Politecnico di Torino

Ingegneria Gestionale
(Management Engineering)
A.Y. 2025/2026

Graduation session: March 2026

Analysis On Airbnb Cross-Side Network Externalities:

The region of Nouvelle-Aquitaine (France)

Professor:

Prof. Francesco Luigi Milone

Student:

Simone Brandimarte

Table of Contents

Table of Figures

Table of tables

1. Introduction

1.1 Goal of the Thesis

This thesis offers an empirical assessment of the strength of cross-side externalities on the home-sharing platform Airbnb. Whereas much of the existing literature treats digital platforms as global and delocalized marketplaces, this study investigates whether network effects vary across space. Because Airbnb's services are delivered in specific physical locations, the platform effectively operates through multiple geographically bounded sub-markets. This work therefore evaluates Airbnb through the lens of a spatially localized digital platform.

The findings indicate that differences in local market conditions and industry structures give rise to distinct patterns of network effects across sub-markets. To examine these mechanisms, the study focuses on the region of Nouvelle-Aquitaine (France), a context that makes it possible to estimate the magnitude of cross-side externalities by leveraging two concurrent shocks affecting both supply and demand.

Together, these two events, the LGV opening (positive demand shock) and the Bordeaux STR (Short Term Rent) regulation (negative supply shock), create a unique natural experiment framework. The simultaneous occurrence of a demand side push and a supply side restriction within the same region allows this thesis to isolate, quantify, and compare the bidirectional CNEs that characterize a spatially localized multi-sided platform like Airbnb. In particular, the analysis focuses on the demand side shock generated by the opening of the LGV high speed rail, which is used as an exogenous source of variation to estimate how increases in demand affect host entry. This setting also helps address the endogeneity and reverse causality problems typically found when studying network effects in two-sided markets.

1.2 Airbnb

Airbnb is the most popular home share platform in the world. It was founded in August 2008 in San Francisco by Brian Chesky, Joe Gebbia, and later Nate Blecharczyk, derived from a simple idea: renting out spare space to help pay rent during a design conference when all hotels were sold out. The founders launched AirBed & Breakfast, offering three air mattresses and breakfast to attendees, the prototype of what would become a global platform. After struggling to gain traction, the founders entered Y Combinator, the biggest startup accelerator and venture capital firm, in early 2009, where they refined the business model, improved the user experience, and rebranded as Airbnb, shifting from airbeds to all types of private homes and rooms. A key early growth tactic was professionalizing listing photos: after observing that poor images reduced bookings, Airbnb hired photographers to shoot hosts' homes, dramatically improving conversion. Between 2010–2012, Airbnb expanded internationally, opening offices in major global cities and raising venture capital (notably from Sequoia Capital). The platform scaled rapidly as travellers embraced cheaper, more local alternatives to hotels, and as homeowners recognized the opportunity to monetize underused living spaces.

By 2013–2015, Airbnb became a mainstream global hospitality player, reaching over 1 million listings, launching mobile apps, and introducing trust-building mechanisms such as verified IDs, reviews, and secure payments. During these years, the company increasingly faced regulatory challenges as cities debated the impact of short-term rentals on housing markets and neighbourhood externalities, triggering the first waves of local legislation. The period 2016–2019 marked consolidation and diversification. Airbnb introduced Experiences (activities hosted by locals), strengthened partnerships with cities, and became a central player in what scholars call the sharing economy. Academic research documented the platform's effects on hotel competition, tourism demand, housing prices, and neighbourhood dynamics (e.g., Zervas et al., 2017; Filippas & Horton, 2023).

Airbnb's trajectory shifted dramatically in 2020 with the COVID-19 pandemic, which affected global travel and forced major layoffs. However, the platform adapted quickly: domestic, rural, and long-term stays surged as users sought housing flexibility, remote-work locations, and safer alternatives to hotels. These shifts reshaped Airbnb's core market and accelerated product changes such as longer-term rentals and flexible search features. In December 2020, Airbnb went public in one of the most successful IPOs in tech history, symbolizing the platform's transformation from a small startup renting air mattresses into a global marketplace operating in more than 190 countries. Since then, Airbnb has continued to evolve through new categories, upgraded host tools, and a strategy increasingly centred on personalization, trust, and sustainable growth, while navigating ongoing debates on regulation, local housing markets, and the role of digital platforms in urban environments.

Airbnb became profitable in 2022, following five consecutive years of losses (with 2017 being the first year for which complete loss data is available). In May 2025, the platform introduced a new suite of features under the category Services, while also enhancing its existing Experiences offering.

In an interview with *Business Insider*, CEO Brian Chesky stated that he envisions transforming Airbnb into an "everything app," openly drawing inspiration from Amazon's evolution, from a small online bookstore into a global e-commerce giant. Chesky highlighted that users can now book a wide range of personalized experiences and on-demand services directly through the platform, including private chefs, personal trainers, photographers, and even in-home massage sessions. [1]

Figure 1 illustrates how Airbnb's activity reflects the characteristics of a *spatially localized digital platform*. The company's revenues consistently peak in Q3 (July–September), which corresponds to the summer period in the Northern Hemisphere, where the majority of global tourism demand originates and where most Airbnb listings are located. This seasonal pattern is economically intuitive: during summer months, travel activity increases sharply as individuals and families take holidays, leading to higher booking volumes and thus higher platform revenues. The repeated Q3 revenue peaks therefore highlight how Airbnb's performance is closely tied to localized, physical consumption of services, even though the matching process between hosts and guests takes place digitally. [2]

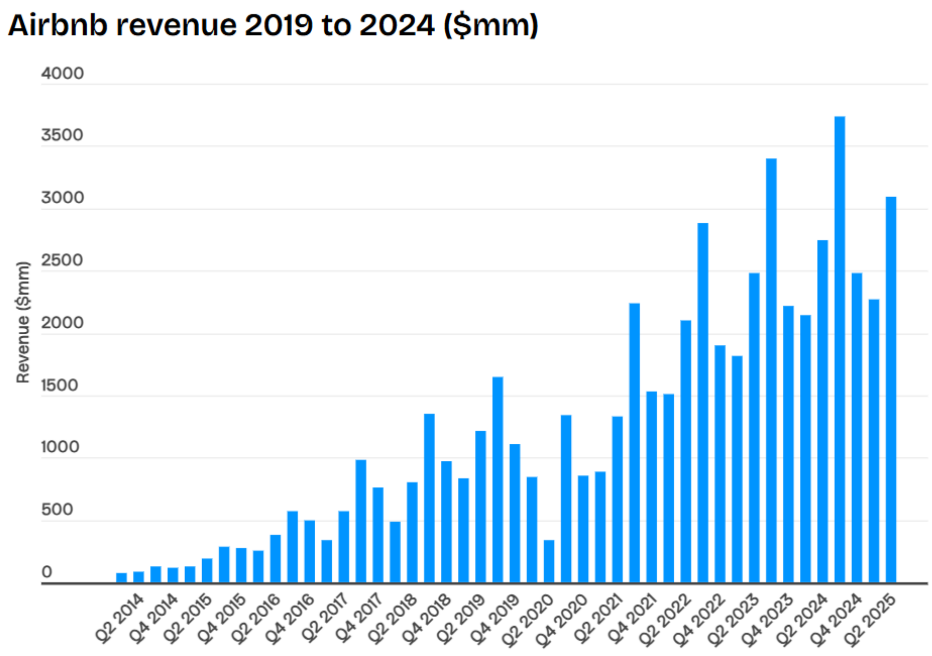

Figure 1, Airbnb Revenue

1.3 Theoretical concepts

This section will propose several theoretical definitions useful for the proper comprehension of the thesis.

Multi-sided digital platforms

By multi-sided digital platforms we refer to online infrastructures that enable interactions between two or more distinct user groups. They function like a traditional physical marketplace that has been translated into a single digital environment, where users can meet, match, and satisfy their respective needs.

In general, a digital platform is composed of four fundamental building blocks (Prof. Milone, lecture on digital platforms):

- **Value Unit:** The information, good, or service exchanged on the platform. Unlike traditional businesses and intermediaries, the platform does not produce the value unit itself.
- **Participants (Producers and Consumers):** The user groups that exchange the value unit through the platform.
- **Core Transaction:** The successful exchange of the value unit between producers and consumers, which typically represents the main source of monetization for the platform.
- **Matchmaking Tool:** The mechanism through which the platform facilitates matching between producers and consumers.

Over the years, as will be discussed in the literature review, researchers have distinguished between two main types of digital platforms.

Delocalized digital platforms

Platforms such as Amazon or Alibaba, where matching, interaction, and the exchange of value units or services can occur at any time and from any location. The value exchange is not tied to a specific geographic space.

Spatially localized digital platforms

Platforms such as Airbnb or Uber, where transactions depend on local market characteristics. The matchmaking and buying process is carried out online while the service is consumed physically in a specific place and time, so users are affected by local regulations, constraints, and geographic features.

Network externalities

Network externalities, also known as network effects, describe the phenomenon whereby the value of a product, service, or platform increases or decreases as more users participate in the system. Originating from the seminal work of Katz and Shapiro (1985), network externalities represent a fundamental principle in the economics of digital markets and are particularly prominent in multi-sided platforms. Unlike traditional goods, whose utility is largely independent of the number of consumers, digital platforms exhibit feedback mechanisms that amplify or diminish user value based on platform participation. These dynamics can manifest within the same group of users (same-side network externalities) or across different groups interacting on the platform (cross-side network externalities).

Chicken-egg problem

The chicken-and-egg problem describes the fundamental challenge that digital multi-sided platforms face during their early stages. Because the value of a platform depends on the presence of two user groups, for example hosts and tourists on Airbnb or drivers and commuters on Uber, each side is reluctant to join unless the other is already active. Producers, such as sellers or service providers, do not want to participate if there are no consumers, while consumers see no benefit in joining if there is no available supply. This circular dependency creates a coordination failure that slows down platform adoption. Caillaud and Jullien (2003) highlight that overcoming this initial imbalance is crucial for triggering the positive feedback loops generated by network externalities, and thus for enabling early-stage platform growth.

2. Literature Review

2.1 Delocalized Digital Platforms

Thanks to the increasing availability of fine-grained data, the literature on digital platforms has flourished in recent years. Although drawing an exhaustive review on the topic of digital platforms is beyond the aim of this study (see Sanchez-Cartas & León, 2021 and Sriram et al., 2015 for efforts in this direction), I hereby provide a summary of the main research strands, as well as an in-depth focus on empirical research on cross-sectional externalities, with the aim to better unveil the novel aspects of the work.

While early contributions on network economies date back to the 1980s (see Katz & Shapiro, 1985, 1986, and 1994; Farrell & Saloner, 1986), the first articles to theoretically define digital platforms appeared in the early 2000s, namely Evans (2003) and Rochet & Tirole (2003). Evans (2003) defines a platform as a market based on the presence of cross-side dependency between two or more groups. Rochet & Tirole (2003) propose a definition of platforms based on the correlation between asymmetric pricing structures (i.e., the different prices charged to each side for the "participation" in the platform) and transaction volume.

One of the first contributions in the empirical literature on cross-side externalities dates back to Rysman (2004), who estimates the presence of significant cross-side externalities in the Yellow Pages MSM, pointing out that an increase in the advertisers' installed base has positive effects on users and vice versa. In the same fashion, the contributions of Nair et al. (2004) and Ackerberg & Gowrisankaran (2006) highlight the existence of positive externalities, respectively, between the hardware installed base and software availability in the personal digital assistant markets, and between banks and customers in the Automatic Clearing House (ACH) market. Following Rysman (2004), the empirical contribution of Wilbur (2008) emphasizes that cross-side effects can trigger not only positive feedback loops but also negative ones. Indeed, according to Wilbur (2008), in the television industry, while audience growth generates a positive externality on advertiser demand, an increase in advertising creates a negative effect on audience size.

One of the most recent and relevant contributions to the field is the research by Chu & Manchanda (2016). This important work underscores the significance of CNE within a consumer-to-consumer platform named Taobao, owned by the Chinese conglomerate Alibaba. The research demonstrates that there is a subsidy side, as the effect of new suppliers in attracting new customers is four times stronger than the reverse effect, introducing a significant asymmetry in externalities. Furthermore, by analysing the variation of CNE over time, the authors observed that demand-to-supply externalities remain stable, while supply-to-demand externalities initially increase before declining over time.

A different perspective is offered by Hinz, Otter and Skiera (2020), who argue that most previous empirical studies on two-sided markets overlook a crucial aspect: they do not separately account for all the dynamic changes occurring on both sides of the platform. Traditionally, research on CNEs and SSNEs (Same Side Network Externalities) relies on net changes, computed as the difference between inflow (new users joining) and outflow (users leaving). However, this aggregate measure can mask important dynamics. For example, a net increase of six suppliers may in reality result from an inflow of ten and an outflow of four, meaning that the net figure hides the true magnitude and direction of the underlying processes.

To address this limitation, the authors propose an influx - outflow model that decomposes platform growth into eight distinct network effects, capturing how the existing installed base influences both acquisition and churn on each side of the market. Specifically, they estimate:

1. Same side effects on buyer acquisition
2. Cross-side effects on buyer acquisition
3. Same side effects on buyer churn
4. Cross-side effects on buyer churn
5. Same side effects on seller acquisition
6. Cross-side effects on seller acquisition
7. Same side effects on seller churn
8. Cross-side effects on seller churn

The authors demonstrate that this model performs particularly well when inflow and outflow are positively correlated, because in such cases net values risk understating or even cancelling out true network effects. Conversely, when inflow and outflow move in opposite directions (negative correlation), the traditional net-change approach may still provide a reasonable approximation. Finally, their paper also offers a comprehensive literature review of previous work on de-localized digital platforms, as summarized in Figure 1.

Author(s)	Main research topic(s)	Industry / data set(s)	Economic dependent variable(s)	Considers		
				CNE	SNE	Influx vs. Outflow
Brynjolfsson and Kemerer (1996)	Installed base on price	Spreadsheet software	Prices	Yes	No	No
Gandal, Kende, and Rob (2000)	Hardware, prices and software on diffusion	CD players and titles	Change in variety and sales	Yes	No	No
Shankar and Bayus (2003)	Network strength in competition	Video game consoles	Network strength	Yes	No	No
Asvanund, Clay, Krish-nan, and Smith (2004)	Incremental value of additional users	Peer-to-peer networks	Network value	No	Yes	(Yes) using proxy
Nair, Chintagunta, and Dubé (2004)	Indirect network effects in competition	PDAs and software	Hardware demand, software provision	Yes	No	No
Rysman (2004)	Importance of cross-side network effects	Yellow Pages	Consumer and advertiser demand	Yes	No	No
Clements and Ohashi (2005)	Indirect NEs, hardware diffusion	Video game systems	Hardware and software adoption	Yes	No	No
Ackerberg and Gowrisankaran (2006)	NEs for banks and customers	ACH banking	Number of transactions	Yes	No	No
Mantrala, Naik, Sridhar, and Thorson (2007)	Marketing invest on profits	Newspapers	Subscriptions, ad revenue, sales	Yes	No	No
Rysman (2007)	Card usage and acceptance	Payment card transactions	Choice of favorite network	Yes	No	No
Wilbur (2008)	Ads on audience size and vice versa	TV ads	Viewer and advertiser demand	Yes	No	No
Liu (2010)	Pricing strategies	Video game consoles	Software and hardware demand	Yes	No	No
Tucker and Zhang (2010)	Installed base on listing behavior	Classifieds platform	Number of listings	Yes	Yes	No
Sridhar, Mantrala, Naik, and Thorson (2011)	Optimal marketing invests with cross-side network effects	Local newspaper	Demand from both sides	Yes	(Yes)	No
Chao and Derdenger (2013)	Network effects on optimal price structure	Portable game consoles	Associated prices	Yes	No	No
Lee (2013)	Effect of vertical integration	Video game industry	Demand from both sides	Yes	No	No
Voigt and Hinz (2015)	Network effects on revenue; revenue-optimal user split	Online dating platform	Revenue	Yes	Yes	No
Chu and Manchanda (2016)	Quantification of CNE and SNE	C2C platform	Growth of installed bases	Yes	Yes	No
This Paper	Separation of influx and outflow with respect to network effects	B2C platform	Growth of installed bases	Yes	Yes	Yes

CNE: Cross-side network effects; SNE: same-side network effects; PDA: personal data assistant; ACH: automated clearinghouse. Influx: number of customers that flow to the market, i.e., are new to the market; Outflow: number of customers that drop out of the market, i.e., churn from the market.

Figure 2, List of previous work on delocalized digital platforms from Hinz, Otter and Skiera (2020)

2.2 Spatially Located Digital Platforms

One of the key contributions provided by this thesis is the analysis conducted on Airbnb, which can be classified as a spatially located digital platform. Indeed, all the works mentioned so far have analysed digital platforms as unlocalized entities where trade can take place globally without the influence of local conditions. However, Airbnb is subject to local factors that can affect both the supply and demand sides of the company. The reality is that in sharing economy companies such as Airbnb, BlaBlaCar, or Uber, while the matching between supply and demand happens online, the actual consumption of the service is tied to a specific geographic location.

The purpose of this work is to assess the impact that local conditions may have on the CNE of both sides of the platform, an impact that is entirely irrelevant in a delocalized entity such as Taobao. Following this line of reasoning, this study builds upon Stallkamp & Schotter (2021), who first argue for the presence of 'geographic boundaries' within an online platform, and Kim et al. (2022), who provide a pioneering contribution on the spatial identification of network effects, estimating that adoption in a fantasy sports market platform is significantly asymmetric across income levels of different US counties.

Another relevant research study that shares a similar objective with this thesis is the work by Stourm and Albuquerque (2024), which introduces the new concept of Flowers and Bees in the context of a sharing economy platform. This recent and detailed study analyses the impact of local network effects within the same city and spatial network effects across different cities in a car-sharing platform. This new paradigm shifts the focus from the traditional chicken and egg problem to a flowers and bees analogy, where flowers represent providers as they remain stationary in a specific location, while consumers act as bees, traveling to access the service. The same logic can be applied to Airbnb. The main findings of this research assess the impact of local conditions across cities in France and how they shape the supply and demand of the service. Indeed, it has been demonstrated that depending on geographic location, an increase in supply can lead to different effects. For example, higher supply in Paris leads to an increase in local consumer adoption, strengthening local network externalities, while an increase in supply in more remote areas of France, such as Corsica, results in greater demand from people in other regions, amplifying spatial network externalities. Another significant takeaway is the presence of information asymmetry, which heightens supply to demand externalities and make consumers less powerful on attracting new providers. This information asymmetry is explained by the fact that consumers have full access to provider data, including the location of the car, price, and car details, while providers do not have access to customer data. Through "seeding" experiments, it has been demonstrated that increasing supply in larger cities leads to the fastest growth for the platform.

