Rivista Italiana di Economia Demografia e Statistica *Volume LXXIX n.4 Ottobre-Dicembre 2025*

THE GEOGRAPHY OF GREEN INNOVATION

Margherita Gerolimetto, Stefano Magrini, Alessandro Spiganti

Abstract. In the last decades, the geography of innovation activity became much more concentrated. By focusing on the metropolitan statistical area of residence of the inventors of patents granted by the United States Patents and Trademark Office between 1976 and 2020, we investigate whether this is true also for "green" innovation, *i.e.* patents covering mitigation or adaptation to climate change. We find a sharp increase in concentration across areas after the beginning of the 2000s, with areas that are generally more innovative also producing more green patents. Moreover, for an in-depth analysis of spatial dependence patterns, we firstly survey the literature to find tests suitable for comparing spatial patterns and then apply these tests to the data on green patenting in comparison with non-green patenting. We find significant differences between the two phenomena.

1. Introduction

It is an established result in the literature on innovation and agglomeration that innovation activities are spatially concentrated (e.g., Audretsch & Feldman, 1996; Buzard *et al.*, 2017). Some authors, in particular, view innovation essentially as an urban phenomenon (e.g., Florida, 2009). One of the underlying explanations, dating back to Marshall (1890), is that geographic proximity facilitates the transfer of knowledge. The idea that informal interactions are central to innovation and knowledge spillover has then become a fundamental element of recent theories if economic growth.

A sizeable literature has then provided empirical estimates of the size and the properties of local knowledge and productivity externalities. Jaffe *et al.* (1993) find that patents citations display a significant bias towards patens that were produced in the same state and metropolitan area. Greenstone *et al.* (2010) estimate significant agglomeration spillovers on Total Factor Productivity by comparing winning and losing countries bidding to attract large plants. Kerr and Kominers (2015) propose a theory of cluster formation based on firm's location and interaction choices and confront its predictions using data on patents citations by technology class, finding out that the geographical properties of innovation clusters are controlled by the spatial range of knowledge transmission which is specific to each technology class.

Several studies focus on the role of specialization and diversity in driving innovation and economic outcomes in cities (Glaeser and Gottlieb, 2009; Delgado *et al.* 2014). Berkes and Gaetani (2021) focus on the geography of unconventional innovation and on how economic geography shapes the creative content of innovation. In particular, they show that high-density areas tend to have an advantage in producing unconventional ideas and that the combination of ideas embedded into inventions is determined by the local technology mix.

In this paper we focus on “green" innovation, loosely speaking patents covering mitigation or adaptation to climate change and see if the way this type of innovation distributes over space shows any significant difference compared to “non-green” innovative activity. More in detail, we firstly create a database of utility patents granted to inventors residing in the conterminous US from 1976 to 2016 at the Metropolitan Statistical Area level. Then, we measure spatial concentration in green and non-green innovative activities, finding a sharp increase in concentration across areas after the beginning of the 2000s, with areas that are generally more innovative also producing more green patents. Then, we examine the evolution of the cross-sectional distribution of green and non-green patents per capita across MSAs. Furthermore, we survey the literature to find tests suitable for comparing spatial patterns and then apply two of these tests to the data on green patenting in comparison to non-green patenting; here, we find significant differences between the two phenomena. Finally, we analyse the intensity of the spatial association of green and non-green patents. Our results document the presence of positive spatial correlation and indicate the existence of a High-High cluster in the North-East of the nation.

2. Data

We measure innovation with the flow number of patents (an exclusionary right conferred for a set period to the patent holder, in exchange for sharing the details of the invention) filed between 1976 and 2016 and eventually granted by the United States Patents and Trademarks Office (USPTO). We associate a patent to a year using the application date, which is the year when the provisional application is considered complete by the USPTO.[1]

[1] We follow the patent literature in focusing on application year rather than on the award year. As noted by Lerner and Seru (2022), the motivation is that, whereas firms will generally tend to file for patents as soon as the discoveries are made in order to protect their intellectual property, the time at which the patent is granted depends on many external factors, like the technological area or the state of the patent office.

