Title	Hops, Skip and a Jump: The Regional Uniqueness of Beer Styles
Authors(s)	Hynes, Ryan M., Buarque, Bernardo S., Davies, Ronald B., Kogler, Dieter Franz
Publication date	2023-06-12
Publication information	Hynes, Ryan M., Bernardo S. Buarque, Ronald B. Davies, and Dieter Franz Kogler. “Hops, Skip and a Jump: The Regional Uniqueness of Beer Styles” (June 12, 2023).
Series	UCD Spatial Dynamics Lab Working Paper Series
Publisher	University College Dublin. Spatial Dynamics Lab
Item record/more information	http://hdl.handle.net/10197/24460

Hops, Skip and a Jump: The Regional Uniqueness of Beer Styles[1]

Ryan M. Hynes, School of Economics, University College Dublin, Dublin, Ireland

Bernardo S. Buarque, Spatial Dynamics Lab, University College Dublin, Dublin, Ireland

Ronald B. Davies, School of Economics, University College Dublin, Dublin, Ireland

Dieter F. Kogler, Spatial Dynamics Lab, University College Dublin, Dublin, Ireland

dieter.kogler@ucd.ie

Abstract

Perhaps more than any other product, beer evokes its place of origin. Part of what makes every pint of Guinness or stein of Paulaner so memorable is what sets them apart and gives them their unique "taste of place." This chapter explores the geographical differentiation of beer. To do so, we collect data on regional beer recipes, styles, and ingredients from a homebrewing website. We then employ Evolutionary Economic Geography (EEG) methods and create weighted co-occurrence networks for the ingredients within each style. We use these networks to identify which ingredients are most important to each beer style, measure a style's robustness, and compare differences between geographically close and distant styles. While previous literature focuses on the related diversification of regions, we use these methods to examine the differences within the same product and across many regional styles and flavours. Combining the EEG methods with this unique ingredients dataset, we show that almost all beer styles rely on only a handful of key ingredients. Yet some regional beers are more robust than others due to readily available substitute ingredients in their region. Likewise, we demonstrate that styles originating in close geographic proximity are more similar in their use of ingredients.

Keywords: Regional Analysis, Evolutionary Economic Geography, Network Analysis, Beer, Ingredients

[1] The Appendix for this chapter is available online at: http://hdl.handle.net/10197/24458

Introduction

In this chapter, we study the regionalization of products using beer styles as an example. We collect an inclusive and extensive dataset on beer recipes - their ingredients, styles, and historical origins - to study how this final good diverges and diversifies across regions and time. Combined with ideas and methods from Evolutionary Economic Geography (EEG), we use these data to compare the co-occurrence of ingredients within styles originating at close and distant physical locations. We measure the variability of inputs used in each style and how much they depend on a few vital local ingredients. In doing so, we aim to extend the literature by demonstrating the significance of regional resources, knowledge, and practices in the differentiation of products.

Past literature focuses on the proximity between products to study the "related diversification" of regions (Hidalgo et al., 2018). These studies rely on the co-occurrence of exports, patent codes, skills, jobs, music genres, and even Olympic sports to expose that "entities are more likely to become successful in activities that are related to their current activities in terms of knowledge and other resources" (Knuepling and Broekel, 2022, p.183). However, they seldom examine how local competencies influence the quality and nature of products made in particular locations (Hummels and Klenow, 2005; Henn et al., 2015). In other words, the current literature lacks empirical studies stressing and uncovering regional differences within the same product. We seek to address this gap. Namely, since we expect these within-product disparities to reinforce and result from the same processes that lead to the related diversification of regions, in this chapter, we borrow and build upon the available methods in the literature to map the regional uniqueness of beer.

As Boschma and Franken (2006, p.278) write, EEG "aims to understand the spatial distribution of routines over time. It is especially interested in analysing the creation and diffusion of new routines in space and the mechanisms through which the diffusion of fitter routines occurs." By replacing routines with recipes, the literature offers us the background to understand how local resources shape the popularity of beers and pave the way to new combinations of ingredients (Metcalfe et al., 2006). In other words, EEG teaches that "experiences and competencies acquired over time by individuals and entities in particular localities to a large degree determine the present configurations as well as future regional trajectories" (Kogler, 2015, p.705). Thus, grounded in this literature, we expect the local beer recipes to gradually change as agents face unique "experiences and competencies" to the point where these become "fundamentally different goods" (Schott, 2004). That is to say, the different experiences and endowments across space eventually transform these beers - or any other final good - into new variations or styles that are particular to a place.

Therefore, EEG offers the right setting to study and measure the within-product differences resulting from local "experiences and competencies." And so, instead of focusing on regions branching into new activities close to their portfolio, we use the insights from EEG to examine differences within the same good - but which originate from distant places. We use the co-occurrence method not to measure product relatedness but to compare the "current arrangement" of beer recipes. We borrow and adjust the EEG methods to answer what sets these local beer styles apart. Following this logic, our approach allows us to examine how the unique local endowments transformed the brewing experience in space - even as we now brew and consume these beers worldwide.

To sum up, our goal is beyond exploring the geography of beer styles. Instead, we use it as a case study to learn about the regionalization of products through the EEG lenses. Beer is one of the oldest and most widespread products of human civilization. Yet it is surprisingly local. Hence, it offers unique insights into the diffusion and regionalization of products. Think about the California Commons, a beer style born out of German brewers struggling with the lack of refrigeration in the American West during the Gold Rush. Or the Bavarian Weissbier, which under the Reinheitsgebot could only be produced by those with exclusive royal rights. And the Belgian Saison that was traditionally brewed by isolated farmhouses in Wallonia using locally grown ingredients. All these examples show the role of climate, regulations, social arrangements, and much more in developing these regional styles and their peculiar recipe configurations. More importantly, the dataset we collect allows us to capture, to estimate the differences across these beer styles. We have detailed information on every recipe - including the weights and measures of individual ingredients - and we use this data to make co-occurrence networks for every beer style in our sample. Besides, since we have precise information about the historical development of beer styles from the Beer Judge Certification Program (BJCP), we can place these in space and group them into families according to their origins. Therefore, we treat the networks as the "fingerprints" of brewing know-how specific to each style's region and history. Succinctly, we argue that beer recipes make an excellent case study, and our data coverage provides the right setting to study the differentiation and regionalization of products.

While the results presented in this chapter remain exploratory, it highlights possible paths for the related diversification and EEG literature. First and foremost, as we argued before, it shows the need to consider how local competencies shape the quality and nature of products. Following this path, future efforts will enhance our very grapes of the concept of relatedness (Juhász et al., 2021). They will provide support for Smart Specialization Strategies pushing forward quality upgrading (Amiti and Khandelwal, 2013; Fernandes and Paunov, 2013; Crespi et al., 2014; Radosevic, 2017) and possibly introduce entirely new research questions and goals to the literature. Consider, for example, past work looking at place-of-origin labels as a branding instrument. These lean on the marketing literature and consider the place of origin as a "competitive differentiation" label that influences consumers' attitudes and purchase intentions (Verlegh and Steenkamp, 1999; Dinnie, 2004). Indeed, writing about the German beer market, Lentz et al. (2006, p.251) evaluate that the "place of production of a good [...] are important attributes by which customers assess their quality." Yet, in this chapter, we used another approach to distinguish between the unique "taste of place" associated with the different beers. We look at the internal composition, the combination of ingredients within every recipe, to measure and understand the role of local inputs in differentiating between styles. And we suspect that future research could do the same for other products with strong regional ties. Hence, as we push the EEG methods to new cases, the regional science literature will provide the means to examine and possibly substantiate the desire to attribute authenticity to a location via appellations.

Another contribution stems from expanding the relatedness methodology into a new area and data source (Whittle and Kogler, 2020). Patent data gives researchers a wealth of information on every invention - e.g., priority year, citations, technological codes, inventors and their locations, etc. Thus, unsurprisingly it has become a valuable source for regional studies literature to graph the dynamics of knowledge creation and integration across space (Griliches, 2007; Neffke et al., 2009; Leydesdorff et al., 2017). Nonetheless, there are equally known limitations to applying and interpreting data from patents and their inherent classification schemes. And the knowledge space method is not exempted from these concerns (Kang and Tarasconi, 2016; Jaffe et al., 2019). As such, the growing availability of alternative data sources, text mining, and other tools could prove essential to circumvent some of these well-established issues from relying on patent and CPC codes. Future efforts following this path could prove valuable to strengthen, validate, and appreciate the limits of the existing practices (Klement and Strambach, 2019). Expanding these methods to new data will advance the field because knowledge does not exist just in the ivory tower but also in everyday products around us. Even something as deceptively simple as beer is full of complex relationships ripe for thorough analysis - something we sought to demonstrate in the current contribution. Indeed, compared to previous work using patent data, we can create highly detailed recipe-style networks because we collect the weights and measures of each ingredient within recipes. Because we know the exact proportions used in each recipe, we no longer must weigh the components equally. Our precise measures allow us to properly reweigh the edges between ingredients based on their relative proportions and thus use the full power of the co-occurrence networks. This may seem like a minor detail, but it allows a level of precision not currently possible in the innovation and regional studies literature. Two beers may have nearly the same list of ingredients yet taste widely different because of the ratios of their combinations. The recipe, like a patent, is not simply a bill of materials but a set of instructions on how to combine them. The level of detail, missed in the current literature due to limitations of patent data, opens a rich new set of questions.

Our work also contributes towards efforts to measure and understand local knowledge capabilities. We stress the regional differences regarding the structure of knowledge networks (Buarque et al., 2020). Instead of plotting a universal knowledge space to measure relatedness, as usual in the literature (Kogler et al., 2013; 2017), we build and compare the co-occurrence networks of each style. We attach each beer style to a unique network and thus allow the graphs' structure to change from place to place. These become the "fingerprints of inventive activity" and carry valuable data regarding how the different styles combine, recombine, and transform recipes. And the growing availability of means to compare networks (Tantardini et al., 2019) can fuel our analysis of what distinguishes the inventive regions. Future research can benefit from these methods to examine what makes a place unique or what drives its similarities. They can study how geography and social proximity explain the similarity across groups or focus on the role of history and institutions in shaping these different networks - like we showed that beer ingredient networks are far closer within regions than within families. Indeed, although still rare, we start to observe investigations using this insight to understand, for example, how the "technology network structure conditions the economic resiliency of regions" (Toth et al., 2021). In summation, examining regional variations in the network structure will promote robust inferences about the causes and consequences of innovation. And once again, the effort outlined in this chapter suggests valuable research paths that can foster our grasp of regional knowledge.

Our final contribution to the literature challenges the traditional definition of a region. We often represent regions as administrative or territorial demarcations with rigid borders, like the NUTS2 zones in Europe. These strict definitions are helpful but imperfect. And they commonly ignore fuzzier concepts like industrial clusters, job markets, and commuter zones (Parr, 2007; Conttineau et al., 2018). Many regions of cultural and economic importance span borders and therefore require a more nuanced definition. Besides, recent publications show that "off the shelf" territorial delineations "vary significantly and [...] may introduce noise or regional bias that merits consideration in any analysis conducted with these units" (Fowler and Jensen, 2020, p.1396). As follows, a growing body of work proposes combining socioeconomic data along with geographical divisions to ensure better-fitting demarcations (Ratti et al., 2010; Thiemann et al., 2010; Piccardi and Tajoli, 2012; Barber and Scherngell, 2013; Hawelka et al., 2014; Sobolevsky et al., 2014; Pei et al., 2014). In light of these recent advances, we choose to take an uncommon definition of regions but one well in line with our topic of interest. Because beer styles span borders/cultures and often intermingle within traditional statistical units, we take a more flexible definition of regions, equating each beer style to its historical place of origin. These can sometimes precisely identify a city - such as Kölsch or Munich Helles. But they can also cover large swaths of places and people like the American IPA and its English counterpart. We imagine these regions accurately reflect our research question, and they highlight the value of diverse and outside-the-box definitions to study human organization across and within territories.

At last, although not our primary objective, we also contribute to the literature on the geography of beers. A large body of work discusses the regional aspect of beer (Patterson and Hoalst-Pullen, 2014). These cover the spatial diffusion of beer from its origin in the Fertile Crescent (Sewell, 2014) and the development of geographic appellations (Mittag, 2014). Several papers discuss the clustering of microbreweries in space (Elzinga et al., 2015; Cabras and Bamforth, 2016; Dennett and Page, 2017; Gatrell et al., 2018; Wojtyra et al., 2020). And, most relevant for this chapter, we also find qualitative investigations regarding the history, importance, and sensitivity of local ingredients that give each beer its unique "taste of place" (Yool and Comrie, 2014; Kind and Kaiser, 2020; Knudson et al., 2020). However, there has yet to be an empirical study of beer ingredients and regional variation. That is likely because it is difficult to get relevant and inclusive data on the subject. Thus, our contribution is assembling data on beer recipes, styles, ingredients, and their locations. These allow us to quantitatively explore what sets beer styles apart. We can then easily measure which ingredients are most important to a beer style and region. We can also observe how robust the networks are to losing local ingredients. That is to say, we bring highly detailed micro-data to longstanding questions in the geography of beer. We also marry this literature with analysis from EEG and innovation studies. In doing so, we shed light on old questions and pave the way for others to ask and answer new ones.

