Short-Term Rental and Urban Change : An Empirical Systemic Approach

Calum Robertson

LA ROCHELLE UNIVERSITÉ

ÉCOLE DOCTORALE EUCLIDE - 618

LABORATOIRE : Usages du Numérique pour le Développement Durable - NUDD

THÈSE présentée par :

Calum ROBERTSON

soutenue le : **11 juin 2024**

pour obtenir le grade de : **Docteur de La Rochelle Université**

Discipline : **Sciences Économiques**

Short-Term Rental and Urban Change: An Empirical Systemic Approach

---

JURY :

Éric DARMON	Professeur, Université Paris Nanterre	Président
Pierre-Alexandre BALLAND	Professeur, Utrecht University	Rapporteur
Julie LE GALLO	Professeure, L'Institut Agro Dijon	Rapporteur
Francesca ARTIOLI	Maîtresse de conférences, EUP Université Paris-Est Créteil	Examinatrice
Sylvain DEJEAN	Maître de conférences, HDR,La Rochelle Université	Co-directeur
Raphaël SUIRE	Professeur, Nantes Université	Co-directeur

Sous la codirection de :

Sylvain DEJEAN, Maître de conférences, HDR, La Rochelle Université
Raphaël SUIRE, Professeur, Nantes Université

Laboratoire :

NUDD

Usages du Numérique pour le Développement Durable

23 Avenue Albert Einstein
17000 La Rochelle

Financement :

Région Nouvelle-Aquitaine

AquiAirbnb

Remerciements

"*A PhD is something you do because there is nothing else you want to do more than learn about some aspect of how the world works*" — Ian Bearden. De toutes les raisons qui inspirent à s'engager dans une thèse, c'est la curiosité qui l'emporte. Tout au long de ce parcours, j'ai eu la chance d'être accompagné par des personnes dont la bienveillance ainsi que le discernement m'ont permis de cultiver cette curiosité. Je tiens à leur exprimer ma gratitude.

Une profonde reconnaissance s'adresse aux encadrants de cette thèse, Sylvain Dejean et Raphaël Suire. La complémentarité de leurs compétences et l'habileté de leur intuition scientifique en ont fait un binôme extrêmement instructif sur le plan technique. À cela s'ajoutent une pédagogie et des qualités humaines remarquables. Leur disponibilité, leur écoute, leur patience, leur rigueur et leurs encouragements ont façonné un environnement particulièrement propice à l'apprentissage du métier de chercheur. Ils ont su tant appuyer ma curiosité que l'encadrer et ont fait de ces années une expérience enrichissante sur tous les plans. J'espère avoir la chance de les côtoyer de nouveau à l'avenir.

Mes remerciements vont aussi aux rapporteurs de cette thèse, Pierre-Alexandre Balland et Julie Le Gallo, ainsi qu'à ses examinateurs, Francesca Artioli et Éric Darmon. La qualité de leur recherche ne laisse aucun doute sur la pertinence des remarques qu'ils feront vis-à-vis de ce travail. Merci également à Stéphanie Peltier et (de nouveau) à Éric Darmon pour avoir été membres de mon comité de suivi individuel. Leurs retours fins et constructifs m'ont permis de prendre du recul non seulement sur le travail dans son ensemble, mais aussi sur ma situation de doctorant.

Je remercie la région Nouvelle-Aquitaine, commanditaire du projet *AquiAirbnb* et financeuse de ce doctorat, ainsi que Fabian Candeau et Lauriane Belloy avec qui ce projet a été partagé. La communauté d'agglomération de La Rochelle, soutenante du projet, a aussi favorisé un échange dont je suis reconnaissant avec des acteurs institutionnels directement concernés. J'adresse également des remerciements sincères aux chercheurs et aux acteurs administratifs de La Rochelle Université, du laboratoire NUDD et de l'école doctorale EUCLIDE avec qui j'ai pu avoir la chance d'interagir.

Ce doctorat a été rythmé par des déplacements des plus enrichissants. Ils en sont deux qui

l'ont particulièrement marqué. Merci à l'équipe d'économie géographique d'Utrecht University qui a proposé une école d'été mémorable et ponctuée d'excellentes interventions. Celle-ci a sans aucun doute eu un impact fort sur la suite de ma thèse. Merci à David Wachsmuth, à l'UPGo, à l'Université McGill et à ses étudiants qui m'ont chaleureusement accueilli pendant deux mois dans un environnement stimulant.

Je tiens à remercier toutes et tous les docteurs et doctorants avec qui j'ai pu échanger tout au long de ces trois années, à la fois en conférence et autour de terrasses de cafés. Ces échanges m'ont été précieux.

Un immense merci à ma mère et à mon père. Vous m'avez toujours encouragé avec bienveillance ma curiosité, nourri mon désir d'apprendre et m'avez permis d'élargir mes horizons. À ma sœur et à mon frère, votre présence d'esprit, votre humour et votre sympathie m'ont été et me resteront essentiels.

À mes amis rencontrés au lycée, nos conversations ont illuminé les périodes les plus intenses de mon adolescence, et je suis très reconnaissant qu'elles aient également animé celles de ma thèse. À ceux rencontrés à la faculté, en licence comme en master, merci d'avoir enrichi mon expérience universitaire de tant de façons. À tous, je mesure la chance qui est la mienne de vous savoir à mes côtés.

Enfin, à Louise. Je suis reconnaissant et heureux de la complicité qui nous lie. Ta patience à toute épreuve, ton soutien, ton intelligence et ton humour ont fait pour moi de ces années de doctorat, comme des autres, un plaisir. J'espère que nous continuerons à partager ces moments qui me sont chers.

Résumé

Cette thèse s'appuie sur un cadre de pensée évolutionniste et interdisciplinaire inspiré de travaux en économie urbaine, en économie géographique et en aménagement urbain pour appréhender les transformations urbaines induites par les locations courte durée (LCD). Portées par l'essor de plateformes numériques telles qu'*Airbnb*, les LCD ont pris une place importante dans le paysage urbain ces dernières années. Ayant facilité le tourisme, parfois dans des zones qui en étaient jusqu'alors relativement dénuées, les LCD suscitent de vives réactions à la fois citoyennes et institutionnelles. De nombreuses villes s'attellent donc à réguler leur développement. L'appréhension des transformations urbaines induites par les LCD et la conceptualisation des régulations sont cependant rendues difficiles par le caractère systémique des villes et de leurs quartiers. Ce travail adopte une vision des espaces urbains comme des systèmes complexes spatialisés. Au sein de ceux-ci, les éléments, les acteurs et leurs interactions innombrables et rétroactives contribuent continuellement à faire évoluer leurs situations mutuelles. En s'implantant dans ces systèmes complexes, les LCD affectent nécessairement les acteurs de la ville et leurs interactions directement ou indirectement. Elles mènent à des transformations qui sont difficiles à quantifier dans leur globalité étant donné la diversité des composantes de la ville et la non-linéarité des interactions qu'elle héberge.

Le travail est divisé en quatre chapitres. Le premier chapitre[1] évalue quantitativement l'efficacité de la régulation des LCD Bordelaise à l'aide de méthodes d'inférence causale et la confronte à ses motivations telles qu'explicitement déclarées par le conseil municipal.

[1]qui est une itération d'un article copublié avec Sylvain DEJEAN et Raphaël SUIRE : https://doi.org/10.1007/s00168-023-01215-4

Il en ressort que comme escompté, la régulation Bordelaise a induit une baisse du nombre de nuitées et d'annonces. Cependant, elle n'a pas complètement épargné les loueurs qui n'étaient pas visés et a des effets spatialement différenciés. De plus, il est difficile de lier les niveaux d'activité de LCD directement à certaines motivations de la régulation. Il en résulte qu'une approche systémique placée sous l'égide du changement urbain pourrait contribuer à mieux comprendre le phénomène des LCD et à conceptualiser sa régulation de manière plus intelligente.

Le deuxième chapitre explore les villes en tant que systèmes et s'inspire du paradigme évolutionniste pour les formaliser et les caractériser. Il postule que la distribution spatiale des commerces et services à buts lucratifs dans les différents quartiers d'une ville dépendent d'un procédé de sélection qui est conditionné à la fois par les consommateurs et par les environnements que ceux-ci affectionnent. De ce fait, la présence d'établissements commerciaux dans un espace donné reflète non seulement les préférences des consommateurs qui fréquentent cet espace, mais reflète aussi holistiquement les caractéristiques de cet espace. Une branche particulière de l'économie géographique évolutionniste s'intéresse à employer des méthodes empiriques pour caractériser les systèmes complexes à travers leur activité économique. Une méthodologie est déployée pour adapter l'*Economic Complexity Index* et le *Product Complexity Index* au milieu urbain via le prisme des établissements commerciaux. L'*Amenity Complexity Index* (ACI) et l'*Amenity-Type Complexity Index* (TCI) qui en découlent sont des réductions du réseau de commerces dans les quartiers d'une ville. Elles sont respectivement des caractérisations unidimensionnelles des sous-systèmes spatiaux urbains et des types de commerces. Il est avancé qu'en séparant les types de commerces via le TCI, les consommateurs de ces types de commerces (tels qu'anticipés par les commerçants) sont différenciés. Parallèlement, l'ACI sépare les espaces à travers les consommateurs qui les fréquentent (tels qu'anticipés par les commerçants). Le chapitre, en proposant ces indicateurs, avance une perception du changement urbain via les compositions locales de demande au sein de sous-systèmes complexes, et donne l'opportunité d'explorer les LCD sous ce prisme.

Le troisième chapitre est une illustration empirique des indicateurs développés dans le chapitre 2 sur la ville de Paris et à l'aide d'un jeu de données compréhensif. Il met en évidence le fait que l'ACI et que le TCI peuvent contribuer à la compréhension de la formation de systèmes urbains complexes, de leur dépendance historique, de leur propension au changement, de leur résilience, ou encore de leurs propriétés. Les séparations induites par les indicateurs étant unidimensionnelles, le chapitre 3 émet également des hypothèses micro-fondées sur la dépense des consommateurs comme caractéristique dominante derrière la séparation spatiale de ce qu'ils consomment — Non seulement au sens commercial strict, mais également au sens large de l'environnement urbain. Les indicateurs sont présentés comme des témoins utiles, bien qu'imparfaits et ne reflétant pas des équilibres stables, des situations urbaines et de leur changement.

Le quatrième et dernier chapitre étend l'utilisation des indicateurs à l'échelle nationale. Il établit leur robustesse à un changement de jeu de données ainsi qu'aux changements méthodologiques qui doivent l'accompagner. Il exploite ensuite les rangs TCI des aménités commerciales au sein des différentes villes pour évaluer les différentes répartitions spatiales de la demande. Il explore la convergence entre les rangs de TCI dans des paires de villes en fonction de l'intensité de l'activité de LCD qu'elles hébergent. Il suggère empiriquement le TCI en tant qu'instrument utile pour déceler un potentiel phénomène d'homogénéisation des espaces urbains induit par les LCD.

Il découle de ce travail dans son ensemble qu'un point de vue systémique et évolutionniste peut s'articuler avec une micro-fondation économique des phénomènes urbains pour construire des instruments empiriques précieux. Ces derniers peuvent contribuer à une compréhension d'ensemble des changements urbains et de phénomènes comme les LCD qui les sous-tendent, à une politique urbaine plus adaptée, et à concevoir des régulations intelligentes.

Abstract

Short-term rentals (STRs), propelled by online platforms like *Airbnb*, have taken over the tourism accommodation industry. In doing so, they have faced strong opposition from residents and policy-makers because of their impacts on urban spaces. This has implied a couple of research challenges. First, what impact does STR have on cities? Second, how can regulation be designed accordingly? The first chapter of this thesis leverages causal inference techniques to empirically evaluate the efficacy of Bordeaux's STR regulation, and it outlines the need for a more holistic approach of STR impacts. The rest of this work argues that STRs modify cities and impact lived experiences in a systemic way that cannot be isolated to single outcomes, and that is a part of a broader process of change. Thus, a perspective of cities as complex systems is fitting. The second chapter draws on evolutionary economic geography and urban planning to shape this perspective. Critically, it leverages retail locations as reflective of urban systemic situations through market mechanisms, and provides a methodology to characterize these situations empirically. The economic complexity literature has supplied valuable ways of characterizing systems, most notably through their economic activities, and its measures are adapted here to fit into a framework of urban systemic change. The *Amenity Complexity Index* (ACI) and *Amenity-Type Complexity Index* (TCI) are at the heart of this work. They separate sub-city level urban spaces and types of commercial amenities one-dimensionally. Chapters 3 empirically consolidates and micro-founds their interpretations as reflective of spending-driven compositions of demand. It explores how they can be used to assess urban change under a new light. Chapter 4 tests for the robustness of the indicators against new data and at a broader spatio-temporal scale, and applies them to the specific issue of STR. It empirically suggests that STR are homogenizing the spatial distribution of retail supply across major French cities. All together, this thesis demonstrates that a holistic empirical approach to urban systems and their change is a valuable addition to current understandings of STR-driven transformations and regulatory efforts.

Keywords: Urban transformations, Short-term rentals, Complex systems, Economic complexity, Evolutionism, Amenities

Contents

List of Tables

List of Figures

General introduction

Amidst the uncertainty that characterizes cities and their evolution, short-term rentals (STR) have emerged as both a symptom and a catalyst of urban change. Not only are they at the heart of a substantial subset of interdisciplinary literature and of regulatory efforts, they also prominently shape modern urban living narratives in as many ways as they impact different segments of urban life. This work considers cities and their neighborhoods from a complex systemic perspective, where the confrontation of STR activity with existing agents, interactions, and dynamics leads to uncertain outcomes. In a bid to contribute clarifying empirical and conceptual procedures to assess how STR relates to urban change, the following thesis situates itself at an evolutionary-motivated intersection between urban economics, economic geography and urban planning.

The ambivalent effects of short-term rentals

The rise of short-term rental (STR) online platforms has drastically changed the way tourists, nomad workers and dwellers experience cities. Thanks to online intermediaries and platforms like *Airbnb*, it has become easy to lease accommodations for short periods of time in flexible ways. This has granted dwellers the opportunity for additional revenue when their dwelling is not occupied. It has offered both tourists and nomad workers a wide variety of lodging in different neighborhoods of cities, leading to an "off-the-beaten-track" (Maitland, 2010) experience that fuels the authenticity perceived by visitors (Füller and Michel, 2014; Lalicic and Weismayer, 2017). It is undeniable that the so-called "*sharing economy*" STRs have claimed to be a part of[2] can theoretically con-

[2] This is a rightfully disputed claim, see Cocola-Gant and Gago (2021).

tribute to solving market inefficiencies, provide value for tourists and hosts, contribute to economic development (A. Hidalgo et al., 2024), and redistribute revenue streams linked to tourism differently. But their impact on urban spaces and on residents has also sprung constroversy.

In fact, most touristic agglomerations in the world face important negative externalities. Barcelona, Lisbon, and Berlin are just some examples of iconic cities where central and historic neighborhoods have been transformed under pressure from short-term rentals. STRs are blamed for many troubles, namely for a pressure on housing prices, for transforming neighborhoods and their available amenities, for nuisance and nightlife that affects the well-being of inhabitants, and for an unfair competition with the hotel industry because of their looser norms and the lack of an adequate tax system. While STRs are likely part of broader urban change and are not the only ones to blame for these phenomena, the new urban tourism they facilitate makes change a lot more visible (Füller and Michel, 2014), and some of these troubles have strong empirical grounding in the literature. The positive association between lodging activity through platforms like *Airbnb* and raises in house and rent prices has been well established. Different causal designs, different datasets, different spatial settings and different time frames have all been leveraged to prove this (Ayouba et al., 2020; Barron et al., 2021; Duso et al., 2020; Garcia-López et al., 2020; Horn and Merante, 2017; Koster et al., 2021; Lee, 2016; Valentin, 2020), although there are suggestions of spatial heterogeneity in the effects (Franco and Santos, 2021), and that these effects are not exclusive to touristic urban agglomerations (Allam, 2023, Chapter 4). By making investments in rental housing more flexible and efficient, STR platforms enhance ongoing trends of housing financialization and eventually lead to displacements of residents (Cocola-Gant and Gago, 2021). They do not allow for a more equal redistribution of revenues linked to tourism as much as they provide additional tempting opportunities for investors.

There is also evidence that STR activity changes the overall urban landscape and has a sizeable impact on commercial amenities. STRs have of course affected the competing hotel industry (Zervas et al., 2017), but more interestingly, they are contributing to reshape the local economy of neighborhoods beyond direct competition. Basuroy et al. (2020) suggest a positive relationship between *Airbnb* and local restaurant revenue in Texas. A. Hidalgo et al. (2023) expand upon this in Madrid and find that STR activity tends to displace resident-oriented businesses in favour of tourism-oriented ones. These causal assessments of STR impacts on specific variables (housing prices, rents, selected retail businesses) provide valuable insights into the dynamics at play.

Policy-makers have not been blind to the trends outlined above, nor to the increasingly negative perception residents have of STRs. *Airbnb* has become the face of touristification of neighborhoods in Barcelona, Amsterdam, Athens, Split, New York and many other cities, paving the way for a proliferation of anti-*Airbnb* actions and protests. The motivation for cities to regulate is clear, and legislation on STRs is now widespread. Some regulations have notably mitigated the specific impact of STR on housing, and have provided opportunities for informative natural experiments in the process. Koster et al. (2021) show that regulation in Los Angeles has been associated with a 2% reduction in house prices and rents, while Duso et al. (2020) also find reductive effects of STR regulation on the housing market in Berlin. Valentin (2020) finds up to a 30% regulation-induced reduction in property values in the most STR-intensive neighborhoods of New Orleans, although his results also suggest spillover effects of regulation pushing STR towards unregulated areas. However, and despite the strength of these causal results, it would be wrong to consider them universally valid. Cities are not homogeneous in the way they entice and adapt to STR, and legislation regarding STRs is not homogeneous either. Cities are more or less strict and more or less specific regarding the hosts and areas they target. From outright bans to quasi-*laissez-faire*, they each articulate different qualitative and quantitative rules to fit their objectives (Aguilera et al., 2021; Bei and Celata, 2023; Nieuwland and Melik, 2020), pre-existing legal frameworks (Briel and Dolnicar, 2021), and overall environments. There is no unique STR regulation that would fit every environment. The plethora of approaches motivates a need to better understand the impact of different types of regulation, especially because there is no guarantee regarding the transposability of a particular regulatory effort from one city to another.

More importantly in the context of this work, these different regulatory approaches outline the spatial heterogeneity of STR impacts that need to be mitigated, as perceived by policy-makers. There is no one-size-fits-all STR regulation because there is no one-size-fits-all geography of STR activity, nor of how this activity relates to places that are singular by nature. On top of this, and while part of this introduction insisted on the very clear trend of STR making housing less accessible to long-term residents, this trend is an outcome of an underlying transformative process. The vehement resistance of residents towards STR specifically in the wider contexts of touristification and gentrification reflects a sentiment that would seldom be captured by rent or property prices alone. *Airbnb* and the "*new urban tourism*" it has facilitated transforms places that were up until now less prone to touristification, and by expanding the "*tourist bubble*" (Ioannides et al., 2019), it changes communities, environments, and interpersonal interactions within them in ways that are difficult to measure. STR activity inherently modifies the systems within cities by being a part of them, and thus impacts the lived experiences of residents. I argue that while outcome-focused causal approaches have been and will continue to be informative, a systemic perspective of cities along with holistic empirical tools are a fitting contribution to further collective comprehension of the STR phenomenon and of surrounding regulatory efforts.

A systemic perspective on urban change

This thesis's view is that cities are to be seen as complex systems that emerge through dynamic interactive processes between all of their constitutive agents. This is far from being a novel perspective within of itself, but the articulation of complex systemic planning perspectives, evolutionary economic geography and economic complexity provides new tools to unpack the process of change within cities – Which STRs have been an important part of in the last decade.

Over 60 years ago, Jane Jacobs (1961) contributed to a perspective of cities wherein urban spaces are shaped as a result of the people that experience them, and of their interactions. This perspective inadvertently co-functioned with emerging generalizations of cybernetic theory and their application to urban science that Batty (2009) aptly describes. In this paradigm, cities were to be seen as both social and spatial systems whose equilibria came as a high-order aggregation of agents' low-order objectives, and were dependent on centralized action. However, these ideas have since evolved into an understanding of cities as *complex* systems where decentralized action (that is, of individual components at a micro level) plays a larger role, in part thanks to interdisciplinary efforts (Allen, 1997; Anderson, 1972; Batty, 2000; Krugman, 1996; Nicolis and Prigogine, 1989). In fact, any division of space (neighborhoods, districts, blocks) or of activities (transportation, retail, health) within a city also holds the characteristics of complex systems itself, leading to a multilevel view of cities as complex systems of complex systems that themselves operate within broader regional, national and international systems. Path-dependence, feedback effects, and the indefinable amount of components and of their boundedly rational interactions constitute continuous disturbances to the environments of these systems, and they all prevent any stable equilibrium from ever being reached. In this context, infrastructure, housing, jobs, innovation, education, and health involve interrelated dynamics among various agents with different objectives. They are difficult to observe and inform given the non-linear, chaotic nature of the outcomes they produce. Planning is inherently concerned with future outcomes, and the complex dynamics underlying cities' systems breed uncertainty and unpredictability over these outcomes. Still, the planning-oriented adaptation of concepts such as self-organization (Krugman, 1996; Moroni et al., 2020; Portugali, 1999), flows (Batty and Cheshire, 2011), and scaling (Batty, 2007, 2012) have offered systemic explications of how cities emerge, adapt and evolve.

Another complementary way in which to incorporate principles of change within these complex systems is to take an evolutionary stance. The generalization of biological evolutionary theory in the last decades, particularly through its modern applications in economics (pioneered by Nelson and Winter, 1985), has allowed for evolutionary economic geography to thrive (Boschma and Lambooy, 1999; Kogler et al., 2023). Planning has not been immune either to biological analogies throughout its history (as explained in Batty and Marshall, 2009) nor, more recently, to concrete applications of evolutionary theory (Bertolini, 2010). In fact, one could argue that the view of cities as complex systems that are built from the bottom-up is fundamentally biological, and that changes in these complex systems can be characterised as evolutionary through concepts of inheritance, variety, selection and accompanying environments, retention, and adaptation. Evolutionary economic geography has leveraged the biologically-motivated perspective of Generalized Darwinism (Essletzbichler and Rigby, 2010; Essletzbichler and Rigby, 2007) to characterize the unevenness of economic activity in space and its change. It is also a promising and suitable way of understanding urban spaces and change.

This work takes the view that short-term-rental platforms have induced many small changes in the presence, objectives, actions, and interactions of the elements and agents that constitute cities as complex systems. In doing so, they have led to significant changes in the macro-level organization of urban systems. These systems' unpredictability however makes the nature and scale of STR-induced changes (and with them STR policy-making) particularly challenging to assess. While focusing on single outcomes such as housing or individual commercial amenities as existing literature has done provides valuable insight, it falls short in capturing urban change in a broader way, and in addressing key impalpable concerns surrounding STR. Finding ways of characterising and typologizing urban systems, and of confronting their evolution to STR, is a promising endeavour towards understanding the difficult-to-observe urban transformations linked to STR. STRs not only reflect but also reshape urban geography in ways that can seldom be captured without a holistic lens.

Networks of commercial (retail) amenities provide a holistic lens to tackle the issue of urban change systemically. They are representative of broader systemic characteristics and can allow for a micro-founded perception of urban change. Through generalized

Darwinism's analogies (Essletzbichler and Rigby, 2010; Essletzbichler and Rigby, 2007), uneven locations of different industries' commercial amenities are dependent on a selection process that is conditioned by retailer's consumption anticipations, and motivated by environments' (systems') abilities to accommodate said consumption. Retailers locate in places they anticipate to find consumption in. However, and through a "*consumer city*"-like (Glaeser et al., 2001) perspective, they are only a part of the reason potential consumers visit these places. A network of commercial amenities by industry is both representative of how consumer preferences are distributed across space, and of how unobserved characteristics that motivate consumption are distributed – As anticipated by retailers. Through this process, the nature and scale of urban change induced by STRs is passed onto neighborhoods' and cities' commercial amenity compositions, both directly and indirectly. STR tourists are likely to have different consumption habits to residents and to directly incentivize the presence of certain types of retail over others. Indirectly, they also change the way non-tourists relate to environments through other mechanisms, and these changes are also reflected as outcomes in commercial amenity networks. Although empirical observations of these networks hold valuable information over systemic states, unpacking their information into readable indicators is not trivial.

Empirically characterizing complex economic systems through their components has been the focus of one of the more recent bifurcations of evolutionary economic geography, namely economic complexity (Balland et al., 2022; C. A. Hidalgo, 2021; C. A. Hidalgo and Hausmann, 2009). This paradigm leverages network-driven readings of components and of interactions within systems to yield holistic perspectives. The methodological tool set it provides has found undeniable empirical success in describing the unevenness of various economic indicators, and of their change, at different scales. One of these tools, the economic complexity index, characterizes and uncovers change within complex systems through the observation of their components. With methodological adaptations and a solid conceptual grounding, an application of economic complexity indexes to commercial amenity networks fits a systemic framework of urban spaces and allows for a readable reduction of those spaces. These indexes, if conceptually unpacked, have the potential to enhance current knowledge regarding systemic urban change at large. They are thus an appropriate candidate to tackle critical questions surrounding short-term rentals and their uneven impacts, as well as to contribute to designing smart STR regulation.

Outline

This thesis is articulated in four separate chapters that together deliver an overview of what a systemic understanding of urban spaces can bring to the STR literature. The first chapter addresses the crucial question of STR regulatory design. STR regulation is made particularly difficult by policy-makers' different rationales, objectives, and understandings of how STR affect urban spaces. It is also made difficult by the fact that there is little history and hindsight regarding the efficacy of different regulatory designs. By using STR regulation in Bordeaux as a natural experiment through a difference-in-differences and its complementary triple difference models, it offers insight over the localized effect of a specific regulation. In doing so, it also outlines how difficult it is to align regulation design with the inherently systemic impacts of STR. There is a need for a systemic stance towards the regulation of STRs, and towards how they relate to key urban change.

Chapter 2 acts upon the need for a systemic outlook on cities to capture urban situations and change, of which STRs are a part. To do so, it adopts an evolutionary frame of reference wherein retail locations come as an imperfect but holistic reflection of broader urban complex systems. By situating itself at the neighborhood level, it takes particular interest in the uneven spatial distribution of different types of retail activity, and leverages it to understand the unevenness of their broader environments. Then, by adapting the *Economic Complexity Index* and the *Product Complexity Index* to urban needs and by unpacking the indicators' complicated formulation, it presents a new and promising way of characterizing urban systems empirically and of observing their change. The *Amenity Complexity Index* (ACI) and the *Amenity-Type Complexity Index* (TCI) are continua that use data on commercial amenity locations (by industry, or type) to respectively separate places and commercial amenity types along a single axis. By extension, the indexes separate underlying consumers by their preferences for types of goods and services, and for overall environment characteristics. Focusing exclusively on amenities that require market coordination allows for micro-founded interpretations. On these continua, the relative proximity of places (amenity types) reflects the relative similarity of their environmental characteristics and of the preferences of their present consumers.

Chapter 3 empirically illustrates the ACI and the TCI as conceptualized in the previous chapter in Paris, using *APUR*'s survey-based retail location dataset in three different years. Its separation of retail goods and services into over two hundred types of commercial amenities provides a rich and dynamic perspective. It builds and expands upon the previous chapter's micro-foundation, and finds that realized spending of consumers (as anticipated by suppliers) is the most credible candidate to associate with TCI and ACI-driven separations. Spending tends to be intertwined with differentiated consumption of amenities, and higher spenders have more influence on retailer locations. As such, this chapter presents an economic interpretation of the indicators as revealing the preferences of higher spenders both for types of commercial amenities and for other place characteristics. Various descriptive applications demonstrate how these reflections of spending-driven compositions of demand can be articulated into urban transformation narratives. For example, it hints that traditional tourism hotspots in Paris are becoming less compositionally different from other less-tourism oriented neighborhoods over time, and it outlines neighborhoods undergoing spending-driven environmental changes. The spatial distribution of ACI values is a particularly apt window into the evolution of Parisian neighborhoods, and makes for a new reading of urban change. Quantitative links between various socio-demographic variables and ACI values are also highlighted, further suggesting that the indicators are unpacking real-world situations.

The fourth and final chapter begins with a solidification of Chapter 3's findings by expanding the spatial and temporal scopes of the ACI/TCI methodology. It observes commercial amenities within 10 of the largest French cities using a national-level business registry over a 15-year window. This dataset's lesser granularity and the adapted methodology it requires are first confronted to the previous chapter's results to test for the resilience of the indicators and of their interpretations. It finds that the ACI and the TCI and their previously unearthed interpretations are robust to the new spatio-temporal setting and to the less granular data, granting them additional external validity. The chapter then uses the indicators for a preliminary empirical investigation of how STR activity changes the composition of local demand within cities. Correlations of TCI ranks between cities stand out as spatially sticky, informative indications of cross-city compositions of demand. Using these correlations and their evolution through time, it empirically suggests that STR tourism is linked to a homogenization of urban spaces across cities. In doing so, the chapter also highlights the need for policy-makers and researchers alike not only to think of STR impacts in a new way, but also of doing so on a relatively underexplored inter-city scale.

All together, this work combines evolutionary stances on urban economics, on economic geography and on planning to contribute towards a complex systemic understanding of urban change and of the processes that underlay it. The perspective and methodology it provides can help researchers in the study of contemporary urban crises such as STR. The approach also underlines the argument for policy-making to take into account the complex systemic nature of cities and neighborhoods. As readable, reductive, and data-informed tools, the ACI and the TCI pave the way for a practical strengthening of the complex cities paradigm.

Chapter 1

Assessing short-term rental regulation in Bordeaux[1]

1.1 Introduction

The rise of short-term rental (STR) online platforms has drastically changed the way tourists, nomad workers and dwellers experience cities. STR tourists (and hosts) seek "off-the-beaten-track" (Maitland, 2010) experiences throughout cities' different neighborhoods, and they are at the heart of ongoing urban changes that have led to tensions. As outlined in the general introduction, existing literature has clearly established the negative impact of STR accommodations on housing and on the rental market, legitimating the need to regulate the development of these new forms of rentals (Ayouba et al., 2020; Barron et al., 2021; Garcia-López et al., 2020; Horn and Merante, 2017).

As a result, a global movement of regulation is rising over affected cities, starting with metropolitan and mostly touristic destinations (Nieuwland and Melik, 2020). However, no dominant design seems to exist in the rules implemented by governments (Bei and Celata, 2023; Nieuwland and Melik, 2020), whether they are local or national. Each city appears to pick in a bunch of heterogeneous measures, which range from a limit on the number of days authorized for STR to a mandatory registration or even a strict ban

[1]This chapter is an iteration of a published work co-written with DEJEAN Sylvain and SUIRE Raphaël: Robertson et al. (2023), available at https://doi.org/10.1007/s00168-023-01215-4. The co-authors have agreed to the following modifications and use.

in specific areas (von Briel and Dolnicar, 2020). In the end, even laissez-faire remains an option for many cities. The ability of local and national authorities to enforce these regulation attempts is also very heterogeneous, which makes the assessment of the overall effectiveness of these public policies difficult.

However, analyzing a specific regulatory effort alone could provide insight about the method it uses. More importantly in the context of this work, it can help open up a discussion about the scale and way these efforts are led in. As such, I investigate the particular rule change regarding STR in Bordeaux, France, and assess its effectiveness in regard to the policy objectives behind it.

This chapter's research is a part of a growing stream of literature that evaluates the consequences of specific STR regulatory efforts. Empirical research endeavours in Berlin (Duso et al., 2020), across American cities (Bekkerman et al., 2023; Chen et al., 2021), in Los Angeles County (Koster et al., 2021) and in Geneva (Falk and Scaglione, 2024) all highlight reductive effects of STR regulation, albeit in different forms and settings, and at different scales. New Orleans' regulation is also particularly suited to natural experiments, and research has suggested regulatory inefficiencies driven by dilution over time and of spillover effects onto unregulated areas (Valentin, 2020; van Holm, 2020). Still, the stark differences between adopted regulations across cities (Bei and Celata, 2023; Nieuwland and Melik, 2020) and their diverging rationales (Aguilera et al., 2021) motivate the need for further analysis in individual cities. The specific nature of regulation in Bordeaux underlines an opportunity to evaluate the individual consequences of a qualitatively and quantitatively targeted regulation. Crucially, Bordeaux's targeting framework is enabled by a national-level law that is also available to other French cities that wish to adopt (and adapt) it. Understanding its impacts is thus helpful towards informing other French cities that are in a position to leverage the same framework.

The context of Bordeaux's willingness to control the development of STRs and limit their negative consequences by implementing an original compensation rule provides a unique opportunity to evaluate the causal impact of a regulation attempt. In July 2017, and enabled by changes in national-level legislation, the city hall of Bordeaux decided to pass a rule which strongly increased the cost of turning dwellings into furnished tourist accommodation. Following a huge increase in STR supply, local authorities decided on a strict "compensation rule" that targets hosts who lease their primary residence for over 120 days a year, as well as hosts that lease secondary residences. Regulation only allows new entries on the STR market of such "commercial" hosts if they compensate for their STR activity by acquiring a comparable commercial dwelling in the same area and turning it into a primary residence for a long-term stayer. In other words, it imposes a "one in, one out" principle to new commercial STR accommodation. In the continuum of regulation measures, the city of Bordeaux chose a comparatively moderate action (see Bei and Celata, 2023 for a perspective of that continuum) by encouraging a targeted change in individual behavior of leasing. Their objective is to prevent or at least limit some undesired urban dynamics (nuisance, negative externalities, emergence of *Airbnb* clusters, etc.) by restricting *Airbnb* activity, while preserving desired "sharing economy" aspects of STR activity intact.

In the following, I estimate how the supply of STR is affected over the city in the most touristic area (city center) but also in the more residential districts (peripheral areas), and take a look at who are the hosts that suffer the most from the novel restrictions to evaluate how efficient the regulation is. To this aim, 3 categories of hosts are distinguished: those who are targeted by regulation, those who are not targeted (STR hosts that follow the "spirit" of the sharing economy by renting all or part of their dwelling for a limited period of time), and those of inhabitants who newly enter the market of STR after the regulation.

Two complementary empirical designs are used. First, a difference-in-differences to estimate the overall effect of regulation, using a comparable set of cities not affected by the regulation as a counterfactual. Then, to estimate the effect for the most peripheral and residential districts specifically, I leverage a triple difference design which compares the difference between peripheral districts in Bordeaux and districts in adjacent municipalities to the difference between peripheral districts in the control group cities and districts in their adjacent municipalities. This second design has the advantage of fixing some issues of the first, while highlighting the potential different effects of the regulation in the more residential and peripheral districts of Bordeaux. The main results are twofold:

1. Regulation in Bordeaux has had a net strong negative effect on STR activity of 322 rented days per month per district on average, representing just under half of a pre-regulation standard deviation. It has also led to 49 fewer listings per month and per district on average. This aggregate result also holds in the city's peripheral neighborhoods, albeit less intensively. Peripheral districts inside Bordeaux have rented over 47 fewer days within the given bandwidth on average than their counterparts outside the area covered by the regulation, in the two years following the regulation.
2. These effects are for the most part driven by a decrease of the activity of those who were targeted by the regulation, as well as by the decrease in the number of entrants in the STR market. However, and importantly in the context of this work, the activity of those who were not targeted by the regulation has also decreased, suggesting some unexpected informational effects that highlight the need for a different approach.

These results also suggest that the regulation's spatial delimitation induces changes in the distribution of STR activity on both sides of Bordeaux's border, encouraging a larger share of outside activity. If not a result, another important observation also arises: It is impossible to fully confront Bordeaux's regulation to its objectives, because some objectives it supports are vague and reflect systemic phenomena. This chapter therefore argues both for the need of new ways to regulate digital platforms devoted to STR, and for the need of new ways of understanding how STR relate to urban change. In particular, a one size fits all policy comes with undesired and indirect effects, and a more targeted and smart regulation is undoubtedly more accurate. But a smarter regulation is difficult to design without a framework that captures the systemic nature of STRs' impact.

The remainder of this chapter is as follows. Section 2 describes the context of Bordeaux's regulation and the expected effects of the compensation rule. Section 3 is devoted to the empirical strategy with two different econometric designs. Section 4 reports commented results which are then discussed in section 5.

1.2 The context of Bordeaux's regulation

Bordeaux is the 9th largest French city with over 270,000 inhabitants in 2022, and is at the heart of a metropolitan area that hosts 1 million people. Known worldwide for its wine industry, the city of Bordeaux also benefits from a mild climate, is near the Atlantic Ocean and offers a renowned gastronomy. According to official statistics of the city, in 2019, 6.35 million nights had been officially rented (in hotels and in registered STR). This figure has steadily increased over the last few years, suggesting that domestic but also international demand for STR has been growing in recent times. This however comes along with various negative externalities. Bordeaux's response to the STR phenomenon was relatively quick, as it was the second French city after Paris to adopt laws limiting it. The motivations behind the need to regulate are summed up by the July 10th 2017 deliberation of the city council[2], which considers five main issues and negative externalities that are created by the development of STR – Issues that echo those explored in this work's general introduction:

- Transformation of some neighborhoods into exclusively touristic areas with a disappearance of community life associated to a degradation of social capital;
- High pressure on housing prices because of the high profitability of STRs;
- Degradation of properties and shared spaces because of the high rotation of guests;
- Competition with traditional tourism accommodation;

French law already addressed the issue of tourist tax collection, which was the fifth objective. It has made it mandatory for hosts to declare their listings to city councils for

[2] http://www.bordeaux.fr/images/ebx/fr/CM/12461/12/acteCM/69313/pieceJointeSpec/148521/file/acte_00046189_D.pdf

taxation purposes, thereby obtaining registration numbers that authorized them to lease their properties on STR platforms. On July 10th 2017, the city of Bordeaux released a deliberation regarding the declaration and registration of STRs[3]. The city, with its deliberation, decided that it would no longer authorize property owners to change the use of their dwelling from "meant for habitation" to "furnished tourist accommodation" unless they abided by extremely strict rules. For new changes in use, the rules in Bordeaux include the need to "compensate" by purchasing another property of a similar size and in the same area that was until then used for commercial purposes. Incoming hosts that want a change in the use of their listing (from "meant for habitation" to "furnished tourist accommodation") would then need to transform their newly-acquired commercial premise into a long-term habitation if they want to be able to lease a secondary residence or any residence for more than 120 days per year.

The new regulations were officially enforced in Bordeaux on March 1st, 2018. It is important to note that existing listings at the time the regulation was announced have had time to register for changes in use without the need for compensation. The new rules are not designed in a way that will phase commercial listings out, but they will limit new entries. Similarly to Paris, public agents were hired to track down illegal listings. On January 1st, 2019 and as a result of the newly passed ELAN law[4], *Airbnb* showed willingness to cooperate as it announced that listings declared as primary residences in Bordeaux would be blocked after having been rented out for 120 days in a calendar year[5]. Together with that came laws making it easier to hand out fines to both hosts and platforms that would not respect the rules[6]. Figure 1.1 serves as a recapitulating timeline for the evolution of STR regulation and enforcement in Bordeaux, and for the time window considered in this chapter.

[3] See the preceding footnote

[4] The ELAN law was passed in November 2018 and also raised the potential fines for non-abiding hosts. Its effect is controlled for in Appendix 1.7.4. https://www.legifrance.gouv.fr/loda/id/JORFTEXT000037639478/

[5] This agreement was established together with other cities and negotiated at a national level, which possibly alleviated the lack of bargaining power medium-sized cities can sometimes have with leading platforms https://immobilier.lefigaro.fr/article/airbnb-vous-ne-pouvez-plus-louer-plus-de-120-jours-par-an-a-paris_47a64ea6-0cda-11e9-ba7e-e142f78c8ed1/

[6] Up to 10.000€ for hosts and 50.000€ for platforms per non-compliant listing https://www.legifrance.gouv.fr/eli/loi/2018/11/23/TERL1805474L/jo/texte

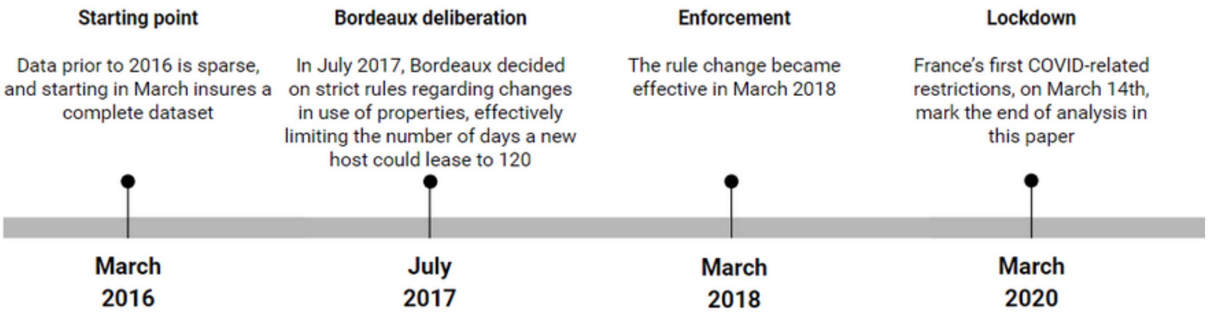


Figure 1.1: Timeline of regulatory events in Bordeaux, France, and of this chapter's analysis.

1.2.1 Expected effects of the regulation

The compensation rule is designed to provide different incentives according to the intensity of the short-term rental's usage. "Non-targeted" hosts are those who live in their dwelling and lease it less than 120 days per year as well as those who lease only a part of their habitation. They should not be impacted by the regulation. As described earlier, they are not the target of the regulators. However, It would be naive to believe that this new regulation has been perfectly understood and integrated by hosts who have very different objectives and face various economic and social contexts. Some could have believed that they had to stop renting their home on *Airbnb* even if they were not affected by the regulation while others could have not been aware of the risk taken in renting more than 120 days and thus capitalized on the success of their accommodation, thereby increased their activity regardless of the new regulation.

For hosts defined as "targeted", the compensation rule applies. Those who live in their habitation and rent more than 120 days have to lease more sparsely, but those who rent a habitation where they do not live have to stop doing it (unless they were already registered) or make the costly compensation. The expected decrease in the overall number of reservation nights in STR could be the consequence of the diminution of days available for rental, but also of listings who exit the market because they became illegal, or because renting one's primary residence less than 120 days a year is no longer profitable.

However, targeted hosts' willingness to conform to the compensation rule is directly related to the ability of the local government to enforce the regulation. It is also, as exposed previously, conditional to the fact that the renters have not understood that between the announcement of the rule and its effective enforcement, a period of 8 months enabled them to change their dwelling from "meant for habitation" to "furnished tourist accommodation" without the need to compensate. This possibility, even if not advertised, could have lowered the effect of the compensation rule and the expected decrease in targeted leasers after the regulation.

A last category of renters has to be considered, those who entered the market after the regulation. They are referred to as the "entrants". The way regulation could have changed the behavior of entrants is twofold. First, a clear and transparent legal framework could favor new entry of non-targeted hosts because they now know better what is legal or what is not, reassuring those who were reluctant to lease their accommodation. The second explanation is driven by market dynamics. As the targeted hosts, who are by definition those with the most important STR activities, leave the market or decrease their leasing activity, they give extra opportunities for new law-abiding entrants. At a constant demand, and if entry is costless, it could be that the number of new entrants compensates "at most" the exit of targeted hosts.

From the regulator's perspective, the main objectives seem to be twofold. First, there should be a global decrease in STR activity. Less STR reservations could directly address some issues raised by the deliberation of July 2017. Second, there should be a change in the composition of STR activity, wherein the aforementioned decrease is driven by commercial hosts whereas non-commercial ones are not effected. In terms of the framework discussed above, the magnitude of the expected effect can be described as follows:

$$\Delta\text{STR} = \Delta\text{non-targeted} + \Delta\text{targeted} + \text{entrants} \leq 0$$
$$\text{with } \Delta\text{targeted} < 0$$

$$\text{and } 0 \leq \text{entrants} \leq |\Delta\text{targeted}|$$

It is worth noting that as the regulation only impacts the city of Bordeaux, spatial spillovers with cities in the suburb of Bordeaux could exist. The discontinuity created by this localized regulation can create opportunities for hosts outside of Bordeaux but sufficiently close to the border to attract the demand eventually left vacant in Bordeaux. In this regard, the success of the compensation rule in Bordeaux could be misleading from a wider metropolitan scale. At least, the existence of potential spillovers could question the way dwellers in areas where the regulation does not apply change their behavior. The aim of the following empirical strategy is to evaluate the overall causal impact of this regulation and to disentangle the different effects at stake. To this aim, I propose two different and complementary empirical designs that assess the effect of the STR regulation.

1.3 Empirical strategy

The uniqueness of Bordeaux's regulatory situation both within its urban area and within its region creates an opportunity for a natural experiment setting that can help evaluate the causal impact of the regulation.

Firstly, a difference-in-differences (DID, or DD) design will allow for an estimation of the aggregate net effect of regulation on activity in Bordeaux. Bordeaux's regulated districts are compared to districts in three other cities of the same region (La Rochelle, Biarritz, and Bayonne) where the regulation did not occur. The DID specification is indeed particularly useful when attempting to identify the global causal effect of a given policy. This global causal effect is for a large part driven by tourism-intensive neighborhoods in the city center, where *Airbnb* accommodations are abundant. However, given the structure of the selected cities, it is at risk of being center-inflated and of poorly reflecting what is happening in peripheral districts. Also, selecting control cities that are far enough from Bordeaux makes regulation-related spillover contamination towards control group cities a very weak threat, but it stresses how prone the DID can be to confounding factors that affect Bordeaux while sparing control cities. This is the design's key weakness.

This is where the second step of the empirical strategy comes in, with a triple differences (TD, or DDD) design. In the TD, I go slightly deeper than in the DID by comparing the difference of STR activity between Bordeaux and its close periphery to the difference of STR activity between the control group cities and their close periphery. The main advantage of this setup is that it mitigates the first design's primary weakness. The underlying assumption in the TD is no longer a parallel trend of STR activity, but rather a parallel trend of the difference between STR activity within cities and within their close peripheries. The structural similarity between Bordeaux and the control cities is not as central to the design as in the DID, what matters is that relative outcomes inside and outside of cities trend in the same way around Bordeaux as they do in control areas (Olden and Møen, 2022). This means that a shock that affects Bordeaux without affecting treatment cities is no longer a threat to validity as long as the outside periphery of Bordeaux is also affected.

The TD also uses a different data set to the first design. Only areas close to the borders of cities are included, to keep accommodations in the treatment and the control group as similar as possible. With this design, the ability to gauge the aggregate causal impact of regulation is lost, but the effect of the regulation in the city outskirts can be better investigated. These peripheral areas might experience disproportionate negative impacts from STR activity given their more residential structure, an effect which can be magnified by the proximity with adjacent areas where the regulation did not occur[7]. As such, combining both of these designs provides further robustness to this chapter's results, since they make up for each other's weaknesses and require different assumptions. The different spatial areas at hand could also provide a different kind of robustness, as I

[7] The use of a spatial regression discontinuity design was considered in a previous iteration of this work. While its results were promising, the spatial scrambling of listings on *Airbnb* (and thus in our data) significantly weakened the design and led to it being dropped.

would be able to show they also hold in more residential peripheral neighborhoods.

1.3.1 Data

This chapter relies on STR listing and activity data acquired from specialized American company AirDNA, one of the standards for research in the field (Ayouba et al., 2020; Chen et al., 2021; Cocola-Gant and Gago, 2021; Valentin, 2020). The data include monthly web scraps beginning in 2016 of all Nouvelle-Aquitaine listings from both *Airbnb* and *HomeAway*[8]. For every listing in every month for which it is active, the dataset provides a number of variables relevant to this research. The estimated number of nights that were booked in a given month[9], which will further be referred to as "Reservation days", is an indication of the influx of STR tourists. It is the dominant variable behind the STR-related problems raised earlier in this chapter, and the one upon which policy-makers have imposed a limit. The estimated number of days for which a listing is available for rental, regardless of whether it is rented or not, is also at the practitioner's disposal in the data[10]. Additionally, information about the spatial coordinates of listings (the shortcomings of these spatial coordinates are discussed hereafter), their price, and their type is used. The "type" variable qualifies the listing as an entire home, a shared room, a private room or any other category of accommodation. The aforementioned data is aggregated at the district level. This allows for the introduction covariate district[11]-level structural data[12] from the French national institute for statistical and economic studies[13]. These data are available at the census tract level, and contribute to the use of census tract (which will further be called "district") aggregation further along in this chapter. The time window of this work's observed data closes at the beginning of March 2020 because of the strong heterogeneous impacts the COVID-19 pandemic along with its restrictions

[8]The French branch of *HomeAway* is known as *Abritel*

[9]This number relies on *AirDNA*'s estimations, which themselves rely on machine learning algorithms using different booking signals and ground truth data from partnered property managers. While the methodology is not fully open, it is explained in surface at this link: https://www.airdna.co/airdna-data-how-it-works, see "Booked vs. Blocked Days".

[10]It is obtained by subtracting estimated blocked days from total calendar days. See the previous footnote.

[11]IRIS are "aggregated units for statistical information", their aim is to divide cities into units of about two thousand inhabitants.

[12]2016 "Filosofi" scheme for median available income, 2018 facilities database ("Base permanente des équipements") for the number of restaurants and hotels.

[13]Institut National de la Statistique et des Études Économiques, INSEE hereafter.

might have had on the observed units.

To evaluate the impact of the compensation rule on the different categories of hosts (commercial or not), I attempt to differentiate the listings that are presumably targeted by the regulator from those that are not.

It is clear that the rule intends to impose a very important cost to those whose entire home listing is not used as anyone's primary residence. The regulation discriminates against the latter through the number of reservation days; the 120-day rule is in fact designed to seek out secondary residences. This implies that a listing rented on STR platforms for over 120 days in a calendar year cannot be anyone's primary residence. In this chapter, the principle is extended to any day a listing is made available for rental and the need for calendar years is dropped. As such, a listing is defined as targeted if it is an entire home that is at least available for ten days per month on average during its stay on the market, with a minimum stay of 6 months. This does not follow the way regulation is enforced, but it follows its principles. I seek to find accommodations that are not peoples' primary residences, albeit in a more comprehensive way. Reservation days are the regulation's only observed variable, but availability for lease is a fair representation of the policy-makers' intention to cut down on non-primary residence listings: a host that lists an entire home on a STR platform is likely not planning on sleeping in it.

Table 1.1 shows descriptive statistics for both categories of listings in Bordeaux, both before and after regulation took place. At first glance, both categories seem to have grown in quantity and in activity over time, with a stronger growth for non-targeted listings. However, it has to be kept in mind that this global increase is an average over a 2-year period, it is thus possible that STR activity varies along the post-regulation period.