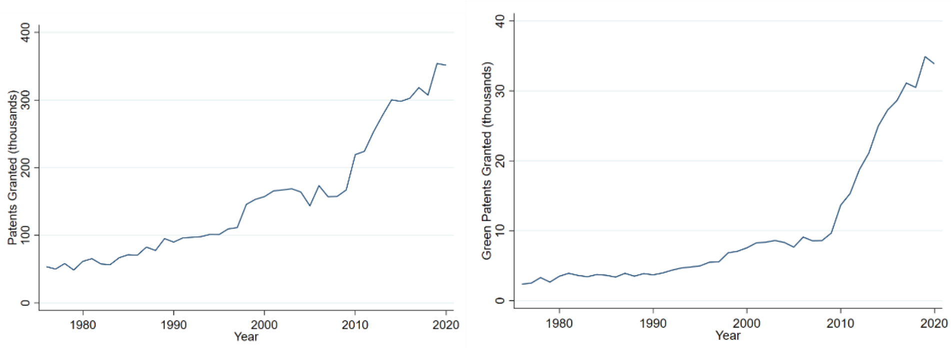
Withdrawn applications are excluded from the analysis. As common in the literature, we restrict our attention to utility patents (thus excluding design patents and plant patents), which cover the creation of a new or improved product, process or machine; these represent approximately 90% of all patents granted by USPTO.

We assign patents to areas according to the location in which the inventor resides as in, e.g., Castaldi and Lobs (2017), Aghion *et al.* (2019), Berkes and Gaetani (2021), Moretti (2021), which is extracted from patent text and used to determine latitude and longitude. We use the residential addresses of the inventors and not the one of the assignee (usually, the company that first owned the patent), because we are mainly interested in the location of processes that lead to inventions, whereas the assignee address often reflects the address of the corporate headquarters and not the R&D facility (Moretti, 2021). When a patent is coauthored by more than one inventor, we split it equally among them, as in Aghion *et al.* (2019), Berkes and Gaetani (2021) and Moretti (2021). Henceforth, we attribute a fraction m/n of a patent to an area a, where n is the total number of inventors reported in that patent and m is the number of inventors in that patent who reside in area a.

In 2013, the USPTO and the European Patent Office introduced a new system of Cooperative Patent Classifications (CPC) that, unlike existing patent classifications such as the International Patent Classifications, can also be indexed with a focus on emerging technologies (Veefkind *et al.*, 2012). These new classifications have been backtracked into the existing databases. We exploit this system by classifying a patent as "green" if it belongs to at least one subclass in the Y02/Y04S scheme, like e.g. Corrocher *et al.* (2021). Within the CPC, the Y02 class covers technologies which control, reduce, or prevent anthropogenic emissions of greenhouse gases (GHG), in the framework of the Kyoto Protocol and the Paris Agreement, and technologies which allow the adaptation to the adverse effects of climate change, whereas the Y04S covers systems integrating technologies related to power network operation, communication, or information technologies for improving the electrical power generation, transmission, distribution, management, or usage.

We focus on inventors residing in the conterminous United States (i.e., the 48 adjoining states and the District of Columbia). In our database, there are 3,836,007 utility patents granted to inventors residing in the conterminous US, of which 249,501 belong to at least one green subclass. It is well known that the number of patents issued by the USPTO annually has steadily increased since the 1990s, as shown in the left panel of Figure 1. The right panel of Figure 1 shows that the number of green patents granted has been growing slowly up until 2006, when an impressive acceleration can be observed, in line with the findings by Corrocher *et al.* (2021).

Figure 1 – *Evolution over time of the patents granted by USPTO.*



(a) All patents *(b) Green patents*

The first (second) panel shows the number of utility (green) patents granted by USPTO in any given year between 1976 and 2020. (Own elaborations using data from USPTO).

In terms of areas, we focus on Metropolitan Statistical Areas (MSAs), i.e. regions consisting of a large urban core together with surrounding communities that have a high degree of economic and social integration with the urban core. We consider MSAs for various reasons. First, MSAs represent economic spatial units and so are considered more appropriate to study economic dynamics than states, regions, or even counties Drennan (2005). Second, innovation is mainly an urban phenomenon; for example, the vast majority of patents in our dataset come from inventors residing in a metropolitan area (approximately 85% for utility patents and 83% for green patents). Third, there is large heterogeneity across MSAs in terms of capacity to innovate.