The rest of the chapter proceeds as follows: Section 3 discusses how we fetch, parse, and normalize the recipe-level data. Section 4 transforms recipe ingredient data into style networks. Section 5 introduces eigenvector centrality, our main measure of ingredient importance. Section 6 details our targeted deletion strategy. Section 7 defines the ability of certain styles to weather losses of key ingredients. Section 8 posits that geography, and the abundance of ingredients area key determinants of robustness. Section 9 concludes.

Data Collection and Mapping

We gather data on 126,256 beer recipes and map them to individual styles, which in turn can be historically linked to countries, regions, and even cities. We use the authoritative BJCP Style Guide to define broad styles of beer, then match beer recipes to styles. We get our beer recipes and their ingredients by downloading BeerXML files from BrewersFriend.com. BrewersFriend allows home brewers and small craft breweries to record and manage their recipes. Recipe ingredients are broken down into hops and malts, each of which detail the types and amounts of ingredients added to the recipe. Figure A (Appendix)[2] provides an abridged example of the BeerXML file for one such recipe. BrewersFriend allows recipes to be made publicly accessible or otherwise marked private. The 126,256 public recipes on BrewersFriend form the basis of our sample. We download these public recipes in BeerXML format, then parse the ingredients in each of the five categories into separate tables. Once parsed, we spend considerable effort disambiguating ingredient names so that they may be matched to multiple recipes.[3]

[2] The Appendix for this chapter is available online at: http://hdl.handle.net/10197/24458

[3] For example, one recipe may use "CaraPils" malt and another "carapils" malt, even though these are the same underlying ingredient.

We then turn to refining our sample. We first restrict our sample to include only recipes using whole ingredients. Some recipes in BrewersFriend use pre-mixes from brewing kits that already combine ingredients and therefore offer little information about the choice or combination of ingredients. This restriction leaves us with 109,015 unique recipes, or 86% of our original sample. We then turn to regionalizing our recipes through their styles.

Each recipe is associated with a single official BJCP style. BJCP styles are an international standard used to group and evaluate beers at brewing competitions worldwide. We group our recipes into 144 different BJCP styles and drop 3,159 recipes that do not specify a style. We drop these BJCP "Specialty Beers" styles including mead, cider, and other non-beers and lose an additional 4,821 recipes (4% of our remaining total). We are left with 101,034 recipes covering 111 styles. Table B (Appendix) lists these styles and the number of recipes in each.

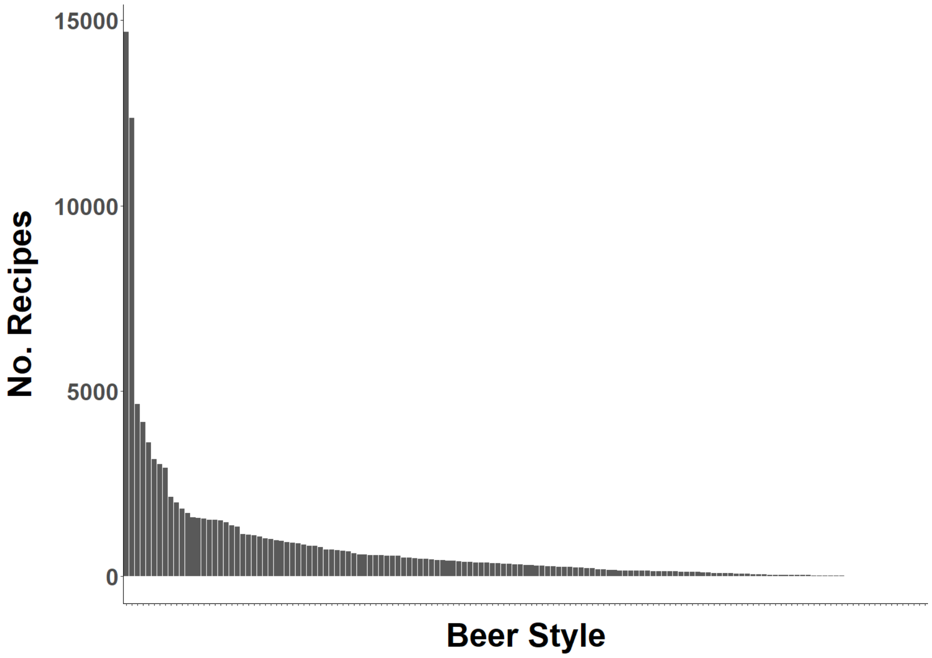


Figure 1: Count of Recipes per Style

Note: The histogram presents the distribution of recipes per beer style. We collect more than 100,000 unique beer recipes from BrewersFriends. All recipes belong to one and only one beer style. The x-axis displays nearly 150 distinct beer styles, where each bar corresponds to one beer style. The users of BrewersFriend catalog their recipe's style following the Beer Judge Certification Program (BJCP) classification scheme. And we order the beer styles according to their number of recipes, where the American IPA is the most popular style with around 15,000 different recipes.

The distribution of recipes across styles is highly skewed. Two styles, American IPA and American Pale Ale, represent more than 25% of all recipes. This may represent an underlying bias in our data as

BrewersFriend is based in the United States, or it could also reflect the tremendous popularity of these styles.[4] However, there are thousands of international users of BrewersFriend and over 100 styles with at least one thousand recipes each. Figure 1 shows the distribution of recipes by individual style. There are a handful of styles, such as New Zealand IPA, that only have one recipe associated with them. To ensure adequate variation within styles, we further restrict our sample to styles that have at least 100 recipes. We lose only 596 recipes (less than 1% of our sample) with this additional restriction but do miss out on a few valuable regional styles like Rauchbier, which is particular to the town of Bamberg in Germany. After cleaning and regionalizing our style sample, we have 100,438 recipes covering 90 styles and spanning 13 countries. We map these styles into 25 regions with varying levels of precision.[5] We then turn to our ingredient types of interest: hops and malts.[6]

We identify 4,882 different malt names across our sample, however not all these malts are truly unique due to minor variations in their names. We disambiguate these malts by first removing all nationality and company information from the name.[7] We then remove special characters and lowercase all names. We fuzzy match our cleaned list of malts back to the recipes and confirm the matches by hand. Like styles, we remove infrequently used malts appearing in fifteen or fewer recipes. We are then left with 170 unique malts used in 99,943 recipes. Table C (Appendix) lists all these disambiguated malts.

Like Malts, we begin with a list of 5,023 Hops that appear at least once across our recipes. We repeat the name normalization process above, removing brand names and indications of origin. We once more confirm these results by hand, paying particular attention to code names and translations. For example, one common hop, Saaz, is known in the Czech Republic as Žatec, the name of the town where it is produced. After normalizing we perform a fuzzy match to a list of well-known hops provided by both Barth-Haas and Hoplist.com. Barth-Hass is one of largest producers of hops worldwide and has developed a ubiquitous Tasting Guide detailing the flavour profile, alpha acid, and location of global hops (BarthHaas, 2018). Hopslist.com similarly maintains a global reference of hops and their locations (Healey, 2016). We disambiguate our 5,023 hops from our recipes to just 229 global hops from the Barth-Haas and Hopslist.com lists. Unmatched hops are almost all due to misclassification such as listing fruit or spices as hops, or other user data-entry errors when creating the recipe.

---

[4] As a robustness check, we randomly draw a sub-sample of American IPA recipes in proportion to the number of recipes in the styles we compare with American IPA. Our results are largely unchanged, so we present the full network in our comparisons.

[5] Some beers like the California Common can be located to a specific city and even a particular brewery: Anchor Brewing in San Francisco. Others have less precise origins. The American IPA is primarily attributed to the West Coast of the United States but is also fairly ubiquitous across the country. Finally, most British beers can only be mapped to the national level, i.e., Scotland or England.

[6] BrewersFriend.com provides five categories of ingredients: Hops, Malts, Yeasts, Water, and Miscellaneous. We focus on hops and malts because they: 1) are arguably the most important ingredients in the recipes; b) are almost always combined with different varieties in recipes, as opposed to yeasts; 3) are the most readily identifiable and easy to accurately localize.

[7] For example, "US - Castle Malting - Pilsner Malt" simply becomes "pilsner."

Table 1: Recipes and Styles by Country

Country	Recipes	Styles	Avg. Color	Avg ABV	Avg IBU	Avg OG	Avg FG	Avg Carb.	Avg Malts	Avg Hops	Avg Oth.
United States	56602	36	11.12	6.10	50.15	1.06	1.01	0.94	2.71	2.52	1.38
United Kingdom	15770	34	24.16	6.05	37.30	1.06	1.02	0.80	3.35	1.88	1.47
Belgium	11577	17	8.95	6.43	25.81	1.06	1.01	1.01	2.51	1.85	1.42
Germany	8307	34	10.07	5.66	21.64	1.06	1.01	1.15	2.58	1.56	1.30
Specialty Beer	4164	37	12.37	6.51	27.39	1.06	1.01	0.61	2.90	1.76	1.58
Ireland	2778	3	18.60	5.20	28.00	1.05	1.01	0.95	3.27	1.68	1.17
Russia	1302	1	47.48	9.90	60.98	1.10	1.02	0.60	4.30	2.01	1.55
Czech Republic	857	5	6.23	5.52	35.37	1.05	1.01	0.95	2.24	1.68	1.08
International	525	3	7.24	5.62	26.62	1.05	1.01	1.26	2.21	1.87	1.22
Austria	492	1	11.08	5.33	25.20	1.05	1.01	0.89	2.85	1.79	1.10

Note: The table shows summary statistics for the beer styles grouped according to their country of origin. We infer the style's country of origin from their historical development, as determined by the Beer Judge Certification Program (BJCP). The first column shows the country of origin, followed by the total number of recipes and styles that belong to the country. The remaining columns present the averages for several beer characteristics. Avg Color refers to the coloring units according to the European Beer Convention (EBC). ABV stands for Alcohol by Volume. IBU is an abbreviation for the International Bitterness Units scale, a standard measure of beer's bitterness. OG and FG indicate the original and final gravity, respectively. These provide brewers with estimations for the amount of sugar in the beer and its potential alcoholic volume. Avg Carb describes the amount of CO_2 present in the beers. Avg Malt, Avg Hops, and Avg Oth. represent the average amount of ingredients used in the recipes.

We similarly restrict our sample to hops appearing in 15 or more recipes to ensure variation across our sample. We lose only 362 recipes with this restriction, leaving us with 229 unique hops used in 92,813 different recipes. Table D (Appendix) lists these hops.

After parsing, disambiguating, and cleaning our sample we are left with 92,813 recipes made from 170 malts and 161 hops across 90 different styles. Table 1 summarizes these data. We now use these data to create recipe-ingredient networks for each style.

Beer Style Networks

We create ingredient co-occurrence networks for all 90 beer styles in our sample. Each recipe represents a unique combination of hops and malts, at the extensive margin if the ingredients appear in a recipe, and at the intensive margin based on the relative proportions used of each input. These style networks are graphical representations of the distinct combinations of ingredients and their volumes.

Each ingredient is a node in the style network. We draw an edge between two ingredients whenever they co-occur in the same recipe. Each edge is weighted in proportion to the amounts used in the recipe. For example, if a recipe uses 1kg of Pale 2-Row malt for every 100g of Chocolate malt, we value the edge between these two ingredients as 1/10.[8] Because every beer belongs exclusively to one style, we can combine the nodes and add their weighted edges to form unique style networks. If the same ingredient pair appears in more than one recipe of the same style, we sum up their weights.

Our style networks describe the relationship between the ingredients used in every beer recipe of a given style. The networks allow us to visualize the unique combinations of ingredients that make up a beer style. We can also represent these style-ingredient relationships algebraically:

$$S_{ij} = \begin{bmatrix} s_{11} & \cdots & s_{1n} \\ \vdots & \ddots & \vdots \\ s_{n1} & \cdots & s_{nn} \end{bmatrix}$$

[8] One does not usually observe the volume of each input used in the end product when using patents or other data sources to build co-occurrence networks. Instead, this literature typically weights edges based on the shares of the nodes. For example, if four technological codes appear in the same patent each gets a weight of ¼. For this reason, we also reproduce our analysis weighting the edges of the style networks by the ingredients' share. Our results are robust to using this more common weighting method.

where S_{ij} is the style's adjacency matrix and every entry s_{ij} measures how often ingredients i and j appear together weighted by their relative proportions. The adjacency matrix above can also be visualized as a style network.[9]

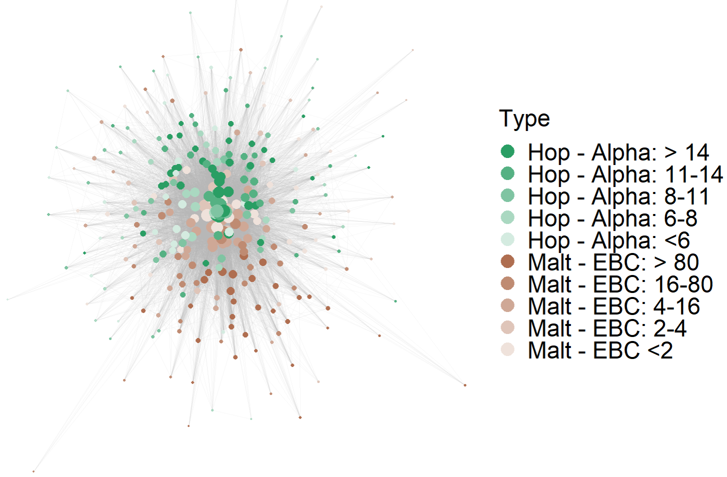
Figure 2 plots two such style networks. Panel A shows the style network for American IPA, the most popular style in our sample. Panel B shows the style network for Kölsch, a beer style named for the German city where it was first created, which perhaps best captures the regional nature of styles.[10] We create both graphs using the Kamada-Kawai force-directed drawing algorithm, which minimizes total path length by placing ingredients commonly used together next to one another (Kamada et al., 1989). Likewise, Kamada-Kawai puts the most connected nodes at the centre of the network. The size of each node is proportional to that node's degree centrality, or how many connected links a node has. The width of each edge is proportional to the weights of the ingredients' as they co-occur in recipes. Hops are coloured in green and shaded by their alpha acid intensity, a proxy for bitterness. The darker the green, the more bitter the hops. Malts are coloured brown and shaded by their European Beer Convention (EBC) coloration.[11]

These two styles and their graphs are quite different. American IPA uses many more unique ingredients than Kölsch (321 nodes against 175). American IPA's ingredients are also more connected to one another with more than 17,000 total edges between its nodes, each of which has 108 edges on average. Kölsch, on the other hand, has only 2,000 total edges and an average of 20 edges per node. The American IPA has a very high network density, which is the number of actual edges between nodes out of all theoretically possible edges. In fact, the American IPA has a relatively high network density of 0.34 or 34% of all possible edges, while the Kölsch has a network density of only 0.13.