One of the limitations of this paper is that the "seeding" experiments may not fully reflect real world conditions. As stated in the conclusion, the authors assume the absence of unobserved shocks that could impact consumer and provider adoptions. However, this is precisely the strength of this thesis , as two exogenous events are taken into consideration, each affecting the demand and supply sides.

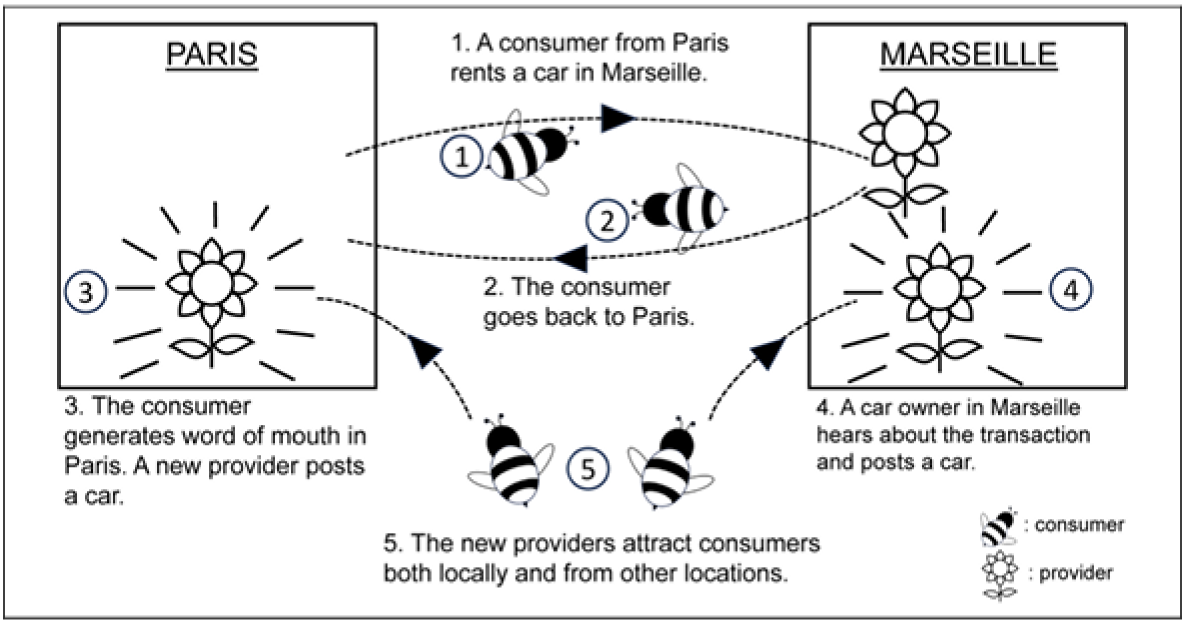

Figure 3, Illustration of Local and Spatial Network Effects from Stourm and Albuquerque (2024)

A further study on another two sided spatially located platform is Cullen and Farronato (2021). It examines the growth dynamics of TaskRabbit, a peer-to-peer marketplace for local, time-sensitive domestic services. Using detailed city-level data, the authors investigate whether the platform exhibits network effects and why some geographic markets scale more successfully than others. Their central finding is that TaskRabbit does not generate meaningful network effects, despite being a two-sided platform. Three mechanisms explain this result.

First, labour supply on TaskRabbit is extremely elastic on the intensive margin: when buyer demand doubles, active sellers nearly double the number of offers they submit, allowing the platform to accommodate more transactions without requiring additional seller entry. Because increased demand does not induce more sellers to adopt the platform, but instead leads existing sellers to work more hours, demand shocks do not permanently raise the platform's value. As a result, the platform does not experience the positive adoption feedback loops characteristic of cross-side network externalities.

Second, the platform is constrained by two-sided geographic proximity. Both buyers and sellers must be simultaneously close in space and time, as services are provided on-site and usually within a very short horizon.

This implies that each city is fragmented into many micro-markets defined at the neighbourhood level. Additional sellers in one part of the city do not increase the matching probability for buyers located elsewhere. Because these gains do not spill over space, the market cannot thicken in a way that generates increasing returns to scale.

Third, the services exchanged on TaskRabbit are highly heterogeneous. Tasks differ substantially in required skills, duration, and complexity, and sellers are not perfectly substitutable. Consequently, additional sellers only increase platform value for specific task categories and locations. Empirically, Cullen and Farronato show that doubling the number of buyers and sellers leads to a proportional (but not super linear) increase in matches, confirming the absence of economies of scale in matching, meaning that matching productivity does not increase with market size.

These structural characteristics sharply contrast with platforms such as Airbnb and Uber, which do exhibit strong cross-side network effects and scale advantages. On Airbnb, supply is capacity-constrained, hosts cannot instantaneously increase the number of rooms available by "working more." When guest demand rises, the resulting increase in occupancy and revenue attracts new hosts to join the platform. Because guests are mobile and travel to the listing, proximity constraints apply primarily to supply, not to both sides simultaneously. This one-sided proximity enables supply-side density to increase the platform's attractiveness for many potential guests, generating measurable cross-side network effects and increasing returns to scale. Similarly, Uber benefits from high service standardization and driver mobility. A ride is a highly homogeneous service, and drivers can reposition themselves across the city. Increases in rider demand attract new drivers to adopt the platform, which in turn reduce waiting times and improve reliability for all riders. These improvements feed back into increased demand, creating a virtuous cycle typical of platforms with strong network effects.

TaskRabbit differs fundamentally from Airbnb and Uber. Its two-sided geographic constraints, highly elastic seller effort, and service heterogeneity prevent the emergence of network effects and limit the scalability of local markets. By contrast, Airbnb and Uber's one-sided proximity, capacity-constrained supply, high substitutability, and mobile demand facilitate the propagation of cross-side network effects, leading to increasing returns and scalable growth. This distinction highlights why platforms offering local, time-sensitive, and heterogeneous services behave very differently from those offering standardized, mobile, or inventory-based services when it comes to market growth and spatial diffusion.

Another relevant study by Sun, Zhang, and Wang (2021) reinforces the selection of these two exogenous events. Their research analysed the correlation between the growth of Airbnb listings and the characteristics of the Chinese city of Nanjing. It was found that most Airbnb properties are located within walking distance of the nearest subway or train station. This finding supports the hypothesis that demand is likely to increase near a new train station following the opening of the LGV (high-speed rail) line. On the other hand, their study also demonstrated that Airbnb's business model has led to a reduction in long-term rentals due to the rise of "professional hosts": individuals who manage at least two properties on the platform. Additionally, the influx of tourists into residential areas has disrupted urban equilibrium. Furthermore the income brought by Airbnb properties is much higher than classic long term rents, leading landlords to swap from the old business model to the modern one. This analysis provides further justification for the Bordeaux Council's decision to introduce new regulations aimed at increasing the barriers to entry for new property providers, thereby reducing supply.

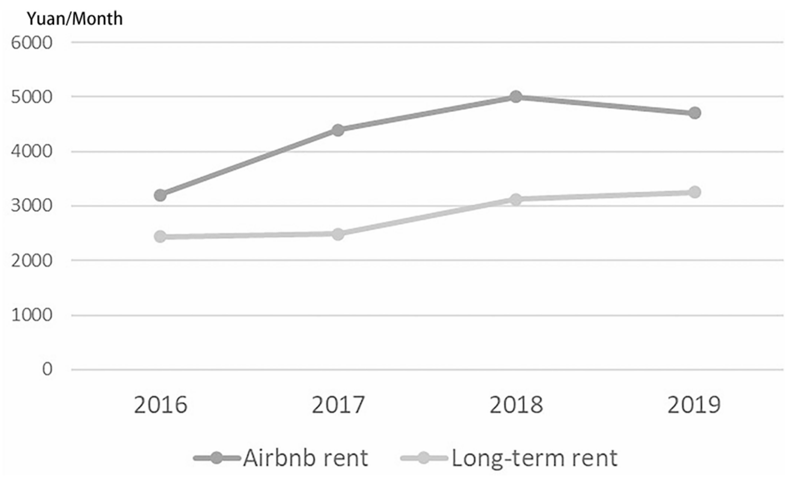

Figure 4, rent gap between short- and long-term in Nanjing. Short-term rent data source: AirDNA, long-term rent data source: China Housing Market.

3. Definition of the Analysis Scope

3.1 Exogenous shocks

As mentioned in the introduction, two main events happened in the region of Nouvelle-Aquitaine, one positively affecting the demand side while the other negatively affecting the supply side.

On the supply side, the region of interest was characterized by a strong regulation on STR activities in the capital city, Bordeaux. On July 10th, 2017, indeed, the council of Bordeaux released a deliberation regarding the declaration and registration of STRs, significantly increasing the entry costs onto the Airbnb platform. The primary motivation for the introduction of such regulation, which was reflected in its design, was to relief the pressure of STRs on the contiguous long term rental market. In fact, as of the entry into force of the legislation, property owners will no longer be allowed to change the use of their home from residential to furnished tourist accommodation (i.e. rented for more than 120 days on an STR platform), unless they comply with specific rules. In particular, the council of Bordeaux opted for a ‘compensation rule’, such that property owners should have to compensate the change in use by purchasing another property of a similar size and in the same area that was until then used for commercial purposes and transform this newly-acquired commercial premise into a long-term habitation (Robertson et al., 2023). The policy became effective in March 2018, and generated some significant effects. Roberson et al. (2023), indeed, find that the share of ‘targeted’ listings on the Airbnb platform, i.e., those rented for more than 120 days per year, significantly declined after the policy enforcement, thereby suggesting that the regulation resulted in a significant (negative) shock for Airbnb.

On the demand side, a second shock characterized the region. On July 2nd, 2017, the first section of the LGV South Europe Atlantique opened. Also known as LGV Sud-Ouest, this is a high-speed rail railway linking Bordeaux and Tours (and consequently Paris in about 2 hours, 1 hour and 10 minutes less than the traditional railway), through intermediate stops in Angouleme and Poitiers (always in the region of New Aquitaine). The preliminary assessment project of the line started in 2001-2003, while starting from 2012 the construction of the railway began. The line started its operations on the July 2nd, 2017, with the route Angouleme-Bordeaux and became fully operative on July 31st, 2017, including the Tours-Angouleme link. The line represents a crucial link in the French transportation organization. As a matter of facts, in 2023 the Paris-Bordeaux LGV line became the second most important high-speed rail line in France, with 95 million passengers passing through the line between July 2017 and October 2023, with summer peaks testifying to the importance of this line for tourism in the New Aquitaine region (LISEA, 2023).

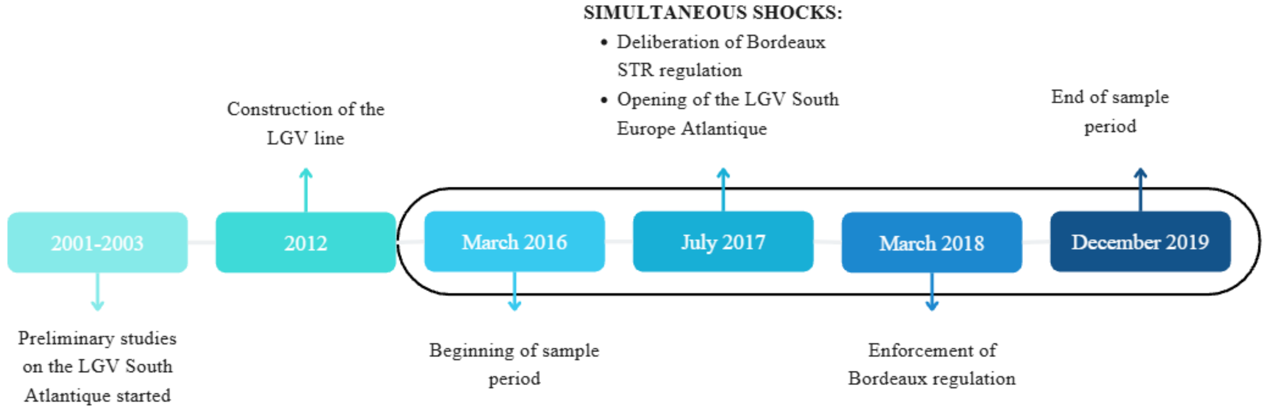

Figure 5, LGV project timeline

3.2 Descriptive analytics

The dataset is sourced from AirDNA, a leading provider of short-term rental market data. Observations are recorded at the IRIS - month level, where an IRIS represents a small area within a municipality, identified by a unique code. To be more precise the dataset represent the classic panel data where entities, in our case IRIS codes, and their specific attributes, are represented over times, in this case at monthly level. As a result, a single municipality may contain multiple IRIS units, allowing for a fine-grained spatial analysis. The dataset includes detailed geographical and temporal information, such as the municipality identifier, reporting month, geographic coordinates, and the distance to the nearest LGV rail station. It also contains comprehensive Airbnb platform metrics, including the number of active listings, total revenues, reserved nights, and related performance indicators.

In addition, the dataset incorporates socio-demographic characteristics at the IRIS level, covering population age distribution, employment structure, and ethnicity. The dataset comprises 313,204 observations, covering 4,060 municipalities and 4,948 IRIS units, and spans the period from March 2016 to December 2021. While the descriptive analysis is conducted on the full sample to provide a comprehensive overview of platform dynamics, the econometric analysis is restricted to observations up to December 2019. This restriction is imposed to avoid potential biases arising from the structural shock induced by the COVID-19 pandemic.

As a first step, descriptive analyses were conducted using Python and Excel to gain a preliminary understanding of the dataset and to inform expectations regarding the subsequent empirical analysis. An examination of the population age distribution reveals that the working-age population represents the dominant demographic group.

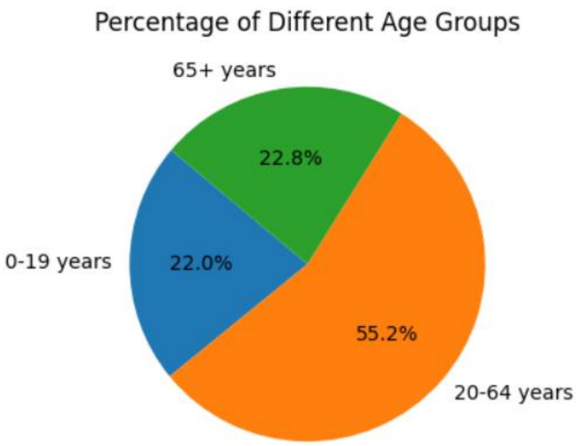

Figure 6, Population's age

When analysing the exogenous demand shock generated by the opening of the LGV stations, it is a good approach to split the dataset into treated and control IRIS. To this end, I computed the median distance to the closest LGV station (74.3 km) and assign a value of 1 (Treated) to IRIS located within this threshold, and 0 to those located beyond it. In addition, I split the dataset over time using a Post dummy, equal to 1 for observations after the LGV opening and 0 for observations prior to the event.

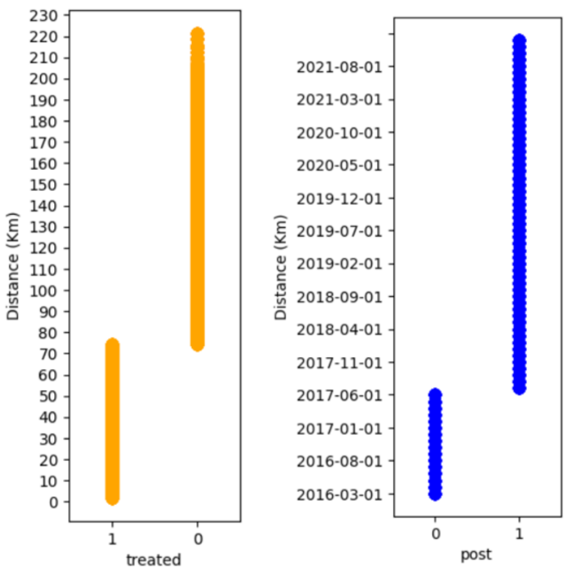

Figure 7, Treated and Post IRIS division

It is informative to examine how the three age groups are distributed between the treated and untreated populations. For the 0-19 age group, approximately 56.3% of individuals belong to treated IRIS, compared to 43.7% in the control group. A similar pattern emerges for the working-age population (20-64), where 55.3% are classified as treated and 44.7% as untreated. These figures indicate that younger and working-age individuals are more likely to reside near the three cities of Bordeaux, Angoulême, and Poitiers, where the exogenous event occurred.

By contrast, the distribution for the 65+ age group is nearly balanced, with 50.2% of individuals in treated areas and 49.8% in control areas. This suggests that older individuals are less spatially concentrated around the LGV-connected cities and therefore less exposed to the demand shock affecting Airbnb activity.

This demographic pattern is consistent with expectations, as the strongest growth in Airbnb demand is observed in urban and well-connected areas, which tend to attract younger populations. Importantly, the treated group is largely composed of students and young professionals, who are more likely to rent housing rather than own it. As short-term rental activity expanded in these cities, this segment of the population was disproportionately exposed to rising long-term rental prices. These dynamics likely contributed to housing market pressures in Bordeaux and help explain the introduction of short-term rental regulations by the city council.