	March 2016 to February 2018 (pre-regulation)				March 2018 to February 2020 (post-regulation)			
	Number of listings	**Mean Res. days**	**SD Res. days**	**Median Res. days**	**Number of listings**	**Mean Res. days**	**SD Res. days**	**Median Res. days**
Targeted	10,082	97.54	113.84	56	12,387	113.07	145.83	58
Not targeted	3869	71.36	94.70	36	4859	90.34	130.48	42

Table 1.1: Summary statistics of Bordeaux's listings according to whether I define them as targeted or not.

Dealing with data scrambling

The data provided by *AirDNA* have many advantages and are very comprehensive. However, like every STR data solution, they scrape public-facing data and therefore only have access to what is visible on the platforms' websites. As a matter of host privacy, *Airbnb* and *HomeAway* both offer hosts the ability scramble listings' coordinates within a 500-m-wide circle on the maps they provide to users.

For the spatial coordinates of listings, *AirDNA* gives the center point of these circles, which means listings could in reality be up to 250 meters away from where they are observed in the data. Further along in this chapter, listings are aggregated at the district (census tract) level. The scrambling should not be too threatening towards the models through the misattributing of listings to different districts. It is very likely that listings will be assigned to districts they do not actually belong to. I however believe this bears little impact on the results: listings may only be assigned to the wrong district if they are very close to the district's administrative border to start with, and they will only ever be wrongly assigned to a district that is contiguous to their actual unobserved district.

As controls, I use numbers of amenities (restaurants, hotels) and the median level of income within a district. These variables will naturally tend to show high spatial autocorrelation in these small non-enclosed spaces. Districts are thus already meant as an imperfect proxy of local (or semi-local) characteristics. While it is true that the *AirDNA* data scrambling might make these proxies slightly more imperfect than they are to start with and smooth the effect over districts in a sense, it should not significantly jeopardize the design.

However, the misplacement of a listing becomes a significant problem if it reflects the wrong treatment status, that is, if a listing outside of Bordeaux is wrongfully placed inside of Bordeaux in the dataset or vice versa. This makes spatial causal analysis difficult, especially in the TD design. 250 meters, at the scale of Bordeaux, is a small distance. For this reason, I decide to ignore listings that are fewer than 250 meters away from the border in both models. In the TD design, it is done on both sides of the border (inside and outside the cities and their adjacent municipalities). This cautionary approach ensures that no unit is falsely assigned to the control or the treatment group and thus enables the bypassing of the scrambling issue.

1.3.2 Difference-in-differences

Identification

Bordeaux's regulatory intervention is unique within its region, *Nouvelle-Aquitaine*. A difference-in-differences (DID) approach allows to evaluate the causal impact of the intervention on the STR market by using other cities as reference points for how STR would have evolved in Bordeaux without treatment, that is, without the regulation. This research design is an efficient way of observing the aggregate impact of treatment. As described previously, The most adequate dependent variable at hand for measuring the strain of STR activity on long-term residents is the aggregate number of nights spent by guests in STR accommodation. The model is specified like a DID as follows:

$$STR_{i,c,t} = \alpha + \beta_{D-D}(T_t \times D_c) + \text{Month}_t + \text{City}_c + X_i + \varepsilon_{i,c,t} \quad (1.1)$$

where $STR_{i,c,t}$ is the number of reservation nights or the number of listings for any given district (that is, census tract) i in city c during month t. T_t is equal to 1 for any observation made after the regulation's enforcement (is $t \geq 2018/03/01$) and to 0 otherwise. D_c is the treatment dummy. It is equal to 1 if city c is Bordeaux and to 0 otherwise, and the coefficient β_{D-D} of its interaction with T_t is the difference-in-differences estimator. $Month_t$ and $City_c$ are dummy vectors respectively used to control for month-specific fixed effects and city-specific fixed effects. Performing the analysis at the more granular district level i rather than only at the city level c allows for better control of structural diversity in the dataset through X_i. The number of restaurants, the number of hotels, the median level of income and the population are taken into account by X_i at the census tract level. They are fixed in time. In order to avoid downward bias of standard errors because of serial correlation, standard errors are clustered using the method described by Cameron et al. (2011) at the level at which regulation happens, that is, the city level (Abadie et al., 2023; Bertrand et al., 2004).

This research design relies on a few key assumptions. The underlying idea is that without treatment, the difference in activity on the STR market between treated and control units would have stayed the same, conditioned on covariates. To validate this assumption, the effect of underlying differences between the treated and control units has to be constant over time. Minimizing these differences is key when building a control group. Bordeaux is however a rather unique city in the Nouvelle-Aquitaine region, as it is both the biggest and the most economically diverse. This makes the prospect of finding a control group that fulfills the parallel trends assumption challenging. The four cities selected to be a part of the control group are Anglet, Bayonne, Biarritz, and La Rochelle. They all host important STR markets and range from being tourism-intensive (Biarritz) to being economically diverse (La Rochelle and Bayonne). All of them offer a proximity to the Atlantic Ocean's seashore, as shown on the map in Appendix 1.7.1, and they are all part of the same administrative region: Nouvelle-Aquitaine. These cities are used to simulate a random assignment for treatment. Appendix 1.7.2 offers a representation of the districts analyzed in the DID, both in Bordeaux and in the control group.

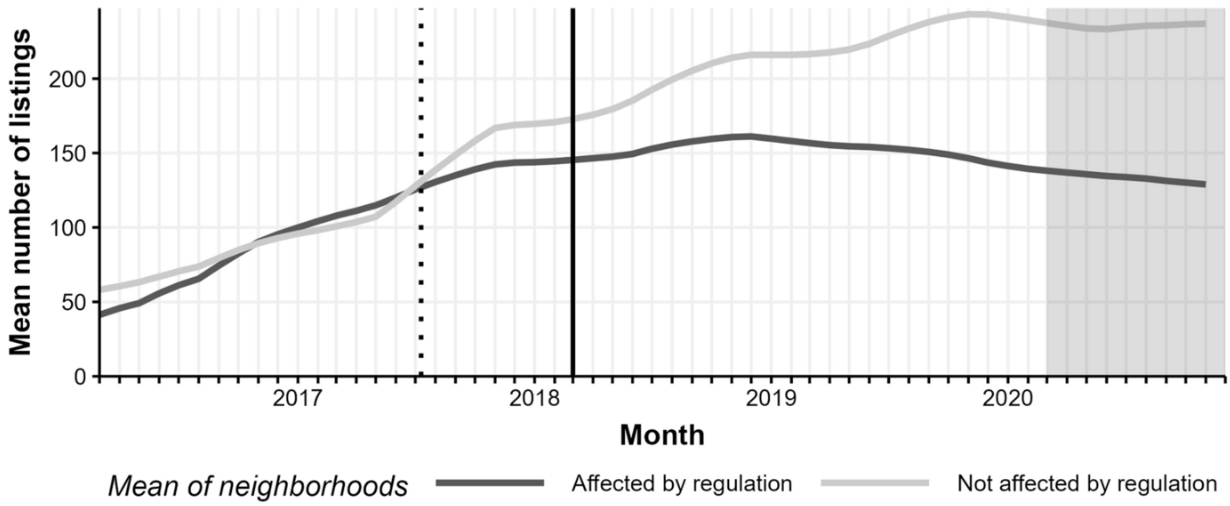
The control group was chosen as a way to strike a balance between the need for separate enough cities that are immune to spillovers from Bordeaux and for cities similar enough (demographically, administratively and geographically) for the likelihood of a STR shock affecting only Bordeaux to be lower. The designs are, however, not completely safe from these potential shocks. For example, a new high-speed train line reduced the time it takes to go from Paris to Bordeaux to two hours[14] (down from three) in July 2017. Since this happened at the same time regulation was announced, it is impossible to rule out an effect of the new train line on STR activity and to perfectly single out the causal impact of regulation with a DID. This kind of threat to the validity of this chapter's results is a good illustration of the value the TD design can bring in addition to the DID, as the former is more immune to such a threat.

Figure 1.2 depicts the 6-month rolling mean number of nights spent in STR accommodation over time in Bordeaux's (treated) districts and in other cities' (control) districts. Treatment districts are less sensitive to seasonal peaks but also record more activity in the winter. Control group districts experience growth in both peak activity and off-season activity even after the treatment, whereas the monthly number of booked nights in Bordeaux gradually declines over time once regulation is enforced. This is consistent with the expected effect of a regulation that blocks new targeted hosts from entering. Robustness checks through variations in the control group composition (see Appendix 1.7.4) reinforce the idea that this endeavour's results do not depend strongly on the observed structural differences between treated and untreated units. The potential influence of anticipation effects after the announcement will also be checked for with a double-DID design (see Appendix 1.7.4). Anticipation effects are especially a talking point following the findings of Gonçalves et al. (2022) about how hosts changed their behavior in Lisbon after regulation was announced and before it was enforced.

Graphically, the parallel trends assumption for the number of listings between Bordeaux and the control group is more salient than it is for reservation days. The exceptional unseasonality of STR activity in Bordeaux in the sample, however, makes the parallel trends assumption a lot more difficult to assess graphically for reservation days. The DID

[14] https://objectifaquitaine.latribune.fr/business/2017-01-06/bordeaux-paris-ce-que-la-lgv-va-changer.html

method is aggregative by design, which means this work will be looking at differences in averages between pre- and post-regulation periods. Since exactly 2 years both before and after regulation are used, there is no direct threat of seasonality toward the estimators. However, seasonality remains an indirect threat to its validity in two ways. Firstly, it affects the reader's ability to assess the parallel pre-trends. If these pre-trends hold, it is a lot likelier that they would have held post-regulation too. With seasonality, it is harder to tell if Bordeaux was trending in the same way as the control group pre-regulation and if it therefore should be expected to have trended in the same way afterward if it had not been regulated. Secondly, there is a risk that time-varying confounding factors that could be season-dependent, like the summer weather, might impact groups differently.



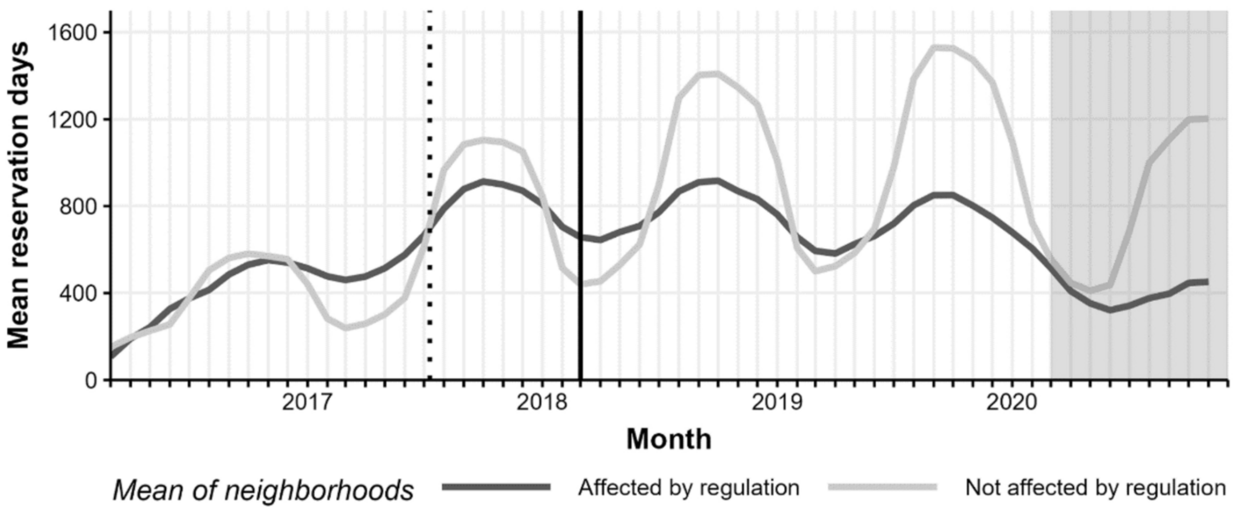


Figure 1.2: Mean number of reservation days per month (bottom) and of the number of active listings (top) for treated (Bordeaux) districts and control group districts. 6-month rolling averages are used to attenuate seasonality. The dotted vertical line indicates the announcement of regulation in Bordeaux (2017-07-10) while the regular vertical line represents regulation enforcement (2018-03-01). The gray area indicates the beginning of COVID-related travel restrictions, a timeframe which is ignored in this analysis.

The first problem is a control group selection issue, as discussed above. Control group selection in quasi-natural experiments is always tricky, and I understand that the spatial, structural and administrative proximity between Bordeaux and the control group cities might invite for further convincing. As such, an event-study-following methodology is included and described in Clarke and Tapia-Schythe (2021) that uses yearly rolling averages in Appendix 1.7.3. I find no evidence of significant pre-regulation differences in STR activity between Bordeaux and the control group cities once covariates and fixed effects are controlled for. This helps the DID estimator gain credibility (Roth et al., 2023) as it means groups did not appear to be different given what was observed before regulation. Despite this, the threat of unobserved time-varying confounding factors subsists. The way it is resolved is by using the TD model to supplement the DID, where validity would be very unlikely to be threatened by the same time-varying confounding factors the DID is. See the Triple difference section for further explanation.

Descriptives

As noted in the previous subsection, and as with most experiments, the DID setting requires the control group and the treated group to be structurally homogeneous. Heterogeneous units across groups would make it difficult to attribute any effect to the regulation, rather than to different STR evolution for different types of cities. A few control group observations might be more rural than any unit one could find in Bordeaux and slightly skew some statistics, while other districts could be more tourism-intensive than the center of Bordeaux. As shown in Table 1.2, the 76 districts from the control group are reasonably similar on average to Bordeaux's 88 districts, both structurally and in terms of STR activity prior to regulation. Bordeaux is both slightly richer and slightly more populated by primary residents on average than the coastal cities. Treated districts are diverse in terms of STR activity, as estimated by the standard deviation of reservation days, but the same can be said about control group units. This is consistent with the idea that STR accommodation tends to be organized in hubs around cities, near tourist landmarks and in trendy neighborhoods.

	Bordeaux districts				Control group districts			
	N	**Mean**	**SD**	**Listings**	**N**	**Mean**	**SD**	**Listings**
Reservation days	88	596.34	569.82	439.10	76	567.80	660.08	304.54
Share of targeted listings	88	0.69	0.13	0.69	76	0.74	0.17	0.76
Share of Primary residences	88	0.89	0.04	0.90	76	0.81	0.18	0.87
Median income (euros)	88	22,284.35	4756.86	22,402	76	20,976.96	3696.19	21,016

Reservation days are a sum per district, "Share of Targeted" listings use this chapter's definition of "Targeted". Share of primary residences is the number of primary residences (as opposed to secondary or vacant) over the total number of dwellings Data from *INSEE* and *AirDNA*.

Table 1.2: Monthly STR activity and structural summary statistics of treatment (Bordeaux) and control group (Anglet, Bayonne, Biarritz, La Rochelle) districts, from March 2016 to March 2018 (Before regulation).

1.3.3 Triple difference

Identification

A triple difference (TD) design is employed along with the DID to make up for some of its shortcomings and to provide further insight. The TD is very close to the DID in the way it is designed, but it is important to point out where they are different. Unlike the DID, it will not allow for the measure of aggregate effect of regulation since just a subsection of the dataset is used (district in the outskirt of Bordeaux with only listings that are between 250 and 250 meters away from the border). It can provide further evidence of causal impact and allow for a more nuanced reading of results across the city thanks to the sectioned dataset. The fact that this design relies on observations close to the border yields results that hold for peripheral, less-tourism oriented neighborhoods, which could be disproportionately threatened by spikes in touristic activity.

Table 1.3 is a small representation of the differences in composition between the two groups of the Bordeaux area (those within Bordeaux, and those just outside of it). Districts that are included in the TD, that is, those that host at least one listing within 250 and 250 meters of the border, are different to those that do not fill these criteria but are still found in the DID setting. They have more primary residences as a share of total housing, listings are less likely to be targeted, and they host less tourism-oriented amenities.

	Central districts				Peripheral districts (used in the triple difference)			
	N	**Mean**	**Median**	**Max**	**N**	**Mean**	**Median**	**Max**
Share of targeted listings	27	0.77	0.69	0.90	61	0.66	0.68	1
Share of primary residences	27	56.41	53	172	61	12.72	10	36
Number of restaurants	27	56.41	53	172	61	12.72	10	36
Number of hotels	27	1.44	0	10	61	0.59	0	11

Central districts are those that are no closer than 250 meters to the border. The triple difference design excludes these central districts and focuses on the more residential (peripheral) ones. Data from *INSEE* and *AirDNA*.

Table 1.3: Monthly STR activity and structural summary statistics of districts used in the TD, which highlight more residential-oriented peripheral districts.

Instead of comparing cities to one another, the TD compares the difference between Bordeaux and its close periphery to the difference between the control group cities and their own close periphery. In that sense, it is a DID between two DIDs. I leverage the same control group cities as in the DID setting (Anglet, Bayonne, Biarritz, and La Rochelle).

The model specification is very similar to that of the DID, but it incorporates additional interaction terms since the TD implies a new subclassification of units which corresponds to them being inside or outside their respective cities:

$$\text{STR}_{i,c,b,t} = \alpha + \beta_{TD} \left(T_t \times D_c \times B_b \right) + \left(C_c \times B_b \right) + \left(M_t \times B_b \right) + \left(M_t \times C_c \right) + \ \ldots$$
$$\ldots \ \ \text{Month}_t + \text{City}_c + X_i + \varepsilon_{i,c,t} \tag{1.2}$$

B_b is the group unit i belongs to, it is equal to 1 if the district is inside of the city's border (meaning inside Bordeaux, La Rochelle, Biarritz, Bayonne or Anglet) and to 0 if not (meaning close to the border but outside these cities). D_c is therefore not a treatment dummy alone: it denotes whether a district belongs to the Bordeaux area (1) or any control area (0). A treated district is one for which both D_c and B_b are both equal to 1. The dummy T_t takes the value 1 when month t is post-treatment, and 0 otherwise. The TD estimator β_{TD} is therefore a coefficient for districts inside of Bordeaux after regulation. Month and city fixed effects M_t and C_c are interacted with each other and with the border dummy B_b independently, so that the model is saturated (Olden and Møen, 2022). Controls X_i include the number of restaurants, the number of hotels and the number of inhabitants in the district. Income data is unavailable for many of the districts and are thus not included. $\text{STR}_{i,c,b,t}$ is still the number of reservation days or of listings in district i of area c on the side of the border b during month t.

In the TD, keeping only listings that are close to the border on either side makes them more comparable to each other. They are thus also more comparable in their receptivity to potential shocks. However, it is important that I avoid pollution to the dataset that could stem from scrambled listings. As a reminder, spatial scrambling can go up to 250 meters. This is why only listings that are at least 250 meters away from the border can be kept, to make sure they are not assigned to the wrong group. Similarly, I want to make sure I do not select listings that are too far away from the border. A 1-kilometer bandwidth seemed appropriate, and the data scrambling forced for the bandwidth to be stopped at 250 meters to exclude any listing that could extend beyond 1 kilometer of the border. Figure 1.3 plots the listings that are filtered into the triple difference analyses as well as the districts they belong to and to which they are aggregated. The maps show that STR activity is less consistently dense around control group cities' borders than it is in Bordeaux. The general diffusion of economic activity is very continuous beyond Bordeaux as the surrounding cities share the same urban area, which is further evidenced by how far the tramway lines extend[15] and by the fact Bordeaux's ring road goes around those cities. This difference between Bordeaux and control group cities is, however, not necessarily a problem under the assumption that the evolution of the ratio of inside to outside activity should be parallel between Bordeaux and control group areas (Olden and Møen, 2022).

In this design, the key assumption is no longer that STR activity in Bordeaux would evolve in a parallel manner to the control group without regulation. Instead, I consider that listings on either side of the border of a city, but still close to it, operate in similar environments when it comes to shocks that could affect the area and should therefore react in similar ways. This means that listings inside and outside of Bordeaux (and similarly inside and outside of control cities), without regulation, should face the same challenges and opportunities as their counterparts on the other side of the border. In the absence of regulation, the difference between Bordeaux and its outskirts would have evolved in the same way the difference between control group cities and their outskirts has. Differences that are measured in the evolution of that balance are therefore attributed to regulation in this setting. This means that this design does not rely on a parallel trends assumption of STR activity (either inside or outside the border) but rather on a parallel trends assumption of the ratio of inside to outside STR activity (Olden and Møen, 2022).

In other words, the TD is used here to make up for the threats to validity in the DID. Because it leverages a different set of assumptions, it is not affected by the same potential time-varying confounding factors the DID could be. This design is not interested in whether activity in the Bordeaux area grows at a faster or at a slower pace than in the control group —– This is allowed to happen. What STR activity is not allowed to do

[15]Figure at the following address: https://fr.maps-bordeaux.com/img/0/bordeaux-tram-carte.jpg
The white area is the municipality of Bordeaux, colored lines are tramway lines.

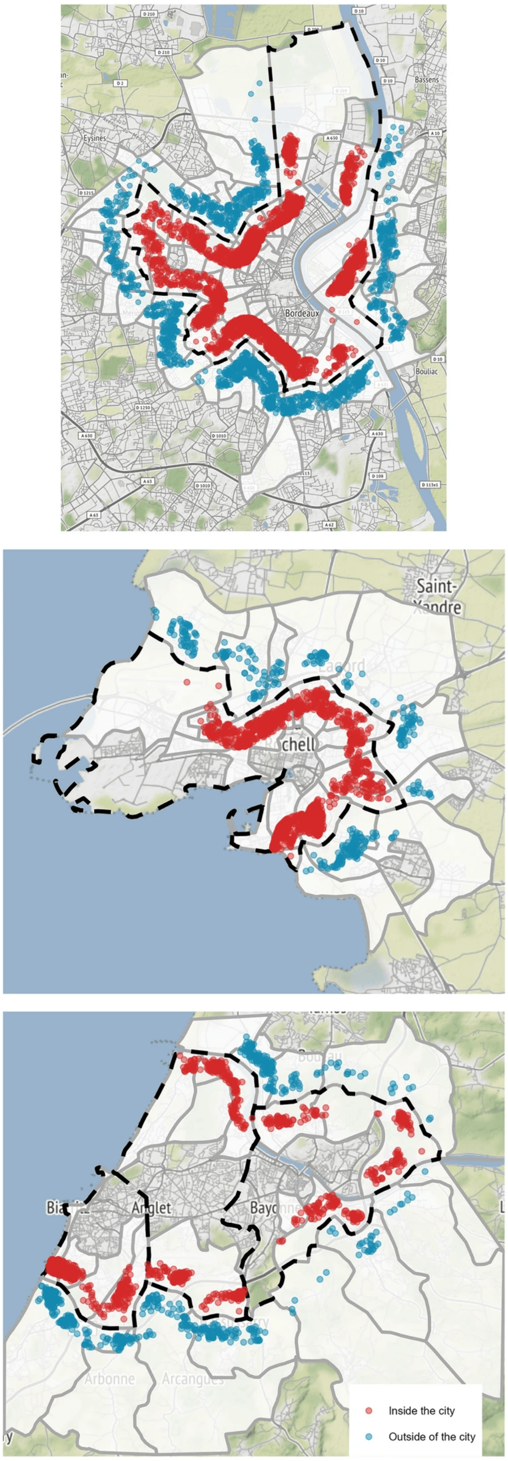


Figure 1.3: Maps of STR listings used in the triple difference design in Bordeaux (top), La Rochelle (middle) and the Biarritz-Anglet-Bayonne (BAB) area (bottom). All of these listings are within 250 and 250 meters of the border according to the dataset. Red listings belong to the "inside" group and blue listings to the "outside" group. Dashed lines are city borders. In the BAB area, I only use the common outside borders of the three cities. White areas are the districts at which listings are aggregated, with gray borders separating them.
Map tiles: ©OpenStreetMap Contributors

in this setting is to (had there been no regulation) change its allocation on one side or another of the border differently in Bordeaux than it does in the control group. This means that a shock such as the arrival of high-speed trains in Bordeaux is no longer a threat to validity so long as it does not only affect STR listings inside the border, which seems highly unlikely. Likewise, a hot summer that encourages tourists to opt for sea-side resorts instead of big cities has no reason to impact one side of either city's border more than another.

Figure 1.4 serves to visually evaluate the parallel pre-trends of STR activity inside-to-outside ratios for the two main variables of interest. Note that the data are once again aggregated at the district level in order to be able to compute an effect on the number of listings. These plots exhibit a durable gap between the ratio in Bordeaux and in the control group, which seems to have appeared after the regulation. More importantly, they strongly suggest both a common trend for the inside/outside activity ratio in the 2 years leading up to the regulation and a common structure to the data around Bordeaux and around other cities, since the ratios lie within very close values prior to the regulation. The lack of seasonality on these graphs is very encouraging toward the capacity of the TD to make up for some problems the DID might have, and the parallel pre-trends are graphically quite clear.

1.4 Results

1.4.1 Difference-in-differences

The first set of findings of the DID regressions are produced in Table 1.4. The number of observations N reflects the number of districts times the number of months in the sample. The lack of median income data in some places leads to fewer observations in regressions that include covariates. Regulation in Bordeaux has, on average for any given district, had a strong negative impact on reservation days. Results are robust to covariate

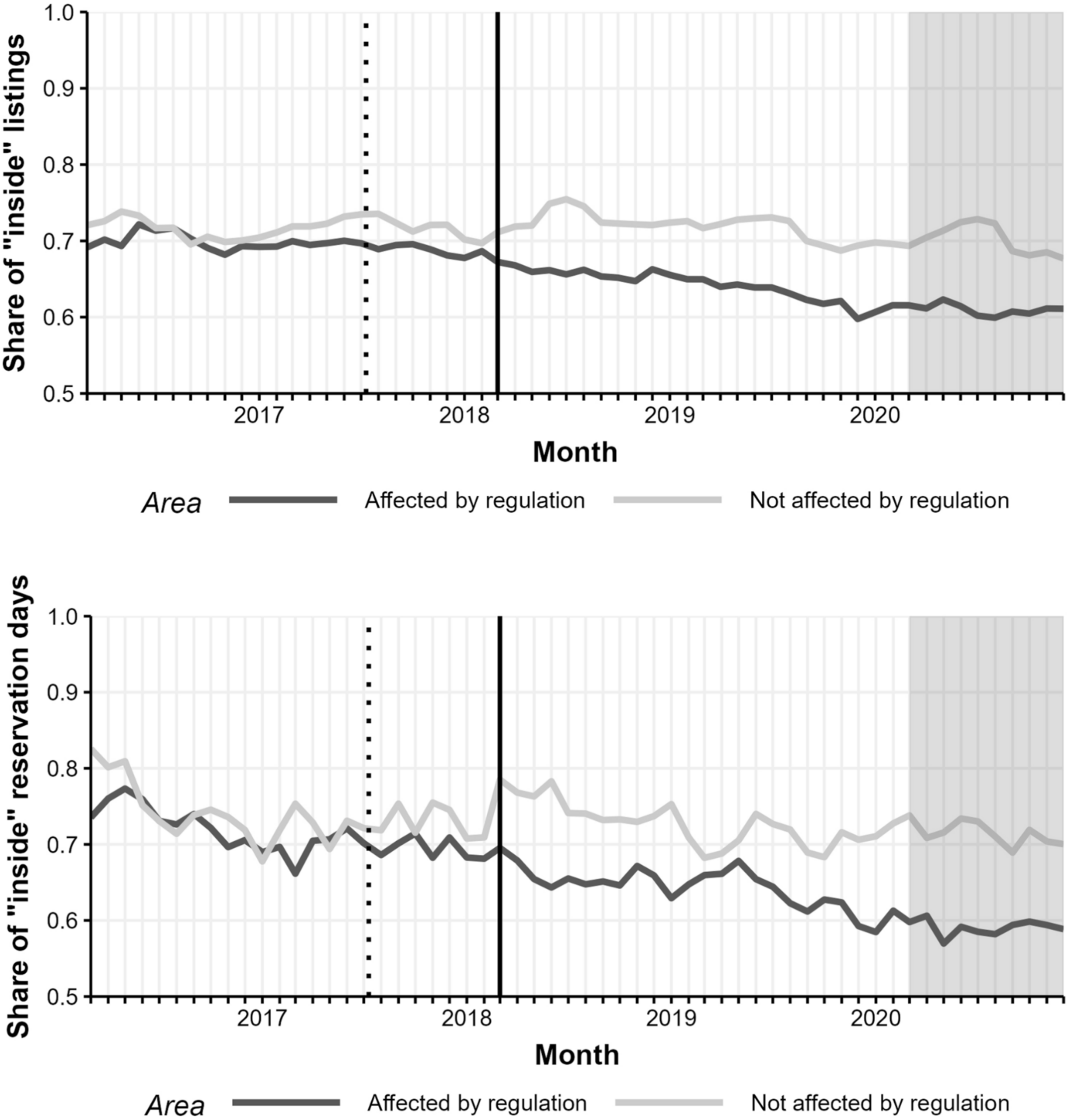


Figure 1.4: Share of STR activity (number of listings and reservation days) that is inside of cities' borders for Bordeaux (affected by regulation, black line) and for control group cities (not affected, gray line) as opposed to "outside" of them, respectively. Only listings that are within 250 and 250 meters of the border are used. The dotted vertical line indicates the announcement of regulation in Bordeaux (2017-07-10) while the straight vertical line represents regulation enforcement (2018-03-01). The gray area indicates the beginning of COVID-related travel restrictions, a timeframe which we ignore in our analysis

introduction both in terms of significance and in terms of scale, as shown in the preferred regression (column 2). Its DID estimator coefficient yields a net negative effect of 322 days. On average, for any Bordeaux district, regulation has had a negative impact of over half of a pre-regulation standard deviation (just under half of a post-treatment standard deviation) on the number of nights spent per month over the 2 years following treatment. Over Bordeaux's 88 districts, regulation has had a cumulative negative impact of 28,352 reservation days per month according to the DID estimation.

The effect of the regulation can also be observed in the number of listings which have strongly decreased (column 3), representing on average 51 fewer listings in Bordeaux districts per month in the post-regulation period. Interestingly the reservation days per listing in Bordeaux have increased, suggesting that while the global STR activities have declined, accommodations which remain in Bordeaux after regulation have experienced an increase in their reservations (column 4).

Table 1.4: difference-in-differences regressions results

	Reservation Days		Number of Listings	Res. Days per Listing	Number of New Entrants	Number of Withdrawals
	(1)	(2)	(3)	(4)	(5)	(6)
After Enforcement × Treatment units	−298.94* [124.02]	−322.18** [123.97]	−51.05** [18.38]	1.13*** [0.227]	−1.89** [0.588]	0.427** [0.135]
Effect as a % of mean	−40.74%	−43.91%	−63.95%	12.37%	−51.91%	22.94%
Covariates		x	x	x	x	x
Year Fixed effects	x	x	x	x	x	x
Pair Fixed effects	x	x	x	x	x	x
N	7872	7680	7680	7663	7680	7680
Districts	164	160	160	160	160	160
Adjusted R^2	0.25	0.52	0.65	0.52	0.32	0.11
F-Statistic	51.75*** (df=52; 7819)	152.27*** (df=55; 7624)	258.30*** (df=55; 7624)	151.53*** (df=55; 7607)	67.62*** (df=55; 7624)	18.41*** (df=55; 7624)

*p<0.05; **p<0.01; ***p<0.001

Effect as a % of mean refers to the Mean in treated (Bordeaux) districts after regulation; i.e., in districts subject to the DID estimator.

Reservation days per listing data are unbalanced, as I am unable to assign values for a given month in districts that do not have any listings for that month.

Introduction of controls affects the number of districts because income data is missing for 4 of the districts.

The difference-in-difference interaction for each dependent variable, noted by "After Enforcement × Treatment units" in the table, is the estimator. The general specification can be found in the "Methodology" section. Standard errors are clustered at the city level and controls include median income in the district as well as the number of restaurants and the number of hotels per inhabitant.

Reading (2): Bordeaux's districts had 322.18 fewer STR reservation days per month on average than they would have had without regulation. This equates to 43.9% percent of both the mean STR activity in treated districts, and to 688 thousand nights over the city and the treated period

As expected and because the regulation mainly aims to increase the cost of entering the STR market, the number of new entrants has decreased following the regulation. 1.89 fewer new listings per month are observed in treatment districts, which is about half of the usual number of new entrants. However, it cannot be ruled out that the decrease in the number of entrants can also be explained by the maturity of the STR market in Bordeaux. Over time, fewer accommodations are available to enter the STR market. Thus, the increase of the number of withdrawals could be a better indicator of the effect of the regulation as it reflects those which were leasing accommodation before the regulation and stopped in any given post-regulation month.

Table 1.5 identifies which kinds of listings have been impacted by the regulation. As a reminder, regulation in Bordeaux was targeted at hosts that rented their dwelling on a STR platform enough for it not to be considered their primary residence anymore. "Targeted" listings as presented in Table 1.5 are entire home listings that are active for at least 10 days per month on average over at least 6 months (120 days a year). "Not targeted" listings are those that are not a part of the "Targeted" group, either because they are not entire homes or because they are not active enough. "Entire homes" as the name indicates only includes entire home listings regardless of their targeted status. "No entire homes" only includes private or shared rooms, which cannot be a part of the targeted group by definition. As expected, those which were targeted have experienced the most important decline in the number of reservation (column 1) days but more importantly in the number of available listings. The decline of 47 listings per month in the post-regulation period represents almost one standard deviation and 87% of the mean number of listings. More surprisingly and unexpectedly given the regulatory design, non-targeted accommodations have also been affected by the compensation rule, even if they have been to a lesser extent given the 1500 monthly reservation days and 315 listings reduction. While it is difficult to explain this unexpected result, it is plausible that some hosts did not perfectly understand the compensation rule and have preferred to stop leasing on the STR market. Regardless of the reason, this outlines one of the regulation's shortcomings.

Another potential effect of the regulation is that it could have increased the incentive to lease just a room, as the limit of the 120 days matters only to entire home. However, the entire effect of the regulation seems supported by entire homes, as no increase is observed for shared rooms. This goes against the hypothesis that the loss in professional activity might be compensated by compliant home-sharing activity. If anything, compliant home-sharing activity has also decreased. These results are also robust in scale and in significance to changes in the control group, to anticipation effects[16] and to

[16] With the regulation being announced just under 8 months before it was actually enforced, some soon-to-be non-compliant hosts were able to change the use of their dwelling for free as long as they did it before March 1st, 2018. A double DID is deployed in Appendix 1.7.4.

	Reservation Days				Number of Listings			
	Targeted	Not Targeted	Entire Homes	No Entire Homes	Targeted	Not Targeted	Entire Homes	No Entire Homes
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
After Enforcement × Treatment units	−295.73* [121.15]	−26.44** [8.69]	−322.72** [124.68]	0.55 [3.21]	−47.47** [18.01]	−3.58*** [0.85]	−50.92** [18.51]	0.13 [0.36]
Effect as a % of mean	−57.30%	−12.15%	−54.27%	0.39%	−87.24%	−14.10%	−82.15%	−0.75%
Covariates	x	x	x	x	x	x	x	x
Year Fixed effects	x	x	x	x	x	x	x	x
Pair Fixed effects	x	x	x	x	x	x	x	x
N	7680	7680	7680	7663	7680	7680	7680	7680
Districts	160	160	160	160	160	160	160	160
Adjusted R^2	0.53	0.38	0.52	0.29	0.68	0.25	0.67	0.13
F-Statistic	157.06*** (df=55; 7624)	72.42*** (df=55; 7624)	149.96*** (df=55; 7624)	55.99*** (df=55; 7624)	290.18*** (df=55; 7624)	33.22*** (df=55; 7624)	279.04*** (df=55; 7624)	20.33*** (df=55; 7624)

*p<0.05; **p<0.01; ***p<0.001

Effect as a % of mean refers to the Mean in treated (Bordeaux) districts after regulation; i.e., in districts subject to the DID estimator.

Standard errors are clustered at the city level.

Reading (5): A district that was affected by treatment hosts 47.47 fewer targeted listings on average than a district that was not affected, in the two years following the treatment.

Table 1.5: Additional specific difference-in-differences regressions cases

COVID-related changes. More details on these different robustness checks can be found in Appendix 1.7.4.

1.4.2 Triple difference

The results yielded by the triple difference in Table 1.6 confirm and complete the DID's results by showing strong evidence of an effect of regulation on reservation days, as well as on the number of listings in the city's peripheral districts. All regressions include clustered standard errors at the area (Bordeaux, La Rochelle or Biarritz/Bayonne/Anglet) level.

Table 1.6: Triple difference regressions results

	Reservation Days		Number of Listings	Res. Days per Listing	Number of New Entrants	Number of Withdrawals
	(1)	(2)	(3)	(4)	(5)	(6)
After Enforcement × Bordeaux area × Inside the city	−47.57*** [14.77]	−47.57*** [13.88]	−7.79*** [1.53]	0.56 [0.44]	−0.39*** [0.18]	0.01 [0.06]
Effect as a % of mean	−35.33%	−35.33%	−53.9%	5.95%	−60.18%	1.99%
Covariates		x	x	x	x	x
Year Fixed effects	x	x	x	x	x	x
Pair Fixed effects	x	x	x	x	x	x
N	8064	8064	8064	8064	8064	8064
Districts	168	168	168	168	168	168
Adjusted R^2	0.22	0.31	0.27	0.26	0.23	0.09
F-Statistic	12.50*** (df=194; 7869)	19.19*** (df=197; 7866)	16.51*** (df=197; 7866)	14.94*** (df=197; 7423)	13.35*** (df=197; 7866)	4.91*** (df=197; 7866)

*p<0.05; **p<0.01; ***p<0.001

Effect as a % of mean refers to the Mean in treated (Bordeaux) districts after regulation; i.e., in districts subject to the DID estimator.

Reservation days per listing data are unbalanced, as I am unable to assign values for a given month in districts that do not have any listings for that month.

The triple difference interaction for each dependent variable, noted by "After Enforcement x Bordeaux area x Inside the city" in the table, is the estimator. Standard errors are clustered at the area level and controls include population in the district as well as the number of restaurants and the number of hotels per inhabitant.

Reading (2): Bordeaux's districts had 47.57 fewer STR reservation days per month, relative to outside districts (in the 500m radius of interest), on average than they would have had without regulation. This equates to 35.33% percent of the mean STR activity in those districts, within that radius.

As a reminder, the triple difference design compares the difference between Bordeaux's peripheral districts and the districts in the adjacent municipalities to the difference between the peripheral districts in the control group cities and the districts of their adjacent municipalities. Specification (2) is the most general form and is this chapter's preferred specification for the general result of regulation on STR activity in peripheral districts.

Any coefficient here should be read quite carefully. Observations are for one district for one month, but these do not paint the full picture of the districts given the selected bandwidth for listings (250–750 m). As such, interpretation of coefficients alone can be quite tricky and does not allow for a direct comparison to the DID's results. The "Effect as a % of mean" line which should help the reader in that aspect.

Overall and using the coefficients as a share of the existing mean, the results are slightly weaker in the triple difference setting than in the DID for the number of reservation days and of listings, but they push in the same direction and remain on the same scale: I respectively find 35 and 54% reductions in reservation days and number of listings, compared to the DID's 44 and 64%. If anything, these serve to strengthen the main results of the DID, showing that they hold up well even with different assumptions and outside of the very active central districts. Interestingly, the concentration of activity found in the DID with the "Reservation days per listing" variable is not replicated here, nor is the number of withdrawals. However, the scale of the estimator for the number of new entrants is even stronger than in the DID. It strongly suggests that the effect of the regulation could be different in the peripheral district as it seems that accommodations which have remained active did not incur any decrease in STR activity. This decrease is more driven by the fewer accommodations that have entered the market than by those that have left it.

Much like it has been done in the DID, the most relevant triple difference regressions are replicated on sectioned parts of the data in Table 1.7 to investigate how different groups of hosts were affected by the regulation. Targeted listings are, as expected and as was found in the DID setting, still disproportionately affected by regulation especially when it comes to the number of listings. However, just as in the DID estimates, the presence of peripheral non-targeted listings (i.e., home-sharing hosts) seems to have also suffered from regulation (regression (6)). The reduction in their activity is weaker than that of targeted listings, but it remains significant at the 1% level.

Another important result of the TD estimates is that it compares accommodations in districts inside and outside Bordeaux from both sides of the city border. The fact that the regulation impacts reservations days and listings inside Bordeaux as compared to those districts outside Bordeaux shows that the compensation rule creates a change in the distribution of STR activity at the city border, probably resulting in external effects with the municipalities adjacent to Bordeaux but not concerned by the STR regulation. In other words, there is a threat of an under-estimation of the effect in this subsection, because regulation in Bordeaux might have had positive effects on STR activity just outside of it.

	Reservation Days				Number of Listings			
	Targeted	Not Targeted	Entire Homes	No Entire Homes	Targeted	Not Targeted	Entire Homes	No Entire Homes
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
After Enforcement × Bordeaux area × Inside the city	−46.18*** [11.13]	−4.43 [4.45]	−50.45*** [12.61]	−0.99 [3.44]	−7.38*** [1.27]	−0.91*** [0.42]	−7.87*** [1.37]	−0.45** [0.35]
Effect as a % of mean	−33.43%	−11.79%	−33.98%	−8.9%	−26.54%	−13.77%	−29.65%	−5.19%
Covariates	x	x	x	x	x	x	x	x
Year Fixed effects	x	x	x	x	x	x	x	x
Pair Fixed effects	x	x	x	x	x	x	x	x
N	7824	7776	7824	7680	7824	7776	7824	7680
Districts	168	168	168	168	168	168	168	168
Adjusted R^2	0.31	0.22	0.31	0.17	0.30	0.18	0.30	0.14
F-Statistic	18.98*** (df=197; 7626)	12.12*** (df=197; 7578)	19.03*** (df=197; 7626)	8.93*** (df=197; 7482)	17.85*** (df=197; 7626)	9.46*** (df=197; 7578)	17.77*** (df=197; 7626)	7.25*** (df=197; 7482)

*p<0.05; **p<0.01; ***p<0.001

Effect as a % of mean refers to the Mean in treated (Bordeaux) districts after regulation; i.e., in districts subject to the TD estimator.

Standard errors are clustered at the area level.

I am interested in the triple difference interaction for each dependent variable, noted by "After Enforcement x Bordeaux area x Inside the city" in the table.

Table 1.7: Additional specific triple difference regressions cases

1.5 Discussion

The results presented in the previous sections have multiple implications. Firstly, both models point toward a decline in the number of reservation days induced by regulation. Reservation days are relevant to a number of objectives set by the city, and results are thus consistent with some of these objectives. The DID setting points toward an aggregate impact across the city of just under half of a pre-regulation standard deviation. The fact that the effect is still significant in scale in the periphery (as estimated in the TD) is also paramount, given the structural nature of peripheral neighborhoods and their propensity to be transformed by high numbers of short-term stayers. If reducing the number of reservation days (especially those who were targeted) was the sole criterion for evaluating regulation in Bordeaux, then it would have to be deemed a success and the net effect would be positive. In some sense, some *Airbnb* tourism clusters could have deflated, and regulation might have limited a potential expansion of the tourism bubble induced by STR (Ioannides et al., 2019). Reservation days are also decomposed into the number of listings and their occupancy, that is, the number of reservation days per listing. The effect of the regulation is stronger on the number of listings than on reservation days. With a change in the supply of STR but no reason for a decrease in total demand, the occupancy of listings that remain on the market has gone up. The DID suggests that in central neighborhoods, regulation has led to a concentration effect that has benefited incumbent targeted listings. However, the DID and the TD coefficients cannot be directly compared, given the fact they come from separate designs using separate data and are subject to different assumptions.

Secondly, the distinction between targeted and non-targeted hosts enables a look at the different effects of regulation according to the motivations of the hosts. As a reminder, Bordeaux's regulation was aimed at STR hosts that had a commercial use of their listing, that is, at listings that did not otherwise serve as the host, or anyone's, primary residence. The effect sought by these discriminatory measures is a reduction in activity exclusively driven by the hosts that are discriminated against, all the while keeping the benefits of home-sharing produced by hosts of primary residences. These benefits include a better use of accommodation resources as well as a better redistribution of tourism-generated revenue across the city. The regulator's decision to protect home-sharing showcases its importance. Results from both the DID and the TD setting however suggest that this implied objective was not exhaustively met, as regulation has had a reductive effect on the activity of non-targeted listings in addition to its impact on targeted listings. To put it simply, hosts who are following the tacit manifesto of the sharing economy ("Not Targeted") consisting in looking for additional income and meeting new persons are perhaps penalized. This is contrary to the expected outcome of such regulation, as a *de facto* ban on commercial STR should create market opportunities for home-sharers and therefore boost their activity or their number. The main plausible explanation is that regulation has had a direct effect on compliant hosts because of informational spillovers. It might be too complicated and bureaucratic, but institutional and media communication around the subject in Bordeaux could have also created a negative sentiment around STR hosting in the city.

The fact post-regulation differences in the triple difference design are observed also highlights threats posed by such city-wide regulatory efforts. The low spatial scale of the decision appears to push potential problems linked to STR activity slightly further away, rather than suppress them. It pushes them onto spaces that are an extension of Bordeaux's urban area and that likely share the same problems faced by residential neighborhoods on the city's inner outskirts. These externalities can be positive when they come along with a spread of income linked to tourism, new dynamism stimulating new localized amenities (restaurants, bakeries, etc.) but they can also be negative when nuisance, high rotation of guests, and increases in rent can push away some residents or damage their well-being.

Moreover, peripheral cities also share a political space and decision-making processes with Bordeaux regarding a variety of fields through the Métropole[17], which is relevant when it comes to urban planning policy-making. This is an institutionalized indication of the common interests Bordeaux has with neighboring cities, and it casts questions over the decision to regulate at the city level given the need for STR policies to be in line with the wider frame of housing and economic policies. Bordeaux, however, of course operates in its own specific context, and the control group is necessarily an imperfect counterfactual given the nature of quasi-experimental settings. This limits the external validity of this analysis. Still, the results this chapter outlined allow for opening up a discussion about one-size-fits-all regulation. In any case, a uniform and one-size-fits-all regulation, both in spatial terms, that is, related to geographical characteristics of neighborhood and in behavioral terms, that is, related to types of hosts seems only partially efficient to keep the control of urban transformations associated with *Airbnb*-like platforms. Indirect and undesirable effects also stem from this uniformity, and a more targeted or smart-regulation should probably be designed.

As a final note of discussion, let me go back to the initial reasons for which Bordeaux's regulation was initially voted by local policy-makers[18] in order to explore whether these goals are reasonably met or not. As far as competition with traditional tourism accommodation and pressure on housing prices go, existing literature points towards less STR activity having a positive effect (Valentin, 2020; Zervas et al., 2017). Less STR activity and more reservation days per listing also imply that the rotation of guests has gone down, in turn helping alleviate the degradation of properties and shared spaces.

However, and most importantly in the broader context of this thesis, the first STR-caused problem brought up in the official document is a "*transformation of neighborhoods*" that comes along with a "*disappearance of community life*" and a "*degradation of social*

[17]Bordeaux Métropole is an agglomeration of 28 different cities in the Bordeaux area. These cities yield some of their decision-making power to the consortium when it comes to certain policy issues https://www.bordeaux-metropole.fr/Metropole/Organisation-administrative/Competences-de-Bordeaux-Metropole)

[18]See Subsection 1.2, or https://www.bordeaux.fr/images/ebx/fr/CM/12461/12/acteCM/69313/pieceJointeSpec/148521/file/acte_00046189_D.pdf

capital". This issue is less clearly defined than the others in the literature. Its vague formulation is evidence of how difficult it would be to evaluate and quantify its relationship to STR, but the concepts it touches upon are at the heart of the urban tensions and crises STR epitomizes. Ioannides et al. (2019) touch on the relationship between *Airbnb* and neighborhood atmospheres, but it is difficult to imagine robust ways in which a practitioner could engage with issues of "neighborhood transformations", of "degradation of social capital" or of "disappearance of community life" in an empirical manner. The study of housing- and retail-oriented impacts of STR touches upon these issues in some aspect, but there is a lack of a more holistic approach that could better echo the sentiment behind regulation and behind the urban tensions that motivate it. Residents are not just unhappy because their rent is marginally growing, they are upset because the places they live in are changing. These changes happen in deeply systemic ways that would be impossible to estimate from individual characteristics (such as rent) alone.

1.6 Conclusion

The aim of this chapter was to assess the effect of the attempted regulation of short-term rental accommodations in Bordeaux, and confront its effect to its objectives. The so-called compensation rule provided an opportunity to evaluate the causal impact of the regulation on aggregate in the city but also on the different categories of hosts.

The main result shows that the number of nights spent in a district of Bordeaux has been reduced by 322 nights per month in the post-regulation period. While this effect is mainly driven by the most touristic and central district of the city, I also show using a triple difference design that the average peripheral district in Bordeaux lost 35% of their reservation days. It is then demonstrated that this decrease in the activity of short-term rental is also driven by those that were not *a priori* affected by the regulation, suggesting the existence of unintended effects in the consequences of the regulation. This research obviously has some limitations, the main one being that the results are place-dependent and contextualized. These results are also time-dependent and sensitive to the local institutions in charge of the regulation design, and of its practical enforcement. Further assessments are needed in various contexts and for more cities to understand what the interrelated external effects on the short-term rental market are.

The debate on regulating digital platforms is a concern for many local, national or continental governments facing many multi-dimensional but interrelated effects (on competition, on geography, on labor, on technology, on mobility, on society as a whole, etc.). There is a growing need to design and adapt better regulations (Frenken et al., 2015). This contribution is a first step toward designing what is suggested as a smart-regulation, considering that a one size fits all approach produces many undesired and ambivalent externalities. It lays the foundations for further work on the way planners in charge of urban transformations, digital platforms, local institutions and regulators should work hand in hand for mutual benefits. It however also begs further questions regarding regarding the unexpected effects of STR regulation on "by-the-manual" home-sharers. In addition, it highlights the need for new methods to understand urban crises STR are a part of through the difficulty in thoroughly confronting regulation efforts to their initial objectives, as diverse as they can be (Aguilera et al., 2021), notably when these objectives are focused on the urban experience of dwellers. A different and new approach is needed to assess transformations of urban spaces, and the rest of this thesis presents and expands upon a promising alternative to existing methods.