We assign an inventor location to a MSA using the 2010 Cartographic Boundary Files provided by the United States Census Bureau.[2]

3. Concentration

There is ample evidence that research and development activities tend to be more concentrated than manufacturing activities (Buzard *et al.*, 2017). Moreover, US patenting activities have become more geographically concentrated since the end of the last century (Castaldi and Lobs, 2017; Andrews and Whalley, 2021; Forman and

[2] Source: https://www.census.gov/geographies/mapping-files/time-series/geo/cartographic-boundary.html. The definition used for the identification of MSAs has evolved over time, with significant changes made especially around census years: results are qualitatively identical independently of the boundary files used.

Goldfarb, 2021; Magrini and Spiganti, 2024). For example, Andrews and Whalley (2021) and Forman and Goldfarb (2021) report a particularly pronounced increase in the geographic concentration of patenting at the US county level starting from the 1990s, while Castaldi and Lobs (2017) highlight that concentration is even more pronounced for highly-cited patents at the state level.

Here, we first measure the spatial concentration in green innovative activities and then examine the evolution of the cross-sectional distribution of green patents per capita across MSAs, comparing these to those for non-green patents. We do so using the Herfindahl–Hirschman Index and the Dartboard Index.

The Herfindahl–Hirschman Index (HHI) provides a measure of the size of innovation in an area in relation to the overall amount of innovation in the nation. For each year t and each area a, where $a = 1,\ldots, A$, the Herfindahl–Hirschman Index concentration index is:

$$HHI_t = \sum_{a=1}^{A} (SharePat_{at})^2 \quad (1)$$

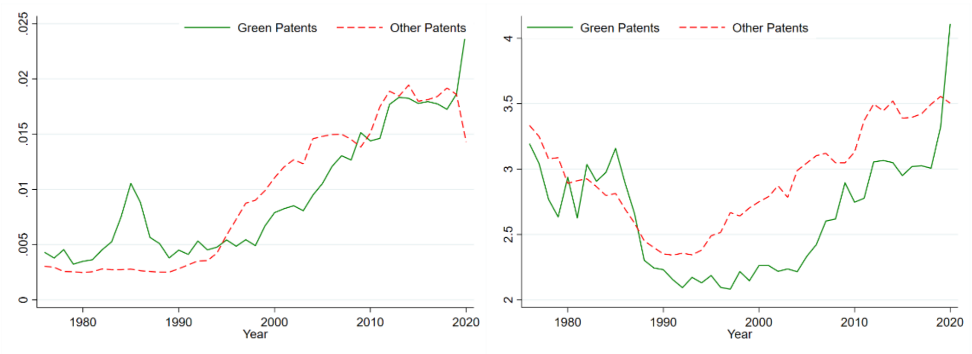
where $SharePat_{at}$ is the patent share of area a. The scale of the index is such that a value of zero can be interpreted as indicating a complete lack of concentration, whereas a value of one would indicate that all patenting occurs in one area.

The spatial concentration index by Ellison and Glaeser (1997), also known as Dartboard Index (*EGI*), compares the observed spatial distribution of innovation to what it would have been if it was proportional to population distribution. In particular, for each year t and all areas a, where a = 1,…,A, the Dartboard Index is

$$EGI_t = \frac{\sum_{a=1}^{A} (SharePat_{at} - SharePop_{at})^2}{1 - \sum_{a=1}^{A} SharePop_{at}^2} \quad (2)$$

where $SharePat_{at}$ and $SharePop_{at}$ are, respectively, the share of patents granted and the share of population living in area a and year t. The scale of this index is such that a value of zero can be interpreted as indicating a complete lack of agglomerative forces, whereas a value of one would indicate that all patenting occurs in one area.

Figure 2 – *Evolution over time of the concentration of the patents.*



(a) Dartboard Index *(b) HH Index*

The first (second) panel shows the Dartboard (Herfindahl-Hirschman) innovation intensity concentration index across metropolitan statistical areas in the United States between 1976 and 2020. (Own elaborations using data from USPTO).

