Using sequence analysis to uncover the spatio-temporal emergence of bioeconomy innovations

Sebastian Losacker * [a b]
Nicolas Befort [c d]
Lukas Kriesch [a]
Stéphane Lhuillery [c e]

* corresponding author (sebastian.losacker@geogr.uni-giessen.de)

[a] Department of Geography, Justus Liebig University Giessen, Germany
[b] CIRCLE—Center for Innovation Research, Lund University, Sweden
[c] Chair in Bioeconomy and Sustainable Development, The World We Want AoE, NEOMA Business School, France
[d] LISIS (INRAE, CNRS, University Gustave Eiffel), France
[e] BETA-CNRS, Bureau d'Economie Theorique et Appliquee & Centre National de la Recherche Scientifique, Universite de Strasbourg, France

Abstract

This paper aims to uncover regional trajectories towards a bioeconomy with a focus on bio-based technologies. We build on the geography of innovation literature and argue that potential pathways towards regional bioeconomies are quite heterogeneous, thus requiring place-based policy strategies to advance the bioeconomy and its innovations. Empirically, the paper combines two unique patent datasets to reveal how the bioeconomy patent portfolio of 617 regions from 27 OECD countries has changed from 1982 to 2014. We utilize geographical sequence analysis, a novel tool recently introduced to geographical research, and shift-share techniques to categorize and better understand the regional trajectories.

Key words: Bioeconomy; Sequence Analysis; Geography of Innovation; Sustainability Transitions

JEL: O31, R11, R12

Acknowledgements: Sebastian Losacker and Lukas Kriesch acknowledge financial support from the German Federal Ministry of Education and Research (BMBF 031B1281).

1 Introduction

In the face of global ecological crises such as climate change, biodiversity loss, and resource depletion, governments worldwide have introduced strategies and agendas aimed at transitioning from a linear economy, which relies mainly on fossil and non-renewable resources, to a more circular, bio-based economy (Patermann & Aguilar, 2021; Proestou et al., 2023). The *bioeconomy* has thus received increased interest from scholars and policymakers, leading to several attempts to study and better understand its progress (Kriesch & Losacker, 2024b; Wydra, 2020). It utilizes biological resources and knowledge to produce goods, services and energy across all sectors. In this regard, knowledge of bio-based technologies is a key pillar in the bioeconomy transition, as (novel) knowledge is a prerequisite for developing innovations that enable and accelerate this transition (Befort, 2023).

In this paper, we contribute to a better understanding of the progress of the bioeconomy transition by studying the development of bio-based technologies at the regional level. We build on the literature in economic geography and the geography of innovation, arguing that regional transition trajectories are likely to be very heterogenous, depending on various region-specific factors and the regional sectoral mix. Empirically, we use patent data as a proxy for knowledge development and innovation activities, combining two unique patent datasets: one indicating the precise geographical location of patent inventors (de Rassenfosse et al., 2019), and the other indicating whether a patent is related to the bioeconomy (Kriesch & Losacker, 2024a). We utilize this data to study how 617 regions (functional urban areas) from 27 OECD countries have changed their patent portfolios regarding bioeconomy patents from 1982 to 2014. We develop a typology of bioeconomy progress and study the unique regional trajectories using geographical sequence analysis, a novel method recently introduced to the toolkit of economic geographers (Losacker & Kuebart, 2024). In the next step, we use shift-share techniques to better understand whether region-specific factors related to the sectoral composition of the regional patent portfolio contribute more to the bioeconomy transition. The findings of our paper extend the literature on geography and innovation in the bioeconomy, and we develop recommendations for policymakers on implementing place-based policies to support the bioeconomy transition.

The remainder of this paper is structured as follows. In Section 2, we review the literature from the fields of economic geography and the geography of innovation to understand how regional transition paths may develop and how they might differ from each other. We also review the literature on regional bioeconomy transitions. In Section 3, we describe our empirical approach in detail and provide initial insights into regional bioeconomy trajectories. In Section 4, we present and discuss the main results. Section 5 concludes and provides policy recommendations.

2 Theoretical background

2.1 Establishing typologies of green path developments

The theoretical background that motivates this paper mainly comes from research in the fields of evolutionary economic geography and the geography of innovation. In this stream of literature, numerous studies aim to understand how regions and their knowledge bases change over time, providing an important foundation for the empirical aims of this paper.

The current debate on green industrial paths raises questions about the conditions for the emergence of environmental innovations (and thus bioeconomy innovations) and the potential for regions to develop competitive advantages in these fields over time. To this end, some authors are seeking to establish typologies of regional trajectories to account for similarities in otherwise unique and heterogenous developmental paths. Against this background, Trippl et al. (2020) have introduced the notion of green path development to describe the ecological transformation of industries on the regional level. More specifically, this notion is applied at the regional level from a regional innovation systems (RIS) perspective, accounting for the role of regional preconditions, institutional structures, and modes of innovation in both enabling and constraining (green) regional development. Based on an initial review of paths, the authors identify four: path renewal, path diversification, path importation, and path creation. Path renewal corresponds to a dynamic driven by the renewal of industry practices through the gradual adoption of environmental innovations. Path diversification corresponds to a dynamic in which brown industries use their dynamic capabilities to create new, complementary activities (Ayoub & Lhuillery, 2024; Le Bas & Scellato, 2014). Path importation involves the arrival of firms from outside a region, helping to structure a new trajectory. Path creation corresponds to the emergence of a new industry within the region, notably through academic spin-offs. This typology underlines the combined role of knowledge availability, institutional conditions, agency, and the availability of local assets for steering regional trajectories. From the perspective of the role of institutions, following Martin and Coenen (2014), Grillitsch et al. (2023) have also established a five-route typology, highlighting both the role of internal dynamics under the effect of institutional investments, entrepreneurial investments, and external dynamics linked to the existence of crises impacting the regional economy. Other works have sought to create typologies of regions in relation to the mode of development of environmental innovations. For example, Grillitsch and Hansen (2019) have established a typology distinguishing between peripheral regions, regions specializing in green industries, regions specializing in brown industries, and metropolitan regions. Each of these regions exhibits particular configurations, notably in terms of available knowledge and actor configurations.

Building on this basis, Hansmeier and Losacker (2024) have highlighted the existence of particular configurations depending on the types of actors present in a region. Through an exploration of regional trajectories, they found that most regions were non-specialized in green domains, and their dynamics could ultimately be driven by start-ups, incumbents, or a combination of both. In the case of specialized regions, their trajectory is sustained either by incumbents capable of generating spillover effects or by newcomers triggering change.

2.2 The role of knowledge in the developments of green paths

Beyond attempts to classify regional trajectories, this literature raises the question, in the wake of reflections in evolutionary economic geography (EEG), of the elements generating regional differentiation. The literature shares the diagnosis of an entanglement between available knowledge (Grillitsch & Hansen, 2019; Tanner, 2016), relations between actors and agency (Grillitsch et al., 2023; Hansmeier & Losacker, 2024; Trippl et al., 2020) and institutional structure (Grillitsch et al., 2023; Martin & Coenen, 2014), the configuration of which is path-dependent, thus dependent on the pre-conditions of the regional economy. EEG has focused in particular on the role of knowledge in the emergence of environmental innovation trajectories. Consistent with authors who have sought to establish typologies of regional trajectories (see above), the literature emphasizes the path-dependent, endogenous and localized nature of knowledge production (Antonelli, 2006, 2017). This means that the development of new knowledge is dependent both on the past knowledge available and accessible to the actors, and on possible recombinations.

The literature points out that a broad knowledge base is positively correlated with innovation intensity (Apa et al., 2018). The high availability of knowledge, supported by diversified and innovative players, would favor the ability to produce innovations. Nevertheless, this enhanced capacity to innovate would tend to lead to the development of incremental rather than radical innovations. Similarly, this effect is reinforced by the positive role of technological proximity on collaboration between players (Calignano et al., 2024). Hence, a broad knowledge base could lead to a trajectory of specialization in one technological field to the detriment of radical innovations on another trajectory. The strengthening of this trajectory is reinforced by the key role played by technological relatedness in the development of innovation trajectories at the regional level (Boschma, 2017; Boschma & Frenken, 2011). This notion refers to the idea that activities develop through the existence of complementarities (Boschma, 2023; Breschi et al., 2003; Teece et al., 1997) in the vicinity of their technological trajectory (Ayoub & Lhuillery, 2024). The process of knowledge production for the development of environmental innovations thus corresponds to a process of re-exploration of existing knowledge stocks to select the relevant knowledge necessary to produce environmental innovations. Nevertheless, the possibilities to reinterpret the knowledge produced are limited by the skills that actors possess but also by the field of possibilities opened by institutions, thus joining the perspectives developed in qualitative approaches to regional innovation systems (Asheim & Coenen, 2005; Isaksen et al., 2022).

In parallel with the relatedness property, the literature has focused on the issue of diversification. Indeed, the existence of a regional trajectory leads actors to reinforce the development of incremental innovations. Some authors put forward the idea that relatedness could make it possible to diversify regional knowledge bases, promoting their cross-fertilization, considering that the possible overlap between available knowledge would foster possible recombinations and the differentiation of one region from others (Balland et al., 2019; Boschma, 2023). In the context of smart specialization strategies, regions with an already diversified knowledge base would benefit from a deepening of their knowledge bases (Rigby et al., 2022). Conversely, as Boschma (2023) points out, peripheral regions tend to benefit from a strategy favoring activities with high relatedness and low complexity. If the associated benefits are relatively low, these peripheral regions could benefit more strongly from comprehensive diversification strategies by investing in knowledge that is uniquely complex, i.e. radically new to the territory. Nevertheless, Apa et al. (2018) point out that while the existence of a diversified knowledge base does generate innovation trajectories, this does not mean that these innovations will be of good quality. It is therefore necessary to take into account not only the heterogeneity of the knowledge available at the scale of a region, but also the level of maturity and complementarity of the knowledge available, enabling technologies to be recombined (Kalthaus, 2020; Turkina et al., 2021). At the scale of a regional innovation system, actors have a common knowledge base consisting of a set of generic and specific knowledge (Giuliani, 2007). This specific knowledge constitutes the source of differentiation between actors. The production of new knowledge (produced within the RIS or outside) and its reinterpretation generates innovations whose diffusion transforms the RIS (Angelakis & Galanakis, 2017; Antonelli, 2017; Frigon et al., 2020). From the point of view of innovation policies, regions must therefore choose between strategies of cross-fertilization between technologies and strategies of investment in complex/related or complex-related knowledge bases, and in generic or specific knowledge.

2.3 The bioeconomy from a regional perspective

Much of the theoretical background described above also applies to the regional transition towards bioeconomies and to changes in the regional knowledge bases that support this transition. Since the adoption of the bioeconomy concept at the European level, the local dimension has been emphasized (European Commission, 2018, 2012), particularly as a means of fostering rural development. This has been one of the arguments presented by agro-industries and forestry actors to promote the bioeconomy. The study of these strategies has primarily focused on the perceptions of actors to highlight shared visions of the bioeconomy or on the technological fields developed, revealing a diversity of bioeconomy forms and difficulties in understanding the modes of bioeconomy development (Andersson & Grundel, 2021; Siekmann & Venghaus, 2024). In this vein, other works have sought to highlight the importance of stakeholder networks, the availability of a knowledge base, a suitable institutional framework, and the multiplication of "helixes" (developing quadruple and quintuple helix approaches) to foster the development of the bioeconomy on a regional scale (Grossauer & Stoeglehner, 2020; Halonen et al., 2022). Despite their interest, this line of research has struggled to provide a generalizable approach for capturing regional trajectories of bioeconomy development.

