UNIVERSITY OF MANNHEIM
Department of Economics

Spatial Competition and Hotel Pricing: Evidence from the 9-Euro Ticket

Master's Thesis

Submitted to:
Prof. Dr. Nicolas Schutz

Submitted by:
Guilherme Schultz Lindenmeyer
Student ID: 1906999
Degree Program: Master of Science in Economics

-

gslindenmeyer@gmail.com

Mannheim, 04.07.2024

Contents

List of Figures

List of Tables

1 Introduction

In the summer of 2022, the German government introduced an innovative transport program, the 9-Euro Ticket, which significantly reduced the cost of public transport across the country. For nine euros a month, during the months of June, July, and August, residents and visitors were able to have unlimited travel on all municipal and regional transport systems, with the exception of the IC and ICE express trains. This promotional ticket was designed to encourage the use of public transport rather than personal vehicles, in line with environmental objectives.

The rationale behind the one-time special offer of the 9-Euro ticket is multifaceted. According to the German Federal Government, the promotion was part of a larger package of financial relief for the German population in the face of rising fuel, energy, and heating prices. The discount was approved and promoted by the coalition committee in Germany, which also aimed to propose this measure to promote the use of public transport as opposed to the use of private cars. In other words, it had the aspect of reducing carbon emissions, encouraging the use of public transport, and also attracting tourists to Germany during the summer. The funds to finance it were provided by the Regionalization Act.

The advantages of the ticket went beyond the low price of transportation. The ticket simplified the complex and fragmented public transportation network within Germany's 16 states, each featuring multiple transport associations. This means that the same state, for example, Baden-Württemberg (BW), can have several transport networks operated by different authorities. And sometimes the authorities cross state borders. The Verkehrsverbund Rhein-Neckar (VRN) association operates in the north of BW but also covers part of Rhineland-Palatinate. The complexity and fragmentation of the different transport regions and their different ways of crossing Germany have been abruptly and brutally simplified by the German 9-Euro ticket. The price and simplicity also meant that Germany had a surge of tourists for the summer of 2022.

For around three months, the ticket has seen a notable response, with more than 52 million units sold[1], not counting automatic sales to subscription holders, and approximately saved 1.8 million tons of CO_2. This widespread distribution has enabled around one billion journeys to be made each month, underlining the popularity of the ticket and the great mobility it has enabled. Our study aims to analyze whether this increased mobility has translated into an increase in the perceived value of German hotels, as well as disturbing the nature of competition between German hotels.

Despite its benefits, the ticket also presented challenges, such as overcrowding and delays, highlighting the infrastructure's inability to meet increased demand. The

[1]Federal Government of Germany, `https://www.bundesregierung.de/breg-de/themen/tipps-fuer-verbraucher/faq-9-euro-ticket-2028756`.

reaction from the public and politicians was mixed, with significant debate about the sustainability of such a system. Initial promises of a permanent low-cost ticket turned into the introduction of a more expensive 49-euro ticket, reflecting the financial difficulties of the subsidy.

The wide-ranging implications of the 9-Euro ticket have given rise to numerous studies examining its environmental impacts, economic balance, and political effectiveness (Gohl and Schrauth, 2022; Aydin and Kürschner Rauck, 2023; Bissel, 2023). Our research focuses on a specific but meaningful effect: the impact of the 9-Euro ticket on competition between hotels in Germany. This study investigates whether hotels have adjusted their pricing strategies in response to the increased demand provided by the ticket.

To this end, we carried out a detailed analysis using daily price data extracted from the web for hotels in the main German cities and other European cities as a control group, covering the months of August and September 2022. We extended the dataset to capture price changes for the same German hotels from 2022 to 2023, where we created the second control group. In addition, we enriched the database with geolocation data, as well as the hotel's distance from its center and distance from the nearest public transport stop. Our aim is therefore to examine three key dimensions of hotel competition, both locally and globally: geographical proximity, online reputation, and quality (proxied by the number of stars). For each of these dimensions, we investigated the effect of the ticket on the estimation of hotel best-responses parameters. The results of this research offer insight into the impact of this transportation policy on the German hotel sector.