1.7 Appendices

1.7.1 Map of city locations

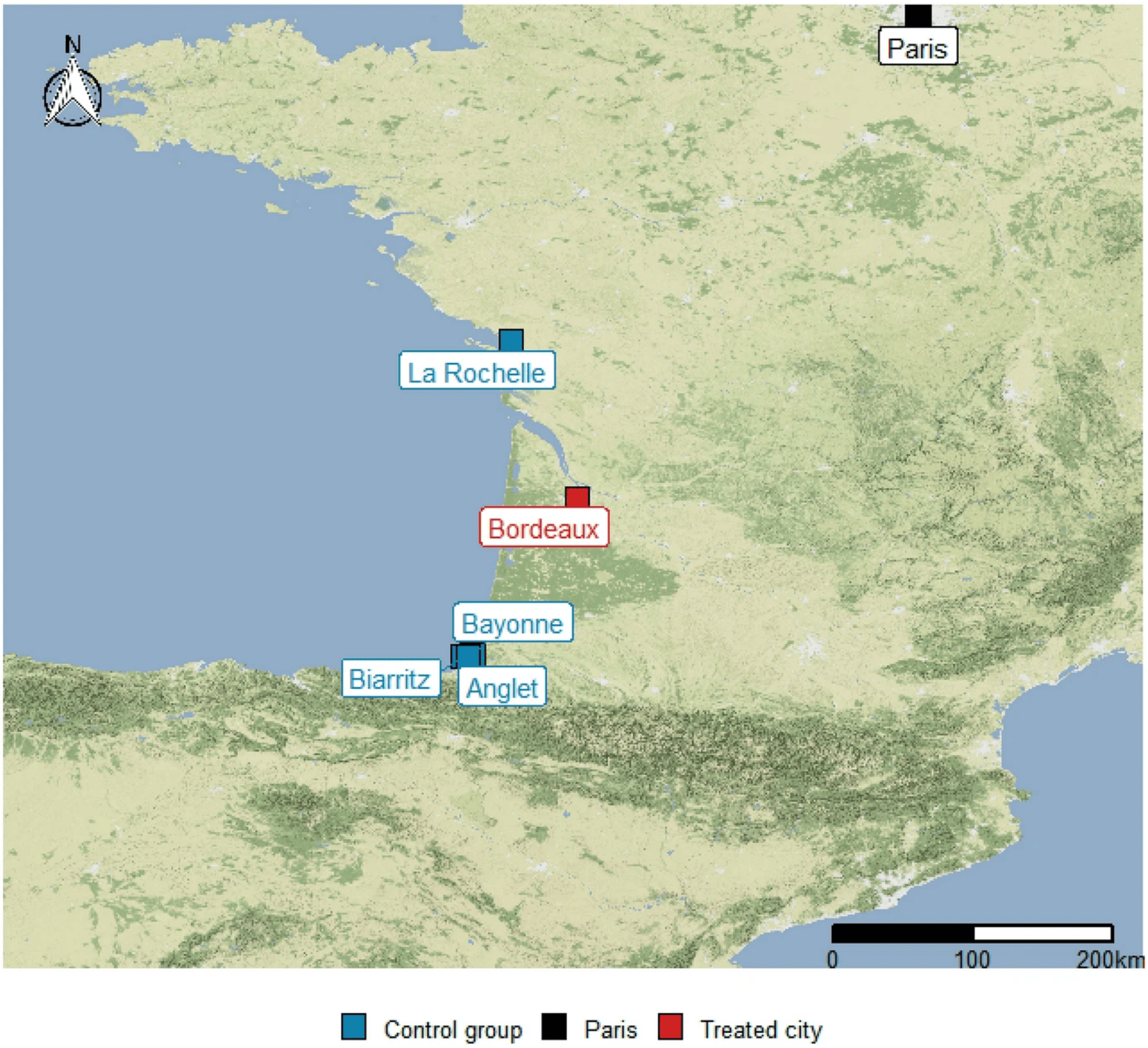


Figure 1.5: Bordeaux and the control group cities are located in sunny and touristic southwestern France. Map tiles: ©OpenStreetMap Contributors

1.7.2 STR in selected cities

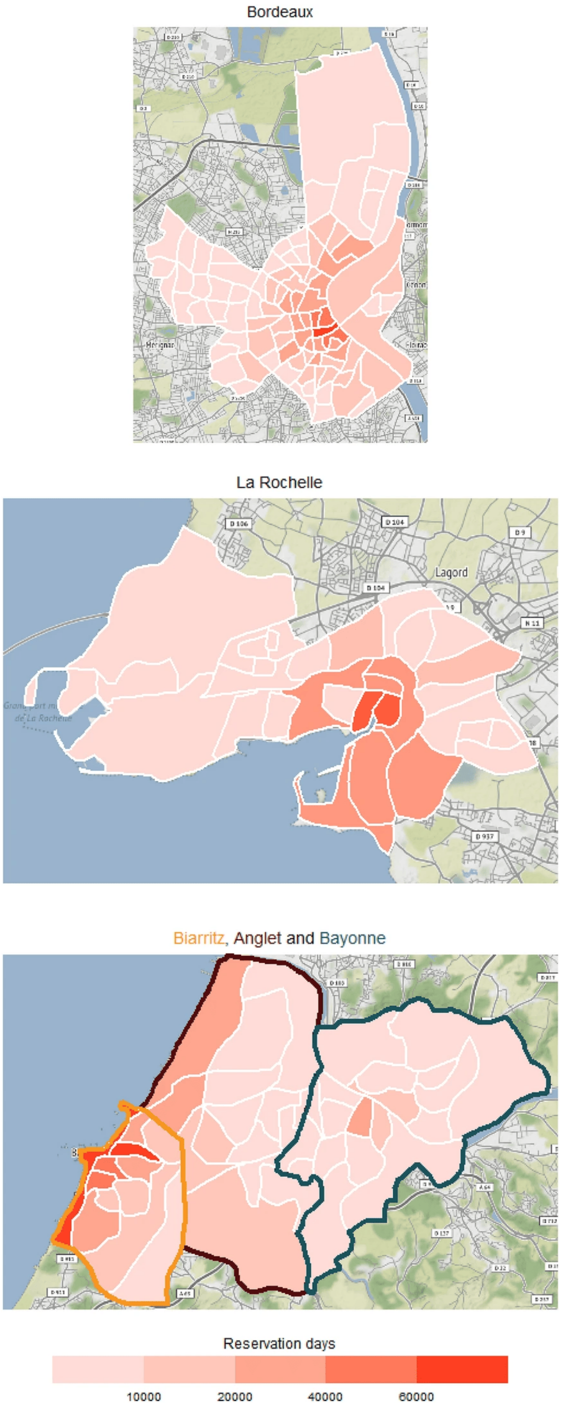


Figure 1.6: Districts of Bordeaux and of the control group cities, colored according to the STR reservation days they hosted in the two years leading up to regulation (March 2016–March 2018)
Map tiles: ©OpenStreetMap Contributors

1.7.3 Event study

I use the event study methodology detailed by Clarke and Tapia-Schythe (2021) to evaluate parallel pre-trends in the DID setting. This is very similar to the use of event studies in Miller et al. (2021). The idea behind this method is often useful to better understand the timing of effects; however, given the seasonality differences between Bordeaux and the control group cities in the data, these are still difficult to make out in this setting. To even out these seasonality issues while preserving the Clarke and Tapia-Schythe (2021) method, I use yearly rolling averages. This means that any lead or lag evaluated relative to the baseline is an average that includes the 11 previous periods. The event study regression's specification is given as follows:

$$\text{STR}_{ict} = \alpha + \sum_{j=2}^{j=12} \beta_j \left(\text{Lag}_j\right)_{it} + \sum_{k=1}^{k=24} \gamma_k \left(\text{Lead}_k\right)_{it} + \text{Month}_t + \text{City}_c + X_i + \varepsilon_{ict}$$

Just like in the main DID specification, STR_{ict} is the number of reservation nights or the number of listings for any given district i in city j during period t. Month_t and City_c are controls for month and city specific fixed effects and is a vector of covariates at the district level. j and k are lag and lead binary variables, reflecting how many months away from regulation (before with j or after with k) a given observation is. Since I am using yearly rolling averages, the 12 first months of the data is not used and there are only 12 lags. I use a standard baseline at $j = 1$. I am comparing deviation between treatment districts and control districts based on the month preceding regulation (February 2018). The event study regressions are only meant as a supportive supplement to the main DID model, and the regressions' estimates of interest are plotted in Figure 1.7.

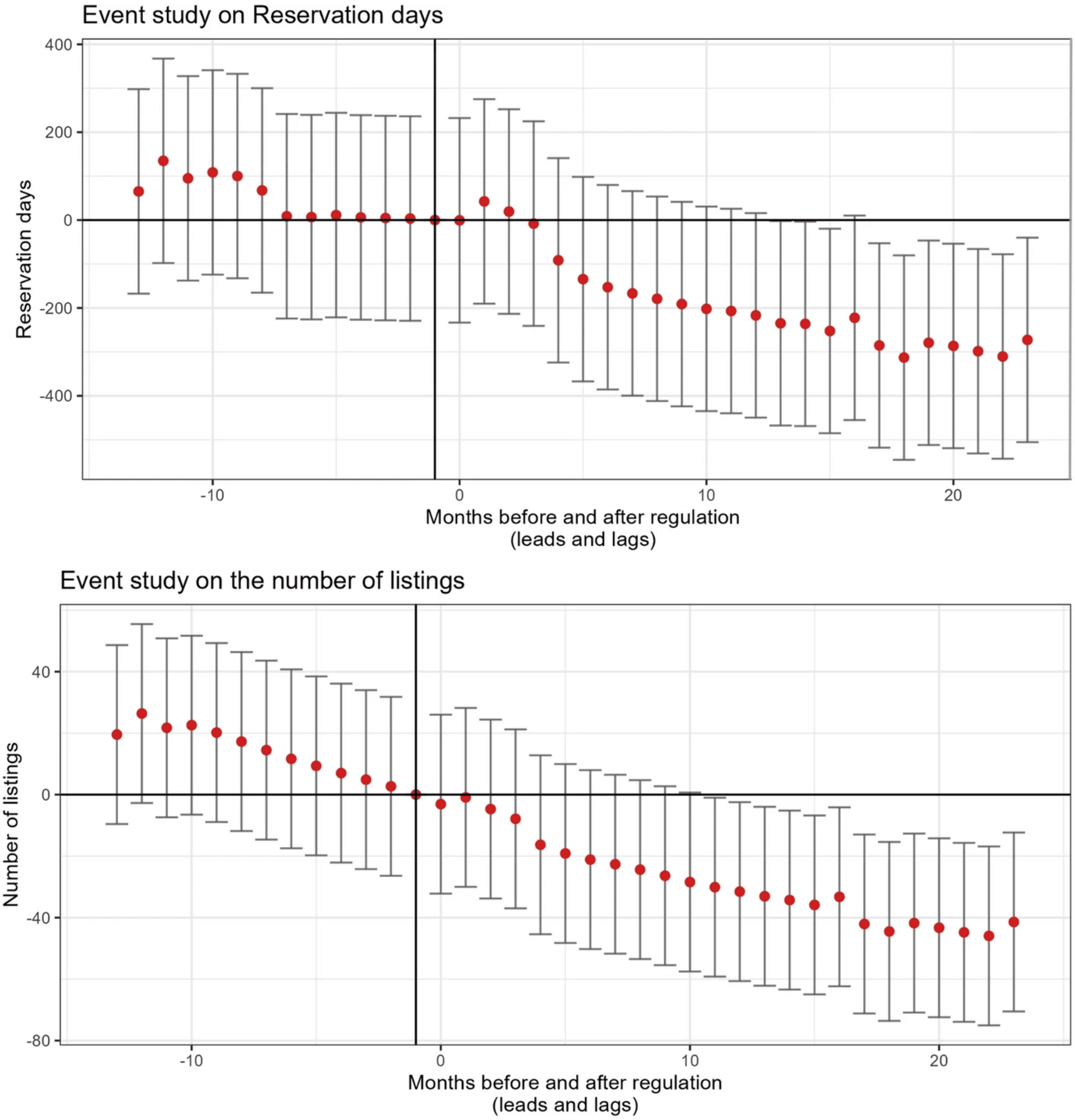


Figure 1.7: Event study plots of STR activity in the DID setting with yearly rolling averages following Clarke and Tapia-Schythe (2021). Red dots are the regression coefficients for the treatment's effect at a given lead/lag and gray bars are these coefficients' 95% confidence intervals

These two plots show that in the year leading up to regulation, there was never any significant difference (at the 95% confidence level) between the expected level of activity in treatment districts and in control group districts, conditioned on observed covariates and fixed effects. This is strong support toward the parallel trends assumption of the

DID setting. Before regulation, I was able to deduce the level of STR activity in Bordeaux using control group cities satisfyingly. The fact that the coefficients only become significant a few months after regulation is a direct result of yearly rolling averages. As a reminder, the first lead that does not include any pre-regulation data is lead number 11.

Event studies are usually apt tools for gaining a better understanding of the timing of a treatment's effect. In the context of the seasonality in the data, however, the timing is trickier to unpack. Given what Gonçalves et al. (2022) found for Lisbon, one might expect that targeted hosts have decided to enter the market at a higher rate between the regulation's announcement and its enforcement. This might well be the case in Bordeaux too, but it cannot be established on the evidence of the event study.

Significant negative effects also appear to become greater over time. This is quite intuitive, and expected, given the way regulation is designed. The more time goes by, the more potentially targeted new entrants are not entering the market when they otherwise would have.

Table 1.8: Tests regarding the difference-in-differences setting

	No COVID time restriction	Anticipation effect	Different control group	Placebo DID	ELAN law effect
	(1)	(2)	(3)	(4)	(5)
After Enforcement × Treated	−385.12*** [29.51]	−275.85*** [65.77]	−332.20* [143.90]	139.70 [95.40]	−305.01*** [40.72]
After Announcement × Treated		−65.40 [82.21]			
After ELAN × Treated					−29.42 [44.79]
Effect as a % of mean	−61.85%	−37.59%	−45.27%	18.16%	−41.57%
Covariates	x	x	x	x	x
Year Fixed effects	x	x	x	x	x
Pair Fixed effects	x	x	x	x	x
N	9280	7680	7056	3680	7680
Districts	160	160	147	160	160
Adjusted R^2	0.55	0.56	0.56	0.62	0.56
F-Statistic	177.3*** (df=65; 9214)	178.5*** (df=56; 7623)	166.7*** (df=55; 7000)	183.5*** (df=28; 3053)	178.5*** (df=56; 7623)

*p<0.05; **p<0.01; ***p<0.001

Effect as a % of mean refers to the Mean in treated (Bordeaux) districts after regulation; i.e., in districts subject to the DID estimator.

Standard errors are clustered at the city level.

1.7.4 Robustness tests

I lead the same robustness checks in both of the settings. They can be found in Table 1.8 for the DID setting and in Table 1.9 for the triple difference setting. As shown in regression (1), COVID-19 restrictions have not evened out the effects linked to regulation. In that model, I use data up until December 2020. If anything, it seems the gap between Bordeaux's districts and control group districts has gotten larger after COVID. Regression (2) intends to check for the existence of an anticipation effect through a twofold DID and finds no evidence of such an effect. As a reminder, the regulation was announced in July 2017, which is the cutoff for the "After Announcement x Treatment" DID estimator. That estimator being insignificant encourages the belief that hosts did not fundamentally change their ways before regulation was actually enforced.

Table 1.9: Tests regarding the triple difference setting

	No COVID time restriction	Anticipation effect	Different control group	Placebo DID	ELAN law effect
	(1)	(2)	(3)	(4)	(5)
After Enforcement × Bordeaux area × Inside the city	−46.65*** [0.53]	−40.43*** [5.70]	−19.66*** [0.00]	−1.24 [3.86]	−52.92*** [2.7]
After Announcement × Bordeaux area × Inside the city		−4.55 [5.74]			
After ELAN × Treated					15.89*** [1.89]
Effect as a % of mean	−39.89%	−30.02%	−14.6%	−1.01%	−39.30%
Covariates	x	x	x	x	x
Year Fixed effects	x	x	x	x	x
Pair Fixed effects	x	x	x	x	x
N	9860	8064	5904	3772	7621
Districts	170	168	123	164	168
Adjusted R^2	0.31	0.31	0.26	0.32	0.31
F-Statistic	19.57*** (df=237; 9622)	15.46*** (df=246; 7817)	14.97*** (df=148; 5755)	15.57*** (df=120; 3651)	15.47*** (df=246; 7817)

*p<0.05; **p<0.01; ***p<0.001

Effect as a % of mean refers to the Mean in treated (Bordeaux) districts after regulation; i.e., in districts subject to the TD estimator.

Standard errors are clustered at the are level.

Royan and Arcachon replace Biarritz and Anglet in regression (3), as the former are two of the other most important STR cities in the region. In regression (4), a placebo DID is manufactured by using a March 2016 to January 2018 dataset, with a fake treatment starting in March 2017. This serves to reinforce the parallel trends assumptions, as it shows no significant difference in development prior to the introduction of regulation. The ELAN law, deployed in November 2018, might also have had an effect on host behavior as potential fines went up under the new nation-wide law. I therefore present another twofold DID in order to confirm that the effect of regulation cannot be solely attributed to national regulation as opposed to city-level regulation. This is presented in regression (5). What is important in these models is that the "After Enforcement x Treatment" remains significant despite the introduction of a second DID estimator.

Chapter 2

Adapting the Economic Complexity Index to urban issues[1]

2.1 Introduction

This chapter answers some of the points discussed and concluded upon in Chapter 1, namely, the difficulty in assessing and regulating urban phenomena that stem from systemic processes. Here, I explore the view of cities as complex and interaction-driven systems, and confront this perspective to that of evolutionary economic geography. This chapter contributes to an interdisciplinary and evolutionary trend in urban planning that could help tackle in tackling urban crises such as short-term rentals. Evolutionary economic geography (Boschma and Lambooy, 1999; Kogler et al., 2023) and the paradigm of economic complexity (Balland et al., 2022; C. A. Hidalgo, 2021) are brought forward. I show how, in tandem, they can assist to uncover and explain urban issues as systemic phenomena through the former's frameworks and the latter's empirical tools.

The Economic Complexity Index (ECI), one of the paradigm of economic complexity's important methodological contributions, is a complicated measure but leveraging it through urban economic interactions can help assess cities in a reductive, holistic way. In turn, this can be useful for policy-makers to understand and address contemporary urban crises from a different point of view. I therefore propose an evolutionary but micro-

[1]This chapter is an iteration of an ongoing work co-written with DEJEAN Sylvain and SUIRE Raphaël: http://econ.geo.uu.nl/peeg/peeg2315.pdf. The co-authors have agreed to the following modifications and use.

founded market-driven framework within which to understand an ECI of cities, before detailing a slightly modified methodology and discussing the implementations of such an index when applied to cities and their commercial amenities. This index, the Amenity Complexity Index (ACI) as well as its counterpart, the amenity-Type Complexity Index (TCI), are then applied in Chapter 3 and expanded upon in Chapter 4.

2.1.1 Cities, evolutionary science and complex systems

Robert Moses, an eminent city planner in 1950's New York, faced strong opposition from Jane Jacobs (1961), who criticized his top-down approach to urban planning. Her work inspired later planning literature to emphasize the systemic and bottom-up nature of urban issues through an evolutionary turn (that is aptly described by Bertolini, 2010, and was pioneered in part by Pumain, 1997), and sometimes within complexity frameworks (Batty and Marshall, 2012; Boonstra and Boelens, 2011; Moroni, 2015; Pumain, 2003; Rauws et al., 2016). In this context, infrastructure, housing, jobs, innovation, education, and health involve interrelated dynamics among various agents with different objectives, which are difficult to observe and inform given the non-linear nature of their outcomes. The complex systems perspective has advanced the theoretical understanding of these dynamics, drawing from interdisciplinary sources (Allen, 1997; Anderson, 1972; Batty, 2000; Krugman, 1996; Nicolis and Prigogine, 1989). Concepts such as self-organization (Krugman, 1996; Moroni et al., 2020; Portugali, 1999) and flows (Batty and Cheshire, 2011) offer systemic explications of how cities emerge, adapt and evolve.

This work takes the systemic nature of cities for granted, but it seeks to build upon it in two separate but ultimately related ways. Firstly, current planning literature still leaves a gap in empirical measures available to city-level policy-makers to understand local-level (that is, inner-city) organization while taking into account the complex systemic nature of the processes that define it. Boeing (2018) provides a network-driven framework within which to unpack the macro-level complexity of cities by outlining existing empirical methods, but these urban design-oriented approaches struggle to provide holistic indices that would characterize sub-city systems. Other empirical contributions leverage mathematics-inspired complex systems to explain human organization (Batty, 2001; Zhong et al., 2014) that are not necessarily easily micro-founded. Moroni et al. (2020) still note the inefficiency of complex system theory when it comes to finding proactive and effective instruments, and complexity in planning is in truth seldom thought of in terms of outcomes in favour of systemic, process-focused emergence paradigms. To be clear, the argument here is not that these paradigms are not important or useful, but that outcome-led backward analysis of cities and of their systems could provide extra insights that are currently less covered.

Secondly, systemic planning approaches could benefit from even further interdisciplinarity. Frenken et al. (2007) define economic geography as "*[dealing]* with the uneven distribution of economic activity across space". Following the seminal work of Nelson and Winter (1985), economics and in turn economic geography have embraced evolutionary theory (Boschma and Lambooy, 1999; Kogler et al., 2023). As a matter of fact, the increasing recognition of the complexity of cities and of their systems in planning echoes evolutionary thinking in economic geography, albeit at different scales and with different approaches. As cited in Kogler et al. (2023), evolutionary economic geography (EEG) focuses on "*the processes by which the economic landscape–the spatial organisation of economic production, circulation, exchange, distribution and consumption–is transformed from within over time*" (Boschma and Martin, 2010). While some approaches take interest in the spatial unevenness of economic activity within cities (Boschma and Weltevreden, 2004; Spencer, 2017), there is still a clear gap within EEG in its assessment of changes, processes and organization in sub-city scale urban spaces.

Methodologically, a profusion of different theoretical and empirical endeavours have transpired throughout EEG (Kogler et al., 2023), of which one is of particular interest to this chapter. Darwin-inspired biological analogies to economic systems have paved the way for a general understanding of economic systems' evolution and transformations through processes of inheritance, variety, selection, and retention, and while these concepts were first adopted in evolutionary economics, they have since also been of interest to EEG research (Dopfer, 2005; Essletzbichler and Rigby, 2010; Hodgson, 2002).

On the one hand, a lack of intra-city empirical and systemic approaches provides an opportunity for an expansion of systemic understandings of cities. On the other hand, evolutionary economic geography, despite a paradigmatic and epistemological resemblance to systemic thinking in planning, provides a different economics-oriented tool set that it scarcely leveraged *within* cities and to tackle issues relating to urban life. This two-way opportunity, in a context of urban tensions and crises, motivates this work. The following section presents economic complexity as a branch of evolutionary economic geography that has gained strong momentum in the last decade, and from which I borrow the *Economic Complexity Index* to typologize urban spaces in a manner I motivate as holistic.

2.1.2 The Economic Complexity Paradigm

I focus on the modern network-based interpretation of economic complexity in evolutionary economic geography spheres (Hausmann et al., 2014; C. A. Hidalgo and Hausmann, 2009) that has gained much traction over the past 15 years (Balland et al., 2022; C. A. Hidalgo, 2021). In the same way systemic approaches to planning go against orthodox views of cities, economic complexity has come as a reaction to traditional approaches to growth (Balland et al., 2022; C. A. Hidalgo, 2021). As a branch of evolutionary economic geography, it is also heavily focused on the importance of systemic interrelations. In economic complexity, the quasi-infinite hidden interactions between interrelated economic agents lead to emergent and constantly out-of-equilibrium systems that evolve non-linearly. Economies are not viewed as the sum of their parts, but as a whole that holds collective properties. In practice, economic complexity supports the idea that "*Growth, development, technological change, income inequality, spatial disparities, and sustainability are the visible outcomes of hidden systemic interactions*" (Balland et al., 2022).

However, one key contribution of economic complexity that sets it apart from other evolutionary approaches (in planning and in economic geography) is how much the former has revolved around providing empirical methods to reduce the dimensionality of complicated problems. The premise is that one cannot predict outcomes through agents' interactions and characteristics in a genotypic way, that is, through the information contained within the processes, because these interactions are inherently too complex. The economic complexity literature takes a phenotypic approach by apprehending systemic interactions from the perspective of what can be observed, that is, by focusing on observables. Reducing the dimensionality of observables through a network perspective can help characterize processes that would otherwise be too complicated, and reveals latent properties of systems. The reduction of systemic interactions and interrelations in the field of economic complexity can be tightened down to two related but separate movements: Relatedness measures and Complexity indices.

Both of these tools are deeply intertwined. They rely on the concept of a product space, or network, in which countries, regions and cities all operate (C. A. Hidalgo et al., 2007). Measuring the relatedness between products in that space provides information about how similar these products are in the inputs, knowledge, and routines they require to be produced (C. A. Hidalgo, 2021). The combinations in places of localized events or attributes influences the likelihood of related events or attributes entering. These combinations can contribute to explaining path-dependent evolutions at various spatial scales (Boschma, 2017; Frenken et al., 2007; Neffke et al., 2011). The Economic Complexity Index (ECI), an index of places' leverage within the same product space, is the measure I translate to cities' amenities and is the focus of this chapter.

2.1.3 The ECI

Economic complexity assumes that all products are not equal within the product space; Their position in the space will depend on what kind of knowledge and capabilities they require. Its core strength however is that its formulation provides indices that are agnostic, *ex-ante*, to what those positions may be. Measuring a countries' economic complexity with the ECI is meant to measure the leverage it has in the product space, and thus its ability to generate future growth in a more complete way than traditional growth models based on aggregate inputs and total factor productivity (Hausmann et al., 2014; C. A. Hidalgo, 2021; C. A. Hidalgo and Hausmann, 2009). In economic complexity, differences in outcomes are due to the distinct kind of goods countries are able to produce.

C. A. Hidalgo and Hausmann (2009) (*HH* hereafter) presented a first version of the Economic Complexity Index. They argued that what countries are able to export relative to the rest of the world is representative of their relative skills and knowledge.

In practice, this index is an applied spectral method that solves a reduction of the global export network of countries and products, and that thus does not require prior assumptions about the relative positions of goods or of countries. Through this, they assign a given complexity index and rank to every country and every product. *HH* assume that products that are non-ubiquitous (that are rare, exported by few countries) and that are produced in diverse countries (that export many different and rare products) require a rich set of enviable capabilities, and that complexity measures reveal this. Effectively, the ECI is computed conjointly with the Product Complexity Index (PCI). The complexity of countries is reflective of the complexity of the products it exports competitively, and vice versa.

The philosophy and the strength of results yielded by the Economic Complexity Index, along with its desirable properties and its agnostic methodology, have granted it a strong stand in the modern economic geography discourse both methodologically and in its very flexible applications (Balland et al., 2022; C. A. Hidalgo, 2021). This empirical success has led the ECI (sometimes in tandem with relatedness measures) to inform policy in very direct ways (Balland et al., 2019; Pugliese and Tacchella, 2020). Indeed, the ECI methodology is very good at explaining differences in various economic outcomes and at multiple spatial scales. It provides usable and information-preserving reductions of complicated data.

Still, it has not been immune to criticism since its foundation. Beyond methodological issues (some of which will be addressed in Section 2.3), questions have arisen in regard to the vaporous meaning of economic complexity as an index. First, the term “complexity” is inherently confusing in the ECI given the meaning of the word in the rest of the literature. The ECI is not *intrinsically* (that is, through its formulation) a measure of the non-linearity of interactions, of systemic dynamism or of emergence effects (Kogler et al., 2023) – It is simply a reductive describer of compositions, and the index has remained somewhat of a black box for a long time in terms of its economical grounding. In their initial effort, *HH* equated economic complexity to being a representation of countries' capacities to export diverse and non-ubiquitous goods, but this interpretation was always only empirically led given the complicated mathematical formulation used to build their index through the *Method of Reflections*. Since then, it has been proven that their *Method of Reflections* was in fact an estimation of a fixed point (Caldarelli et al., 2012; Hausmann et al., 2014; Kemp-Benedict, 2014) that is crucially orthogonal to the diversity of countries' productions or to the ubiquity of its exports (Kemp-Benedict, 2014; Mealy et al., 2019).

Ten years on, and following in the footsteps of Kemp-Benedict (2014), Mealy et al. (2019) have provided critical insight into how to interpret the ECI and the PCI. While the formulation itself does not hold any intrinsically economic meaning, its efficiency in clustering observations and representing differences between countries (or products) in their co-specialization makes it suited to analyze economic phenomena, and I find it to also be promising in an urban context. The insights Mealy et al. (2019) provide can be shoehorned into the capability-driven framework of C. A. Hidalgo and Hausmann (2009), and knowledge is a intuitive enough driver of differences in outcomes that the ECI can now be applied as a standard measure (Balland and Rigby, 2017). But this standardization has still been due to the undeniable observed effectiveness of the measure in uncovering hidden systemic phenomena, rather than due to deductions from its formulation.

In the following, I argue that the economic complexity paradigm and its network-driven approach is suited to address systemic planning-oriented issues, and in turn to help understand urban crises. Using economic complexity indices to granularly understand cities is also the focus in the work of Juhász et al. (2023), where they show that an economic complexity index of neighborhoods' amenities is correlated to socio-economic mixing, and relate it to urban segregation.

However, this work's perspective is quite different from that of Juhász et al. (2023). Beyond the slight methodological differences, this chapter seeks to lay the foundations of what is and should be treated as a novel and very different application of the ECI. I advance that the empirical success of the ECI and its intuitive underlying assumptions of compositional complexity are not relevant enough alone to standardly apply it to such a different setting. Levels of "Complexity" or of "Sophistication" of spatial entities and of activities have come to bear meaning in economic complexity spheres, yet their meaning in these spheres is quite different from how those words are used in other academic contexts[2]. Their colloquial meaning can in no way be directly transferred to an "ECI of neighborhoods". The amenity-driven approach that will be leveraged here (and that Juhász et al. (2023) also leverage) requires a careful assessment of the latent characteristics that drive the index, and of its formulation, in order to unpack any meaningful economic interpretation.

Crucially, interpretations in this chapter as a whole are built on modern understandings of the ECI formulation's, that is, it considers the ECI straightforwardly as a minimization of the squared distance between nodes in a similarity matrix wherein edges depend on relative co-ocurrences (Mealy et al., 2019). In that sense, it goes against the normative and absolute neighborhood complexity interpretations of Juhász et al. (2023), who propose that neighborhoods' complexity values depend on the "sophistication" of local amenity

[2]The most relevant definition of "complexity" to this work is the one found in planning, and it is very different from the one implied by the ECI. Likewise, in EEG at large, "complexity" bears a different meaning Kogler et al., 2023; Martin and Sunley, 2007

supply, on the diversity of neighborhoods, and on the ubiquity of amenities. Here, the index adapted at an inner-city level will simply separate inner-city spatial entities unidimensionally into how similar or dissimilar they are in what they host, and economic meaning has to be derived from what motivates the unevenness of differentiated retail presences. Therefore, defining co-ocurrences (that is, a binary network representing a city) with a tidy background is critical to obtaining an index that mirrors real-world issues. It is also critical to presenting a way of reducing the information contained within cities that is not a convoluted black box, and to uncover meaning from the separation of urban spaces.

Having described the bilateral epistemological grounding of this chapter, the following will focus on carefully building a complexity measure for neighborhoods in a ground-up way. First, I will argue through EEG reasoning and terminology that market dynamics are inherently suited to reducing the complex nature of cities. Then, a framework within which to understand these dynamics and how they can relate to a complexity index will be presented, and economic complexity's product space will be translated into a strictly commercial amenity space. A set of methodological recommendations will finally be detailed to present a specific adaptation of the economic complexity methodology, the Amenity Complexity Index (ACI) and its accompanying amenity-Type Complexity Index (TCI), to reduce neighborhoods' commercial amenity spaces. Using a consumption-focused framework, I find that the ACI is a promising tool to uncover urban dynamics through its underlying segregation of places, of commercial amenities, and of consumers, provided it is built in a way that it can be adequately understood. I argue that, through market dynamics, separating places through their commercial amenities can be leveraged to understand cities in the context of systemic urban transformations.

2.2 The commercial amenity space

"*Cities must attract workers on the basis of quality of life*" (Glaeser et al., 2001). In the same way differences in outcomes between cities are driven by their amenity space in Glaeser et al. (2001), different outcomes within cities' places are also driven by and represented in their overall amenity space; And a commercial amenity space is an apt reduction of that space from a systemic standpoint.

Glaeser et al. (2001) specifically concentrate on the demand for cities from the widest possible angle. This section's objective is to build a network-driven framework within which to understand urban systems, and it ultimately limits itself to the observable results of market-driven consumption dynamics (more specifically, to commercial amenities, or retail businesses). Economic complexity's product space's (C. A. Hidalgo et al., 2007) concentration on products alone is natural and understandable; They do not mix trade data with socio-demographic or institutional variables for example, and the reasons to do so appear straightforward. The reasons not to include other amenity classes might however need more explaining in this application, especially as past research efforts of the ECI on amenities have not taken this approach (see Juhász et al., 2023).

2.2.1 Focusing on Commercial amenities

It should be reiterated that although I borrow a Glaeser et al. (2001)-like consumption-driven view of peoples' experiences within cities, the focus here is on differentiating spaces within cities and not on differentiating cities among each other. The authors outline four main classes of amenities that dwellers and visitors alike could seek to enjoy: public goods and services, speed (or accessibility), aesthetics, and services and consumer goods. These four categories are not necessarily seen as exhaustive within this chapter – But they outline the underlying idea is that everything people experience within cities can be typologized as consumption and constitutive of their attractiveness. However, and as a nod to the evolutionary economic geography literature, any application of the ECI or the

PCI that seeks to reduce cities as systems needs to be able to uncover hidden processes and characteristics to be of any use.

This chapter's application of the ECI will concentrate exclusively on market-driven, or retail, or commercial amenities. The main reason for this is that a consistent set of amenities needs to be used, and that commercial amenities provide the most insight among available options. The ECI and the PCI are deeply intertwined indices that rely on a bipartite network with spatial nodes (countries in the ECI) and activity nodes (product export specialization). By separating neighborhoods based on co-occurrences, there is a simultaneous separation of the selected observables that occur within neighborhoods. To grant this separation meaning, it is critical that observables (that is, "activity nodes") themselves are comparable in the way they interact with the rest of the system. The reason this is critical is that it would be difficult to provide any clarity over what ECI or PCI results imply by mix-and-matching amenity classes[3] on the base of their similarity. Ultimately, a PCI application that for example assigns rankings to city halls, restaurants, benches, and monuments through their co-occurrences within a single network would be curiosity-inducing, but it would lack any explainability regarding underlying processes because the processes that lead to their individual occurrences are too diverging. A focus on commercial amenities allows for a better micro-foundation of the mechanisms behind the ECI/PCI application than alternatives, and thus for a better reading of "complexity" outcomes.

What makes retail (or commercial) amenities particularly interesting and comparable from a theoretical standpoint, and the reason they are selected in this work as opposed to other amenity classes, is their long-term presence being conditioned on economic interactions. Their existence in neighborhoods requires market coordination and business decisions that come as a result of many codependent interactions and interrelations, rather than *directly* from conscious planning decisions. Economic interactions and the firm-level

[3] In their multidimensional economic complexity, Stojkoski et al. (2023) combine different complexity sources, but they do not mix different datasets when building the indicators.

evolutionary process they lead to are extremely promising in regard to how the results of an ECI or PCI application to cities can be explained.

On top of this, the systemic and complex nature of cities touched upon in Section 2.1.1 both drives and is reflected in the spatial distribution of market and non-market consumption within them, and notably of retail amenities. The links between these transformations and commercial amenities are empirically solid. Current important discourse regarding urban transformations through the lens of the demand for cities (Florida et al., 2023), overtourism and place alienation (Diaz-Parra and Jover, 2020), and transnational gentrification (Sigler and Wachsmuth, 2016, 2020) are to various degrees all intertwined with commercial amenities. Empirically, Glaeser et al. (2018b) find that leveraging commercial amenities using *Yelp* data can help predict gentrification, albeit with caution regarding causal interpretations. Couture and Handbury (2020) find that non-tradeable service amenities play an important role on location choices of college graduates. Recent tourism-oriented literature has also been investigating links between commercial amenities and changes in the composition of local demand (Particularly short-term rental literature, Basuroy et al., 2020; A. Hidalgo et al., 2023; Ioannides et al., 2019).

With adequate tools, commercial amenities can teach us about urban dynamics and the modern challenges cities are facing. Existing empirical literature leverages fundamentally different approaches to this chapter's understanding of cities and the challenges they face, but it also opens up a conversation about location choices of retailers as dependent on the composition of local demand.

2.2.2 Commercial amenity locations

"*The explanation to why something exists intimately rests on how it became what it is*" (Dosi, 1997 as cited in Frenken and Boschma (2007)). A PCI/ECI application on commercial amenities aims to typologize them and the neighborhoods they locate in through their common occurrences. However, any compositional insight the ECI/PCI method can provide relies on qualifying the mechanisms behind commercial amenity locations, in other words, it relies on providing insight on the processes underlying the amenity space.

As the opening quote to this subsection might have given away, the view of this chapter is fundamentally driven by evolutionary economics. Some modern understandings of firm and industry locations, following in the steps of evolutionary economic geography (EEG) as a whole, have taken an evolutionary turn in recent decades (Boschma and Frenken, 2003; Spencer, 2017). While the overwhelming focus of EEG has been to apply evolutionary thinking to industry or general firm location dynamics, and to the competitive advantages of cities, regions and countries, there is reason to believe many of the same tenets apply to retail establishment (plant-level) location dynamics (and, to some extent, to the locations of non-commercial amenities too). But even within EEG, many different parallels, metaphors and epistemological stances (Kogler et al., 2023) could be translated to the problem of amenity location within cities in their own ways. Out of the available options, I argue that Generalized Darwinism provides an intuitive perspective into intra-city retail location dynamics.

Generalized Darwinism in EEG advances that core principles of biological evolution are relevant to assess change in economic systems, albeit in a way that is specific to these systems (Essletzbichler and Rigby, 2010; Essletzbichler and Rigby, 2007). The approach of Essletzbichler and Rigby (2010) is more nuanced than earlier applications of Darwinian economics (Hodgson, 2002), and it does not argue for the unaltered use of biological analogies, but it still adopts core ideas of variety, selection and continuity in shaping economic landscapes in coordination with more classical economic thinking. Drawing from this, I set out analogies that I deem to be relevant towards understanding the unvenen spatial distribution of retail activity within cities.

The premise here is that changes in their systems stem from variety (and innovation) among retailers, from market and consumption-conditioned selection of retail establishments (or plants), and from the retention and adaptation of incumbents. These processes are themselves very dependent on the relevant environment (in its most holistic definition) around retailers and on path-dependency (Boschma and Lambooy, 1999; Martin and Sunley, 2006) of both retailers as firms and of their environments. A short translation of the most retail-relevant evolution-driven analogies to the distribution of commercial amenities (and changes thereof) within cities follows.

Variety refers to the heterogeneity of agents, and *New Variety* is a primary driver of long term evolution in spatial economic systems. Changes in variety can come as a result of technological innovations, of marketing innovations, or even as an indirect result of institutional intervention. For example, e-cigarettes have provided opportunities for entrepreneurs to change the retail landscape of cities with new types of stores that sell new types of goods. But existing products can also be marketed differently and provoke change; The widespread availability of new technologies has for instance allowed for a rise in "click&collect" market interactions and a different way of consuming known products that undoubtedly affects the selection process.

Selection refers to the ability of commercial amenities to withstand economic activity in time and in their spatially defined environment. The selection forces within retail are strong (Boschma and Weltevreden, 2004) and they shape the entry, exit and adaptation of firms. With the light assumption that retailers are profit-driven, their selection process is conditioned on the localized availability of individuals willing to consume them and on the aggregation of these individuals in places (Waldfogel, 2008), that is, what is further called their selection environment. This is the key behind most economic approaches to consumption, from Hotelling's most basic models (Hotelling, 1929) to those that rely on complementarity across supply, and it might might come across as a basic point. But it is a fundamental difference with other amenity classes (public services, aesthetics, accessibility) that people also consume in non-monetary ways within cities. In the context of an outcome-driven exercise that typoligizes spatial systems based on their commercial amenities, the conditioning of commercial amenities' selection on consumer presence is key because it allows for drawing parallels between types of spaces and the presence of consumers that attend these types of spaces, and between types of retailers and the localized presence of consumers that purchase their goods and services. It should be noted that the observed network of retail amenities within cities is not seen as the one-to-one outcome of a selection process within complex systems, because the rationality of retailers is bounded and their information is imperfect. These complex systems are out-of-equilibrium and constantly adapting through changes in variety and changes in selection environments. In the e-cigarette analogy, suppliers opened up many stores once the product was available, but a number of them had to close within the following few years (Lanza and Pittman, 2019). This is a good example both of the selection process and of the imperfect ability suppliers have to anticipate future compositions of local demand where they open up. But observing a network of commercial amenities can still provide valuable (if imperfect) information regarding how these systems have evolved through time, and where retailers anticipate relevant future consumption.

The *Selection Environment* within which commercial amenities operate is spatially bounded. The relationship between each individual retail establishment and its selection environment is complex and singular by nature, and it includes strong feedback effects. Retailers obey to and shape their environments, and they compete simultaneously within multiple different environments at different scales. It is however clear that many aspects of selection environments are shared by different retailers, even if they relate to them in different ways. This includes logistical supply questions and the institutional setting alluded to in Essletzbichler and Rigby (2007) and Nelson (1995). But more practically and because selection is consumer-conditioned, and to go back to the Glaeser et al. (2001) consumer city setting, it includes all amenities available to consumers and is not limited to other economic agents or firms. The accessibility of places, their perceived safety, their aesthetics, their historical significance, their built form, and their spatial relationship with other possible complementary or competing economic activities are a non-exhaustive list of elements that all play into consumer presence and, in turn, into retail location decisions in ways that are quasi-impossible to measure. The way in which retailers perceive this environment (and how their potential consumers might relate to them) in light of their own information provides *cause* for their location decisions, a central element of Darwinism (Hodgson, 2002). In short, history matters because it shapes environments, which in turn shapes future environments. Inertia is a strong feature of retail location decisions, which in turn strengthens selection (Essletzbichler and Rigby, 2010) and nourishes path dependencies (Boschma and Lambooy, 1999; Martin and Sunley, 2006) that enhance how observed networks of commercial amenities represent selection environments. It is the different ways in which retailers relate to their environments that leads to uneven outcomes, but their relation with their environment is conditioned by their anticipation of consumer behavior. This chapter makes no direct attempt at assessing the relationship between environments and causes, but it advances that observing an uneven distribution of commercial amenities within cities is intimately tied to an observation of the uneven characteristics of spatial selection environments within cities.

2.2.3 A network perspective

This subsection presents the aforementioned selection mechanism in a formal network manner mimicking economic complexity' network representations of underlying interactions (C. A. Hidalgo and Hausmann, 2009). However, before that, the evolutionary framework explored above lacks in one key question: What is it that makes people consume different commercial amenities? Part of this is the environment commercial amenities operate in, hence consumers' preferences for these environments. Another part is their preference for the services or goods offered by retailers, and this has until now not been detailed in this chapter.

This problem can find answers in various fields. In this specific setting, grasping the mechanisms behind individual preferences is not necessary, and we do not need to take an epistemological stance on consumer choice (see Hands, 2010 for a perspective) or consumer rationality. Preferences for amenity consumption are accepted holistically and abstractly as a result of individuals' income, cultural influences, perceived identities, status, tastes; Overall, as a result of what would be an unquantifiable *habitus* in a sociological framework (Bourdieu, 1977). It is still important, within a conceptual framework, to note that these preferences exist. Especially as I attempt to characterize the groups of people behind different amenities' consumption later in Chapter 3. To that aim, I argue that a strong common characteristic to people that consume similar commercial amenities is how much they consume as a whole, as a function of both their spending power and their will to spend it on commercial amenities. This ties in to the idea of budget constraints in consumer choice theory – there is strong empirical evidence that spending power is linked to observed differentiated consumption (two examples of which are Aguiar and Bils, 2015; Jackson, 1984), and to the idea that economic capital is intrinsically linked to *habitus* (Bourdieu, 1987).

It is also critical to outline that this subsection marks a switch from plant-level analysis (individual commercial amenities) to industry-level analysis (commercial amenity types) that is necessary to any economic complexity reduction. As such, it makes the admittedly stronger assumption that consumer preferences are for industries rather than for individual retailers. In other words, commercial amenities should be grouped in a way that consumers' preferences for retailers (regardless of their selection environments) are equal within types, or industries – That is, different establishments within types should be a substitutable as possible. How valid this assumption will be in a observed setting will depend on the practitioner's ability to granularly define types of commercial amenities in ways that are as close as possible as consumers envision them. Keep in mind that this does not imply selection environments are irrelevant in consumers' choices. On the contrary, the combination of their preference for a type of good or service and a selection environment will allow for that type of retailer to be selected or not.

For now, I formalize a setting wherein consumers simultaneously have different preferences for types of commercial amenities they want to consume and different preferences for places (or the spatial selection environment commercial amenities operate in) to consume in. Because selection is conditioned on consumer interaction, the network linking commercial amenities and their selection environments (places) can be decomposed into two constitutive bipartite networks, or matrices. One of the preferences consumers have for commercial amenity types, and one of their preferences for places. The third network links commercial amenities to places through their presence. In a perfectly rational and information-full world, it would come as a function of the last two and represent the outcome of a selection process. Of course, this assumption is not made here, and these depictions are meant for understandability. Still, and this was the argument made in the previous subsection, the third network still contains important information about selection, even in an observed setting.

Take a consumer-amenity matrix AC with a rows (commercial amenity types) and i (consumers) columns where $AC_{a,i}$ is equal to the amount consumer i is willing to spend on amenity type a. The row sums of this matrix is the overall spending power of consumers, and their preferences define how they allocate that spending across different columns. Conversely, the column sums of the matrix are supply-side revenue, provided consumption is realized. For the purpose of visual clarity, this perspective is simplified by assuming binary relationships between individuals and types of amenities, where $AC_{a,i} = 1$ when i is willing to spend a unit of income in a and $AC_{a,i} = 0$ otherwise. This caps per-amenity consumer spending to 1 unit of income in a way that is not reflective of the real world, but the assumption is not kept beyond visualizations. A bipartite network illustrates AC in Figure 2.1.

Parallelly, the same individuals distribute their spending in different places within the city. Where they consume depends on place characteristics that throw us back to the selection environment alluded to above. For example, accessibility to a place for a consumer likely influences these links strongly. Urban planners have relative control upon these characteristics. I note a consumer-place matrix $CP_{i,p}$ with I consumers in rows and P places in columns. Different links between consumers and places do not all hold the same value in regard to consumption allocation, but once again for visual simplicity, CP

is binarized into 1 when a given consumer can spend a single unit of spending in a place and to 0 otherwise. A network representation of CP can be found in Figure 2.1. CP is a representation of the demand for consumption in places.

Figure 2.1 gives a visual representation of a third matrix $PA_{p,a}$, where places p share edges with amenity types a when a is selected in p. PA is the commercial amenity space and it what will later be observed with data. The underlying intuition here is that PA is a function of CA and CP (even if, in an observed setting, PA would be an imperfect function of CA and CP). For simplicity's sake, I assume that all different commercial amenities require the same level of revenue in a place to be present in it. The level of required revenue is set to 2 in this example. If at least two units of relevant consumption for a are found in p, then presence in p is sustainable for a. The observed commercial amenity space PA is a function of the hidden relationships between consumers and amenity types (CA) and between consumers and places (CP): When a has a high enough number of relevant consumers in its selection environment p, it is selected.

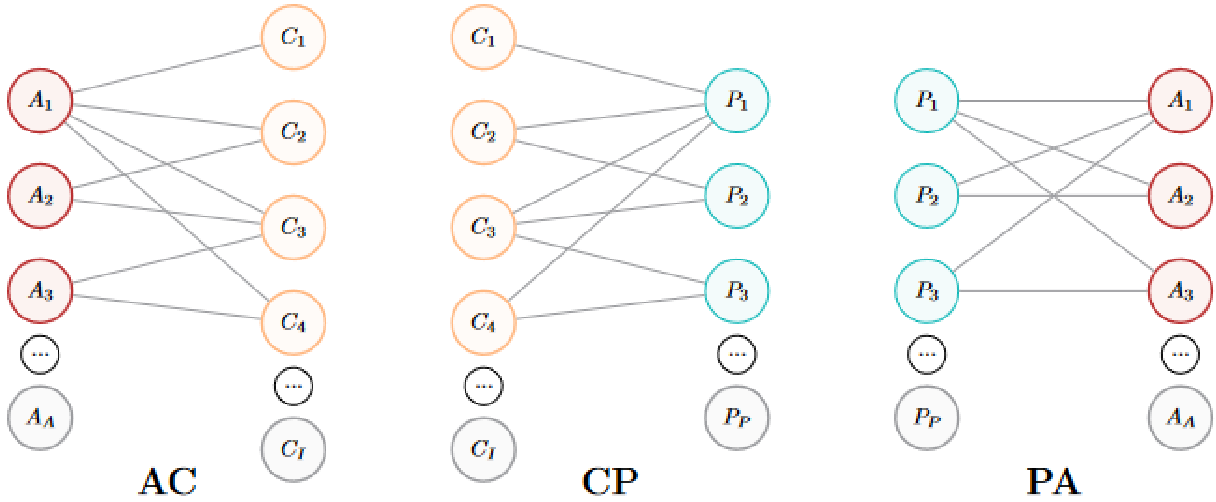


Figure 2.1: Network visualizations of the coordination processes behind commercial amenity consumption.
Adapted from Hidalgo, C. A. and Hausmann (2009, Fig. 1.A)

(AC): Network of individual consumers and commercial amenities.
Consumers C_I share an edge with commercial Amenity A_a if they allocate consumption to it. Here, A_1 is a necessary amenity, as everyone spends on it. In practice, this amenity is likely to be essential to everyday life, like a supermarket for example. C_3 has the highest spending power and shares it equally between all three of the presented amenities. C_2 and C_4 both share the same spending power, but they have different preferences.
(CP): Network of individual consumers and places within cities.
Consumers C_I share an edge with Places P_P if the place is attractive enough for them to consume there. In this specific example, P_1 is frequented by everyone. In practice, P_1 is likely to be central, providing access to various (commercial and non-commercial) amenities in a way that is accessible to all the consumers. C_3 spreads her spending the most across the city, while C_1 only spends in P_1. C_2 and C_4 each provide selection opportunities in two different places.
(PA): Network of commercial amenities and places within cities.
Commercial Amenity types A_A share an edge with Place P_P if they are selected within a place. I argue that PA comes as a result of subnetworks AC and CP. Amenity type A_1 is present in all 3 Places because it is desired by all consumers, and all 3 places host at least 2 consumers. A_2 is not present in P_3 because C_2 is missing from P_3. A_3 is not present in P_2 because C_4 is missing from P_2.

The PA matrix, that is meant to represent the outcome of the selection process, is what we can observe with data. It is also what Juhász et al. (2023) and C. A. Hidalgo et al. (2020) observe. Edges between places and amenity types signal selection, and hence the availability of a sufficient amount of relevant consumers (that are willing and able to consume it) in a place.