To address this challenge of understanding the local level, work around the notion of clusters and its extension in terms of bio-clusters (Ayrapetyan et al., 2022; Hermans, 2018) has highlighted the role of alliances between regional economic and political players in structuring bioeconomy clusters. These clusters help mobilize political resources, knowledge, and funding for bioeconomy infrastructures (Befort, 2023; Birch, 2008; Hellsmark et al., 2016; Martin et al., 2023). These studies share a common interest in analyzing the flow of knowledge from the local to the global level for the production of innovation, but have mainly focused on regions marked by the existence of a bioeconomy cluster centered on a biorefinery, adopting a research strategy aligned with the broader field of the geography of sustainability transitions (Hansmeier & Kroll, 2024). However, these biorefinery clusters have been structured mainly by the gradual addition of complementary activities around a paper mill or distillery, exploiting only a limited share of the bioeconomy's available technological variety (Bauer et al., 2018; Befort, 2020; D’Amato et al., 2020; Hellsmark et al., 2016). Additionally, these biorefineries tend to develop incremental innovations because they belong to industries (food and paper) that are more technology-intensive and have a knowledge base with a high degree of relatedness and low complexity (Befort, 2023). This regional specialization effect of the bioeconomy is confirmed by d'Adamo et al. (2022), who highlight inter-regional differences between northern and southern Italy, based on the heterogeneous pre-eminence of bioeconomy sectors in the regions. Regional analysis of the bioeconomy should therefore be based on an analysis of the composition of the regional bioeconomy knowledge base, aiming to identify the dynamics of trajectories.

3 Data and Methods

3.1 Data

For the empirical part of this paper, we merge information from two original patent datasets. As is standard in innovation research, we use patents as a proxy to measure knowledge development and innovation activity (Archibugi, 1992; Griliches & Zvi, 1990). The first dataset comprises patents pertaining to the bioeconomy, while the second dataset offers highly granular and detailed geographical information on patent inventors. By merging these datasets, we aim to gain a comprehensive understanding of the geography of innovation in the bioeconomy.

The bioeconomy patent dataset was recently published by Kriesch & Losacker (2024a). These authors employed a sophisticated natural language processing approach to identify patents related to the bioeconomy from patent abstracts. This method overcomes the limitations of rule-based approaches used for identifying bioeconomy patents, such as simple keyword searches or combinations of technology codes, which cannot fully capture the bioeconomy within existing classification systems. The dataset by Kriesch & Losacker (2024a) promises to provide a more complete picture of bioeconomy-related inventions, and we are among the first to make use of it. In a prior endeavor, using text data from firm websites, the same authors demonstrated the applicability of natural language processing approaches for measuring bioeconomy-related activities (Kriesch & Losacker, 2024b). The dataset by Kriesch & Losacker (2024a), derived from PATSTAT 2022, offers information on whether each unique patent (identified by the application ID) is linked to the bioeconomy. This allows for simple matching with data from PATSTAT or other patent databases.

The second dataset utilized in this paper furnishes geographical details about patent inventors. Generally, the geographical analysis of patenting activity is subject to various biases when relying on applicant information for geolocation, such as the influence of headquarter locations. To mitigate these issues, it is common practice to utilize inventor addresses, which offer a more accurate representation of where an invention was developed. However, the inventor information available in PATSTAT is often incomplete and may not accurately reflect the true geographic origin of patent developments. To address these challenges, we utilize the geocoded patent dataset published by De Rassenfosse et al. (2019). This dataset enhances PATSTAT's address data with additional information from major patent offices, allowing for the precise determination of geographic coordinates for the majority of inventor addresses. We match the bioeconomy patents to this dataset using the unique application ID of each patent. At this stage, our dataset encompasses a total of geocoded 16,534,658 patents. Among these, 1,093,281 (6.61%) patents relate to the bioeconomy. In a next step, we utilize the inventor locations provided by De Rassenfosse et al. (2019) to match each patent to one or several functional urban areas (FUAs) in OECD countries as defined by Dijkstra et al. (2019). FUAs encompass not only city centers but also their economically integrated surrounding areas, as evidenced by commuting patterns. Utilizing FUAs in geographical patent analysis offers several advantages over traditional administrative boundaries, which may obscure the true locations of inventions due to factors like cross-boundary commuting by inventors. Furthermore, reliance on administrative boundaries can introduce issues often associated with the modifiable areal unit problem, such as scale and aggregation effects that can distort statistical results. FUAs are increasingly utilized in empirical research on innovation geography (Broekel et al., 2023; Fritsch & Wyrwich, 2021), making them highly suitable for the analysis presented in this paper. In attributing patent counts to FUAs, we adhere to a full count approach, recognizing that knowledge is an indivisible good. This allows us to straightforwardly tally the number of bioeconomy patents as well as the total patents per region and year. We compute three-year moving windows for the number of patents per region and year to smooth out fluctuations due to outliers and shocks, as well as to reflect the cumulative nature of knowledge development. Subsequently, we exclude all regions where the total number of patents is less than five in any year within our observation period, resulting in the removal of several smaller regions. The final dataset used for the empirical analysis in this paper consists of a balanced panel of 617 functional urban areas from 27 OECD countries, covering the period from 1982 to 2014.

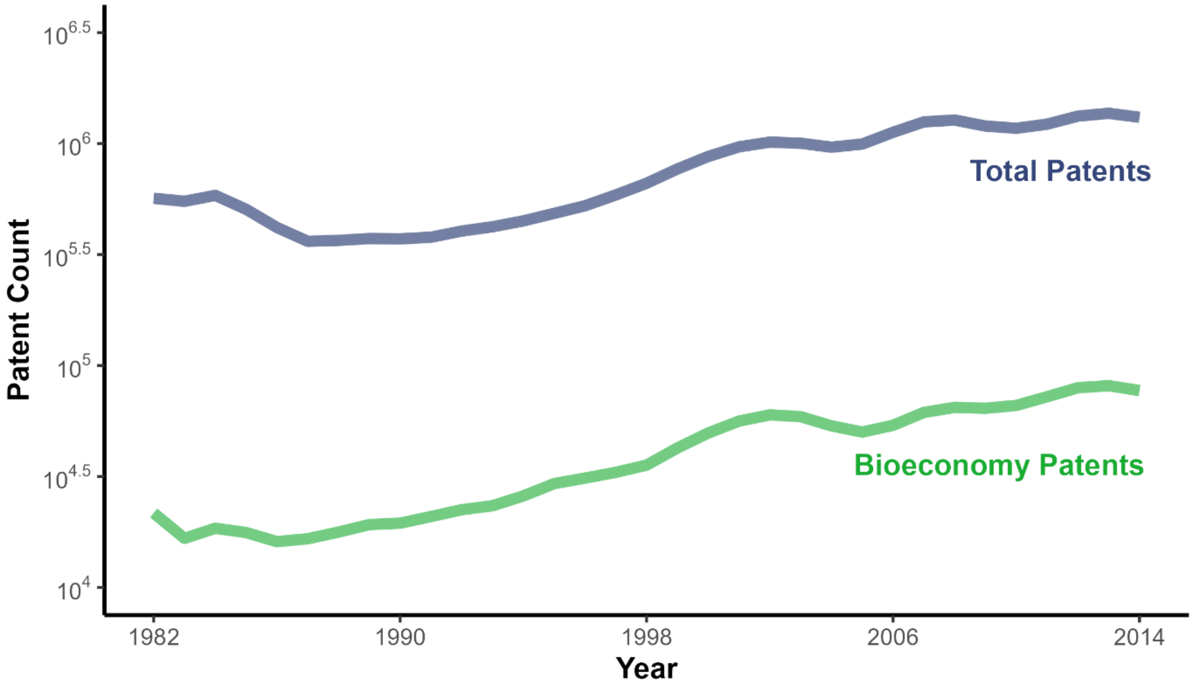


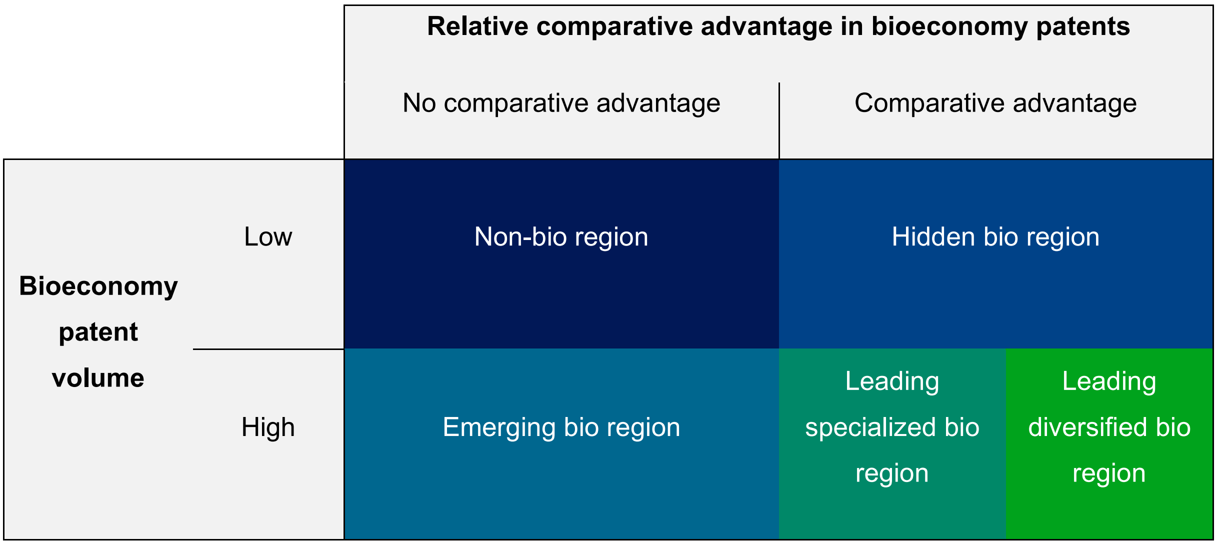
Figure 1: Annual number of patents (bioeconomy and total) on log scale, global sum, 1982-2014

For the final dataset, figure 1 visualizes the development of the number of bioeconomy patents as well as the total patents on a logarithmic scale with the 3-year moving window counts. It is evident that there has been an upward trend in both the overall patent count and the number of bioeconomy patents, albeit with minor fluctuations in the proportion of bioeconomy patents over time. Moreover, there is considerable heterogeneity in the prevalence of bioeconomy patents across different countries (and also across different regions), as further illustrated by Figure A1 and A2 in the appendix.

3.2 Methods

Utilizing the aforementioned unique dataset, our analysis focuses on the evolution of regional bioeconomy patent portfolios over time. Our primary objective is to delineate the regional pathways towards a bioeconomy, using data from the most recent year in the dataset (2014) as a reference point. Additionally, we aim to conduct comparative analyses between regions, facilitating an understanding of each region's bioeconomy development trajectory over time.

To address these empirical goals, we propose a methodology based on three main empirical steps. In the first step, we create a typology that facilitates the comparison of the progression of regional bioeconomies over time by categorizing each region for each year into a distinct state. This typology will help us assess the progress of the bioeconomy. In the second step, we use sequence analysis techniques to understand how the state of each region has changed over time, drawing region-specific evolutionary trajectories. We then compare the similarity of regional sequences based on cluster analysis, providing us with typical patterns of how regional bioeconomies change. In the final step, we uncover the mechanisms behind these changes through a shift-share analysis for each cluster. This final step allows for a better understanding of the regional trajectories identified.