The American IPA network seems to be more robust and complex than that of Kölsch. The American IPA network includes more ingredients with stronger connections between them. Still, one might argue that because our sample of American IPA recipes is much larger than any other style, and more than ten times greater than Kölsch, we misrepresent its network connectivity.[12]

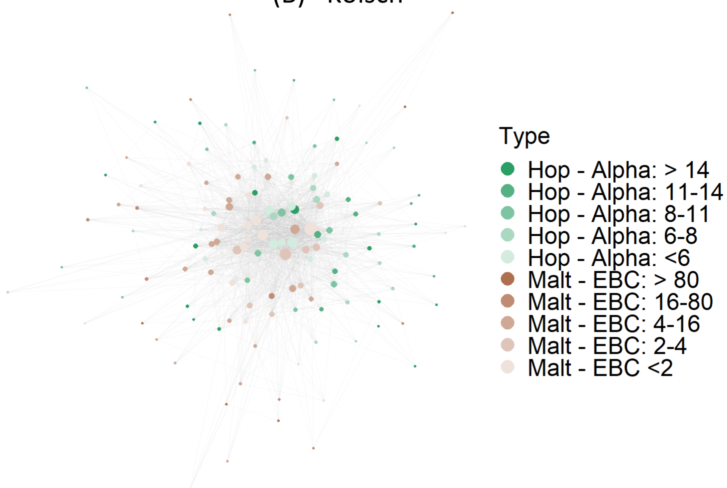
[9] The adjacency matrix, edge list, and networks are different ways to represent the same relationship between nodes and edges. We provide definitions for all three in the Appendix. See primary references Wasserman et al. (1994) and Barabási et al. (2016) for further information.
[10] Since 1997, Kölsch has held a Protected Geographical Indication (PGI) within the European Union.
[11] EBC coloration is a grading scale based on the colour a particular malt imparts on the beer. Pilseners and other light beers have an EBC of 4, whereas darker malt beers such as stouts have an EBC of 70.
[12] We test if the differences between the American IPA and Kölsch are the result of sample size. To do so, we take 1,000 random subsamples of American IPA consisting of 1,000 recipes each, approximately the same number of Kölsch recipes. Although on average the American IPA random subsample

(A) American IPA



Nodes	Density	Diameter	Clustering	Avg. Degree
321	0.34	1.20	0.54	108

(B) Kölsch



Nodes	Density	Diameter	Clustering	Avg. Degree
175	0.13	1.50	0.39	23

Figure 2: Example Networks

Note: The graphs show the co-occurrence of ingredients within every recipe that belongs to the style. We use a Kamada-Kawai algorithm to plot the ingredients as nodes and draw an edge whenever they co-occur in the same beer recipe. We weigh the edges in proportion to their relative amounts. The color of the nodes depends on their ingredients' type and intensity. We show the malts in shades of brown and hops in green. Below each graph, we display a table containing descriptive statistics of the networks.

networks are not as connected as the full sample American IPA, they remain more connected than the Kölsch one. The random sub-samples have more nodes (22) and edges between them (4,600), more than the Kölsch network. The sub-sample networks also have a higher density (0.17), average clustering coefficient (0.40), and degree (40), as well as a shorter diameter (1.4).

Nevertheless, these differences exist across all styles in our sample, and are evident even when considering networks with a similar number of recipes.

For example, California Common, another typical beer from the American West Coast, has a similar number of recipes as Kölsch at about 900 each. Despite the similar number of recipes, California Common lists more ingredients (220 nodes) and a higher average number of edges per node (38). Also, the California Common's average clustering coefficient is 0.46 compared to Kölsch's 0.39, meaning it includes relatively more "three-way" connections between ingredients. The California Common's network is also smaller than Kölsch's in the sense that it takes fewer steps to traverse the network. Indeed, California Common's maximum shortest path, or diameter, is 1.0 compared to Kölsch's 1.5.

These style networks provide us with a tractable method to visualize and model the relationship between recipes and ingredients. We now use these models and their properties to identify key ingredients, robustness, and relatedness across beer styles.

Eigenvector Centrality

While there is clearly significant variation among styles and their networks within our sample, our goal is to identify which ingredients set these networks apart and differentiate the beer styles. In other words, we are looking for the most important ingredient nodes in a given style network. And we turn to eigenvector centrality as a measure of each node's relative importance within a network. We follow the seminal work of Bonacich (1972) and calculate centrality as the weighted sum of the centrality of all adjacent nodes. Mathematically we can express eigenvector centrality as:

$$\lambda c(v_i) = \sum_{j=1}^{n} s_{ij}\, c(v_j)$$

where λ is the eigenvalue scale factor, $c(v_i)$ represents the centrality score of node vector v_i and s_{ij} is the weighted edge between nodes i and j. Algebraically this represents every element in the adjacency matrix (S_{ij}).

Eigenvector centrality differs from traditional degree measures of importance because it also accounts for the relevance of a node's neighbours. Or as Ruhnau (2000, p.360) explains: "the centrality of nodes does not only depend on the number of adjacent nodes but also their value of centrality." Eigenvector centrality awards points for being linked to very central nodes even if the node itself has just a few connections. And, for this reason, it is often used in social sciences to measure the influence of agents (Abbasi et al., 2011; Li et al., 2016; Parand et al., 2016).

Table 2 shows the top ten ingredient nodes by eigenvector centrality in our original American IPA and Kölsch networks. We normalize the centrality scores between zero and one, such that the most central node in each network will always have a score of one.[13]

Table 2: Top Ten Nodes by Eigenvector Centrality

American IPA			**Kölsch**		
Ingredient	Type	Eigenvector	Ingredient	Type	Eigenvector
Citra	Hop	1.00	Pilsner	Malt	1.00
Pale 2-Row	Malt	0.91	Hallertau	Hop	0.73
Cascade	Hop	0.77	Tettnanger	Hop	0.54
Amarillo	Hop	0.75	Saaz	Hop	0.44
Simcoe	Hop	0.74	Vienna	Malt	0.39
Centennial	Hop	0.74	Hersbrucker	Hop	0.36
Mosaic	Hop	0.63	Wheat	Malt	0.29
Columbus	Hop	0.49	Perle	Hop	0.27
Chinook	Hop	0.47	Pale 2-Row	Malt	0.22
Maris Otter	Hop	0.36	Magnum	Hop	0.20

Note: The table shows the top ten ingredients for two example beer networks according to their eigenvector centrality. The first column shows the ingredient names. The second column displays their type - malt or hop - and the third the ingredients' eigenvector score. We measure centrality according to the eigenvector formula developed by Bonacich (1972) We normalize the centrality scores between one and zero, such that the most central node always has an eigenvector centrality

Once again, there are considerable differences between these two styles, this time in key ingredients. The most central nodes for American IPA are mostly bittering hops with high-intensity alpha acids from the Yakima Valley in Washington State: Citra, Cascade, Amarillo, Centennial, etc. On the other hand, Kölsch relies heavily on aromatic hops traditionally found in Pilsners and Lagers from the Bavaria and Bohemia regions such as: Hallertau, Tettnanger, and Saaz. There are likewise significant differences in eigenvector centrality between the top ten ingredients of both styles. The distance between the first ranking ingredient in Kölsch and the rest is much greater than that in American IPA, implying the German style relies more heavily on a single malt source: Pilsner. Figure 3

[13] We consider hops and malts together as both are fundamental ingredients to recipes, which we use each in different combinations.

further illustrates this difference and plots the histogram of eigenvector centrality for all ingredients in both beer styles.

Both histograms in Figure 3 show signs of long tails common in power-law and Pareto distributions, which confirm that our beer networks display scale-free properties prevalent in many social, biological, and physical systems (Newman, 2005). In scale free networks, there are often a small number of highly connected nodes with most other nodes having little to no edges. This unequal distribution persists even when the system expands or contracts, hence the name scale-free.

Because the number of edges per node is so skewed, a common trait across scale free networks is their resiliency to "errors" or the loss of nodes. Because most nodes have few connections, deleting a random node from a scale-free network does little to change the network's overall structure and function. Conversely, scale-free networks are extremely vulnerable to "the selection and removal of a few nodes that play a vital role in maintaining the network's connectivity" (Albert et al., 2000, p.379). The concepts of error tolerance and attack vulnerability are fundamental for designing and understanding communication networks such as the World Wide Web. Beer is definitely not the Internet, so instead it helps to imagine a scenario where due to climate change or diseases we are no longer able to produce one or two varieties of hops. Depending on the centrality of these lost hops in the style network, we ought to expect different effects on the network structure and number of feasible recipes.

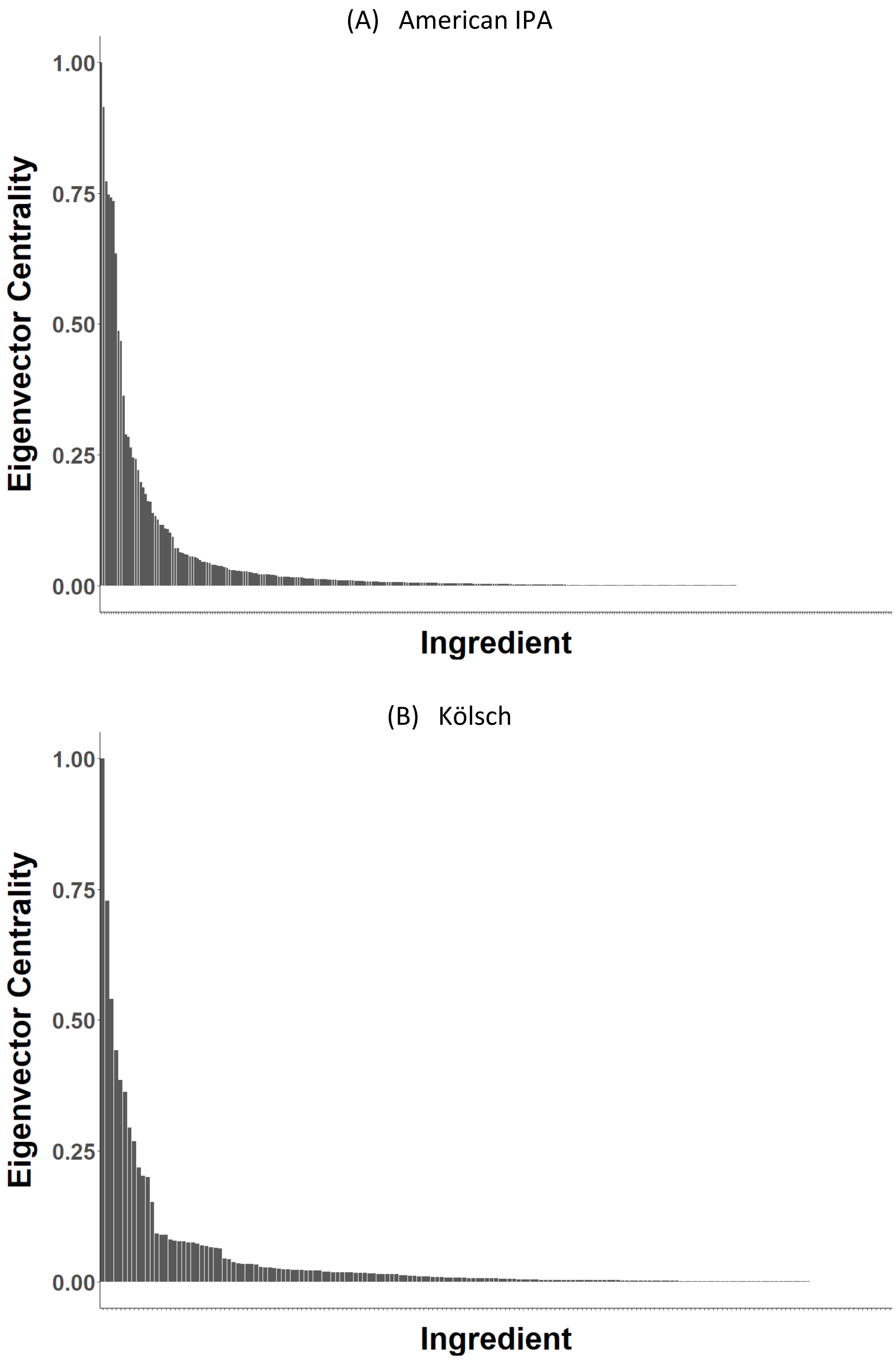


Figure 3: Eigenvector Centrality Distribution within Styles

Note: The plots are the eigenvector centrality distributions for every ingredient in the American IPA and Kölsch networks. The x-axis lists ingredients ranked by centrality. The y-axis is the eigenvector centrality score. We measure centrality according to the eigenvector formula developed by Bonacich (1972). We normalize centrality scores between one and zero, such that the most central node always has a score of one.