The networks presented in Figure 2.1 are voluntarily oversimplified views of reality.

Still, they suffice to perceive consumption in cities from a complex systemic standpoint. Interactions between and among all agents are permanent, and small changes can have large and non-linear ripple effects. Let me play out a made-up scenario to demonstrate these ripple effects. A planner decides to replace a parking lot with a park in city center P_1. This might make the place less accessible to suburban consumer C_2 for whom it will be a lot more costly to drive there, but encourage C_4 to spend a larger share of her time there because she finds it more desirable. As a result, C_2 visits P_3 instead. Parallelly, C_4 now meets up with her friends in P_1 instead of P_3. With time, A_3 will be lacking available revenue in P_3 because of C_4's absence, and its presence will not be sustainable in the long term. The changes in place habits of these consumers and of the commercial amenity-mix distribution could in turn impact other places' relative ability to attract consumers and their own selection process. In fact, this scenario can not be fully played out as the multiplicity of agents and interactions will never leave enough time to reach an optimal situation. This is inherent to complex systems: they are permanently emerging and out-of-equilibrium. Attempting to fully grasp the process of places' emergence within cities through their characteristics alone is quasi-impossible. Yet the commercial amenity mixes of places tell us about a number of different things that can influence both the presence of consumers and their spending habits. In its most holistic interpretation, the retail amenity space and the selection process it implies contain information about flows and interactions, ranging from supply and demand dynamics, to top-down zoning, to specialization, to quality of life, to overall atmosphere.

Economic complexity methodology is particularly well-suited to solving network-driven problems and to unlocking hidden underlying patterns. Leveraging the full information of the observable network in PA is important, because it can uncover profiles of places and types of amenities through the lens of consumer-driven selection. Associating these profiles to profiles of consumers can in turn help understand changes within cities on a broader scale (Florida et al., 2023; Sigler and Wachsmuth, 2016). In the following section, an adaptation of the Economic Complexity Index to places and to their commercial amenities is presented as a way to reduce the observable network PA while preserving information over AC and CP.

2.3 A methodology to measure the ECI of Places and the PCI of Amenities

2.3.1 Specifying the ECI

Economic complexity has given economists information-preserving ways of summarizing productive structures. The flexibility of these principles has been thoroughly explored in the last few years of evolutionary economic geography research, as Balland et al. (2022) and C. A. Hidalgo (2021) summarize well. After having presented the groundbreaking work of C. A. Hidalgo and Hausmann (2009) (HH) and explored the ECI's interpretations, I will discuss its conceptual and practical challenges in the amenity space and propose a specification for the Amenity Complexity Index (ACI). HH explain differences in complexity for countries as emanating from sets of productive capabilities that determine which goods a country is capable of exporting. The ECI is meant to reduce a bipartite network (like PA in Figure 2.1) that can be represented in a binary matrix $M_{c,p}$ with $M_{c,p} = 1$ where country c is specialized in product p and $M_{c,p} = 0$ otherwise. It does so in a way that preserves information well and is agnostic about the complexity of both products and industries (C. A. Hidalgo, 2021). In their original work, HH presented the Method of Reflections (MOR) as a way of estimating country (ECI) and product (PCI) complexity indexes. The MOR uses an average-based iterative algorithm based on countries' diversity and products' ubiquity (or non-rarity) as initial conditions. Diversity, noted $K_{c,0}$, is number of industries each country is specialized in: $K_{c,0} = \sum_a (M_{c,p})$. Ubiquity $K_{p,0} = \sum_n (M_{c,p})$ is the sum of countries each industry is specialized in. The MOR interacts ubiquity and diversity measures over i orders, with $i \in \{1, ..., I\}$ and I being the order at which $K_{c,i}$ (ECI) and $K_{p,i}$ (PCI) are stable. It defines the ECI $K_{c,i}$ and the PCI $K_{p,i}$ simultaneously as:

$$K_{c,i} = \frac{1}{K_{c,0}} \times \sum_{p} (M_{c,p} \times K_{p,i-1}) \tag{2.1}$$

$$K_{p,i} = \frac{1}{K_{p,0}} \times \sum_{c} (M_{c,p} \times K_{c,i-1}) \tag{2.2}$$

HH originally argued that $K_{c,i}$ was a generalized measure of diversification or of ubiquity depending on the parity of the iteration, but these interpretations have since been mathematically disproved (Kemp-Benedict, 2014). In fact, it was later proven that the MOR converges towards estimations of two specific fixed points, as demonstrated in the literature (Caldarelli et al., 2012; Hausmann et al., 2014; Kemp-Benedict, 2014). Using square similarity matrices $W_{c,c'}$ and $W_{p,p'}$ between countries (products) weighted by their diversity (ubiquity), that is:

$$W_{c,c'} = \frac{M}{K_{c,0}} \times \frac{M^T}{K_{p,0}} \tag{2.3}$$

$$W_{p,p'} = \frac{M^T}{K_{p,0}} \times \frac{M}{K_{c,0}} \tag{2.4}$$

Values in $W_{c,c'}$ can be interpreted as probabilistic measures of reaching country c' from c through a path of their shared exported products (Kemp-Benedict, 2014). See Appendix 2.6.1 for a visual explication of how these matrices are built. Namely, $\frac{M_{c,p}}{K_{c,0}}$ is the chance likelihood of drawing product p from the export basket of c, and it is combined with a $\frac{M^T}{K_{p,0}}$ (otherwise noted $\frac{M_{c',p}}{K_{p,0}}$) likelihood of drawing c' among the countries p is exported in. The sum of those combined probabilities for every product in country yields a row-stochastic matrix that is interpreted as the relative similarity between countries (or products in the case of $W_{p,p'}$). The eigenvectors $\upsilon^C[2]$ and $\upsilon^P[2]$ associated with the second largest eigenvalues of $W_{c,c'}$ and $W_{p,p'}$ correspond to the fixed points estimated by high-order iterations of the MOR. It is then standard practice (C. A. Hidalgo, 2021) to normalize $\upsilon^C[2]$ and $\upsilon^P[2]$ by subtracting them by their means and dividing them by their standard deviations, yielding the ECI and the PCI as it is now used. Although the intrinsic ties between products' and countries' complexities remain, complexity can no longer be thought of as a "generalized measure of diversity" (as in *HH*), and it is in fact orthogonal to diversity (Kemp-Benedict, 2014; Mealy et al., 2019). Interestingly, the eigenvector procedure has however opened up new, clearer interpretations than those at high orders of the initial MOR (C. A. Hidalgo, 2021; Kemp-Benedict, 2014; Mealy et al., 2019). The ECI and the PCI tell us about how different countries (or products) are from each other in the similarity matrix, both in sign and in magnitude, by laying them on a spectrum (or on a continuum). Here, they are a way of separating observations based on relative co-occurrences in $M_{c,p}$ that minimizes the differences in export baskets of similar countries (Mealy et al., 2019). It helps to imagine the ECI (and conversely the PCI) as a procedure that assigns countries to a position on a line, and that it is up to the economist to provide meaning to either ends of the line.

Because we know that $\upsilon^C[2]$ and $\upsilon^P[2]$ are respectively equivalent to $-\upsilon^C[2]$ and $-\upsilon^P[2]$, the fundamental interpretation of the ECI (PCI) should be relative[4], and "high" or "low" complexity mean nothing through their formulation alone. This is a very important point: Complexity indices are valuable metrics that hold latent information regarding the product space and that have desirable properties. But in no way are they normative by nature, nor are they constitutive of underlying economic characteristics through their formulation alone. They are simply measures of relative distances in a similarity network between countries or products, and ultimately outline differences in productive compositions.

I use the standard (C. A. Hidalgo, 2021) normalized eigenvector method on the commercial amenity-place network, and reiterate the relativity of complexity because it is paramount for interpreting the indexes. There is nothing absolute about low or high complexity values. A highly complex place is however similar to other highly complex places and very dissimilar to low complex places in its composition of commercial ameni-

[4]See Appendix 2.6.2 for why this is, and how possible sign flips are dealt with in my methodology

ties. It helps to imagine complexity indices as positions on a line that simply indicate how close observations are to each other. The intuition of this work is that (dis-)similarities yielded by complexity can be interpreted through the consumption-driven selection lens of the framework outlined in Section 2.2.2. Having provided a basic understanding of what the ECI means, the focus is now on the challenges that are specific to the amenity space in building the network.

2.3.2 Amenity complexity's conceptual and practical challenges

Two separate dimensions have to be articulated to define each place's amenity-mix; Space, that is, what a place is, and what intensity (amenity count) is enough to justify binary adjacency between places n and amenity types a in the network $M_{n,a}$ that links them.

Defining places

Existing network-driven amenity literature evacuates the spatial problem by using predefined administrative boundaries (Juhász et al., 2023) or clustering amenities (C. A. Hidalgo et al., 2020). In this work, places are implied to be a spatially defined selection environment of retail plants, and those two methods are not satisfactory in that sense. I seek an alternative spatial division of cities that would:

(1) Allow for heterogeneity within administrative boundaries that are not intrinsically relevant to consumption nor to retailers' location decisions;

(2) Be respectful of the urban landscape because this landscape shapes the ability of consumers to move;

(3) Provide continuity across space so that selection environments can overlap because commercial amenities' selection environments are best represented by the space directly around them.

All three of these issues can be solved by considering practitioner-defined spatial coordinates as the centers of places, while allowing for overlapping places and with a large enough number of places. The downside of this added granularity is it might require more work to clean the data (which is already a problem at less granular country levels, see C. A. Hidalgo (2021)) and to select the appropriate global space of the network (In other words, a higher-scale selection environment that would be common to all commercial amenities within the network). In practice I find that using residential buildings as the center of places strikes a good balance between granularity and practicality, solving (1) by filtering out a priori non-urban places better than gridded points would[5], and thus also providing a better solution to (2). I call c the cutoff used to define places around buildings, with $c < d$ the distance between buildings, allowing for the introduction of spatially correlated places (solving (3)). Commercial amenities are a part of a building's place when they are within c of it. Both the metric used to define c (for example meters or minutes) and the value of c depend on the practitioner's understanding of local context, and on the size of agents' perceived neighborhood-scale selection environments. A general rule of thumb is that c should define sizes of places in a way that is as close as possible to how place-goers envision them in a homogeneous way across the city. As a side note, c can be a metric measure of distance, but an even more adequate application would see it be a practical measure of distance that is more respectful of the urban landscape, such as travel time.

Having defined places and c, I now focus on how to binarize the links between places and their commercial amenity types.

Defining binary co-occurrences

Computing the ECI typically requires a binary matrix. The problem of binarization is equivalent to finding the best way of reducing a bipartite weighted network $O_{n,a}$, that holds the number of retail amenities of type a in place n, into $M_{n,a}$, a matrix of 0s and 1s. $O_{n,a}$ can be viewed as a weighted bipartite network between places and amenities where weights are discrete counts. I argue for a ideal-scenario methodology wherein $M_{n,a}$ is defined through a simple measure of presence instead of the ECI-standard Revealed

[5]Gridded points can of course be a valid alternative, as can any method that satisfies $c < d$ in a sufficiently granular way while also satisfying (2). But in practice, I found that they can only work as long as they are filtered and snapped to a street network

Comparative Advantage (results using RCA in application remain close but are noisier, see Chapter 3's appendices).

Using the Balassa index (Balassa, 1965) of Revealed Comparative Advantage (RCA) to binarize $O_{n,a}$ is considered straightforward in economic complexity. Only a few papers ever discuss it beyond its application (Mealy et al. (2019), for example, discuss RCA threshold selection in their supplementary materials), and Juhász et al. (2023) also use the Balassa index for amenities. I question the validity of the RCA in the amenity space and argue that an alternative presence-based indicator holds both more practical and conceptual ground. Using the RCA, for a place n in which amenity type a represents a proportion at least equal to (for R* = 1, the standard) the proportion of a over the city, $M_{n,a}$ will be equal to 1;

$$RCA_{n,a} = \frac{O_{n,a}/\sum_a O_{n,a}}{\sum_n O_{n,a}/\sum_n \sum_a O_{n,a}} \tag{2.5}$$

$$M_{n,a} = \begin{cases} 1 & \text{if } RCA_{n,a} \geq R* \\ 0 & \text{otherwise} \end{cases} \tag{2.6}$$

The RCA is a measure of relative, not absolute, strength: a place n might have a RCA in restaurants when n' does not, despite n' having more restaurants. In a consumer-driven selection setting, it is difficult to believe restaurant-going consumers are more likely to attend a place because that place does not have a jewelry store, for example. $M_{n,a}$ should map what is where, not which places are *relatively* good at hosting things.

Moreover, count data are not comparable from one amenity type to another, nor from one specific amenity to another within a type. A restaurant has different implications to a places' economic interactions, different relative importance in the network, and is subject to different selection forces when it can seat ten people compared to when it can seat one hundred, which there is no way of observing (for now). Because there is reason to believe both different types of amenities and different places will have different size patterns, defining a binary matrix in a relative way is likely to be skewed — On top of being overly sensitive to thresholds[6]. Conceptually, attracting consumption is not about the relative strength places have in hosting amenities, it is about their absolute strength, and this framework is therefore emancipated from specialization and diversification.

Moreover, and opting for an absolute definition, the difference between consumers being attracted to a place because of a commercial amenity, or the evidence a commercial amenity can be sustained in a place, is never bigger between two discrete values of count than it is between 0 and 1. There is more to be learned about places' commercial amenity-mixes through the absolute absence of retail types than through any given relative intensity of presence as long as amenity types are defined granularly enough. When types of businesses are missing, I can be absolutely certain there is no realized consumption for their type in a place.

I thus conclude that the best count threshold is 1, that is, that the simple presence of an amenity type within a place in $O_{n,a}$ at any intensity is enough for it to be linked to the place:

$$M_{n,a} = \begin{cases} 1 & \text{if } O_{n,a} \geq 1 \\ 0 & \text{otherwise} \end{cases} \tag{2.7}$$

This methodology however densifies $M_{n,a}$, and in applications where amenity types are not sufficiently granularly defined[7], $M_{n,a}$ risks being too saturated to be able to extract any information using complexity indices. In such cases, the RCA is a good candidate as a fallback option. Still, using simple presence is both the most agnostic decision and the one that allows for the most straightforward interpretations of "complexity".

[6]Mealy et al. (2019) show that R*=1 is a relatively suitable standard, but they also show that the ECI is not perfectly stable with respect to $R*$.

[7]These include *Google Places* and *OpenStreetMap* data applications

2.3.3 The Amenity Complexity Index

To determine the Amenity Complexity Index (ACI), a particular specification of green complexity found in Mealy and Teytelboym (2022) is followed, but I use it for different reasons. Using very granular places (i.e large N in $M_{n,a}$) comes with potential drawbacks that are mitigated through this specification. First, the level of hardware needed to compute W_N (the place similarity matrix based on $M_{n,a}$ in Equation 2.3) grows fast enough with N that building-level indices rapidly becomes impossible. Thankfully, as demonstrated by Mealy et al. (2019) in their supplementary materials, a place's ACI $\upsilon^N[2]$ corresponds to the average of the Amenity-Type Complexity (TCI, our equivalent of the PCI) of the amenities it is associated to in $M_{n,a}$. This is what C. A. Hidalgo (2021) refers to as biology-inspired reciprocal averages. Therefore, the ACI can be deduced through $\upsilon^A[2]$ and $M_{n,a}$ by summing the TCI of present amenities and dividing this sum by the place's diversity and $W_{n,n'}$ does not require direct computation.

The amenity data should be as granular as possible, that is, it should ideally not exclude amenity types based on a scarcity threshold as can sometimes be the case in economic complexity (C. A. Hidalgo, 2021). The downside of this is that it makes place indexes susceptible to outliers and quasi-random noise induced by amenity suppliers' imperfect market decisions and the granularity of places. Considering places have very different market sizes and diversities ($K_{n,0}$), and that some are very small, distortions brought by extreme TCI values in a place with low diversity would automatically make it an outlier if the ACI was defined as an average of TCIs. I therefore opt for a specification of the ACI that is additive regarding the TCI of present amenities, similarly to Mealy and Teytelboym (2022) for countries' green complexity:

$$ACI_n = \sum_a (TCI_a \times M_{n,a}) \tag{2.8}$$

where $TCI = \frac{\upsilon^A[2] - \overline{\upsilon}^A[2]}{\sigma(\upsilon^A[2])}$, a normalized version of $\upsilon^A[2]$ the second eigenvector associated to the second-largest eigenvalue of $W_{a,a'}$ (equivalent to $W_{p,p'}$ in equation 2.4). $W_{a,a'}$ is thus central to the methodology. Its construction and interpretation are further detailed in Appendix 2.6.1. Likewise, the eigenvector-driven TCI it yields is subject to potential quasi-random sign flipping across observed years and I invert the vector if it is negatively correlated to amenities' ubiquities. See Appendix 2.6.2 for more detail. The TCI yields the (dis)similarity between amenity types based on a component that best explains the differences across $W_{a,a'}$ while being agnostic about what that component might be in the real world. The ACI therefore does not directly measure places' selection similarity, but it gives an aggregation of their present amenities' (dis)similarities. I am measuring the ability of places to accumulate amenities associated with complex places, not their similarity to complex places – although the two are strongly correlated[8]. For that reason, the ACI in its essence should be treated as a measure of the bias towards lower or higher TCI values of its present commercial amenities.

Understanding the ACI in the real world and in the context of selection environments consequently depends on providing meaning to the component behind $\upsilon^A[2]$. In a simplified way, this component represents what most differentiates commercial amenity types given their absence/presence within the network. This also means ACIs and TCIs are not comparable across cities, and to make them comparable, one would need to treat all sub-city places as part of the same amenity space (or overall selection environment) with complicated and framework-breaking implications.

2.4 Discussion

This chapter has borrowed notions from planning, economics and evolutionary economic geography and articulated them with economic complexity indices to motivate a new perspective of urban neighborhoods that could be helpful in the context of urban transformations. Using these notions, I have sought to present an ECI/PCI application within cities that is less dependent on backward-inducing of interpretations through data, and instead builds their meaning in a ground-up way with the following key takeaways:

[8]See Chapter 3's Appendix 3.8.1

- The ACI (TCI) is a relative reduction of a network of similarities between observed places (commercial amenities) that tells us how different they are relative to each other in their commercial amenity (place) co-occurrences.
- Observed occurrences of commercial amenities within places are intimately linked to an ongoing selection process that articulates retailers, places as selection environments, and potential consumers.
- In this context, the ACI and TCI can be further interpreted beyond a simple measure of differences in co-occurrences to benefit research in at least two separate (but complementary) ways:
  1. The more different the ACI values of places, the more different the environment within which commercial amenities operate, and the more different their overall urban space is perceived by consumers. Likewise, the more different the TCI values of commercial amenity types, the more different the characteristics of the environments within which they can be selected.
  2. The more different the ACI values of places, the more different the consumers that operate the selection (that is, consume) process are in their consumption preferences within those urban spaces. Likewise, the more different the TCI values of commercial amenity types, the more different the consumption of consumers that select them.

TCI/ACI measures depend on retail locations, and they thus mark an important difference with traditional ECI/PCI applications in EEG. EEG and standard ECI/PCI applications tend to be more focused towards the uneven distribution of economic activity in productive terms. Knowledge and the relative ability differences countries, regions, and cities have of combining it, are important towards how production is distributed in space. In seeking to map urban systems and transformations through the way people experience cities, this chapter's endeavour has instead shifted towards a focus on economic activity through consumption, not through production. Understanding the TCI/ACI requires understanding what motivates uneven retail locations. While knowledge and technology are intuitive and strong predictors of geography in production frameworks, I argue that retail locations are primarily dependent on consumer selection regardless of these elements. Types of retail are not inclined to share the same spaces on the basis of an elusive shared level of "*sophistication*" (as Juhász et al. (2023) imply by interpreting their version of the TCI that way), but rather on a shared required selection environment. Assume a simple example in which ice cream stores and high-end hotels tend to locate in similar places. They would thus be close in terms of their TCI. Basing interpretation on "sophistication" would imply that ice cream shops and high-end hotels share similar levels of technology and of knowledge combinations. The interpretations above argue that, instead, their ability to share space would depend on them sharing consumers (that is, they are complementary for tourists) and/or sharing required environmental characteristics (proximity to monuments, for example) – With disregard for the production processes underlying these retailers.

By providing new means to identify and typologize both the composition of sub-city level local demand and the characteristics of urban spaces, and despite not necessarily being process-focused by nature, I advance that complexity indices are a promising way of engaging with urban changes and the processes that underly them.

However, before exploring further, I wish to clarify a few of the key points. First, and despite finding the use of residential buildings as centers of places practically sound, the indices are not meant to exclusively represent selection (or consumption) driven by residential consumption. The presence of residents affects the selection environment, and it *can* be a driver of selection itself, but so can any non-resident (even tourists, obviously). A useful exercise is to imagine residents of places to likely have stronger links (as represented in Figure 2.1) with the places they live in, but not as bounded by the places they live in.

Second, and to reiterate, a drawback of these measures is that they depend on how suppliers with imperfect information think they can coordinate with consumers, not on *actual* selection processes within places. The market for retail goods and services is not frictionless. Suppliers of amenities might misjudge consumption, be over-resilient, or miss opportunities. Consumers are not omniscient or perfectly rational either. In the short-term, the observed place-amenity network is a slightly polluted representation of the hidden processes that are implied in this chapter. Once again borrowing from EEG, this pollution could be representative of "chance" (Boschma and Lambooy, 1999). Commercial amenities' locations are probabilistic, not deterministic for given stable optimal location choices. This problem is mitigated by the fact commercial amenity locations likely depend on centuries of historical dynamics within systems, and suboptimal suppliers are unlikely to profit long-term. Observed retail locations are thus *satisficing* solutions to location choice problems. But for complexity changes over shorter periods, the risk of "chance" misallocating a supplier is higher. The wider and longer the data and the denser the city, the better the ACI will be at representing underlying processes.

One specific methodological point should also be further discussed: This chapter advocated for the use of a presence-based indicator that, effectively, seeks to deduce selection process characteristics through the absence of commercial amenities in places. This was done as a lesser evil to the threshold-dependent RCA (see Chapter 3's Appendix 3.8.1 for empirical evidence of threshold dependence), but it still has implications. The main disadvantage of a presence-based indicator (other than it's impossibility without access to top-tier data) is its greater susceptibility to *chance*, that is, to suppliers' misallocations. Values, in time, will carry a lot less momentum than they would with a RCA, and could be polluted by variations that are not reflective of actual underlying processes. Conversely, this also allows for the ACI and the TCI to be more reactive to actual changes and to detect them earlier than a RCA-based method would. All in all, I feel there is a strong argument for future research to further distance itself from the economic complexity methodology and to explore the implications of discarding the binarization process. It is understandable why the ECI is binarized as this mitigates population size effects, but it is not as clear that the heterogeneous aggregate volumes of commercial amenities hosted by places is irrelevant to this work's endeavour of typologizing neighborhoods.

With that in mind, the ACI could help answer important questions relevant to both EEG and to planning: What kind of path dependency is place-based consumption built on? How important are lock-in effects, and how are they affected by shocks? Who, what and where do changes in city paradigms stem from? How does the distribution of public goods and services relate to the local composition of demand and its changes? How do consumption structure changes affect people, and who do they affect? These questions can interest researchers and policy-makers who want to make cities better. The ACI and TCI cannot provide straight answers to them alone, but they can enrich our perception of these issues, especially alongside prior knowledge, solid methodological grounding and other data.

Path-dependence is a key focus of evolutionary research, and will be further explored in the context of neighborhoods later in this work. Policy-makers might relate the ACI to its causes and effects they want to enhance or mitigate, especially in the context of urban crises. For example, the COVID-19 pandemic is a hot topic for city structure transformations (Florida et al., 2023) that came outside the planner's control, and lockdown policies and COVID-driven consumption shifts could relate to changes in the composition of demand that complexity indices can outline. Labor phenomena such as digital nomadism, working from home or from everywhere, and AI-induced job market shifts are all factors that should affect complexity as I measure it in cities. The ACI and TCI could also add to short-term rental (STR) literature that explores links between modern mass tourism and urban transformations. In this context and that of gentrification, the ACI and the TCI could be valuable proxies to observe STR-induced urban transformations.

In an applied sense, place complexity could also be studied from a policy evaluation perspective. For example, regulation that limits STR may seek to limit changes in some places' composition of local demand (Robertson et al., 2023; Chapter 1), assuming that tourists and residents have different preferences. This chapter's methodology provides a measurable approach for evaluating these efforts. Modern planning issues like chrono-urbanism and the 15-minute city (Khavarian-Garmsir et al., 2023; Moreno et al., 2021) or the rise of B2C commerce (Boschma and Weltevreden, 2004) also imply amenity structure shifts that the ACI could outline. Chrono-urbanism is about controlling flows. As far as flows make cities and places as systems (Batty and Cheshire, 2011), the ACI as an outcome of consumer flows makes sense. By limiting non-residents' flows across places, chrono-urbanism would increase residents' consumption capacity importance — maybe with unintended effects on the distribution of commercial amenities as measured by the ACI.

One last time, the ACI is not a normative and absolute measure of place quality, nor is the TCI a normative and absolute measure of retail quality. Unlike the consensus on regions, cities and countries in economic complexity (Balland et al., 2022), more complex places are not always better. Complexity can come with inequality, and I stress that increasing it should not be a policy goal by itself because effects are ambivalent. But complexity could be seen as normative within specific objectives. Mealy and Teytelboym (2022) use economic complexity-inspired methods with a custom list of green products to evaluate countries' transition towards them. Similarly, an objective could be pre-defined (for example, avoiding place alienation by tourists and lifestyle migrants (Diaz-Parra and Jover, 2020)) and select amenity types linked to it (those sought by tourists and lifestyle migrants) to evaluate the risk of places transitioning towards undesired states.

2.5 Conclusion

Seeking to apply ECI/PCI indices to cities in order to assess urban transformations is a complicated issue that it is not made easier by the indicators' full names. Metrics of economic complexity do not measure the complexity of systems, instead, they outline patterns in compositions. This does not take anything away from their efficiency in doing so, but it means applying these methods to cities as systems is not as trivial as existing literature lets it seem. It has to be theoretically articulated with what drives change in urban spaces and with how people experience them.

Articulating evolutionary principles with modern understandings of planning has helped build the Amenity Complexity Index (ACI) and the Amenity-Type Complexity Index (TCI) from the ground up in a manner that grants these indices meaning. By reflecting relationships between consumers and urban spaces, the ACI and the TCI can be key in future research to unlock new cognizance regarding local compositions of demand and the processes that motivate urban changes. These concepts are crucial to apprehending ongoing and future urban crises, and complexity indices are a fitting candidate to fill the gap in understandings of STR that was designated in the discussion of Chapter 1.

In the following two chapters, I apply the ACI and the TCI. Chapter 3 presents an application of the ACI and of the TCI in Paris with ideal data that highlights potential uses of the ACI at a city-level and further unpacks the meaning of place and amenity complexity using the framework mapped out in this chapter. Chapter 4 tests the resilience of the indices to less ideal but more widely available data, and presents an example application of the ACI and of the TCI at an inter-city level to present a way in which the methodology could contribute to shaping STR discourse and policy.

2.6 Appendices

2.6.1 Interpretations of the similarity matrix

The explication of $W_{c,c'}$ and of $W_{a,a'}$ as normalized similarity matrices were kept brief in the main body. They are further elaborated upon here. I specifically detail how $W_{a,a'}$ is derived, the "second largest" eigenvector of which yields the TCI (see 2.3.3).

Given an amenity type a, $W_{a,a'}$ yields the probability that a random walker leaving from a in $M_{n,a}$ ends up at a' within two steps. As formulated in 2.4 for products, $W_{a,a'}$ is dependent on the number of common places a and a' are present in, on how many places a belongs to, and on the number of different amenities present in the same places. I call $W_{a,a'}$ the *relative* similarity between a and a' in the places they are found in common. The random walker has to end up somewhere, and $W_{a,a'}$ is row-stochastic (sums of rows are equal to 1).

Figure 2.2 (A) can help understand how $W_{a,a'}$ is built as a normalized similarity matrix using the example place-amenity network found in 2.1 (PA). Note that because probabilities are relative and dependent on the places a is present it, $W_{a_1,a_3} \neq W_{a_3,a_1}$. It is relatively more likely to find a_1 in any place a_3 is present in than the opposite, because a_1 is present in all of a_3's places, but the opposite is not true. As such, $W_{a,a'}$ is not symmetrical.

Figure 2.2 (B) shows a network representation of the example matrix $W_{a,a'}$ derived from the procedure in Figure 2.2 (A) (Eq. 2.4). Note that the edges leaving each note sum to 1 because the matrix is row-stochastic, and they are directed because it is not symmetrical. $W_{a,a'}$ can be interpreted as a Markov chain, wherein the directed edges between nodes represent transition probabilities between states, or types of commercial amenities[9].

Real-world interpretations of $W_{a,a'}$ are thus:

- The relative likelihood of finding a' versus other amenity types in a place in which

[9] Also note that the diagonal of $W_{a,a'}$ is not equal to 1 – Although in a binary setting the diagonal will always host the highest (or equal highest) value. If it were equal to one, it would mean that all other commercial amenity types would need to have a 0 probability of facing the same selection processes as a

a is found.

- The probability that a *transitions* towards a' through their common places. The higher $W_{a,a'}$, the higher the likelihood that a faces a selection process which also allows for a''s presence, as a share of all possible selection process similarity likelihoods.

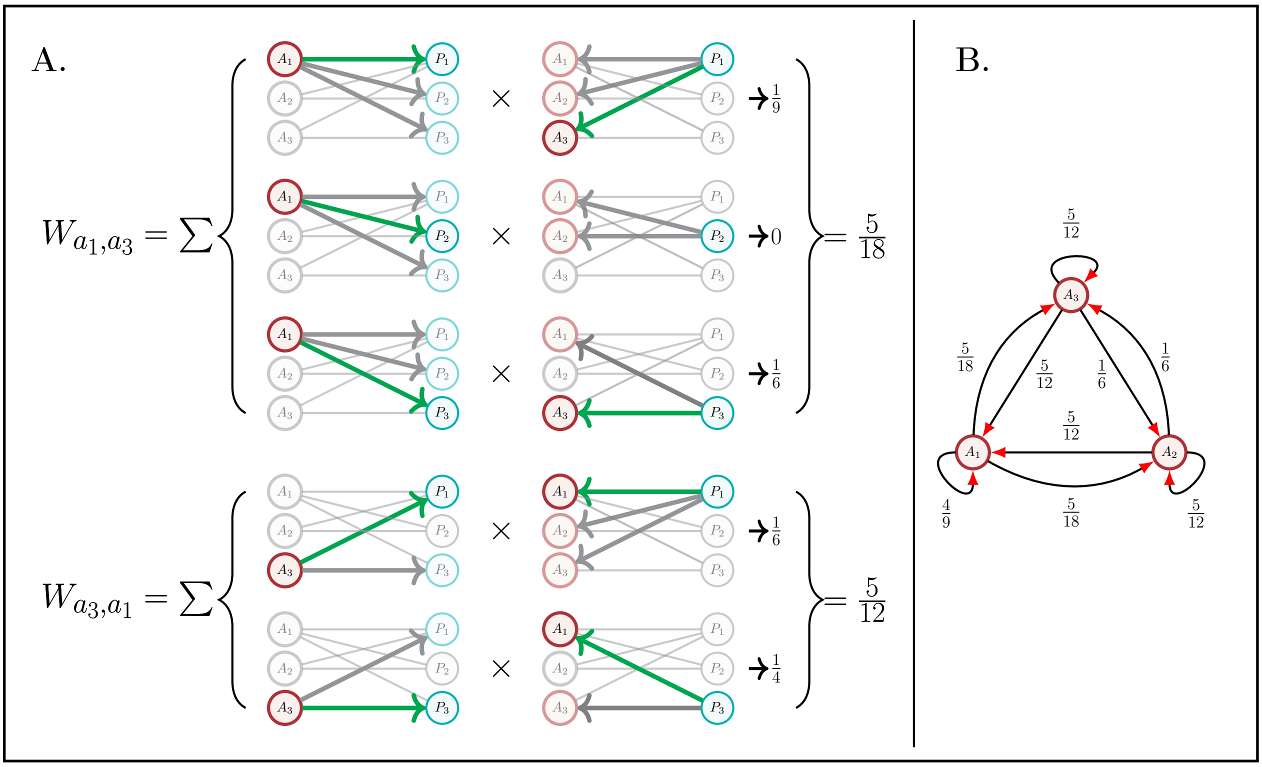


Figure 2.2: Visual representations of an example transition matrix between commercial amenities, $W_{a,a'}$. The example matrix it is derived from is found in Figure 2.1.
(A) How $W_{a,a'}$ is built using the initial place-amenity network. A random walker walks from a_1 to a_3, and from a_3 to a_1 respectively. The associated values in $W_{a,a'}$ are the sums of every relative probability of reaching a' for each place a is found in.
(B) Visual representation of $W_{a,a'}$ as a Markov chain. Directed edges denote the probability of transitioning from one node to another.

2.6.2 On the reversibility of the ECI and the TCI

In Subsection 2.3.1, I stated that "$\upsilon^C[2]$ (the ECI) *and* $\upsilon^P[2]$ (the PCI) *are respectively equivalent to* $-\upsilon^C[2]$ *and* $-\upsilon^P[2]$". Take a matrix A and one of its eigenvectors υ, with an associated eigenvalue λ. Eigenvectors, when multiplied by their associated eigenvalues, can be candidly seen as decompositions of a matrix that preserve its direction when the eigenvalue transforms it. By definition, an eigenvector is a vector that satisfies the equation $Av = \lambda v$. Because $A(-v) = -(Av) = -(\lambda v) = \lambda(-v)$, $-v$ thus fits the same definition as v with the same associated eigenvalue, and is also a possible eigenvector of A. There is in fact no inherent correctness to the *orientation* of the eigenvector (that is, the relative signs of the values it holds). Different algorithms' implementations can sometimes even yield different eigenvector signs.

Within the confines of this chapter's interpretation of the TCI, multiplying the eigenvectors by -1 does not matter either. Recall that Mealy et al. (2019) show (through the work of Shi and Malik (2000)) that the ECI (and PCI) provides real numbers that "*minimize the sum of the squared distances between nodes, where nodes are weighted according to the similarity matrix*", under the constraint of assigning non-zero values that are balanced in terms of their sign. Formally, this can be written as providing values for y_i that minimize the following under the constraint of $\sum_i y_i d_i = 0$:

$$\frac{\sum_{ij}(y_i - y_j)^2 W^C_{ij}}{\sum_i y_i^2 d_i} \tag{2.9}$$

Where i and j are rows and columns of the similarity matrix W^C. It is as trivial within this interpretation that multiplying y_i by -1 has no impact on the similarity-weighted minimization of 2.9, because $(y_i - y_j)^2 = ((-y_i) + y_i)^2$ and $y_i^2 = (-y_i)^2$.

This Appendix outlines the fact that the sign of any ECI/PCI value can only be evaluated *relative* to other ECI/PCI values and holds no meaning alone, and this "freedom to invert" should reaffirm the reader in the idea that such "complexity" measures, while they hold desirable properties, are not *inherently* tied to any positively or negatively interpretable economic meaning in their formulation.

However, when comparing positions of commercial amenities or of places across time on their ACI/TCI continua, and as economic meaning is assigned to the indices by the practitioner, it is useful to outline the same characteristics on the same side of the continua consistently through time – This is not guaranteed by formulation. There is thus a need for a data-driven way of keeping the orientation of eigenvectors consistent, that is, keeping the same sign assigned to the same underlying characteristics consistently.

Instead of engaging in more complicated formulas (see Damask, 2020, for a perspective), I follow the example Balland (2017) provides in his R implementation of the PCI as I find it leads to consistent economic underlying meaning. After having computed the PCI, he checks whether it is negatively correlated to product ubiquity (from the binary matrix), and returns an inverted PCI if it is not. I follow the exact same procedure for the TCI. In practice, this method allows for a stable index wherein places and amenity types are guaranteed not to switch from one extreme to the other year-on-year. By using ubiquity, I imply that commercial amenity rarity has a non-zero and consistently oriented correlation to where retailers locate within cities. While this is not a ground truth through the TCI's formulation, it holds up well in practice, and the following chapter will explain this selection through an application on Paris.

In the following applications, high TCI values will thus be positively associated with retail amenities' rarity, and high ACI values will be likely be positively associated with the rarity of places' hosted commercial amenities. But as explained in the methodology section and demonstrated in this Appendix, TCI values being "high" (positive) or "low" (negative) has no internal explanatory power aside from "high" being very different from "low" in consumption and selection processes.

Chapter 3

The case of amenity-driven complexity in Paris[1]

3.1 Introduction

Chapter 2 presented the Amenity Complexity Index (ACI) and the Amenity-Type Complexity Index (TCI) as adaptations of metrics of economic complexity to urban spaces. It argued that they are representations of relative compositions of commercial amenities across a city's neighborhoods, and it laid out tools within which to understand these relative compositions; But it has until now remained purely conceptual.

The next natural step is to expand upon the way complexity indices can relate to urban changes and crises is an empirical application. As such, this chapter produces a dynamic implementation of the ACI and the TCI to the city of Paris. This implementation mostly remains descriptive, but it simultaneously serves as a proof of concept, as support and amplification of the indices' methodology and interpretations as Chapter 2 presented them (see Section 2.4), and as a tool through which to further discern a research agenda built on top of the indices.

It has been established that the presence of commercial amenities depends on the coordination of suppliers (and their information) and consumers (and their preferences)

[1]This chapter is an iteration of an ongoing work co-written with DEJEAN Sylvain and SUIRE Raphaël: http://econ.geo.uu.nl/peeg/peeg2315.pdf. The co-authors have agreed to the following modifications and use.

in a selection-like process that materialize in spatially defined selection environments (places). The separation of places and of commercial amenities operated by ACI and the TCI are one-dimensional separations of underlying consumption that imperfectly depends on consumer preferences and on selection environments. It is important to restate that where this is the case along this chapter, the word "complex" is used to note contrast in compositions with other observed places (or types of amenities). Before I begin, let me therefore make a quick note on terminology. It can be difficult to emancipate the ACI and the TCI from traditional notions of "complexity"[2], and I once again outline the need to be careful when describing them, which is the objective of this chapter. As such, I avoid as much as I can referring to the ACI and the TCI in a top-down absolute manner, and try to describe places, amenity types not in terms of their complexity, but rather in terms of their positions on the spectra that are the ACI and the TCI. This terminology echoes that used by biologists with the Leaf Economics Spectrum (I. J. Wright et al., 2004) given their use of principal components analysis that echoes the ECI procedure (C. A. Hidalgo, 2021).

Building on descriptive assessments, I postulate that separating commercial amenities through the TCI equates to separating the selection environments in which they can prosper, and by deduction, to separating the consumers that elect to consume goods and services within these environments through their preferences. This is based in the premise that the presence of commercial amenities reveals the anticipation of suppliers, and that despite being imperfectly informed and boundedly rational, their anticipation yields a satisfying picture of actual present consumers.

In turn, the ACI as an aggregation of TCIs present in places tends to reflect the heterogeneity of places not just in their commercial amenity composition but in the composition of everything that constitutes them as selection environments (that is, amenities and institutions holistically and the consumers that are attracted to them).

While Chapter 2 seldom made observations or assumptions about what separated con-

[2]Chapter 2, and its Appendix 2.6.2 should hopefully help inform this issue, which Kogler et al. (2023) also refer to.

sumers in their ability to affect the selection process, this chapter's application to Paris draws another parallel to the Leaf Economics Spectrum (I. J. Wright et al., 2004). Like it, it interprets the ACI and TCI spectra as each underlying a specific continuous meaning. In the following, I defend the proposition that the amount of spending people are able and willing to dedicate to commercial amenities is strongly associated with their heterogeneous impact on selection processes, both through their different preferences and through their gaps in leverage. Because high-spending consumers can pay premiums for goods and services and because their spending ability strongly affects their realized consumption preferences (that is, their role in the selection process), they have stronger leverage on the presence or absence of the amenities they prefer by being present themselves. The ACI, by typologizing the composition of local demand and because this composition is intimately linked to spending, thus informs us on which places have overall characteristics that are attractive to higher-spending consumers.

The illustration of these notions through an application of the ACI/TCI methodology to the city of Paris reaffirms their promise as indicators of urban situations and changes, and their potential role in shaping future economic geography and planning discourse. I elect to illustrate the spectra in Paris, a particularly interesting use-case of the ACI/TCI for multiple co-dependent reasons. Most importantly, the availability of exceptional data that separates commercial amenities into extremely granular types allows for the best possible translation of the processes outlined above, and it is a good starting point to assess their validity empirically. It is also a large city in which multiple activity hubs coexist. There is no single place to which everyone would go to consume, and the evolution of different hubs can therefore be compared. Paris is also still one of the leading world tourist destinations, and an international destination for world-class business congresses and fairs. On top of this, France is very centralized in its qualified job market for economic and administrative centers of decisions. Overall, it is a city that experiences large, global and diverse flows of people that are willing and able to consume, which in turn makes it particularly vulnerable to urban tensions. All of these reasons coupled with the proactive nature of its institutions in shaping urban spaces make it a prime candidate for a first empirical assessment. Let me first present the dataset and the binary matrix $M_{n,a}$, representing the commercial amenity space. Then, I will investigate the association of the TCI and of the ACI with other network characteristics of $M_{n,a}$, before assessing the way the ACI is distributed both in a spatial and in a non-spatial way. I will finally confront ACI values to other place characteristics within a random forest procedure and discuss the implications of the ACI and of the TCI for future research.

3.2 Data

A strong argument for the study of Paris is the quality of available data. I leverage *BD-COM*[3] data from the Paris Urbanism Agency (*APUR*)[4]. *BDCom* is an on-field survey-based documentation of retail and private commercial services in Paris that is carried out once every 3 years, the objective of which is to provide an exact, exhaustive and detailed view of Parisian commercial businesses. Data are available for years 2014, 2017 and 2020 at the time of writing. Unlike *Google Places* (used notably C. A. Hidalgo et al., 2020; Juhász et al., 2023; Kaufmann et al., 2022) or *OpenStreetMap* data, *BDCom* is independent of users and of businesses, and its repeated surveys provide a dynamic perspective. This is key to unlocking ongoing urban transformations and to an evolutionary analysis. On top of this, *BDCOM*'s documentation of market-oriented amenities into categories is a lot more detailed than that of off-field alternatives or of public business registries that tend to group together amenities with poor substitution properties from consumers' perspective. The systemic nature of consumption coordination indeed implies that amenities should ideally be grouped into types in a way that is entirely indistinguishable to consumers. *BDCOm* comes a lot closer to that ideal than alternatives. Its categories are also focused on commercial amenities and tend to be better divided. For example, a 5-star hotel does not serve the same consumers as a 1-star hotel does, it does not go

[3]Available online for 2017 at: https://opendata.apur.org/datasets/Apur::bdcom-2017-1/about
And for 2020 at: https://opendata.apur.org/datasets/Apur::bdcom-2020-1/about
[4]https://www.apur.org/fr
APUR is an independent agency that is primarily funded by the city of Paris.
The *BDCom* survey is publicly funded, which is part of the reasons it is openly available.

through the same selection process and does thrive in the same environments. Likewise, designers' clothes shops are different from generalist clothes shops. These are the kinds of subtleties that give *BDCOM* an advantage over alternatives. I use 202 separate commercial amenity types that are listed in Appendix 3.8.3 along with their corresponding TCIs.

3.2.1 Building the binary matrix

Using official data (namely a *BDNB*[5] aggregation of *BDTOPO*[6] data) and following the methodology set out in Chapter 2's Subsection 2.3.3, I set the coordinates of each Parisian residential building[7] as the center of places. These building-centered spatial units can be aggregated up later on to include socio-economic data at the census level, and it should be clear that using buildings as the center of places serves no interpretative purpose beyond their practicality to efficiently grid out the city.

From these building coordinates, I use the Pereira et al. (2021) adaptation of the *Conveyal* R5 routing engine[8] along with different OpenStreetMap street networks for each year[9] to compute a travel time matrix of itineraries by foot between all buildings and all amenities. I use cutoff $c = 15$ minutes with a walking speed of 4km/h to determine the count-based matrix $O_{n,a}$ linking n places (centered around buildings) to a amenity types for each year. Because the ACI and the TCI are eigenvectors of a similarity matrix of places, their absolute values are only relevant relative to others within the same eigenvector, and comparing raw values across years has complicated implications. The most common way of dealing with this issue while observing change is to compare ranks of values, and it is the method selected in Chapter 4. However, because the span of observed years is fairly short and because the absolute difference between values remains relevant

[5] https://www.data.gouv.fr/fr/datasets/base-de-donnees-nationale-des-batiments/
[6] https://geoservices.ign.fr/documentation/donnees/vecteur/bdtopo
[7] The handful of buildings outside the ring road have a hard time finding itineraries towards the inside, and we therefore exclude them/
[8] https://github.com/conveyal/r5
[9] This allows the index to be robust against changes in the city landscape. Historical street network data corresponding to each year is provided by Geofabrik
http://download.geofabrik.de/europe/france/ile-de-france.html

to conceptual distance between places (or amenity types), I decide to include all three observed years in the same matrix and in the same ACI/TCI computation. Commercial amenities within every place (or spatial selection environment) n are thus observed 3 times: Once in 2014, once in 2017, and once in 2020. Thus, each n is actually expanded into 3 separate rows in $M_{n,a}$, the binary presence-based simplification of $O_{n,a}$. While this allows to outline changes in ACI distributions, it implicitly uses the average TCI across all 3 observed years instead of assigning per-year TCI ranks. This would not be suitable to longer time frames (see the one used later on in Chapter 4). As such, $M_{n,a}$ is a binary matrix with roughly 168,000 rows (55,985 residential buildings observed in 3 different years) and 202 columns in which cells hold a value of $M_{n,a} = 1$ when an amenity type can be found within 15 minutes of a place's center.

A quick precision should be made regarding the selection of c. The initial use of a 15-minute cutoff stems from growing literature around the 15-minute city (Khavarian-Garmsir et al., 2023; Moreno et al., 2021; Pisano, 2020). It reasonably implies that 15 minutes is an acceptable walking distance to consume, and that any element of the environment is part of the selection-anticipated location choice of retailers. It appears to be a good compromise for the city of Paris that will be large enough to uncover spatial trends, but small enough that the city is not artificially over-homogenized. Appendix 3.8.2 provides evidence that the ACI is not overly sensitive to the selection of c in Paris. Figure 3.1 gives the reader an idea of how much space the 15-minute radius represents and of how amenity-dense Paris is. It also serves to remind her of the amount of spatial autocorrelation in the data: One should imagine a continuum of buildings between the two selected points. As such and by design, a retail establishment that appears only once in the data may belong to multiple thousands of places. In this example, green dots count as present in both buildings' places.

Figure 3.1: Visual representations of 2 buildings' (that is, places') commercial amenity spaces in 2020, and of how these spaces can overlap. Map A shows all amenities within a 15-minute walk of the buildings ($O_{n,a}$), while Map B shows only those that are the closest of their type to the building ($M_{n,a}$).
Map tiles: ©OpenStreetMap Contributors

3.3 The local composition of demand

In this subsection, I evaluate key properties of the ACI and of the TCI and relate them to the initial selection framework set up in Chapter 2. To unpack this with data, I start by evaluating the TCI of commercial amenity types (TCI) to better grasp how they are separated, and to build an economic interpretation of complexity indices. Past research has focused on (and interpreted) complexity indices as directly reflective of place diversity and average amenity ubiquity. The ACI is not dependent on them through its formulation, nor is it reflective of them by design. Still, they remain correlated. I therefore follow up by explaining why these correlations exist using this thesis' framework, and use the correlations to further illustrate intuitions regarding how the indices separate consumers and environments.

3.3.1 The TCI as a separation of consumers

The ACI of places is a function of the TCI of the types of retail they host, and vice versa. The separation between commercial amenity types operated by the TCI is dependent on where amenities are found — Or where they are *not* found. In fact, the latter statement is truer in the case of Paris: it is an amenity-dense city, and $M_{n,a}$ is 80% filled. This means that at least one establishment of most commercial amenity types can be selected (and locally available) in most places in Paris, and that more information about the composition of places and the ability of retailers to be selected is gained from the lack of selection than through selection itself. This is only made possible by the exceptional granularity of amenity types in the *BDCom* dataset, and applications in less ideal scenarios should look into other ways of binarizing the matrix[10].

Types of retail that are less ubiquitous (more rare) provide more information about places as selection environments and about the consumers that consume them, and they are thus more likely to be strongly separated from the rest. Table 3.1 presents a sample of the 15 most divisive amenity types in each direction, that is, those with the highest and lowest TCIs.

Table 3.1: Top 15 and bottom 15 commercial amenity types by their TCI ranks in Paris. See Appendix 3.8.3 for all the ranks.

Rank	Name	TCI	Rank	Name	TCI
1	Luxury general food > 300 m²	6.103		...	
2	Tourist hotel - Palace	5.494	188	DIY	-0.544
3	Tourist hotel with 5 stars	2.999	189	Nurse's office	-0.545
4	Department store	2.995	190	Sale of pets	-0.591
5	Coffee shop	2.801	191	Moving / Storage	-0.666
6	Large cultural multispecialist	2.488	192	Tattoo - Piercing	-0.727
7	Ticketing - Booking shows	2.454	193	Telecommunication in store	-0.779
8	Smile Bars	2.334	194	DIY and home equipment rental	-0.792
9	Sale of coins and medals	2.06	195	Youth Hostel	-0.822
10	Haute couture - Designers	1.856	196	Discount store	-0.904
11	Gambling	1.81	197	Sale of automotive equipment	-0.962
12	Sale of erotic items and sex shop	1.524	198	Ambulances	-1.612
13	Concert hall	1.269	199	Discount supermarket	-1.94
14	Philately	1.251	200	Car technical control center	-2.301
15	Sales room	1.226	201	Specialized supermarket	-3.138
	...		202	Hypermarket	-4.919

A qualitative look at Table 3.1 reveals patterns that drive TCI values and ranks. On

[10]Chapter 4, because of its internally less ideal scenario, will define the binary matrix through an RCA.

the right-hand side, one finds goods and services that are either cheap, primarily associated with residential use, with necessity, and/or with what consumption economists might call inferior goods[11]. On the left hand side, amenities with the highest TCIs are strongly associated with tourism, high prices and margins, leisure, and could be qualified as luxury goods. The presence of coffee shops at the top is also interesting: Bantman-Masum (2020) uses them to qualify ongoing transnational gentrification in Paris. In Figure 3.2, it is clear that as amenities tend to be less ubiquitous they tend to be more informative of TCI and more distinctive from each other. The bright blue quasi-vertical line in the middle represents types of commercial amenities that are observed to be present almost everywhere. It is also clear that at an aggregate level, more ubiquitous amenity types tend to be associated with negative TCI values. The separation it presents, which can also attempt to perceive qualitatively through Table 3.1 (or Appendix 3.8.3 with the example of food stores), is the basis for defining the TCI of places as indicative of the composition of underlying demand.