		Relative comparative advantage in bioeconomy patents		
		No comparative advantage	Comparative advantage	
Bioeconomy patent volume	Low	Non-bio region	Hidden bio region	
	High	Emerging bio region	Leading specialized bio region	Leading diversified bio region

Figure 2: Classification of bioeconomy regions

Classification of bioeconomy regions

Figure 2 presents our typology for assessing the advancement of the bioeconomy across regions. It encompasses three primary dimensions. (1) The first dimension focuses on the patent volume within a region, comparing the count of bioeconomy patents in region r during year t against the median count of bioeconomy patents across all regions in the benchmark year of 2014. (2) The second dimension pertains to the relative comparative advantage RCA of a region in the bioeconomy domain. We calculate the RCA to assess the relative strength of region r in year t in the bioeconomy, using the proportion of bioeconomy patents among all patents in 2014 as a benchmark. This enables us to gauge the advancement of the bioeconomy within a region relative to the global state in the most recent year observed in our study. An RCA greater than 1 signifies that a region possesses a higher share of bioeconomy patents relative to the benchmark, indicating a comparative advantage. Conversely, an RCA lower than 1 suggests that the region's share of bioeconomy patents is below the benchmark, reflecting a comparative disadvantage. The RCA can be denoted as follows:

$$RCA_{r,t} = \frac{\left(\frac{BE\ Patents_{r,t}}{Total\ Patents_{r,t}}\right)}{\left(\frac{\sum_r BE\ Patents_{r,2014}}{\sum_r Total\ Patents_{r,2014}}\right)}$$

Based on the two dimensions of bioeconomy patent volume and relative comparative advantage in bioeconomy patents, we categorize regions into four distinct states that signify their development in the bioeconomy for each year covered in our dataset:

- **Non-bio regions**: These are regions whose bioeconomy patent volume falls below the global median of 2014 and concurrently possess an RCA value less than 1. This classification suggests that such regions have not specialized in the bioeconomy nor have they demonstrated significant activity in bioeconomy innovation.
- **Hidden bio regions**: These regions are characterized by an RCA greater than 1, indicating specialization in bioeconomy patents, yet their bioeconomy patent volume remains low. This implies that despite their smaller scale, they have a focused and potentially impactful bioeconomy knowledge base.
- **Emerging bio regions**: Regions are designated as emerging bio regions if they have a bioeconomy patent volume exceeding the global median of 2014, but an RCA lower than 1. This indicates that while these regions are active in bioeconomy patenting, they are not as specialized in it relative to the global benchmark.

The fourth category, denoted as **leading bio regions**, comprises those regions that exhibit a comparative advantage in bioeconomy patents, with an RCA greater than 1, and concurrently possess a high volume of bioeconomy patents. (3) At this stage, we introduce a third dimension to further differentiate between two distinct types of leading bio regions. That is, we study the concentration of bioeconomy patents across technological domains, as classified by CPC codes, within a region. This allows us to further differentiate between two types of leading bio regions:

- **Leading specialized bio regions**: These are regions with both a high volume of bioeconomy patents and an RCA exceeding 1, indicating a comparative advantage in bioeconomy patents. Additionally, these regions display a concentration of patents in a few CPC classes, suggesting a focused expertise in specific areas of the bioeconomy.
- **Leading diversified bio regions**: Similar to specialized bio regions, diversified bio regions also have a high volume of bioeconomy patents and an RCA greater than 1. However, unlike their specialized counterparts, diversified bio regions have a broad distribution of bioeconomy patents across a wide array of CPC classes, indicating a varied and extensive patent portfolio within the bioeconomy.

For this dimension, we employ the Shannon Index H as a measure of the diversity within a regional portfolio of bioeconomy patents. This index assesses the distribution of bioeconomy patents across different CPC classes (4-digit), capturing both the variety of patent classes represented (richness) and the relative number of patents in each class (evenness). For a specific region *r* and year *t*, *H* is calculated as follows:

$$H_{r,t} = -\sum\nolimits_{j} \left(p_{r,t,j} \cdot ln\left(p_{r,t,j}\right)\right)$$

where $p_{r,t,j}$ is the proportion of bioeconomy patents in patent class *j* out of the total bioeconomy patents in region *r* during year *t*. A high value for $H_{r,t}$ suggests that the region's bioeconomy portfolio is more diverse, meaning that patents are more evenly spread across different CPC classes. This implies that the bioeconomy portfolio in that region is not heavily concentrated in a few areas but is instead spread out across a wider range of technological domains. A lower Shannon Index suggests that the region's bioeconomy portfolio is less diverse, with a higher concentration of patents in fewer CPC classes. This implies specialization in a few areas within the bioeconomy. A Shannon Index of 0 would indicate no diversity, meaning all bioeconomy patents in the region for that year fall into a single CPC class. We categorize leading bio regions into specialized bio regions and diversified bio regions based on the median Shannon Index in 2014 as a threshold. In summary, our classification assigns each region (FUA), for every year within our panel dataset, to one of the five distinct states.

Sequence analysis

The approach described above allows us to monitor the evolution of the bioeconomy in each region over time, resulting in distinct regional trajectories. For this purpose, we employ sequence analysis techniques, a collection of methods suited for analyzing and exploiting this type of data. Sequence analysis handles multivariate categorical time series data and is extensively applied across various research domains within the social sciences (Abbott, 1995; Liao et al., 2022). Despite its widespread use, it has only recently been adopted in geographical studies focusing on regional innovation activities (Hansmeier & Losacker, 2024; Kuebart, 2022; Losacker & Kuebart, 2024).

Sequences are defined as the orderly arrangement of states for each observational unit over time. In a more formal context, a state sequence of length *l* is a sequentially ordered list comprising *l* elements (states) selected from a set *A* (referred to as the alphabet). Thus, a sequence *x* can be represented as an array of states $x = (x_1, x_2, x_3, \ldots, x_l)$, where each x_j is an element of *A*. In our study, sequences are operationalized as regional trajectories (see Losacker & Kuebart, 2024). Sequences are thus the chronological arrangement of states from our typology depicted in Figure 3 (our alphabet) over time. In other words, we classify each region for each year and assign one of the five types, thus obtaining sequences depicting the change in regional bioeconomy patent portfolios. The change in type between two years then indicates a change in the regional bioeconomy portfolio. However, we limit these changes to situations where a type is maintained for at least two consecutive years to smooth out short-term variations resulting from proximity to the defined thresholds. Figure 3 illustrates the distinct bioeconomy trajectories of three German regions: Wetzlar (DE079), Marburg (DE053), and Frankfurt am Main (DE005). The figure is intended to depict what sequences look like, showcasing the evolution of bioeconomy patent activities in the three regions.

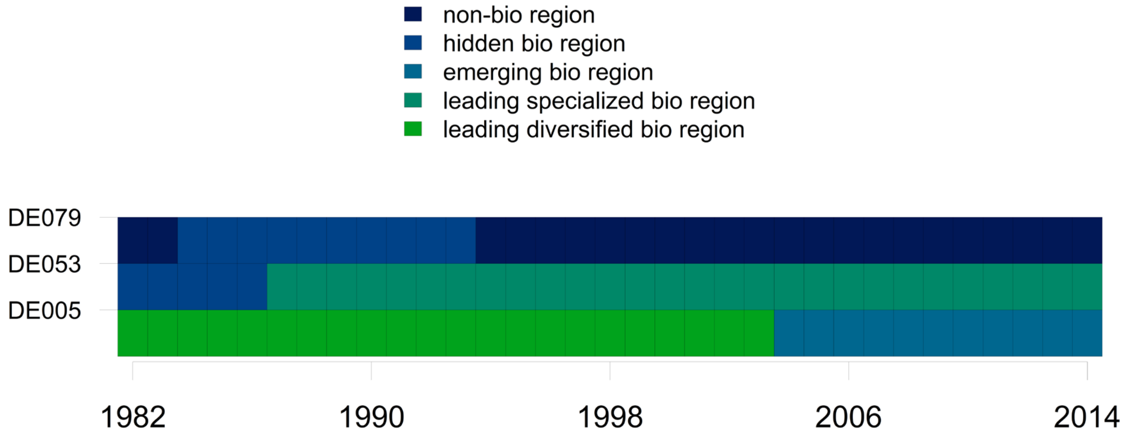


Figure 3: Sequence index plots for Wetzlar (DE079), Marburg (DE053) and Frankfurt a. M. (DE005)

Wetzlar (DE079), initially a non-bio region in the early 1980s, established a comparative advantage in bioeconomy patents until the mid-1990s (hidden bio region), eventually losing this advantage again and stabilizing as a non-bio region. Marburg (DE053) began as a hidden bio region, characterized by a comparative advantage in bioeconomy patents despite a low overall patent volume. Over time, Marburg not only retained its comparative advantage but also increased its patent volume, establishing itself as a leading bio region. Today, it is home to numerous research institutions and companies specializing in biotechnologies, making it a frontrunner in this sector both in Germany and globally. Frankfurt am Main (DE005) has consistently maintained a high volume of bioeconomy patents throughout the observed period, a factor largely attributed to the region's size and thick regional innovation system. Despite its overall patent strength in diverse bioeconomy related fields, Frankfurt has also displayed a comparative advantage in bioeconomy patents for most of the time, but lost it in the early 2000s and can now be considered an emerging bio region with a high patent volume but without a comparative advantage.

Sequence analysis thus enables a detailed examination of individual regional trajectories but, more crucially, it also facilitates the comparison of all 617 regional sequences with one another. To achieve this comprehensive comparison, we utilize an optimal matching (OM) algorithm that assesses the similarities and differences across all 617 regional sequences. An OM algorithm, often used in sequence analysis, is a method for quantifying the similarity between sequences. It calculates the distance between sequences by determining the minimum number of operations required to transform one sequence into another. These operations include insertion, deletion, and substitution of elements within the sequences. The algorithm assigns costs to each operation (insertion, deletion, substitution), and the total cost of the optimal set of operations needed to match two sequences constitutes the 'distance' between them. This distance measure can then be used to cluster sequences into groups with similar patterns, identify outliers, or analyze the overall structure and dynamics within the dataset (Abbott & Tsay, 2000). For the empirical analysis presented in the next section, we have designated the cost for insertion and deletion operations as two for our OM algorithm. Additionally, we employ a substitution matrix, whereby all substitutions (altering one sequence element to another) are attributed an identical cost of one. However, substitutions of non-bio regions with leading bio regions (and vice versa), which are arguably more dissimilar, are assigned a higher substitution cost of two. For the three sequences shown above, it is thus evident that the optimal matching distance between Marburg and Frankfurt is smaller than the distance between Wetzlar and Frankfurt. This suggests a closer similarity in the bioeconomy trajectories of Marburg and Frankfurt compared to the similarity in the trajectories of Wetzlar and Frankfurt.

The findings discussed in the subsequent section demonstrate stability with respect to variations in these parameters. Furthermore, alternative algorithms were explored in place of optimal matching, yielding results that closely align with those obtained through the OM algorithm. The distance matrix generated by our OM algorithm serves as the foundation for a hierarchical cluster analysis (specifically, Ward clustering). This method effectively groups regions with similar developmental trajectories into clusters, ensuring high within-cluster similarity and pronounced distinctions between clusters, thereby highlighting between-cluster heterogeneity.

Shift-share analysis

To better understand the factors driving changes (growth) in regional bioeconomy portfolios, in the last step, we will delve deeper into the differences between the clusters through a shift-share analysis. Shift-share analysis is a simple analytical tool used primarily in the fields of regional science and economic geography to examine how specific sectors within a region contribute to economic changes over time relative to a larger reference economy (like a global standard) (Dunn, 1960). It helps to identify whether the growth or decline of a region is due to global trends, the sectoral mix of a region, or specific regional factors. Shift-share analysis breaks regional growth intro three components: (1) Global share: This component measures what the region's growth would have been if it had grown at the same rate as the global economy. It is calculated by applying the global growth rate to the region's initial output levels. (2) Technology Mix: This factor assesses how the region's sectoral structure contributes to its performance. If a region is concentrated in fast-growing sectors, this effect will be positive. It is calculated by looking at the difference between the global growth rate of each sector and the overall global growth rate, then applying this difference to the region's initial output in those sectors. (3) Regional share: This component analyzes the region's performance relative to the global trend for each sector. It highlights the region's specific advantages or disadvantages by considering how much the actual growth rate of each sector in the region differs from its expected global growth rate.