If the lost hops are very central to the style network, we would expect its structure to change significantly. If instead the hop is peripheral, the network structure and its observable characteristics would not change much at all. To put this idea into practice, imagine the world is no longer able to produce Citra hops. Kölsch beers would not fundamentally change, whereas the network structure and frontier of possible recipes within the American IPA network would be significantly reduced. We explore this network robustness and sensitivity to particular ingredients in Section 5 below.

Stress Test

Across the many applications of network models, a common concern is measuring their robustness or the networks' ability to withstand failures without significant loss of function or connectivity. Indeed, robustness is a valuable concept whether one wants to measure the systemic risk of the banking sector (Haldane and May, 2011) or study the stability of entire ecosystems (Montoya et al., 2006). In regional studies, network robustness gained attention since it often embodies the local capacity to manage and recover from economic, technical, or cultural shocks - the so-called resilience of regions (Martin et al., 2016). Past publications using different data sources show the importance of industrial and employment connectivity on performance following an economic crisis (Moro et al., 2021; Duan et al., 2022). And the same holds true for technological networks (Toth et al., 2022). Furthermore, these networks help scholars to forecast regional vulnerability to automation or simulate their recovery after the Covid-19 lockdown (del Rio-Chanona, 2020 et al.; 2021).

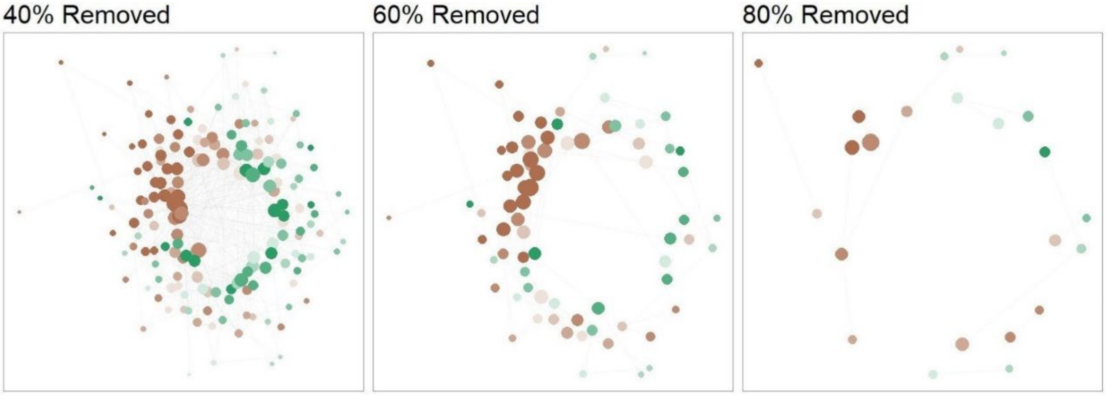
When it comes to robustness, all these examples illustrate that the structure of the underlying network largely determines "the system's ability to survive random failures or deliberate attacks" (Barabasi et al., 2016, p.274). And one can capture the robustness of different networks by measuring the impact of node removal. That is, rather than observing the centrality of a node, we can ask: what if the node had never existed in the first place? Or, in the case of beer, one might ask what would happen when growing certain hop strains is no longer possible due to extreme weather? Along these lines, we follow Albert et al. (2000) and iteratively remove nodes from our style networks. Then we recalculate the key statistics to measure how much the network changes in the node's absence. This approach allows us to assess aggregate network statistics like density or average path length as a function of one particular node. It tells us the relative influence of each node on the network's connectivity. In our case, this approach reveals how sensitive beer styles are to losing any one ingredient, which in turn reveals the ingredients' importance to the style.

We run this stress test in two ways. First, we delete nodes in rank order according to their eigenvector centrality. Second, we delete nodes at random as a baseline comparison. To truly randomize this deletion process, we run 10,000 iterations of random deletions for each network and report the average changes in network statistics. We provide a glossary of these statistics and their definitions in Table A (Appendix).

Figure 4 shows the consequences of both targeted and random deletion in the American IPA network. Panel 4A shows the resulting network after targeted deletion of 40%, 60%, and 80% of the most important nodes according to eigenvector centrality. Panel 4B shows the same process, this time removing nodes randomly. Like Figure 2, we use the Kamada et al. (1989) network plotting algorithm. To better visualize the effects of deleting nodes, we fix the network at its original layout, then remove nodes and edges from it. However, we properly re-scale the network after each deletion when re-calculating network statistics. As before, the node sizes are proportional to the weighted number of connections, and their colours depend on the ingredient type and intensity.

Panel 4A clearly shows the sensitivity of the American IPA network to targeted deletion. In contrast, Panel 4B shows American IPA's relative resilience to random deletion. Even if we delete 40%, 60% or 80% of the nodes, the resulting networks from the random attacks have more connections and shorter paths relative to the targeted attack networks. To measure how much variation, we obtain from the deletions, we compute four key network statistics and compare them to the full network. We reproduce the randomization order 10,000 times and save the density, diameter, average clustering coefficient, and average degree from the resulting networks.

(A) Targeted Attack



(B) Random Attacks

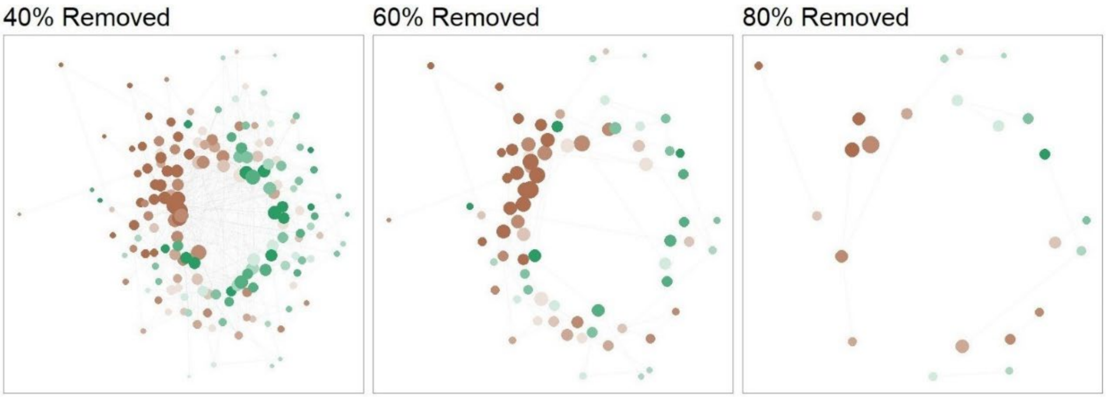


Figure 4: Stress Test - American IPA

Note: Both panels depict the impact of removing 40%, 60% or 80% of the nodes from the American IPA network. Panel A shows the effect of targeted deletion according to eigenvector centrality. Panel B shows a random attack where the nodes are deleted in random order.

Figure 5 shows the distribution of the absolute percentage change in our four network statistics after randomly deleting 50 nodes. We also highlight the changes in those statistics from a targeted deletion of 50 nodes with a dashed red line. Figures B and C (Appendix) repeat this targeted and random deletion exercise for Kölsch to much the same effect. The effects of the targeted attack are clearly much greater than its random counterpart. Even though the American IPA is the largest style in our sample and perhaps the most connected network out of all styles, it relies on just a handful of keystone ingredients without which the entire style network crumbles. These keystone ingredients are what differentiate styles and create a unique, identifiable flavour.

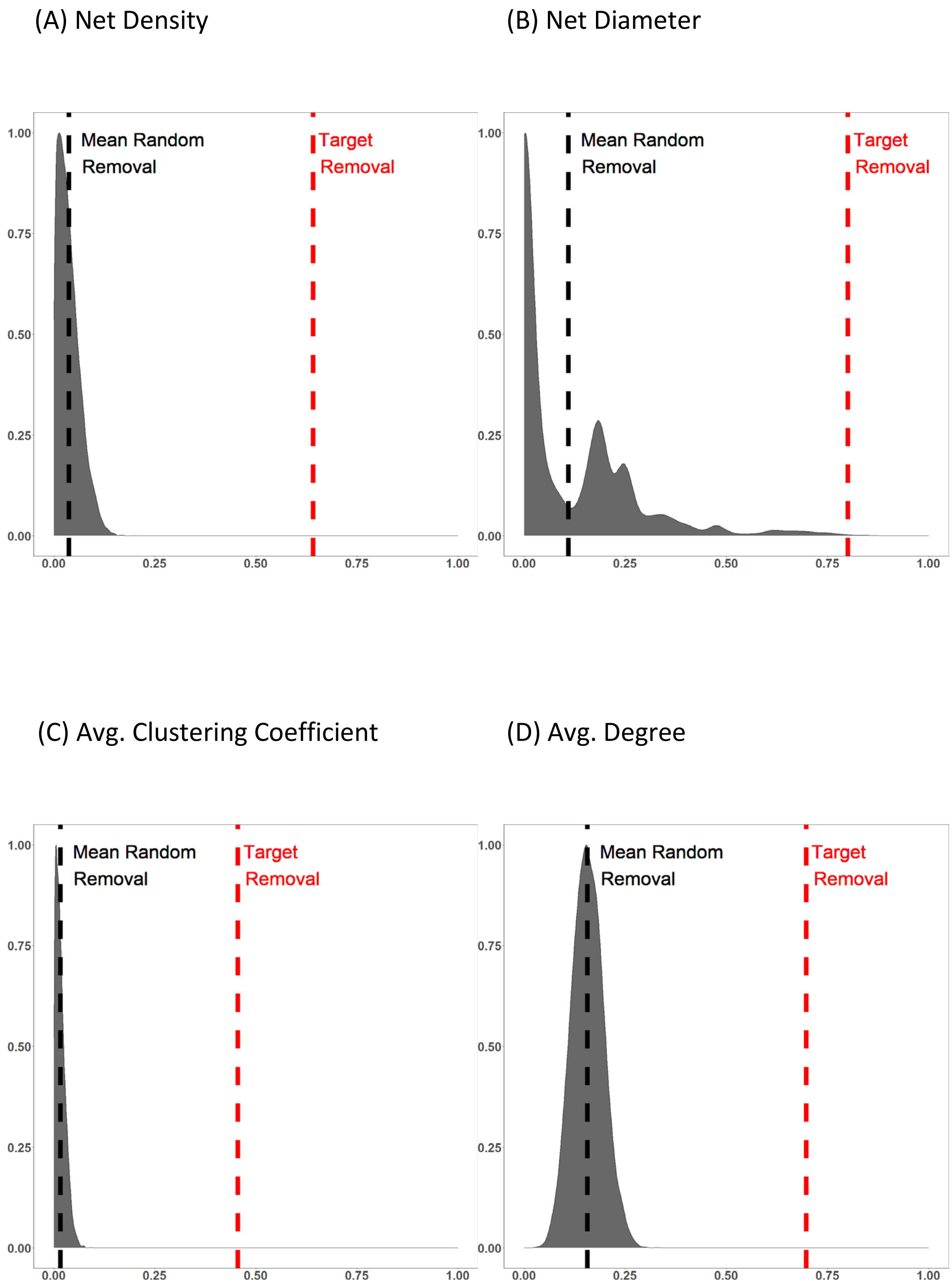


Figure 5: Resiliency Against Random Removal

Note: The figures plot the probability density distribution of the effect of deleting 50 nodes from the American IPA network. The y-axis is the probability density scaled between zero and one, such that the most frequent effect is equal to one. The x-axis is the absolute value of the percentage change of a given network statistic. Density refers to the number of edges out of total possible links. Net Diameter is the maximum shortest path. Clustering coefficient is the fraction of total three-way connections out all possible ones. Average degree is the average number of edges each node has. A network is no longer connected when Average Degree falls below two. The dashed black line is the average effect of 10,000 random deletions. The dashed red is the effect of targeted deletion according to eigenvector centrality.