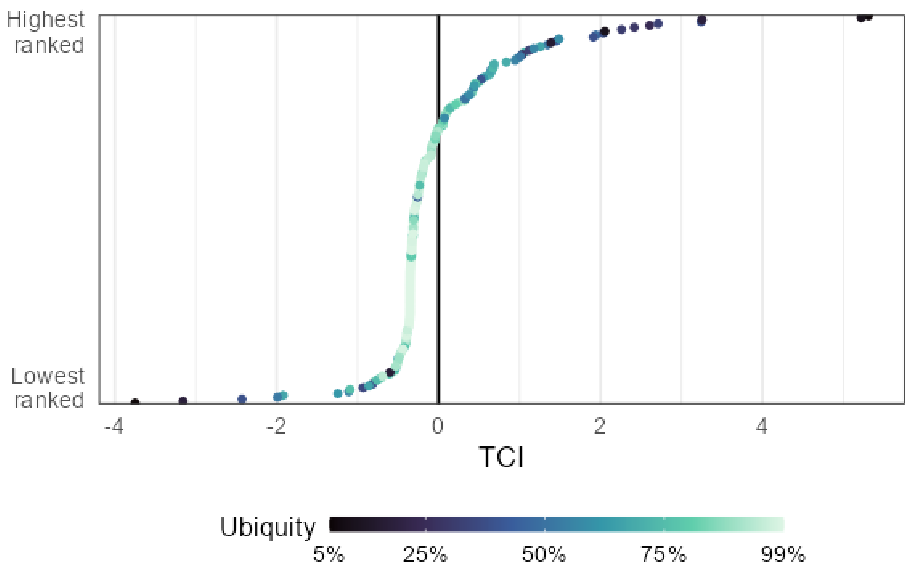


Figure 3.2: TCI values, ranks and ubiquities (as a share of places covered) of amenity types

[11]In an Engel-like framework, inferior goods are those for which consumption goes down as income goes up

In economic complexity, the separation of products is thought to be made through the sophistication of their underlying technology (C. A. Hidalgo, 2021), and, by simplification, of how difficult they are to produce. In this context, the presence of high or low TCI market-driven amenities is a reflection of the presence of consumers more associated with one type of commercial amenities than with another, both because amenities need the right consumers in order to be sustained through a selection process and because they are susceptible to attract consumers that desire them within the places they are present in. Understanding the TCI and the ACI requires an appreciation of what it is that motivates a uni-dimensional separation in consumption, that is, peoples' propensity to (co-)consume different goods and services from different commercial amenities. Although the answer is necessarily intricate, I argue that consumers' spending power is intrinsically linked to the nature of their consumption, hence that the type of commercial amenities as separated in the TCI is broadly representative of the spending power of locally present consumers.

First, spending power separates realized consumption as a direct factor. For given preferences, people might allocate more to leisure-oriented amenities when they have more ability and willingness to spend. This specific example is consistent with Engelian frameworks of consumption, where income-driven preferences for leisure and diversity are found (Jackson, 1984; Table 2 in Aguiar and Bils, 2015). Second and indirectly, it does so as a confounder of possible social, cultural, and other consumption determinants that would be difficult to quantify alone. In Paris specifically, consumers' residential status[12] could be a relatively important determinant because non-residents are likely to consume different amenity types than residents do. However, tourists also come with a lot of spending power allocated to amenities relative to residents for the short time they are present, and separating consumers by spending power will tend to separate tourists from residents. This is how and why tourists transform the commercial amenity landscape (and arguably the overall amenity landscape) as outlined by Bordeaux's regulators (see Chapter 1) and the literature (A. Hidalgo et al., 2024). The ties between culturally

[12] To be understood in the context of tourists vs. residents

or socially differentiated consumption practices and peoples' ability to spend are not a groundbreaking idea either, and they fit into a Bourdieu-indulging view of economic capital associated with consumption-determining *habitus* (Bourdieu, 1987).

In fact, what determines consumption preferences is likely a set of deeply interrelated factors and complex interactions both internal to consumers and external with the selection environments they frequent. However, as uni-dimensional separation, any separation of consumption preferences the TCI implies should be characterized by a uni-dimensional concept[13]. This section's argument that spending is as strong a representative differentiating characteristic of consumption preferences as any is both conceptually and observationally grounded.

Moreover, the objective here is not to directly infer spending power levels from the indicators. But they can unveil tendencies within places' and amenities' compositions of local realized demand, and the likelihood of spending power playing a large role in the process of selection in any setting is very high. Thus these operational confoundings between spending power and other consumer preference characteristics are a strength rather than a weakness. They can help characterize consumption in a broad way that allows for more flexibility of the indicator across different cities with different local contexts. High-TCI amenities are consumed by consumers that select them, and the preferences of these consumers are strongly related to their spending power in part because spending power enables these preferences to be realized. In the following, I evaluate the underlying structure of places' amenities in Paris to consolidate the understanding of how the ACI separates places, and of what is associated with the consumption of higher or lower TCI amenities.

3.3.2 The ACI and network characteristics

Unearthing links between complexity indices and network measures of diversity and ubiquity was at the heart of early ECI approaches (C. A. Hidalgo and Hausmann, 2009), and

[13] This echoes biology's similarly derived leaf economics spectrum (I. J. Wright et al., 2004), which represents leaves' relative efficiency in utilizing nutrients.

they have even been used to fundamentally interpret the indices. Recent related urban work by Juhász et al. (2023), when applying a standard ECI to neighborhoods' overall amenities, proposes that:

> *A neighborhood has a complex amenity mix in case it offers a diverse set of amenities of those types that other locations are not specialized in. On the contrary, complex amenities are those that only few neighborhoods are specialized into and are co-located with diverse sets of similarly non-ubiquitous amenities.*

This proposition, while tempting given earlier ECI-driven research, is not adopted by this work because it is unrelated to the indices' formulations (Kemp-Benedict, 2014; Mealy et al., 2019)[14]. Still, it remains true that Juhász et al. (2023) find correlations between complexity values, diversity, and ubiquity[15]. These network measures do not motivate an interpretation of complexity values alone, but it is also unlikely to find relationships as strong as they did by chance[16]. These correlations also replicate in this chapter's Parisian data (albeit at a lower scale). ACI/TCI values do not come as a result of diversity nor of ubiquity, but both of these measures are affected by retail location dynamics that also affect the ACI/TCI. In other words, while diversity and ubiquity do not *inherently* affect the ACI, the selection process and its underlying market forces that is separated in the TCI/ACI also commonly affects diversity and ubiquity values. Using the understanding of ACI/TCI outcomes and of retail locations as it has been built up until now in this thesis, I unpack the reasons for which these relationships exist.

Places' diversity ($K_{n,0} = \sum_a (M_{n,a})$) and average ubiquity ($K_{n,1}$ as equivalent to $K_{c,1}$ in equation 2.1) are key characteristics of the amenity presence network. The two leftmost panels in Figure 3.3 tell us that places' ACI levels are (weakly) correlated with the

[14] Appendix 2.6.2 should also remind the reader that any characteristic, even the ones uncovered in this work, cannot straightforwardly lead to high or low complexity values. They simply lead to separations along the continua. Under the hypothesis that a characteristic affects the separation the complexity methodology drives (which diversity and ubiquity do not do intrinsically), it is up to the practitioner to arbitrarily define whether that characteristic will lead to "high" or to "low" levels of complexity. As such, an absolute proposition like the one above of what higher complexity is linked to is misleading without appropriate assumptions and clarifications.

[15] But notably, they find these based on a matrix of Revealed Comparative Advantages, not of presences. More accurately, their complexity measure is correlated to places having more different competitive advantages, and to amenities being hosted with a competitive advantage by fewer places.

[16] Although as will be shown in Chapter 4, the authors' use of a RCA to determine the binary matrix has strong implications towards these correlations

diversity of commercial amenities they host, and with how rare their amenities are. The rightmost subfigure of Figure 3.3 echoes an inverse relationship between the diversity of places and their average ubiquity that hints at nestedness, a concept that emanates from almost century-old ecosystemic analyses of species in biology (Hausdorf and Hennig, 2003) and is a staple of early economic complexity analysis (Hausmann and Hidalgo, 2011).

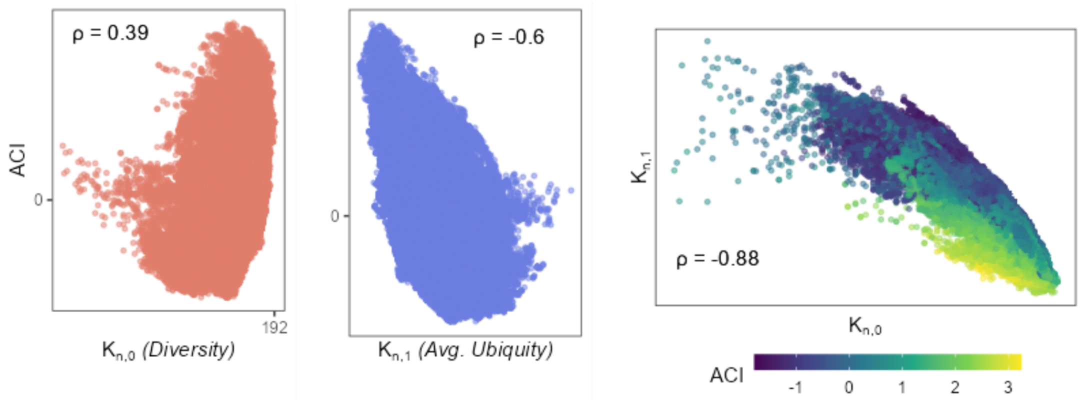


Figure 3.3: Diversity, ubiquity, and complexity of places with c = 15 in all 3 studied years. More diverse places tend to have a lower average ubiquity of amenity types, and the ACI is correlated to both of these measures. Spearman correlations are shown within each plot.

In unpacking how diversity and ubiquity measures relate to the ACI, I seek to answer how patterns in place similarities relate to these metrics. In other words, why is it that places that host many different types of amenities tend to host the same kinds of amenities? Conversely, why do places that host rarer kinds of amenities also tend to host the same ones? I argue that a spending-power based explanation of the TCI/ACI spectra offers suitable answers to these questions.

Take the premise that retail-diverse places tend to intake larger flows of population, and that they also tend to intake a more socio-demographically diverse set of people. The relatively straightforward underlying intuition is that, because different people have different preferences, their agglomeration in places makes it likelier for more types of businesses to flourish. This has conveniently been empirically proven by Chong et al. (2020) using credit card data and customer information.

This implies that environments have a tendency to be commonly attractive for consumers to exert selection in. This so-called attraction to places stems from a complex process that includes everything that is a part of the selection environment. But indeed, accessibility, aesthetics, safety, and walkability are all among characteristics of places that are likely to be enjoyed by many regardless of their preferences for types of retail, which leads to an unevenness of place diversity.

However, as commonly desirable as they may be, attractive environments do not have enough space to accommodate everyone's retail preferences. Crucially, higher spenders have more leverage in the selection process because they are willing to pay more for goods and services. In competitive environments that are tougher and that lead to higher costs for firms through retail property and rent prices, businesses that cater to higher spenders will be likelier to be selected and to thrive. As such, places that possess characteristics that everyone seeks (including high spenders) will tend to host a disproportionate amount of high-spending-catered amenities. Conversely, low spenders will tend not to have the leverage required to select amenities in them. As seen through the application of Chong et al. (2020), these also happen to be places that tend to have higher amenity diversity. Hence, there is a positive observed relationship between the ACI and diversity – Not because diversity differentiates places in the ACI's formulation, but because diverse places tend to disproportionately host amenities that higher spenders consume[17] (and to disproportionately *not* host shops only lower spenders consume in).

This ties into the question of the ACI's relationship with ubiquity. What the middle quadrant of Figure 3.3 tells us is that the kind of amenities places host and the average rarity of amenities places host are related. But as shown earlier in Figure 3.2, the relationship between the TCIs of amenity types and their ubiquity is non-monotonic. In fact,

[17]Note that in their work, Juhász et al. (2023) suggest a non-null relationship between places' complexity values and socio-economic mixing, but they also suggest a null relationship between diversity and socio-economic mixing. The first relationship is one I also imply here: Higher spenders locate in places that are commonly attractive, and that other spenders therefore also visit. But their lack of a relationship between socio-economic mixing and amenity diversity is very surprising at first glance, especially given previous results (Chong et al., 2020). This can however be explained by the fact Juhász et al. (2023) base their diversity measure on the number of different revealed comparative advantages in places, not directly on the number of different types of amenities they host. In their matrix construction, a commonly attractive place that hosts many different types of amenities also mechanically makes it less likely for each present amenity type to gain a comparative advantage. On the other hand, and while unattractive places host less different types of commercial amenities, the few they host will be likelier to each exhibit comparative advantages because their markets are smaller.

there are two clear and polar opposite sets of low-ubiquity amenities that locate in very different places: Those that relatively pull the indicator up and those that pull it down. Call the first kind of rare commercial amenities that are associated to high TCI values *luxury* amenities, for which relative consumption goes up with spending. The second kind are low-TCI *inferior* amenities, for which relative consumption goes down with spending. The ACI being negatively correlated to ubiquity means that luxury amenities tend to be selected together in places that on average host rarer amenities, whereas inferior amenities tend to be selected in places that otherwise tend to host non-rare amenities. Let me explain why this is the case.

Income-driven preferences for consumption diversity (Aguiar and Bils, 2015; Jackson, 1984; as seen in the previous subsection), the structure of the dataset in how it separates amenities into types, the higher dependence on agglomeration effects of leisure-oriented goods and services, and the more granular tailoring of commercial activities to high spenders' tastes are all partly reasons for this. But I want to focus on how the extra leverage granted to higher spenders in retail locations is likely to lead to their preferred amenities being closer together. Crucially, higher-end rare amenities are better tailored to specific preferences, and they thus require larger combinations of people with different tastes. Many of them will tend to locate in areas that are commonly attractive because they require mixing.

For the same reasons of higher-spender leverage in the selection process mentioned above, consumers of luxury amenities have more weight in their consumption locations than those of inferior amenities. On the one hand, luxury amenities need to locate where their higher spending consumers want to be, and these tend to be the same places for different types of luxury retailers. On the other hand, and because low-spending consumers have less leverage, inferior amenities tend to base their locations more on other factors that are less commonly shared by different types of retailers, such as logistical advantages. In other words, different types of luxury retailers have more common ground in where they can be selected than different types of inferior retailers have. Inferior amenity types thus end up being spread out in the city into different places, whereas luxury amenities concentrate in higher-rarity associated and commonly attractive places. As Figure 3.3's rightmost quadrant exhibits, higher rarity also has a strong tendency to be found in the same areas that exhibit higher diversity.

Diversity and ubiquity as concepts bear no impact on the TCI nor on the ACI through their formulations. But they do happen to be correlated to the indices in this example. Market mechanisms lead higher spending-associated amenities (high TCIs) to be skewed towards neighborhoods with a broader set of different available retailers (higher diversity), and in places that have retailers others do not (lower average ubiquity). Note that the TCI/ACI's spending-based separation is not based on diversity nor on ubiquity, but that spending attributes are also expected to impact those network measures. The fact these relationships are observed is thus reassuring regarding the ability of the TCI (and hence of the ACI) to separate consumers through spending. This subsection has also allowed to explore intuitions underlying the idea that high-TCI (or luxury) amenities, through the relative leverage of their high-spending consumers, locate in more desirable places.

Some of this section has been descriptive and explanatory of a Parisian-specific context that holds questionable external validity, but it has also delimited assumptions within which to inform a deeper interpretation of both the TCI and the ACI. The TCI, in its representation of the composition of demand for amenities, is a spectrum of the co-consumption (or co-selection) of commercial amenities their suppliers individually anticipate. The single element that best defines consumers' differentiated realization of consumption is their amount of spending. A spending spectrum is the most valid holistic reason on which the selection process conditioned by consumers, because it determines what they can and cannot consume. On top of consumption coming as a function of spending, it comes as a function of many other unobservable elements that spending itself is a function of.

Moreover, people who are able to consume a wider variety of commercial amenities through their spending exert a stronger force of selection upon retailers. They therefore have more leverage in the selection environments their preferred commercial amenities can thrive in, and these high-TCI commercial amenities end up being located in places high spenders are most attracted to. By deduction, the ACI can be interpreted as revealing the anticipated presence of high spenders, and in turn, revealing the places these high spenders are anticipated to be attracted to. In short, this section has amplified the interpretations outlined in Chapter 2's Section 2.4 in the following simplified way: The TCI is a spectrum of revealed preferences for amenities based on relative spending, and the ACI is a spectrum of economically revealed (in the sense of Samuelson (1948)) relative place attractiveness based on relative spending.

The following section confronts the ACI and the TCI as spectra along with their new interpretations to a spatial analysis within Paris, and delineates means through which it can be related to urban changes and tensions.

3.4 The ACI's distributions

In the previous section, I have presented how the ACI can help unlock consumption structures in a broader way. I proposed that, as an indicator that is intrinsically linked to local consumers' spending power, the ACI is therefore intrinsically linked to the relative attractiveness of both commercial and non-commercial amenities (examples of which are nature, accessibility, public services, or aesthetics), or of place characteristics, or of selection environments. I now illustrate this with map-based approaches to Parisian ACI representations of places, and leverage the dynamic data at hand to show how the ACI can be a valuable proxy for urban transformations on a broader scale than just commercial amenities. I then abstract from space to assess the ACI from at an overarching city scale and explore how it can provide new perspectives for local-level institutions.

3.4.1 Mapping the spectrum

To help elucidate the distribution of the ACI spectrum and its evolution, I look at the ACI in Paris through a qualitative spatial analysis. Prior knowledge of local context is key to making the best out of the spectrum the ACI represents, and mapping is a good place to start. Figure 3.4 maps the spatial distribution of mean ACI ranks[18] throughout the observed period (A) and the absolute difference between the 2020 ACI and the 2014 ACI for every place[19] (B).

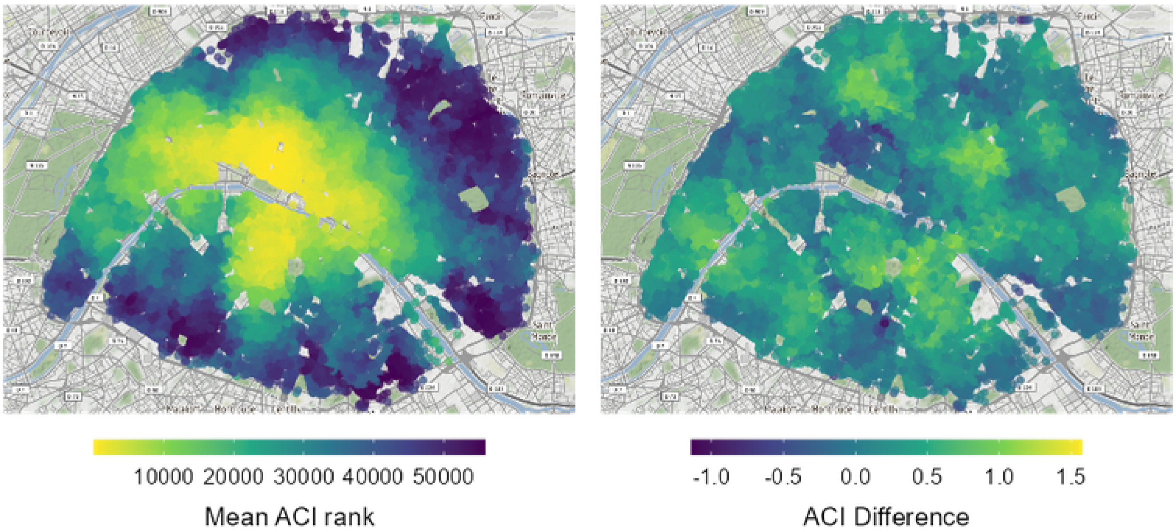


Figure 3.4: **(A)** Map of the mean ACI ranks of places in Paris in 2014, 2017 and 2020. Lower ranks are associated to higher ACI values; **(B)** Absolute change in ACI over the observed period (2020-2014) Map tiles: ©OpenStreetMap Contributors

The assumptions behind the TCI and the ACI that grant them economic meaning are that they typologize localized demand, but many factors may affect such demand. Having a central position is evidently an important differentiating factor of places in the ACI. In an inland city like Paris, being further away from the city's ring road is positively correlated to higher ACI values. This is consistent with previous attempts to map human activity in cities (Zhong et al., 2017), with Juhász et al. (2023)'s implementation of amenity complexity where they find a direct link between geographical centrality and their version of the ACI, and with urban central place theory (Christaller, 1966) – Although the approach of this work is fundamentally different.

In this work's framework, the centrality of locations can be seen as a proxy for acces-

[18] ACI ranks, much like ACI values, should not be understood as absolute top-down distinctions. It helps to imagine them as a right-to-left of 0 classification instead of a top-to-down one.

[19] As a reminder, this can only be done because the TCI was computed from a single matrix wherein each place was represented in 3 rows, one for each year. See Subsection 3.2.1

sibility. Accessibility is however not the only explanatory factor behind the relationships people have with selection environments, nor is it the only latent the ACI can reveal. The spatial distribution of the ACI in Figure 3.4 (A) is not homogeneously decreasing from a point in the center to the periphery. Very high ACI-ranked places in the center-West tend to be hotspots for traditional tourism and internationally renowned brands; On top of being well located in terms of their accessibility, they offer aesthetically pleasing architecture (For example the *Champs-Elysee*), are safer than peripheral alternatives, host museums (*Le Louvre*) and monuments (*Arc de Triomphe, Palais Royal, Opéra Garnier*). They are also not too far away from the river banks of *La Seine*. I argue that, like other non-commercial amenities of consumer city thinking (Glaeser et al., 2001) that include aesthetics and public goods and services, accessibility induced by central locations is a desirable trait of places that suits the idea of a spending-spectrum based selection process of amenity locations. Accessibility is attractive, and as such, high spenders elect to consume in accessible environments, and they select commercial amenities associated with their own preferences that hold higher TCIs.

Attractiveness is not assumed to be universally defined across consumers, and different people might prefer different things. Still, under the assumptions of TCI separation being strongly linked to spending and of spending granting leverage in selection, the ACI yields the place preferences of those with more spending power *on balance*, as anticipated by suppliers of commercial amenities. While this interpretation seems to fit well in Paris from a qualitative look at Figure 3.4, it begs the question of how places come to be attractive through time. Paris has had time, throughout history, to develop relative place attractiveness in ways that can make it akin to path-dependent and lock-in phenomena that are insisted upon in evolutionary economic geography (Boschma and Lambooy, 1999; Kogler et al., 2023; Martin and Sunley, 2006). It is notable that the "*rive droite*", the part of the city that is north of *La Seine*, exhibits higher ACI ranks than the other side. This echoes the historical development of the city that already had its trade, its economic activity and its population development focused on the *rive droite* throughout the Middle

Ages, with the south side of the river hosting convents and universities. These universities were however located in the bright part of the *rive gauche* (South of the *Seine*, near the *Pantheon* monument), and high ACI ranks on that side appear to struggle to break the old city boundaries.

The role of urban centrality (and of other non-commercial amenities) is a lot less obvious as a feature of ACI changes in Figure 3.4 (B). Central (and high-ACI) places have not necessarily been getting more extreme in terms of their ACI between 2014 and 2020. In fact, there is no visible indication of a self-reinforcing process of high (and low) ACI places in the observed period, nor of a polarization of compositions of demand. Some relatively central places have ranked up, but not those that held the highest ACI ranks. Instead, Western traditional tourism hotspots appear to be closer to the rest of the ACI spectrum. At the same time, some initially less attractive places in the periphery have been getting closer to them. It is critical to note however that changes outlined in the figure are low in scale and are not paradigm-shifting – Commercial amenities (and amenities as a whole) are sticky, they require a lot of capital, and even if strong changes had happened in the local composition of demand over the observed period, they would take time to be reflected in the ACI.

Figure 3.5 combines the 2014 ACI ranks and changes in ACI into a single map in order to further unlock spatial patterns, and demonstrates how the ACI can be related to urban transformations.

There is a lot to unpack from this map. The first element is a confirmation that there is no clear pattern of high-ACI places moving relatively rightwards on the spectrum (that is, become more extreme) in time than low-ACI places (Spearman $\rho = -0.13$ between ACI ranks and ACI changes), or else the blue parts would be very bright. While self-reinforcing processes have almost certainly played a role historically, they do not seem to be at play in the observed period.

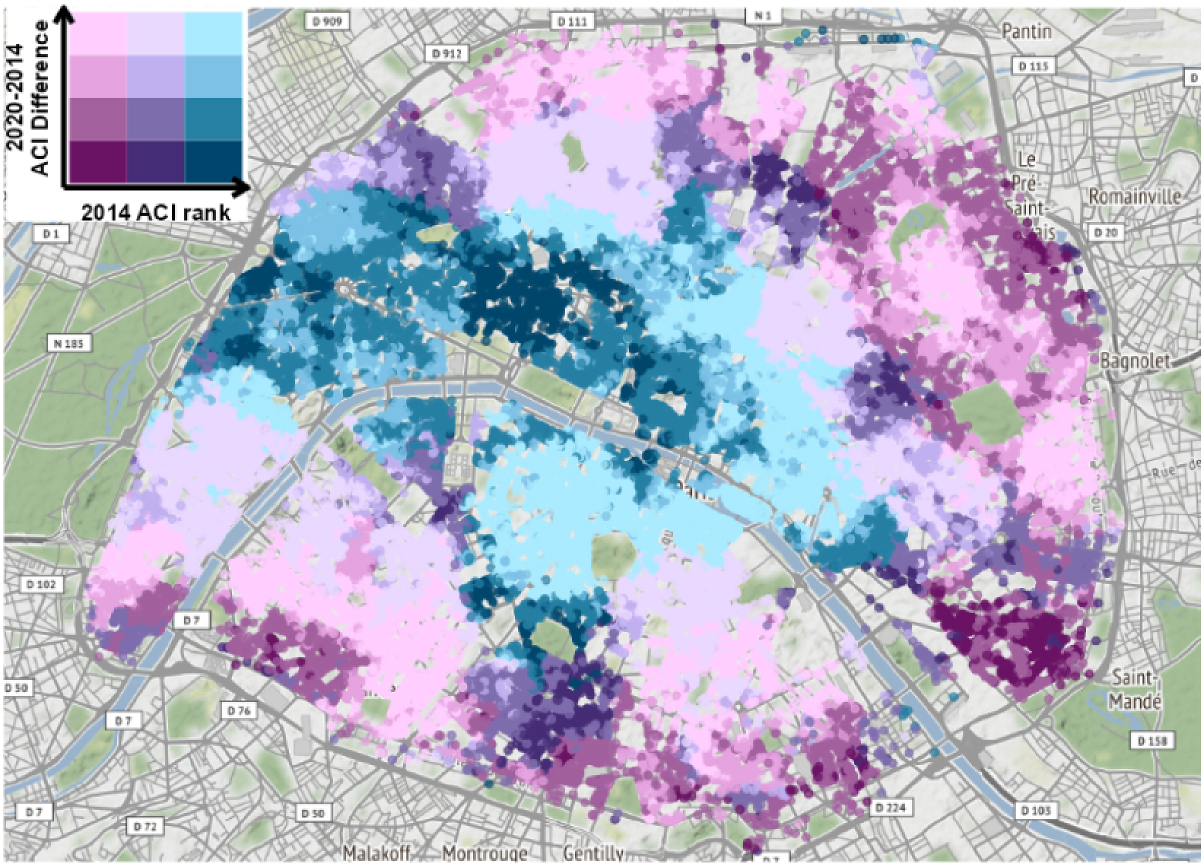


Figure 3.5: This map combines the rank of places on the 2014 ACI spectrum and their shift on the ACI spectrum through time into a single map. Darker colors denote ACI downward shifts and brighter colors denote ACI upward shifts. Different hues denote different tertiles of 2014 ACI. Brightness divides changes in ACI levels into 4 different groups, where the darkest are downward shifting (towards the lower-spending associated side of the spectrum), the second darkest are stable, and the 2 brightest are upward-shifting at different speeds.

With ΔACI_n the 2020-2014 difference in ACI in place n and σ the standard deviation of ACI differences:
Downward shift (dark) if $\Delta ACI_n < -\frac{\sigma}{2}$
Stable (moderate) if $-\frac{\sigma}{2} \leq \Delta ACI_n \leq \frac{\sigma}{2}$
upward shift (bright) if $\frac{\sigma}{2} \leq \Delta ACI_n \leq \sigma$
Strong upward shift (very bright) if $\Delta ACI_n > \sigma$

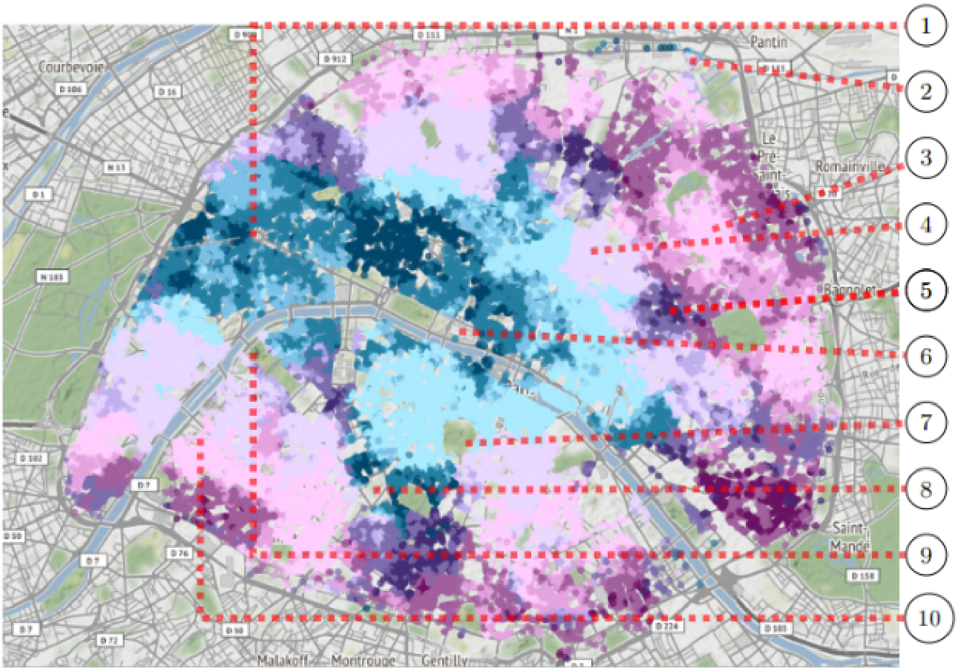


Figure 3.5: Annotated version to illustrate examples in the text
Map tiles: ©OpenStreetMap Contributors

Among average-to-high ACI-ranked places (Purple and Blue), there has been an upward shift in the ACI of places located towards the East and the North. Center-west places, the more traditional tourist center that expands from the city's main intra-travel transport hub *Châtelet-Les-Halles* and the nearby *Louvre* museum (Pointer [6] in Figure 3.5) westwards to the city's periphery, have either retained their values or shifted towards lower ACIs between 2014 and 2020 and thus towards lower relative spending of local consumers. This darker-to-moderate blue area includes important historical and touristic landmarks like the *Louvre Museum* ([6]), the *Palais Royal*, *Concorde* square, the *Champs-Elysée*, the *Arc de Triomphe* ([1]) and the *Palais Garnier* opera. These remain some of the city's highest-ACI places places in 2020, but they are being caught up to by places in multiple other neighborhoods. I notice a large semicircle of upward-shifting (towards higher spending) ACI values around the Eastern side that extends roughly from the *Canal St-Martin* ([4]) to just past the *Luxembourg Gardens* ([7]) into the *Quartier Latin*. Although Paris is a tourism-oriented city as a whole, what opposes this semicircle the most to places in the Centre-West is the fact that these are not traditionally as tourism-oriented. They host landmarks that can be visited, but they are also places with strong local implementation that include universities, important nightlife and theaters. The peaks in ACI shifts on the rive droite are located just North of *Le Marais*, around the *Canal Saint-Martin* ([4]). It is clear that some of these rapidly shifting places are disproportionately prone to urban transformations and gentrification. This is especially credible on a transnational scale for the city of Paris, in generally explained by Sigler and Wachsmuth (2020).

The *Canal Saint-Martin* ([4]) is a perfect example of lifestyle-driven gentrification, even if it is not necessarily only transnational. Changes in consumption observed through the ACI in a spending-spectrum interpretation are backed up by a gentrification sentiment that has been ongoing for years and is expanding in space, enough so to earn the neighborhood a gentrification-linked nickname outlined by journalist Hasse (2021) in Le Parisien[20]. A case study by Bantman-Masum (2020) uses a specific area at the junction between the 4th and the 11th district that is just south of the *Canal* to better understand ongoing commercial gentrification through the example of coffee shops.

The hypothesis that the rise in Short-term rentals through platforms such as *Airbnb* also has had a role to play in these transformations in the local composition of demand (as anticipated by suppliers) is not far-fetched either (A. Hidalgo et al., 2024). The importance of this element, and hence of the information the ACI can provide, is exacerbated by the policy-level recognition of these phenomena (see Chapter 1).

The map alone cannot tell us if these places' ACI shifts are due to tourists visiting them instead of traditional western hotspots. However, despite the lack of polarization of the ACI at a macro city-wide scale, neighboring places could be polarizing relative to each other. In fact, while the spatial distribution of the ACI is quite continuous (see Figure 3.4), I find stark contrasts in ACI shifts on the outward edges of some strongly

[20]"[...] like in the southern area of the canal Saint-Martin (10th district), which is already known as «Boboland»"–Le Parisien, 2021
"Boboland", or land of the "Bobos", a pejorative word for bourgeois-bohemian people

shifting places. A great example of this is the shrinking of ACI in places at the junction between *Belleville* and *Folie-Méricourt* ([5]). This dark area is stuck between the rightward-shifting *Canal Saint-Martin* ([4]) and the *Village Jourdain* ([3]).

Jourdain ([3]) is a historically lower-income neighborhood that has been in the process of "accelerated" gentrification in recent years. Communities within that atmospheric neighborhood describe it as an "*enclosed area*" with an "*invisible frontier*" to *Belleville* in the press (Georges, 2020), motivating the concept of the *Jourdain Village.* The same *Le Monde* article asserts that the *Jourdain Village* ([3]) "brand" is an asset to real estate agents to differentiate the neighborhood from others in the area. From a consumption standpoint, it is plausible that the area around pointer [5] is struggling to attract outsiders with high spending power precisely because of its proximity to the attractive *Canal Saint-Martin* ([4]) and *Village Jourdain* ([3]).

The ACI can be attached to real-world meaning, historical path dependencies and perceptions of local communities and development in an extensive way. The post-industrial development around *Porte de Pantin* ([2]), the separation of consumption created by the *Luxembourg Gardens* ([7]) and the *Montparnasse Station* ([8]), the lack of complexity around the *Eiffel Tower* ([9]), and the rejuvenation around the new mall in *Beaugrenelle* ([10]) are just some examples of the ability ACI measures can have to go hand in hand with local knowledge to better understand narratives surrounding urban situations and transformations.

The ACI and the TCI, as holistic spectra, could help objectivize narratives surrounding real estate prices, short-term rentals, community life, atmosphere, and in a general sense be confronted to processes of gentrification that are at the heart of urban tensions. However, I must reiterate that they are the result of centuries of systemic evolution. The scale of shifts over such a short period of time will therefore seldom be paradigm-shifting. Despite this, viewing the ACI and the TCI through a macroscopic lens provides insight over cities' evolutions in a way that is less sensitive to noise and to slow evolution than the localized lens used until now. To do so, I present the distributions of the ACI in Paris, their evolution, and relate them to a more general view of cities as complex systems of places.

3.4.2 Non-spatial distributions

Here, I explore ACI *coherence* within cities, that is, how wide the spectrum of ACIs is and how places are placed along it relative to each other. In other words, I look at the distribution of the ACI, which as a reminder holds information about the revealed relative attractiveness of places as anticipated by suppliers of commercial amenities. Policy-makers directly or indirectly affect this through planning strategies, but it is also the result of market forces, laissez-faire, and centuries of historical developments.

Looking at how these distributions change over time (ideally over longer windows than those used here) is a good indicator of how much more or less centralized a city is becoming in its ability to generate high-or-low TCI consumption. Figure 3.6 plots the Amenity Complexity Index distributions in Paris for the 3 observed years. Here, positions of indexes are relative to the indexes of other places and over all years thanks to the way the network was constructed.

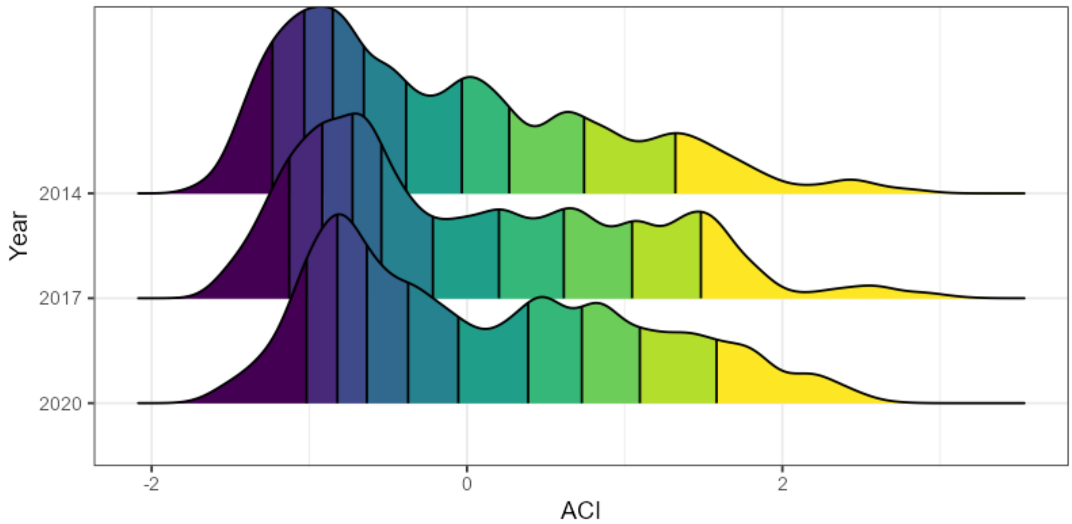


Figure 3.6: Density distributions of the ACI in Paris in 2014, 2017 and 2020, colored by decile.

As consumers with higher relative spending have more leverage in commercial amenity locations, policies that affect how non-market characteristics of places are distributed across cities necessarily affect relative attractiveness of places. This is the case of intended efforts to change how accessible different places are, which the Paris-embraced 15-minute city concept is a part of (Khavarian-Garmsir et al., 2023; Moreno et al., 2021; Pozoukidou and Chatziyiannaki, 2021). But it is also intertwined with an infinity of unintended and sometimes hidden phenomena that affect the preferences of consumers for amenities and for selection environments. For example, paradigm shifts like that of short-term rental platforms are likely to have changed how attractive different places are for tourists to consume (or exert selection) in. These potential shifts would be reflected in the ACI as long as suppliers react to, or anticipate, tourists' presence. The relative concentration or dilution of the ACI reflects the spatial diffusion of high-spender consumption, and thus, the relative homogeneity of places in their ability to attract consumers with higher relative spending power. The ambivalence of homogeneous or heterogeneous attraction effects fits into important, broader discussions about urban transformations and how we want to live in cities.

In Paris specifically, it is noticeable that places have slowly moved away from lower

ACIs towards higher ACIs over time. This is true for every ACI decile, but it seems primarily driven by the leftmost (low ACI) peak moving rightwards. Places with poor abilities to generate consumption of higher spenders are, in general, slowly gaining more ability to attract them. On the other hand, the most attractive places are less segregated from the rest in 2020 than they were in previous years. In concrete words, it seems that suppliers anticipated consumers to be diffusing their spending more equally across the city in 2020 than in 2014, leading to a de-concentration of the ACI, with the possible ambivalent effects mentioned in the previous subsection.

This is however an average result that has to be moderated by its size, and qualitative spatial analysis (such as that in Figure 4.3) should help the practitioner understand these changes. For example, the leftward move of high-ACI places between 2014 and 2020 can possibly be put down to traditional tourism hotspots comparatively losing out in high-spending consumption. The reader should once again however keep in mind that changes in commercial amenity structures are a part of slow ongoing processes that take decades to unfold.

3.5 Random forest

This chapter has related the ACI and the TCI of Paris to latent patterns of consumption-conditioned selection processes that are themselves related to local consumer characteristics (most importantly, to spending), and motivated place characteristics (attractiveness) as determinants of consumer presence. To further cement the idea that the ACI spectrum is determined by latent place and socio-demographic variables, I naively relate the ACI to other observables that should contribute to explaining both consumers' preferences for different goods and services and place characteristics. I find that the ACI correlates well to a number of different socio-demographic variables, but seldom in a linear way and often in combination with others. Using the M. N. Wright and Ziegler (2017) implementation of random forests algorithms (Breiman, 2001), I show that proxies for local place and demand characteristics are apt at predicting the ACI. I use the share of primary residences built before 1919, the median income of the place's census tract[21], the Gini coefficient of income in the census tract, the share of dwellers aged 25 to 39 and a measure of nearby short-term rental visits (see Appendix 3.8.3 for data sources).

As shown in Appendix 3.8.3, the mean squared prediction error on out-of-bag observations is 0.0174, which is about 1.7% of a standard deviation of the ACI. The model's variables therefore offer excellent predictions of the ACI. Random forests are not linear at an aggregate level, and they do not assign overall coefficients nor signs to variables, that is, it could not have predicted the ACI without the prior knowledge of how the ACI related to these variables in Paris through training. However, these models can still assign importance measures based on how much prediction accuracy is lost when the variables' values are randomly permuted. These measures are meant to outline how impactful a variable is within the model and are presented in Table 3.2, along with the Spearman correlation of the variables to the ACI. For example, on average, the R-squared of the regression only decreases by 0.15 when year dummies are randomly permuted in a tree, as opposed to 0.79 when the Gini coefficient of income is permuted. Features however combine in ways that are difficult to quantify to yield different values; For example, the share of 25-to-39 year olds among primary residents is negatively correlated to the ACI, but it could very well contribute positively in higher-income places that have older buildings in the random forest setting. The contribution of this regression to this chapter is mainly to demonstrate that the ACI is representative of elements that constitute a place and its consumers – The ACI and the TCI contain information about socio-demographic and place characteristics. Going back to chapter 2's conceptual mapping of relationships between consumers and places as represented in Figure 2.1, these results outline that the ACI/TCI methodology is informative of edges between consumers C_c and places P_p, both

[21]There are 992 census tracts in Paris, called "IRIS". Spatially interpolating census-level data to obtain place-level granularity using Kriging methods yields even better random forest estimates – That is, the models presented here models do not overfit on census tracts. Interpolating data however makes our models less transparent and more convoluted, so I opt to keep them at a census level.

through consumer compositions and preferences, and through place characteristics. But these elements are intrinsically complex in the way they interact with each other, and they impossible to individually generalize from this exercise alone.

Still, and critically in the broader context of this thesis, Short-Term Rentals (STR) nights spent is a variable that comes out as strongly linked to ACI/TCI levels, with a 0.51 Spearman correlation and the third-highest feature importance. Higher STR activity is associated with places that exhibit higher spending. An important underlying assumption of the ACI/TCI endeavour in this thesis has been that STR activity is linked to changes in compositions of demand (that is, of C_c) and of retail within cities that reflect broader urban change. While there is no assumption of causality in the results presented here, they still suggest that the ACI/TCI methodology holds information regarding compositions of local demand, that some of the difference in these compositions is informed by STR activity, and that the ACI/TCI are therefore fitting tools to assess STR-driven urban transformations.

	Feature Importance for predicting the ACI	Spearman correlation to the ACI
Median income	0.38	0.53
Gini coefficient of income	0.79	0.71
Share of primary residencies built pre-1919	0.58	0.73
Short-term rental nights spent	**0.38**	**0.51**
Share of 25-39 year olds	0.24	-0.09
Year dummies	0.15	$\emptyset$

Table 3.2: Importance of different variables within the random forest model (using the full data), correlations of the same variables to the ACI, and random forest model characteristics.

To add to the regression and to double-check that data is not overtrained, I train the same random forest model on a subset of 70% of the data[22]. The other 30% are then used solely for prediction purposes, the results of which (plotted in Figure 3.7) confirm that the predictor is very accurate. Moreover, Figure 3.7 also makes the point of a

[22]Results outlined above are already on out-of-bag data in a similar procedure that is integrated to the algorithm, but this double-checking provides an additional layer of caution

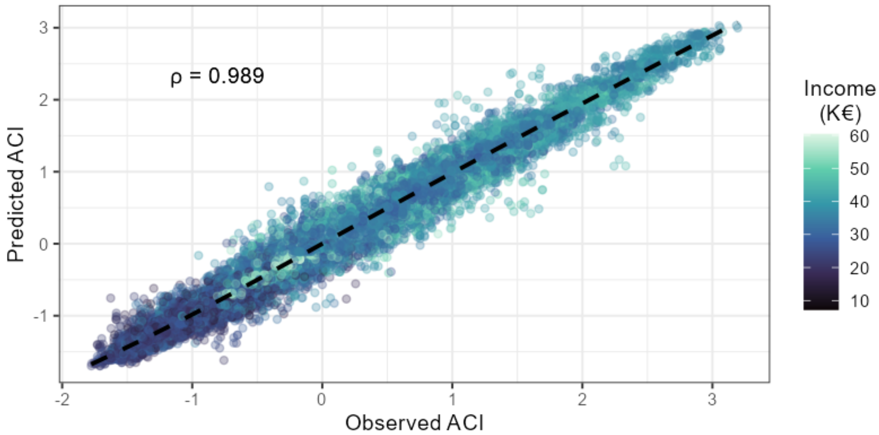


Figure 3.7: Random forest prediction results on a test subset of places, and tested places' actual ACI values. The dashed line is a loess fit of actual observed ACI values against predicted values. Observations are colored by income to outline the imperfect positive relationship it has with the ACI.

non-linear (and non-monotonic) relationship between census tracts' median incomes and their places' ACI values. A place with richer residents does not systematically equate to higher TCI anticipated consumption (that is, richer people do not live in places wherein high spending consumption is observed to be anticipated by suppliers), but combinations between median income and other variables are useful towards finding high-TCI-linked consumption, as evidenced by the figure. This is an important enough point to be reiterated, especially because the buildings-led definition of places might have confused the reader: The ACI and the TCI are never assumed to be solely based on the anticipated consumption of residents. Residents of a place could be likelier to relatively consume more in their own spaces in some contexts, but this framework completely allows them to seek different place characteristics within which to live (tranquility, for example) and within which to consume. In fact, the use of median income and of the Gini coefficient of median income here could just as well be related to the ACI as confounding characteristics of the selection environment.

3.6 Discussion

Following in the footsteps of Chapter 2 that laid the grounds of an economic complexity-driven methodology applied to urban spaces, this chapter has applied ACI and the TCI to the city of Paris. In doing so, it has outlined associations between the spectra and extremely relevant urban transformations, in part by providing further economic intuition and assumptions to the latent characteristics comprised in the ACI and in the TCI:

- Heterogeneous realized preferences of consumers for commercial amenities and for places are latent characteristics that separate commercial amenities and places onto one-dimensional continua represented by the TCI and the ACI
- Consumers' realized spending (as anticipated by suppliers) is both a strong direct factor and a strong indirect confounder of other characteristics that collectively determine what people consume. Thus, it is the best candidate for an economic interpretation of the TCI and of the ACI as continua.
- The tendency of spending to be intertwined with differentiated consumption, combined with the assumption that higher spenders have more leverage in where commercial amenities locate through market mechanisms, leads to two new assumption-dependent interpretations of the continua:
    1. The more different the ACI values of places, the more different the overall spending of consumers that operate selection within them tends to be. Likewise, the more different the TCI values of commercial amenity types, the more different the overall spending of consumers that select them tends to be.
    2. The higher the ACI of a place, the more its overall characteristics are revealed to be attractive to higher spenders. Conversely, the lower the ACI of a place, the less it is able to attract levels of spending associated with other, higher ACI places. Likewise, the higher the TCI of a commercial amenity type, the more it requires selection environments attractive to higher spenders to exist. Conversely, the lower the TCI of a commercial amenity type, the less it is dependent

on selection environments that are revealed to be attractive to higher spenders.

Beyond the assumptions specific to these interpretations and those overarching the entire method (especially suboptimal decisions of suppliers), a caveat to the second interpretation needs to be outlined, yet this caveat also clarifies the interpretation. The place (or cluster of places) with the highest ACI value is indeed, under the assumptions above, *revealed* to be preferred by the highest spenders as a consumption environment. But the conceptual reasoning behind the interpretation of attractiveness is only revealed at the margin of each ACI value, and it can only be led downwards in terms of spending. In other words, it can only be said that it any given place is revealed to be more attractive than the place with the ACI value that is directly below it. Two elements are at play here. First, there is no way of knowing how homogeneous preferences for places are across the spending spectrum. Second, spenders only get to choose their selection environments out of those that are not taken up by consumers with higher spending, so there is no way of knowing if they would prefer the environments that lie above their own in the ACI.

As such, an additional assumption of homogeneity in place preferences needs to be made in order to generalize the interpretation of the ACI as revealed attractiveness of places (as anticipated by suppliers). I do find it reasonable to assume that many place characteristics are traits that are commonly desirable (or undesirable) across the spending spectrum, and that preferences for these traits are less heterogeneous than realized consumption (as anticipated by suppliers). The (weak) association between the ACI and diversity hints at this. I argue that environment preferences across spending are reasonably homogeneous except for market-driven characteristics[23] (this includes commercial amenities but is also extended to every interaction that involves monetary exchanges, such as for example distance-weighted real estate prices, or the price of transit). My reasoning is that spenders that do not consume in places are essentially priced out of them because they have less leverage. Note that this does not require assuming they would spend on different commercial amenities if they could. More attractive consumption envi-

[23]In other words, the amount of spending is not a good indicator of peoples' preferences for accessibility, atmosphere, safety, institutional presence, aesthetics, the quality of social interactions, etc.

ronments are more competitive, but they are not larger. Suppliers that wish to operate in commonly attractive environments thus require consumers with higher spending abilities to cover for the costs of more expensive real estate the competition for space induces. The additional assumption upon which the ACI relies as a general interpretation of revealed attractiveness can therefore be summarized as follows: In the absence of higher spenders, lower spenders would rather consume in the environments higher spenders occupy.