We perform the shift-share analysis on our empirical data as follows: we analyze the growth of regional output of bioeconomy patents from the start of the period to the end. Specifically, we calculate the total number of patents for the first five years starting in 1982 and for the last five years ending in 2014. We then compute the growth rate between these periods for each region, which indicates whether and to what extent the bioeconomy patent output in a region has changed over time. We then calculate the difference between these periods for each region, which indicates whether and to what extent the bioeconomy patent output in a region has changed over time. This difference is referred to as the total shift TS_r. More formally, the total shift is represented by

$$TS_r = X_r^{t_2} - X_r^{t_1}$$

where X refers to the number of patents in region r either in period one, t_1, or in period two, t_2. As mentioned above, we can break down the total shift into three components: the global share GS_r, the technology mix TM_r, and the regional share RS_r. The sum of these components equals the total shift.

$$TS_r = GS_r + TM_r + RS_r$$

We calculate the global share GS_r based on the global growth rate, denoted as γ_g. This rate is used to estimate how the number of bioeconomy patents in a region would have changed if it had grown at the growth rate observed globally. By comparing the total shift TS_r and the global share GS_r, we can determine whether a region has performed better or worse than expected given the global growth.

$$GS_r = X_r^{t_1} * \gamma_g$$

Next, we assess the technology mix effect TM_r, which helps us understand the extent to which the sectoral composition of a regional bioeconomy portfolio has contributed to regional growth. In our case, we examine the distribution of regional bioeconomy patents across the 4-digit CPC classes. We use the global growth rate γ_{gj} for each technology class j to estimate how the patent output of region r in technology class j would have changed if it had matched the global growth rate.

$$TM_{r,j} = X_{r,j}^{t_1} * (\gamma_{g,j} - \gamma_g)$$

Aggregating the technology mix effect from the region-sector level, denoted as TM_{rj}, to the regional level TM_r, provides the overall impact of the sectoral composition on regional growth. Lastly, we can calculate the regional share, which is considered a residual effect. It captures all region-specific factors contributing to growth, net of the global share effect and the technology mix effect. This component of the analysis isolates the unique contributions of the region to its own growth, providing insights into how local conditions, policies, or capabilities have enabled the region to outperform or underperform relative to global trends and sectoral compositions.

$$RS_r = TS_r - TM_r - GS_r$$

The results of the shift-share analysis will aid in comprehending the mechanisms behind changes in regional bioeconomy portfolios among the clusters identified in the preceding step.

4 Results

Figure 4 displays sequence index plots for all 617 regions. In the index plots, each row corresponds to a regional sequence. Thus, the complete index plot offers a comprehensive overview of how all regions have evolved over time. The left panel presents an unsorted index plot, while the middle and right panels exhibit index plots sorted by the first state and by the last state, respectively. This sorting facilitates the visual analysis of the sequence data. When comparing the index plots, we observe an overarching trend: regional patent portfolios have progressively shifted towards a more biobased focus over time. This trend is discernible through the increasing number of regions classified as leading bio regions, accompanied by a corresponding decrease in the number of regions categorized as non-bio regions. Furthermore, a noteworthy decline is apparent in regions classified as hidden bio regions, denoting regions with a comparative advantage in biobased technologies but exhibiting a low patent volume. This decline is understandable in light of the general upward trend in biobased patents over time, as illustrated in Figure 1. This trend also contributes to the rise in the number of regions classified as emerging bio regions, reflecting the increased biobased patent output observed in many regions. Most intriguingly, the sudden emergence of regions classified as leading specialized bio regions stands out. At the onset of the observation period, only a handful of regions fell into this category, yet there was a steady increase in the prevalence of this regional type throughout the 2000s. While many leading bio regions boast a diversified patent portfolio spanning multiple technological fields, there are also leading regions where biobased patents are concentrated within a narrow technological domain. In Figure A3 in the appendix, we also present a clearer depiction of the changes in the frequency of each state over time. In the same Figure, we also present the most frequent sequences, which are sequences that are similar across multiple regions. This provides another interesting insight, revealing a noteworthy degree of stability with a considerable number of regions maintaining the same state throughout the entire period. This stability is particularly pronounced for the leading diversified regions and, to a lesser extent, for the non-bio regions. Upon comparing this observation with Figure 4, we notice that leading regions are highly likely to retain their state over time. While some non-bio regions appear to remain in this state, many successfully transition towards more bio-oriented states. Indeed, this descriptive visual assessment of changes in regional patent portfolios offered by sequence analysis provides substantial insight into the advancement of the bioeconomy.

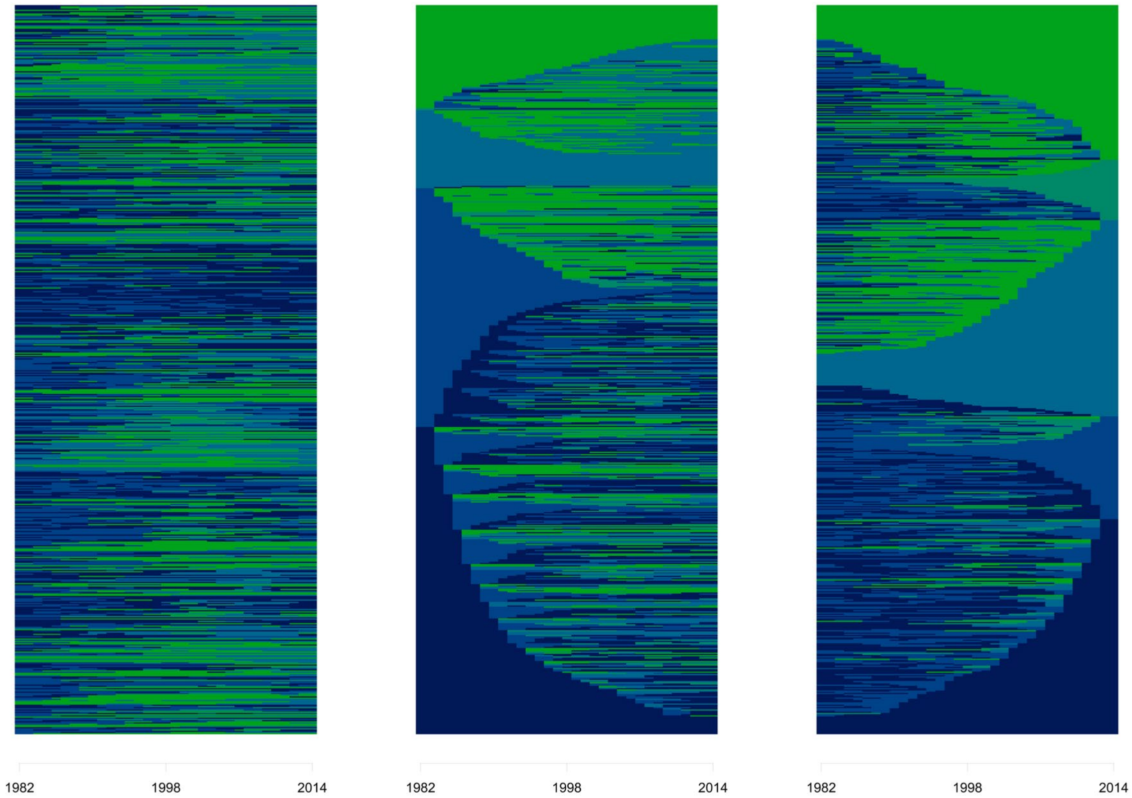


Figure 4: Sequence index plots: unsorted (left), sorted by first state (middle), sorted by last state (right)

As evidenced by the full sequence index plots, regional sequences exhibit considerable heterogeneity, with regions following diverse pathways towards the bioeconomy. While these pathways are to some extent unique—indeed, approximately 85% of all sequences are distinct—there are also discernible general patterns that apply to several regions. For instance, transitioning from hidden bio regions to leading bio regions over time is a trend observed across multiple regions. To systematically explore these patterns, we computed a distance matrix that captures the dissimilarities between each sequence and all other sequences using optimal matching algorithms (see Section 3) and conducted a hierarchical Ward cluster analysis. This method facilitated the identification of six clusters characterized by similar patterns within each cluster, yet distinct differences between clusters. The decision to adopt a six-cluster solution was informed by considerations such as cluster size, interpretability, and was supported by clustering metrics such as the within-cluster sum of squares and the elbow criterion. The results are presented in Figure 5, depicting unsorted index plots by cluster. Additionally, sorted index plots are provided (refer to Figure A4 and Figure A5 in the appendix), along with state frequency plots (refer to Figure A6 in the appendix).

Cluster 1: This cluster predominantly features a consistent pattern of specialized bio-regions over time. These regions have maintained a strong specialization in bioeconomy activities, marking them as long-term leading bio-regions. Notably, 98 out of 617 functional urban areas belong to this cluster.

Cluster 2: This cluster includes regions whose sequences are characterized by a transition from a non-bio region state to an emerging bio-region state. Although these regions do not consistently exhibit a comparative advantage throughout the period, they have managed to increase the number of bioeconomy patents. Some regions in the cluster have also become leading regions. The cluster consists of 79 regions.

Cluster 3: Dominated by non-bio region sequences, this cluster exhibits minimal change over time, with the majority of regions showing a stable absence of significant bioeconomy activity throughout the observed period. Some regions manage to establish a comparative advantage in biobased patents despite having low total outputs (hidden bio regions). It is the largest cluster, comprising 199 regions.

Cluster 4: In this cluster, we observe a pattern similar to that in Cluster 2, characterized by a transition towards increased bioeconomy activities in terms of total patent output. Primarily, this cluster includes regions that were initially considered hidden bio regions and then transitioned to becoming leading bio regions. Consequently, they have sustained their competitive advantage while simultaneously increasing the overall output of biobased patents. The cluster features 67 regions and is the smallest cluster identified.

Cluster 5: Regions in this cluster are quite distinct and their sequences are markedly different from sequences observed in other clusters. Primarily, this cluster includes regions categorized as hidden bio regions for most of the time. However, some manage to transition into specialized leading regions during the observed period, particularly in the 2000s. This is quite peculiar, given that most of the leading regions found in other clusters, and overall, typically have a diversified patent portfolio rather than a specialized one. Consequently, regions in this cluster tend to have a high concentration of bioeconomy patents in one or a few technological domains. The cluster includes 89 regions.

Cluster 6: While many clusters, particularly Cluster 2 and 4, feature success stories of regions transitioning towards a stronger bioeconomy knowledge base, the regions in Cluster 6 exhibit a different pattern. Initially categorized as leading regions at the beginning of the observation period, these regions later lose their comparative advantage and fall into the state of emerging bio regions. Despite still producing a high volume of biobased patents, the output is less than what would be expected given the global ratio.