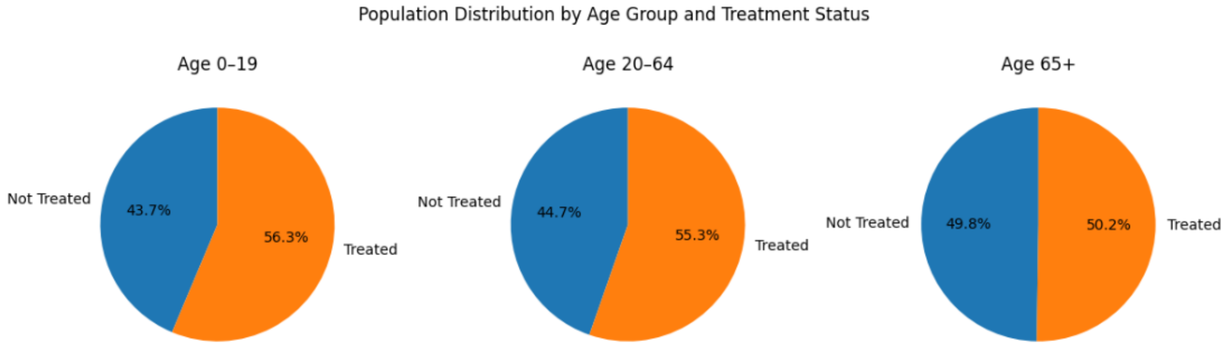

Figure 8, Population distribution between treated and non treated IRIS

Another analysis involved visualizing the distribution of Airbnb registrations across income groups. To this end, five income categories, very low, low, medium, high, and very high, were constructed using a quintile-based approach. The income groups were then placed on the x-axis, and a scatter plot was generated to illustrate the distribution of Airbnb registrations within each income category. As expected, the number of Airbnb listings increases with higher income levels. This pattern suggests that participation on the supply side of the short-term rental market is strongly associated with wealth. Hosting on Airbnb typically requires ownership of a residential asset, such as a house or an apartment, that is not used as a primary residence or that can be withdrawn from the long-term rental market. Higher-income individuals are therefore more likely to possess such assets.

This distributional pattern supports the interpretation that Airbnb supply is not evenly accessible across income groups, but rather concentrated among wealthier households. As a result, lower-income groups face structural barriers to entry on the supply side of the platform, reinforcing the role of income and asset ownership as key determinants of participation in the short-term rental business.

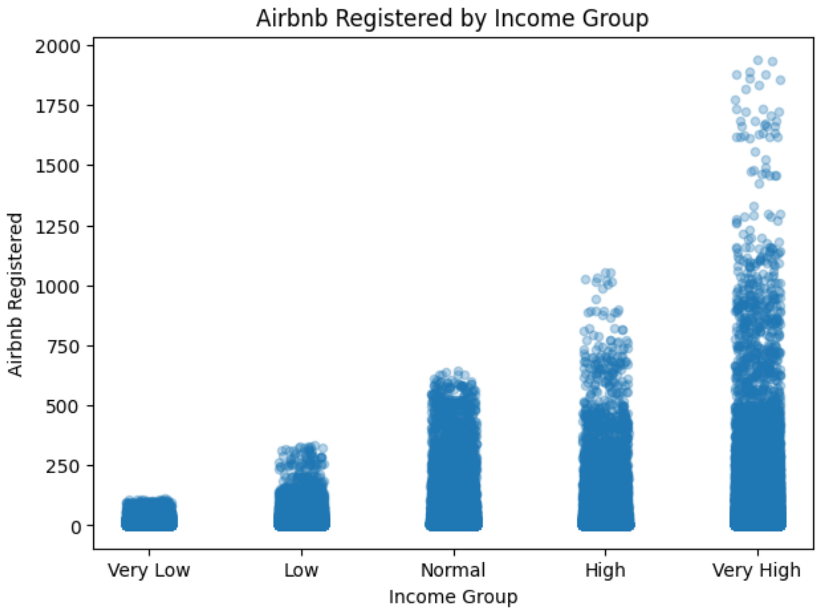

Figure 9, Airbnb registered by income group

The dataset was also analysed using Excel. Figure 9 presents a graph illustrating the year-over-year (YoY) growth in reserved nights across the entire Nouvelle-Aquitaine region over time. For each month, demand is compared with the corresponding month in the previous year, and the resulting percentage change is reported. The figure clearly highlights the impact of the exogenous shock associated with the opening of the LGV high-speed rail line on the demand side of Airbnb. In particular, between March and July 2017, monthly YoY growth rates average around 160%, more than double the growth observed in the following year. This sharp increase indicates a substantial and immediate boost in tourist demand that in part can be explained by the improvement in regional accessibility, but also by Airbnb market expansion within Europe.

By contrast, during 2018 and 2019, demand growth appears more stable and moderate, suggesting a normalization phase in the absence of major exogenous shocks affecting travel demand and a consolidation of the Airbnb market. Finally, the abrupt collapse observed in March 2020 reflects the onset of the COVID-19 pandemic, which generated an unprecedented negative shock to tourism and short-term rental activity. A similar pattern emerges between March and April 2021, when the gradual relaxation of pandemic-related restrictions generated a renewed surge in tourism demand. This sharp increase reflects the recovery of travel activity as mobility constraints were progressively lifted and consumer confidence improved. Once again, the data point to the presence of a positive exogenous shock affecting the demand side of the platform. This evidence further reinforces the idea that spatially localized digital platforms, such as Airbnb, are highly sensitive to social, geographic, and broader human factors. Variations in mobility conditions, infrastructure availability, and public health dynamics translate directly into localized fluctuations in platform demand, despite the digital nature of the matching process.

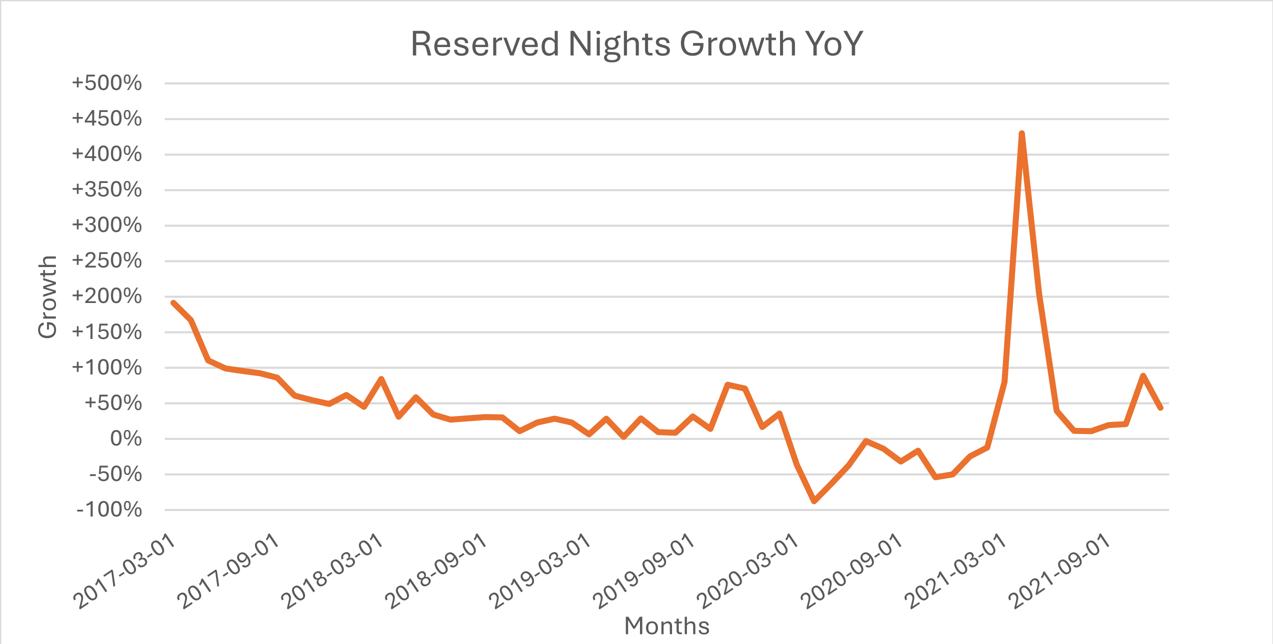

Figure 10, Demand growth YoY

It is also insightful to examine the differences in reserved nights between treated and control IRIS units. As shown in Figure 11, the most attractive locations in terms of reserved nights within the Nouvelle-Aquitaine region are predominantly situated outside the three cities connected by the LGV line. This pattern indicates that areas farther from Bordeaux, Poitiers, and Angoulême tend to record higher numbers of reserved nights.

Importantly, this evidence does not imply that these cities lack tourism appeal. Rather, it can be explained through two complementary mechanisms. First, the introduction of high-speed rail substantially reduced travel times, encouraging shorter stays and day trips in well-connected urban centres. As a result, visitors may travel more frequently but stay for fewer nights. Second, larger cities offer a wider range of alternative accommodation options, such as hotels and hostels, which intensifies competition for Airbnb. The presence of a more developed traditional hospitality sector in urban areas likely diverts some demand away from short-term rental platforms. In contrast, more remote areas, which are harder to reach and offer fewer accommodation substitutes, tend to attract longer stays, leading to higher numbers of reserved nights per booking. Third, cities like Bordeaux can be often visited for work purposes and again leading to shorter stays and opting for traditional accommodation paid by companies.

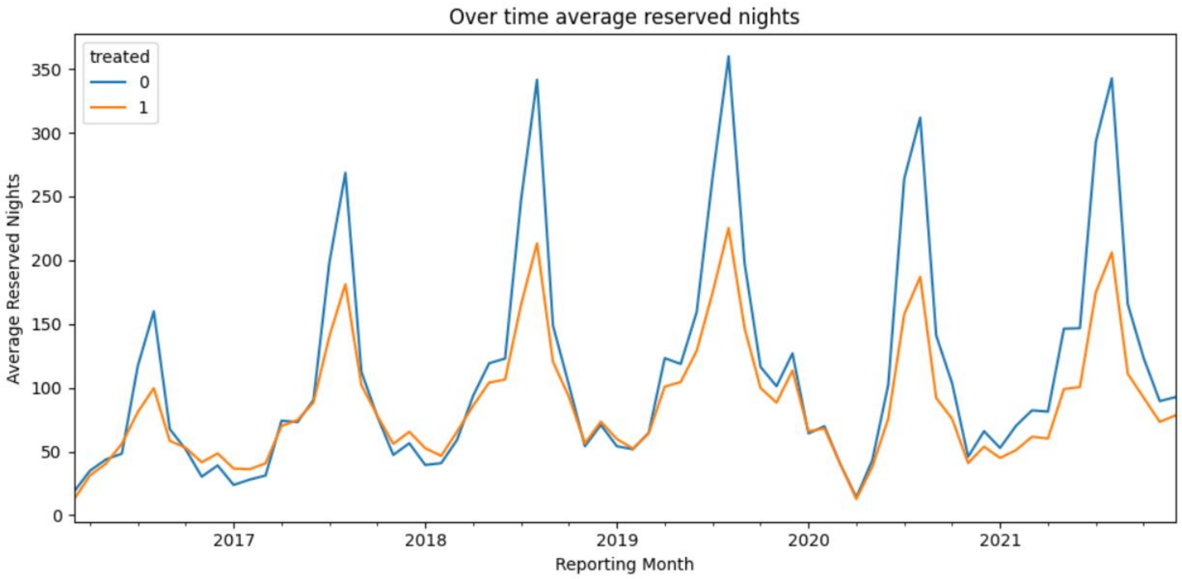

Figure 11, Average reserved nights (Treated / Control)

Following the same approach, a parallel analysis was conducted on the supply side, focusing on the growth over time of registered Airbnb listings. As before, the full dataset was considered, including both treated and control IRIS and all available months.

A first notable observation is that supply dynamics appear more stable and less volatile than demand dynamics. Although supply does respond to exogenous shocks, these reactions are generally more gradual and controlled. In addition, negative shocks seem to have a weaker impact on supply than positive ones. While it is still premature to draw definitive conclusions, the data suggest that the opening of the LGV, by positively affecting demand, also generated a spillover effect on the supply side. In particular, in March 2017, the number of registered Airbnb listings increased by approximately 123% compared to the same month of 2016, providing preliminary evidence of cross-side network externalities operating from demand to supply. On the other hand, the increase of hosts at the beginning of 2017 is more likely to be the effect of market conditions (see housing and loans data next) or Airbnb growth. Indeed through descriptive analytics, we are not able to distinguish whether the effect was generated by one force or the other one.

Conversely, the regulatory intervention appears to have had a more limited effect on supply dynamics. Despite the introduction of stricter regulations, the number of Airbnb listings continued to grow, even though with a negative trend, until the onset of the COVID-19 pandemic, indicating that the regulatory shock was less disruptive than the demand-driven effects associated with the LGV opening.

Finally, the pandemic period offers a striking contrast between demand and supply responses. While demand collapsed sharply, with reserved nights declining by approximately 88%, supply growth largely stagnated rather than reversing. This pattern suggests that, for many property owners, maintaining a listing on Airbnb remained a preferable strategy compared to exiting the platform or shifting to alternative uses. This behaviour is consistent with Airbnb's emergency measures introduced during the pandemic, which were aimed at retaining hosts and stabilizing the supply side of the platform. For instance Airbnb launched their program of open homes helping health workers and caregivers to find an accommodation close to their workplace. [3] These descriptive analyses of demand and supply can be meaningfully related to the "flowers and bees" framework proposed by Stourm and Albuquerque (2024). The results clearly indicate that the demand side (bees) is more sensitive to variation and fluctuation, as consumers are inherently mobile and travel across locations, meaning being subject to external risk such as flight cancelation, local market condition, social emergencies, etc… Consequently, demand reacts more strongly to exogenous shocks, whether positive or negative, leading to pronounced volatility in observed outcomes.

By contrast, the supply side (flowers) is geographically fixed and comparatively stable. Because providers cannot easily relocate or adjust their physical assets in response to short-term shocks, supply exhibits greater resilience to external disturbances. This rigidity is even more pronounced in the context of home-sharing platforms when compared to car-sharing platforms. Housing assets are significantly more expensive, less mobile, and subject to higher fixed and regulatory costs than cars. Moreover, the stock of vacant or underutilized housing units is structurally more limited than the stock of privately owned vehicles, which further constrains the potential expansion of supply in response to demand fluctuations. In addition, platform-level interventions can effectively support host retention with relatively limited effort, even during adverse periods.

The trend also shows that the effects of short-term rental (STR) regulations are less pronounced in the data. While supply growth gradually decelerated over time, it remained positive until the onset of the pandemic, indicating that regulatory constraints brought a more moderate influence on hosting decisions.

Finally, the post-pandemic recovery provides an instructive contrast. In April 2021, demand surged sharply, with reserved nights increasing by approximately 430%, yet supply expanded by only 9%. The result underscores the presence of binding capacity constraints on the supply side and highlights how not all positive demand shocks generate comparable supply responses, particularly in markets characterized by rigid and spatially fixed assets such as housing.

Figure 12, Supply growth YoY

The following visualizations further clarify the reasoning developed thus far for the supply side. As will be shown, the sharp increase in Airbnb registrations observed at the beginning of 2017 can be also partly explained by an additional exogenous shock affecting housing market conditions in France. Beyond the opening of the LGV line, 2017 marked a particularly favourable year for Nouvelle-Aquitaine, and more broadly for France, especially with regard to mortgage financing conditions. The French housing market and lending environment closely followed the interest rate policies of the European Central Bank (ECB). In the aftermath of the 2008 global financial crisis, the euro area experienced a prolonged period of economic distress, with several member states, including Italy and Spain, facing severe financial instability.

To prevent systemic collapse and preserve the integrity of the monetary union, the ECB implemented a series of unconventional monetary policy measures. Following the well-known statement by former ECB President Mario Draghi, in which he pledged to do “whatever it takes” to safeguard the euro, the central bank embarked on an aggressive easing cycle. Policy interest rates were progressively reduced, eventually reaching historically low levels by 2019.

These measures were aimed at fighting deflationary pressures by lowering the cost of borrowing and stimulating investment. Reduced financing costs supported economic activity, contributing to higher employment levels and, in turn, increased aggregate demand. This macroeconomic environment helped sustain rising asset prices, including housing prices, and likely reinforced demand dynamics in the real estate and short-term rental markets. [4,6]

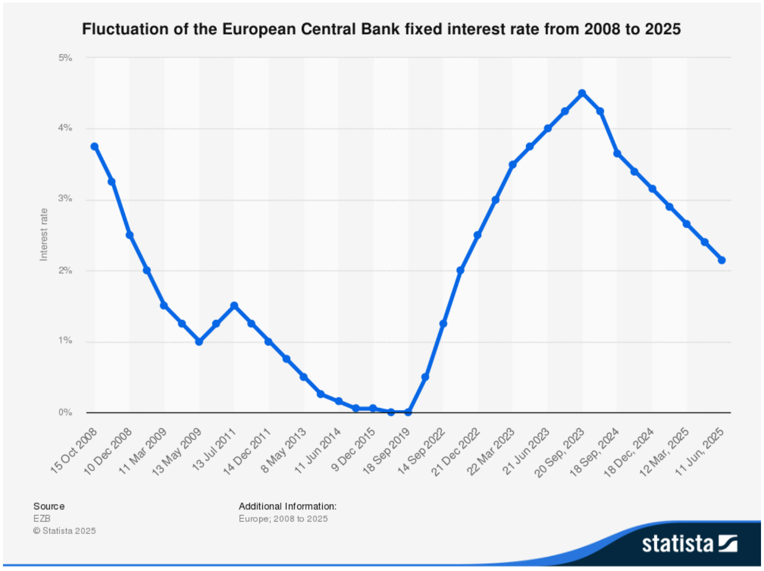

Figure 13, ECB Interest Rate trend

As shown in Figure 14, the impact of the ECB’s monetary policy actions is clearly reflected in the French housing market. Applying the same analytical approach used for the supply and demand growth charts, the data indicate that in 2017 the region of Nouvelle-Aquitaine experienced a substantial increase in both primary and secondary residences. This expansion is consistent with exceptionally favourable lending conditions, which lowered borrowing costs and facilitated housing investment across the region.