How valid the practitioner deems this assumption to be in her setting will define the degrees of care and of generality with which she interprets the ACI and the TCI. This of course also has to be articulated with other assumptions. While defining spending as the leading and most representative component behind the TCI separation is one of the weaker assumptions, it should be kept in mind that other components are also at play. Spending by itself only outlines a tendency in positioning on the spectra, and it is not a universal rule. The assumption that higher spenders, through price mechanisms, have more leverage was also made. But what is probably the most important assumption, and one that was outlined in Chapter 2's discussion (section 2.4), is the dependence of retailers' locations on their own anticipations instead of on actual consumption.

A good way of assessing, refuting or validating these assumptions would be to expand the scope of the analysis, and this can be done over three dimensions: time, space, and data sources. Expanding upon this chapter's relatively short time window to better observe change would be a good place to start in further validating the ACI/TCI method. The changes observed in Paris were seldom paradigm-shifting. The stickiness of commercial amenities because of the high level of capital they imply, while a desirable trait within evolutionary analysis (Essletzbichler and Rigby, 2007) as it provides for more observable systems, leads to change being a very slow process. One aspect of this that is not fixable with commercial amenity data alone is that, as anticipation is a difficult exercise for suppliers, the ACI of places might tend to lag behind the composition of demand it implies[24].

[24]Retailers do not only have to be reactive, they can also be proactive and provoke changes in compositions of demand. This is because they are a part of what makes people attracted to places as a whole, that is, to what motivates links between consumers and places in Chapter 2's Figure 2.1.

It is not unlikely that it takes years of non-selection for retailers to exit the market, and that it can take a long time to open up shop in reaction to a change in the composition of local demand too. As such, and in the spirit of evolutionary economic geography as a means to observe changes in systems, there is a clear need for more complete historical data than the ones used here.

A second necessary expansion of the ACI/TCI is spatial. While the interpretations presented in Chapter 2's discussion (section 2.4) were not reliant on an observed setting, those presented here were built in tandem with the application in Paris. But what if some of the characteristics that were observed and that were assumed to be universal are actually Paris-exclusive? The endeavour of Juhász et al. (2023) in Budapest is reassuring in that aspect, but presenting the principles outlined here as universal (in the context of metropolitan spaces) require replicating the procedure on different cities.

Third, the quality of the BDCom data used in this chapter's application is an exception, not the rule, and it is not realistic to expect other applications to have access to BDCom-like levels of granularity. Thus, questions can and should be asked about the robustness of the results found here with a different, less ideal dataset of commercial amenities.

In light of these three main limits, Chapter 4 will present a new application of the ACI/TCI. This will be done through a open business registry dataset that will include Paris, and will allow for testing the resilience of the ACI/TCI in outlining what this chapter presented with less ideal data. This dataset is national-level, and I will therefore take this opportunity to apply the ACI/TCI to 10 additional French cities in the hope of granting the procedure more spatial external validity. Another advantage of the business registry is that it provides more historical data, namely, a usable time window from 2008 to 2022, which could allow for the better observation of structural change. Because this section has primarily been descriptive and micro-focused, Chapter 4 then takes a macroscopic approach to assess changes in structures of compositions of demand across cities, and to evaluate how these changes could relate to the rise of short-term rental activity in the last decade.

3.7 Conclusion

All things considered, I propose that the ACI and the TCI are spectra that can amplify researchers' understanding of cities because they indirectly reflect characteristics of places and the composition of their local demand.

Under their strictest interpretations, the ACI (TCI) puts places (commercial amenity types) on a continuum that tells us how different they are in their commercial amenity (place) co-occurrences relative to each other. I argue that commercial amenity locations reflect suppliers' anticipations of a consumer selection that depends on these consumers' preferences for types of goods and services and on the characteristics of the selection environment. The ACI (TCI) thus differentiates places (commercial amenity types) according to the underlying selected preferences of people that consume within them (that consume them). Selected (or realized, or consumed) preferences for goods and services are heavily intertwined with consumers' overall spending, and the continuum of the TCI thus tends to reflect a continuum of spending. Because pricing mechanisms grant higher spenders more leverage in selection, and because the ACI is a function of TCIs within places, the continuum of the ACI tends to reflect the attractiveness of places as revealed by suppliers' anticipations of spending-tied consumer presence.

Regardless of how strictly the TCI and the ACI spectra are interpreted, they provide novel, systemically motivated perceptions of urban situations and dynamics that answer policy-relevant issues regarding relative compositions of demand with an original outlook. The following chapter, after providing additional external validity to the spectra, will illustrate their use in answering policy-relevant questions.

3.8 Appendices

3.8.1 Alternative binary network definitions

The TCI as defined here uses a simple presence indicator to determine whether an amenity is within a place or not. Standard economic complexity applications use the Balassa (1965) index of Revealed Comparative Advantage (RCA). Its definition is clarified in Chapter 2's Section 2.3.1.

This application is thus also an opportunity to evaluate the differences between a presence-based matrix and a RCA-based one in terms of their impact on the ACI and on the TCI. First, Figure 3.8 outlines that results obtained through the RCA and through a presence-based definition correlate well. I find similar spatial trends in RCA-driven ACI measures as those found in Figure 3.5, but also notice the index as a whole is considerably more noisy than with a presence-based indicator. Moreover, Figure 3.8 also outlines the sensitivity and non-monotonic nature of an RCA-based ACI with respect to the definition of the threshold R* (equation 2.6), which, while usually assigned a value of 1, could be arbitrarily set to any other value with no fundamental loss in economic reasoning.

A second alternative in the methodology could have been to apply the ACI in an ECI-like way, as an average of TCIs instead of as their aggregation (I further call this ACI'). As seen in the right panel of Figure 3.8, these two measures are in fact very close. The Spearman correlation between the measures is 0.99, with many points overlapping on the plot. I also see that, naturally, The less diverse places are, the more the ACI' deviates from our ACI. The conceptual grounding behind this is that the more amenities a place has, the more information we have over which direction on the spectrum that places' consumers tend towards. On top of this, an aggregation method is more resilient towards lower quality data (either because of a poor dataset, or a result of poor routing between buildings and amenities because of imperfect street networks).

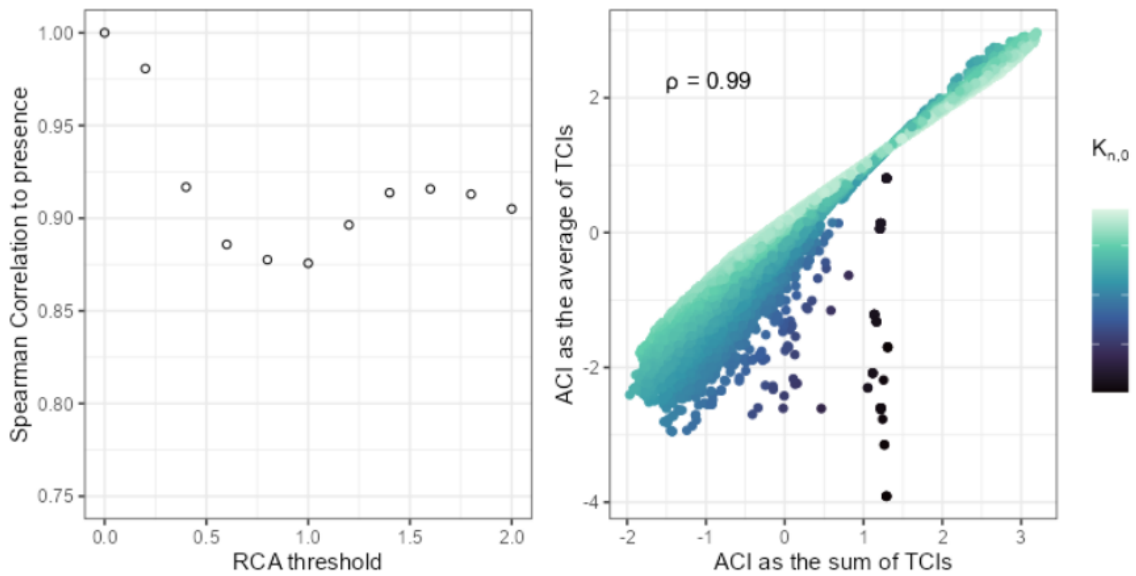


Figure 3.8: (left) Correlations between the presence-based ACI and the RCA-based ACI, at given thresholds R* (0.2 increments); (right) Relationship between the ACI as a sum of TCIs and the ACI as an average of TCIs colored by place diversity.

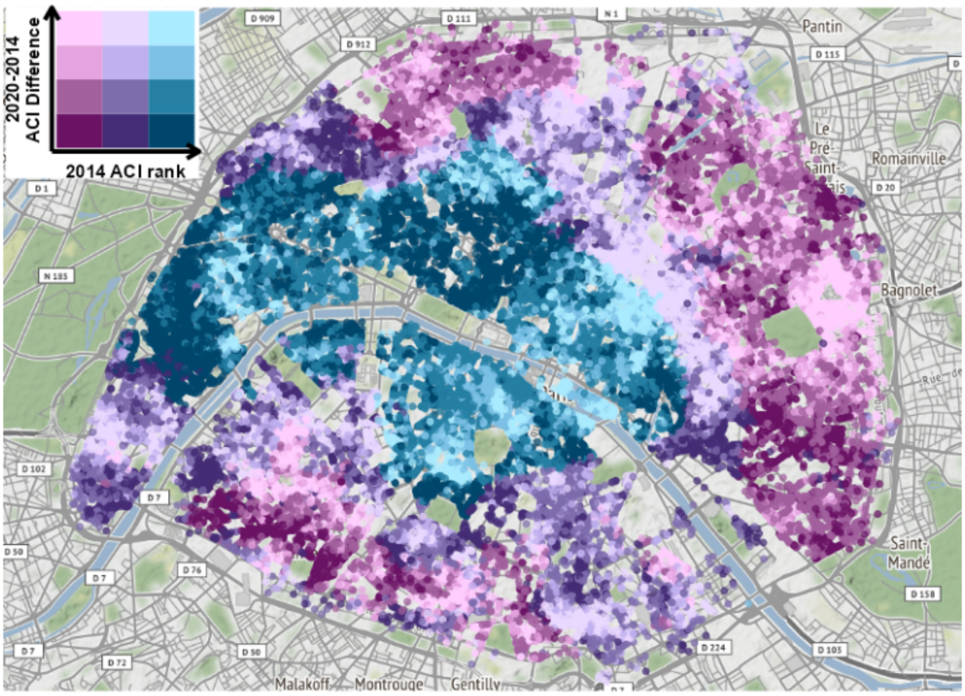


Figure 3.9: Level and evolution of ACI ranks (2014) and levels (2020-2014) using a RCA with R* = 1 to construct the binary network. To be compared to Figure 3.5, using the same color scale.
Map tiles: ©OpenStreetMap Contributors

3.8.2 Sensitivity to cutoff selection

Figure 3.11 shows the equivalent of Figure 3.5 for different selections of c, the cutoff in minutes used to determine the size of places. I find that our indicator is not overly sensitive to cutoff selection in order to determine overall spatial patterns beyond $c = 7$. The cutoff selection issue is akin to finding a balance between a measure that segregates places enough through their patterns in order to flatten out noise (which is predominant in low cutoffs), but that also allows for heterogeneity within close places (which can get lost in larger cutoffs). As a reminder, a cutoff of $c = 15$ was used throughout this chapter to define places. Figure 3.10 shows changes in TCI ranks for given cutoffs c. Barring small differences, changes in the selected cutoff beyond $c = 7$ also do not have large bearings over the overall separation of amenities.

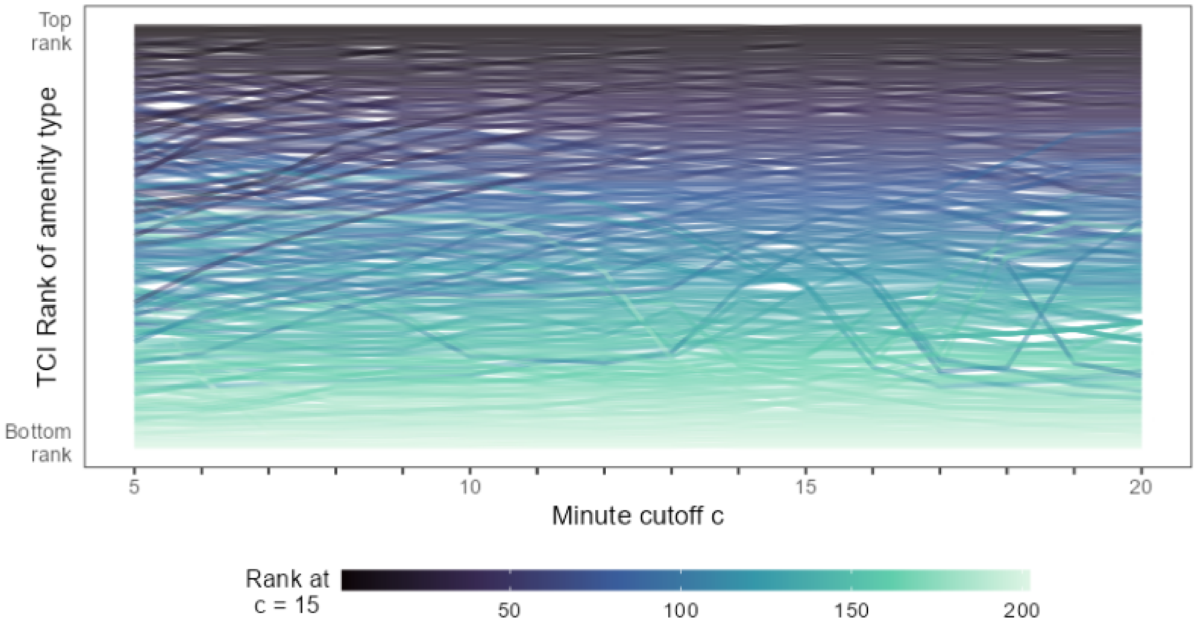


Figure 3.10: The evolution of TCI ranks throughout different cutoffs c. Every line is an amenity type, and they are vertically ordered by their ranks at $c = 15$ for comparison.

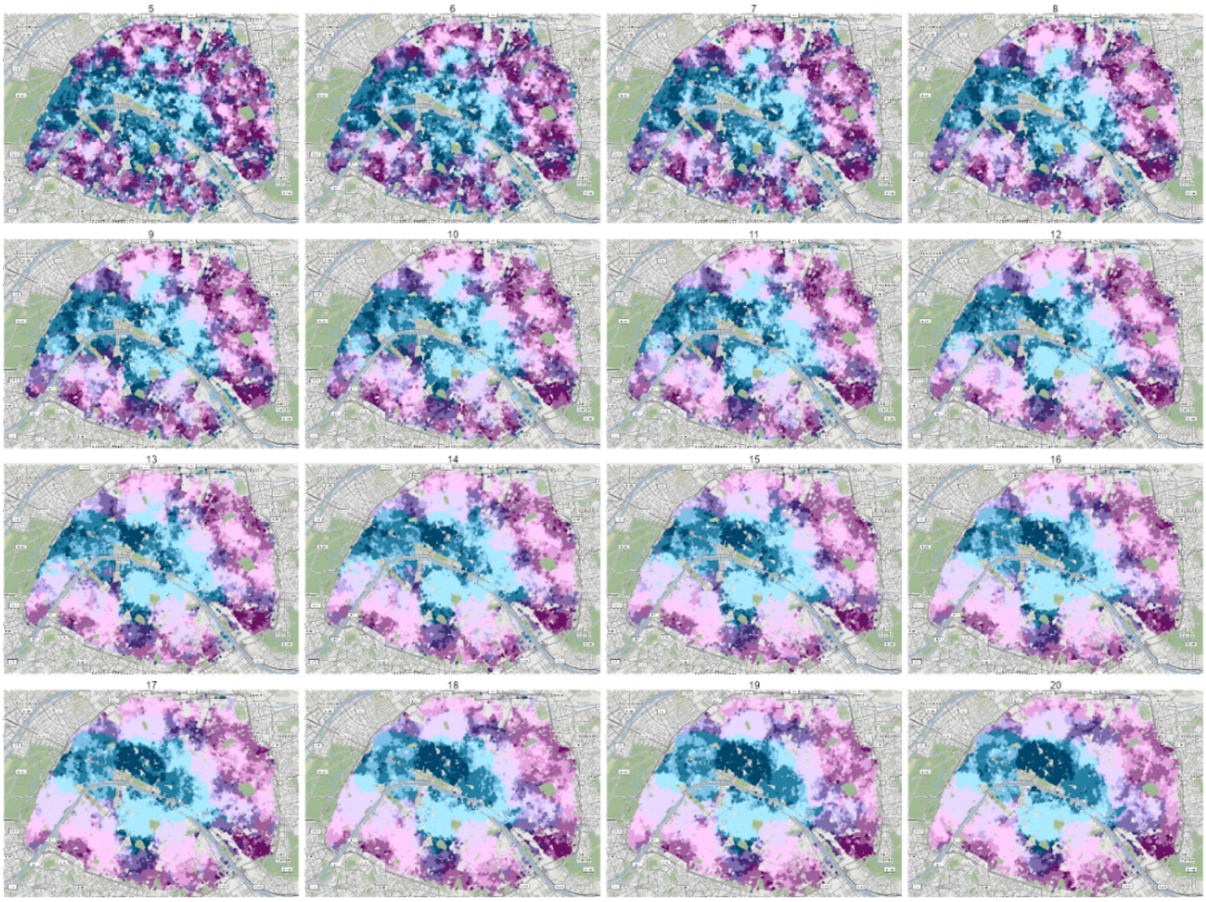


Figure 3.11: ACI level and change between 2014 and 2020 in Paris. These maps are replications of Figure 3.5 for different values of cutoff c, from $c = 5$ (top left) to $c = 20$ (bottom right) in increments of 1. Map tiles: ©OpenStreetMap Contributors

3.8.3 Random Forest

Sources and different granularities of the data used in the random forest regression are presented in Table 3.3. Model performance and characteristics are presented in Table 3.4. They relate to the version of the regression that relies on internalized splits of training and test data, but results with an additional manual split are consistent with them, as outlined by Figure 3.7 in the chapter's main text.

Variable	Median Income	Gini Coef. of Income	Share of pre-1919 housing	Share of 25-39 year olds	STR Res. Days within 500m
Data Source	INSEE	INSEE	INSEE	INSEE	AirDNA
Dataset name	"Revenu disponible"	"Revenu disponible"	"Logement"	"Population"	-
Granularity	Census tract	Census tract	Census tract	Census tract	Building

Table 3.3: Data sources of variables selected in the random forest regressions.

Observations	Number of trees	Max depth	Features at each split	Minimum samples per split	**R-squared**
163,693	500	∅	3	1	0.982

Table 3.4: Random Forest Model Characteristics

TCI Values

The BDCom data used in this chapter originally contains over 220 amenity types. I remove wholesale amenities; B2B businesses; non-selling offices and workshops; empty, under construction, stocking and equipment premises; government offices; and unemployment offices. I also aggregate two variables that were otherwise over granularized. “Car related” aggregates all combinations of car dealerships, garages, and fuel pumps when they are found in combination with each other. Car dealers and service stations that do not offer any other services are left independent. Likewise, “Motorbike related” aggregates simple motorbike repairs and motorbike repairs that are also dealers.

This yields the following list of 202 commercial amenities, ranked by their TCI values over all 3 observed years. U is the presence rate of each amenity, that is, its ubiquity (non-rarity) divided by the total number of places. Paris is an amenity-dense city, and the overall presence rate is around 80% even with this very granular dataset. The more an amenity type is present across places, the less information it holds, and therefore the less extreme its TCI will be.

A good way to illustrate the separation yielded by the TCI is to check for ubiquitously needed characteristics of goods and services that people consume through different types of retailers. In Table 3.5, I highlight stores related to non-service food shopping. The link between non-ubiquity and information gained over the amenity is clear, and it is also qualitatively clear that there is a spending power pattern associated with their separation.

Table 3.5: List of amenity types used to compute the ACI and the TCI along with the share of places they cover, and ranked by their TCI

Rank	Name	U	TCI	Rank	Name	U	TCI	Rank	Name	U	TCI
1	**Luxury general food > 300 m²**	9.57	6.103	69	Photocopies	96.36	-0.151	136	Shoe repair - "Minute" repair (keys, heels...)	99.98	-0.281
2	Tourist hotel - Palace	15.36	5.494	70	Sale of living room and bedroom furniture	97.98	-0.157	137	Laundry - Pressing	99.99	-0.281
3	Tourist hotel with 5 stars	49.22	2.999	71	Patisserie	98.39	-0.17	142	Workshop in the store	99.99	-0.281
4	Department store	30.25	2.995	72	**Specialized grocery store**	80.91	-0.176	142	Asian caterer	99.99	-0.281
5	Coffee shop	44.22	2.801	73	Leather goods - Travel items	98.31	-0.177	142	Butchery - Butchery	99.99	-0.281
6	Large cultural multispecialist	38.49	2.488	74	Furniture sales and multi-specialists	98.3	-0.179	142	Bookstore	99.99	-0.281
7	Ticketing - Booking shows	42.5	2.454	75	Roaster - Tea and coffee retailer	98.17	-0.183	142	Brasserie - Continuous restaurant with tobacco	99.99	-0.281
8	Smile Bars	15.88	2.334	76	Women's shoes	98.61	-0.188	142	Clothing alterations	99.99	-0.281
9	Sale of coins and medals	42.58	2.06	77	Parapharmacy	96.85	-0.192	142	Laundromat	99.99	-0.281
10	Haute couture - Designers	62.23	1.856	78	Thrift store - Clothes sale - Depot-sale	98.8	-0.198	142	Beauty institute - Thermal and thalasso activities	99.99	-0.281
11	Gambling	53.85	1.81	79	Specialist in a sport	95.98	-0.199	142	General building work (electricity, plumbing, painting, insulation...)	99.99	-0.281
12	Sale of erotic items and sex shop	51	1.524	80	Sale of tableware - Household utensils - Art of the table	98.5	-0.2	142	Bank - Savings Bank	99.99	-0.281
13	Concert hall	64.88	1.269	81	Hardware and drugstore	98.98	-0.203	148	European Restaurant	99.99	-0.281
14	Philately	47.53	1.251	82	**Organic food store**	98.48	-0.204	149	Brasserie - Continuous catering without tobacco	100	-0.281
15	Sales room	42.1	1.226	83	Framing - Sale of paintings - Posters	98.36	-0.206	152	**General food <120m²**	99.99	-0.281
16	Generalist Sport	59.14	1.201	84	Smoke-free bar or café	99.15	-0.215	152	Pharmacy	99.99	-0.281
17	Cabaret - Dinner and show	73.73	1.148	85	Computer self-service - Cybercafé	88.02	-0.219	152	Optician	99.99	-0.281
18	Pop-up store†	20.18	1.082	86	Furniture craftsmanship (upholsterer, cabinetmaker...)	99.18	-0.225	152	Real estate agency	99.99	-0.281
19	Ice cream shop	78.54	1.024	87	Coffee - Tobacco	99.13	-0.226	154	Seated fast food	99.99	-0.281
20	Aparthotel	74.05	0.873	88	Telephony (Major operators + resellers)	99.24	-0.226	155	Bakery - Pastry shop	99.99	-0.281
21	ATM (not linked to a bank)	68.81	0.825	89	School / extracurricular courses (Acadomia...)	98.59	-0.228	156	Traditional French restaurant	99.99	-0.281
22	Engraving	69.38	0.785	90	General household equipment	99.26	-0.234	157	Hairstyle	99.99	-0.281
23	Men's shoes	81.87	0.782	91	Second-hand goods - *Brocante*	97.2	-0.239	158	Asian restaurant	99.99	-0.282
24	Sale of video games (+ video game room)	71.32	0.734	92	Sale of toys and games	99.4	-0.24	159	Cultural and leisure activities courses (pottery, dance...)	99.95	-0.282
25	Exchange office	80.91	0.727	93	Mixed shoes	99.26	-0.241	160	Insurance	99.97	-0.282
26	Sale of religious articles	79.57	0.629	94	African restaurant	79.26	-0.242	161	Sale of frozen products	99.94	-0.283
27	Binding and finishing	77.42	0.573	95	Sale of hearing aids	99.01	-0.245	162	Creamery - Cheese factory	99.7	-0.284
28	Other venue	64.21	0.558	96	Fishmonger's	98.84	-0.249	163	Sale of fruits and vegetables	99.93	-0.285
29	Sale of old books -	80.5	0.478	97	Fabrics - Textile -	99.36	-0.249	164	Veterinarian	99.84	-0.286

30	Garden center - Nursery	58.44	0.449	98	Service station	96.01	-0.253	165	Sale of medical articles - Prostheses and orthopedic insoles	95.53	-0.287
31	Discotheque and private club	82.6	0.433	99	Car rental	98.89	-0.254	166	Household appliance specialist	92.47	-0.288
32	Costume or accessory rental - Leisure	74.68	0.427	100	Maghreb restaurant	99.02	-0.256	167	Driving school	99.89	-0.288
33	Generalist household appliances - Radio - TV - Hi-Fi	81.06	0.39	101	Fashion jewelry - Fashion accessories	99.68	-0.257	168	Pet grooming and equipment	95.93	-0.289
34	Childcare	82.47	0.368	102	Stationery - Office Supplies	99.56	-0.257	169	Sale of records, cassettes, CDs, DVDs	81.23	-0.291
35	Sale of luminaries	89.04	0.351	103	Men's Ready-to-Wear	99.69	-0.26	170	Multi-sports hall	99.53	-0.292
36	Watches	84.89	0.35	104	Sale, repair, rental of bicycles / electric bikes	95.85	-0.26	171	Sale of newspapers	99.63	-0.298
37	West Indian restaurant	64.27	0.334	105	Repair of electrical or electronic items	97.2	-0.262	172	Personal services (cleaning, help for the elderly...)	99.11	-0.302
38	Other world restaurant	88.81	0.321	106	Tourist hotel with 2 stars	99.72	-0.264	173	Dental office	99.59	-0.305
39	Cinema	86.75	0.271	107	Tobacco	99.72	-0.267	174	Funeral homes	97.75	-0.306
40	Graphic arts materials - Creative leisure	77.91	0.247	108	**Regional and foreign specialty food products**	99.81	-0.267	175	Tanning salon - Solar / UV	90.42	-0.309
41	Tourist hotel with 4 stars	91.49	0.237	109	Physiotherapist's office	99.78	-0.268	176	Radiology Center	96.89	-0.318
42	Sale of cameras	72.94	0.213	110	Tourist hotel with 3 stars	99.83	-0.268	177	Home delivery of food dishes	99.33	-0.319
43	Sale and manufacture of bridal wear	88.29	0.187	111	Perfumery - Beauty products	99.84	-0.27	178	Sale of computer equipment	98.02	-0.321
44	Games room or club	71.31	0.156	112	Medical analysis laboratory	99.86	-0.272	179	Printing	94.85	-0.322
45	Custom tailor	92.58	0.131	113	Chocolate - Confectionery	99.87	-0.274	180	Bazaar	99.18	-0.33
46	Buy - Sell gold	91.83	0.091	114	Sale of electronic cigarettes	99.8	-0.275	181	Carpentry - Glazing - Mirrors	96.52	-0.34
47	Household linen	94.01	0.074	115	Massage parlour	99.9	-0.277	182	**Curbside grocery pickup†**	22.38	-0.358
48	Shopping and express mail	68.22	0.035	116	Delicatessen - Catering - Delicatessen	99.93	-0.278	183	Car dealer	94.21	-0.373
49	Radio - TV - Hi-Fi Specialist	92.46	0.025	117	Ready-to-wear Mixed	99.95	-0.278	184	Motorbike related	97.27	-0.395
50	Video Club (Cassette and DVD rental)	57.41	0.006	118	Art Gallery	99.91	-0.278	185	Temporary employment agency	87.03	-0.441
51	Floor and wall coverings	95.97	-0.036	119	Watches - Jewelry	99.91	-0.278	186	Car related	97.72	-0.452
52	Sale of kitchen and bathroom furniture	95.86	-0.044	120	Nail care	99.93	-0.278	187	Motorcycle dealer	85.31	-0.473
53	Sports - Clothing and footwear	96.45	-0.055	121	Locksmithing	99.95	-0.279	188	DIY	90.53	-0.544
54	Rapid development - Photo film sale	95.21	-0.064	122	Travel and tourism agency - Airlines	99.96	-0.279	189	Nurse's office	95.07	-0.545
55	Bimbeloterie - Souvenirs	94.71	-0.065	123	**Monoprix**	98.83	-0.28	190	Sale of pets	52.92	-0.591
56	Central and South American restaurant	89.62	-0.066	124	Other physician assistant activity - Speech therapist	99.97	-0.28	191	Moving / Storage	87.49	-0.666
57	Children's shoes	91.46	-0.066	125	Professional training courses (languages, computers...)	99.89	-0.28	192	Tattoo - Piercing	80.38	-0.727

Table 3.5: (Continued)

58	Theater	93.74	-0.076	126	Florist	99.98	-0.281	193	Telecommunication in store	83.7	-0.779
59	Jewelry	95.84	-0.082	127	Specialized gym	99.98	-0.281	194	DIY and home equipment rental	83.41	-0.792
60	Tourist hotel with 1 star	84.29	-0.101	128	Retail trade of beverages	99.98	-0.281	195	Youth Hostel	53.09	-0.822
61	Ready-to-wear Lingerie	97.42	-0.102	129	Indian, Pakistani and Middle Eastern restaurant	99.99	-0.281	196	Discount store	68.44	-0.904
62	Antiques	96.89	-0.112	130	Women's ready-to-wear	99.99	-0.281	197	Sale of automotive equipment	79.53	-0.962
63	Discount telephony and accessories (no particular brand)	97.62	-0.119	131	Newspaper kiosk	99.98	-0.281	198	Ambulances	74.94	-1.612
64	Tourist hotel without star	93.84	-0.12	132	Medical practice	99.98	-0.281	199	**Discount supermarket**	66.59	-1.94
65	Children's clothing	97.83	-0.128	133	**Classic convenience store**	99.98	-0.281	200	Technical control of the car	53.96	-2.301
66	Manufacture and sale of musical instruments	90.97	-0.131	134	**Classic supermarket**	99.98	-0.281	201	**Specialized supermarket**	25.64	-3.138
67	Studio of photographic reports	95.55	-0.136	135	Fast food standing up	99.99	-0.281	202	**Hypermarket**	7.66	-4.919
68	Tea room	97.99	-0.144								

† Grocery curbside pickups and pop-up stores were new additions to the 2020 database. They are absent from all places in 2014 and 2017.

Table 3.5: (Continued)

Chapter 4

Extending urban complexity analysis to study short-term rental dynamics

4.1 Introduction

In the previous two chapters, I explored how indicators from the economic complexity literature can be applied to urban spaces and what they can reveal. A major reason why economic complexity literature is valuable in modern economic geography discourse is its versatility and adaptability both in methodology and in applications (Balland et al., 2022; C. A. Hidalgo, 2021). However, the new setting within which I have adapted these methods is fundamentally different to their traditional applications, and it requires further validation. This validation cannot depend solely on the standardization of economic complexity indices that results from their widespread use in economic geography. The amenity-based method and its consumer-driven interpretations need to be expanded to cement them as revealing of underlying economic and non-economic systemic non-linear interactions within cities.

So far, and although a conceptually different approach to amenity complexity was undertaken in Budapest (Juhász et al., 2023), the only case study of the specific methodology developed in Chapter 2 has been the city of Paris during a single 6-year period, undertaken in Chapter 3. This was in part thanks to exceptional data, but this work's idea of an Amenity Complexity Index (ACI) and amenity-Type Complexity Index (TCI) would benefit from a broader expansion towards non-exceptional data for various reasons and dimensions. To be practical and reliable, it also needs to be compatible with other data sources. First, a different dataset could support or challenge the findings from the previous chapter. Second, a dataset of amenities in a different city and time period would enhance its external validity, especially considering the uniqueness of Paris as a city.

Furthermore, a more comprehensive empirical examination of the index, beyond the simple correlations shown in the previous chapter, would benefit its future use. The first chapter of this thesis investigated the impact of Bordeaux's local policies on short-term-rental (STR) activity, and the literature that informed it strongly emphasizes the impacts of short-term rentals on residents. Their effects are complicated and multifaceted, and most empirical studies on them focus on direct impacts on rental and property markets (Barron et al., 2021; Wachsmuth and Weisler, 2018) or on pre-assumed roles of commercial amenities (A. Hidalgo et al., 2023; Zervas et al., 2017).

The consumer-driven framework introduced in Chapter 2 aligns with the complicated and multifaceted impacts of platforms like *Airbnb*. They entail significant shifts in the composition of place demand through systemic and evolutionary processes that can have powerful spillover effects on other actors in the network. As such, the ACI and the TCI are suitable candidates to characterize how STRs contribute to shaping cities.

The spectrum of the ACI and of the TCI tend to reflect the relative composition of local demand in an imperfect but systemically motivated way, both at the city level and the place level. Illustrating how they can be used to observe changes in consumer composition along with specific events (such as rises in STR activity) is a promising way of employing these spectra to better understand cities and the underlying processes that shape them, especially in the context of STR. In its context, changes as outlined by the ACI and the TCI are crucially motivated by and for residential urban life, as their impacts are dependent on a plethora of characteristics that, alone, would not allow for characterization. The holistic approach of the ACI/TCI solves this.

This chapter will consist of two main sections. In the first section, I advocate for the use of French business register data to expand the amenity complexity's scope in both space and time. I explain the methodological adjustments that are needed with the new data. I then employ this data to evaluate the validity of the previous chapter's interpretations in an internally less optimal but externally more promising setting. In the second section, I concentrate on amenity-type complexity (TCI) ranks and how they differ across cities. I use these comparisons, and how they vary over time, to credibly pose the question of a possible *Airbnb*-induced homogenization of major French cities.

4.2 A different dataset

Expanding the scope of the ACI/TCI spectra requires several decisions that need to be justified and assessed. First, I will introduce the data used in this chapter and argue for its choice over alternatives. Decisions regarding the data structures and how the methodology needs to be adjusted to new data will follow. Then, a validity assessment of this more general method will be conducted in two steps, where I will compare results in Paris to the previous chapter's before examining the indicator's consistency across the other observed cities.

4.2.1 Data selection

I am seeking a dataset of commercial amenities along with their precise coordinates that achieves the best balance in expanding the spatial scope of the method and its time window under the constraint of it being comparable to the Parisian setting. This has to be articulated with data quality assessments. Ideally, the categorization of commercial amenity types would be as fine-grained as goods and services commercial amenities provide are substitutable for consumers. *APUR*'s *BDCom* dataset from Chapter 3 was of exceptional quality in that regard. Replicability is also of course a consideration. I argue that the *SIRENE* database, an *INSEE*-led business directory, achieves the best balance out of all of these criteria despite a few other options were considered, especially *Google Places* and *OpenStreetMap*.

Google Places data are a common choice among researchers (C. A. Hidalgo et al., 2020; Juhász et al., 2023; Kaufmann et al., 2022). These data depend on suppliers' declarations, but businesses have very strong incentives to be visible on *Google* especially in major cities. They are regarded as high quality and are extensively available in space. However, they rely on one-time collections. In other words, getting dynamic data over time requires collecting the data multiple times, as amenities' opening dates or closed amenities are not comprehensively returned. Additionally, these datasets may be more challenging to replicate at a larger scale because they are not free. In terms of granularity (how well businesses are classified by type), these data are far from what can be found in Chapter 3 (see Appendix 3.8.3 for a full list of *BDCom* categories), but they offer a solid picture of the amenity landscape within which categories are created, all of this with a suitable consumer-oriented definition and worldwide availability.

A popular alternative, although perhaps not as academically prominent, is *OpenStreetMap* (*OSM*). *OSM* is free to use and is functionally similar to *Google Places* in its definition of amenity categories, except the data are crowdsourced and face less editing. Their main and crucial advantage over *Google Places* data, besides the fact they are free to use and replicable, is their availability in time. Historical datasets span back over a decade in major cities. On the downside, and despite interesting efforts at using historical point of interest datasets (Hochmair et al., 2018; Zhang and Pfoser, 2019), its quality is disputed especially in less dense places.

Despite the spatial advantages of *OpenStreetMap* and *Google Places*, I prefer to use *SIRENE* data. *SIRENE* is an open and public-facing French business register that is maintained by *INSEE*, the national institute for statistics. Its main (and, in contrast to the alternatives, only) downside is that it is limited to France. Business register data is of course also accessible in other countries under different forms, and the potential for replication of such a process elsewhere remains strong. For example, official public-facing and free data has been used to map out amenity changes in Madrid in the context of short-term rental impacts (A. Hidalgo et al., 2023).

Moreover, registering business establishments is a compulsory part of the fiscal process, and this data is the closest to a ground truth as opposed to alternatives. In other words, I consider it to be high quality. *SIRENE* has recordings up to almost eight decades old, but frequent changes in the nomenclature of business classification make it challenging to make cross-year comparisons when these changes occur. Still, I end up with data on all French business establishments and their registration codes for years ranging from 2008 to 2022, which is a better time span than any alternative could provide. Within these data, after extensive reordering and cleaning, I select French cities with over 200,000 inhabitants. One main drawback of using non-*APUR* data is that, similar to *OSM* and *Google Places* alternatives, *SIRENE* is far from being as fine-grained as *APUR*. This has methodological implications that are discussed below.

4.2.2 Methodological implications of SIRENE data

As a reminder of Chapter 2, computing the ACI and the TCI is done through a bipartite network. To this aim, I replicate the process of 15-minute areas around building coordinates using *BNDB* data is used to build a preliminary matrix with count data of commercial amenities within places. But the less granular definition of amenity types makes the binarization process more complicated. Using a simple presence indicator as was argued for in Chapter 2's Section 2.3.2 scenario is no longer possible. The less precise division of amenity types means amenities are grouped up into fewer categories, and count per category is higher. Thus, a presence-based bipartite network would be saturated.

As a fallback, the Balassa index of Revealed Comparative Advantage (Balassa, 1965) is used, as is usually the standard in Economic Complexity literature and as is used by Juhász and his coauthors in their own Budapestian approach (Juhász et al., 2023). As a reminder, and with a base count matrix $O_{n,a}$ of n places and a amenity types, the binary matrix $M_{n,a}$ is thus defined as follows:

$$RCA_{n,a} = \frac{O_{n,a}/\sum_a O_{n,a}}{\sum_n O_{n,a}/\sum_n \sum_a O_{n,a}} \tag{4.1}$$

$$M_{n,a} = \begin{cases} 1 & \text{if } RCA_{n,a} \geq R* \\ 0 & \text{otherwise} \end{cases}$$

Because I use $R* = 1$ as is standard[1] in the literature. $M_{n,a}$ contains a 1 where a place n is specialized in amenity type a, that is, where it hosts its "fair share" of type a relative to the network as a whole. This methodological change should be seen as a necessary concession that comes as a result of the less granular data. This concession comes at a cost of a more convoluted interpretation.

Let me refer to Section 2.3.3 and to the definition of the TCI as it is used in this work (Hausmann et al., 2014; Mealy et al., 2019). The TCI is built using $W_{a,a'}$, a similarity matrix of $M_{n,a}$, and it is the eigenvector associated with the second-largest eigenvalue of $W_{a,a'}$. This eigenvector is representative of a component, or of a dimension, of $W_{a,a'}$. Because it is associated with the most segregating component of the similarity matrix, it has been interpreted as revealing a separation of commercial amenities based on their co-presences.

Because $M_{n,a}$ is no longer a function of amenity co-presences but rather of the co-specialization of places for amenity types, the second-largest eigenvector of $W_{a,a'}$ no longer represents a co-presence based similarity component between amenity types, but rather a co-specialization based one. What might seem like a small nuance fundamentally obfuscates the index, and could hurt the ability to interpret it as clearly as in the previous chapter.

One more important methodological feature of note is how time is dealt with in the base matrix. The intent of this chapter is to develop a way of observing the evolution of the TCI and of the ACI in time, but this is notoriously difficult to do using absolute values[2].

[1] Although results can be sensitive to $R*$, as outlined in Chapter 3's Appendix 3.8.1

[2] The orientation of the TCI is kept consistent through different matrices in different years thanks to the ubiquity-driven

With *APUR*'s *BDCOM* data, I pooled all three observed years together into a single matrix. Here and while keeping cities' networks separated from each other, I cannot pool all 25 observed years together. This would not fit a consumer-facing framework, because the assumption that consumer preferences are constant over the entire time span is unreasonable. The implications of this is that cross-time value-on-value comparisons are not possible in this setting, and that ranks have to be used instead.

Autocorrelation is introduced in the form of a 3-year time averaging of the specialization matrix $M_{n,a}$. This is a common method (C. A. Hidalgo, 2021) that reduces the impact of noise on the specialization matrix and mitigates some of the issues related to threshold selection. In practical terms, the original count data for any given year is modified to include the years before and after. Where $O^t_{n,a}$ is the original count matrix of n places and a amenity types in year t, this means that:

$$O^t_{n,a} = (O^t_{n,a} + O^{t-1}_{n,a} + O^{t+1}_{n,a})/3$$

SIRENE data, through its differences with the first TCI/ACI application in this work, implies different methodology. This different methodology risks obfuscating what the indicator can reveal. Testing for robustness of the new data was already needed, and is now even more important given the changes outlined in this section.

4.3 Robustness

The robustness of the TCI/ACI method is tested for in two complementary ways. First, I confront the results obtained using the *SIRENE* dataset within Paris to those obtained using the *APUR* dataset in the previous chapter. I show that results are qualitatively similar, and that the underlying tendencies revealed by the TCI/ACI method are resilient to lower-quality data, and to methodological differences. Second, I leverage the spatial expansion granted by *SIRENE* data and find that tendencies outlined in Paris are consis-

inversion procedure described in Chapter 2's Appendix 2.6.2, but its scale is not.

tent with those outlined elsewhere, which enhances the external validity of the TCI/ACI interpretations.

4.3.1 A comparison in Paris

In this subsection, I examine how the nationwide *SIRENE* data used to generalize the TCI/ACI spectra compares to the *APUR* data presented in the previous chapter in an attempt to check for the validity of a generalization. Beside the data itself, the different methodology it implies could hinder interpretations.

Similar results across datasets would allow for more flexibility in interpreting the modified version of the spectra (that is, this chapter's version) more similarly to the full-scale survey and presence-based version. Computing the Amenity Complexity Index (ACI) with *SIRENE* data (noted **sACI** hereafter, with accompanying **sTCI**) has benefits and drawbacks when compared to the original method (noted **oACI** hereafter, with accompanying **oTCI**). It enables for exploring a wider space and a wider time frames. By comparing it to the *APUR* dataset and the presence-based method to test its validity, I focus only on the sACI/sTCI's potential drawbacks by limiting its area to Paris and its timeframe to the same years available in the *APUR* dataset, that is 2014, 2017 and 2020 exclusively[3]. I adopt the same perspective as the previous chapter by first exploring the sTCI's results, and second assessing how it is distributed in space.

First, I explore the underlying structure of the sTCI and how it relates to ubiquity to outline important differences in network structures implied by the RCA. In the previous chapter, the use of presence allowed ubiquity to be defined as non-rarity of presence (or rarity of absence), which led to an interpretation of rare amenities having more weight in defining the oTCI (and the resulting oACI). In other words, commercial amenities that were more often absent were stronger contributors to the latent information as revealed

[3]Outside of co-occurrence definitions, the full methodology of Chapter 3 is used in this section. This means that, as was the case in that application, $M_{n,a}$ is a single matrix that contains observations in all 3 years, and within which each of the places is represented by 3 rows. This methodological quirk is not kept beyond this subsection for reasons outlined in Subsection 4.2.2

by the oTCI because they were the strongest contributors to differences in observed occurrences.

With an RCA-based method, amenities that are the most ubiquitous are not necessarily those that are the most present, but they are those that the most places have a comparative advantage in. This implies an important shift in the concept of ubiquity. In the oTCI setting, ubiquity was representative of how easy it was to sustain at least one business of a type given consumer preferences across spaces, and thus, of the relative non-influence of selection environments on types of amenities. In the sTCI setting, ubiquity should be interpreted differently. The sTCI distribution in Figure 4.1 mirrors the oTCI's of Figure 3.2 (a replication of which is found below), and it outlines two major distinctions between the two methods.

The RCA method smooths the distribution of sTCIs. Back in chapter 3's Figure 3.2, only the top and bottom oTCI values were markedly different from those ranked just above or below them. In other words, a large majority of retail types were average and had little impact on the oACI of places that hosted them. Figure 4.1, on the other hand, shows a less spiky distribution of amenity values relative to their rank; This is evidenced by the more linear upward curve of points. Amenity types tend to be more distinct from each other within the sTCI than they were within the oTCI, and each competitive advantage gained or lost by a place will on average influence its sACI score more than in the oACI counterpart. This is a double-edged sword, because while it implies that the sTCI might be less susceptible to information loss in its structure, it is also dependent on a seemingly arbitrary RCA threshold. In other words, because sTCI values are so separated, minor changes in RCA thresholds would likely lead to significant shifts — Both in terms of ranking and of absolute values.

Second, the relationship between the sTCI and RCA-defined ubiquity is vastly different from that between the oTCI and presence-defined ubiquity, and this can be explained with economic intuitions. In chapter 3's Figure 3.2, amenity types with extreme oTCI

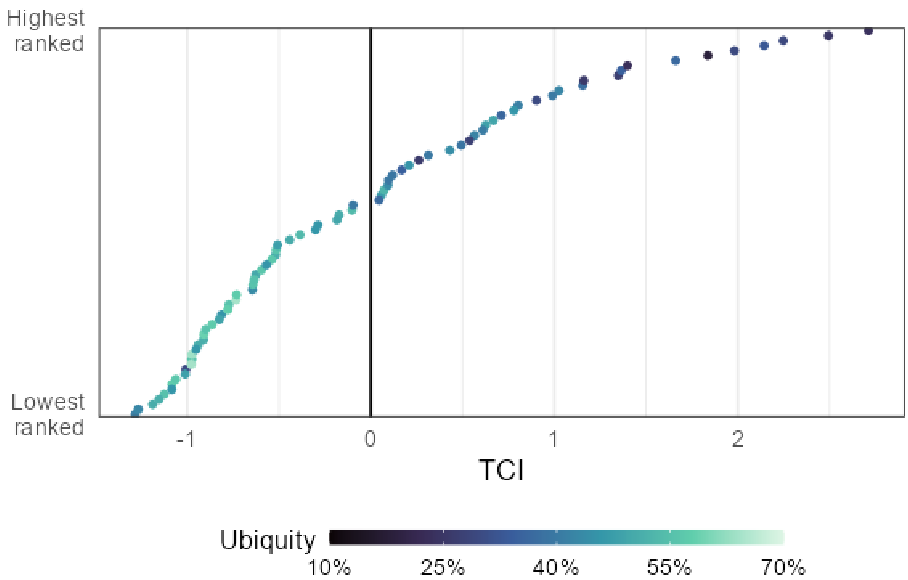


Figure 4.1: sTCI values, ranks, and ubiquities (as a share of places that have a comparative advantage) of amenity types. A parallel to Figure 3.2. The lower an amenity type's ubiquity, the higher its complexity tends to be. Complexity values are also a lot less normally distributed than they were in the previous chapter.

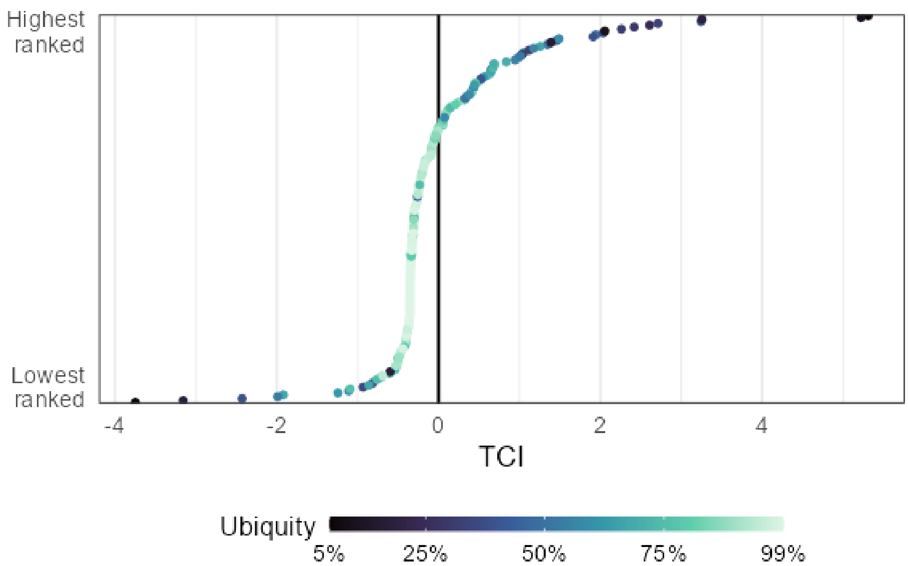


Figure 3.2: The distribution of oTCI values was quite different, and their relationship with ubiquity was non-monotonic (repeated from page 107)

values exhibited very low ubiquity, that is, they were rare. Since most amenities had very high presence rates, they were very similar to each other in their co-presence patterns and had similar oTCI values. The relatively few rare (non-ubiquitous) amenities were eventually split at either extremity of the oTCI's distribution. With an *Engel*-inspired view of amenity consumption, I derived that non-ubiquity was indicative of exclusive products and services, with exclusivity driven either by very low ("inferior" amenities, with low presence-based ubiquity and low oTCI values) or by very high consumer preference ("luxury" amenities, with low presence-based ubiquity and high oTCI values). This view was further reinforced by qualitative analysis in Chapter 3, but the RCA-based methodology of this chapter does not allow for such an idea of rarity. Instead, the relationship between ubiquity and the sTCI is observed to be linear.

In fact, and mimicking Economic Complexity analyzes, the "easier" an amenity type is to host, the higher its ubiquity tends to be in a RCA framework. By design in the RCA, ubiquity is intimately tied with the overall supply of commercial amenities within places and with their RCA-driven diversity. As such, it also tends to be representative of non-linear and unobservable differentiating place characteristics, and it is not surprising to see that, despite a few exceptions, the overall inverse relationship between ubiquity and the sTCI is clear in Figure 4.1. Let me make the following assumption to unpack this: Places that are better at attracting people (through all of their characteristics) make for selection environments that tend to be more suitable to hosting more amenities as a whole, and to be hosting more different types of amenities (in count terms but not necessarily in RCA terms). Meanwhile, the RCA is formulated in a way that nestedness is inherent as long as the overall supply (in total count) of commercial amenities is uneven, and that it is not neutral with respect to its diversity. Thus, amenities that exhibit higher ubiquity tend to be systematically observed in selection environments that attract less consumption as a whole. This fits well with the previous chapters' interpretation of the TCI as tending to revealing underlying attractiveness of commercial amenities, and is a sign of the method's robustness to new, worse data.

