance, shirts and blouses are related because they are manufactured using similar materials and technologies. In practice, however, which inputs and knowledge are used in a production process are at best imperfectly observed. The modern methods used to *infer* relatedness are agnostic about the specific inputs or complementarities that drive relatedness among activities. They look, for instance, at the co-export of products [2], the flow of labor among industries [3], or combined measures of input-output links and shared labor pools [4]. This methodological flexibility, has allowed scholars from a variety of fields to document a robust and reproducible relationship between the probability that a location will develop expertise in a new industry [5–7], technology [8,9], research area [10], product [2], or occupation [11], and the number of related activities that are already present in that location.

Hidalgo et al. (2007)[2] introduced the idea of the product space—a network connecting products that are likely to be exported in tandem—to show that the probability that a country will start exporting a product increases with the number of related products that this country already exports (Figure 1.B). Neffke et al. (2011)[5], and Zhu et

al. (2017)[6], looked respectively, at Sweden and China to show that the probability that an industry will enter a region increases with the number of related industries present in it. Kogler et al. (2013)[8] and Boschma et al. (2015)[9] connected technology classes that co-occur in patents to show that cities in the United States were more likely to begin patenting in a technology class when they had expertise in related technologies (Figure 1.C). Guevara et al. (2016)[10], connected research areas to show that the probability that a scholar, university, or country, starts publishing in a new research area increases with the number of related areas in which a scholar, university, or country, has expertise (Figure 1.D).

These efforts have generalized the principle of relatedness from a simple observation, to a formal empirical result that is valid for multiple spatial scales, economic activities, and institutional backgrounds. Yet, these results have also demonstrated the empirical strength of the principle. In Figures 1b-1d, the probability of entering an activity rises between eight-fold and twenty-fold when we move from an unrelated activity to a related activity. This shows that the principle of relatedness is not only robust and ubiquitous, but also, strong.

But does the principle of relatedness teach us something about economic development? And is it a principle that matters?

The high degree of reproducibility of this principle hints at something fundamental: the variety of mechanisms by which economies and organizations learn. The principle of relatedness is not the result of trivial input output relationships, like needing iron to make steel. Material input-output relationships used to be important when transportation costs were relatively high. In the world of atoms, every city had a newspaper, and every paperboy had a route. Today, each country has only a few newspapers, and a few search engines are enough to serve the whole world. In the last decades the relative cost of moving knowledge, vis-à-vis the cost of moving the fruits of knowledge (products), has increased. Knowledge continues to be embodied in networks of people [12], and is concentrated in a few places, while bits and products travel swiftly around the world.

So why should we care about this “sticky knowledge principle?” On the one hand, new measures of relatedness are helping advance a more pragmatic industrial policy. The specificity of the principle is an antidote to the traditional temptations of industrial policy: the “cathedrals in the desert” built in naïve attempts to engineer development¹, or the gradualism that emerges when private sector companies capture public investment for sectors that are already well developed. Also, the principle is a remedy for the “everyone wants to be the new Silicon Valley” fever. But inoculation from erroneous ideas is not all that the principle provides.