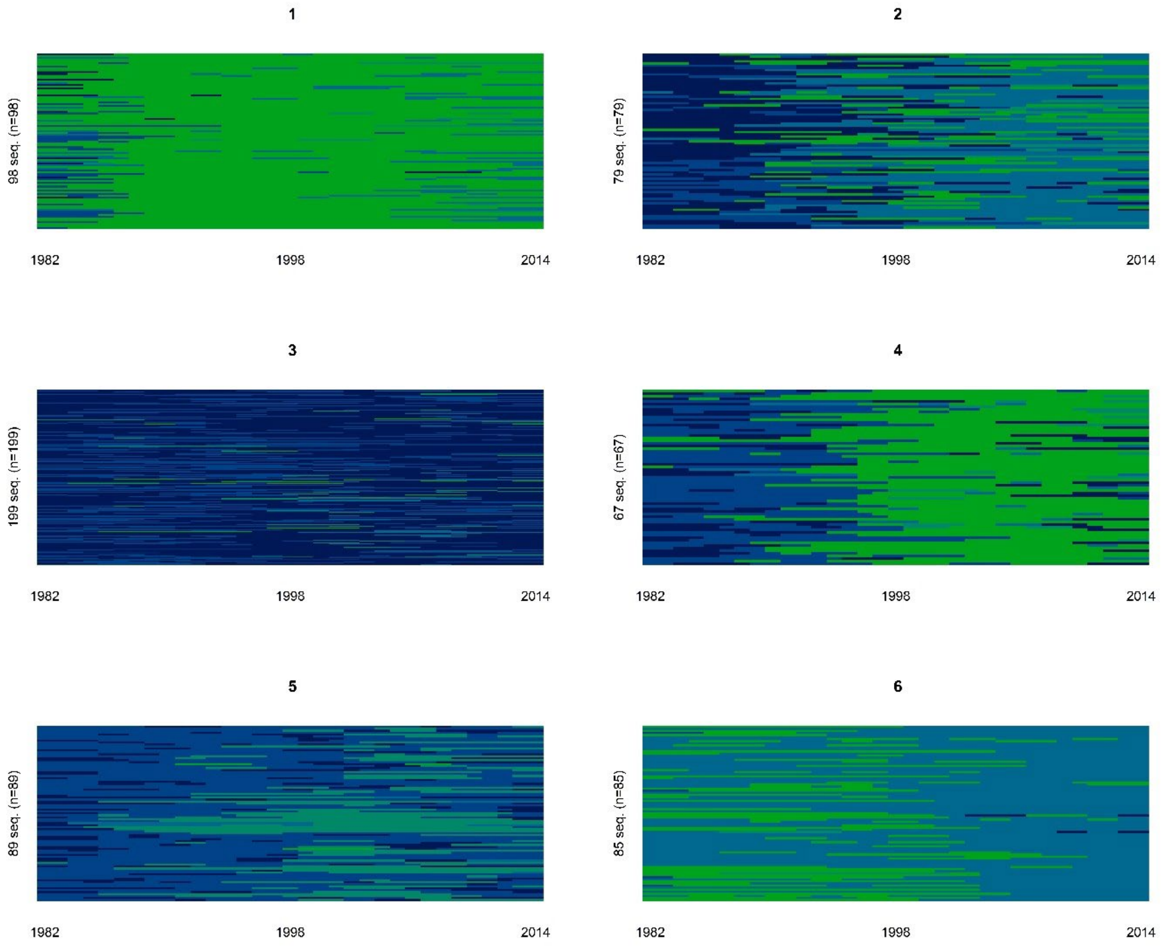


Figure 5: Sequence index plots by cluster, unsorted

The use of functional urban areas limits our empirical analysis because we cannot include additional structural variables as covariates in econometric analyses due to the lack of available secondary data at the FUA level. However, we utilized population data to examine differences between the clusters using an ANOVA (see also Figure A7in the appendix). The results indicate a significant effect of the cluster on log population sizes (F (5, 609) = 81.99, $p < 0.01$). Given the significance of the overall model, a Tukey HSD post-hoc test was employed to investigate pairwise differences between clusters. The post-hoc analysis revealed significant differences ($p < 0.01$) in population sizes among most clusters, with the exception of differences between clusters 4 and 2, as well as clusters 5 and 3, where the pairwise comparisons did not show significant variation.

Lastly, the results of the shift-share analysis help us make sense of the sequencing patterns described above. The sequence analysis has already revealed much about how regions have changed over time, i.e., from a non-bio region to a leading one, and how a region compares to the global average, but we know little about the underlying effects of these changes. Hence, we examine the technology mix effects as well as the regional share effects for each region, analyzed by cluster, in order to provide an overall picture. In Figure 6, we show the relative effects segmented by cluster.

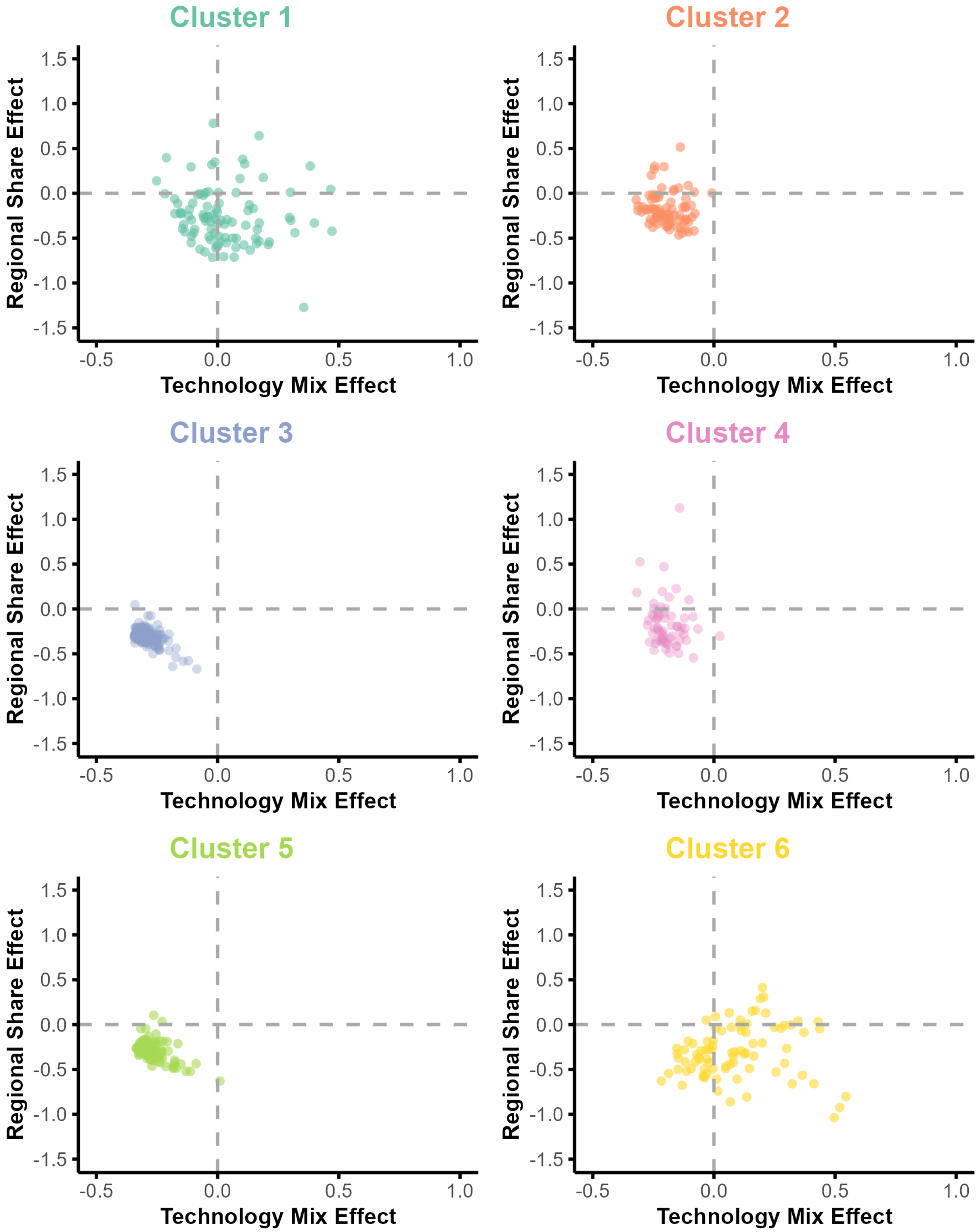


Figure 6: Shift-Share effects by cluster

For Cluster 1, we do not really see a coherent pattern. Some regions display a positive regional share effect, while others show a negative one, and the same holds true for the technology mix effect. Overall, the regions in Cluster 1 have sustained their leading state throughout the entire period, as explained earlier. However, it appears that some—namely those with negative regional share and/or technology mix effects—have performed worse than the global trend. This reflects that these regions might also lose their leading position in the future. In Cluster 2, we see a clearer picture. In this cluster, all regions exhibit a negative technology mix effect, indicating that their bioeconomy portfolio is not optimal and includes many slow-growing technology fields. While regions in this cluster have indeed increased their bioeconomy patent output, they could have achieved even higher outputs if they had a different sectoral composition. The same essentially holds true also for regions in Cluster 4. In Clusters 3 and 5, with very few exceptions, all regions exhibit both a negative technology mix effect and a negative regional share effect. This indicates that regions in these clusters have a technology mix centered on slow-growing (or declining) sectors. This is particularly striking for the regions in Cluster 5 that have transitioned to a leading specialized bio region state. Although these regions are specialized and leading, they seem to concentrate on sectors that are not growing at a high rate globally. These regions are at high risk of becoming locked-in, potentially reverting to a non-bio region state. For Cluster 6, we also observe a quite mixed pattern. However, this is interesting to note because Cluster 6 includes many regions that have lost their comparative advantage over time while still maintaining a high overall output of bioeconomy patents. It appears that some regions have managed to maintain this high output because they are home to fast-growing sectors, as indicated by positive technology mix effects. Meanwhile, others also exhibit a positive regional share effect, suggesting that they are performing better due to unique locational factors.

Overall, we find that differences in the technology mix effect between clusters are more pronounced than regional share effects. To support these descriptive results, we conducted an ANOVA to compare the technology mix effects across the different clusters. The results indicate a significant difference in the technology mix effects between clusters ($F(5, 611) = 241.3$, $p < 0.01$). Given the significance of the overall model, a Tukey HSD post-hoc test was employed to investigate pairwise differences between clusters. The post-hoc analysis revealed significant differences ($p < 0.01$) in the technology mix effects among most clusters, except for clusters 2 and 4, as well as clusters 3 and 5, where the pairwise comparisons did not show significant variation. For the regional share effects, we did not find significant differences between clusters ($p > 0.05$). Changes in regional bioeconomy patent output can thus be largely attributed to the sectoral composition of the regional patent portfolios, which favor either fast-growing or slow-growing technology fields. Meanwhile, unique locational factors seem to play a supporting role in explaining growth in bioeconomy patents. Moreover, we observe that for many clusters, the shift-share analysis yields strikingly similar effects across numerous regions within these clusters. This indicates that the clusters of regional trajectories identified through the sequence analysis not only appear similar in terms of how regions have changed over time, but these changes can also be explained by the very same mechanisms within the clusters.

5 Conclusion

In this paper, we aimed at enhancing the understanding of the bioeconomy transition by examining the development of bio-based technologies at the regional level. Our analysis built on the literature from economic geography and geography of innovation research, suggesting that regional transition paths are highly heterogeneous due to various region-specific factors and the regional sectoral mix. Using patent data as a proxy for knowledge development and innovation activities, we combined two unique datasets: one showing the precise geographical locations of patent inventors (de Rassenfosse et al., 2019) and the other identifying whether a patent was related to the bioeconomy (Kriesch & Losacker, 2024a). This data allowed us to examine how 617 regions (functional urban areas) across 27 OECD countries altered their patent portfolios concerning bioeconomy patents from 1982 to 2014. We created a typology of bioeconomy progress and analyzed the distinct regional trajectories using geographical sequence analysis, a novel method recently adopted by economic geographers (Losacker & Kuebart, 2024). Additionally, we applied shift-share techniques to determine the extent to which region-specific factors related to the sectoral composition of the regional patent portfolios influenced the bioeconomy transition trajectories.

Our study revealed distinct regional trajectories towards a bioeconomy based on bio-based technologies, categorized into six clusters. Cluster 1 comprised long-term leading bio-regions, maintaining a consistent specialization in bioeconomy activities. Cluster 2 included regions transitioning from non-bio to emerging bio-regions, increasing their bioeconomy patent output despite lacking a consistent comparative advantage. Cluster 3 was dominated by non-bio regions with minimal change, though some managed to gain a comparative advantage in bio-based patents. Cluster 4 showed regions evolving from hidden bio regions to leading bio regions, maintaining a competitive advantage while increasing patent output. Cluster 5 featured regions with unique sequences, often transitioning from hidden bio regions to specialized leaders, with high concentrations of bioeconomy patents in specific technological domains. Cluster 6 included regions initially leading in bioeconomy patents but later losing their comparative advantage. The shift-share analysis indicated that differences in technology mix effects between clusters were more pronounced than regional share effects. Regions with a negative technology mix effect typically had slow-growing technology portfolios, while those with a positive effect were in fast-growing sectors. The ANOVA and post-hoc tests confirmed significant differences in technology mix effects among most clusters, highlighting the impact of sectoral composition on regional bioeconomy patent output. Unique locational factors played a supporting role in this growth. Overall, the study suggests that place-based policy strategies are crucial for advancing regional bioeconomies, considering the heterogeneous nature of regional transition trajectories

For long-term leading bio-regions (Cluster 1), it is crucial to maintain their competitive edge through continuous innovation and development in bioeconomy technologies. Policies should encourage diversification within these regions to prevent potential lock-in effects. Emerging bio-regions (Clusters 2 and 4) require targeted support to optimize the sectoral composition of their patent portfolios, focusing on fast-growing bioeconomy sectors. Hidden bio-regions (Clusters 3 and 5) need policies that unlock their potential by providing resources and incentives for innovation in underrepresented bioeconomy sectors. Declining bio-regions (Cluster 6) can benefit from support aimed at re-establishing their comparative advantage by investing in emerging and fast-growing bioeconomy sectors. General strategies applicable to all regions include integrating bioeconomy strategies into broader regional development plans to ensure a holistic approach that leverages existing strengths and opportunities. These place-based policy strategies can help regions navigate their unique transition paths towards the bioeconomy.