This robustness should however also be tested through the assessment of sTCI and of sACI results, and the comparison to their oACI and oTCI counterparts. The rest of this subsection will compare the *SIRENE*-based measures with the *APUR*-based ones and show how they produce similar outcomes despite their differences. This will support the broader use of the *SIRENE* database for this chapter's implementation and future endeavours. Table 4.1 gives a qualitative overview of the sTCI, mirroring Table 3.1. However, the categories in the *SIRENE* database are not the same as those in the *APUR* database, so a direct match is not possible. Nevertheless, the sTCI reveals patterns that are consistent with the oTCI.

The *SIRENE* database is not designed to provide information from a consumer perspective, but rather from an institutional one that helps with taxation and planning. For instance, the oTCI distinguished between different types of restaurants and hotels by their star ratings, while the *SIRENE* database only has one category for hotels and one for fast food versus traditional restaurants.

Table 4.1: Top 15 and bottom 15 commercial amenity types by their complexity (TCI) ranks in Paris. The first table uses the sTCI (*SIRENE* data). The second table uses the oTCI (*APUR* data, see Table 3.1)

Rank	Name	sTCI	Rank	Name	sTCI
1	Retail sale of leather and travel goods	2.71		...	
2	Second-hand goods in stores	2.49	64	Funeral services	-0.91
3	Watches and jewelry	2.25	65	Motorbike related	-0.94
4	Clothing in specialized stores	2.14	66	Other personal and household goods	-0.95
5	Other miscellaneous specialist retailers	1.98	67	Frozen food retailing	-0.97
6	Department stores	1.84	68	Teaching sports and leisure activities	-0.98
7	Footwear retail	1.66	69	Telecommunications equipment	-0.98
8	Watch and jewelry repairs	1.4	70	Repair of household appliances	-1.01
9	Perfume and beauty products	1.37	71	Fruit and vegetable specialized stores	-1.01
10	Textiles in specialized stores	1.35	72	Supermarkets	-1.06
11	Film screenings	1.16	73	Other stalls and markets	-1.08
12	Hotels and similar accommodation	1.16	74	Meat and meat products in specialized stores	-1.08
13	Other household equipment	1.03	75	Computer and electronics repair	-1.13
14	Restaurants	0.99	76	General grocery	-1.15
15	Books in specialized stores	0.9	77	Textiles, clothing and footwear at stalls, markets	-1.19
	...		78	Food retailing on stalls, markets	-1.27

Rank	Name	oTCI	Rank	Name	oTCI
1	Luxury general food $>$ 300 m^2	6.103		...	
2	Tourist hotel - Palace	5.494	188	DIY	-0.544
3	Tourist hotel with 5 stars	2.999	189	Nurse's office	-0.545
4	Department store	2.995	190	Sale of pets	-0.591
5	Coffee shop	2.801	191	Moving / Storage	-0.666
6	Large cultural multispecialist	2.488	192	Tattoo - Piercing	-0.727
7	Ticketing - Booking shows	2.454	193	Telecommunication in store	-0.779
8	Smile Bars	2.334	194	DIY and home equipment rental	-0.792
9	Sale of coins and medals	2.06	195	Youth Hostel	-0.822
10	Haute couture - Designers	1.856	196	Discount store	-0.904
11	Gambling	1.81	197	Sale of automotive equipment	-0.962
12	Sale of erotic items and sex shop	1.524	198	Ambulances	-1.612
13	Concert hall	1.269	199	Discount supermarket	-1.94
14	Philately	1.251	200	Car technical control center	-2.301
15	Sales room	1.226	201	Specialized supermarket	-3.138
	...		202	Hypermarket	-4.919

Still, some features stand out in Table 4.1. First, the amenity categories are less differentiated than they were in the oTCI. Extreme values are less pronounced, but there is still a clear trend of leisure (and tourism) oriented goods and services at the top, and necessity-focused activities at the bottom. Moreover, as in Table 3.1, types of amenities at the top are intuitively associated with luxury goods and services, while the ones at the bottom are linked to inferior goods and services. This suggests that the sTCI is also closely related to consumers' spending, and it helps confirm the validity of this chapter's index. The computation does not have the same data as the oTCI to separate amenity types, but it reaches similar results based on qualitative analysis.

The ACI (place complexity) rankings can complement this and provide a more objective way to assess the same problem, as they can be used to determine whether places within cities are sorted consistently. Appendix 3.8.1 showed that an RCA version of the *APUR* ACI was consistent with its presence-based counterpart. In practice, common patterns are also observed across the sACI and the oACI. The sACI distributions in Figure 4.2 reproduce those in Figure 3.6 for the oACI. For each year, the height of the graph represents the relative density of places around a given sACI value. Each graph is divided and colored by decile.

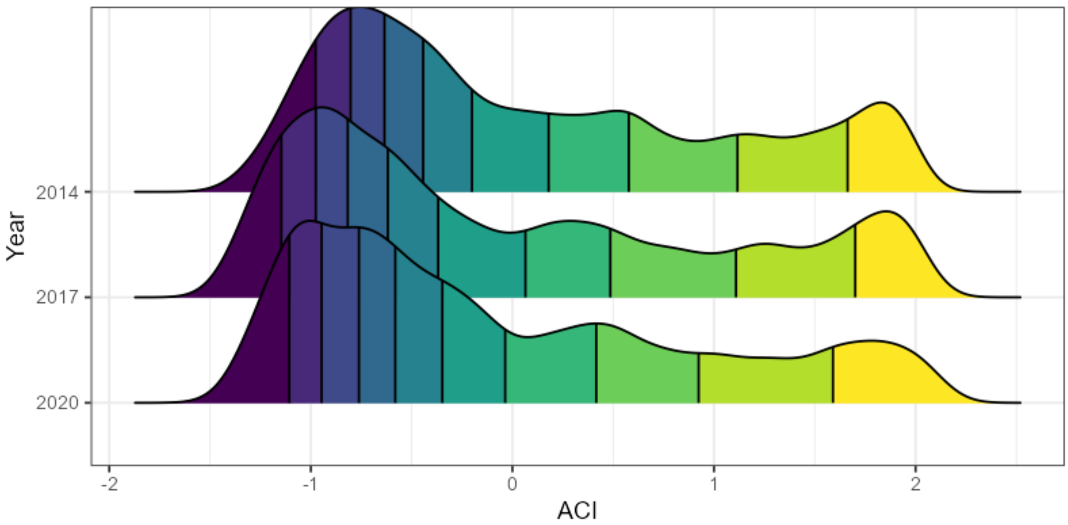


Figure 4.2: sACI distributions in 2014, in 2017 and in 2020, colored by decile.

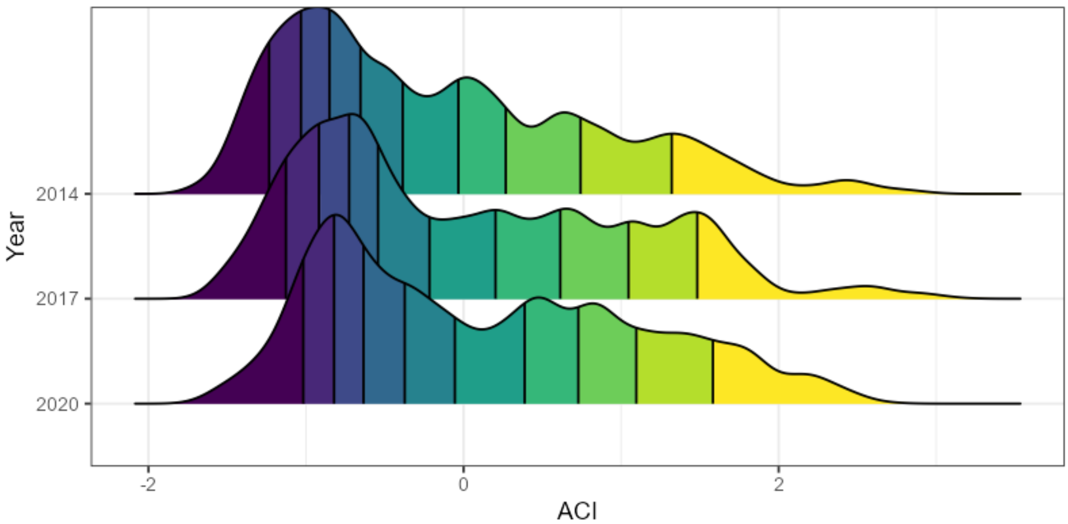


Figure 3.6: (repeated from page 124)

Structurally and regardless of their evolution over time, the distributions in Figure 4.2 are extremely similar to those from Figure 3.6. They are right-skewed with a mode around the second decile, they are otherwise uniformly distributed rightward of the median until the top decile, and both versions operate within a similar range. The oACI's distribution is perhaps slightly more spiky, and its long tail on the right-hand side is replaced by a small spike in the 10th decile in the sACI.

As for the changes over time, they are less consistent across both ACI measures in their overall structure. In the oACI, a general rightward shift could be observed, that is, every decile was moving rightward and becoming more like the highest-sACI places over time. This is true for some of the sACI's deciles, but not for all of them. The potential volatility of ACI measures over such short timeframes was discussed in chapters 2 and 3 and the robustness of the oACI's changes in distributions to different, less precise data was expected. Still, the fact that the distributions as a whole (that is, time-independently) are so similar is very promising regarding the flexibility of the indices. But are the same places in the same areas of both distributions?

Because I use the same buildings as centers in the sACI as I did in the oACI, this critical question can be answered straightforwardly by comparing complexity values at the place level. The first important element of note is that the Spearman correlation between both complexity measures is $\rho = 0.895$: Places' sACI ranks have a sturdy tendency to be near their oACI ranks. Generally, high-sACI places in the *APUR* method are also high-sACI with *SIRENE* data. This is backed up by Figure 4.3, which mirrors Figure 3.4. Patterns in change are a lot noisier in the sACI than they were in the oACI, despite some similarities. The lower data granularity and the sensitivity to the RCA's threshold certainly do not help. However, this concern is alleviated by the larger timeframe *SIRENE* data covers. Importantly, over large periods of time, the sACI and the oACI appear to converge – As illustrated by the quasi-mirrored spatial patterns in the figures' mean ACI ranks.

4.3.2 Consistency across time and space

While the *SIRENE* results agree with the *APUR* data used earlier, the data's scope has until now been restricted to Paris. Expanding the idea to a larger group of cities is essential to validate this method's results, to further explore them, and to cement the ACI/TCI method as relevant to future research. I use the methodology described in Section 4.2.2

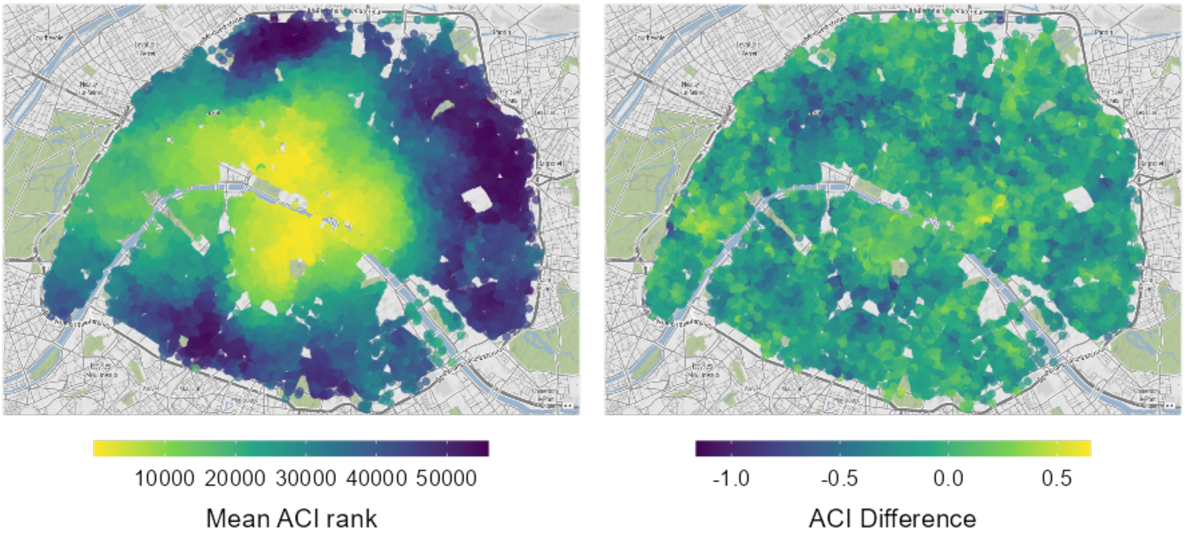


Figure 4.3: Maps of the mean sACI rank in Paris in 2014, 2017 and 2020 (left), and of the absolute change in sACI over the observed period (right).
Map tiles: ©OpenStreetMap Contributors

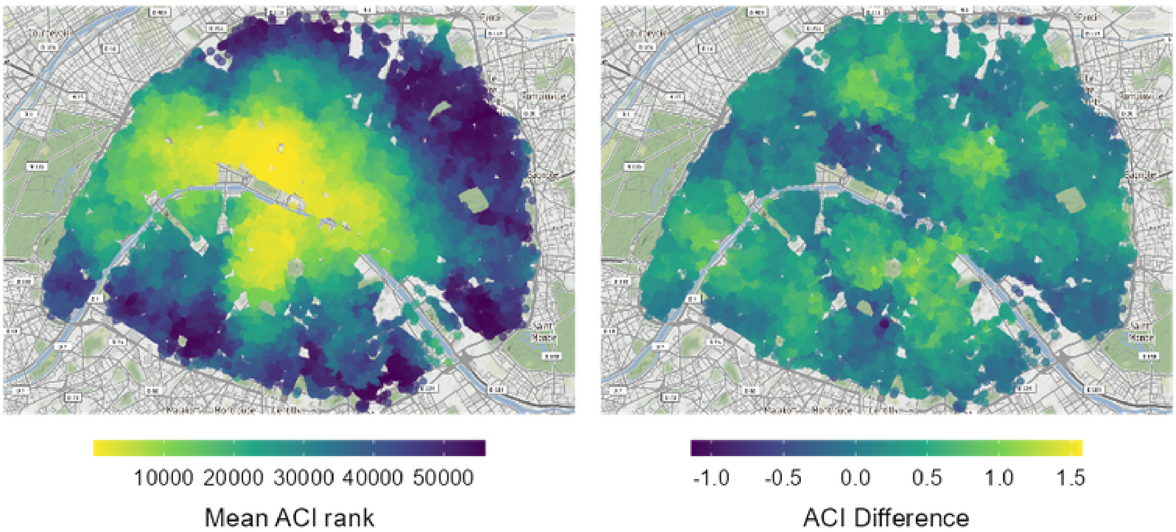


Figure 3.4: Maps of mean oACI ranks and oACI 2020-2014 differences (repeated from page 116)

to calculate the TCI and ACIs of all French cities that have over two hundred thousand inhabitants with the exception of Nice[4]. This threshold yields a list of 10 cities: Paris, Marseille, Lyon, Toulouse, Nantes, Montpellier, Strasbourg, Bordeaux, Lille, and Rennes.

To assess the relevance of the ACI and the TCI in these new settings, I focus on a few different elements from a macroscopic lens and compare the results they produce to the consumer-driven framework of the two previous chapters. First, how does the TCI change over time? It was discussed at the end of Chapter 3 that Paris, and cities in general, evolve organically through processes that take centuries. On top of this, and from a consumer-motivated perspective, TCI changes over time would imply either changes in consumer preferences, or changes in the non-market environment of places, or changes in how suppliers anticipate these elements to suit them. As such and while cities are permanently evolving, this evolution is thought to be slow and there should be no *drastic* change of TCI ranking over the 15-year period analyzed here. Second, how homogeneous is the TCI across cities, and how consistent is it with the Parisian example? From a consumer-derived perspective, and following one possible interpretation outlined in Chapter 3, it would be expected for the TCI ranks to depend on consumer preferences for types of goods and services within cities as entire selection environments. It is reasonable to expect for these tastes to be broadly similar within a country given the influence of culture, of social ties and of mobility.

These preferences were held constant in time with the previous chapter's methodology because the timeframe was short enough for the assumption to be valid. Over longer periods of time, the assumption loses validity, as various external factors could influence consumption preferences in the long term. This is why TCIs are computed as averages

[4]The city of Nice is very spread out and would need an alternative definition for its global network, or overarching selection space, than the simple city boundaries used in this work. A large amount of buildings and of the space within city boundaries is essentially rural. In order to keep the methodology as general as possible, Nice is excluded. However, Nice was also quite consistent with other cities (especially South-Eastern ones) in terms of its TCI ranks. The problem of Nice as found out in this chapter lied in maps of absolute ACI values, where the sheer amount of rural places made the distribution extremely uneven, and was a poor representation of "urban" complexity. Appendix Table 4.3 further supports that the inclusion of Nice does not affect the results outlined further along in this chapter.

over only 3-year windows for each year with the longer span *SIRENE* data, as detailed earlier. However, as unrealistic as it would be to assume that preferences are completely static over time, it would be equally problematic to observe commercial amenity TCIs to be very unstable in time. If that were the case, this framework would make the assumptions that co-preferences are volatile and that the supply of commercial amenities adapts quickly to follow them. Both would be difficult to justify. Thankfully, yearly TCI values are remarkably robust in time, as shown in Appendix Figure 4.8. A few specific amenity types seem to be undergoing structural changes in how they are distributed across cities. These dynamics would be interesting to examine on a case-by-case basis, but they are beyond the scope of this work. Nevertheless, they are rare and isolated, and the vast majority of commercial amenity types end up with a 2022 TCI rank that is close to its 2008 counterpart.

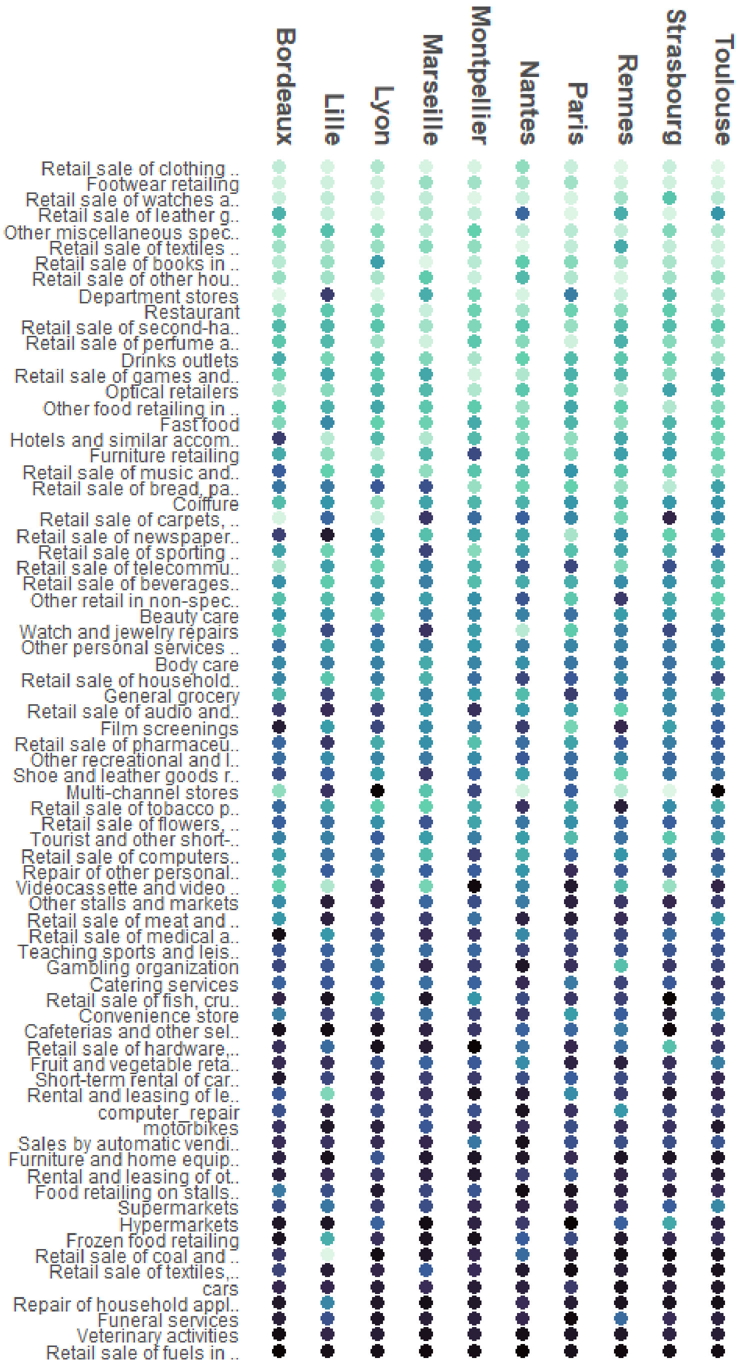
Perhaps more importantly, I now look at how TCI ranks vary across cities. This spatial problem is conceptually similar to the problem of time explained just above. Only here, TCI computations across cities are completely separate from each other, that is, no commercial amenity location within a city can affect its rank within another city. With that in mind, what can be expected from TCI ranks across cities? The consumer-framework-based answer to this question depends on how heterogeneous consumer co-preferences and the distributions of place characteristics are expected to be across cities. On the one hand, cities have different settings, backgrounds, they host different groups of people who presumably have different preferences and would generally be expected to have somewhat different co-location patterns. On the other hand, these cities all share a large part of their administration at a country level, with people that speak the same language and consume the same media. Expectations wise, this leads to a nuanced conclusion that echoes the one above: Across cities, TCIs should be mostly homogeneous yet still have some room to vary. Figure 4.4 presents the average TCI ranks of amenity types within each given city. They are ordered by highest to lowest average rank (across all cities and years), and colored according to their rank within a given city.

First, the overall amenity ranking observed in Figure 4.4 reflects a tendency that could qualitatively be estimated by a realized spending spectrum, which is consistent with the index's previous application. In other words, leisure-oriented amenities tend to be at the top, as do tourist-oriented ones or luxury ones. Second and most importantly, while different cities present different TCI ranks for certain amenity types, the overall takeaway of this figure and of its underlying data is that all cities *together* have a strong tendency for a top-down bright-to-dark gradient. In other words, given commercial amenities tend to be found in similarly high-ACI (or low-ACI) places across cities. Once again, these TCI values were computed completely independently of each other. This kind of result across so many cities is *extremely* unlikely to happen by chance, even less so with intuitive amenity types at the top and at the bottom of the rankings. It is a hugely promising indication that the TCI, and hence the ACI, is uncovering real-world phenomena.

This section demonstrated the validity of the ACI and TCI spectra at a larger spatial scale than Paris by using the *SIRENE* data in three different ways. First, I showed that *SIRENE* data, although less comprehensive and granular than *APUR* data, could provide a very similar picture of both the TCI and the ACI in Paris, showcasing the stability of the indices even in less ideal empirical settings. Second, the different methodology required by the longer time span matched the expectations set out in the previous chapter: TCIs may change over time, but these processes are slow. Third, and most importantly, TCI ranks were remarkably consistent across cities when measured independently. This paves the way for a broader and stable consumer-led interpretation of the TCI that can echo that of the previous chapter.

The following sections of this chapter will dive deeper into TCI ranks from a macroscopic lens, and into how they compare across both time and space. This section was meant to evaluate the robustness and resilience of the TCI/ACI methodology on new data, and in new settings. It has granted this work more external validity, and in doing so, it had inadvertently touched on the idea of homogeneity of TCI ranks across

Figure 4.4: Mean TCI ranks of commercial amenity types over the observed period, for each city. The brighter the dot, the higher the amenity's mean rank in the column city. Rows are ordered from top to bottom by the overall TCI rank of amenities across cities, with the best ranks at the top.



cities. In the following, I present an example of how the TCI can be used for applied research by confronting the homogeneity of TCIs to a contemporary urban challenge: The development of short-term rentals.

4.4 The ACI/TCI methodology applied to a short-term rental question

Making up for the shortcomings "complexity" measures have in direct explainability is a complicated task, and I have attempted to alleviate them throughout this work. However, a non-negligible aspect of the Economic Complexity Index's success has been its broad empirical applications. Ultimately, the use of TCI/ACI typologies will also primarily be assessed through their empirical applications. The value of this section should not be to find data that help build an interpretation of the ACI/TCI. Rather, it should showcase how the indices can help understand a process through the lens of how they were constructed.

4.4.1 Urban tourism and commercial amenities

The amenity-Type Complexity Index and the Amenity Complexity Index, in chapters 2 and 3, were constructed and explained as separation measures that segregate underlying consumers through their preferences, as anticipated by suppliers. Additional layers of interpretation, namely the spending-based nature of this consumer separation and the place-based preferences it can reveal are not as settled, but they can still serve as guides in confronting ACI/TCI spectra to real-world issues. A very neat feature of the *SIRENE* data used here, as opposed to *OpenStreetMap* or *Google Places* data, is the broad time window it covers. At a very basic level of this consumption framework, changes in retail locations in time imply changes in the composition of local consumers. By extension, a reductive separative measure like complexity can reflect how consumer compositions within different places comparatively evolve.

The idea of changes in commercial amenity locations resonates with multiple concepts in urban studies, some more readily testable than others. First and foremost, with the idea of gentrification, and in a two-way relationship. As residents of a place gentrify, the observed local commercial amenities should follow suit as high-spending-requiring suppliers can become selected. Likewise, certain patterns of local amenities can be more or less suited to attracting gentrifying populations (Couture and Handbury, 2020; Glaeser et al., 2018a). However, and to be clear, the framework used in this work is not only resident-dependent; It presents selection outcomes that also depend on non-residing consumers. Consumers that affect the place-amenity network can reside inside or outside the network (that is, the city), and no assumptions are made about their provenance. While this distinction has not been made in this chapter until now, it matters in the real world. Tensions can arise when outsiders come and alter place-amenity networks in ways that are not desirable to residents (Sigler and Wachsmuth, 2020). While this is true of gentrifiers of any kind, it is a particularly sensitive and ongoing issue regarding urban tourism. *Airbnb* and other short-term rental (STR) platforms have exacerbated an urban crisis motivated by a struggle between residents and non-residents for space (Ioannides et al., 2019; Wachsmuth and Weisler, 2018).

STR platforms, as explained in the general introduction, have contributed to stirring up mass touristification in different ways. First, by making tourism more accessible, they have intensified activity in tourism-oriented cities or areas. Second, through the flexibility they provide, they have allowed for an extensive development of tourism across cities and within areas that were until now less affected (Franco and Santos, 2021; Ioannides et al., 2019). Amenities as a broad concept (see Glaeser et al., 2001) are at the heart of discussions about the resident-tourist tensions platforms like *Airbnb* exacerbate – And it is part of the official motivation for Bordeaux's regulation, as noted in Chapter 1. The link between commercial amenities specifically and STR activity is itself a strong feature of the literature (Diaz-Parra and Jover, 2020; A. Hidalgo et al., 2023; Zervas et al., 2017). However, existing empirical literature has either focused on specific commercial amenities benefiting or losing out following STR implementation, or it has predefined amenity types as residential or tourism-oriented. On top of the lack of a systemic approach that is agnostic about the role of specific types of commercial amenities, little attention has been brought towards the effect of STR activity on commercial amenities at a cross-city scale.

This section explores a specific aspect of how the intensification of urban tourism can affect the inner-city distribution of commercial amenities, and I use the TCI to assess its effect in a holistic fashion. In particular, I examine whether urban tourism can lead to the homogenization of commercial amenity distributions across cities, implying a loss of city-specific authenticity. The underlying assumption is that the cultural heritage of cities is reflected in the current spatial consumption patterns of consumers, and thus in the co-location of commercial amenities that anticipate these patterns. In other words, the premise is that local consumers have slightly different preferences from city to city, and that these preferences are captured by my ECI-inspired procedure. The differences in consumption patterns are both indicative and constitutive of cities' uniqueness.

It is difficult to know how different tourists' consumption preferences are from each other, and from one city's visitors to another's. The presence of STR tourists could affect both the overall supply of commercial amenities within cities and its distribution through a change in the composition of consumers that is then reflected in ACI and TCI spectra. Through this mechanism, I aim to use them to empirically investigate whether mass urban tourism (more specifically, STR tourism) is associated with the loss of cities' uniqueness, as illustrated by their commercial amenity co-location patterns. This relates closely to important research on mass tourism. It also relates loosely to other urban transformation concepts, such as transnational and lifestyle-led gentrification (Sigler and Wachsmuth, 2020). I explain a simple empirical strategy that relies on correlations between TCI ranks across cities and a Two-Way Fixed Effects (TWFE) design in the following subsection. I then present and discuss the results it produces.

4.4.2 Empirical strategy

This subsection aims to find a way of observing how the co-location patterns of commercial amenities change across cities, and how these changes relate to STR activity. This work has been based on a simplified presentation of commercial amenity co-locations through two related measures: The ACI and the TCI. Let me reiterate what the changes in these measures over time would mean within the framework developed in chapters 2 and 3.

It should be noted that ACI and TCI values are not directly comparable across years. This was possible in Chapter 3 because the network (or base matrix) stacked all 3 observed years, but the assumptions that came with that were not applicable to the wider time frame of this design. Here, when tracking the changes of a given spatial unit or an amenity type, ranks should be used instead of absolute values. However, there is still a lot of information that can be obtained from both the ACI and the TCI ranks at a city-wide level.

From the ACI's perspective, changes in the spatial distribution of high-ranked places versus low-ranked ones are indicative of how the composition of local demand is changing within cities, and of how the attractiveness of places is changing. However, cross-city comparisons can be made challenging by the lack of consistent units across cities. There are no natural place-for-place comparisons, and they would have to be created. It could be interesting to compare ranks across specific spatial divisions of cities, or to match places based on prior socio-demographic (or other) characteristics. The designs these questions would require are beyond the scope of this chapter, but the ACI paves the way for answering them in interesting ways.
Comparisons using the TCI are much simpler. Because the same *SIRENE* database is used across all the observed cities, commercial amenity types themselves are cross-city comparable units. In the previous section, Figure 4.4 and Appendix Figure 4.8 both used

the coherence of amenity types to present visual comparisons. The strategy to reveal an association between the TCI and touristification, however, requires a two-way comparison across time and space.

This subsection presents a simple method to compare TCIs across cities, based on Spearman correlations. The units of observation are pairs of cities, with the correlation coefficient calculated for each year. Some amenity types may have missing values in some city-year observations; in those cases, the ranks of those amenities are discarded in the pair for that year only. Before using these correlations in a regression model, it is important to explain what they imply and what their change implies.
Let me call $S_{t,i,j}$ the similarity between the TCI ranks of two cities in year t, with

$$S_{t,i,j} = \rho(TCI_{t,i}, TCI_{t,j})$$

As a reminder, TCI values (and ranks) are simply indicative of coherence in co-occurrences between units. In other words, amenities with high ranks have similar location patterns as other amenities with high ranks. The larger the rank difference between amenities, the more different their location patterns are. As established earlier in this manuscript, this framework assumes that commercial amenity locations reveal the presence of their consumers (as anticipated by suppliers), and that the TCI is thus a segregation algorithm that separates consumers by how similar they are to each other. Therefore, TCI ranks within cities indicate the co-preferences of consumers, or groups of consumers that form a coherent consumption unit through the places they elect to consume in. Alternatively, commercial amenity type A and commercial amenity type B will be found in similar places if people consume them together, or if consumers of A and of B consume in similar places.

In the context of $S_{t,i,j}$, the interest is in whether the same types of amenities co-locate in the same way across cities. $S_{t,i,j}$ changes over time when types of retail locate differently within cities. Thus, it can change over time if consumers' preferences for amenities change, if other amenities (in the broad sense) influence them to consume in different places, or if the people that consume themselves change. The implication behind this research strategy is that urban tourism affects the composition of consumers within and across cities, and that in doing so, they affect the relative locations of commercial amenities within cities. This can be detected by $S_{t,i,j}$. There is a lot to say about these correlations within city pairs, even descriptively, and the next subsection will show examples of TCI correlations between cities. However, the overall focus of this section is to investigate a link between $S_{t,i,j}$ and STR tourism.

As a measure of urban tourism, I use data from *AirDNA*, a private provider that offers monthly observations of *Airbnb* and *HomeAway* listing activity at a very granular spatial scale. These observations are aggregated at the city-year level. This is the same data source as that from Chapter 1. The results presented later are therefore not generalizable to non-STR-driven tourism (such as hotel-driven tourism), but this concern is mitigated by the prominence of STR platforms in recent years and by their relevance in the urban discourse. Since this research does not focus on the rent gap or on housing-specific issues, but rather on the presence of tourists and their impact on places, it uses the number of reservation days (i.e., nights spent) within cities instead of the number of listings, which can also be found in the literature.

One important challenge to overcome is that the explained variable ($S_{t,i,j}$) is pair-based, whereas the reservation days are per-city. Therefore, there needs to be a way of grouping the reservation days on a city-pair basis. In other words, the goal is to discover whether the similarity $S_{t,i,j}$ of cities that are gaining STR activity at the same time changes. In the same spirit of simplicity as $S_{t,i,j}$ was defined, I choose to use a simple minimum between the yearly reservation days of city i and city j:

$$STR_{t,i,j} = min(STR_{t,i}, STR_{t,j})$$

This makes the interpretation of STR presence being linked to $S_{t,i,j}$ depend on the "weakest" STR city. In doing so, it prevents a very "strong" city from falsely driving the estimator alone in the way a mean would. However, as shown in Appendix 4.7.1, using the mean of reservation days within the pair yields similar results.

The model is set out as follows. I use a standard TWFE estimator to assess the relationship between the STR pair construction $STR_{t,i,j}$ and the Spearman correlation of TCI ranks within city pairs, $S_{t,i,j}$. Spatial fixed effects are therefore included at the pair level, is a set of covariates:

$$S_{t,i,j} = \theta_t + \eta_{i,j} + \beta^{TWFE} STR_{t,i,j} + X_{t,i,j}$$

Where θ_t are time fixed effects, $\eta_{i,j}$ are city pair fixed effects, $STR_{t,i,j}$ is considered a continuous treatment and β^{TWFE} is the estimator, my coefficient of interest. A vector of covariates $X_{t,i,j}$ is also constructed at a pair level. It includes the per-year (or closest available year) population difference within the pair and the difference in the 2013 number of hotels.

It should be kept in mind that any and all results depend on the parallel trends assumption being valid in such a setting. This is not a very straightforward assumption to defend, and weakens identification. In other words, would the "weak" cities' (in terms of STR activity) similarities to others have trended in the same way comparatively "stronger" cities' did in the absence of STR? This is a counterfactual, but it is important to understand it as an assumption. Covariates help somewhat alleviate this threat, as both the number of hotels and the population difference within a pair could tell us about how a pair's similarity would have evolved. On top of this, and while there are a sufficient number of different city pairs, they were all treated at roughly the same time; As such, some effects of STR might be diluted into yearly fixed effects, and vice versa. This design cannot lead to strong causal interpretations alone, but it can nonetheless remain an interesting exploration of a TCI-driven empirical urban research opportunity.

4.4.3 Descriptives

Before I present the results, let me shortly provide insight into the dependent variable $S_{t,i,j}$. First, all values of S are naturally bounded by 0 and 1, as S is a Spearman correlation of TCI ranks. The average similarity (that is, correlation) of pairs over all years is 0.68, with a 0.09 standard deviation.

The geography of TCI rank similarities is an intuitive angle to understand the underlying point being made in this section. Figure 4.5 maps the median $S_{t,i,j}$ for selected cities. Beyond interesting individual dynamics that are outside the scope of this chapter, values appear to exhibit spatial stickiness, especially within the southernmost cohort of cities. This stickiness is consistent with this work's understanding of the TCI as reflective of local consumption preferences, and grants it further credibility. As cities are closer together, they are likelier to share commuters, social circles, cultural practices and ultimately local dwellers and visitors with more similar consumption habits. If people who consume within cities are more alike, then it will be reflected in the co-location patterns of their commercial amenities, and the TCI appears to reflect them even with just simple correlation metrics. Beyond direct links between consumers, it is also credible that closer cities tend to share more similar distributions of place characteristics, for example through common historical development or from sharing common natural resources. While only descriptive, this is encouraging regarding the following section's results and especially promising for future research opportunities involving the TCI and the ACI.

The spatial distribution of economic activity and its changes is obviously a very important facet of evolutionary economic geography (Kogler et al., 2023), and how knowledge diffuses in space is itself a key focus. Despite being unrelated to knowledge and built from a consumption rather than a productive perspective, these maps are an indication that qualifying cities using the TCI can teach us about how information, institutions, and norms relate at cross-city scales. A consumption-driven perspective to economic ge-

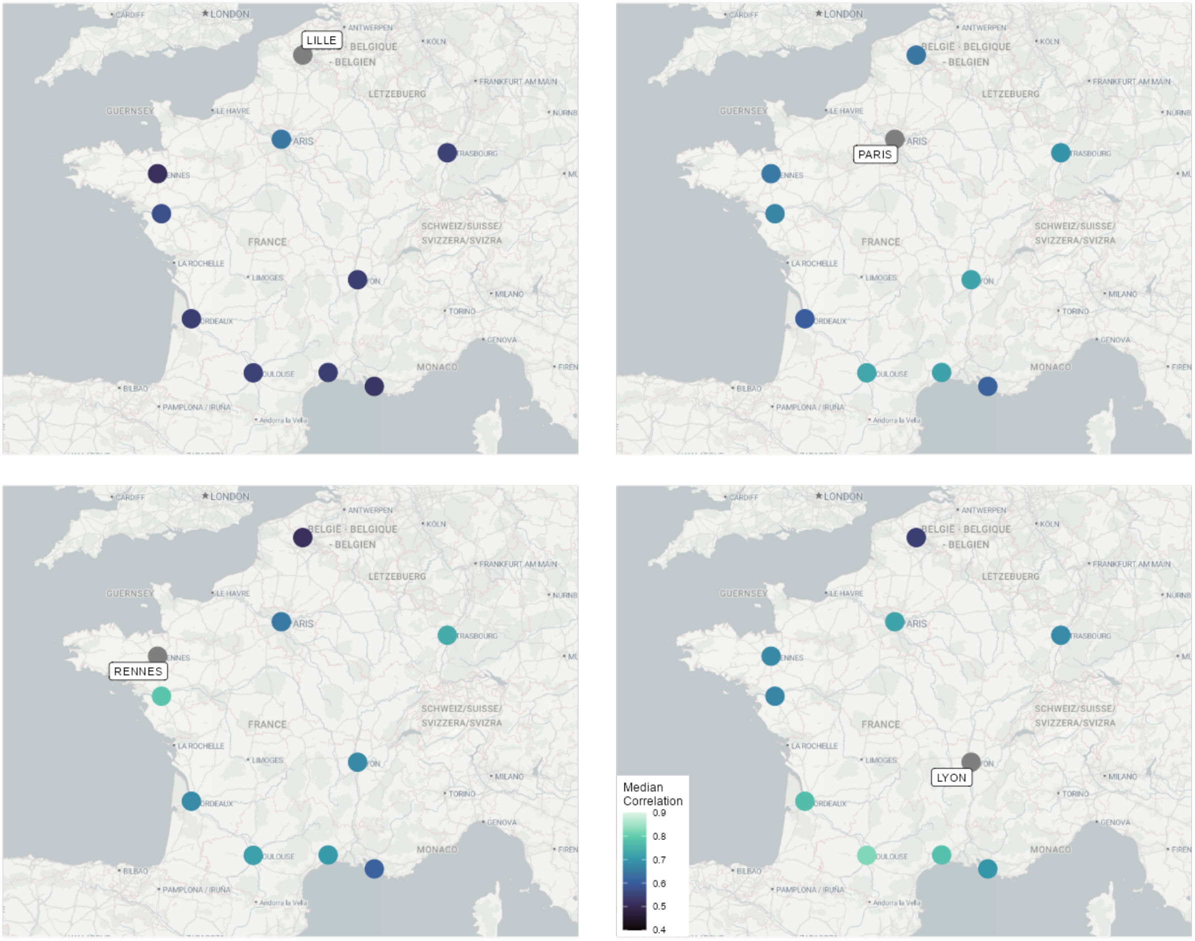


Figure 4.5: Maps of median TCI correlations (from 2008 to 2022) between the labeled (gray) city and the others. Brighter dots are cities that have a more similar TCI ranking to the labeled city's.
Map tiles: ©OpenStreetMap Contributors

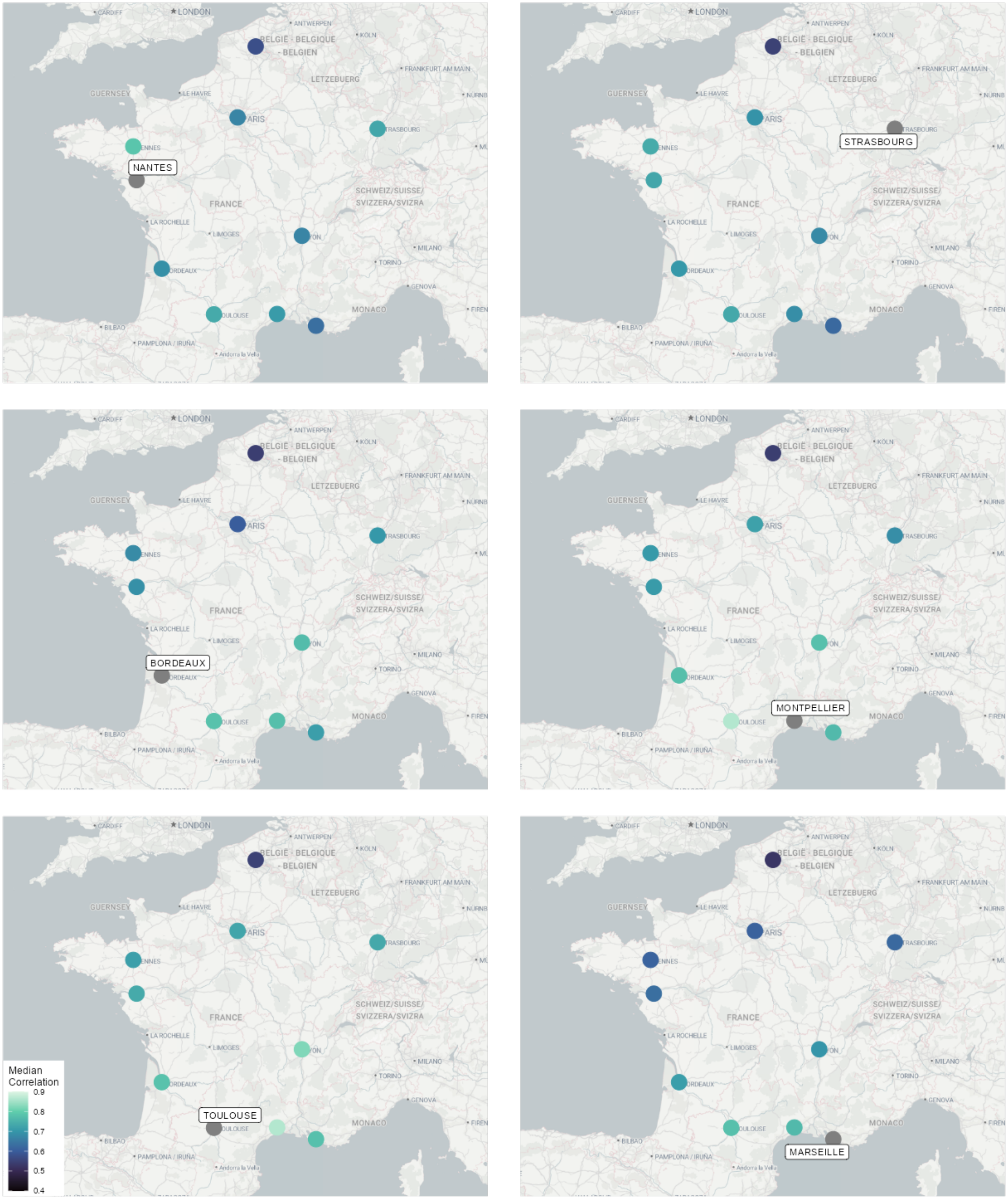


Figure 4.5: (Contunued)

ography is not dominant, and the methods expanded upon in this work would be fitting candidates to expand upon it.

In the following subsection, changes of $S_{t,i,j}$ in time will be leveraged and the relationship between these changes and STR activity will be assessed. In that context, it is important to note that the geography depicted in Figure 4.5 is mostly persistent in time. In other words, there are no *drastic* changes in how cities relate to each other through what is estimated here. Figure 4.6 depicts the evolution of the mean, median, and quartiles of $S_{t,i,j}$ throughout all observed years at an aggregate level. Altogether, the TCI-based similarity of cities has gone up over time, albeit not drastically. Inter-year variations, even at either quartile, tend to be small. The following section will present the results and attempt to untangle the role of STR activity in this uptrend.

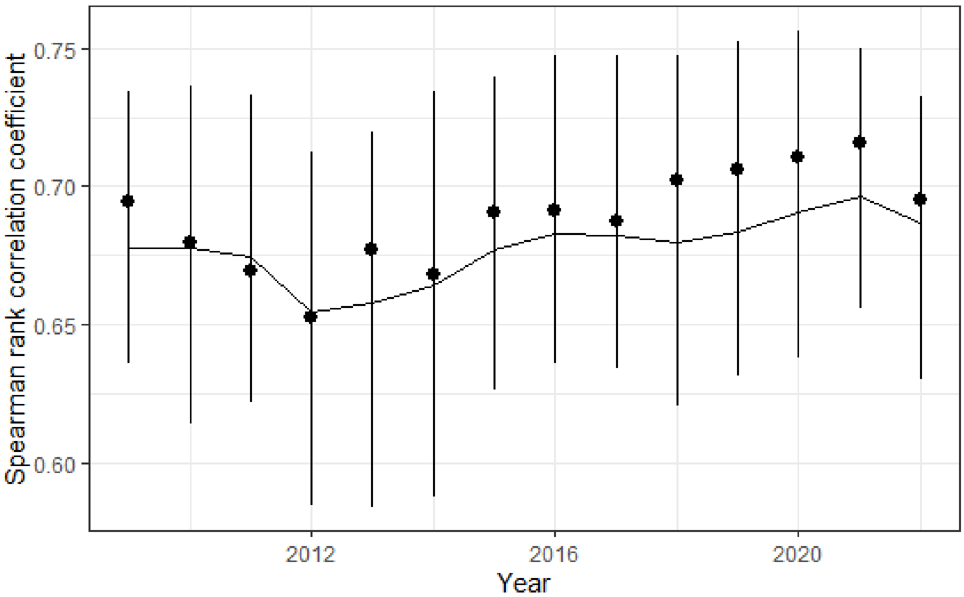


Figure 4.6: Mean (line), median (points) and quartiles (vertical lines) of $S_{t,i,j}$ in each observed year.

4.4.4 Results

The main findings of the TWFE design are presented in Table 4.2 and in an event study in Figure 4.7. Standard errors are two-way clustered at the level of each city that is a part of the pair. In both of the sets of results presented, certain regressions use STR indicators are led and lagged. This allows for more intricate causal intuitions. Concretely, for similarity values in year t between city i and city j, a one-year lag is the relationship $STR_{t+1,i,j}$ (the STR indicator in $t+1$) has with a pair's similarity. Paris is a unique city within the data. It has both a consistently high number of Reservation days and is a heavy amenity-dense tourism hotspot among the cities that are observed. Because this uniqueness might be strong enough to drive estimators alone, robustness regressions are repeated with a modified dataset of which Paris is excluded in Appendix 4.7.1, along with alternate STR indicator definitions.

A few main results stand out from Table 4.2. Positive significant relationships between the minimum number of reservation days within the pair and the Spearman correlation of their TCI rankings are present in the absence of time deferral, as seen in (1). This is robust to the introduction of covariates in regression (2). While coefficients are significant, their scale is however quite small, at under a tenth of a standard deviation.

There is however an important element to note when contextualizing these results: The hypothesis that STR visitors homogenize cross-city commercial amenity structures depends strongly on suppliers of commercial amenities. More specifically, it depends on the ability of suppliers to coordinate with the composition of local demand, which can be difficult for them to anticipate. Their reaction might not be perfect nor instant. Moreover, the model presented here carries a threat of reverse causality; That is, the similarity of cities could contain information over their attractiveness to STR tourists that the covariates miss out on. One way of addressing this issue is to lag the STR indicator.

Regression (3) uses $STR_{i,j,t-1}$ as the short-term rental indicator, and offers a view of the relationship between the previous year's minimum reservation days of a pair and its current year's TCI rank similarity. The mechanism behind the STR-led homogenization of cities is hypothesized to be reactive to the presence of STR tourists more than it is proactive. Therefore, allowing suppliers time to adapt to the composition of local

Table 4.2: TWFE Regression results

	Dependent variable: $S_{t,i,j}$			
	(1)	(2)	(3)	(4)
$\beta^{TWFE} STR_{t,i,j}$	2.90e-08**	2.94e-08**		
	[1.30e-08]	[1.20e-08]		
$\beta^{TWFE} STR_{t-1,i,j}$			**3.73e-08*****	
			[9.94e-09]	
$\beta^{TWFE} STR_{t+1,i,j}$				-4.72e-08**
				[2.25e-08]
Effect scale	7.19%	7.29%	9.25%	-11.72%
Covariates		x	x	x
Year Fixed effects	x	x	x	x
Pair Fixed effects	x	x	x	x
Sample size	312	312	312	312
Adjusted R^2	0.89	0.90	0.90	0.89

*p<0.05; **p<0.01; ***p<0.001

$S_{t,i,j}$ is the Spearman correlation of TCI ranks between cities i and j in year t.

$STR_{t,i,j}$ is the number of STR reservation days in the least STR-active city between i and j in year t.

Effect Scale corresponds to the change in the dependent variable's standard deviation for a standard deviation change in the STR indicator.

Standard errors are two-way clustered at the level of each city in the pair, and controls include the per-year population difference within the pair as well as the 2013 minimum number of hotels.

Reading (3): An extra reservation day in the least STR-intensive city of a pair was linked to a 3.73e-08 higher correlation of TCI rankings within the pair the next year.

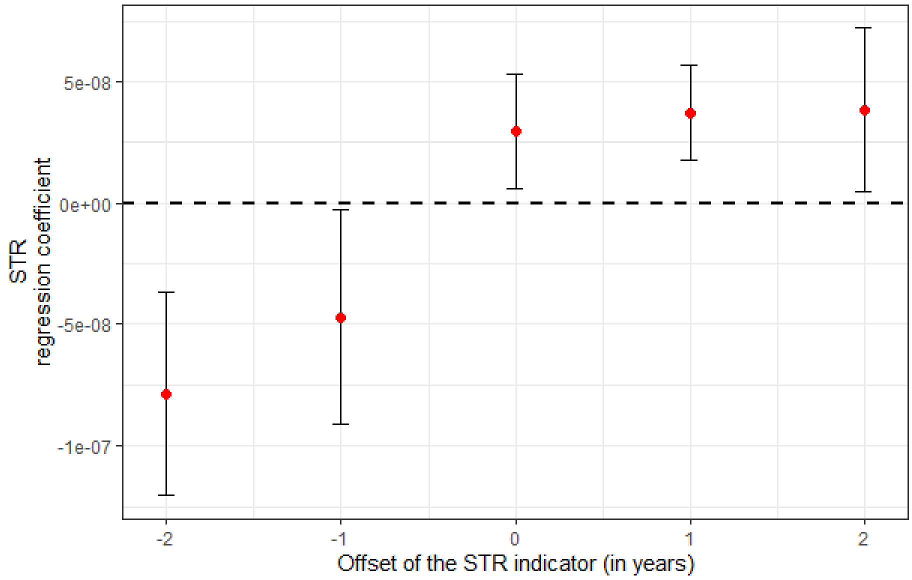


Figure 4.7: TWFE estimator leaded and lagged from $STR_{i,j,t+2}$ (leftmost) to $STR_{i,j,t-2}$ (rightmost), and accompanying 95% confidence intervals.

demand is better, and regression (3) is the preferred estimator. It remains positive, gains in significance as opposed to the same-time measure in regression (2) and adds a small amount of magnitude. It is also robust to the exclusion of Paris from the dataset, as noted in Appendix 4.7.1.