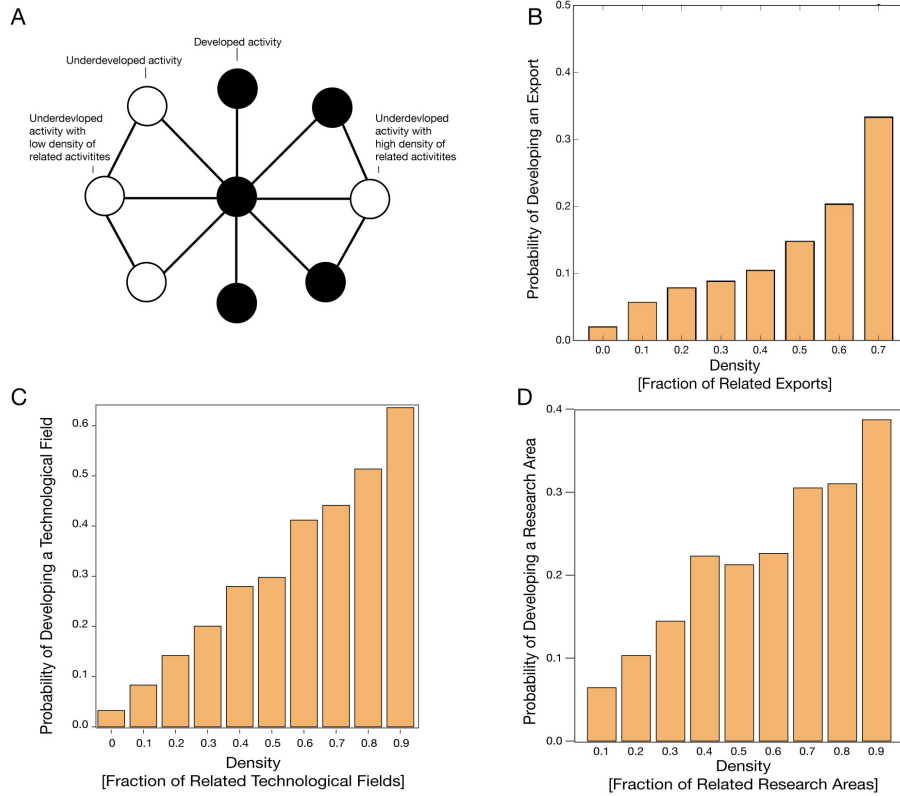


Figure 1. A. The number of related activities present in a location can be measured using density: a weighted average of the fraction of related activities that are present in that location. Relatedness is represented using a network, where nodes represent activities, such as products, research areas, or technology classes, and links connect related activities (products that are co-exported, or research areas that share authors). Black nodes indicate the activities that are present in a location. The activities that are not present in that location are shown in white. The node on the far left represents an activity with low density (few related activities present in that location, 1 out of 3). The node on the far right is an activity with high density (with many related activities present in that location, 3 out of 3). **B.** Probability that a country will develop comparative advantage in a product in a four-year period (that it will export more of that product than what should be expected from that country's total exports and the size of that product's market) as a function of the density of related products (fraction of related products already exported by that country). **C.** Probability that a city (U.S. Metropolitan Statistical Area) will produce more patents in a technology class than what is expected from that city's total patenting volume and the size of the technology class as a function of the density of related technological fields. **D.** Probability that a country will develop a research area in a period of three years (that it will publish more papers than what is expected from that country's total publication volume and the size of the field) as a function of the density of related research areas.

What is sometimes counterintuitive, is that the principle of relatedness is not about over-specialization [13]. It is about understanding the unique paths that lead to diversification. Just like our understanding of gravity is linked to our dream to fly, our understanding of relatedness is linked to our desire to break the shackles of path dependency. The principle is linked to the desire of diversifying Chile away from Copper, or to have an Arab peninsula that is not solely reliant on exporting oil.

The principle also suggest that industrial policy should not be centered solely on identifying promising industries, but on identifying mechanisms that facilitate knowledge flows among industries and regions [14]. The policies supported by the principle are those focused on attracting the knowledge that regions are missing by facilitating the flow of the people who carry that knowledge [15], and by creating social bridges to the places where that knowledge is present. Some of these policies may involve traditional but important industrial policy instruments, such as industrial parks [16–18]. Other policies, focused on knowledge flows and collective learning, can sometimes suggest counterintuitive results. For instance, much of the world has human capital development policies that subsidize graduate studies while demanding the immediate return of students to their homeland. Yet, the bridges created by these scholars—and the respective “brain flows” they generate—can have a bigger impact in their homeland’s long-term collective learning than the repatriation of the human capital they absorbed in four years of college [19]. The principle of relatedness invites us to evaluate policies not based on short-term winners and loser, but on their ability to contribute to collective learning.

The principle of relatedness also warns us about some pitfalls. For instance, the principle of relatedness is a force that increases spatial inequality and can reduce the ability of peripheral cities to develop. That is, the principle of relatedness is better news for Boston than for Buffalo.

So where should this research move next? Exploring the channels that promote collective learning, unpacking the idea of relatedness, and identifying how and when countries and regions deviate from this principle are all fertile avenues of research. Recently, Pinheiro et al. showed that countries tend to deviate from this principle in about 7% of all cases, and that they are more likely to do so at an intermediate level of diversification [20]. Lee and Malerba, and Lee and Ki, showed that industry life-cycles provide windows of opportunities that can reshuffle industrial leaders [21,22]. Boschma et al. showed that relatedness is a stronger force in countries with more coordinated forms of capitalism, than in countries with more liberal forms. Zhu et al. showed that policy interventions have been effective at moving Chinese regions into relatively unrelated areas [6]. Petralia et al (2017) showed that technological relatedness matters more for developing than advance economies [23]. Murray et al. showed that strong intellectual property restrictions can reduce the number of new entrants in innovative activities [24]. And Uhlbach et al. [25] showed that in Europe R&D subsidies are more effective when they target areas with an intermediate level of relatedness, not too unrelated, but neither too closely related.

There are also theoretical questions, such as identifying the optimal diversification strategies that a country or region should follow. After all, recent work by Alshamsi et al. has shown that targeting the most related product is not always the optimal strategy,

and that optimal diversification strategies need to worry about the right time to target relatively unrelated activities [26].

The good news is that learning appears to pay off for the economies who master it. Knowledge intense economies tend to grow faster [27,28] and be less unequal [29] at the national level, than less knowledge intense economies, at the same levels of income, human capital, and similar institutions. Also, economies with a diverse set of related industries tend to experience faster employment growth in those sectors [30], and more entrepreneurship [32]. Workers in industries that enjoy the benefits of relatedness also are more robust to displacement shocks, since they find related work in their region more easily than comparable workers displaced in regions with few related activities [33]. These benefits tell us that efforts focused on promoting collective learning, among industries and regions, seem to be the right way to go.