References

Abbott, A. (1995). Sequence Analysis: New Methods for Old Ideas. *Annual Review of Sociology*, *21*(1), 93–113. https://doi.org/10.1146/annurev.so.21.080195.000521

Abbott, A., & Tsay, A. (2000). Sequence Analysis and Optimal Matching Methods in Sociology. *Sociological Methods & Research*, *29*(1), 3–33. https://doi.org/10.1177/0049124100029001001

Andersson, I., & Grundel, I. (2021). Regional policy mobilities: Shaping and reshaping bioeconomy policies in Värmland and Västerbotten, Sweden. *Geoforum*, *121*, 142–151. https://doi.org/10.1016/J.GEOFORUM.2021.02.005

Angelakis, A., & Galanakis, K. (2017). A science-based sector in the making: the formation of the biotechnology sector in two regions. *Regional Studies*, *51*(10), 1542–1552. https://doi.org/10.1080/00343404.2016.1215601

Antonelli, C. (2006). The Business Governance of Localized Knowledge: An Information Economics Approach for the Economics of Knowledge. *Industry and Innovation*, *13*(3), 227–261. https://doi.org/10.1080/13662710600858118

Antonelli, C. (2017). Endogenous innovation: the creative response. *Economics of Innovation and New Technology*, *26*(8), 689–718. https://doi.org/10.1080/10438599.2016.1257444

Apa, R., De Noni, I., Orsi, L., & Sedita, S. R. (2018). Knowledge space oddity: How to increase the intensity and relevance of the technological progress of European regions. *Research Policy*, *47*(9), 1700–1712. https://doi.org/10.1016/J.RESPOL.2018.06.002

Archibugi, D. (1992). Patenting as an indicator of technological innovation: a review. *Science and Public Policy*, *19*(6), 357–368. https://doi.org/10.1093/spp/19.6.357

Asheim, B., & Coenen, L. (2005). Knowledge bases and regional innovation systems: Comparing Nordic clusters. *Research Policy*, *34*(8), 1173–1190. https://doi.org/10.1016/J.RESPOL.2005.03.013

Ayoub, M., & Lhuillery, S. (2024). Mix and match: what is the best R&D recipe for eco-innovation? *Industry and Innovation*. https://doi.org/10.1080/13662716.2024.2359380

Ayrapetyan, D., Befort, N., & Hermans, F. (2022). The role of sustainability in the emergence and evolution of bioeconomy clusters: An application of a multiscalar framework. *Journal of Cleaner Production*, *376*, 134306. https://doi.org/10.1016/J.JCLEPRO.2022.134306

Balland, P. A., Boschma, R., Crespo, J., & Rigby, D. L. (2019). Smart specialization policy in the European Union: relatedness, knowledge complexity and regional diversification. *Regional Studies*, *53*(9), 1252–1268. https://doi.org/10.1080/00343404.2018.1437900

Bauer, F., Hansen, T., & Hellsmark, H. (2018). Innovation in the bioeconomy–dynamics of biorefinery innovation networks. *Technology Analysis and Strategic Management*, *30*(8), 935–947. https://doi.org/10.1080/09537325.2018.1425386

Befort, N. (2020). Going beyond definitions to understand tensions within the bioeconomy: The contribution of sociotechnical regimes to contested fields. *Technological Forecasting and Social Change*, *153*, 119923. https://doi.org/10.1016/J.TECHFORE.2020.119923

Befort, N. (2023). *The Bioeconomy: Institutions, Innovation and Sustainability*. Routledge. https://doi.org/10.4324/9781003103011

Birch, K. (2008). Alliance-Driven Governance: Applying a Global Commodity Chains Approach to the U.K. Biotechnology Industry. *Economic Geography*, *84*(1), 83–103. https://doi.org/10.1111/J.1944-8287.2008.TB00392.X

Boschma, R. (2017). Relatedness as driver of regional diversification: a research agenda. *Regional Studies*, *51*(3), 351–364. https://doi.org/10.1080/00343404.2016.1254767

Boschma, R. (2023). Evolutionary economic geography and policy. *Routledge Handbook of Evolutionary Economics*, 332–340. https://doi.org/10.4324/9780429398971-28/EVOLUTIONARY-ECONOMIC-GEOGRAPHY-POLICY-RON-BOSCHMA

Boschma, R., & Frenken, K. (2011). The emerging empirics of evolutionary economic geography. *Journal of Economic Geography*, *11*(2), 295–307. https://doi.org/10.1093/jeg/lbq053

Breschi, S., Lissoni, F., & Malerba, F. (2003). Knowledge-relatedness in firm technological diversification. *Research Policy*, *32*(1), 69–87. https://doi.org/10.1016/S0048-7333(02)00004-5

Broekel, T., Knuepling, L., & Mewes, L. (2023). Boosting, sorting and complexity—urban scaling of innovation around the world. *Journal of Economic Geography*. https://doi.org/10.1093/JEG/LBAD006

Calignano, G., Lee, J., & Kogler, D. F. (2024). Nanotechnology and knowledge relatedness: how to identify optimal regional partners in EU innovation networks? *Science and Public Policy*. https://doi.org/10.1093/scipol/scae032

D'Adamo, I., Falcone, P. M., Imbert, E., & Morone, P. (2022). Exploring regional transitions to the bioeconomy using a socio-economic indicator: the case of Italy. *Economia Politica*, *39*(3), 989–1021. https://doi.org/10.1007/S40888-020-00206-4/FIGURES/2

D'Amato, D., Veijonaho, S., & Toppinen, A. (2020). Towards sustainability? Forest-based circular bioeconomy business models in Finnish SMEs. *Forest Policy and Economics*, *110*, 101848. https://doi.org/10.1016/J.FORPOL.2018.12.004

de Rassenfosse, G., Kozak, J., & Seliger, F. (2019). Geocoding of worldwide patent data. *Scientific Data*, *6*(1), 260. https://doi.org/10.1038/s41597-019-0264-6

Dijkstra, L., Poelman, H., & Veneri, P. (2019). *The EU-OECD definition of a functional urban area*. https://doi.org/10.1787/d58cb34d-en

Dunn, E. S. (1960). A statistical and analytical technique for regional analysis. *Papers in Regional Science*, *6*(1), 97–112. https://doi.org/10.1111/j.1435-5597.1960.tb01705.x

European Commission. (2018). *A sustainable Bioeconomy for Europe: strengthening the connection between economy, society and the environment*. 107. https://doi.org/10.2777/478385

European Commission, D.-G. for R. and I. (2012). *Innovating for sustainable growth - a bioeconomy for Europe*. https://doi.org/https://data.europa.eu/doi/10.2777/6462

Frigon, A., Doloreux, D., & Shearmur, R. (2020). Drivers of eco-innovation and conventional innovation in the Canadian wine industry. *Journal of Cleaner Production*, *275*, 124115. https://doi.org/10.1016/J.JCLEPRO.2020.124115

Fritsch, M., & Wyrwich, M. (2021). Is innovation (increasingly) concentrated in large cities? An international comparison. *Research Policy*, *50*(6), 104237. https://doi.org/10.1016/j.respol.2021.104237

Giuliani, E. (2007). The selective nature of knowledge networks in clusters: evidence from the wine industry. *Journal of Economic Geography*, *7*(2), 139–168. https://doi.org/10.1093/JEG/LBL014

Griliches, & Zvi. (1990). Patent Statistics as Economic Indicators: A Survey. *Journal of Economic Literature*, *28*(4), 1661–1707. https://ideas.repec.org/a/aea/jeclit/v28y1990i4p1661-1707.html

Grillitsch, M., & Hansen, T. (2019). Green industry development in different types of regions. *European Planning Studies*, *27*(11), 1–21. https://doi.org/10.1080/09654313.2019.1648385

Grillitsch, M., Sotarauta, M., Asheim, B., Fitjar, R. D., Haus-Reve, S., Kolehmainen, J., Kurikka, H., Lundquist, K.-J., Martynovich, M., Monteilhet, S., Nielsen, H., Nilsson, M., Rekers, J., Sopanen, S., & Stihl, L. (2023). Agency and economic change in regions: identifying routes to new path development using qualitative comparative analysis. *Regional Studies*, *57*(8), 1453–1468. https://doi.org/10.1080/00343404.2022.2053095

Grossauer, F., & Stoeglehner, G. (2020). Bioeconomy—Spatial Requirements for Sustainable Development. *Sustainability 2020, Vol. 12, Page 1877*, *12*(5), 1877. https://doi.org/10.3390/SU12051877

Halonen, M., Näyhä, A., & Kuhmonen, I. (2022). Regional sustainability transition through forest-based bioeconomy? Development actors' perspectives on related policies, power, and justice. *Forest Policy and Economics*, *142*, 102775. https://doi.org/10.1016/J.FORPOL.2022.102775

Hansmeier, H., & Kroll, H. (2024). The geography of eco-innovations and sustainability transitions: a systematic comparison. *ZFW - Advances in Economic Geography*. https://doi.org/10.1515/ZFW-2022-0010/ASSET/GRAPHIC/J_ZFW-2022-0010_FIG_002.JPG

Hansmeier, H., & Losacker, S. (2024). Regional eco-innovation trajectories. *European Planning Studies*, 1–22. https://doi.org/10.1080/09654313.2024.2308027

Hellsmark, H., Frishammar, J., Söderholm, P., & Ylinenpää, H. (2016). The role of pilot and demonstration plants in technology development and innovation policy. *Research Policy*, *45*(9), 1743–1761. https://doi.org/10.1016/J.RESPOL.2016.05.005

Hermans, F. (2018). The potential contribution of transition theory to the analysis of bioclusters and their role in the transition to a bioeconomy. *Biofuels, Bioproducts and Biorefining*, *12*(2), 265–276. https://doi.org/10.1002/BBB.1861

Isaksen, A., Trippl, M., & Mayer, H. (2022). Regional innovation systems in an era of grand societal challenges: reorientation versus transformation. *European Planning Studies*, *30*(11), 2125–2138. https://doi.org/10.1080/09654313.2022.2084226

Kalthaus, M. (2020). Knowledge recombination along the technology life cycle. *Journal of Evolutionary Economics*, *30*(3), 643–704. https://doi.org/10.1007/S00191-020-00661-Z/FIGURES/9

Kriesch, L., & Losacker, S. (2024a). A global patent dataset for the bioeconomy. *Papers in Innovation Studies*, *2024*(08).