The empirical analysis revealed that the 9 euro ticket led to a slight increase in hotel prices during the month of August 2022, revealing the impact of greater mobility on demand and hotel prices. In the competition analysis, geographical proximity played a major role, as analyzing competition by geographical distance most strongly explained the variation in prices. Further, our study concluded that the price competition between hotels is marked by strategic complementarity and that during the validity of the ticket, hotels' pricing strategies became less sensitive to their rivals' prices, indicating a shift towards more independent pricing. Although factors such as quality and online reputation influenced prices, their impact was less significant compared to geographical factors, especially when limited to the finite number of hotels, thus highlighting the importance of location and local competition in the hotel industry. These results suggest that policies such as the 9-euro ticket, even when not directly targeting the hotel sector, can affect other economic sectors such as hospitality.

The rest of this thesis is structured as follows: Section 2 reviews the relevant literature; Section 3 describes the theoretical framework used to support our empirical methodology; Section 4 describes the methodology and data used, followed by an analysis of the results in Section 5; finally, Section 6 discusses the conclusions of this work, its limitations and economic implications.

2 Literature Review

This literature review aims to cover the key relevant papers in the literature on spatial competition, both in theory and in empirical research. One of the first works to consider the aspect of location in the context of price competition was Hotelling (1929). Hotelling's work was not aimed at initiating the literature on spatial competition. On the contrary, he was addressing the thinking of the time, which was to accept the intrinsic instability of duopolies. Cournot innovated by thinking of competition via capacity, in which firms compete by maximizing their profits with the strategic variable q, but stability is only maintained if prices are not independent. Hotelling's view was that discontinuity, present, for example, in Bertrand's model, was an aberration of nature. In practice, prices are adjusted, and demand changes gradually, rather than abruptly. Hotelling's innovation was to see the market as a straight line rather than a point. With this, it was as if each seller was a monopolist up to a certain range and region.

Hotelling's work broke new ground by demonstrating that moderate variations in prices between sellers do not lead to an abrupt change in consumer choice, but rather to a continuous transition, suggesting a more stable competitive equilibrium than previous theories had proposed. And by modeling with a straight line, it was possible to be virtually a limited monopoly without rivals having an explicit price agreement. Hotelling also discusses the social inefficiencies generated by competition. In the model, profits depend directly on transportation costs, which means that agents would have incentives to have higher transportation costs to increase profits. This makes sense in the model since increasing transportation costs is nothing more than increasing the region in which the trader acts as a monopolist, relieving the pressure of competition. Hotelling then models the sequential entry of traders and the location decision. In the context of maximizing private profit, the tendency of the theory proposed by Hotelling is for companies to locate as close to each other as possible, generating agglomeration - or low diversification, not minimizing society's transport costs. In the case of hotels, we could think of them as being concentrated in the center of a city.

The idea of agglomeration became known as the principle of minimum differentiation, which only years later was proven invalid by d'Aspremont et al. (1979). The point to be corrected was that, by assuming linear transportation costs in a simultaneous price decision game between two firms on the line, there is no Nash Equilibrium in pure strategies for all initial location configurations. Furthermore, modifying the model to assume quadratic transport costs[2], there is the conclusion that there is a price equilibrium, regardless of the initial locations of the sellers. Contrary to the principle of minimum differentiation, this modification concludes that sellers will position themselves at a maximum distance from each other. This result shows that space matters more than initially proposed by Hotelling. In the context of our framework, this finding highlights that hotels may choose to distance themselves from competitors rather than grouping together as Hotelling initially proposed.

Adapting Hotelling's original model to a two-stage price and location model, Osborne and Pitchik (1987) showed an equilibrium with linear prices by allowing mixed strategies in prices, considering pure strategies in the first stage (location). This equilibrium has as its solution the location close to the quartiles of the straight line, approaching the solution that minimizes transport costs. A later paper by d'Aspremont et al. (1983) proved that equilibria do not exist with an agglomeration of firms at the same point, as originally proposed by Hotelling in 1929. The argument follows the logic of Bertrand Competition, where firms selling homogeneous products would have no incentive to agglomerate at the same point on the line because the price would tend towards marginal cost and profit towards zero.

Hotelling's model can be adapted with the Cournot assumption, where the strategic variable is quantity. One of the first papers to mix spatial competition with the Cournot assumption was Greenhut and Greenhut (1975), although it did not use Hotelling's modeling directly. Later, in Hamilton et al. (1989) there was modeling a la Hotelling with a comparison of spatial competition comparing Bertrand with Cournot.