Conversely, regression (4) confronts a pair's similarity to the STR indicator in the next year. It is included primarily for robustness reasons, because a positive significant estimator would have been an indicator of reverse causality. However, it turns out to be negative and significant, which might also provide insight. In order to visualize these ideas and to dig further, Figure 4.7 extends the STR indicator from $STR_{i,j,t+2}$ to $STR_{i,j,t-2}$, from left to right respectively.

In Figure 4.7, x-axis values of -1, 0 and 1 are visual representations of regressions (4), (2) and (3) of Table 4.2 respectively. The intuition here is that observations on the left side of 0 present the reverse effect of TCI similarity on STR tourism attractiveness, whereas observations on the right side present the effect of STR tourism on the similarity of cities' commercial amenities.

There is an interesting possible explanation for the negative relationship observed when STR data is leaded: STR tourists are looking for off-the-beaten-track tourism, they are seeking an authentic experience (Diaz-Parra and Jover, 2020). These are more likely to be found in cities where residents have their own distinct tastes and cultures, and where TCI ranks are therefore less correlated to other cities' ranks. This section's results imply, under assumptions, that while STR tourists enjoy cities that have more dissimilar amenity co-location patterns, they contribute through their presence to homogenize these patterns across cities. This ties into existing literature: While STR tourists seek authentic places, they make places inauthentic (Diaz-Parra and Jover, 2020).

These results should however be interpreted very cautiously. First, because other likely but less intriguing explanations could be found to explain what is observed above. Most importantly, the model is polluted by the concordant entry of STR across cities. There is no proper control city where STRs never entered. Cities differ in their STR intensity, and it is something that is attempted to be leveraged here. But it could be that cities were already on commercial amenity similarization paths for other reasons (for example, the rise of online retailing or better cross-city transit), and that these reasons coincided with STR presence and intensity.

There are also exogenous elements that, from a supply-sided perspective, could make both STR presence likelier and cities more similar to each other. For example, student-heavy cities tend to have apartments that are more suited to STR tourism (through their availability in the summer or their sizes), and students that move between cities could likewise influence and homogenize the local demand for commercial amenities.

Last but not least, an important caveat has to be raised about the STR data used here. *Airbnb* first opened its French offices in 2012, and STR listings on offer were likely prevalent before that — especially in Paris. Yet, the *AirDNA* data used here only goes back as far as 2014 or 2016 depending on the city, and the first few years of TCI ranks are effectively made worthless because of missing data. On top of this, the STR data used stops in 2021. Therefore, only the 5-year span from 2016 to 2021 contains STR data for every city, with COVID-ridden 2020 being among that span. That leaves the model with only a few usable years to base itself on. This makes the results less robust and limits how much STR indicators can be leaded and lagged. The design would gain significantly from a longer, more up-to-date analysis.

Despite these limitations in painting a clear-cut story of the link between STR tourism and homogenization of cities, the empirical exercise undertaken in this section suggests that STR-driven tourism has been linked to a meaningful change in the local composition of demand in major French cities. Even more interestingly, to a change that is driving these cities closer together. More consequentially, it suggests an important stream of urban literature (be it planning or economic geography-inspired) can gain from reductive network measures such as the ACI and the TCI, as long as they are adequately understood.

4.5 Discussion

The objective this chapter had set out was to offer a wider scope for the application of the ACI and the TCI developed in chapters 2 and 3. This was done in two separate but related ways that present implications for the use of urban "complexity" indicators in further research.

As a whole, this work has set out to build "complexity" indices from the ground up. The Economic Complexity Index (ECI) methodology this work is derived from has the clear downside of its results being difficult to explain. I argue that, when using new data, this ground up approach has to be kept in mind in order to be able to gain substantive insight out of ECI adaptations in a fresh urban setting. The process of finding new spaces of study and applying them requires a careful understanding of what the ACI and the TCI imply.

Data selection is the first element to attentively take into account. In this chapter, I elected for a granular step-by-step spatial expansion of the ACI's and the TCI's scopes. These data were selected for their cleanliness, their official administrative background that ensures ground truth the most, and their ability to confidently track amenities over broad time periods. This was done with the downside of a narrower replicability, as *SIRENE* is limited to the boundaries of France, and with a focus on carefully validating (or refuting) the previous chapters' analysis. The approach of Juhász et al. (2023) is fundamentally different from this one despite also being an ECI adaptation, and it differs in the data it uses to build its indices too, as it leverages *Google Places* data. Further efforts could be made to deepen this string of research, either by using widely replicable (albeit messier) *OpenStreetMap* historical data or, more tediously, by expanding the use of the index to other official registry data in different countries[5]. Whichever method is selected, the practitioner should be aware of what her data and its subtleties entail, and it is difficult to argue for a one-size-fits-all adoption of the index.

For example, using *APUR* data and its very granular categories allowed for the use of a presence-based indicator in chapter 3 which is not compatible with chapter 4's category-wise more aggregated data, as matrices would be full of 1s. This chapter therefore leveraged a RCA, and has showcased a particularly desirable property of so-called complexity measures regarding different data, categorization, and binarization methods: It has remained resilient. A binary matrix that emanates from datasets with independent categorizations and that was built with a different threshold reasonably led to the same intuitions about the relative composition of local demand within Parisian places, and across Parisian amenity types. The separation the indicator reveals and the information it reduces is resilient both to weaker data and, by extension, to imperfect data. The most promising and important contribution of this chapter has been to document this resilience, as it paves the way for a wider application of the indices that adheres to the initial framework.

[5]Links to openly available business registers can be found on *Opencorporates*' website: https://opencorporates.com/registers

Still, while this work has sought to provide evidence that the ground the ACI/TCI was built on is relevant, it should be insisted on again that "complexity" as it is used here must be understood on a fundamental level as a spectral clustering algorithm. Any interpretation and intuitions attached to the indicator depend on the researcher's understanding of what separates places and types of commercial amenities in their co-occurrences. Chapter 2 has set up a framework within which to understand this separation through market-driven dynamics and suppliers' anticipations of consumer presence. Chapter 3 and this one have linked real-world data to the indices within this thesis' own framework.

While this work as a whole advocates for a careful use of ECI adaptations because of their inherent opacity, the second part of this chapter took the chance to showcase how exciting the TCI and the ACI can be as reductive spectra to understand cities differently. Analyzing the indices at an inner-city level is an important facet of this research to explore, and it can provide insights that are readily linkable to various planning issues. But the focus here has been on TCI measures, and it outlined the relevance of a cross-city approach to the problem of urban tourism and the tensions it creates. These approaches are typically complicated by data-driven intricacies, and the contribution a holistic reductive outlook like the TCI can bring is difficult to understate.

Something that is innovative within STR literature and that was allowed for by these indices is to uncover the relevance of STR externalities at a broader level than singular urban spaces themselves. The typical scale STR-linked phenomena are thought through, talked about and acted upon is either at the city level or at the sub-city place level, with some regulation-focused research mentioning cities' immediate surroundings. The cross-city approach of this chapter permitted by the TCI outlined challenges and implications that could involve urban spaces collectively. Regulating STR activity at the city level is relevant, but understanding it at a global scale could help make such regulation smarter and answer a need for coordination. In other words, local-scale administrations should not be the only ones worrying about phenomena such as STR, and wider (region, country or even continent-wide) policy-makers also have to gain from reductive measures that, while they paint broader pictures, are better hints of underlying systemic interactions. Essentially, the final section of this chapter was geared towards the impact of STR on French cities collectively and co-dependently, which to my knowledge is a new and relevant way of tackling this issue.

Extending further and even if it is pertinent at a city scale, STR regulation would be made smarter by a better understanding of how it can affect cities. The TCI and the ACI measures assist in unraveling structural similarities between cities. By better understanding which cities are structurally similar to each other, as revealed by the indices, we could better understand how similarly cities would react to issues that would otherwise be too complicated to unravel given their systemic nature.

Coincidentally, the second part of this chapter also highlighted an additional problem to assess, both conceptually and methodologically. The event study was presented in the particular setting of urban tourism and its impact on the local composition of demand, but it cannot be detached from one of the TCI's main assumptions: Commercial amenity locations depend on suppliers *anticipation* of consumer presence, not on actual consumer presence. On that note and in a evolutionary sense, the TCI could also be of use to unravel the speed at which different suppliers react to changes in the demand landscape and observe the capacity of different agents with different characteristics to adapt to these changes. The concept of adaptation in this context is fundamentally evolutionary. It motivates a push towards new perspectives to understand of the processes, impacts, and characteristics of agents behind change within urban spaces, and the uneven distribution of economic activity.

4.6 Conclusion

This chapter was meant as an addition to the work of chapters 2 and 3. It expanded the ACI for places and the TCI for retail amenity types further by using a different data source, less well-defined commercial amenity categories, a wider spatial setting and a longer time span, and by providing an example of an empirical research use-case.

The indices shined through their resilience to new data and to data-induced tweaks in the methodology, and through their consistency in a more challenging spatio-temporal setting. They are able to retain meaning across time and space, and the first section's findings are a strong boost to their external validity.

The second section contributed to the credibility of complexity methodology in urban contexts with promising results that should be further explored. I found indications that short-term rental tourism is linked to the homogenization of French cities in terms of their commercial amenities, and offered an underexplored, macroscopic and holistic view of the urban crises STR emphasizes.

I have shown that interpretations for the ACI and the TCI as laid out in chapters 2 and 3 enhance the indices' relevance in tackling urban issues through the lens of STR. STR causes tensions among people in ways that are difficult to quantify because they emanate from complex systemic processes, and the methodology presented in this thesis is suitable to characterize these processes.

4.7 Appendices

4.7.1 Additional regression settings

Two additional settings are considered for robustness tests. The first two regressions in Table 4.3 exclude Paris from the dataset. This is done to ensure results are not only driven by the most tourism-focused and by far largest city. The main result, that is, preferred regression (3) from Table 4.2, yields a positive and significant result. The reverse effect of reservation days on similarity is however not upheld in the absence of Paris.

Regressions (3) and (4) leverage a modified version of the STR indicator. Instead of just using the flat number of reservation days, in these regressions, $STR_{t,i,j}$ uses the minimum ratio of reservation days to population within the pair. Once again, the preferred regression's results are maintained (with a stronger relative effect) but the leaded version is not.

Finally, as mentioned in the data section of this chapter, the structure of Nice's administrative borders render analysis within the city difficult. Still, and as demonstrated in regressions (5) and (6), results would hold in the presence of Nice.

Table 4.3: Additional TWFE Regressions

	Dependent variable: $S_{t,i,j}$					
	Paris excluded		Alternative STR indicator		Nice included	
	(1)	(2)	(3)	(4)	(5)	(6)
$STR_{t-1,i,j}$	3.06e-08***		2.32e-02***		1.80e-08***	
	[1.17e-08]		[6.84e-03]		[3.28e-09]	
$STR_{t+1,i,j}$		-4.70e-08		1.07e-02		-4.47e-08***
		[3.05e-08]		[1.01e-02]		[9.26e-09]
Effect scale	6.33%	-9.72%	13.97%	6.53%	5.27%	-13.12%
Covariates	yes	yes	yes	yes	yes	yes
Year Fixed effects	yes	yes	yes	yes	yes	yes
Pair Fixed effects	yes	yes	yes	yes	yes	yes
Sample size	249	249	291	291	382	382
R^2	0.93	0.93	0.92	0.91	0.90	0.89
Adjusted R^2	0.92	0.91	0.90	0.90	0.88	0.87

*p<0.05; **p<0.01; ***p<0.001

Effect Scale corresponds to the change in the dependent variable's standard deviation for a standard deviation change in the STR indicator.

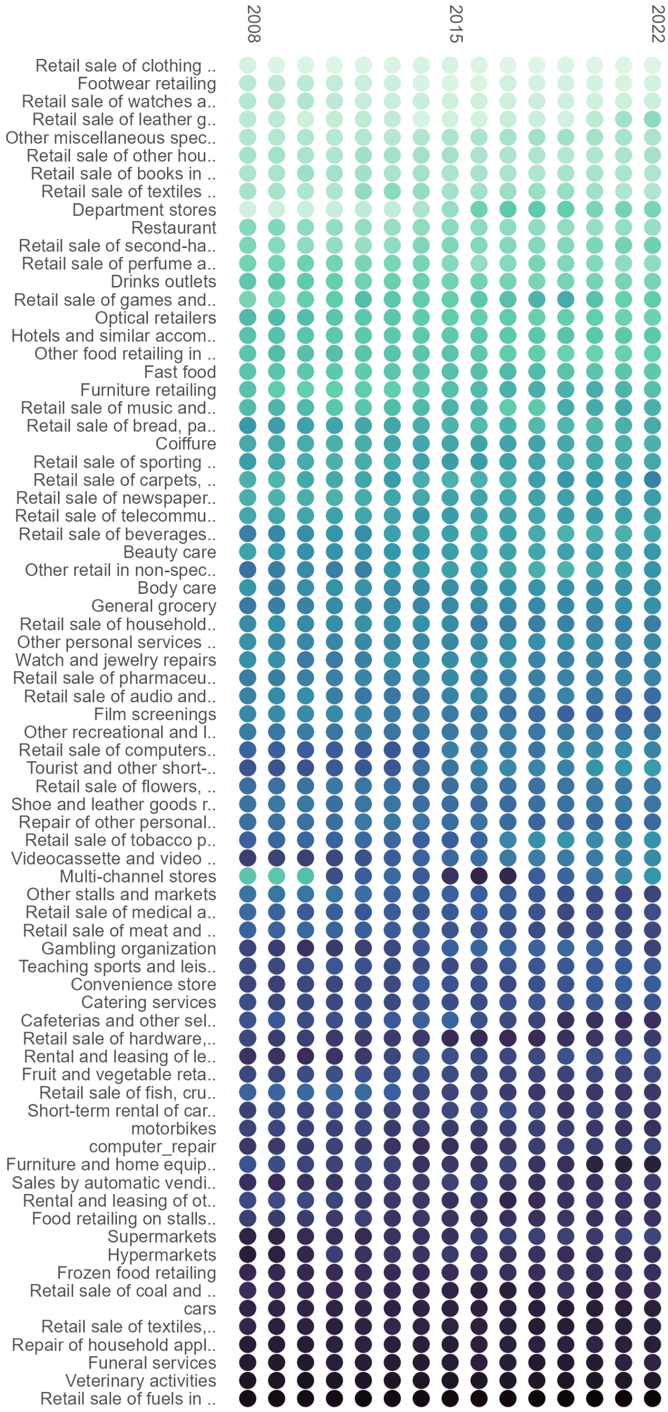
Standard errors are two-way clustered at the level of each city in the pair, and controls include the per-year population difference within the pair as well as the 2013 minimum number of hotels.

Reading (2): An extra reservation day in the least STR-intensive city of a pair was linked to a 3.06e-08 higher correlation of TCI rankings within the pair the next year.

4.7.2 Yearly variations of TCI ranks

Figure 4.8 is analogous to Figure 4.4. Instead of presenting per-city TCI values across years, it depicts per-year TCI values across cities. Over the entire set of cities, on average, TCI ranks are persistent across time. A few examples of commercial amenities experience stronger relative changes in ranks. Department stores and multichannel stores have lost out in terms of complexity since 2008, whereas videocassette stores and watch stores have comparatively gained.

Figure 4.8: Mean TCI ranks of commercial amenity types over major French cities, for each year. The brighter the point, the higher the amenity's mean rank in that year. The y-axis is ordered by the overall TCI rank of amenities across years



General conclusion

This thesis set out to assess how short-term rentals (STRs) relate to change in cities and in their neighborhoods, and in turn to contribute to informing STR regulatory design. The example of Bordeaux's regulation as assessed in the first chapter highlighted the need for a new perspective on how STRs affect urban spaces. STR tourists, hosts, and platforms change how people experience cities, and they change urban environments beyond isolated measurable outcomes. The rest of the work designated the framework of complex systems and evolutionary thinking as fitting to understand change in cities. Critically, it has contributed network-driven empirical tools inspired by economic complexity to characterize complex urban systems at the neighborhood scale. It has articulated the spatial distribution of different types of retail amenities in cities to provide holistic measures of urban situations, and thus of urban change, through the coordination of retailers with consumers. The indexes provided were proven to resiliently unveil useful information about urban systems at both the city scale and at the sub-city scale. They also underline an interesting avenue for endeavours in economic geography that take interest in the spatial distribution of consumption, rather than of production.

The indexes and the framework they are a part of are a promising avenue for future research. Addressing how the work presented here translates to STR policy designs that fit specific rationales is crucial. The ultimate section of this thesis made the unusual point of highlighting the relevance of regional and national scales[6] for STR policy. STR intensifies tourism and changes the relative trajectories of cities, and it would be fitting

[6]Bordeaux's regulation, as analysed in the first chapter, leveraged a national-level law. This law was however not meant to coordinated different cities, it rather allowed cities legislative tools to locally and independently regulate.

to address it in the same way overall spatial economic and tourism strategies are addressed. That is, through the coordination of international, national, regional and local interests. The tools presented here evaluate the systemic impacts of STR activity at the cross-city level, and can thus support various levels of governance in their quest for smarter regulation. Of course, this work's measures are also beneficial at a city-level. Spatial differentiation of regulation *within* cities, notably through licensing differences or outright bans, is also a relatively common occurrence. An empirical separation of cities' subsystems' characteristics allows for a better assessment of how prone places have been, or risk being, to transformations induced by high-spending STR tourists. It thus can be used as an additional tool to inform the city-level spatial control of STR development, be it through a lens of preservation or of economic growth. Designing smart STR regulation is difficult, but building further on the characterizations of urban systems and on their change is encouraging in the way of helping outcomes fit policy-makers' intentions.

Short-term rentals have been a leitmotif throughout this thesis, and they were the original motivation for this research. But by taking a systemic and evolutionary stance on urban spaces, this manuscript's scope was necessarily broadened. The last three chapters were focused on a holistic perspective of urban systems that, while it encapsulates STR-driven transformations, also ambitiously sought to contain everything that defines urban spaces and their evolution. This makes isolating the role of STR in urban transformations more challenging, but it also grants the indicators and the framework very diverse opportunities for future research. They are exciting tools to unravel the dynamics behind systemic evolution within cities, behind the spatial organization of human activity, the resilience of spaces or of activities, and their relative adaptability. Beyond these historical and evolutionary problematics, contemporary practical applications in planning should also not be limited to STR. Understanding modern urban crises, the changes in the geography of work and of consumption new technologies imply, spatial inequalities within cities, cities' abilities to face social and environmental challenges, and their overall resilience as sustainable systems are all research prospects that the consumption-driven holistic endeavour of this thesis can contribute towards.

Bibliography

Abadie, A., Athey, S., Imbens, G. W., & Wooldridge, J. M. (2023). When Should You Adjust Standard Errors for Clustering?*. *The Quarterly Journal of Economics*, *138*(1), 1–35. https://doi.org/10.1093/qje/qjac038

Aguiar, M., & Bils, M. (2015). Has Consumption Inequality Mirrored Income Inequality? *American Economic Review*, *105*(9), 2725–2756. https://doi.org/10.1257/aer.20120599

Aguilera, T., Artioli, F., & Colomb, C. (2021). Explaining the diversity of policy responses to platform-mediated short-term rentals in european cities: A comparison of barcelona, paris, and milan. *Environment and Planning A: Economy and Space*, *53*(7), 1689–1712.

Allam, Y. (2023). *Political decisions, discrimination, and housing market: Insights from french municipalities. economics and finance.* [Doctoral dissertation, Chapter 4, Université Bourgogne Franche-Comté, Available at : https://theses.hal.science/tel-04390099v1].

Allen, P. M. (1997). *Cities and Regions as Self-organizing Systems: Models of Complexity*. Psychology Press.

Anderson, P. W. (1972). More Is Different. *Science*, *177*(4047), 393–396. https://doi.org/10.1126/science.177.4047.393

Ayouba, K., Breuillé, M.-L., Grivault, C., & Le Gallo, J. (2020). Does Airbnb Disrupt the Private Rental Market? An Empirical Analysis for French Cities. *International Regional Science Review*, *43*(1-2), 76–104. https://doi.org/10.1177/0160017618821428

Balassa, B. (1965). Trade Liberalisation and “Revealed” Comparative Advantage1. *The Manchester School*, *33*(2), 99–123. https://doi.org/10.1111/j.1467-9957.1965.tb00050.x

Balland, P.-A. (2017). Economic Geography in R: Introduction to the EconGeo Package. *SSRN Electronic Journal*. https://doi.org/10.2139/ssrn.2962146

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

Balland, P.-A., Broekel, T., Diodato, D., Giuliani, E., Hausmann, R., O’Clery, N., & Rigby, D. (2022). The new paradigm of economic complexity. *Research Policy*, *51*(3), 104450. https://doi.org/10.1016/j.respol.2021.104450

Balland, P.-A., & Rigby, D. (2017). The Geography of Complex Knowledge. *Economic Geography*, *93*(1), 1–23. https://doi.org/10.1080/00130095.2016.1205947

Bantman-Masum, E. (2020). Unpacking commercial gentrification in central Paris. *Urban Studies*, *57*(15), 3135–3150. https://doi.org/10.1177/0042098019865893

Barron, K., Kung, E., & Proserpio, D. (2021). The Effect of Home-Sharing on House Prices and Rents: Evidence from Airbnb. *Marketing Science*, *40*(1), 23–47. https://doi.org/10.1287/mksc.2020.1227

Basuroy, S., Kim, Y., & Proserpio, D. (2020, January). *Sleeping with Strangers: Estimating the Impact of Airbnb on the Local Economy* (SSRN Scholarly Paper No. ID 3516983). Social Science Research Network. Rochester, NY. https://doi.org/10.2139/ssrn.3516983

Batty, M. (2000). Less is More, More is Different: Complexity, Morphology, Cities, and Emergence. *Environment and Planning B: Planning and Design*, *27*(2), 167–168. https://doi.org/10.1068/b2702ed

Batty, M. (2001). Polynucleated Urban Landscapes. *Urban Studies*, *38*(4), 635–655. https://doi.org/10.1080/00420980120035268

Batty, M. (2007). *Cities and complexity: Understanding cities with cellular automata, agent-based models, and fractals*. The MIT press.

Batty, M. (2009). Cities as complex systems: Scaling, interaction, networks, dynamics and urban morphologies.

Batty, M. (2012). Building a science of cities. *Cities*, *29*, S9–S16. https://doi.org/https://doi.org/10.1016/j.cities.2011.11.008

Batty, M., & Cheshire, J. (2011). Cities as Flows, Cities of Flows. *Environment and Planning B: Planning and Design*, *38*(2), 195–196. https://doi.org/10.1068/b3802ed

Batty, M., & Marshall, S. (2009). Centenary paper: The evolution of cities: Geddes, abercrombie and the new physicalism. *The Town Planning Review*, 551–574.

Batty, M., & Marshall, S. (2012). The Origins of Complexity Theory in Cities and Planning. In J. Portugali, H. Meyer, E. Stolk, & E. Tan (Eds.), *Complexity Theories of Cities Have Come of Age: An Overview with Implications to Urban Planning and Design* (pp. 21–45). Springer. https://doi.org/10.1007/978-3-642-24544-2_3

Bei, G., & Celata, F. (2023). Challenges and effects of short-term rentals regulation: A counterfactual assessment of European cities. *Annals of Tourism Research*, *101*, 103605. https://doi.org/https://doi.org/10.1016/j.annals.2023.103605

Bekkerman, R., Cohen, M. C., Kung, E., Maiden, J., & Proserpio, D. (2023). The effect of short-term rentals on residential investment. *Marketing Science*, *42*(4), 819–834. https://doi.org/https://doi.org/10.1287/mksc.2022.1409

Bertolini, L. (2010). Complex systems, evolutionary planning? In *A planner's encounter with complexity* (pp. 81–98). Routledge.

Bertrand, M., Duflo, E., & Mullainathan, S. (2004). How Much Should We Trust Differences-In-Differences Estimates?*. *The Quarterly Journal of Economics*, *119*(1), 249–275. https://doi.org/10.1162/003355304772839588

Boeing, G. (2018). Measuring the complexity of urban form and design. *Urban Design International*, *23*(4), 281–292.

Boonstra, B., & Boelens, L. (2011). Self-organization in urban development: Towards a new perspective on spatial planning. *Urban Research & Practice*, *4*(2), 99–122. https://doi.org/10.1080/17535069.2011.579767

Boschma, R., & Weltevreden, J. (2004). The evolutionary nature of b2c e-commerce in inner cities. *E-commerce, e-business and the dynamics of economic development. Ashgate, Aldershot, forthcoming.*

Boschma, R. (2017). Relatedness as driver of regional diversification: A research agenda. *Regional Studies*, *51*(3), 351–364.

Boschma, R., & Frenken, K. (2003). Evolutionary economics and industry location. *Review for Regional Research*, *23*, 183–200.

Boschma, R., & Lambooy, J. G. (1999). Evolutionary economics and economic geography. *Journal of Evolutionary Economics*, *9*(4), 411–429. https://doi.org/10.1007/s001910050089

Boschma, R., & Martin, R. L. (2010). *The handbook of evolutionary economic geography*. Edward Elgar Publishing.

Bourdieu, P. (1977). *Outline of a Theory of Practice* (R. Nice, Trans.). Cambridge University Press. https://doi.org/10.1017/CBO9780511812507

Bourdieu, P. (1987, October). *Distinction: A Social Critique of the Judgement of Taste*. Harvard University Press.

Breiman, L. (2001). Random Forests. *Machine Learning*, *45*(1), 5–32. https://doi.org/10.1023/A:1010933404324

Briel, D. v., & Dolnicar, S. (2021). The evolution of Airbnb regulation - An international longitudinal investigation 2008–2020. *Annals of Tourism Research*, *87*, 102983. https://doi.org/https://doi.org/10.1016/j.annals.2020.102983

Caldarelli, G., Cristelli, M., Gabrielli, A., Pietronero, L., Scala, A., & Tacchella, A. (2012). A Network Analysis of Countries' Export Flows: Firm Grounds for the Building Blocks of the Economy. *PLOS ONE*, *7*(10), e47278. https://doi.org/10.1371/journal.pone.0047278

Cameron, A. C., Gelbach, J. B., & Miller, D. L. (2011). Robust Inference With Multiway Clustering. *Journal of Business & Economic Statistics*, *29*(2), 238–249. https://doi.org/10.1198/jbes.2010.07136

Chen, Y., Huang, Y., & Tan, C. H. (2021). Short-term rental and its regulations on the home-sharing platform. *Information & Management*, *58*(3), 103322. https://doi.org/https://doi.org/10.1016/j.im.2020.103322

Chong, S. K., Bahrami, M., Chen, H., Balcisoy, S., Bozkaya, B., et al. (2020). Economic outcomes predicted by diversity in cities. *EPJ Data Science*, *9*(1), 17. https://doi.org/https://doi.org/10.1140/epjds/s13688-020-00234-x

Christaller, W. (1966). *Central places in southern Germany* (Vol. 10). Prentice-Hall.

Clarke, D., & Tapia-Schythe, K. (2021). Implementing the panel event study. *The Stata Journal*, *21*(4), 853–884. https://doi.org/10.1177/1536867X211063144

Cocola-Gant, A., & Gago, A. (2021). Airbnb, buy-to-let investment and tourism-driven displacement: A case study in lisbon. *Environment and Planning A: Economy and Space*, *53*(7), 1671–1688.

Couture, V., & Handbury, J. (2020). Urban revival in America. *Journal of Urban Economics*, *119*, 103267. https://doi.org/10.1016/j.jue.2020.103267

Damask, J. (2020). A consistently oriented basis for eigenanalysis. *International Journal of Data Science and Analytics*, *10*(4), 301–319.

Diaz-Parra, I., & Jover, J. (2020). Overtourism, place alienation and the right to the city: Insights from the historic centre of Seville, Spain. *Journal of Sustainable Tourism*, *29*(2-3), 158–175. https://doi.org/10.1080/09669582.2020.1717504

Dopfer, K. (2005). Evolutionary economics: A theoretical framework. *The evolutionary foundations of economics*, 3–55.

Dosi, G. (1997). Opportunities, incentives and the collective patterns of technological change. *The economic journal*, *107*(444), 1530–1547.

Duso, T., Michelsen, C., Schäfer, M., & Tran, K. D. (2020, August). *Airbnb and Rents: Evidence from Berlin* (SSRN Scholarly Paper No. ID 3676909). Social Science Research Network. Rochester, NY. https://doi.org/10.2139/ssrn.3676909

Essletzbichler, J., & Rigby, D. L. (2010). Generalized darwinism and evolutionary economic geography. In *The handbook of evolutionary economic geography*. Edward Elgar Publishing.

Essletzbichler, J., & Rigby, D. L. (2007). Exploring evolutionary economic geographies. *Journal of Economic Geography*, *7*(5), 549–571.

Falk, M., & Scaglione, M. (2024). Effects of regulations on the airbnb market in geneva. *Tourism Economics*, *30*(3), 615–632.

Florida, R., Rodríguez-Pose, A., & Storper, M. (2023). Critical Commentary: Cities in a post-COVID world. *Urban Studies*, *60*(8), 1509–1531. https://doi.org/10.1177/00420980211018072

Franco, S. F., & Santos, C. D. (2021). The impact of airbnb on residential property values and rents: Evidence from portugal. *Regional Science and Urban Economics*, *88*, 103667.

Frenken, K., & Boschma, R. (2007). A theoretical framework for evolutionary economic geography: Industrial dynamics and urban growth as a branching process. *Journal of economic geography*, *7*(5), 635–649.

Frenken, K., Meelen, T., Arets, M., & Van de Glind, P. (2015). Smarter regulation for the sharing economy. *the Guardian*. Retrieved March 25, 2021, from http://www.theguardian.com/science/political-science/2015/may/20/smarter-regulation-for-the-sharing-economy

Frenken, K., Van Oort, F., & Verburg, T. (2007). Related Variety, Unrelated Variety and Regional Economic Growth. *Regional Studies*, *41*(5), 685–697. https://doi.org/10.1080/00343400601120296

Füller, H., & Michel, B. (2014). 'stop being a tourist!'new dynamics of urban tourism in b erlin-k reuzberg. *International Journal of Urban and Regional Research*, *38*(4), 1304–1318.

Garcia-López, M.-À., Jofre-Monseny, J., Martínez-Mazza, R., & Segú, M. (2020). Do short-term rental platforms affect housing markets? Evidence from Airbnb in

Barcelona. *Journal of Urban Economics*, *119*, 103278. https://doi.org/10.1016/j.jue.2020.103278

Georges, C. (2020). A Paris, itinéraire d'un quartier devenu « village ». *Le Monde.fr*. Retrieved May 16, 2023, from https://www.lemonde.fr/m-perso/article/2020/09/11/a-paris-itineraire-d-un-quartier-devenu-village_6051874_4497916.html

Glaeser, E. L., Kim, H., & Luca, M. (2018a). Nowcasting Gentrification: Using Yelp Data to Quantify Neighborhood Change. *AEA Papers and Proceedings*, *108*, 77–82. https://doi.org/10.1257/pandp.20181034

Glaeser, E. L., Kim, H., & Luca, M. (2018b, August). *Measuring Gentrification: Using Yelp Data to Quantify Neighborhood Change* (tech. rep. No. w24952). National Bureau of Economic Research. Cambridge, MA. https://doi.org/10.3386/w24952

Glaeser, E. L., Kolko, J., & Saiz, A. (2001). Consumer city. *Journal of Economic Geography*, *1*(1), 27–50. https://doi.org/10.1093/jeg/1.1.27

Gonçalves, D., Peralta, S., & Pereira dos Santos, J. (2022, November). Short-Term Rental Bans and Housing Prices: Quasi-Experimental Evidence from Lisbon. https://doi.org/10.2139/ssrn.4281261

Hands, D. W. (2010). Economics, psychology and the history of consumer choice theory. *Cambridge Journal of Economics*, *34*(4), 633–648. https://doi.org/10.1093/cje/bep045

Hasse, B. (2021). Paris : Le quartier populaire Louis-Blanc, nouveau «Boboland» ? *leparisien.fr*. Retrieved June 8, 2023, from https://www.leparisien.fr/paris-75/paris-le-quartier-populaire-louis-blanc-nouveau-boboland-12-01-2021-8418662.php

Hausdorf, B., & Hennig, C. (2003). Nestedness of north-west European land snail ranges as a consequence of differential immigration from Pleistocene glacial refuges. *Oecologia*, *135*(1), 102–109. https://doi.org/10.1007/s00442-002-1142-y

Hausmann, R., & Hidalgo, C. A. (2011). The network structure of economic output. *Journal of Economic Growth*, *16*(4), 309–342. Retrieved June 17, 2023, from https://www.jstor.org/stable/41486815

Hausmann, R., Hidalgo, C. A., Bustos, S., Coscia, M., & Simoes, A. (2014). *The atlas of economic complexity: Mapping paths to prosperity*.

Hidalgo, A., Riccaboni, M., & Velazquez, F. J. (2023). When local business faded away: The uneven impact of Airbnb on the geography of economic activities. *Cambridge Journal of Regions, Economy and Society*, rsad009. https://doi.org/10.1093/cjres/rsad009

Hidalgo, A., Riccaboni, M., & Velázquez, F. J. (2024). The effect of short-term rentals on local consumption amenities: Evidence from Madrid. *Journal of Regional Science*, *n/a*(n/a). https://doi.org/10.1111/jors.12685

Hidalgo, C. A. (2021). Economic complexity theory and applications. *Nature Reviews Physics*, *3*(2), 92–113. https://doi.org/10.1038/s42254-020-00275-1

Hidalgo, C. A., Castañer, E., & Sevtsuk, A. (2020). The amenity mix of urban neighborhoods. *Habitat International*, 102205. https://doi.org/10.1016/j.habitatint.2020.102205

Hidalgo, C. A., & Hausmann, R. (2009). The Building Blocks of Economic Complexity. *Proceedings of the National Academy of Sciences of the United States of America*, *106*, 10570–5. https://doi.org/10.1073/pnas.0900943106

Hidalgo, C. A., Klinger, B., Barabási, A.-L., & Hausmann, R. (2007). The Product Space Conditions the Development of Nations. *Science*, *317*(5837), 482–487. https://doi.org/10.1126/science.1144581

Hochmair, H. H., Juhász, L., & Cvetojevic, S. (2018). Data Quality of Points of Interest in Selected Mapping and Social Media Platforms. In P. Kiefer, H. Huang, N. Van de Weghe, & M. Raubal (Eds.), *Progress in Location Based Services 2018* (pp. 293–313). Springer International Publishing. https://doi.org/10.1007/978-3-319-71470-7_15

Hodgson, G. M. (2002). Darwinism in economics: From analogy to ontology. *Journal of evolutionary economics*, *12*, 259–281.

Horn, K., & Merante, M. (2017). Is home sharing driving up rents? Evidence from Airbnb in Boston. *Journal of Housing Economics*, *38*, 14–24. https://doi.org/10.1016/j.jhe.2017.08.002

Hotelling, H. (1929). Stability in Competition. *The Economic Journal*, *39*(153), 41–57. https://doi.org/10.2307/2224214

Ioannides, D., Röslmaier, M., & Zee, E. v. d. (2019). Airbnb as an instigator of 'tourism bubble' expansion in Utrecht's Lombok neighbourhood. *Tourism Geographies*, *21*(5), 822–840. https://doi.org/10.1080/14616688.2018.1454505

Jackson, L. F. (1984). Hierarchic Demand and the Engel Curve for Variety. *The Review of Economics and Statistics*, *66*(1), 8–15. https://doi.org/10.2307/1924690

Jacobs, J. (1961). Death and Life of Great American Cities, 474.

Juhász, S., Pintér, G., Kovács, Á. J., Borza, E., Mónus, G., L\Horincz, L., & Lengyel, B. (2023). Amenity complexity and urban locations of socio-economic mixing. *EPJ Data Science*, *12*(1), 34. Retrieved February 19, 2024, from https://epjds.epj.org/articles/epjdata/abs/2023/01/13688_2023_Article_413/13688_2023_Article_413.html

Kaufmann, T., Radaelli, L., Bettencourt, L. M. A., & Shmueli, E. (2022). Scaling of urban amenities: Generative statistics and implications for urban planning. *EPJ Data Science*, *11*(1), 1–19. https://doi.org/10.1140/epjds/s13688-022-00362-6

Kemp-Benedict, E. (2014). An interpretation and critique of the Method of Reflections. Retrieved June 15, 2023, from https://www.sei.org/publications/an-interpretation-and-critique-of-the-method-of-reflections/

Khavarian-Garmsir, A. R., Sharifi, A., & Sadeghi, A. (2023). The 15-minute city: Urban planning and design efforts toward creating sustainable neighborhoods. *Cities*, *132*, 104101. https://doi.org/10.1016/j.cities.2022.104101

Kogler, D. F., Evenhuis, E., Giuliani, E., Martin, R., Uyarra, E., & Boschma, R. (2023). Re-imagining evolutionary economic geography. *Cambridge Journal of Regions, Economy and Society*, *16*(3), 373–390. https://doi.org/10.1093/cjres/rsad029

Koster, H. R. A., van Ommeren, J., & Volkhausen, N. (2021). Short-term rentals and the housing market: Quasi-experimental evidence from Airbnb in Los Angeles. *Journal of Urban Economics*, *124*, 103356. https://doi.org/10.1016/j.jue.2021.103356

Krugman, P. (1996). *The self organizing economy*. John Wiley & Sons.

Lalicic, L., & Weismayer, C. (2017). The role of authenticity in airbnb experiences. *Information and Communication Technologies in Tourism 2017: Proceedings of the International Conference in Rome, Italy, January 24-26, 2017*, 781–794.

Lanza, H. I., & Pittman, P. S. (2019). A peek past the vape clouds: Vape shop decline in long beach, california during 2015-2018. *Tobacco regulatory science*, *5*(5), 447.

Lee, D. (2016). How Airbnb Short-Term Rentals Exacerbate Los Angeles's Affordable Housing Crisis: Analysis and Policy Recommendations. *Harvard Law & Policy Review*, *10*, 229. https://heinonline.org/HOL/Page?handle=hein.journals/harlpolrv10&id=235&div=&collection=

Maitland, R. (2010). Everyday life as a creative experience in cities (M. D. Alvarez, Ed.). *International Journal of Culture, Tourism and Hospitality Research*, *4*(3), 176–185. https://doi.org/10.1108/17506181011067574

Martin, R., & Sunley, P. (2006). Path dependence and regional economic evolution. *Journal of economic geography*, *6*(4), 395–437.

Martin, R., & Sunley, P. (2007). Complexity thinking and evolutionary economic geography. *Journal of economic geography*, *7*(5), 573–601.

Mealy, P., Farmer, J. D., & Teytelboym, A. (2019). Interpreting economic complexity. *Science Advances*, *5*(1), eaau1705. https://doi.org/10.1126/sciadv.aau1705

Mealy, P., & Teytelboym, A. (2022). Economic complexity and the green economy. *Research Policy*, *51*(8), 103948. https://doi.org/10.1016/j.respol.2020.103948

Miller, S., Johnson, N., & Wherry, L. R. (2021). Medicaid and Mortality: New Evidence From Linked Survey and Administrative Data*. *The Quarterly Journal of Economics*, *136*(3), 1783–1829. https://doi.org/10.1093/qje/qjab004

Moreno, C., Allam, Z., Chabaud, D., Gall, C., & Pratlong, F. (2021). Introducing the "15-Minute City": Sustainability, Resilience and Place Identity in Future Post-

Pandemic Cities. *Smart Cities*, *4*(1), 93–111. https://doi.org/10.3390/smartcities4010006

Moroni, S. (2015). Complexity and the inherent limits of explanation and prediction: Urban codes for self-organising cities. *Planning Theory*, *14*(3), 248–267. https://doi.org/10.1177/1473095214521104

Moroni, S., Rauws, W., & Cozzolino, S. (2020). Forms of self-organization: Urban complexity and planning implications. *Environment and Planning B: Urban Analytics and City Science*, *47*(2), 220–234. https://doi.org/10.1177/2399808319857721

Neffke, F., Henning, M., & Boschma, R. (2011). How Do Regions Diversify over Time? Industry Relatedness and the Development of New Growth Paths in Regions. *Economic Geography*, *87*(3), 237–265. https://doi.org/10.1111/j.1944-8287.2011.01121.x

Nelson, R. R., & Winter, S. G. (1985). *An evolutionary theory of economic change*. Harvard university press.

Nelson, R. R. (1995). Co–evolution of industry structure, technology and supporting institutions, and the making of comparative advantage. *International Journal of the Economics of Business*, *2*(2), 171–184. https://doi.org/10.1080/758519306

Nicolis, G., & Prigogine, I. (1989). *Exploring complexity: An introduction*. W H Freeman & Company.

Nieuwland, S., & Melik, R. v. (2020). Regulating Airbnb: How cities deal with perceived negative externalities of short-term rentals. *Current Issues in Tourism*, *23*(7), 811–825. https://doi.org/10.1080/13683500.2018.1504899

Olden, A., & Møen, J. (2022). The triple difference estimator. *The Econometrics Journal*, utac010. https://doi.org/10.1093/ectj/utac010

Pereira, R. H. M., Saraiva, M., Herszenhut, D., Braga, C. K. V., & Conway, M. W. (2021). R5r: Rapid Realistic Routing on Multimodal Transport Networks with R^5 in R. *Findings*. https://doi.org/10.32866/001c.21262

Pisano, C. (2020). Strategies for Post-COVID Cities: An Insight to Paris En Commun and Milano 2020. *Sustainability*, *12*(15), 5883. https://doi.org/10.3390/su12155883

Portugali, J. (1999). *Self-organization and the city*. Springer Science & Business Media.

Pozoukidou, G., & Chatziyiannaki, Z. (2021). 15-Minute City: Decomposing the New Urban Planning Eutopia. *Sustainability*, *13*(2), 928. https://doi.org/10.3390/su13020928

Pugliese, E., & Tacchella, A. (2020). *Economic complexity for competitiveness and innovation: A novel bottom-up strategy linking global and regional capacities* (tech. rep.). Joint Research Centre (Seville site).

Pumain, D. (1997). Pour une théorie évolutive des villes. *L'Espace géographique*, 119–134.

Pumain, D. (2003). Une approche de la complexité en géographie. *Géocarrefour*, *78*(1), 25–31.

Rauws, W., de Roo, G., & Zhang, S. (2016). Self-organisation and spatial planning: An editorial introduction. *Town Planning Review*, *87*(3), 241–251. https://doi.org/10.3828/tpr.2016.18

Robertson, C., Dejean, S., & Suire, R. (2023). "Airbnb in the City": Assessing short-term rental regulation in Bordeaux. *The Annals of Regional Science*. https://doi.org/10.1007/s00168-023-01215-4

Roth, J., Sant'Anna, P. H. C., Bilinski, A., & Poe, J. (2023, January). What's Trending in Difference-in-Differences? A Synthesis of the Recent Econometrics Literature. https://doi.org/10.48550/arXiv.2201.01194

Samuelson, P. A. (1948). Consumption Theory in Terms of Revealed Preference. *Economica*, *15*(60), 243–253. https://doi.org/10.2307/2549561

Shi, J., & Malik, J. (2000). Normalized cuts and image segmentation. *IEEE Transactions on pattern analysis and machine intelligence*, *22*(8), 888–905.

Sigler, T., & Wachsmuth, D. (2016). Transnational gentrification: Globalisation and neighbourhood change in Panama's Casco Antiguo. *Urban Studies*, *53*(4), 705–722. https://doi.org/10.1177/0042098014568070

Sigler, T., & Wachsmuth, D. (2020). New directions in transnational gentrification: Tourism-led, state-led and lifestyle-led urban transformations. *Urban Studies*, *57*(15), 3190–3201. https://doi.org/10.1177/0042098020944041

Spencer, G. M. (2017). Knowledge neighbourhoods: Urban form and evolutionary economic geography. In *Evolutionary economic geography* (pp. 179–194). Routledge.

Stojkoski, V., Koch, P., & Hidalgo, C. A. (2023). Multidimensional economic complexity and inclusive green growth. *Communications Earth & Environment*, *4*(1), 130.

Valentin, M. (2020). Regulating short-term rental housing: Evidence from New Orleans. *Real Estate Economics*, *n/a*. https://doi.org/https://doi.org/10.1111/1540-6229.12330

van Holm, E. J. (2020). Evaluating the impact of short-term rental regulations on airbnb in new orleans. *Cities*, *104*, 102803.

von Briel, D., & Dolnicar, S. (2020, March). *The Evolution of Airbnb Regulation - An International Longitudinal Investigation 2008-2020* (preprint). SocArXiv. https://doi.org/10.31235/osf.io/t4nqs

Wachsmuth, D., & Weisler, A. (2018). Airbnb and the rent gap: Gentrification through the sharing economy. *Environment and Planning A: Economy and Space*, *50*(6), 1147–1170. https://doi.org/10.1177/0308518X18778038

Waldfogel, J. (2008). The median voter and the median consumer: Local private goods and population composition. *Journal of Urban Economics*, *63*(2), 567–582. Retrieved May 17, 2023, from https://ideas.repec.org//a/eee/juecon/v63y2008i2p567-582.html

Wright, I. J., Reich, P. B., Westoby, M., Ackerly, D. D., Baruch, Z., Bongers, F., Cavender-Bares, J., Chapin, T., Cornelissen, J. H., Diemer, M., et al. (2004). The worldwide leaf economics spectrum. *Nature*, *428*(6985), 821–827.

Wright, M. N., & Ziegler, A. (2017). Ranger: A Fast Implementation of Random Forests for High Dimensional Data in C++ and R. *Journal of Statistical Software*, *77*, 1–17. https://doi.org/10.18637/jss.v077.i01

Zervas, G., Proserpio, D., & Byers, J. W. (2017). The Rise of the Sharing Economy: Estimating the Impact of Airbnb on the Hotel Industry. *Journal of Marketing Research*, *54*(5), 687–705. https://doi.org/10.1509/jmr.15.0204

Zhang, L., & Pfoser, D. (2019). Using OpenStreetMap point-of-interest data to model urban change—A feasibility study. *PLOS ONE*, *14*(2), e0212606. https://doi.org/10.1371/journal.pone.0212606

Zhong, C., Arisona, S. M., Huang, X., Batty, M., & Schmitt, G. (2014). Detecting the dynamics of urban structure through spatial network analysis. *International Journal of Geographical Information Science*, *28*(11), 2178–2199. https://doi.org/10.1080/13658816.2014.914521

Zhong, C., Schläpfer, M., Müller Arisona, S., Batty, M., Ratti, C., & Schmitt, G. (2017). Revealing centrality in the spatial structure of cities from human activity patterns. *Urban Studies*, *54*(2), 437–455. https://doi.org/10.1177/0042098015601599

Calum ROBERTSON

Locations courte durée et changement urbain : une approche systémique et empirique

Les locations courte durée (LCD), propulsées par des plateformes comme *Airbnb*, ont redéfini le paysage de l'hébergement touristique. Ce faisant, elles se sont heurtées à une forte opposition de la part des résidents et des décideurs politiques en raison de leurs impacts sur les espaces urbains. Ce travail soutient que les manières dont les LCD perturbent les dynamiques existantes au sein des villes n'est pas isolable à des conséquences seules, et qu'elles perturbent les villes en tant que système. Il donc une approche systémique. S'inspirant des courants évolutionaires en économie, il exploite les emplacements de commerces comme reflétant des situations systémiques urbaines par le biais de mécanismes de marché. Il fournit une méthodologie pour caractériser ces situations de manière empirique. Cette thèse adapte les mesures d'*economic complexity* pour les inscrire dans un cadre de changement systémique urbain, et propose deux indicateurs : L'*Amenity Complexity Index* (ACI) et l'*Amenity-Type Complexity Index* (TCI). Ceux-ci séparent de manière unidimensionnelle les espaces urbains et les types de commerces à l'échelle de la ville. Leurs interprétations sont ici consolidées comme reflétant les compositions locales de demande, et les situations des systèmes. Ils peuvent être utilisés pour évaluer les transformations urbaines sous un nouvel angle. Ces indicateurs sont résistants aux changements de données et méthodologiques, et il est pertinent de les utiliser pour aborder la question des LCD. Dans son ensemble, cette thèse démontre qu'une approche évolutive et holistique des systèmes urbains et de leurs changements est prometteuse pour étendre la compréhension des transformations urbaines induites par les LCD.

Mots clés: Transformations urbaines, Locations courte durée, Systèmes complexes, Economie de la complexité, Evolutionisme, Aménités

Short-Term Rental and Urban Change: An Empirical Systemic Approach

Short-term rentals (STRs), propelled by online platforms like *Airbnb*, have taken over the tourism accommodation industry. In doing so, they have faced strong opposition from residents and policy-makers because of their impacts on urban spaces. This work advances that STRs disturb existing dynamics within cities in ways that would be impossible to isolate, and that they have the power of affecting seemingly unrelated components of urban spaces as systems. Drawing on evolutionary theory, it leverages retail locations as reflective of urban systemic situations through market mechanisms, and provides a methodology to characterize these situations empirically. Measures of economic complexity are adapted to fit into a framework of urban systemic change, yielding the *Amenity Complexity Index* (ACI) and *Amenity-Type Complexity Index* (TCI). They separate sub-city level urban spaces and types of commercial amenities one-dimensionally. Their interpretations are consolidated in this thesis as reflective of compositions of demand and of systemic situations, and they can be used to assess urban change under a new light. These indicators are resilient to data and methodological changes, and they are adapted to study the specific issue of STR. All together, this thesis demonstrates that an evolutionary and holistic approach to urban systems and their change is a valuable addition to current understandings of STR-driven transformations.

Keywords: Urban transformations, Short-term rentals, Complex systems, Economic complexity, Evolutionism, Amenities

La Rochelle Université

NUDD

Usages du Numérique pour le Développement Durable

23 Avenue Albert Einstein, 17000 La Rochelle

Région Nouvelle-Aquitaine