Figure 2, which reports the evolution of the indexes, documents a sharp increase in concentration across areas; it starts in the mid-1990s for non-green patents and at the beginning of the 2000s for green ones.[3] Specifically, we observe that the HHI exhibits an initial decline in concentration, which Andrews and Whalley (2021) suggest could be due to improvements in transportation, increasing access to higher education, and the expansion of direct federal funding for research. Conversely, the increase in concentration goes hand-in-hand with increasing assortative sorting of skills across cities and the emergence of superstar cities. Towards the end of the 2020s, concentration shows a tendency to decline, at least for non-green patents.

4. Geographical patterns

4.1. Stochastic spatial processes

Spatial data can be thought as resulting from observations of a stochastic process $\{Z(s): s \in D \subset R^d\}$, where D is a set of R^{d}, $d = 2$, and Z(s) denotes the attribute we observe at s. In case of patents, spatial referenced data can be of two types, according to the assumptions made on **D**:

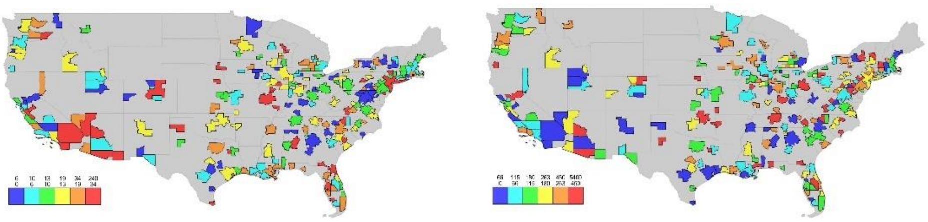
[3] The spike in 1987 in the concentration of green patents is due to an increase of the fraction of green patents coming from innovators residing in Pittsburgh, PA, and in particular in Allegheny County (by 4.2 percentage points) and Westmoreland County (by 1.9 percentage points), which mostly reversed in 1988.

i) areal (or lattice) data: the domain is a fixed countable collection of irregular areal units at which variables are observed. Here patenting of given characteristics is aggregated into areas that form a partition of the study region.

ii) point patterns: the domain is random. The index set gives the location of random events of the spatial point pattern.

Here we will focus on areal-based spatial point patterns. In the graphs below we present the geographical distribution of the patents data analysed in this paper.

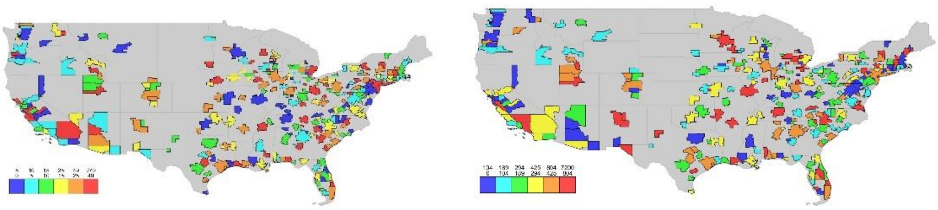
Figure 3 – *Spatial point pattern of patents in 1980.*



(a) Green patents *(b) Other patents*

Notes: patent activity is divided into sextiles going from blue (low activity) to red (high).

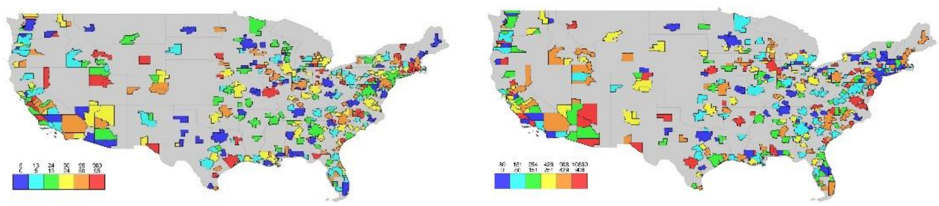
Figure 4 – *Spatial point pattern of patents in 2000.*



(a) Green patents *(b) Other patents*

Notes: patent activity is divided into sextiles going from blue (low activity) to red (high).

Figure 5 – *Spatial point pattern of patents in 2020.*



(a) Green patents *(b) Other patents*

Notes: patent activity is divided into sextiles going from blue (low activity) to red (high).