Network Robustness and Local Resources

A common feature across all beer styles is their high dependence on a few key ingredients. All style networks show scale-free properties and thus are vulnerable to the failure of just a few nodes. However, there is significant variation in ingredient dependence across styles. Beer styles are not equally robust and deleting the most central nodes in one style might have a more powerful effect than in another. Let us return to our original example and compare the network and eigenvector centrality distributions of American IPA and Kölsch. The first is more robust because it has a larger number of connections and many ingredients with a relatively high centrality. So, it can afford to lose more critical nodes than Kölsch. But is this a unique attribute of the American IPA alone, or common to other styles originating from the same region or family? To find out, we compare American IPA within and across style families.[14] Table 3 introduces two new beer styles, English IPA and Munich Helles, and shows the network consequences of targeted deletion of the top fifty ingredients for all four styles. Figure B (Appendix) plots the ingredient networks for these two additional styles.

All four example networks experience a loss in connectivity after deleting the top five, ten, twenty, or fifty most central nodes. After deleting the top twenty nodes, every network is nearly half its original size by density or average degree. Likewise, network diameter nearly doubles after removing the top twenty nodes, meaning all four networks are becoming less connected and more difficult to traverse. Despite these similar trends, Table 3 also shows variation within each style's ability to withstand shocks. American IPA experiences the largest overall drop in average degree after deleting fifty nodes yet remains more connected than the full Kölsch or Munich Helles networks. Moreover, this is not just a function of the IPA family, as closely related English IPA does not exhibit the same robustness. What makes the American IPA so much more robust than other networks? It might not be only a function of sample size, but rather the result of the style's relative fungibility of key ingredients. The American IPA has more close substitutes and it makes use of more diverse ingredients.

[14] The BJCP also defines several "Style Families" that group multiple related styles. These families are listed in Table B (Appendix).

Table 3: Network Robustness

Style	Nodes Del.	% Nodes Del.	Density	Diameter	Avg. Clust. Coeff.	Avg. Degree
American IPA	0	0.00	0.34	1.20	0.54	108.50
	1	0.31	0.33	1.20	0.54	106.04
	5	1.56	0.31	1.50	0.53	96.60
	10	3.11	0.27	1.50	0.52	85.74
	20	6.23	0.22	1.55	0.50	66.80
	50	15.57	0.12	2.16	0.37	32.84
English IPA	0	0.00	0.19	1.16	0.45	49.17
	1	0.38	0.18	1.16	0.44	47.59
	5	1.90	0.16	1.70	0.43	41.77
	10	3.80	0.14	1.47	0.40	35.38
	20	7.60	0.10	1.91	0.35	25.10
	50	19.01	0.05	2.45	0.23	10.64
Kölsch	0	0.00	0.13	1.50	0.39	22.96
	1	0.57	0.12	1.81	0.40	20.89
	5	2.85	0.09	2.00	0.38	16.00
	10	5.71	0.07	2.33	0.34	11.71
	20	11.43	0.05	2.42	0.27	7.62
	50	28.57	0.01	3.30	0.30	1.85
Munich Helles	0	0.00	0.16	1.85	0.42	15.39
	1	1.02	0.14	2.15	0.43	13.63
	5	5.10	0.10	2.14	0.37	9.67
	10	10.20	0.07	2.15	0.35	6.68
	20	20.40	0.04	1.97	0.32	3.23
	50	51.02	0.01	0.83	1.00	0.62

Note: The table reports the effect of deleting the most central nodes on the network structure of four example beer styles. We remove from the beer styles a sequence of nodes, starting from the one with the highest eigenvector centrality. And we report four network statistics resulting from the deletions. The first column displays the style's name, and the second, how many nodes we removed from its network. Because the style's networks have different sizes, column three shows the percentage of nodes removed from the original network. The remaining columns describe the resulting network structure after the deletions. Density refers to the number of edges out of total possible links. Diameter is the maximum shortest path. Avg. Clust. Coeff. measures the fraction of three-way connections. And Avg. Degree reports, on average, how many links reaches a given node. According to the Molloy-Reed criterion, the network will lose its giant component if the average degree falls below two.

Turning back to our hypothetical where Citra hops go extinct, American IPA would still have many alternative selections with similar traits. And this is why the geography of a style matters. The U.S. produces more than 60 different types of hops, and some of them are very similar to Citra. Brewers often refer to Citra and its sister hops as the 7Cs, which includes Cascade, Centennial, Chinook, Cluster, Columbus and Crystal. All these hops are known for their intense and bright citrus flavour. So, while the Citra hop is a key component of American IPAs, it is also easily replaceable. It is then important to understand the correlation between a style's robustness, the availability of related ingredients, and the overall diversity of inputs used. So, we introduce three new variables to measure these factors.

We start by measuring each style's robustness according to Toth et al. (2022, p.13), who study the co-occurrence of patent classes and define technological robustness as the "amount of node removal that a region's technology network could withstand without being fragment into many unconnected components." Along these lines, like the authors, we use the Molloy and Reed (1995) criterion as the threshold below which the network fragments into many separate pieces. Mathematically the criterion is:

$$\Omega_s = \frac{\sum_{i=1}^{N} k_{is}^2}{\sum_{i=1}^{N} k_{is}}$$

where Ω_s is the resiliency score, or the percentage of nodes removed before the Molloy Reed criterion falls below two, and k_{is} is the average degree. Having defined a measure of network robustness, we now introduce two of its key determinants: related and unrelated variety.

EEG discusses the differences between related and unrelated variety and how these shape the ability of regions to diversify, innovate, and grow (Content and Frenken, 2016; Boschma, 2017; Miguelez and Moreno, 2018; Rocchetta and Mina, 2018). We borrow these concepts to understand how the availability of substitutes for key ingredients shapes the robustness of our style networks. We take related variety to represent the presence of similar substitutes - e.g., Citra or Chinook - while unrelated variety is a style's ability to source from multiple and distinct products - e.g., Pale 2-Row and roasted barley.

We measure unrelated variety according to Frenken et al. (2007) and we apply the Shannon Entropy formula (Shannon, 1948) to the incidence of ingredients as follows:

$$UV_s = \sum_{i=1}^{N} P_{is} \log_2 \left(\frac{1}{P_{is}} \right)$$

where P_{is} is the probability of finding ingredient *i* in beer style *s*. The formula applied to our beer styles captures the level of "uncertainty" or "surprise" across each style's recipes. That is, Shannon Entropy measures the likelihood a recipe includes an unexpected ingredient not commonly found in other beers that belong to the same style. As such, it captures how styles source from distinct ingredients. For an example of a surprising ingredient, think of using a Chocolate malt, typically found in dark and robust Stouts, to make an American IPA.

Frenken et al. (2007) exploit the unique hierarchical structure of employment data to distinguish between related and unrelated variety. However, we cannot apply the same approach to our beer recipes as we cannot separate ingredients into hierarchical structures. Instead, we follow Kogler et al. (2013, 2017) and calculate average relatedness of individual ingredients as a measure of related variety.

We first create a global co-occurrence network covering all beer recipes in our sample regardless of style. The global network follows the same structure as the individual styles described in Section 3. We use this network to measure the similarity or relatedness between each ingredient pair. We measure relatedness by standardizing the elements of the adjacency matrix by the square root of the product of the number of recipes in the row and column ingredients of each element:

$$R_{ij} = \frac{s_{ij}}{\sqrt{N_i * N_j}}$$

where R_{ij} measures the relatedness of each ingredient pair, s_{ij} are the elements of the adjacency matrix and measure how often these two ingredients co-occur (weighted by their proportions), and N_i, N_j are the count of total recipes containing each ingredient. Considering the incidence of ingredients within each style to the sum of their proportions we estimate the style's average relatedness as:

$$AR_s = \frac{\sum_i \sum_j R_{ij}\left(N_i N_j\right) + \sum_i 2N_i}{P_s(P_s - 1)}$$

where P_s is the total count of recipes within each style. Therefore, while unrelated variety measures how much each style sources from various ingredients, average relatedness measures the similarity or compatibility of ingredients used within a style, where we first estimate relatedness using the global sample of recipes. In other words, average relatedness measures the availability of substitutes for every core ingredient used in a given style. For example, two similar hops like Citra and Mosaic have a relatively high average relatedness of 3.69, whereas two distant hops such as Citra and Hallertau have an average relatedness of just 0.21. The same is true of malts as well. The delicious Pale 2-Row and Chocolate example above also has a mercifully low average relatedness of only 0.21. We conclude that if a style uses more similar ingredients, it will have a higher average relatedness and more readily available substitutes.

These EEG metrics allow us to measure the diversity of ingredients within a style, as well as the importance of having substitutes. It is important to note, however, that these variables are not mutually exclusive. A style could have both high levels of average relatedness and unrelated variety. That is to say, a style could simultaneously use many ingredients, each with ample substitutes.

After introducing these measures of network robustness, unrelated variety, and average relatedness, we can observe the interplay between them within recipes of a given style. Figure 6 plots this relationship. Styles with higher levels of both related and unrelated variety tend to be more robust. Taking geography into account, American Styles are more robust than the English, Belgian, or German ones, precisely because of their diverse range of ingredients and easily available substitutes.

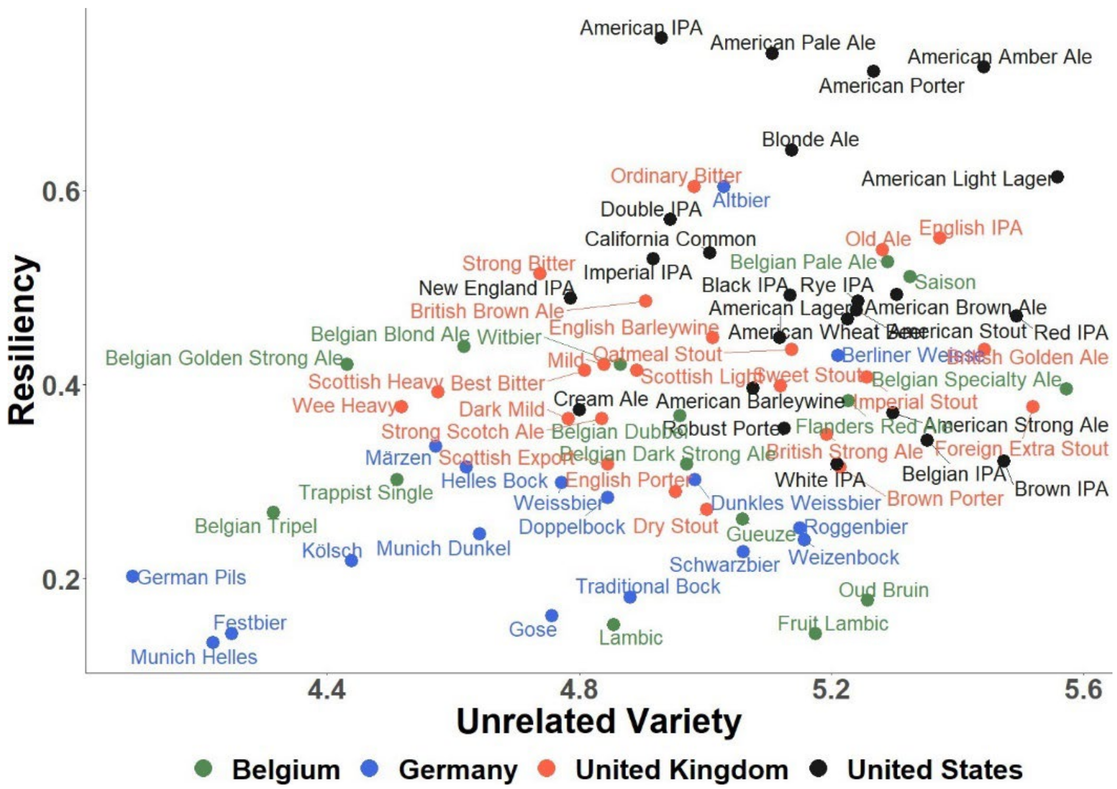


Figure 6A: Resiliency vs. Unrelated Variety

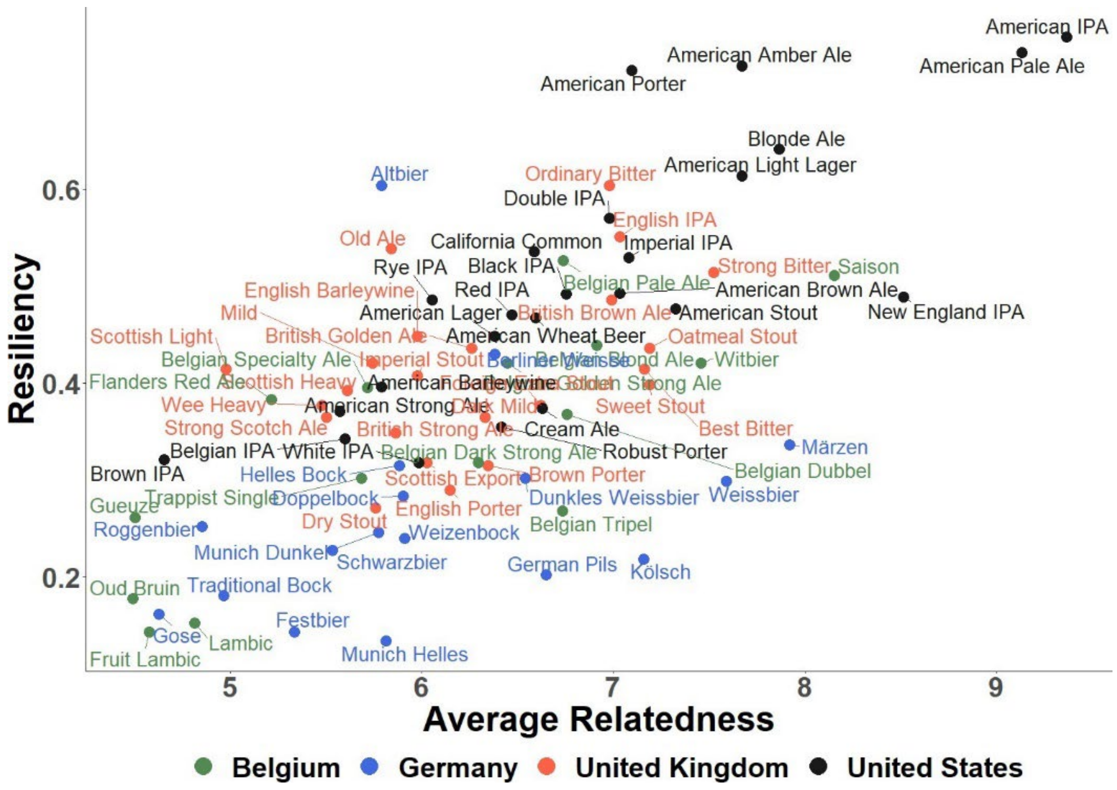


Figure 6B: Resiliency vs. Average Relatedness

Note: Figures 6A & 6B plot the correlation between resiliency and either unrelated variety or average Relatedness. The y-axis is resiliency, which we measure as the percentage of nodes a network can lose before fragmenting into many unconnected components. The x-axis is either unrelated variety or average relatedness. We calculate unrelated variety using the Shannon Entropy formula following (Frenken et al., 2007). We calculate average relatedness following (Kogler et al., 2013). Points are colored according to the country of origin.