Anderson and Neven (1991) further explored the adaptation of Hotelling's model in the context of Cournot competition, with a special focus on firms' location trends. In this work, the authors model a two-stage model with two firms, with location choice as the first stage, and quantity setting in the second stage. Solving by backward induction, they prove that, assuming convex transportation costs and demand such that the market is always served by both firms, with competition a la Cournot at each point on the line. The only equilibrium in this model is symmetrical and assumes agglomeration in the center of the line. Extending the model to n firms and assuming linear demand with linear transportation costs, the only equilibrium is agglomeration in the center. Although this model has the property of zero profit when allowing free entry, its advantage is that a potential entrant attracts consumers from the entire market, not just its immediate region.

The modifications to the Hotelling model suggested so far have been in the direction of changing the strategic variable, the concept of equilibrium, the number of firms and the cost of transportation. Salop (1979) extended the model to be analyzed in a circular

[2]The assumption of quadratic transport costs allowed the discontinuities in the demand curve to be removed when analyzing with linear transport costs.

manner, establishing the basis for future analyses of market competition in other spatial configurations. Neven (1986) paper was one of the first to change the distribution of consumers along the line, which until then had been uniformly distributed. The paper assumes quadratic transport costs and a uni-modal distribution in the center of the line. The trade-off with two opposing forces is highlighted: moving away from the rival firm to weaken competition or moving closer to the center to be closer to consumers. The authors proved that concave distributions of consumers on the line guarantee the existence of price equilibrium. In addition, their result shows that by gradually moving from a uniform distribution to a strong concentration of consumers in the center, firms will gradually choose to move away from the edges and locate themselves closer to the center, making evident the trade-off between more monopoly power vs. market size.

As the theoretical literature on spatial competition has progressed, we can see that we are getting increasingly closer to more realistic settings. In our context, it is reasonable to imagine hotels scattered throughout a city, but their locations are not random. Most hotels choose to position themselves close to the city center, which is in theory where the majority of their demand wants a hotel to be. Even so, the reality is not total agglomeration in the center. Hotels are spread out in other parts of the city. Based on our descriptive analyses, it seems that hotels tend to seek a trade-off between monopoly power, by positioning themselves in regions with few rivals and having greater control over their prices, and positioning themselves in regions with larger consumers but suffering greater pressure on their prices. In addition, we realized that hotels that choose to isolate themselves can take two paths: either they suffer negative pressure on their prices because they are not in the best location based on demand, or they are hotels that use other characteristics to be able to charge a price premium (e.g. resorts, etc.). In the following sections, we explore more about the descriptive analysis of our data.

Caplin and Nalebuff (1991) paper made a broad demonstration in the context of imperfect competition, proving that there is pure price equilibrium in multidimensional product differentiation models, laying the foundations for the advancement of the literature. In other words, in the spatial context, it means that price equilibria exist without the imposition of symmetry for any number of competitors and products.

Tabuchi and Thisse (1995) article was a significant advance on Hotelling's price-location model, showing that as we move away from uniform densities, the assumption of symmetry between firms begins to lose meaning. The authors also model sequential entry, where two firms enter the market in sequence, but choose prices simultaneously. The result interestingly reveals asymmetry, with the first mover having an advantage and locating itself in the center, and the second mover choosing to locate itself further away, guaranteeing monopoly power, but fewer profits due to the distribution of consumers - avoiding the pressure of competition. This issue was further explored by

Anderson et al. (1997), where they augmented the model by considering log-concave consumer densities and proved that, under some conditions, a unique pure-strategy equilibrium exists in the two-stage model. If the density is very concave, there are no symmetric equilibria, only asymmetric ones.

Another direction in which Hotelling's model could be adjusted is towards different production costs that vary according to space. Combes (1997) was the first to mention competition a la Cournot with some differentiation in production costs, but restricted to just two locations. Only in Mayer (2000) there was a modeling a la Cournot and Hotelling considering firms with heterogeneous production costs with a continuum of location choices. Assuming a convex distribution of production costs, the result shows that the agglomeration solution in the center is the result only in the case where the center is the lowest-cost region. If there is a symmetrical and concave distribution of costs, the location equilibrium depends on transportation costs. For sufficiently low transport costs, firms opt for maximum differentiation because they prefer to stay in regions with lower production costs, and for sufficiently high transport costs, firms move closer to the center. A number of other theoretical modifications to Hotelling's initial model were discussed in a critical review of the literature by Biscaia and Mota (2013).