For decades scholars have been trying to understand how cities, regions, and countries develop. The principle of relatedness does not provide a full answer to this question, but it signals a path. More importantly, the principle of relatedness is a beautiful illustration of the collaborative nature of science. Generalizing this principle to multiple scales, activities, institutional backgrounds, and metrics, has not been the work of a lone scholar, but the result of a diverse academic community involving geographers, economists, urban planners, physicists, and more. By building on each other's work, the scholars in this community have created a literature that supports the generalization of this principle. The reproducibility and robustness of this finding, therefore, does not rest on a single paper, instrumental variable, or key experiment, but on a growing literature that has replicated this principle tirelessly, demonstrating its solidity as a foundation for current and future research.

References

1. A. Marshall, *Principles of Economics* (Macmillan and Company, 1890).
2. C. A. Hidalgo, B. Klinger, A.-L. Barabási, and R. Hausmann, *Science* **317**, 482 (2007).
3. F. Neffke and M. Henning, *Strategic Management Journal* **34**, 297 (2013).
4. M. Delgado, M. E. Porter, and S. Stern, *J Econ Geogr* **16**, 1 (2016).
5. F. Neffke, M. Henning, and R. Boschma, *Economic Geography* **87**, 237 (2011).
6. S. Zhu, C. He, and Y. Zhou, *J Econ Geogr* **17**, 521 (2017).
7. J. Gao, B. Jun, A. "Sandy" Pentland, T. Zhou, and C. A. Hidalgo, ArXiv:1703.01369 [Physics, q-Fin] (2017).
8. D. F. Kogler, D. L. Rigby, and I. Tucker, *European Planning Studies* **21**, 1374 (2013).
9. R. Boschma, P.-A. Balland, and D. F. Kogler, *Ind Corp Change* **24**, 223 (2015).
10. M. R. Guevara, D. Hartmann, M. Aristarán, M. Mendoza, and C. A. Hidalgo, *Scientometrics* **109**, 1695 (2016).
11. R. Muneeppeerakul, J. Lobo, S. T. Shuttters, A. Gómez-Liévano, and M. R. Qubbaj, *PloS One* **8**, e73676 (2013).
12. C. Hidalgo, *Why Information Grows: The Evolution of Order, from Atoms to Economies* (Basic Books, New York, 2015).
13. E. L. Glaeser, H. D. Kallal, J. A. Scheinkman, and A. Shleifer, *Journal of Political Economy* **100**, 1126 (1992).

14. P.-A. Balland, R. Boschma, J. Crespo, and D. L. Rigby, *Regional Studies* 1 (2018).
15. F. Neffke, M. Hartog, R. Boschma, and M. Henning, **14.10**, (2014).
16. S. Zheng, W. Sun, J. Wu, and M. E. Kahn, *Journal of Urban Economics* **100**, 80 (2017).
17. Z. Chen, S. Poncet, and R. Xiong, *Journal of Comparative Economics* **45**, 809 (2017).
18. M. E. Kahn, W. Sun, J. Wu, and S. Zheng, *The Revealed Preference of the Chinese Communist Party Leadership: Investing in Local Economic Development versus Rewarding Social Connections* (National Bureau of Economic Research, 2018).
19. A. Saxenian, *The New Argonauts: Regional Advantage in a Global Economy* (Harvard University Press, 2007).
20. F. L. Pinheiro, A. Alshamsi, D. Hartmann, R. Boschma, and C. A. Hidalgo, ArXiv:1801.05352 [Physics, q-Fin] (2018).
21. K. Lee and F. Malerba, *Research Policy* **46**, 338 (2017).
22. K. Lee and J. Ki, *Research Policy* **46**, 365 (2017).
23. S. Petralia, P.-A. Balland, and A. Morrison, *Research Policy* **46**, 956 (2017).
24. F. Murray, P. Aghion, M. Dewatripont, J. Kolev, and S. Stern, *American Economic Journal: Economic Policy* **8**, 212 (2016).
25. W.-H. Uhlbach, P.-A. Balland, and T. Scherngell, (2017).
26. A. Alshamsi, F. L. Pinheiro, and C. A. Hidalgo, *Nature Communications* **9**, 1328 (2018).
27. C. A. Hidalgo and R. Hausmann, *PNAS* **106**, 10570 (2009).
28. R. Hausmann, C. A. Hidalgo, S. Bustos, M. Coscia, A. Simoes, and M. A. Yildirim, *The Atlas of Economic Complexity: Mapping Paths to Prosperity* (MIT Press, 2014).
29. D. Hartmann, M. R. Guevara, C. Jara-Figueroa, M. Aristarán, and C. A. Hidalgo, *World Development* **93**, 75 (2017).
30. M. Delgado, M. E. Porter, and S. Stern, *Research Policy* **43**, 1785 (2014).
31. K. Frenken, F. V. Oort, and T. Verburg, *Regional Studies* **41**, 685 (2007).
32. M. Delgado, M. E. Porter, and S. Stern, *J Econ Geogr* **10**, 495 (2010).
33. Neffke, Frank, Otto, Anne, and Hidalgo, Cesar A., *Papers in Evolutionary Economic Geography* **16.05**, (2016).

ⁱ Classic examples are the Volta River project in Ghana or the construction of a large Iron processing plant in Ipatinga Brazil.