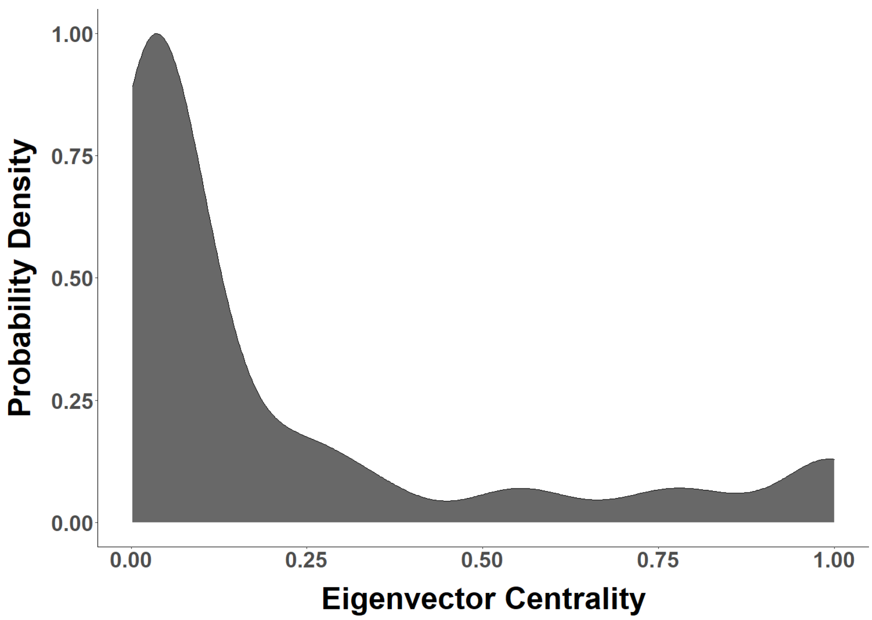
Does Geography Matter?

So far, we have shown that beer styles are highly dependent on a few central keystone ingredients. It remains to be shown that these keystone ingredients differ across styles, otherwise all beers would rely on the same few ingredients. We now turn to demonstrating how central ingredients vary across styles, and that each beer depends on a unique combination of core ingredients. It is these unique combinations that contribute most to a style's network and to its distinctive flavour.

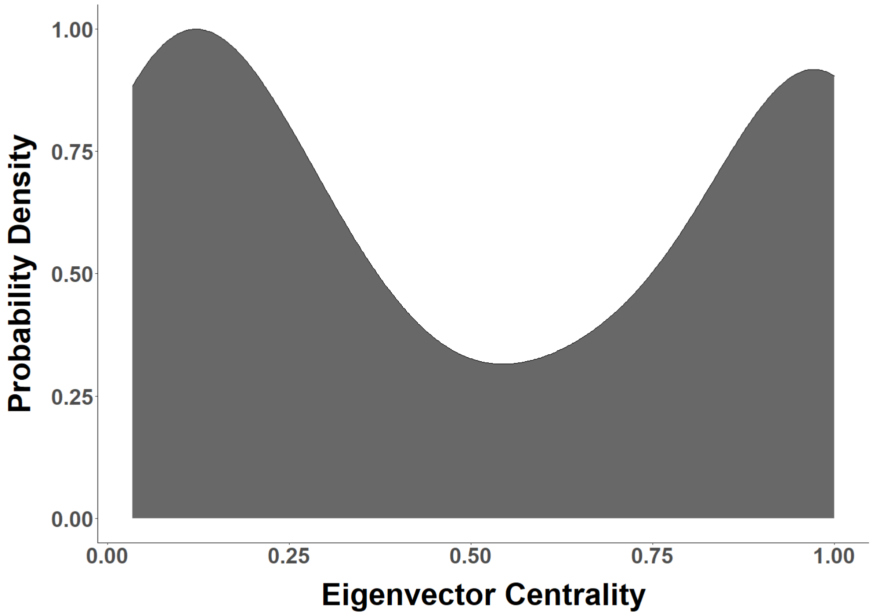
Table 2 highlights that our two sample styles, American IPA and Kölsch, rely on different ingredients with different eigenvector centrality scores. Turning to other style networks, we note how distinct nodes are both highly central to the network and also specific to that style. For example, dark roasted barley is the most central component of Irish Stout, and its most famous variant, Guinness. Dark Munich malt is the most central ingredient for the local Dunkel dark lager. Vienna malt is unsurprisingly the most central ingredient in the Vienna Lager.

Part of what makes these styles so easily identifiable is that their central ingredients are either not used or are of much lower importance in other styles of beer. It is helpful to visualize the distribution of eigenvector centrality for a given key ingredient node across style networks. Figure 7 shows the probability distribution of eigenvector centrality for the two most central nodes in the American IPA and Kölsch: Citra hops and Pilsner malt, respectively. While these distributions are different, they both reveal a bi-modal pattern indicating that while an ingredient may be used in many recipes, it is highly relevant in just a few. Indeed, we find that Citra hops are central components of most American ales but are missing from many European lagers. By contrast, Pilsner is the preferred base malt for many continental lagers from the Bavaria and Bohemia regions but is not as common in English and American ales, which tend to use pale ale malts such as Maris Otter Pale or Pale 2-Row as their base malt.

(A) Citra



(B) Pilsner



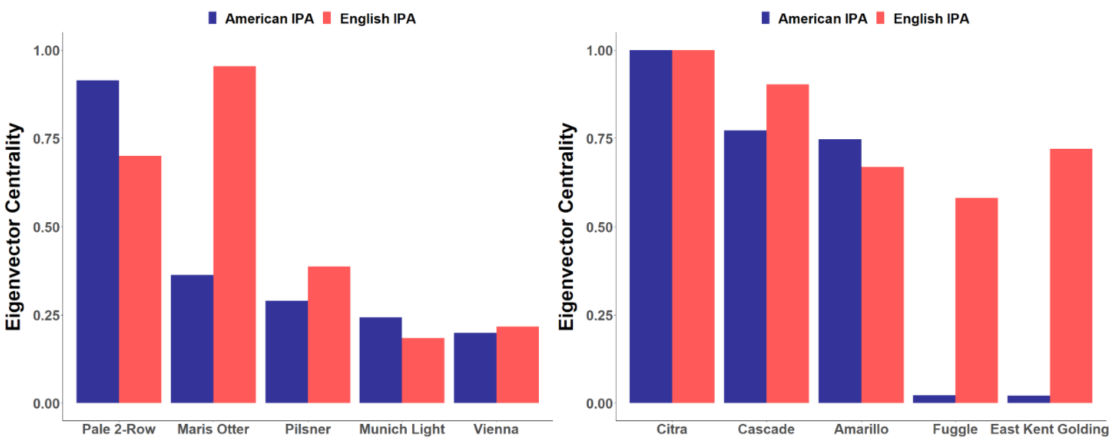
Note: The plots show the probability density function of the eigenvector centrality for two ingredients prevalent in many beer styles: Citra hops and Pilsner malt. The x-axis is the eigenvector centrality of the nodes computed for every style in our sample. The y-axis is the probability density of the centrality score. Both axes are scaled between one and zero such that when an ingredient is the most influential in a network, it will have a centrality score of one. A probability density of one means this is the most frequent centrality score of the ingredient among the beer styles. We measure eigenvector centrality according to Bonacich (1972).

Figure 7: Eigenvector Centrality Distribution across Styles

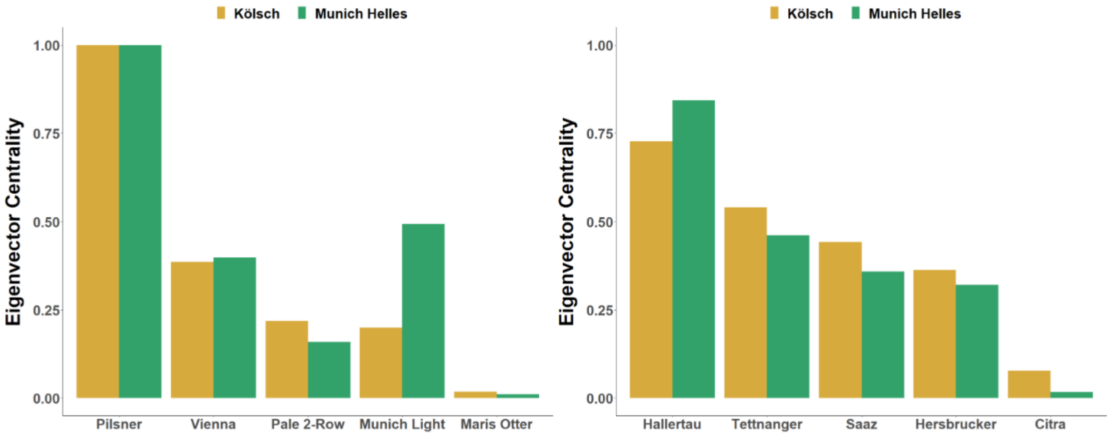
To further understand how geography shapes differences in ingredient centrality it is helpful to think about two examples. Figure 8A shows the eigenvector scores of the top ingredients in two members of the same style family, American and English IPA. Considering just the hops shown in Panel 8A, it is clear that English IPA makes heavy use of American hops. Despite this colonial influence, English IPA also relies heavily on two distinctively English hops, East Kent Golding and Fuggles. These hops are conspicuously absent from American IPA, and their inclusion contributes to English IPA's unique characteristics and flavour. The American hops are bittering hops with high levels of alpha acids and citric flavour, while their English counterparts are mixed purpose hops with fewer alpha acids and are known for their earthy tones (Healey, 2016; BarthHaas, 2018). English IPA also uses Maris Otter malt much more than the American IPA. Looking at the centrality scores of ingredients across the two IPAs, it is easy to see that English IPA has more herbal tones, which captures why the English version “has less hop intensity and more pronounced malt flavours than typical American versions” (BJCP, 2015).

Having considered regional differences in similar styles above, we turn to style differences within the same region. Figure 8B shows the same relationship for two German beers: Munich Helles, a light lager, and Kölsch, a pale ale. Despite belonging to two distinct style families, there is significant overlap in the centrality of their ingredients. Kölsch is the only pale ale brewed in Germany, which makes it distinct from all other beers in the country and unique to Cologne. Yet, compared to ales from other nations, Kölsch uses significantly more of the base malts usually found in German pilsners and lagers. Further, it favours using the German and Czech hops abundant in lagers and known for their aroma, low bitterness, and lightly flowery and spice taste (Healey, 2016; BarthHaas, 2018). These central German nodes contribute to its uniqueness, an ale with pronounced lager traits, which could easily lead the “untrained taster to mistake it for a somewhat subtle Pils” (BJCP, 2015).

(A) American IPA vs. English IPA



(B) Kölsch vs. Munich Helles



Note: These figures plot the eigenvector centrality of the five most central malts and hops in four style networks: American IPA, English IPA, Kölsch, and Munich Helles. Eigenvector centrality measures the importance of each ingredient to a style, which we compute according to Bonacich (1972). Panel a show the comparison between two styles of the same family (IPA) across different countries: the United States and England. Panel b compares the centrality scores for two styles of different families, pale ale and pale lager, within the same country of origin: Germany. We arrange the ingredients in Panel A according to their centrality scores for American IPA. In Panel b, we arrange the ingredients according to their score for Kölsch.