We have focused so far on showing the theoretical advances in the spatial competition literature. But since we need estimation methods, we turn to the advances in the literature when it comes to the empirical side. In the empirical direction of the study of price competition, one of the classic articles in the area is Berry et al. (1995), which develops the empirical method for analyzing supply and demand in markets with product differentiation, assuming Bertrand competition. This method takes product characteristics into account in order to estimate supply and demand coefficients. This allows policy counterfactuals to be carried out to investigate a particular market, such as the U.S. automotive industry. On the other hand, the estimation is structured and there is a great need for data. In addition, the role of geographical proximity is left aside in this empirical framework. In another 1995 study, Feenstra and Levinsohn (1995) developed a method for estimating markups and market conduct in sectors with multidimensionally differentiated products, applying their model to the U.S. automobile market in 1987. The authors considered the Euclidean distance between product characteristics

In Pinkse et al. (2002) there was an improvement in the empirical estimation of prices taking into account the role of location. In this work, they introduced an innovative method for estimating competition via prices in the context of differentiated products. In their work, price is considered the strategic variable, and best-response functions are estimated. They use parametric and semi-parametric methods to estimate reaction function slopes. With this, it is possible to investigate the nature of price competition in a given market: whether it is local or global. Pinkse et al. (2002) proposes characteristics of rivals' nearest neighbors as an instrument for the rival's price, which is the endogenous variable. They study the wholesale gasoline market and investigate whether competition is local or global. Based on their results, they conclude the finding of extreme local competition in the example studied. In the Pinkse and Slade (2004), the authors investigate the UK brewing industry and the effects of mergers. The study is based on a structural model of Pinkse et al. (2002), extending it to include substitution between brands. They study spatial competition in product characteristics and conclude that competition is local. Both articles proposed using the inverse of Euclidean distance between products as a proxy of proximity between products.

This paper aims to add to the literature on spatial competition and the lodging industry. Specifically on the hotel sector, the works of Mazzeo (2002a,b) were among the first in the field of industrial organization to study the hotel market with empirical models. Both articles aim to endogenize the firm's decision-making in an estimation and study the relationship between outcomes and market structure. The article by Kosová et al. (2013) investigates a panel data of hotels, studying whether different ways of organizing the hotel affect the outcomes. Lewis and Zervas (2016) analyzes the effect of consumer reviews on the hotel industry. In Kalnins (2016), the authors present the research and findings initially proposed by Baum and Mezias (1992) on the Manhattan hotel sector to an investigation of various hotel markets in the U.S. and the conclusion that competition is predominantly local in two dimensions: price and geographical location. The paper by Kalnins et al. (2017) investigates the role of mergers in the hotel sector and their competitive effects. In another more recent study, Armona et al. (2021) combines consumer browsing history with hotel data to investigate the effects of mergers. The work by Farronato and Fradkin (2022) investigates the effect of Airbnb on the lodging industry.

In addition, another area that we intend to contribute to is testing the hypothesis that firms might have boundaries on their choice. Firms, in order to set their price, might look at a limited and potentially pre-defined set. In the case of hotels, firms may limit themselves to competing with only a few hotels in the round. The work of Baum and Lant (2003) investigates the Manhattan hotel industry and shows that managers tend to rely on a limited set of rivals to make pricing decisions. In Li et al. (2018) the authors also investigate Manhattan, but using data from online clicks and exposures and estimation via 2SLS. In their work, they use machine learning models of variable selection and conclude that distance or quality can play a role in limiting hotel competition. In the paper by Becerra et al. (2013) the authors study the hotel market in Spain and conclude that differentiation between hotels helps to alleviate competitive pressure. The paper by Lee (2015) investigates the hotel sector in Texas and finds that competition is local and disregards the hypothesis that it could be global within a city or region. Finally, the article by Rezvani and Rojas (2020) investigates price competition between hotels in the dimensions of price, quality, and reputation, and studies the extent to which it is localized.