At a first glance, in 1980 and 2000 green patenting tends to concentrate in area where non-green patenting is not particularly concentrated. A clear example of this is the South-West of the nation. Conversely, in 2020 the spatial patters appear to be more alike.

4.2. Testing for similarity of point patterns

The method that is mostly adopted in the literature to test for similarity of spatial point patterns is Andresen's (2009) test; this test however is not applicable in the present case, due to some peculiarities of the data: the USPTO usually attributed the location of the inventor to the central point of the county of residence, thus leading to more than one observation per a specific spatial point in each area. Therefore, we adopt a sort of combined approach. In a first step, we will resort to the so-called modified T test for assessing the correlation between two spatial processes proposed by Clifford *et al.* (1989) and further modified by Dutilleul *et al.* (1993). The null hypothesis is the absence of correlation between two spatial point patterns; the original test statistics has been modified to improve the approximation of the variance of the sample correlation coefficient.

Then, in a second step, we adopt two similarity tests by Alba-Fernandez *et al.* (2016). The null hypothesis here is the similarity of two spatial point patterns. Both tests move from a preliminary phase of counting the number of occurrences per each areal units and the idea is to find out whether it is obtained a sample of a multinomial distribution. To do this, in one case (the so-called T_1 test) it is adopted a Chi-square test for homogeneity, while the second test (the so-called T_2 test) is based on the negative of Matusita's affinity. For both tests the reference distribution is χ^2. For space reasons, we refer to the papers by Dutilleul *et al.* (1993) and Alba-Fernandez *et al.* (2016) for the formal details of the tests and here we focus instead on the results on our data, summed in the following tables:

Table 1 – *Spatial point pattern correlation test: null hypothesis is absence of correlation.*

Year	T-value	P-values
1980	0.0003	0.938
2000	0.0060	0.750
2020	0.0100	0.704

Notes. The test is performed using the SpatialPack package in R. (Vallejos et al., 2020).

Table 1 reveals that we cannot reject the null hypothesis of absence of correlation between the patterns; this can be considered as signal of lack of similarity between the point patterns. Given this result, we move to the second set of tests to confirm or disconfirm this evidence.

Table 2 – *Similarity tests: null hypothesis is the similarity of two spatial point patterns.*

Year	T_1	T_2	$\chi^2(0.90)$	$\chi^2(0.95)$	$\chi^2(0.99)$
1980	580.330	3540.883	262.117	270.684	287.247
2000	505.392	2944.563	308.614	317.888	335.776
2020	195.379	8102.683	371.719	381.872	401.408

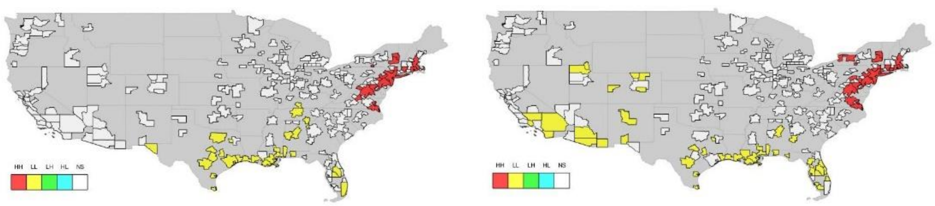
Notes. H0 is rejected for large values of the statistic. The test is run in R following Alba-Fernandez et al. (2016).

Table 2 shows that we can reject the null hypothesis that the patterns are similar for 1980, 2000 for both tests. As for 2020, we cannot reject the null hypothesis for T_1. Looking at these results, it seems that the geographical distribution of patenting activities, somewhat different at the beginning, has become more similar towards the end of the period.

5. Spatial correlation analysis

Local Moran's I is a local spatial autocorrelation statistic developed by Anselin (1995) and based on the Moran's I statistic. It is a Local Indicator of Spatial Association (LISA) and, consequently, it has the following two properties: i) for each observation, it gives an indication of the extent of significant spatial clustering of similar values around that observation ii) the sum of LISAs for all observations is proportional to a global indicator of spatial association. For space reasons, we refer to the papers by Anselin (1995) for the formal details and here we focus instead on the results on our data. Below there is a list of graphs presenting the local Moran significance maps for our data, obtained with a Matlab code. In particular, we show local associations which are significant at the 15% level, using a spatial weight matrix based on 15% nearest-neighbours.