Figure 8: Eigenvector Centrality of Ingredients by Styles

Another way to consider how beer styles differ with geography is to compare similar style networks. Along these lines, we measure the product-moment correlation coefficients between every style adjacency matrix. The correlation coefficient captures how similar the weighted edges between ingredients are across any two styles. Correlation gives us the overlap between style networks where ingredients appear frequently together and combine in similar ways. From our example in Figure 8, we ought to expect a higher correlation coefficient between the two German styles than their American counterpart.

Mathematically, we can express the styles correlation coefficient as:

$$cor(S,S') = \frac{cov(S,S')}{\sqrt{cov(S,S)cov(S',S')}}$$

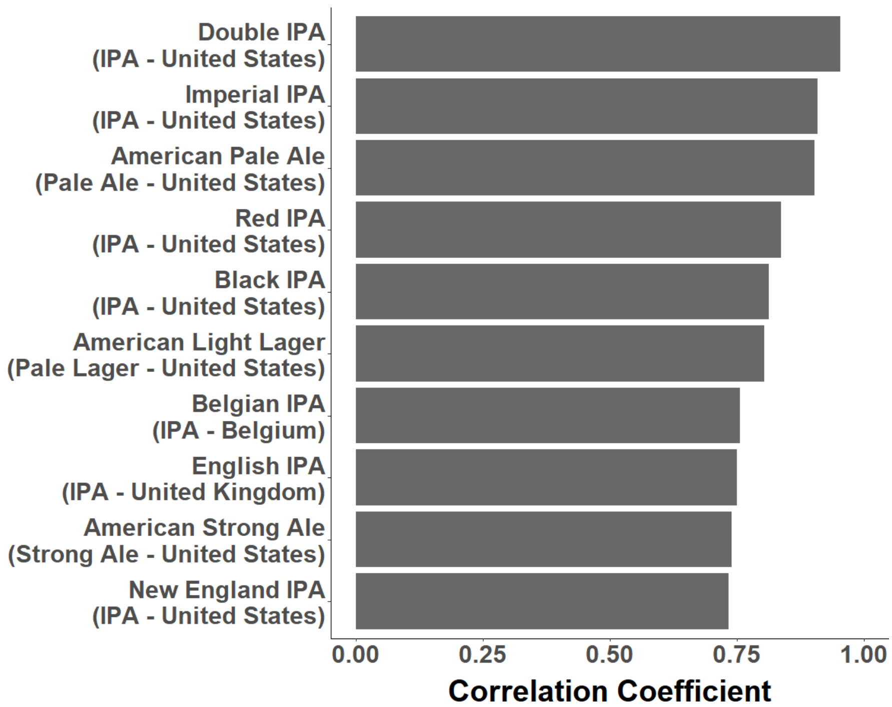
where S and S^0 are two example adjacency matrices, and their covariance is given by:

$$cov(S,S') = \frac{1}{|V_2|}\sum_{i,j}(S_{ij} - \mu_S)(S'_{ij} - \mu'_S)$$

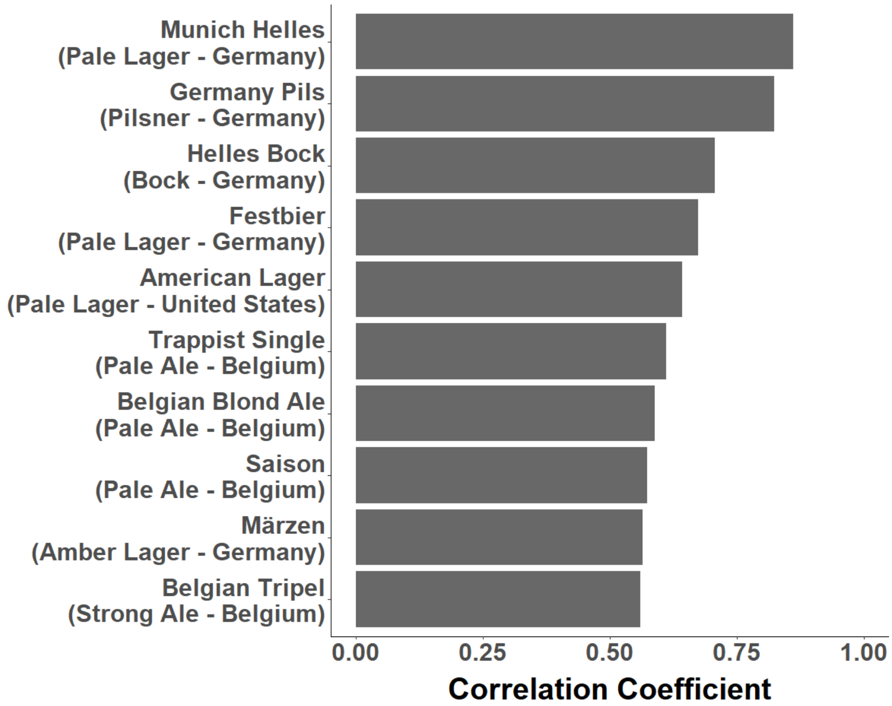
where S_{ij} and S^0_{ij} are the elements within each adjacency matrix, or the weighted edges between the ingredients i and j in both matrices, μ_S and μ^0_S are the average degree, and $|V_2|$ is the variance. If two adjacency matrices have comparable weighted edges between their ingredients, then those styles are similar and will have larger correlation coefficients.

Figure 9 lists the top ten correlations for two networks: American IPA and Kölsch. Perhaps unsurprisingly, we find American IPA to be very similar to other American beers, including the American Light Lager, particularly due to the pronounced use of American hops. Kölsch, on the other hand, is most similar to German and Bohemian lagers, and to a lesser degree to other pale ales from Europe, especially those in Belgium.

(A) American IPA



(B) Kölsch



Note: These plots show the ten styles most similar to American IPA and Kölsch, ranked by correlation coefficient. The x-axis is the correlation coefficient. The y-axis displays the names of the most similar styles, with style family and country of origin in parenthesis. We calculate the coefficients as product-moment correlations between any two styles' adjacency matrices.

Figure 9: Top Ten Similar Styles

Figure 9 highlights that beer recipes and styles are clustered in space. Beer styles are more similar to other styles from the same region, even if those styles belong to very different families. This is true for our American IPA and Kölsch networks, and for other styles in different regions. For example, Bohemian Pilsner is more closely related to its regional neighbour, Czech Pale Lager (correlation coefficient of 0.95), than it is to a beer of its same style, German Pils (correlation coefficient of 0.76). Likewise, Saison, a beer style from French-speaking Wallonia in Belgium, is more similar to other Belgian ales like the Belgian Golden Ale (0.78). In turn, the Belgian Golden Ale is more similar to other styles from Dutch-Speaking Flanders like Belgian Golden Strong Ale (0.91) and Belgian Trippel (0.91).

Therefore, regional ingredients are not only critical to the uniqueness and resilience of a style, but they also transcend style boundaries and link geographically proximate beers together. This makes good sense, as the original brewers primarily had access to local ingredients and made the most with what was available. This lack of variety, be it natural or imposed, as under the German Reinheitsgebot, informed the development of these Classical styles. Even in an era of globalization, these differences persist. New World styles like American IPA benefit from the abundance of ingredients available to them. This results in a large number of ingredients (average relatedness) with a substantial number of ready substitutes (related variety). These factors give New World styles incredible resilience to losing keystone ingredients, as well as the flexibility to adapt and embrace new ones. This adaptability explains the extreme popularity of these styles and why so many brewers are drawn to them.

Concluding Remarks

We bring new data and adapt EEG methods to the debate on beer and place. Our goal is to extend the regional studies literature, validate the EEG models, and provide further evidence on the complex relationship between knowledge, products, and space. Focusing on a single final product that harbours distinct regional styles, our analysis stresses the significant diversity of beers - or the within-product differences. It also brings valuable insights regarding how geography and history shape these remarkable differences. Along these lines, we show that every beer style relies on a few unique, local ingredients. We measure that styles originating in close physical proximity are more similar among them than those sharing family ties. We discuss the availability of substitutes and how beer styles differ regarding how they source from various often unrelated inputs. And finally, we reasoned about the connection between resources and network robustness and its apparent regional contrasts.

Altogether, these results help us compose a story about the regionalization and differentiation of products. They tell a story about the local access to resources and how these transform goods in space. We believe this story will interest the reader. They will strengthen future research on specialization or the evolution of regional knowledge beyond the established related diversification examples. Writing about international trade, Schott (2004) underlines the endowment-driven within-product specialization and its consequences for firms, workers, productivity, and growth. And we feel this is a natural path forward for the EEG literature. Thus, we expect other studies will follow. The power of our analysis comes from our data and case study - we collect detailed information on every component inside a product with strong regional ties. But our results remain exploratory at best. To push forward the literature, we need more studies that can make robust assertions on the role of local knowledge in the development and differentiation of products.

Throughout this chapter, we used unorthodox interpretations of the EEG tools, which might take some readers aback. Still, we imagine these definitions are not necessarily a bad thing. Only by doing so can we extend the EEG methods to new topics and questions. These reimaginations can thus provide context for the limitations and strengths of our models. They can shine on the hidden assumptions and implications behind our theories. Consider, for example, the concepts of related and unrelated variety. For sure, Franken et al. (2007) didn't have beer ingredients in mind when writing their influential piece. Yet, because we reinterpret these concepts as such, we could establish a link between geography, the uses of diverse resources, and network robustness. A line of research we feel could enhance the existing literature on regional resiliency. Besides, we expect our unorthodox application to aid others in understanding/interpreting what variety means. Empirically, one can employ the entropy formula to measure the diversity of species, disciplines, opinions, and so much more. In fact, Claude Shannon (1971, p.180) recounts that von Neumann convinced him to call it entropy because "no one knows what entropy really is, so in a debate, you will always have the advantage." As such, the authors admit they finally understood the principle behind the formula, and the concept of unrelated variety for that matter, when applying it to their favourite beer recipes. For us, entropy now means the delicious and unexpected combination of seemingly incompatible ingredients within a glass of beer. Or even this chapter we are writing, an unusual combination of network science, regions, and beer.

To conclude, we are the first to collect and disambiguate a comprehensive set of beer recipe data, which we hope others will build on. Not only can this beer data answer other longstanding questions in the geography of beer literature, but the highly detailed ingredient information can also be seen as data on intermediate goods used to produce a final product. Because of this, we are able to bring an existing methodology to a new area of inquiry. We hope our application of regional studies methods to new data and fields of study inspires others to do the same. We quantify the benefits of styles having an abundance of ingredients and substitutes in their regions. This conclusion is a sensible one and is by no means specific to beer alone. Especially in today's ever more connected world, embracing the abundance and diversity that globalization offers is useful for everyone, brewers included.

References

Abbasi, A., Altmann, J., and Hossain, L. (2011). Identifying the Effects of Co-authorship Networks on the Performance of Scholars: A correlation and regression analysis of performance measures and social network analysis measures. *Journal of Informetrics*, 5(4):594–607.

Albert, R., Jeong, H., and Barabási, A.-L. (2000). Error and Attack Tolerance of Complex Networks. *Nature*, 406(6794):378–382.

Amiti, M. and Khandelwal, A. K. (2013). Import Competition and Quality Upgrading. *Review of Economics and Statistics*, 95(2):476–490.

Barabási, A.-L. et al. (2016). *Network Science.* Cambridge University Press.

Barber, M. J. and Scherngell, T. (2013). Is the European R&D Network Homogeneous? Distinguishing Relevant Network Communities Using Graph Theoretic and Spatial Interaction Modelling Approaches. *Regional Studies*, 47(8):1283–1298.

BarthHaas (2018). Hop Harvest Guide 2018.

BJCP (2015). Beer Judge Certification Program 2015 Style Guidelines.

Bonacich, P. (1972). Factoring and Weighting Approaches to Status Scores and Clique Identification. Journal of Mathematical Sociology, 2(1):113–120.

Boschma, R. (2017). Relatedness as Driver of Regional Diversification: A Re- search Agenda. Regional Studies, 51(3):351–364.

Boschma, R. A. and Frenken, K. (2006). Why Is Economic Geography Not an Evolutionary Science? Towards an Evolutionary Economic Geography. *Journal of Economic Geography*, 6(3):273–302.

Buarque, B. S., Davies, R. B., Hynes, R. M., and Kogler, D. F. (2020). OK Computer: The Creation and Integration of AI in Europe. *Cambridge Journal of Regions, Economy and Society*, 13(1):175–192.

Cabras, I. and Bamforth, C. (2016). From reviving tradition to fostering innovation and changing marketing: The evolution of micro-brewing in the UK and US, 1980–2012. *Business History*, 58(5):625–646.

Content, J. and Frenken, K. (2016). Related Variety and Economic Development: A Literature Review. *European Planning Studies*, 24(12):2097–2112.

Cottineau, C., Finance, O., Hatna, E., Arcaute, E., and Batty, M. (2019). Defining Urban Clusters to Detect Agglomeration Economies. *Environment and Planning B: Urban Analytics and City Science*, 46(9):1611–1626.

Crespi, G., Fernández-Arias, E., and Stein, E. (2014). *Rethinking Productive Development*, pages 3–31. Palgrave Macmillan US, New York.

del Rio-Chanona, R. M., Mealy, P., Beguerisse-Díaz, M., Lafond, F., and Farmer, J. D. (2021). Occupational Mobility and Automation: A Data-driven Network Model. *Journal of the Royal Society Interface*, 18(174):20200898.

del Rio-Chanona, R. M., Mealy, P., Pichler, A., Lafond, F., and Farmer, J. D. (2020). Supply and Demand Shocks in the COVID-19 Pandemic: An Industry and Occupation Perspective. *Oxford Review of Economic Policy*, 36 (Supplement 1): S94–S137.