Kriesch, L., & Losacker, S. (2024b). Bioeconomy firms and where to find them. *REGION*, *11*(1), 55–78. https://doi.org/10.18335/region.v11i1.523

Kuebart, A. (2022). Open creative labs as functional infrastructure for entrepreneurial ecosystems: Using sequence analysis to explore tempo-spatial trajectories of startups in Berlin. *Research Policy*, *51*(9), 104444. https://doi.org/10.1016/J.RESPOL.2021.104444

Le Bas, C., & Scellato, G. (2014). Firm innovation persistence: a fresh look at the frameworks of analysis. *Economics of Innovation and New Technology*, *23*, 423–446. https://doi.org/10.1080/10438599.2014.895511

Liao, T. F., Bolano, D., Brzinsky-Fay, C., Cornwell, B., Fasang, A. E., Helske, S., Piccarreta, R., Raab, M., Ritschard, G., Struffolino, E., & Studer, M. (2022). Sequence analysis: Its past, present, and future. *Social Science Research*, *107*, 102772. https://doi.org/10.1016/J.SSRESEARCH.2022.102772

Losacker, S., & Kuebart, A. (2024). Introducing sequence analysis to economic geography. *Progress in Economic Geography*, *2*(1), 100012. https://doi.org/10.1016/J.PEG.2024.100012

Martin, H., & Coenen, L. (2014). Institutional Context and Cluster Emergence: The Biogas Industry in Southern Sweden. *European Planning Studies*, *23*(10), 2009–2027. https://doi.org/10.1080/09654313.2014.960181

Martin, H., Grundel, I., & Dahlström, M. (2023). Reconsidering actor roles in regional innovation systems: transformative industrial change in the forest-based bioeconomy. *Regional Studies*, *57*(9), 1636–1648. https://doi.org/10.1080/00343404.2022.2151581

Patermann, C., & Aguilar, A. (2021). A bioeconomy for the next decade. *EFB Bioeconomy Journal*, *1*(1), 100005. https://doi.org/10.1016/J.BIOECO.2021.100005

Proestou, M., Schulz, N., & Feindt, P. H. (2023). A global analysis of bioeconomy visions in governmental bioeconomy strategies. *Ambio 2023*, 1–13. https://doi.org/10.1007/S13280-023-01958-6

Rigby, D. L., Roesler, C., Kogler, D., Boschma, R., & Balland, P. A. (2022). Do EU regions benefit from Smart Specialisation principles? *Regional Studies*, *56*(12), 2058–2073. https://doi.org/10.1080/00343404.2022.2032628

Siekmann, F., & Venghaus, S. (2024). Regional transformation pathways for the bioeconomy: A novel monitoring approach for complex transitions. *Journal of Industrial Ecology*, *28*(3), 603–616. https://doi.org/10.1111/JIEC.13484

Tanner, A. N. (2016). The emergence of new technology-based industries: the case of fuel cells and its technological relatedness to regional knowledge bases. *Journal of Economic Geography*, *16*(3), 611–635. https://doi.org/10.1093/JEG/LBV011

Teece, D. J., Pisano, G., & Shuen, A. (1997). Dynamic capabilities and strategic management. *Strategic Management Journal*, *18*, 509–533. https://doi.org/10.1002/(SICI)1097-0266(199708)18:7

Trippl, M., Baumgartinger-Seiringer, S., Frangenheim, A., Isaksen, A., & Rypestøl, J. O. (2020). Unravelling green regional industrial path development: Regional preconditions, asset modification and agency. *Geoforum*, *111*, 189–197. https://doi.org/10.1016/j.geoforum.2020.02.016

Turkina, E., Van Assche, A., & Doloreux, D. (2021). How do firms in co-located clusters interact? Evidence from Greater Montreal. *Journal of Economic Geography*, *21*(5), 761–782. https://doi.org/10.1093/JEG/LBAA019

Wydra, S. (2020). Measuring innovation in the bioeconomy – Conceptual discussion and empirical experiences. *Technology in Society*, *61*, 101242. https://doi.org/10.1016/j.techsoc.2020.101242

Appendix

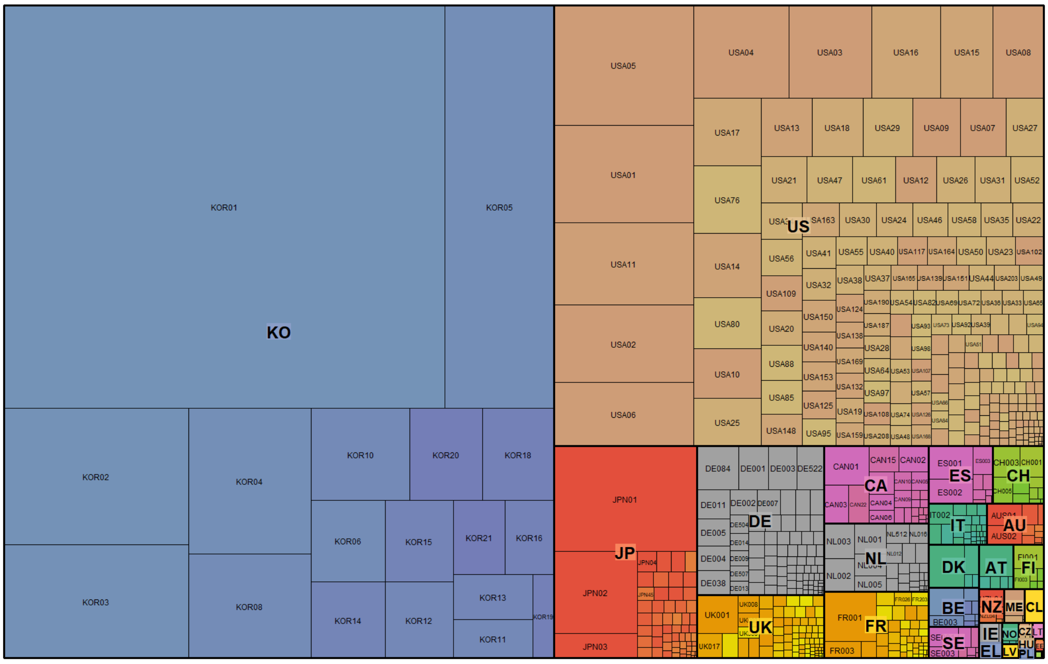


Figure A1: Distribution of bioeconomy patents across FUAs and countries, 2014

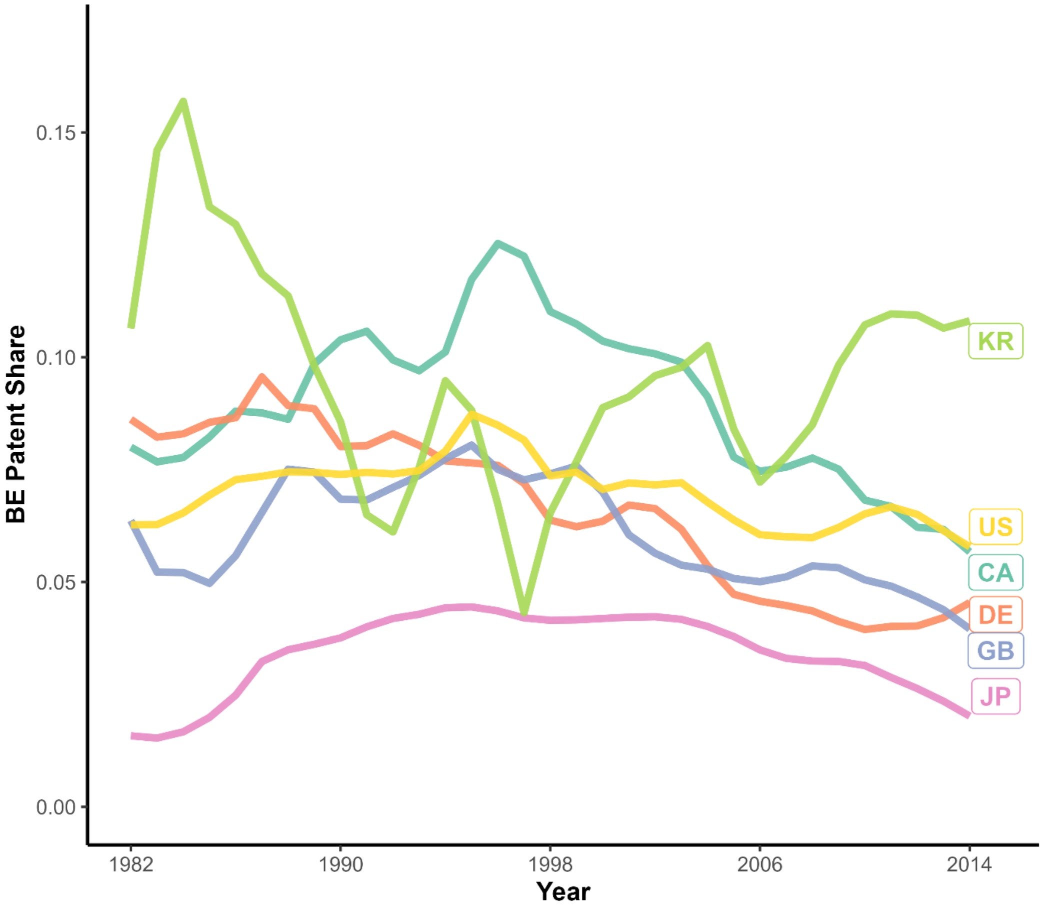


Figure A2: Share of bioeconomy patents in selected countries, 1982-2014

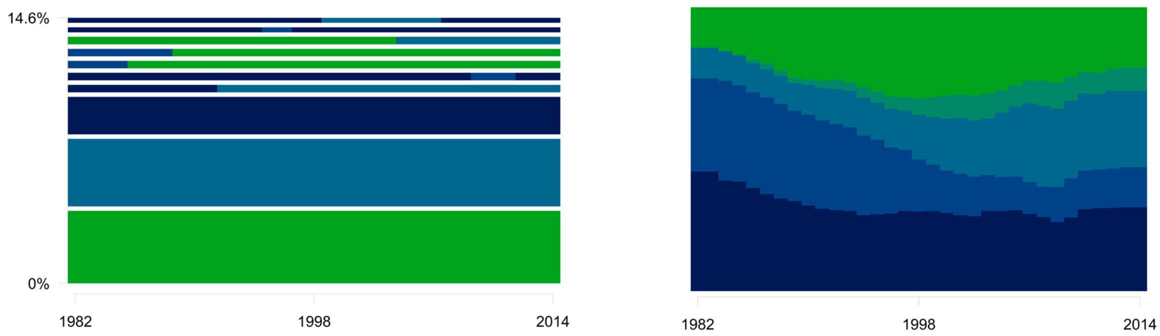


Figure A3: Most frequent sequences (left), state frequency plot (right)