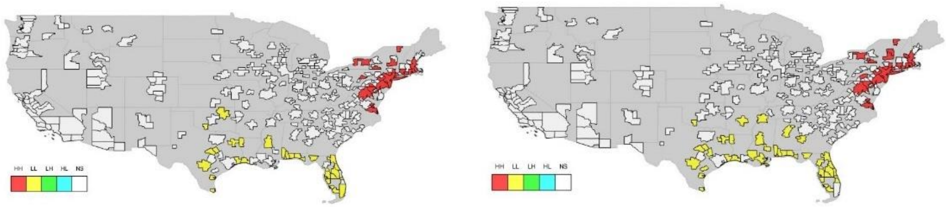
Figure 6 – *Local Moran map 1980.*



(a) Green patents *(b) Other patents*

Notes: red indicates High-High significant local spatial correlations; yellow indicates Low-Low significant local spatial correlations.

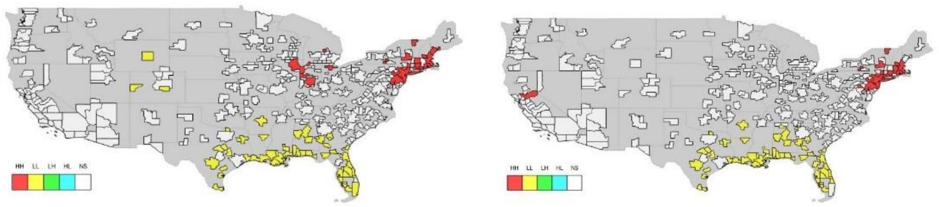
Figure 7 – *Local Moran 2000.*



(a) Green patents *(b) Other patents*

Notes: red indicates High-High significant local spatial correlations; yellow indicates Low-Low significant local spatial correlations.

Figure 8 – *Local Moran 2020.*



(a) Green patents ***(b)*** *Other patents*

Notes: red indicates High-High significant local spatial correlations; yellow indicates Low-Low significant local spatial correlations.

We observe that Moran's *I* test shows the presence of positive spatial correlation and indicates the existence of a High-High cluster in the North-East of the nation. However, the spatial extent of this cluster seems to have reduced by 2020.

6. Conclusions

Patenting activity has increased in general, more so in the case of green patenting after 2010. Concentration shows similar trends overall, but with a visible difference in the last years where concentration of green patents strongly increases while concentration of other patents has remained constant or decreased. The geographical distribution of patenting activities, somewhat different at the beginning, has become more similar towards the end of the period. The Moran's *I* test show the presence of positive spatial correlation and indicates the existence of a High-High cluster in the North-East of the nation.

References

AGHION P., AKCIGIT U., BERGEAUD A., BLUNDELL R., HEMOUS D. 2019. Innovation and top income inequality. *Review of Economic Studies*, Vol. 86, No. 1, pp. 1–45.

ALBA-FERNANDEZ M.V., ARIZA-LOPEZ F.J., JIMENEZ-GAMERO M.D., RODRIGUEZ-AVI J. 2016. On the similarity analysis of spatial patterns. *Spatial Statistics*, Vol. 18, pp. 352-362,

ANDRESEN M.A. 2009. Testing for similarity in area-based spatial patterns: a non parametric Monte Carlo approach. *Applied Geography*, Vol. 29, No. 3, pp. 333–345.

ANDREWS M.J., WHALLEY A. 2021. 150 years of the geography of innovation. *Regional Science and Urban Economics*, Vol. 94, pp. 103-127

ANSELIN L. 1995. Local Indicators of Spatial Association LISA. *Geographical Analysis*, Vol. 27, pp. 93-115.

AUDRETSCH D.B., FELDMAN M.P. 1996. R&D spillovers and the geography of innovation and production. *The American Economic Review*, Vol. 86, No. 3, pp. 630–640.

BERKES E., GAETANI R. 2021. The Geography of Unconventional Innovation. *The Economic Journal*, Vol. 131, pp. 1466–1514.