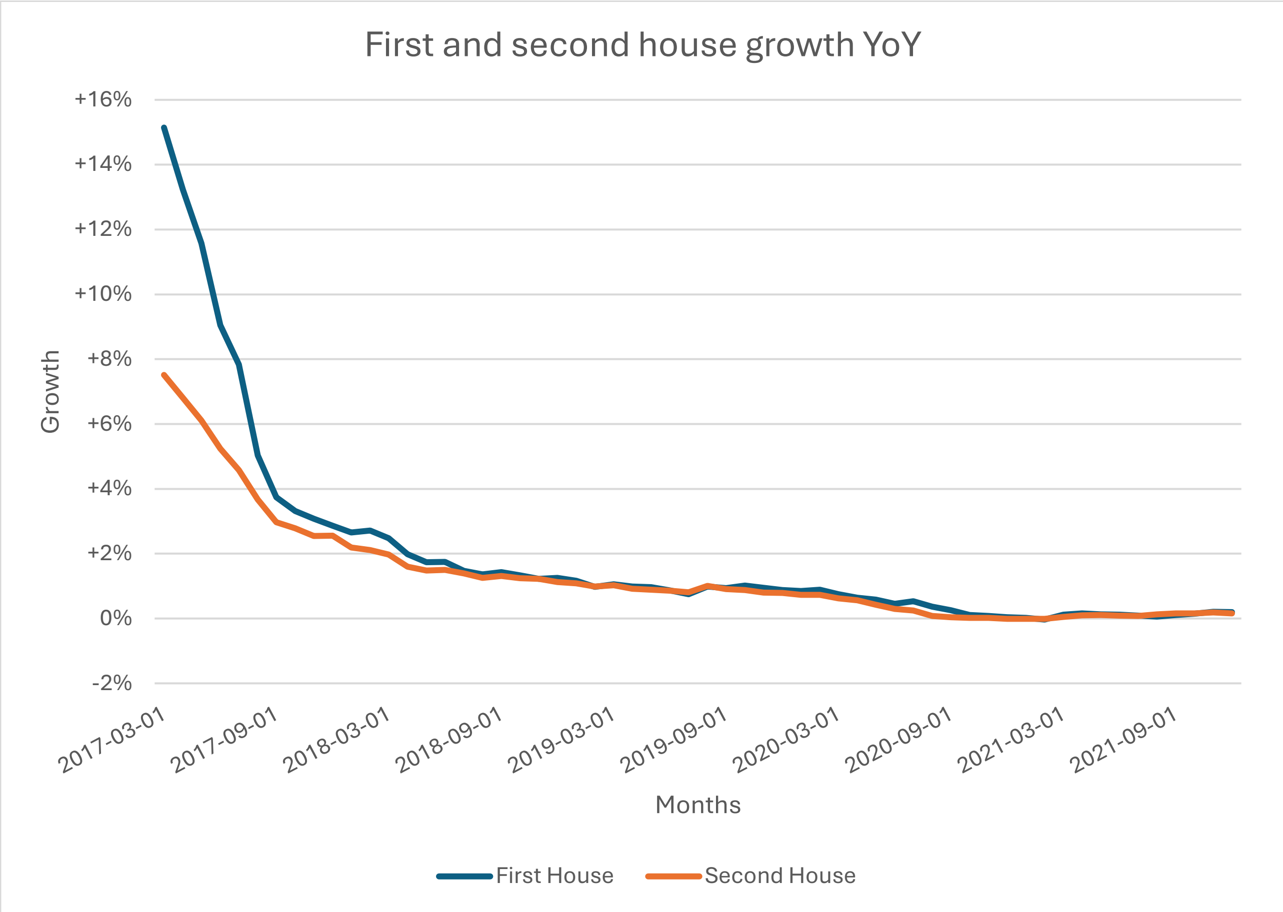

Figure 14, First and second house growth YoY

The figure below illustrates the evolution of home loan interest rates alongside average loan maturities. It is evident that by the end of 2016 and the beginning of 2017, borrowing conditions became exceptionally favourable, characterized by historically low interest rates and extended loan maturities. These conditions translated into lower monthly mortgage payments, significantly improving housing affordability and expanding access to credit, particularly for lower-income households. [5]

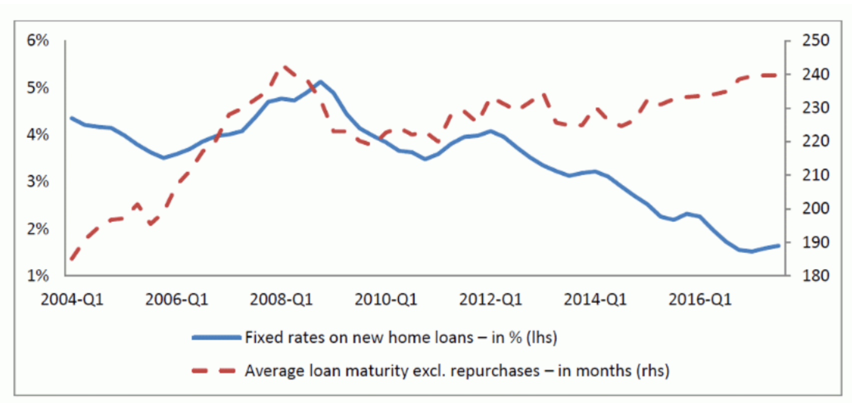

Figure 15, French house loans condition

In conclusion, this chapter analysed the evolution of both Airbnb supply and demand, together with key demographic characteristics of the region. The analysis highlighted the intrinsic differences between the two sides of the market: demand appears significantly more volatile and sensitive to exogenous shocks, while supply remains comparatively stable and fixed. This dynamic reflects the “flower” metaphor introduced by Stourm and Albuquerque. Furthermore, the increase in the number of hosts during 2017 can also be linked to the favourable and unique housing and credit conditions that existed in the period immediately preceding the pandemic. In the following chapter, we further investigate the strength of Airbnb’s cross-network effects (CNE) by isolating the two sides of the market and examining how fluctuations in demand influence the supply side.

4. Data and Methodology Description

The methodology for this particular analysis follows a standard two stage least squares (2SLS) regression. This specific econometric technique uses an instrumental variable to mitigate the impact of endogeneity within the explanatory variables. For our estimation, we must run two different equations: we first need to assess the impact of the instrumental variables on demand (LGV opening). The predicted residuals obtained then serve as control variable in the final stage when considering the two sides of the market as independent and dependent variables.

The first stage is as fundamental as the second, since it isolates demand shock to assess the growth of supply in an unbiased way. Due to time constraints, this thesis will focus on demand to supply network externalities and reserve the opposite side for future work.

To estimate D2S externalities, we run the following regression model in Equation 1:

$$N_{i,t} = \alpha + \beta^{D2S} D_{i,t,k} + \gamma \, SSE_{i,t} + X_{i,t} + \mu + \tau + \varepsilon_{i,t}$$

Equation 1 – Supply Equation

The terms in Eq. 1 are defined as follows:

a) $N_{i,t}$ is the dependent variable. This variable counts the number of new listings joining the Airbnb platform in IRIS *i* at month *t*. As we will see when applying the models, this variable will be tested both on Airbnb registered and bedroom registered in order to test different level of granularity within the supply scope. When included into Eq. 1, the variable $N_{i,t}$ has been log-transformed.

b) $D_{i,t,k}$ is the variable of interest, as β^{D2S} represents the elasticity of prospective Airbnb suppliers' entry decision to demand variations. Due to the presence of temporal lags between the moment of the purchase and the consumption of the service in the touristic sector, the definition of demand has to take into account future prospects of demand at the moment of the entry (*t*) which are typically known in advance. For this reason, I have operationalized our demand variable $D_{i,t,k}$, defined as:

$$D_{i,t,k} = [\sum^{j=t \text{ to } j=t+k} RN_{i,j}] / (k+1)$$

Equation 2 – Definition of $D_{i,t,k}$

where ($RN_{i,j}$) is the count of Reserved Nights in IRIS *i* and month *j*, thereby assuming that the decision to join the Airbnb platform takes into consideration prospective demand up to *k* periods ahead. On the basis of the typical patterns within the platform, we have tested our models employing values of k from 2 to 5. Also this variable has been log-transformed when included into the regression equation.

c) Consistent with the main literature on platforms, I have included in Eq. 1 a specification for modelling same side externalities ($SSE_{i,t}$). These externalities are those generated by individuals acting on the same side. In our case, while SSEs on the demand side can be considered absent, as new joiners on the demand sides will not affect my decision to book on Airbnb, is instead legitimate to expect that the actual level of supply in a given IRIS could influence the decision of prospective suppliers to join the platform and act in a specific local market. Thus, in line with extant literature which highlights the role of spatial concentration of competitors in determining suppliers' entry decision in platforms (e.g. Buzzacchi, Grilli & Milone, 2023) we have included a linear and a quadratic term of the lagged stocks of supply ($Stock_{i,t}$) in an IRIS *i* at time *t-1* to capture the possible presence of non-linear SSEs on the supply side.

d) $X_{i,t}$ accounts for any observable factors, out of the network structure, influencing Airbnb supply. Indeed, I have included several controls at the IRIS level to account for potential sources of endogeneity. First, we control for socio-demographic trends including the log of population and the share of foreign people. Second, we control for local labour market conditions as we included the share of working-active persons and the share of unemployed individuals. Third, we capture the dynamics of the local touristic industry over time with the help of two interaction effects: namely, a linear time trend multiplied by the density of hotel beds to the number of inhabitants (in 2016), to control for the temporal evolution in the local intensity of touristic activity. Finally, we also controlled for the promulgation of the STR regulation implemented in Bordeaux starting from July 2017 through a proper step-type dummy variable.

In conclusion we can identify the variable $X_{i,t}$ as :

$X_{i,t} = logP_{i,t} + For_{i,t} + WA_{i,t} + U_{i,t} + Y_{i,t} + B_{i,t}$

The estimation of β^{D2S} requires further methodological steps to correctly identify the magnitude of D2S externalities. Indeed, although Equation 1 controls for multiple confounders $X_{i,t}$, it is still likely that endogeneity in various forms could still remain in our framework. First, we acknowledge that unobserved IRIS-level characteristics (e.g., the presence of highly attractive touristic amenities) and macro-economic shocks common to all individuals can largely influence our estimates. To account for unobserved heterogeneity across locations and over time, the empirical specification includes both location and time fixed effects. Location fixed effects absorb all time-invariant characteristics at the municipality level, such as intrinsic touristic attractiveness, geographic features, and structural differences across areas. Time fixed effects capture common temporal shocks affecting all locations, including seasonality, macroeconomic conditions, and aggregate trends in Airbnb usage. By controlling for these sources of unobserved heterogeneity, identification relies exclusively on within location variation over time, allowing the estimated coefficients to be interpreted as differential effects net of confounding factors. Second, we acknowledge the potential concern of reverse causality, thereby indicating that without properly accounting for endogeneity of demand, Equation 1 is not able to disentangle whether we are estimating the effect of demand shocks on supply or vice versa.

To deal with this issue we employ an Instrumental Variable (IV) approach, which we detail as follow.

The instrumental variable for $D_{i,t,k}$ should be an exogenous demand shifter, which is not directly correlated with the entries of suppliers onto the platform. In this sense, we leverage on the opening of the LGV South Europe high speed rail and deem it as a candidate instrumental variable in our case. To properly use this instrumental variable I applied a two stage least square approach (2SLS) where I estimated in a first stage the demand variable $D_{i,t,k}$ employing a traditional difference in difference (DID) strategy comparing treated and control IRIS before and after the opening of the high-speed rail. I identified treated IRIS based on the distance to the closest station interested by the opening of the LGV (the three stations are Bordeaux, Angouleme and Poitiers), relying on the rationale that the closer to the station, the higher the magnitude of the treatment effect. Specifically, to ensure a proper balancing between treated and control units, treated units are those IRIS with distance to the closest station below sample median. (See figure 7)

Consistently, our first stage can be summarized as in Equation 3:

$$D_{i,t,k} = \alpha + \rho_1 \, Treat_{i,p} \cdot Post_t + \rho_2 \, Treat_{i,p} \cdot Post_t \cdot (1/Dist_i) + \gamma \, SSE_{i,t} + X_{i,t} + \mu_i + \tau_t + \varepsilon_{i,t}$$

Equation 3 – Instrument for $D_{i,t,k}$

To estimate Equation (3), we begin by preparing the panel dataset at the IRIS-month level. Reporting dates are first checked and converted into a proper datetime format, and observations from January 2020 onwards are excluded in order to avoid distortions related to the COVID-19 shock. The final sample therefore spans the period 2016–2019, which is consistent with the pre-pandemic setting of the analysis.

I then constructed the set of socio-demographic and labour-market controls included in $X_{i,t}$. Specifically, I computed the logarithm of total population to capture scale effects, the share of foreign residents over total population to account for demographic composition, the share of unemployed individuals over the working-age population, and the share of working-active individuals, defined as the proportion of the working-age population that is neither unemployed nor inactive. Observations with non-positive working-age population are excluded from these ratios to avoid invalid values.

To capture time dynamics, I extracted the calendar month from each reporting month (January → 1, February → 2…) and constructed a linear time trend. The time trend is defined as a ranking of reporting months, assigning increasing integer values to each subsequent month starting from the beginning of the sample. This trend captures smooth temporal evolution common to all IRIS. In order to control for heterogeneity in tourism intensity over time, I computed hotel beds per capita for each IRIS and fix baseline exposure at its 2016 level. This baseline measure is then interacted with the linear time trend, allowing IRIS with different levels of touristic infrastructure to follow distinct demand trajectories over time. In other words, It is included to control for the possibility that Airbnb grows faster over time in more tourism intensive areas, so that this trend is not mistakenly attributed to cross-network effects

I also controlled for regulatory effects by constructing a step dummy identifying the introduction of short-term rental regulation in Bordeaux. This dummy takes value one for observations belonging to the municipality of Bordeaux from July 2017 onwards, and zero otherwise, thereby isolating the impact of local regulatory intervention on Airbnb demand.

I captured supply side same-side externalities by exploiting the stock of Airbnb listings already active in each IRIS prior to the realization of demand. Formally, for each IRIS i and month t I constructed the lagged stock of supply as the total number of Airbnb and bedrooms registered in the IRIS at time $t-1$. This variable reflects the pre-existing density of hosts in the local market and proxies the competitive pressure faced by individual suppliers when deciding to operate on the platform.

To allow for potential non-linear effects of competition, I included both a linear and a quadratic term of the lagged supply stock. This flexible specification accommodates different market dynamics: while a higher number of listings may initially enhance the attractiveness of the platform by increasing variety and visibility, beyond a certain threshold further increases in supply may intensify competition and reduce individual hosts' performance. As a robustness check and to reduce the influence of extreme values in highly saturated markets, I also considered an alternative specification based on the logarithmic transformation of the lagged supply stock.

The treatment variables are then defined following the difference-in-differences framework. I constructed a standard average treatment effect term as the interaction between a treated area indicator and a post LGV opening dummy. In addition, we allow for treatment intensity to vary with spatial proximity by interacting the treated indicator and post opening dummy with the inverse distance to the nearest LGV station. This second interaction captures heterogeneous treatment effects within treated IRIS based on their relative accessibility to the high-speed rail network.

The empirical specification includes both municipality fixed effects and time fixed effects. The inclusion of municipality fixed effects implies that the estimation does not compare different locations with one another in terms of their average level of Airbnb demand. Instead, each municipality is compared only with itself over time, considering IRIS belonging to the same municipality as equal. All time-invariant characteristics of each location, such as baseline touristic attractiveness, proximity to the coast, urban structure, or permanent differences in accessibility, are fully absorbed by the fixed effects. As a result, the estimated coefficients are identified exclusively from within-location variation over time.

Time fixed effects play a symmetric role with respect to the temporal dimension. They absorb all shocks and patterns that are common to all IRIS in a given month. This includes seasonality in tourism demand, aggregate trends in the diffusion of Airbnb, macroeconomic conditions, and any other factors that affect all locations simultaneously. Consequently, once time fixed effects are included, the model does not capture whether Airbnb demand increases or decreases in absolute terms after the opening of the LGV. Any general post LGV increase in demand that occurs across the entire region is removed by the time fixed effects.

The joint inclusion of municipality and time fixed effects implies that the estimated coefficients capture relative changes in demand. Specifically, the model identifies whether, after the LGV opening, Airbnb demand in certain locations deviates from the regional average in a given month, compared to the same deviation observed before the shock. In other words, the estimation focuses on differential dynamics across space rather than on aggregate demand movements.

Once all dependent and independent variables were constructed, the empirical analysis was organized as a sequence of progressively richer specifications. At each step, additional controls were introduced in order to assess how the coefficients of interest responded to changes in model specification, until reaching the full specification presented in Equation (3). This stepwise approach allows for a transparent evaluation of robustness and helps identify the contribution of each group of controls.

The baseline specifications are the following:

Model 1 → $D_{i,t,k} = \alpha + \rho_1 \text{ Treat}_{i,p} \times \text{Post}_t + \mu_i + \tau_t + \varepsilon_{i,t}$

Model 2 → $D_{i,t,k} = \alpha + \rho_1 \text{ Treat}_{i,p} \times \text{Post}_t + \rho_2 \text{ Treat}_{i,p} \times \text{Post}_t \times (1/\text{Dist}_i) + \mu_i + \tau_t + \varepsilon_{i,t}$

These initial models capture the post-treatment effect and, respectively, a distance weighted treatment intensity. The two models contain as well time and location fixed effects. Next, I explicitly control for same side externalities (SSE):

Model 3 → $D_{i,t,k} = \alpha + \rho_1 \text{ Treat}_{i,p} \times \text{Post}_t + \gamma \text{ SSE}_{i,t} + \mu_i + \tau_t + \varepsilon_{i,t}$

Model 4 → $D_{i,t,k} = \alpha + \rho_1 \text{ Treat}_{i,p} \times \text{Post}_t + \rho_2 \text{ Treat}_{i,p} \times \text{Post}_t \times (1/\text{Dist}_i) + \gamma \text{ SSE}_{i,t} + \mu_i + \tau_t + \varepsilon_{i,t}$

I then accounted for the regulatory shock affecting Airbnb activity in Bordeaux, which constitutes an important confounding factor:

Model 5 → $D_{i,t,k} = \alpha + \rho_1 \text{ Treat}_{i,p} \times \text{Post}_t + \gamma \text{ SSE}_{i,t} + B_{i,t} + \mu_i + \tau_t + \varepsilon_{i,t}$

Model 6 → $D_{i,t,k} = \alpha + \rho_1 \text{ Treat}_{i,p} \times \text{Post}_t + \rho_2 \text{ Treat}_{i,p} \times \text{Post}_t \times (1/\text{Dist}_i) + \gamma \text{ SSE}_{i,t} + B_{i,t} + \mu_i + \tau_t + \varepsilon_{i,t}$

Subsequently, I introduced controls for local touristic trends, allowing demand dynamics to differ systematically across areas with heterogeneous exposure to tourism:

Model 7 → $D_{i,t,k} = \alpha + \rho_1 \text{ Treat}_{i,p} \times \text{Post}_t + \gamma \text{ SSE}_{i,t} + B_{i,t} + Y_{i,t} + \mu_i + \tau_t + \varepsilon_{i,t}$

Model 8 → $D_{i,t,k} = \alpha + \rho_1 \text{ Treat}_{i,p} \times \text{Post}_t + \rho_2 \text{ Treat}_{i,p} \times \text{Post}_t \times (1/\text{Dist}_i) + \gamma \text{ SSE}_{i,t} + B_{i,t} + Y_{i,t} + \mu_i + \tau_t + \varepsilon_{i,t}$

Finally, starting from Model 8, I increased the specification with demographic controls, capturing differences in population size and composition:

Model 9 → $D_{i,t,k} = \alpha + \rho_1 \text{ Treat}_{i,p} \times \text{Post}_t + \rho_2 \text{ Treat}_{i,p} \times \text{Post}_t \times (1/\text{Dist}_i) + \gamma \text{ SSE}_{i,t} + X_{i,t} + \mu_i + \tau_t + \varepsilon_{i,t}$

5. Results

5.1 Results 1st stage

As expected, progressively enriching the baseline specifications led to improvements in both robustness and statistical significance of the key coefficients. Given the relatively weak performance of the first six models, the empirical results reported in the main text focus on the final three specifications, which provide the most complete and reliable representation of the underlying demand dynamics.