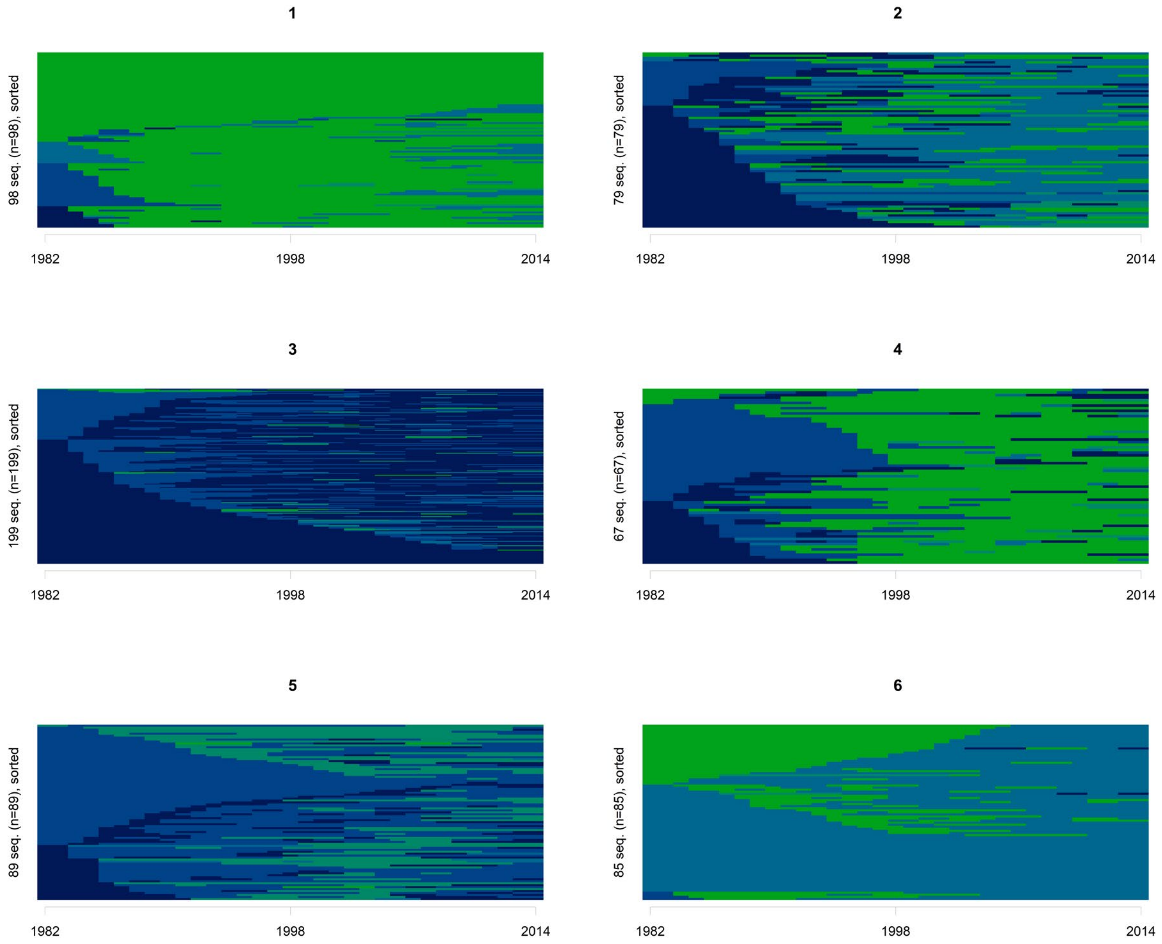


Figure A4: Sequence index plots by cluster, sorted by first state

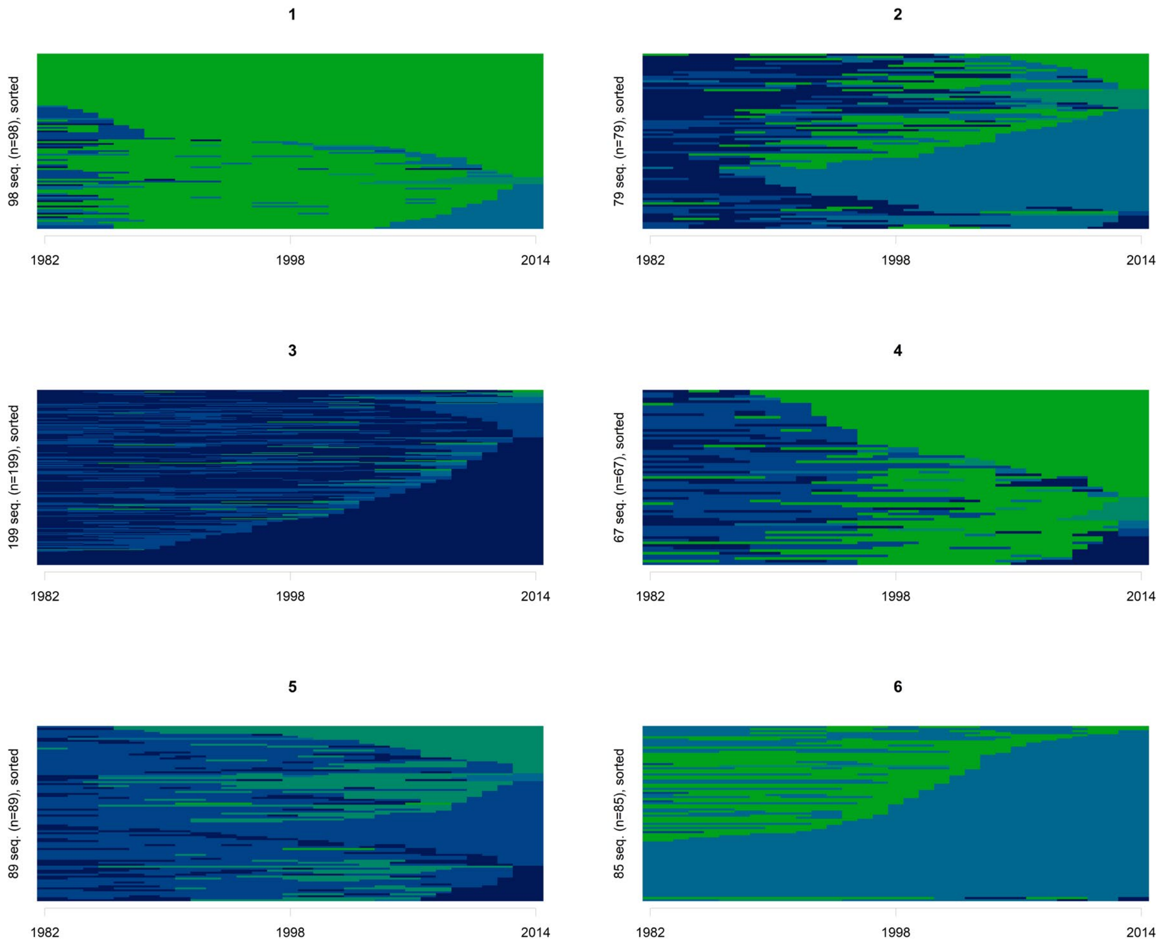


Figure A5: Sequence index plots by cluster, sorted by last state

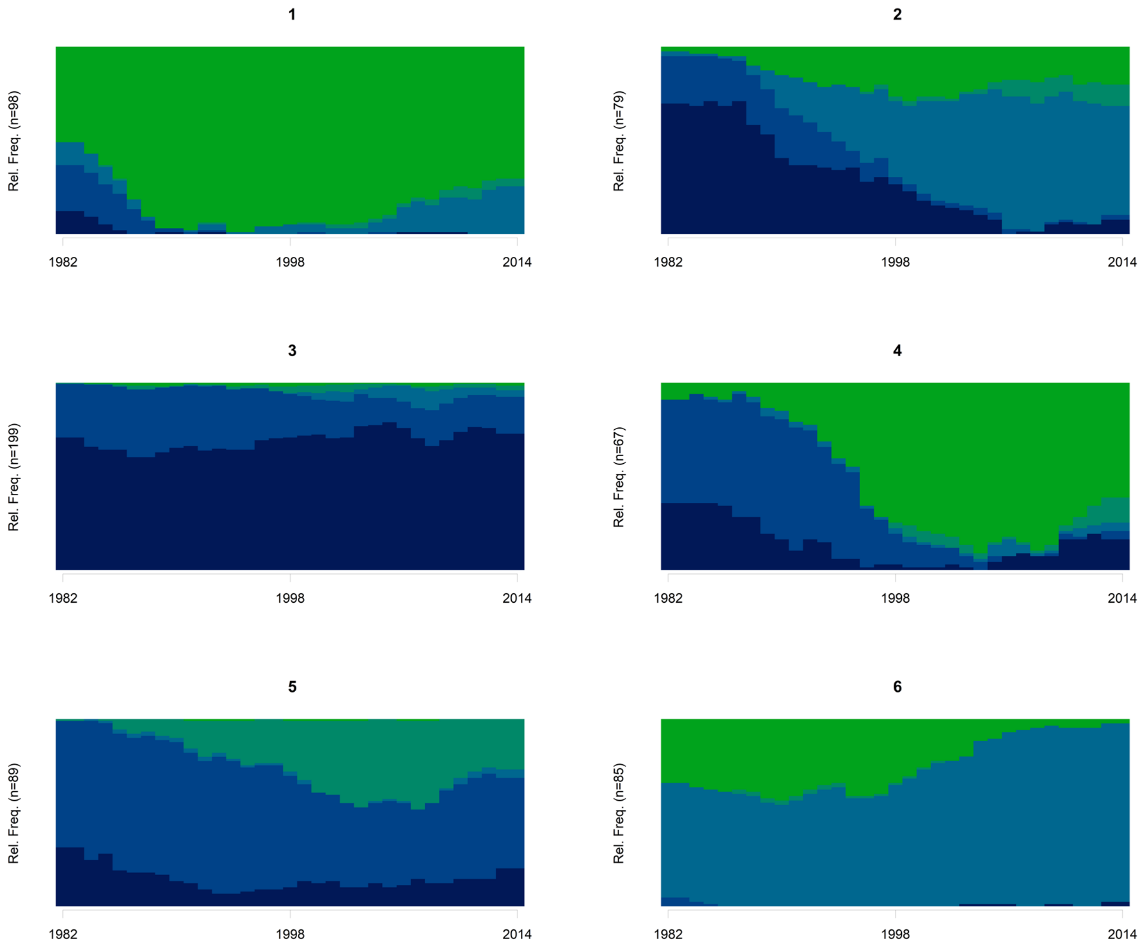


Figure A6: State frequency plot by cluster

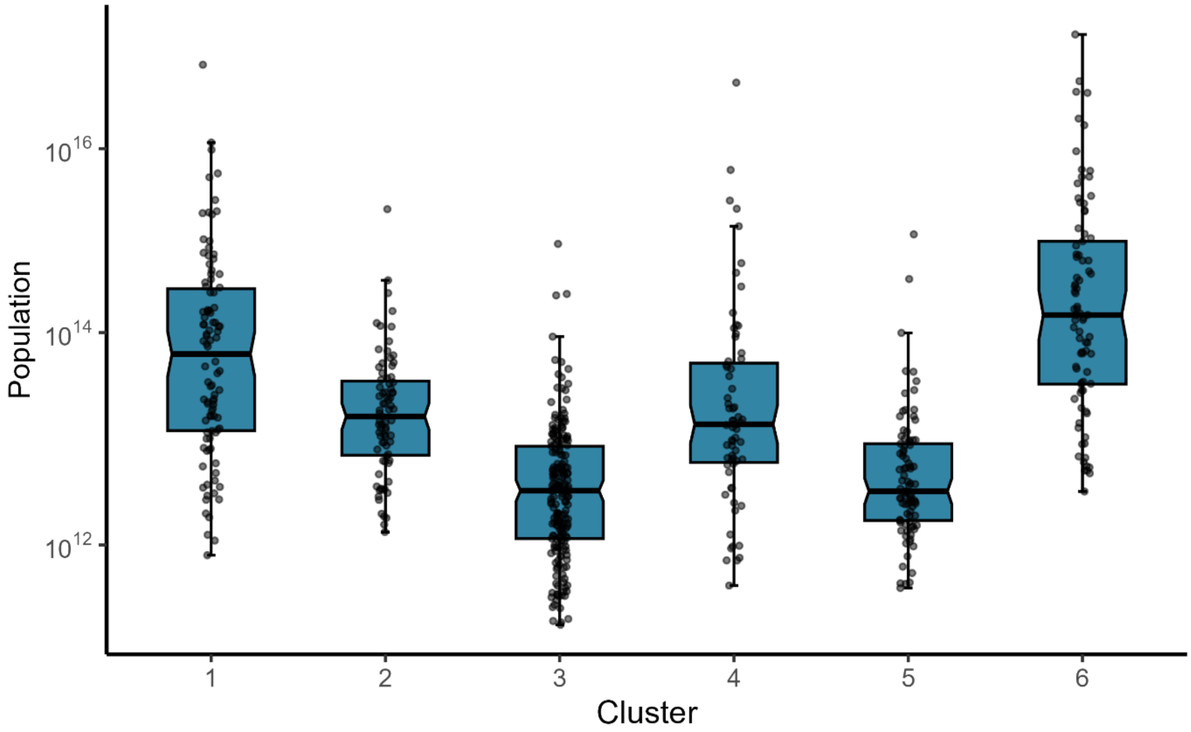


Figure A7: Regional population by cluster, average 2012 to 2014