BUZARD K., CARLINO G.A., HUNT R.M., CARR J.K., SMITH T.E, 2017. The agglomeration of American R&D labs. *Journal of Urban Economics*, Vol. 101, pp. 14–26.

CASTALDI C, LOS B, 2017. Geographical patterns in US inventive activity 1977–1998: The “regional inversion” was underestimated. *Research Policy*, Vol. 46, No. 7, pp. 1187–1197.

CLIFFORD P., RICHARDSON S., HEMON D. 1989. Assessing the significance of the correlation between two spatial processes. *Biometrics*, Vol. 45, pp. 123-134.

CORROCHER N., MALERBA F., MORRISON A. 2021. Technological regimes, patent growth, and catching-up in green technologies. *Industrial and Corporate Change*, Vol. 30, No. 4, pp. 1084–1107.

DELGADO M., PORTER ME., STERN S. 2014. Clusters, convergence, and economic performance. *Research Policy*, Vol. 43, pp. 1785-1799

DRENNAN MP. 2005. Possible sources of wage divergence among metropolitan areas of the United States. *Urban Studies*, Vol. 42, No. 9, pp. 1609–1620.

DUTILLEUL P., CLIFFORD P., RICHARDSON S., HEMON D. 1993. Test for assessing the correlation between two spatial processes. *Biometrics*, Vol. 49, No. 1, pp. 305-314.

ELLISON G., GLAESER E. 1997. Geographic concentration in U.S. manufacturing industries: A dartboard approach. *Journal of Political Economy,* Vol. 105, No. 5, pp. 889–927.
FLORIDA R. 2009. *Who's your city?* New York: Basic Books.
FORMAN C., GOLDFARB A. 2021. Concentration and agglomeration of IT innovation and entrepreneurship: Evidence from patenting.
GLAESER EL., GOTTLIEB J.D. 2009. The Wealth of Cities: Agglomeration Economies and Spatial Equilibrium in the United States. *Journal of Economic Literature*, Vol. 47, No. 4, pp. 983–1028.
GREENSTONE M., HORNBECK R., MORETTI E. 2010. Identifying Agglomeration Spillovers: Evidence from Winners and Losers of Large Plant Openings. *Journal of Political Economy*. Vol. 118, No. 3 pp. 536-598.
JAFFE A., TRAJTENBERG M., HENDERSON R, 1993. Geographic Localization of Knowledge Spillovers as Evidenced by Patent Citations. *The Quarterly Journal of Economics.* Vol. 108, No. 3, pp. 577-598
KERR W., KOMINERS S. 2015. Agglomerative Forces and Cluster Shapes. *The Review of Economics and Statistics*, 2015, vol. 97, No. 4, 877-899..
LERNER J., SERU A. 2022. The Use and Misuse of Patent Data: Issues for Finance and Beyond. *The Review of Financial Studies*, Vol. 35, No. 6, pp. 2667–2704.
MAGRINI S., SPIGANTI A. 2024. A tale of two cities: Communication, innovation, and divergence. *Economic Inquiry*, Vol. 62, No. 1, pp. 390–413.
MARSHALL A. 1890. *Principles of economics*. London: Macmillan.
MORETTI E. 2021. The Effect of High-Tech Clusters on the Productivity of Top Inventors. *American Economic Review*, Vol. 11, No. 10, pp. 3328–3375.
VALLEJOS R., OSORIO F., BEVILACQUA M. 2020. *Spatial Relationships Between Two Georeferenced Variables: With Applications in R*. Cham: Springer.
VEEFKIND V., HURTADO-ALBIR J., ANGELUCCI S., KARACHALIOS K., THUMM N. 2012. A new EPO classification scheme for climate change mitigation technologies. *World Patent Information*, Vol. 34, No. 2, pp. 106–111.

Margherita GEROLIMETTO, Università Ca' Foscari Venezia, Dipartimento di Economia, margherita.gerolimetto@unive.it
Stefano MAGRINI, Università Ca' Foscari Venezia, Dipartimento di Economia, stefano.magrini@unive.it
Alessandro SPIGANTI, Università Ca' Foscari Venezia, Dipartimento di Economia, alessandro.spiganti@unive.it