Furthermore the three models were also tested with different value of prospective demand k. Since an host normally consider future demand, for the purpose of the analysis, I decided to run the three models with k = 2,3,4,5. The following results are obtained using Airbnb registered as supply proxy, accounted in this first stage in the SSE term.

Table 1, Model 7 (Airbnb registered as supply proxy) – Exogenous shock impact on Airbnb demand

$$D_{i,t,k} = \alpha + \rho_1\, Treat_{i,p} \times Post_t + \gamma\, SSE_{i,t} + B_{i,t} + Y_{i,t} + \mu_i + \tau_t + \varepsilon_{i,t}$$

MODEL 7 AIRBNB REGISTERD	***k = 2***	***k = 3***	***k = 4***	***k = 5***
Treat_median × Post	0.0590**	0.0627**	0.0614**	0.0604**
(Std.Err.)	(0.0250)	(0.0246)	(0.0243)	(0.0241)
[p-value]	[0.0182]	[0.0109]	[0.0116]	[0.0122]
$\boldsymbol{B_{i,t}}$	YES	YES	YES	YES
$\boldsymbol{Y_{i,t}}$	YES	YES	YES	YES
SSE$_{i,t}$	YES	YES	YES	YES
Observations	109,281	106,457	103,674	100,906
R² (within)	0.189	0.185	0.188	0.203

Table 2, Model 8 (Airbnb registered as supply proxy) – Exogenous shock impact on Airbnb demand

$$D_{i,t,k} = \alpha + \rho_1\, Treat_{i,p} \times Post_t + \rho_2\, Treat_{i,p} \times Post_t \times \left(\frac{1}{Dist_i}\right) + \gamma\, SSE_{i,t} + B_{i,t} + Y_{i,t} + \mu_i + \tau_t + \varepsilon_{i,t}$$

MODEL 8 AIRBNB REGISTERED	***k = 2***	***k = 3***	***k = 4***	***k = 5***
Treat_median × Post	0.0576**	0.0602**	0.0586**	0.0576**
(Std.Err.)	(0.0253)	(0.0250)	(0.0248)	(0.0246)
[p-value]	[0.0232]	[0.0161]	[0.0179]	[0.0191]
treat × post × inv_dist	0.0221	0.0369	0.0422	0.0433
(Std. Err.)	(0.0392)	(0.0395)	(0.0411)	(0.0438)
[p-value]	[0.5735]	[0.3492]	[0.3045]	[0.3226]
$\boldsymbol{B_{i,t}}$	YES	YES	YES	YES
$\boldsymbol{Y_{i,t}}$	YES	YES	YES	YES
SSE$_{i,t}$	YES	YES	YES	YES
Observations	109,281	106,457	103,674	100,906
R² (within)	0.189	0.185	0.188	0.203

Table 3, , Model 9 (Airbnb registered as supply proxy) – Exogenous shock impact on Airbnb demand

$D_{i,t,k} = \alpha + \rho_1\, Treat_{i,p} \times Post_t + \rho_2\, Treat_{i,p} \times Post_t \times (1/Dist_i) + \gamma\, SSE_{i,t} + X_{i,t} + \mu_i + \tau_t + \varepsilon_{i,t}$

MODEL 9 AIRBNB REGISTERED	***k* = 2**	***k* = 3**	***k* = 4**	***k* = 5**
Treat_median × Post	0.0574**	0.0593**	0.0570**	0.0554**
(Std. Err.)	(0.0256)	(0.0253)	(0.0251)	(0.0248)
[p-value]	[0.0250]	[0.0192]	[0.0229]	[0.0257]
treat × post × inv_dist	0.0208	0.0408	0.0494	0.0538
(Std. Err.)	(0.0338)	(0.0343)	(0.0355)	(0.0376)
[p-value]	[0.5371]	[0.2335]	[0.1636]	[0.1524]
$\boldsymbol{X_{i,t}}$	YES	YES	YES	YES
SSE$_{i,t}$	YES	YES	YES	YES
Observations	109,281	106,457	103,674	100,906
R^2 (within)	0.1892	0.185	0.188	0.2035

A first insight that we can underline is the decrease in the number of observations as the parameter k increases. This is due to the logic of the construction of the prospective demand variable $D_{i,t,k}$, which is defined as the average of reserved nights from period t up to $t + k$. As k grows, the computation of $D_{i,t,k}$ requires the availability of future demand information for a longer horizon. Consequently, observations corresponding to the last k periods of the sample cannot be computed and are mechanically dropped from the estimation. This truncation at the end of the sample window leads to a progressive reduction in the number of usable IRIS / month observations as k increases.

Next, all the three models shows a positive and a statistical significant parameter for the average treatment effect, leading to the conclusion that on average IRIS within the treatment group experienced a rise of Airbnb reserved nights around 5 / 6 % higher than the control group. On the other hand when we look at the variable also controlled by the inverse of the distance, we still find a positive value for the parameter but a p-value higher than 5% meaning that the result can not be consider as statistical significant, therefore the effect of the shock can not be distinguished from zero . Furthermore the growth of the value of k brings to the model more significance by reaching 15% in model 9 with k = 5. In order to correctly asses the impact of the distance to the closest station on Airbnb demand growth, I decided to create a new model (model 10) with the main regressor represented by the variable post * inv_distance and eliminating the ATE variable. By doing so we consider all the IRIS and municipality in the region without splitting the dataset between control and treated. We compare pre and post event Airbnb reserved night across the entire region with the variation of the distance to the closest station.

Table 4, Model 10 (Airbnb registered as supply proxy) – Exogenous shock impact on Airbnb demand

$D_{i,t,k} = \alpha + \rho_3\, Post_t \times (1/Dist_i) + \gamma\, SSE_{i,t} + X_{i,t} + \mu_i + \tau_t + \varepsilon_{i,t}$

MODEL 10 AIRBNB REGISTERD	***k* = 2**	***k* = 3**	***k* = 4**	***k* = 5**
post × inv_dist	0.0481	0.0692*	0.0770**	0.0810**
(Std.Err.)	(0.0333)	(0.0354)	(0.0362)	(0.0369)
[p-value]	[0.1496]	[0.0505]	[0.0335]	[0.0281]
Observations	109,281	106,457	103,674	100,906
$\boldsymbol{X_{i,t}}$	YES	YES	YES	YES
$\mathbf{SSE}_{\boldsymbol{i,t}}$	YES	YES	YES	YES
R^2 (within)	0.187	0.184	0.187	0.2017

As we can see starting from k = 3 the result become significant, underlining that IRIS closer to LGV stations face a stronger positive effects on the Airbnb demand. When both the average treatment effect and the distance-weighted treatment intensity are included simultaneously, the two regressors potentially exhibit a high degree of correlation, as they are both activated only in the post-treatment period and within treated municipalities. This may lead to partial multicollinearity, inflating standard errors and reducing the statistical precision of the distance based coefficient. To further prove this hypothesis, I also ran an alternative model (model 11) including only the ATE term, excluding the variable capturing distance to the nearest LGV station. Specifically, the interaction term post × inv_dist was replaced with Treat_median × Post.

The results from this final specification are consistent with the findings discussed above. However, a closer inspection of the estimated coefficients reveals that p-values are further reduced, suggesting improved statistical precision. This provides additional evidence that, when using Airbnb registered as supply proxy, separating the two regressors mitigates multicollinearity concerns and yields more robust and reliable estimates.

Table 5, Model 11 (Airbnb registered as supply proxy) – Exogenous shock impact on Airbnb demand

$D_{i,t,k} = \alpha + \rho_1\, Treat_{i,p} \times Post_t + \gamma\, SSE_{i,t} + X_{i,t} + \mu_i + \tau_t + \varepsilon_{i,t}$

MODEL 11 AIRBNB REGISTERED	***k = 2***	***k = 3***	***k = 4***	***k = 5***
Treat_median × Post	0.0588**	0.0620**	0.0603**	0.0590**
(Std.Err.)	(0.0253)	(0.0249)	(0.0247)	(0.0244)
[p-value]	[0.0200]	[0.0129]	[0.0145]	[0.0158]
Observations	109,281	106,457	103,674	100,906
$\boldsymbol{X_{i,t}}$	YES	YES	YES	YES
SSE$_{i,t}$	YES	YES	YES	YES
R² (within)	0.189	0.185	0.188	0.203

In conclusion, these final models strengthen the thesis that the opening of the Bordeaux-Paris LGV route created a positive touristic flow within the municipalities closer to the operative train stations. Indeed, the positive and significant coefficient of the inverse distance variable (0.081) in Model 10 (with k=5) indicates a clear distance decay effect. Translated into quantitative terms, for a municipality located 1 km from a station (where 1/dist = 1), the LGV opening is associated with an increase of approximately 8.1% in Airbnb reserved nights. This impact follows a non-linear trend as the distance from the station increases: for instance, the estimated effect doubles to 16.2% for accommodations located at 500 meters (1/dist = 2), while it significantly drops to 4.05% at 2 km (1/dist = 0.5) and becomes marginal (approx. 0.8%) at a distance of 10 km (1/dist = 0.1). These examples confirm that the demand shock is highly concentrated around the infrastructure nodes.

In these first five models, the supply of the Airbnb platform was proxied using the total number of listings, defined as Airbnb Registered. However, the results obtained from the first stage regressions using this metric lacked sufficient statistical significance, especially for the coefficient with inverse of the distance when applied with the ATE term, and exhibited low explanatory power, failing to clearly identify the relationship between the exogenous shocks and demand response. This led to a further methodological refinement in which the supply measure was replaced with the total number of rooms, or Bedrooms Registered. This transition is fundamentally justified by the need to capture the actual renting capacity of the market more accurately than a simple count of units. Given that a single listing can represent anything from a small studio to a large multi-room apartment, the number of bedrooms provides a more granular and faithful representation of the effective accommodative capacity offered within each local market.

Furthermore, this choice addresses and explain the significant spatial heterogeneity observed in the distribution of property sizes across the various IRIS areas as shown in figures 16 and 17. In dense urban centres, the supply typically consists of numerous small units, whereas more residential or peripheral areas may feature a lower number of listings that nonetheless provide a substantial number of rooms. By measuring supply through bedrooms, the model correctly accounts for this variance in mass, ensuring that the competitive pressure and the entry rate in areas characterized by larger properties are not underestimated.

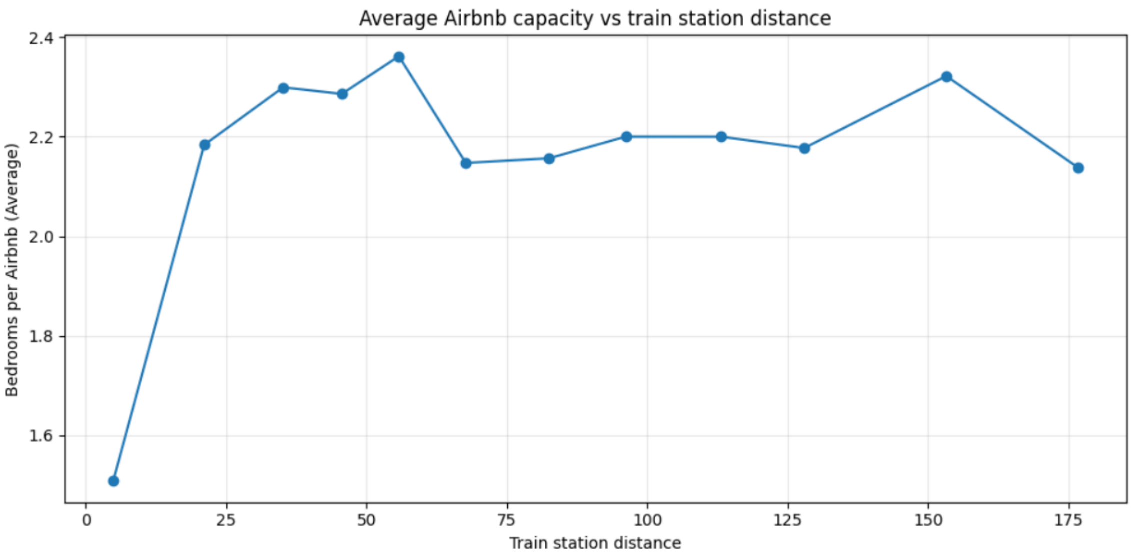

Figure 16, Average Airbnb capacity vs train station distance

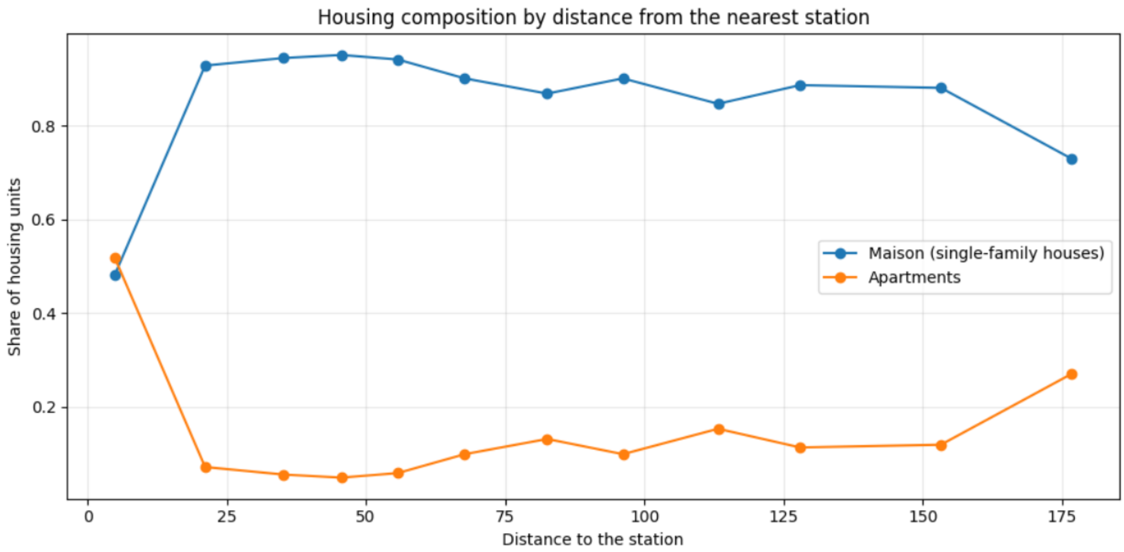

Figure 17, Housing composition by distance from the nearest station

Finally, this approach ensures greater theoretical consistency in both stage of the 2SLS approach. Indeed in the DID the supply is accounted while modelling for SSE. As noted in the relevant literature regarding platform competition, the entry decisions of suppliers are influenced by the concentration of existing competitors in the micro-zone. A quadratic specification using bedrooms captures more effectively the saturation point of the local market, as the marginal value of adding capacity is more closely tied to the available room count than to the number of digital profiles on the platform. Consequently, the use of bedrooms registered provided the necessary statistical clarity and significance, allowing for a more reliable estimation of the supply side response to demand shocks in the final model.

Table 6, Model 9 (Bedrooms registered as supply proxy) – Exogenous shock impact on Airbnb demand

$$D_{i,t,k} = \alpha + \rho_1 \text{Treat}_{i,p} \times \text{Post}_t + \rho_2 \text{Treat}_{i,p} \times \text{Post}_t \times (1/\text{Dist}_i) + \gamma\, \text{SSE}_{i,t} + X_{i,t} + \mu_i + \tau_t + \varepsilon_{i,t}$$

MODEL 9 BEDROOMS REGISTERED	***k* = 2**	***k* = 3**	***k* = 4**	***k* = 5**
Treat_median × Post	0.0565**	0.0587**	0.0567**	0.0554**
(Std. Err.)	(0.0281)	(0.0276)	(0.0272)	(0.0268)
[p-value]	[0.0442]	[0.0336]	[0.0371]	[0.0388]
treat × post × inv_dist	0.1377***	0.1553***	0.1625***	0.1662***
(Std. Err.)	(0.0400)	(0.0407)	(0.0416)	(0.0430)
[p-value]	[0.0006]	[0.0001]	[0.0001]	[0.0001]
$\boldsymbol{X_{i,t}}$	YES	YES	YES	YES
$\mathbf{SSE}_{\boldsymbol{i,t}}$	YES	YES	YES	YES
Observations	109,281	106,457	103,674	100,906
R^2 (within)	0.171	0.170	0.174	0.188

Table 7, Model 10 (Bedrooms registered as supply proxy) – Exogenous shock impact on Airbnb demand

$$D_{i,t,k} = \alpha + \rho_3\, Post_t \times (1/Dist_i) + \gamma\, SSE_{i,t} + X_{i,t} + \mu_i + \tau_t + \varepsilon_{i,t}$$

MODEL 10 BEDROOMS REGISTERD	***k* = 2**	***k* = 3**	***k* = 4**	***k* = 5**
post × inv_dist	0.1648***	0.1837***	0.1903***	0.1937***
(Std.Err.)	(0.0419)	(0.0441)	(0.0447)	(0.0451)
[p-value]	[0.0001]	[0.0000]	[0.0000]	[0.0000]
Observations	109,281	106,457	103,674	100,906
$\boldsymbol{X_{i,t}}$	YES	YES	YES	YES
$\mathbf{SSE}_{\boldsymbol{i,t}}$	YES	YES	YES	YES
R^2 (within)	0.169	0.167	0.172	0.1851

For the sake of simplicity, only models 9 and 10 are reported here, as they incorporate the most comprehensive set of regressors and control variables; the remaining specifications are documented in the Appendix. These results validate the decision to utilize bedrooms as the supply proxy rather than general Airbnb listings. Specifically, while the Average Treatment Effect (ATE) remains stable and consistent with the models accounting for Airbnb registered, the variable including the inverse of the distance to the station finally achieves the expected statistical significance and theoretical role within the model. With a significance level below one percent, the results confirm that proximity to the LGV station is a primary driver of the shock, with the impact on Airbnb demand intensifying as the distance to the infrastructure decreases.