Dennett, A. and Page, S. (2017). The Geography of London's Recent Beer Brewing Revolution. *The Geographical Journal*, 183(4):440–454.

Dinnie, K. (2004). Country-of-Origin 1965-2004: A Literature Review. *Journal of Customer Behaviour*, 3(2):165–213.

Duan, W., Madasi, J. D., Khurshid, A., and Ma, D. (2022). Industrial Structure Conditions Economic Resilience. *Technological Forecasting and Social Change*, 183:121944.

Elzinga, K. G., Tremblay, C. H., and Tremblay, V. J. (2015). Craft Beer in the United States: History, Numbers, and Geography. *Journal of Wine Eco- nomics*, 10(3):242.

Fernandes, A. M. and Paunov, C. (2013). Does Trade Stimulate Product Quality Upgrading? *Canadian Journal of Economics*, 46(4):1232–1264.

Fowler, C. S. and Jensen, L. (2020). Bridging the Gap Between Geographic Concept and the Data We Have: The Case of Labor Markets in the USA. *Environment and Planning A: Economy and Space*, 52(7):1395–1414.

Frenken, K., Van Oort, F., and Verburg, T. (2007). Related Variety, Unrelated Variety and Regional Economic Growth. *Regional Studies*, 41(5):685–697.

Gatrell, J., Reid, N., and Steiger, T. L. (2018). Branding Spaces: Place, Region, Sustainability and the American Craft Beer Industry. *Applied Geography*, 90:360–370.

Griliches, Z. (2007). *R&D and Productivity: The econometric evidence*. University of Chicago Press.

Haldane, A. G. and May, R. M. (2011). Systemic Risk in Banking Ecosystems. Nature, 469(7330):351–355.

Hawelka, B., Sitko, I., Beinat, E., Sobolevsky, S., Kazakopoulos, P., and Ratti, C. (2014). Geo-located Twitter as Proxy for Global Mobility Patterns. Cartography and Geographic Information Science, 41(3):260–271.

Healey, J. (2016). *The Hops List: 265 Beer Hops Varieties from Around the World*. Hoplist.com.

Henn, C., Papageorgiou, C., Romero, J. M., and Spatafora, N. (2020). Export Quality in Advanced and Developing Economies: Evidence from a New Data Set. *IMF Economic Review*, 68(2):421–451.

Hidalgo, C. A., Balland, P.-A., Boschma, R., Delgado, M., Feldman, M., Frenken, K., Glaeser, E., He, C., Kogler, D. F., Morrison, A., Neffke, F., Rigby, D., Stern, S., Zheng, S., and Zhu, S. (2018). The Principle of Relatedness. In Morales, A. J., Gershenson, C., Braha, D., Minai, A., and Bar-Yam, Y., editors, *Unifying Themes in Complex Systems IX. ICCS 2018*, pages 451–457. Springer International Publishing.

Hummels, D. and Klenow, P. J. (2005). The Variety and Quality of a Nation's Exports. *American Economic Review*, 95(3):704–723.

Jaffe, A. B. and De Rassenfosse, G. (2019). Patent Citation Data in Social Science Research: Overview and Best Practices. In B. Depoorter, P. Menell and D. Schwartz, editor, *Research Handbook on the Economics of Intellectual Property Law*, chapter 2, pages 20–46. Edward Elgar Publishing.

Juhász, S., Broekel, T., and Boschma, R. (2021). Explaining the Dynamics of Relatedness: The role of co-location and complexity. *Papers in Regional Science*, 100(1):3–21.

Kamada, T., Kawai, S., et al. (1989). An Algorithm for Drawing General Undirected Graphs. *Information Processing Letters*, 31(1):7–15.

Kang, B. and Tarasconi, G. (2016). PATSTAT Revisited: Suggestions for Better Usage. *World Patent Information*, 46:56–63.

Kind, C. and Kaiser, T. (2020). Heat, Hops, Hallertau: Exploring Implications of Climate Change for the German Beer Sector. In *The Geography of Beer*, pages 103–111. Springer.

Klement, B. and Strambach, S. (2019). How do New Music Genres Emerge? Diversification Processes in Symbolic Knowledge Bases. *Regional Studies*, 53(10):1447–1458.

Knudson, W., Sirrine, J. R., and Mann, J. T. (2020). The Geographic Dispersion of Hop Production in the United States: Back to the Future? In *The Geography of Beer*, pages 113–120. Springer.

Knuepling, L. and Broekel, T. (2022). Does Relatedness Drive the Diversification of Countries' Success in Sports? *European Sport Management Quarterly*, 22(2):182–204.

Kogler, D. F. (2015). Editorial: Evolutionary Economic Geography – Theoretical and Empirical Progress. *Regional Studies*, 49(5):705–711.

Kogler, D. F., Essletzbichler, J., and Rigby, D. L. (2017). The Evolution of Specialization in the EU15 Knowledge Space. *Journal of Economic Geography*, 17(2):345–373.

Kogler, D. F., Rigby, D. L., and Tucker, I. (2013). Mapping Knowledge Space and Technological Relatedness in US Cities. *European Planning Studies*, 21(9):1374–1391.

Lentz, P., Holzmüller, H. H., and Schirrmann, E. (2006). City-of-Origin Effects in the German Beer Market: Transferring an International Construct to a Local Context. In *International Marketing Research (Advances in International Marketing, Vol. 17)*, pages 251–274. Emerald Group Publishing Limited, Bingley.

Leydesdorff, L., Kogler, D. F., and Yan, B. (2017). Mapping Patent Classifications: Portfolio and statistical analysis, and the comparison of strengths and weaknesses. *Scientometrics*, 112(3):1573–1591.

Li, X., Liu, Y., Jiang, Y., and Liu, X. (2016). Identifying Social Influence in Complex Networks: A Novel Conductance Eigenvector Centrality Model. *Neurocomputing*, 210:141–154.

Martin, R., Sunley, P., Gardiner, B., and Tyler, P. (2016). How Regions React to Recessions: Resilience and the Role of Economic Structure. *Regional Studies*, 50(4):561–585.

Metcalfe, J. S., Foster, J., and Ramlogan, R. (2006). Adaptive Economic Growth. *Cambridge Journal of Economics*, 30(1):7–32.

Miguelez, E. and Moreno, R. (2018). Relatedness, External Linkages and Regional Innovation in Europe. *Regional Studies*, 52(5):688–701.

Mittag, R. (2014). Geographic Appellations of Beer. In *The Geography of Beer*, pages 67–74. Springer.

Molloy, M. and Reed, B. (1995). A Critical Point for Random Graphs with a Given Degree Sequence. *Random Structures & Algorithms*, 6(2-3):161–180.

Montoya, J. M., Pimm, S. L., and Solé, R. V. (2006). Ecological Networks and Their Fragility. *Nature*, 442(7100):259–264.

Moro, E., Frank, M. R., Pentland, A., Rutherford, A., Cebrian, M., and Rahwan, I. (2021). Universal Resilience Patterns in Labor Markets. Nature Communications, 12(1):1–8.

Neffke, F. (2009). *Productive Places: The influence of technological change and relatedness on agglomeration externalities*. PhD thesis, Utrecht University.

Newman, M. E. (2005). Power Laws, Pareto Distributions and Zipf's law. *Con- temporary Physics*, 46(5):323–351.

Nooteboom, B. (2000). Learning by Interaction: Absorptive Capacity, Cognitive Distance and Governance. *Journal of Management and Governance*, 4(1- 2):69–92.

Parand, F.-A., Rahimi, H., and Gorzin, M. (2016). Combining Fuzzy Logic and Eigenvector Centrality Measure in Social Network Analysis. *Physica A: Statistical Mechanics and its Applications*, 459:24–31.

Parr, J. B. (2007). Spatial Definitions of the City: Four Perspectives. *Urban studies*, 44(2):381–392.

Patterson, M. and Hoalst-Pullen, N. (2014). *The Geography of Beer: Regions, Environment, and Societies*. Springer, New York.

Pei, T., Sobolevsky, S., Ratti, C., Shaw, S.-L., Li, T., and Zhou, C. (2014). A New Insight into Land Use Classification Based on Aggregated Mobile Phone Data. *International Journal of Geographical Information Science*, 28(9):1988– 2007.

Piccardi, C. and Tajoli, L. (2012). Existence and Significance of Communities in the World Trade Web. *Physical Review E*, 85(6):066119.

Radosevic, S. (2017). Assessing EU Smart Specialization Policy in a Comparative Perspective. In Radosevic, S., Curaj, A., Gheorghiu, R., Andreescu, L., and Wade, I., editors, *Advances in the Theory and Practice of Smart Specialization*, pages 1–36. Academic Press.

Ratti, C., Sobolevsky, S., Calabrese, F., Andris, C., Reades, J., Martino, M., Claxton, R., and Strogatz, S. H. (2010). Redrawing the Map of Great Britain from a Network of Human Interactions. *PloS One*, 5(12): e14248.

Rocchetta, S. and Mina, A. (2019). Technological Coherence and the Adaptive Resilience of Regional Economies. *Regional Studies*, 53(10):1421–1434.

Ruhnau, B. (2000). Eigenvector-centrality a Node-centrality? *Social Networks*, 22(4):357–365.

Schott, P. K. (2004). Across-product Versus Within-product Specialization in International Trade. *The Quarterly Journal of Economics*, 119(2):647–678.

Sewell, S. L. (2014). The Spatial Diffusion of Beer from its Sumerian Origins to Today. In *The Geography of Beer*, pages 23–29. Springer.

Shannon, C. E. (1948). A Mathematical Theory of Communication. *The Bell System Technical Journal*, 27(3):379–423.

Sobolevsky, S., Sitko, I., Des Combes, R. T., Hawelka, B., Arias, J. M., and Ratti, C. (2014). Money on the Move: Big Data of Bank Card Transactions as the New Proxy for Human Mobility Patterns and Regional Delineation. The Case of Residents and Foreign Visitors in Spain. In *2014 IEEE International Congress on Big Data*, pages 136–143. IEEE.

Tantardini, M., Ieva, F., Tajoli, L., and Piccardi, C. (2019). Comparing Methods for Comparing Networks. *Scientific Reports*, 9(1):1–19.

Thiemann, C., Theis, F., Grady, D., Brune, R., and Brockmann, D. (2010). The Structure of Borders in a Small World. *PloS One*, 5(11): e15422.

Toth, G., Elekes, Z., Whittle, A., Lee, C., and Kogler, D. F. (2022). Technology Network Structure Conditions the Economic Resilience of Regions. *Economic Geography*, 98(4): 355-378.

Verlegh, P. W. and Steenkamp, J.-B. E. (1999). A Review and Meta-Analysis of Country-of-Origin Research. *Journal of Economic Psychology*, 20(5):521–546.

Wasserman, S., Faust, K., et al. (1994). *Social Network Analysis: Methods and Applications*, volume 8. Cambridge University Press.

Whittle, A. and Kogler, D. F. (2020). Related to What? Reviewing the Literature on Technological Relatedness: Where We are Now and Where Can We go? *Papers in Regional Science*, 99(1):97–113.

Wojtyra, B., Kossowski, T. M., Březinová, M., Savov, R., and Lančarič, D. (2020). Geography of Craft Breweries in Central Europe: Location Factors and the Spatial Dependence Effect. *Applied Geography*, 124:102325.

Yool, S. and Comrie, A. (2014). A Taste of Place: Environmental Geographies of the Classic Beer Styles. In *The Geography of Beer*, pages 99–108. Springer.

Acknowledgments

We thank participants at the UCD/UCLA Seminars in Economic Geography Series for useful comments. All errors are our own.

About the Authors

Ryan Hynes is a software engineer, attorney, and researcher working at the intersection of data, law, and technology. As an academic, he creates complex datasets to answer new questions on trade, migration, taxation, and more. As an engineer, he builds tools to help professionals find the most relevant information in an ever-growing haystack of documents. Ryan earned his PhD in Economics, specializing in trade and innovation, from University College Dublin.

Bernardo S. Buarque is a Research Scholar at the Max Planck Institute for the History of Science. He joined the institute after earning his PhD from the University College Dublin. His research focuses on complex networks and dynamic systems modeling, particularly in the context of innovation and knowledge diffusion.

Professor **Ron Davies** has been a Professor of Economics at University College Dublin since 2008. After earning his Ph. D. at the Pennsylvania State University in 1999, he joined the University of Oregon faculty where he remained until relocating to Dublin. Specializing in trade and foreign direct investment, he has a particular focus on the tax and trade policies that affect globalization.

Dieter F. Kogler is Prof. of Economic Geography and the Academic Director of the Spatial Dynamics Lab at University College Dublin. His research focus is on the geography of innovation and evolutionary economic geography, with particular emphasis on knowledge production and diffusion, and processes related to technological change, innovation, and economic growth. Dieter’s career path combines professional and research experience acquired in Europe, the United States, and Canada within a variety of areas pertaining to the spatial analysis of socio-economic phenomena.