Finally it is important to recognize the two assumptions on which the instrumental variable relies on. First, the validity of our instrumental variable would be compromised if the allocation of treatment were not exogenous with respect to Airbnb activity, which is the case if policymakers had decided to build and route the LGV through Bordeaux, Angoulême, and Poitiers precisely because Airbnb supply or tourism related activity was already stronger in these areas than elsewhere. However, there are strong reasons to rule out this concern. On the one hand, the preliminary feasibility studies and planning phase for the LGV line took place between 2001 and 2006, a period during which Airbnb did not yet exist. On the other hand, construction began in 2012, when Airbnb activity in France, and more broadly in Europe, was still very limited relative to its subsequent expansion. Taken together, these elements allow us to reasonably and conservatively assume that the decision to open the LGV line was made independently of any local dynamics related to Airbnb.

Second, and more importantly, the instrumental variable would be invalid if Airbnb hosts anticipated the opening of the LGV and strategically entered the platform in advance, based on rational expectations of a future increase in demand. In such a scenario, early entry could confer competitive advantages at the time the demand shock materializes, most notably through reputation building or improved visibility on the platform, see Rossi, 2023. Anticipatory entry of this kind would generate endogeneity concerns, as it could affect pre-treatment supply through same side or cross-side network externalities, thereby violating the exclusion restriction.

To address this issue, figure 17 presents a comparison analysis between controlled and treated IRIS of monthly average Airbnb entry rates at the IRIS level over the pre-treatment period, spanning from March 2016, the first month available in our data, to June 2017, the month immediately preceding the LGV opening. We show that, conditional on observable characteristics, treated IRIS do not exhibit significantly higher entry rates than non-treated ones. The absence of differential pre-treatment entry patterns supports the validity of our instrument and suggests that anticipatory behaviour by hosts is unlikely. More broadly, this finding is consistent with the high degree of supply flexibility characterizing Airbnb markets, Farronato and Fradkin, 2022, implying that hosts are more likely to enter in response to realized, and observable, demand shocks rather than on the basis of long-term anticipatory strategies.

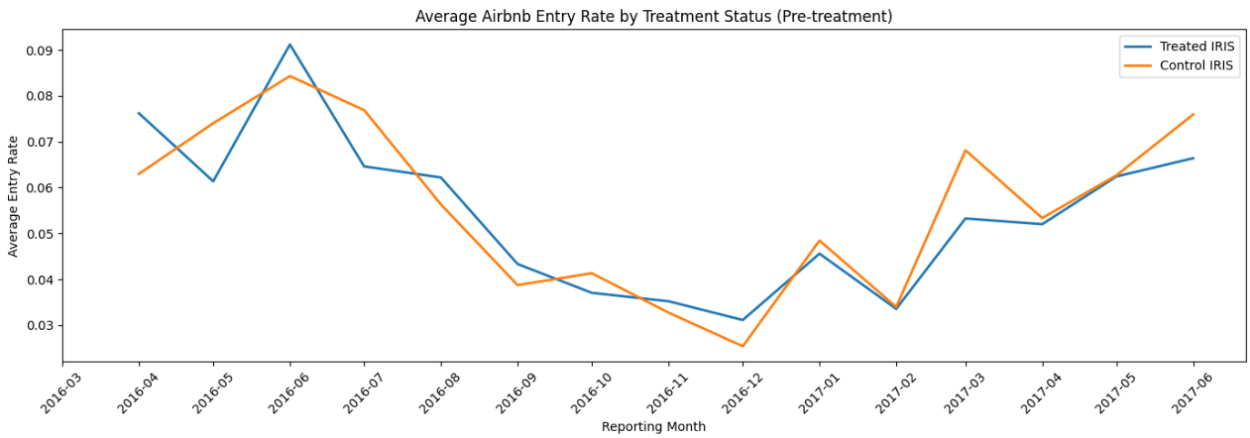

Figure 18, Average Airbnb Entry Rate by Treatment Status (Pre-treatment)

5.2 Results 2nd stage

Now that we have obtained the results of the first stage of the 2SLS approach, we can proceed by estimating the overall supply equation (Equation 1) with Airbnb demand as the main regressor:

$N_{i,t} = \alpha + \beta^{D2S} D_{i,t,k} + \gamma\ SSE_{i,t} + X_{i,t} + \mu + \tau + \varepsilon_{i,t}$

Equation 1 – Supply Equation

First of all, we test the equation through a simple OLS (Ordinary Least Squares) model to verify the impact of demand (independent variable) on supply (dependent variable) without isolating the two sides of the Airbnb market. By doing this, we accept the risk of reverse causality, as both sides are simultaneously affecting each other. Indeed, without accounting for the LGV opening shock as an exogenous instrument, the model is unable to distinguish whether demand is driving supply or vice versa. This leads to an attenuation bias caused by endogeneity between the two variables, which results in an underestimation of the true demand coefficient.

The following tables present the initial Ordinary Least Squares (OLS) estimations, comparing the results using Airbnb Listings versus Bedrooms Registered as proxies for supply.

Table 8, Estimation of D2S CNE through OLS

ESTIMATION OF CNE (OLS)	*k = 2*	*k = 3*	*k = 4*	*k = 5*
D2S CNE (ln Di,t,k)	0.0226***	0.0265***	0.0291***	0.0316***
AIRBNB REGISTERED	(0.0006)	(0.0007)	(0.0007)	(0.0008)
	[0.0000]	[0.0000]	[0.0000]	[0.0000]
D2S CNE (ln Di,t,k)	0.0238***	0.0280***	0.0310***	0.0339***
BEDROOMS REGISTERED	(0.0008)	(0.0009)	(0.0009)	(0.0010)
	[0.0000]	[0.0000]	[0.0000]	[0.0000]
Observations	109,281	106,457	103,674	100,906
$X_{i,t}$	YES	YES	YES	YES
$SSE_{i,t}$	YES	YES	YES	YES
R^2 (within) - AIRBNB	0.930	0.928	0.927	0.926
R^2 (within) - BEDROOMS	0.914	0.912	0.911	0.910

Since both the independent and dependent variables are log-transformed, the coefficients represent the elasticity of supply relative to demand. Specifically, these results suggest that a 1% increase in demand translates to a marginal increase of approximately 0.02% to 0.03% in supply.

As previously discussed, these estimates are significantly lower than expectations due to reverse causality bias. Because the OLS model cannot distinguish the direction of the relationship between demand and supply, the coefficients suffer from attenuation bias caused by endogeneity. Notably, at this stage of the analysis, there is a negligible difference between using Bedrooms Registered or Airbnb Listings as the supply proxy. Both metrics confirm that the OLS approach fails to capture the full magnitude of the supply side response, necessitating the transition to a control function approach to integrate the effect of the exogenous shock.

The following tables present the second-stage results of the 2SLS estimation using the control function approach. Following the methodology of Model 9, the most comprehensive specification regarding control variables and regressors, the residuals from the first stage demand equation were included as an additional regressor to account for endogeneity. Results are presented for both Airbnb Listings and Bedrooms Registered across different k values.

Table 9, Estimation of D2S CNE through 2SLS, using residuals from model 9

ESTIMATION OF CNE (2SLS) M9	*k = 2*	*k = 3*	*k = 4*	*k = 5*
D2S CNE (ln Di,t,k)	0.0340	0.0375	0.0403	0.0430
AIRBNB REGISTERED	(0.0354)	(0.0352)	(0.0368)	(0.0374)
	[0.3375]	[0.2858]	[0.2731]	[0.2501]
D2S CNE (ln Di,t,k)	0.0756*	0.0720**	0.0742**	0.0760**
BEDROOMS REGISTERED	(0.0390)	(0.0365)	(0.0365)	(0.0363)
	[0.0527]	[0.0485]	[0.0418]	[0.0363]
Observations	109,281	106,457	103,674	100,906
Control function	LINEAR	LINEAR	LINEAR	LINEAR
$\boldsymbol{X_{i,t}}$	YES	YES	YES	YES
$\mathbf{SSE}_{\boldsymbol{i,t}}$	YES	YES	YES	YES
R^2 (within) - AIRBNB	0.929	0.928	0.926	0.925
R^2 (within) - BEDROOMS	0.907	0.908	0.906	0.906

The transition to bedrooms registered as the primary proxy for supply led a significant improvement in the model's performance and theoretical consistency. When using simple Airbnb listings, the coefficients for demand lack statistical significance across all time horizons, making the results inconclusive for identifying the supply side response. In contrast, the model utilizing bedrooms achieves statistical significance (at the 5% and 10% levels) for all k coefficients. Notably, the estimated elasticity of supply relative to demand, more than doubles compared to the initial OLS estimations, rising to approximately 0.075 % corresponding to a 1 per cent increase in demand. This confirms that once the endogeneity is addressed through the control function, measuring supply in terms of actual physical capacity (rooms) provides a much more accurate and significant estimation of how the market reacts to demand shocks than the mere count of digital listings.

As a final verification, the residuals from Model 10 were applied to the second stage of the modelling. This specification represents the most rigorous test of our hypothesis, as it confirms that utilizing bedrooms registered as a supply proxy significantly enhances both the statistical significance and the estimated magnitude of the demand impact on supply.

Table 10, Estimation of D2S CNE through 2SLS, using residuals from model 10

ESTIMATION OF CNE (2SLS) M10	*k = 2*	*k = 3*	*k = 4*	*k = 5*
D2S CNE (ln Di,t,k)	0.2551*	0.1804*	0.1637*	0.1578*
AIRBNB REGISTERED	(0.1404)	(0.0999)	(0.0902)	(0.0846)
	[0.0691]	[0.0708]	[0.0694]	[0.0622]
D2S CNE (ln Di,t,k)	0.1399***	0.1269***	0.1243***	0.1247***
BEDROOMS REGISTERED	(0.0350)	(0.0325)	(0.0326)	(0.0315)
	[0.0001]	[0.0001]	[0.0001]	[0.0001]
Observations	109,281	106,457	103,674	100,906
Control function	LINEAR	LINEAR	LINEAR	LINEAR
$\boldsymbol{X_{i,t}}$	YES	YES	YES	YES
$\mathbf{SSE}_{\boldsymbol{i,t}}$	YES	YES	YES	YES
R^2 (within) - AIRBNB	0.7673	0.8618	0.880	0.892
R^2 (within) - BEDROOMS	0.883	0.892	0.894	0.896

The results from Model 10 definitively demonstrate the superiority of the bedrooms registered proxy. With a statistical significance at the 1% level, the estimated elasticity reached a robust value of approximately 0.14. This indicates that a 1% increase in demand leads to a 0.14 % increase in room based supply.

In contrast, the results obtained using the Airbnb registered proxy continue to exhibit weaker statistical significance, consistently falling outside the 5% threshold. Although this model displays a similar pattern, with elasticity values averaging around 0.18, the lack of precision makes these estimates less reliable for causal inference. The transition from the initial OLS elasticity (0.03) to the 2SLS elasticity of 0.14 in model 10 and 0.076 in model 9 confirms that addressing endogeneity via the control function reveals a much more substantial and significant market response than previously observed.

This analysis underscores the critical role of instrumental variables when assessing cross-network externalities (CNE). As evidenced by the results, omitting the IV approach leads to biased estimates that substantially underestimate the strength of these externalities. Once the endogeneity was properly addressed, the elasticity of supply in response to demand increased sharply, validating the importance of isolating one side of the market when investigating network effects. However, even with this correction, it remains evident that while the impact of demand to supply CNE is positive, its magnitude is structurally constrained. The calculated elasticity of approximately 0.14 suggests that even significant growth in touristic flows creates only a limited influence on the total volume of supply.

This empirical observation finds strong theoretical support in the "flowers and bees" framework established by Stourm and Albuquerque (2024). In their research on car sharing platforms, they demonstrate that spatially localized digital marketplaces are fundamentally supply driven, a reality that dictates decision making during the early stages of platform growth. Based on this intrinsic characteristics, decision maker are more prone to invest and subsidies supply and providers, as they will generate a stronger spill over effect on the demand side. This was actually confirmed during the history of Airbnb when the founders, like Brian Chesky, were obsessively helping and demonstrating interest towards host by also showing themselves door to door. [7] Within this paradigm, providers function as the "flowers" entities that are, by their very nature, fixed and stable. Unlike the "bees" (the demand side) who move fluidly, the emergence of new supply is restricted by significant barriers. In the specific context of Airbnb, becoming a host requires a substantial commitment of capital, time, and personal energy. This means that even a disruptive exogenous shock that triggers a massive increase in tourism will not generate an equivalent or immediate reaction on the supply side as also shown in figures 10 and 12 with the Covid demand explosion. The Airbnb host community can be compared to flora that grows in extreme conditions, such as high altitude mountains or colder climates. These environmental challenges represent the high entry costs and effort previously discussed. In this light, an exogenous shock like the introduction of the LGV serves as a unique and rare sunny and warm period. While such an event provides a more hospitable environment for growth, it does not fundamentally change the nature of the flowers that due to the normal harsh condition are limited in expansion. Instead, it allows the existing supply to slowly expand and improve in both quality and quantity over time, rather than forcing an instantaneous or unnatural transformation of the market landscape.

These findings directly inform the decision making processes of platform leaders, particularly regarding the allocation of marketing budgets and subsidies. When financing growth initiatives, leaders must align their targets with the underlying drivers of the ecosystem. In the case of Airbnb, where supply acts as the primary engine for growth, marketing campaigns and financial incentives are most effective when targeted at hosts. Conversely, as the empirical results demonstrate, focusing exclusively on the demand side yields a much lower impact on supply expansion and, by extension, slow down the overall growth trajectory of the platform.

However, the nature of Airbnb necessitates a tailored approach that accounts for sub-market variability. Because Airbnb is a spatially localized digital platform, even within a single region, individual municipalities may exhibit vastly different behaviours and growth patterns. A strategy that succeeds in an urban centre may fail in a rural zone. Consequently, it is fundamental for leaders to move beyond broad regional averages and instead analyse specific local market conditions and demographic data. By adapting the decision making process to these localized realities, platforms can ensure that their interventions are tailored to the unique economic and environmental constraints of each specific zone. To optimize the decision making process, platform leaders must distinguish between several nuanced strategies that account for regional heterogeneity. While Airbnb is generally supply driven, a deeper analysis might reveal that different zones possess distinct DNA. Some areas may exhibit exceptionally strong demand to supply (D2S) cross-network externalities, becoming demand driven hotspots. Conversely, other locations remain infrastructure driven, where the platform's growth is dictated almost entirely by the presence and stability of the hosts. This variance means that even within a single region, the two forces might be equally strong in one municipality facilitating growth, while the latest remains entirely stagnant in another region, where both supply and demand struggle to attract either side of the market despite intervention.

Faced with this diversity, leadership can adopt two primary strategic paths. The first approach is to concentrate time, capital and energy into "fertile" markets where growth is virtually guaranteed by the intrinsic characteristics of the location. In these areas, while supply remains the guiding force, also the demand side produces a powerful spillover effect that amplifies every dollar spent. This strategy focuses on maximizing efficiency by fuelling engines that are already primed to run.

The second strategy involves a more defensive, balancing act. Leaders may choose to disproportionately subsidize "adverse" markets where growth is stuck and cross-network externalities are difficult to ignite. These zones require much higher levels of investment to overcome high entry barriers and reach a critical mass. While this path is more resource intensive, it is essential for platform leaders to reach geographic equilibrium.

6. Local Heterogeneity of Cross-Side Network Externalities

Markets and platforms characterized by cross-side network externalities (CNE) are often regarded as unique global marketplaces where the two sides interact regardless of geographical location. This is the case, for instance, as discussed in our literature review, of C2C e-commerce platforms (e.g., Taobao; Chu and Manchanda, 2016) where producers and consumers can easily interact regardless of their physical proximity. Indeed, Stallkamp and Schotter (2021) define them as "platforms without borders", highlighting the absence of geographical constraints in such markets. Conversely, the market for Short-Term Rentals (STRs) of Airbnb is characterized by local network effects. A guest's utility does not derive from the global size of the platform, but rather from the available supply in the specific neighbourhood or local area where they intend to stay. This condition is closely aligned with the context analysed by Cullen and Farronato (2021) in their study of the digital marketplace TaskRabbit, where user value depends strictly on the local supply of tasks. This spatial dependency suggests the presence of locally heterogeneous cross-side network effects: in certain areas, the supply of accommodations might be highly elastic to fluctuations in tourist demand (demand-driven sub-markets), while in others, the growth of users on the demand side can be highly connected to host variation (hosts-driven sub-markets).

Following these premises, the objective of this section is to estimate the local heterogeneity of network externalities using a Causal Forest DML approach. By mapping the varying strength of the response of supply (registered bedrooms) to demand fluctuations (reserved nights), this analysis aims to provide granular insights and practical implications for targeted urban policy and regional strategies.

6.1 Causal Forest

A causal forest is a machine learning method used to estimate heterogeneous causal effects. In simple terms, it helps us understand not only whether a treatment has an effect on average, but also where the effect has the highest impact and what are the drivers of the latest.

An important aspect in evaluating the heterogeneity of CNE is the shift from the Average Treatment Effect (ATE), which captures the average impact of CNE across the entire region, to the Conditional Average Treatment Effect (CATE). The CATE provides a more granular perspective by estimating the expected demand to supply elasticity for each IRIS based on its specific characteristics. This approach allows the strength of the CNE to vary across local contexts and makes it possible to identify the main drivers of this heterogeneity. In the causal forest framework, this is assessed through feature importance, which measures the power to which each covariate contributes to explaining variation in the estimated effects. Variables with higher importance are those that more frequently generate splits associated with substantial differences in CATE, indicating that the effect of the CNE varies significantly across their values. Consequently, feature importance highlights the local characteristics that most strongly shape the spatial heterogeneity.

In our model, the outcome variable is the log of bedrooms registered, used as supply proxy (log_AR), while the treatment variable is the log of expected demand measured through reserved nights (log_D_k). The model was tested with two months of perspective demands, translated with the value of k equal to two. The estimated effect therefore captures how changes in local demand influence the entry of new hosts, that is, the demand to supply CNE. Instead of assuming that this effect is the same across all IRIS, we allow it to vary depending on local characteristics such as population density, distance from the closest LGV station, lagged stock of supply, share of apartments, share of foreign residents, unemployment rate, income, hotel beds per capita, and other indicators of local housing and tourism structure. We estimate these heterogeneous effects using a Causal Forest approach. The method builds many decision trees (1000 in our case) and divides IRIS into groups with similar characteristics. The splits are chosen to maximize differences in the estimated treatment effect, not differences in the outcome itself. This approach allows us to identify which IRIS are more elastic, where supply strongly reacts to demand, and which are more rigid, thereby revealing the spatial heterogeneity of CNE. Finally, since the causal forest produces one CNE estimate per IRIS - month observation, these estimates were aggregated at the IRIS level by averaging them across months, thereby obtaining a single measure of network effect intensity for each IRIS.

6.2 Results

Examining the distribution of the estimated local elasticities across IRIS (Figure 19), the average effect is 0.0192, which is very close to the baseline OLS estimate of 0.0238. This similarity suggests that both causal forest and OLS regression when deployed, reach very close results, supporting the robustness of the two models. The distribution of estimated effects is entirely positive, indicating that the rise of demand is systematically associated with an expansion of supply across neighbourhoods. This pattern is consistent with the economic intuition underlying platform markets: negative cross-side effects would imply that an increase in demand induces hosts to exit the platform, which would contradict standard supply incentives in the short-term rental market of Airbnb.

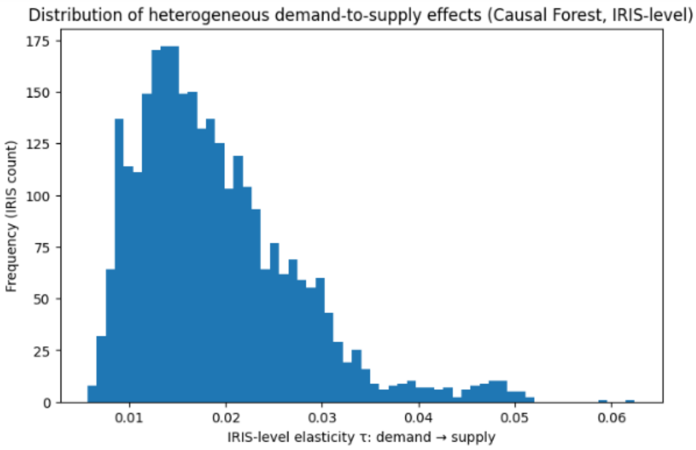

Figure 19, distribution of heterogeneous demand to supply effect

The model created, used as covariates the control variables accounted in the previous regression models, plus other geographic and demographic characteristics of each IRIS within the dataset. As we can see from figure 20 the main drivers of local heterogeneity of CNE are the population density, followed by percentage of apartment as accommodation type and third the number of hotels beds per capita.

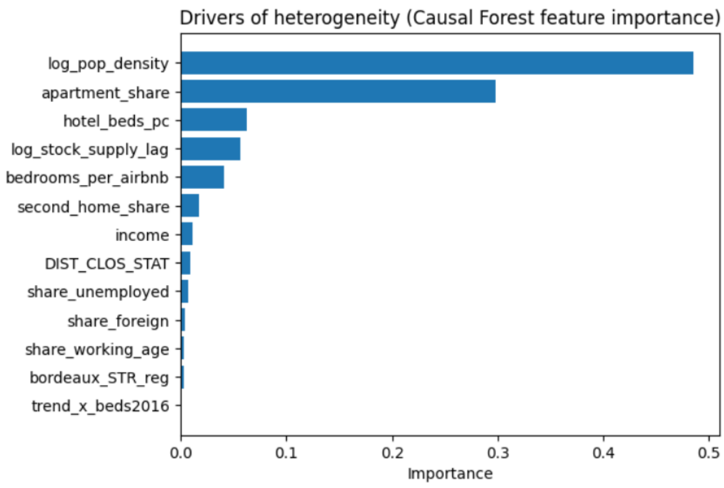

Figure 20, Feature importance

In this context, the causal forest can serve two main purposes. First, it can be used to identify the key features driving the heterogeneity of cross-side network externalities. Second, it can be employed to classify each IRIS according to whether its local platform dynamics are predominantly supply-driven or demand-driven, based on its structural characteristics. Since this thesis focuses on estimating CNEs on only one side of the market, the empirical analysis will concentrate on the first function, uncovering the determinants of heterogeneity in cross-side network effects. Based on the feature importance measures derived from the causal forest, I selected the variables that both exhibit the most informative patterns and contribute most significantly to the heterogeneity of the estimated cross-side network effects.

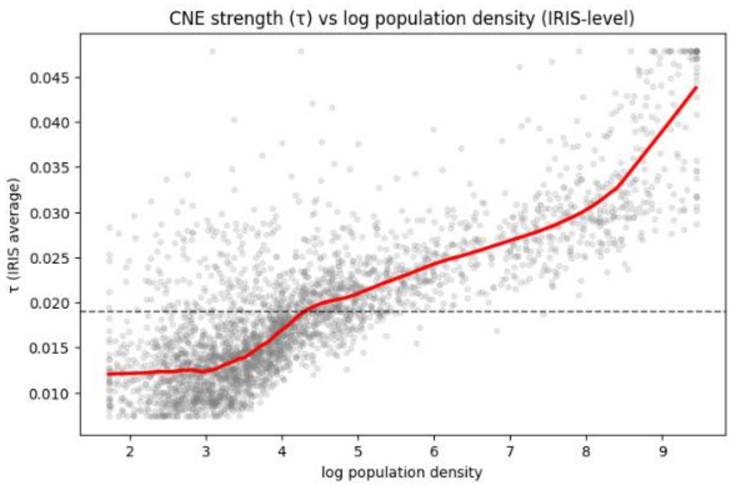

Figure 21, CNE strength vs population density

As previously mentioned, population density represents the main driver of heterogeneity within the dataset. Since this variable was not directly available in the original data, I constructed it prior to estimation.

Specifically, I merged the original dataset with the official IRIS shapefile provided by the GeoServices portal of the "Institut national de l'information géographique et forestière (IGN)", which contains detailed geographical information for each IRIS, including land surface area. By integrating these geospatial data, I was able to compute population density as the ratio of total population to the corresponding IRIS land area. The results clearly highlight the central role of population density in shaping the heterogeneity of CNE elasticity. As population density increases, the magnitude of cross-side network externalities rises, reaching values of approximately 0.03 in the most densely populated areas.

This pattern suggests that in highly dense urban contexts, an increase in demand generates substantially stronger spillover effects on the opposite market side compared to sparsely populated areas. In other words, the responsiveness of supply to demand shocks is amplified in densely populated markets, where economic interactions and matching opportunities are more concentrated.

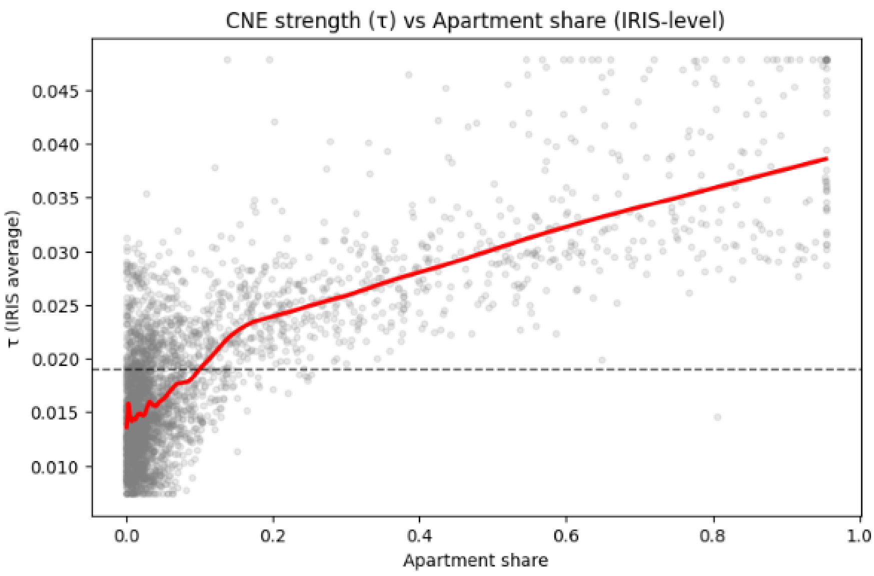

Figure 22, CNE strength vs apartment share

Alongside population density, the share of apartments also exhibits a meaningful pattern in the model. This variable measures the proportion of dwellings classified as apartments relative to the total housing stock within each IRIS. Consistent with the previous findings, apartment share reinforces the urban-density mechanism. Apartments are typically the predominant housing type in densely populated urban areas like bigger cities, as also shown in figure 16, and the results show that CNE elasticity increases as the share of apartments rises. In other words, IRIS characterized by a higher concentration of apartment units display stronger cross-side network effects. On the other hand, the data suggests that, between two densely populated zones, the stronger CNE effect will be registered in the one with the higher share of apartments.

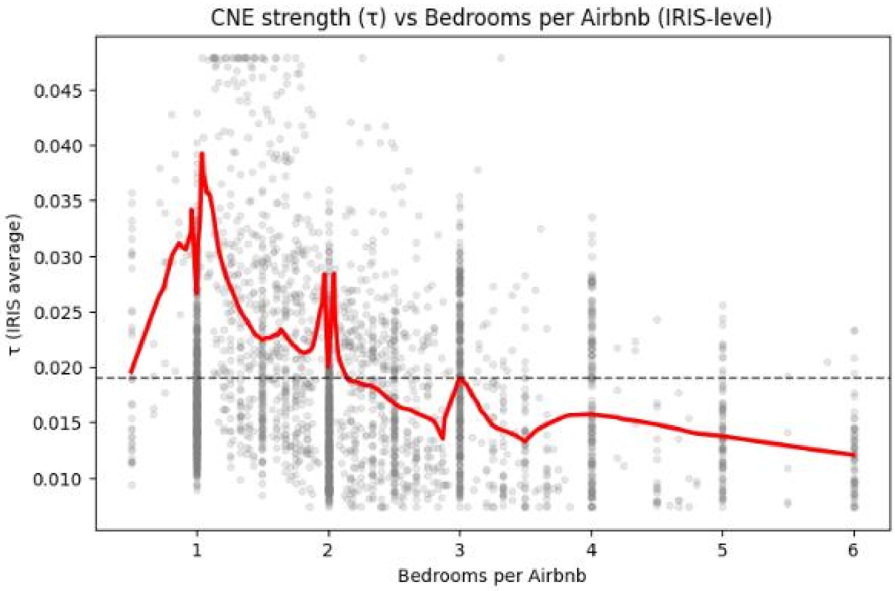

Figure 23, CNE strength vs Airbnb capacity

As a final insight, Figure 23 illustrates the relationship between Airbnb capacity and the strength of cross-side network externalities (CNE). The pattern displayed in the graph further reinforces the population density mechanism discussed previously. As shown earlier in Figure 16, higher Airbnb capacity is predominantly observed in areas located farther from LGV stations, which are situated in larger metropolitan centres.

Consistently, Figure 23 shows that the highest CNE values are concentrated within the 1-2 bedroom capacity range. This category corresponds to standard urban apartments, which are the dominant accommodation type in densely populated cities. Therefore, the results suggest that CNE intensity is maximized in urban contexts characterized by moderate capacity listings typical of metropolitan housing structures such as apartments.

6.3 Conclusions

From a managerial perspective, two strategic paths emerge from this population density pattern. The first consists of leveraging the urban advantage by intensifying marketing investments in high-density areas. Given that CNE elasticity is significantly stronger in these contexts, amplifying demand through targeted campaigns is likely to generate disproportionately larger spillover effects on the opposite market side. In particular, demand-side marketing initiatives aimed at attracting tourists in densely populated zones are expected to trigger a stronger supply response from hosts compared to similar investments in smaller, less populated markets. In other words, because cross-side effects are amplified in urban environments, each marginal unit of demand stimulated in these areas produces a greater indirect impact on platform growth.

On the other hand, Airbnb leaders might adopt a second strategic approach by subsidizing submarkets characterized by weak cross-side network externalities (CNE). In particular, these are IRIS with low population density, characterized by independent houses, where an increase in demand generates relatively small spillover effects on supply compared to more urban areas. In such contexts, market forces alone may not be sufficient to trigger self-reinforcing platform growth. This strategy would be especially relevant if decision-makers aim to achieve a more balanced geographical development within the region, avoiding excessive concentration in urban cores and preventing the marginalization of remote markets. By investing in these weaker submarkets, for example through targeted marketing campaigns or incentives for hosts, platform leaders could stimulate supply responses that would otherwise remain limited. In this perspective, specifically for the region of Nouvelle-Aquitaine in the summer of 2017, the opening of the LGV can be interpreted as a complementary factor. As demand in areas close to LGV stations is expected to increase exogenously due to improved accessibility, these central IRIS will experience platform growth even without additional intervention. This external shock allows decision-makers to redirect resources toward more peripheral IRIS, where endogenous growth dynamics are weaker and where the impact of the exogenous shock is minimal. Otherwise with a more aggressive strategy, focusing on central and urban area will lead to bigger effects caused by both the LGV opening and also by the specific nature of CNE within highly populated area. Consistent with these strategies further implications could also be made relatively to housing types and Airbnb capacity. Indeed, following the results obtained, leaders and decision makers would obtain a larger impact on their marketing campaigns within those IRIS with larger apartment shares which of course lead to smaller Airbnb capacities. These conclusions, tell us that, from a managerial perspective, it is also fundamental to deeply understand the characteristics of the region we are operating and capture systematic differences of cities and urban compositions. To this end, the data tell us that between two cities with the same population density, we will invest more capital in the one with the higher apartment share, if our intention is to obtain stronger return and impact on the overall platform growth.

7. Appendix

Table 11, Model 7 (Bedrooms registered as supply proxy) – Exogenous shock impact on Airbnb demand

$D_{i,t,k} = \alpha + \rho_1 \text{ Treat}_{i,p} \times \text{Post}_t + \gamma \text{ SSE}_{i,t} + B_{i,t} + Y_{i,t} + \mu_i + \tau_t + \varepsilon_{i,t}$

MODEL 7 BEDROOM REGISTERED	***k = 2***	***k = 3***	***k = 4***	***k = 5***
Treat_median × Post	0.0692**	0.0729***	0.0718***	0.0711***
(Std.Err.)	(0.0275)	(0.0270)	(0.0265)	(0.0262)
[p-value]	[0.012]	[0.0068]	[0.0068]	[0.0066]
$\boldsymbol{B_{i,t}}$	YES	YES	YES	YES
$\boldsymbol{Y_{i,t}}$	YES	YES	YES	YES
$\mathbf{SSE}_{i,t}$	YES	YES	YES	YES
Observations	109,281	106,457	103,674	100,906
R^2 (within)	0.172	0.170	0.174	0.188

Table 12, Model 8 (Bedrooms registered as supply proxy) – Exogenous shock impact on Airbnb demand

$D_{i,t,k} = \alpha + \rho_1 \, Treat_{i,p} \times Post_t + \rho_2 \, Treat_{i,p} \times Post_t \times \left(\frac{1}{Dist_i}\right) + \gamma \, SSE_{i,t} + B_{i,t} + Y_{i,t} + \mu_i + \tau_t + \varepsilon_{i,t}$

MODEL 8 BEDROOM REGISTERD	***k = 2***	***k = 3***	***k = 4***	***k = 5***
Treat_median × Post	0.0589**	0.0618**	0.0604**	0.0596**
(Std.Err.)	(0.0280)	(0.0276)	(0.0272)	(0.0268)
[p-value]	[0.0355]	[0.0249]	[0.0263]	[0.0263]
treat × post × inv_dist	0.1536***	0.1669***	0.1716***	0.1723***
(Std. Err.)	(0.0379)	(0.0375)	(0.0383)	(0.0402)
[p-value]	[0.0001]	[0.0000]	[0.0000]	[0.0000]
$\boldsymbol{B_{i,t}}$	YES	YES	YES	YES
$\boldsymbol{Y_{i,t}}$	YES	YES	YES	YES
$\mathbf{SSE}_{i,t}$	YES	YES	YES	YES
Observations	109,281	106,457	103,674	100,906
R^2 (within)	0.172	0.170	0.174	0.187

Table 13, Model 11 (Bedrooms registered as supply proxy) – Exogenous shock impact on Airbnb demand

$$D_{i,t,k} = \alpha + \rho_1\, Treat_{i,p} \times Post_t + \gamma\, SSE_{i,t} + X_{i,t} + \mu_i + \tau_t + \varepsilon_{i,t}$$

MODEL 11 BEDROOM REGISTERED	***k = 2***	***k = 3***	***k = 4***	***k = 5***
Treat_median × Post	0.0657**	0.0691**	0.0676**	0.0665**
(Std.Err.)	(0.0276)	(0.0271)	(0.0267)	(0.0263)
[p-value]	[0.0172]	[0.0108]	[0.0112]	[0.0114]
Observations	109,281	106,457	103,674	100,906
$X_{i,t}$	YES	YES	YES	YES
$SSE_{i,t}$	YES	YES	YES	YES
R^2 (within)	0.171	0.170	0.174	0.188

Table 14, Estimation of D2S CNE through 2SLS, using residuals from model 11

ESTIMATION OF CNE (2SLS) – M11	***k = 2***	***k = 3***	***k = 4***	***k = 5***
D2S CNE (ln Di,t,k)	0.0260	0.0244	0.0243	0.0254
AIRBNB REGISTERED	(0.0336)	(0.0315)	(0.0321)	(0.0324)
	[0.4384]	[0.4374]	[0.4489]	[0.4325]
D2S CNE (ln Di,t,k)	0.0283	0.0265	0.0274	0.0277
BEDROOMS REGISTERED	(0.0347)	(0.0329)	(0.0335)	(0.0338)
	[0.4141]	[0.4194]	[0.4127]	[0.4120]
Observations	109,281	106,457	103,674	100,906
Control function	LINEAR	LINEAR	LINEAR	LINEAR
$X_{i,t}$	YES	YES	YES	YES
$SSE_{i,t}$	YES	YES	YES	YES
R^2 (within) - AIRBNB	0.929	0.928	0.927	0.926
R^2 (within) - BEDROOMS	0.914	0.912	0.911	0.910

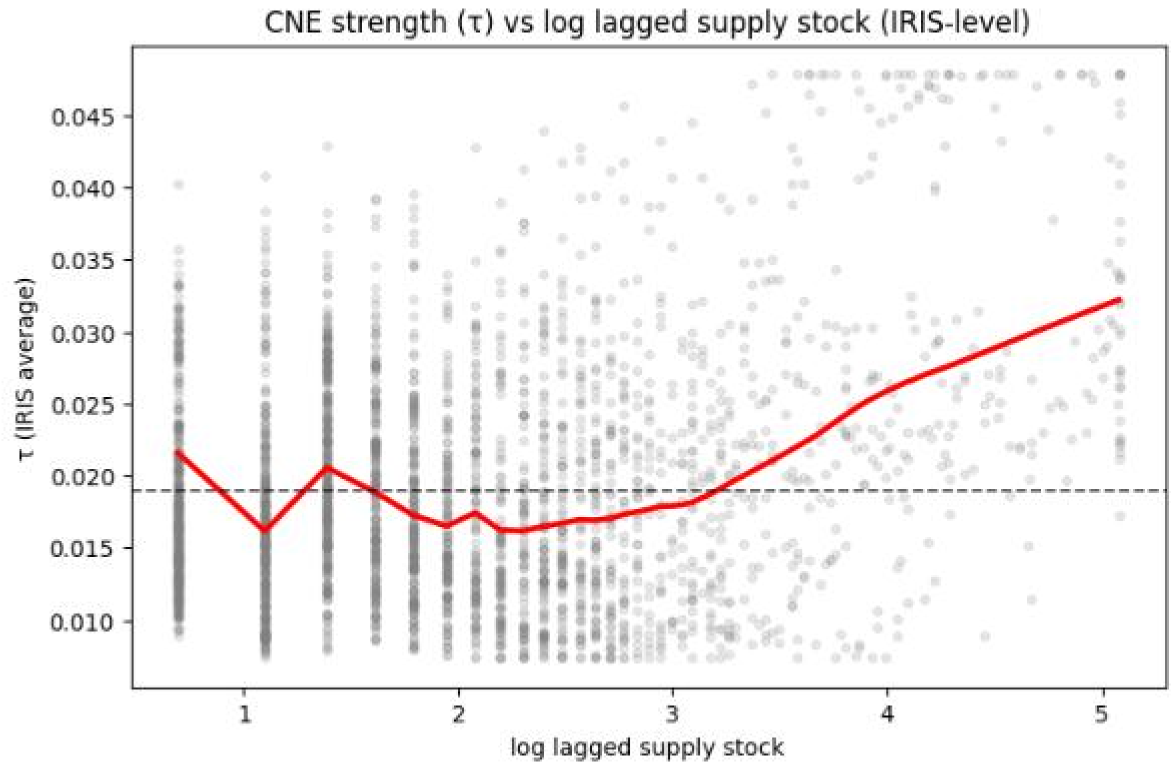

Figure 24, CNE strength vs SSE

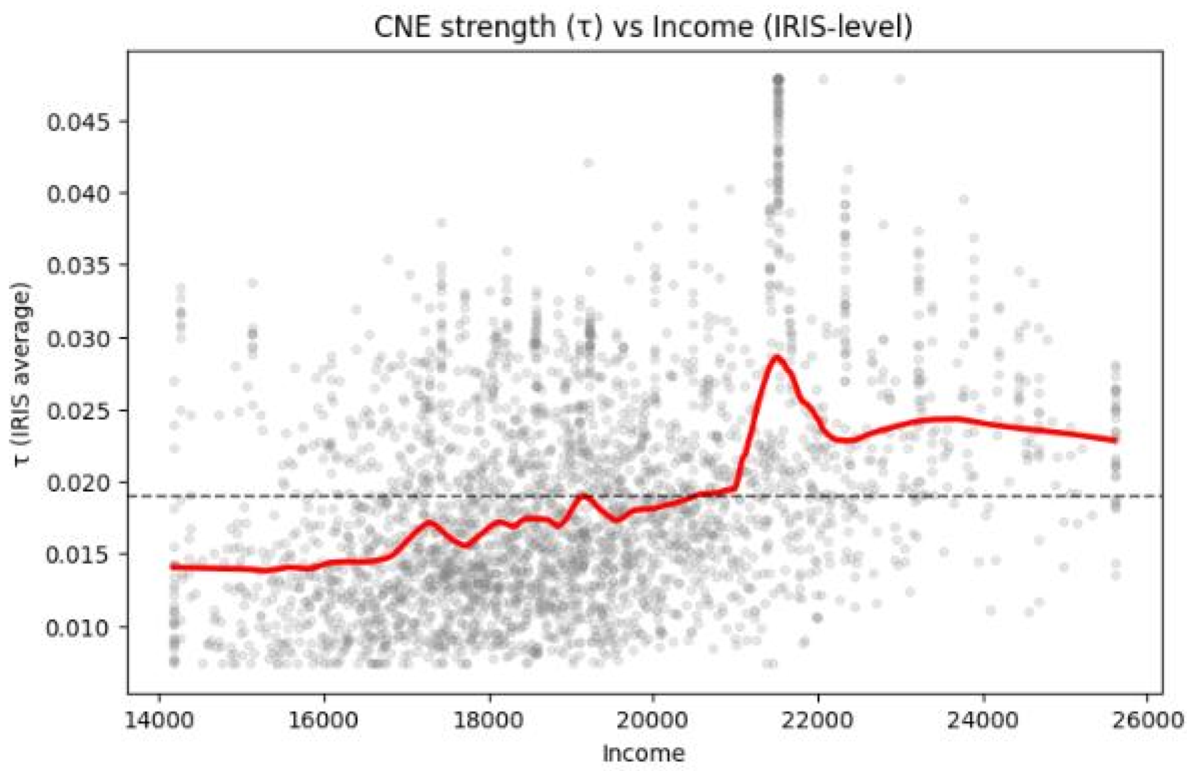

Figure 25, CNE strength vs income

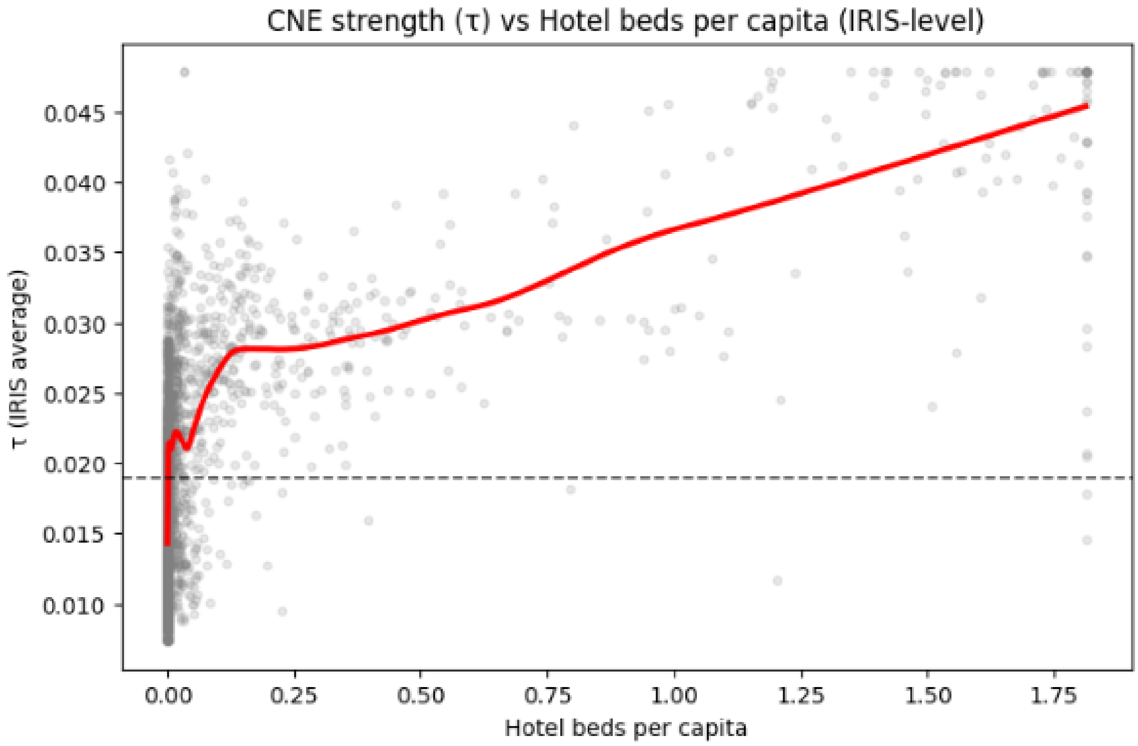

Figure 26, CNE strength vs hotel beds per capita

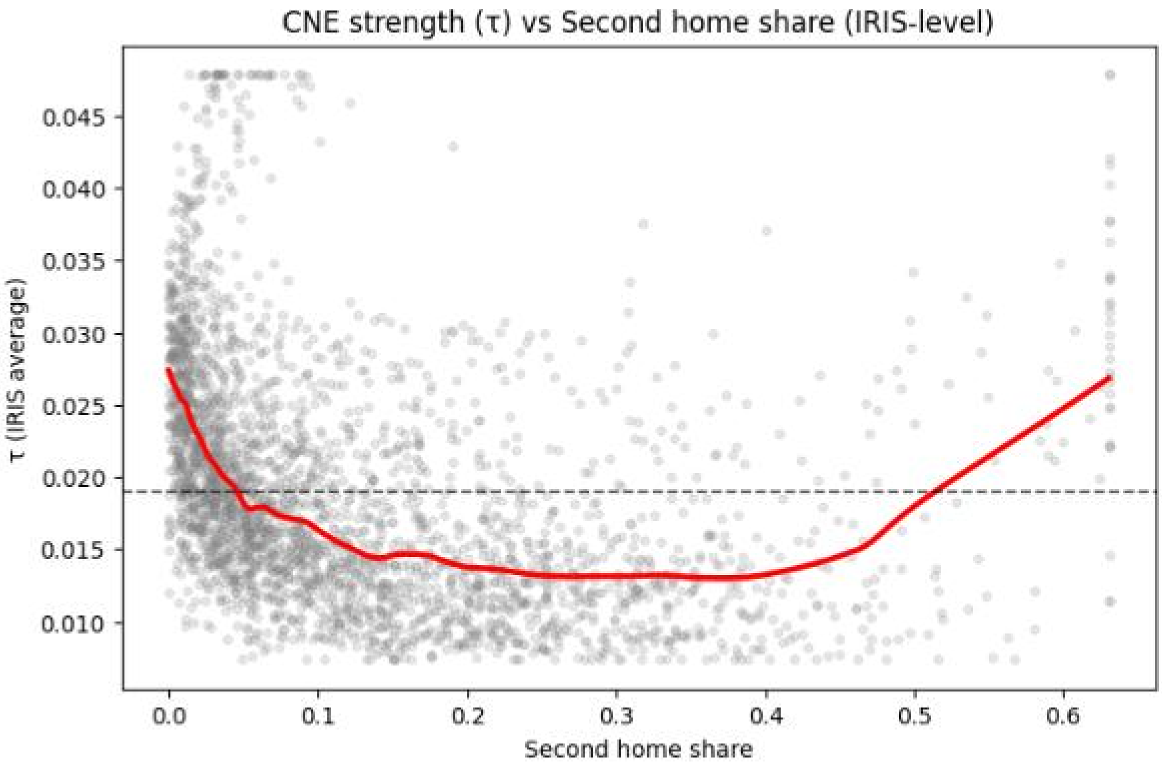

Figure 27, CNE strength vs second home share

8. References

1. Vlamis, K. (2025). *Brian Chesky explains what's needed to accelerate Airbnb's growth*. [online] Business Insider. Available at: https://www.businessinsider.com/brian-chesky-airbnb-growth-rate-whats-needed-2025-9.
2. Curry, D. (2025). *Airbnb Revenue and Usage Statistics*. [online] Business of Apps. Available at: https://www.businessofapps.com/data/airbnb-statistics/.
3. Airbnb (2020). *Free Places to Stay for Medics & Caregivers Fighting Coronavirus in France*. [online] Airbnb Newsroom. Available at: https://news.airbnb.com/free-places-to-stay-for-medics-and-caregivers-fighting-coronavirus-in-france/
4. Statista (2024). *ECB interest rates trends 2020*. [online] Statista. Available at: https://www.statista.com/statistics/621489/fluctuation-of-fixed-rate-interest-rates-ecb/.
5. Banque de France. (2017). *Stable borrowing capacity and dynamic real estate market in 2017 | Banque de France*. [online] Available at: https://www.banque-france.fr/en/publications-and-statistics/publications/stable-borrowing-capacity-and-dynamic-real-estate-market-2017
6. ECB (2025). *Key ECB Interest Rates*. [online] European Central Bank. Available at: https://www.ecb.europa.eu/stats/policy_and_exchange_rates/key_ecb_interest_rates/html/index.en.html.
7. Hoffman, R. (2018). *How to Scale a Magical Experience: 4 Lessons from Airbnb's Brian Chesky*. [online] Medium. Available at: https://reid.medium.com/how-to-scale-a-magical-experience-4-lessons-from-airbnbs-brian-chesky-eca0a182f3e3.
8. Ackerberg, D. A., & Gowrisankaran, G. (2006). Quantifying equilibrium network externalities in the ACH banking industry. The RAND Journal of Economics, 37(3), 738–761.
9. Buzzacchi, L., Grilli, L., & Milone, F. L. (2023). Spatial competition and entry decisions in platform markets.
10. Caillaud, B., & Jullien, B. (2003). Chicken & egg: Competition among intermediation service providers. The RAND Journal of Economics, 34(2), 309–328.

11. Chu, J., & Manchanda, P. (2016). Quantifying cross and direct network effects in online consumer-to-consumer platforms. Marketing Science, 35(6), 870–893.
12. Cullen, Z., & Farronato, C. (2021). Outsourcing tasks online: Matching supply and demand on peer-to-peer internet platforms. Management Science, 67(10), 6071–6087.
13. Evans, D. S. (2003). The antitrust economics of multi-sided platform markets. Yale Journal on Regulation, 20(2), 325–381.
14. Farrell, J., & Saloner, G. (1986). Installed base and compatibility: Innovation, product preannouncements, and predation. American Economic Review, 76(5), 940–955.
15. Filippas, A., & Horton, J. J. (2023).
16. Farronato, C., & Fradkin, A. (2022). The welfare effects of peer entry in the accommodation market: The case of Airbnb. American Economic Review, 112(9), 3032–3063.
17. Hinz, O., Otter, T., & Skiera, B. (2020). Estimating network effects in two-sided markets. Journal of Marketing Research, 57(1), 12–29.
18. Katz, M. L., & Shapiro, C. (1985). Network externalities, competition, and compatibility. American Economic Review, 75(3), 424–440.
19. Katz, M. L., & Shapiro, C. (1986). Technology adoption in the presence of network externalities. Journal of Political Economy, 94(4), 822–841.
20. Katz, M. L., & Shapiro, C. (1994). Systems competition and network effects. Journal of Economic Perspectives, 8(2), 93–115.
21. Kim, J., et al. (2022). Spatial network effects in platform adoption.
22. LISEA. (2023). Traffic results for the LGV Sud Europe Atlantique line.
23. Nair, H., Chintagunta, P., & Dubé, J.-P. (2004). Empirical analysis of indirect network effects in the market for personal digital assistants. Quantitative Marketing and Economics, 2(1), 23–58.
24. Rochet, J.-C., & Tirole, J. (2003). Platform competition in two-sided markets. Journal of the European Economic Association, 1(4), 990–1029.
25. Robertson, D., et al. (2023). Short-term rental regulation and platform supply: Evidence from Bordeaux.
26. Rossi, A. (2023). Entry timing and reputation advantages in platform markets.

27. Rysman, M. (2004). Competition between networks: A study of the market for yellow pages. Review of Economic Studies, 71(2), 483–512.

28. Sanchez-Cartas, J. M., & León, G. (2021). A literature review of digital platform research. Telecommunications Policy, 45(6), 102–116.

29. Sriram, S., et al. (2015). Network effects in two-sided markets: A review. Marketing Science, 34(1), 1–14.

30. Stallkamp, M., & Schotter, A. P. J. (2021). Platforms without borders? The international strategies of digital platform firms. Global Strategy Journal, 11(1), 58–80.

31. Stourm, E., & Albuquerque, P. (2024). Flowers and bees: Local and spatial network effects in sharing economy platforms.

32. Sun, S., Zhang, X., & Wang, Y. (2021). Urban infrastructure and Airbnb location choice: Evidence from Nanjing. Tourism Management, 83, 104–214.

33. Wilbur, K. C. (2008). A two-sided, empirical model of television advertising and viewing markets. Marketing Science, 27(3), 356–378.

34. Zervas, G., Proserpio, D., & Byers, J. (2017). The rise of the sharing economy: Estimating the impact of Airbnb on the hotel industry. Journal of Marketing Research, 54(5), 687–705.