Our work aims to apply the Rezvani and Rojas (2020) framework, but differs from it in the context that we extend it to investigate the nature of competition in Germany. That is, with our novel database collected during the years 2022 and 2023, we intend to study the underlying structure of hotel price competition in Germany. Furthermore, the 9-Euro Ticket is an event that lowered transportation costs widely across Germany during 2022. Our work aims to assess whether this event has provoked any disturbances in the way hotels compete in any of the three dimensions: geographic, quality, and online reputation.

3 Theoretical Framework

One of the aims of this work is to estimate the optimal price response equations for hotels in Germany and how the 9-Euro ticket event impacts this relationship. In this section, we will delineate the theoretical framework on which the model we estimate is based. Using the results of Pinkse et al. (2002), Li et al. (2018) and Rezvani and Rojas (2020), we demonstrate that, by assuming a Bertrand-Nash conduct and a linear demand function, the optimal best-response function of hotels is linear in the prices of other rivals. Later, we present the reasoning of Li et al. (2018), where they show that even assuming a multinomial logit demand, the linear best-response approximation in rivals' prices is still suitable.

3.1 Linear Demand

Assuming N hotels in a given market, we define a market for contextualization purposes as a combination of city and night. The demand for hotel j can be expressed as:

$$q_j = \alpha_j + \gamma_{j1}p_1 + \gamma_{j2}p_2 + \ldots + \gamma_{jN}p_N \tag{1}$$

Here, q_j determines the aggregated quantity of rooms demanded by consumers (tourists, business travelers, etc.), α_j is the base demand for hotel j regardless of price, and γ_{ij} is the degree of influence that the price of hotel i has on the demand for hotel j. Assuming that the hotels act rationally, each hotel will then set a price p_j for each market that maximizes profits, also taking into account the prices set by other hotels in the vicinity. The profit maximization problem can be written as:

$$\max_{p_j} (p_j - c_j)\, q_j = p_j\,(p_j - c_j)\,(\alpha_j + \gamma_{j1}p_1 + \gamma_{j2}p_2 + \ldots + \gamma_{jN}p_N) \tag{2}$$

This translates into a best-response pricing strategy where hotel j sets its price p_j based on the observed prices of the other hotels in the same market. The best response function can be derived from the first-order condition of the maximization problem, where p_{-j} is the vector of rivals' prices:

$$p_j(p_{-j}) = \frac{c_j}{2} - \frac{\alpha_j}{2\gamma_{jj}} - \frac{\gamma_{j1}}{2\gamma_{jj}}p_1 - \frac{\gamma_{j2}}{2\gamma_{jj}}p_2 - \ldots - \frac{\gamma_{j,j-1}}{2\gamma_{jj}}p_{j-1} - \frac{\gamma_{j,j+1}}{2\gamma_{jj}}p_{j+1} - \ldots - \frac{\gamma_{jN}}{2\gamma_{jj}}p_N \tag{3}$$

Therefore, under the assumption of linear demand, the optimal response function shows that each hotel's pricing strategy is linearly influenced by competitors' prices, providing an initial approach to understanding price setting in a competitive hotel market. To operationalize Eq.3, we can simplify the first two terms, which are specific to firm j as $\delta_j = \frac{c_j}{2} - \frac{\alpha_j}{2\gamma_{jj}}$. Additionally, we can define each component multiplying the price p_i as $\beta_{ji} = -\frac{\gamma_{ji}}{2}\gamma_{jj}$. We then have the estimable equation:

$$p_j(p_{-j}) = \delta_j + \sum_{\substack{i=1 \\ i\neq j}}^{N} \beta_{ji}p_i \tag{4}$$

In Section 4, we will discuss in more detail the methodology to be employed, discuss the possible problems that can arise during the estimation process and, based on the literature, suggest possible solutions to effectively estimate the parameters of interest.

3.2 Multinomial Logit Demand

The multinomial logit model, based on McFadden (1978) work, is prevalent in empirical research across both academic and industry settings. The advantage of this model is largely due to its analytical ease, i.e. there is a closed-form solution, avoiding complex cases involving multivariate integration. Multinomial logit demand is a step towards complexity relative to linear demand, with it, it is possible to assume a form of heterogeneity in consumers and products. Here, we present the Li et al. (2018) reasoning, which shows that even assuming multinomial logit, the assumption of linear best-responses in rival prices is still adequate.

Assuming again N hotels, with each hotel supplying one room. Define the potential demand for each market as $\mathcal{M}$. Consumer i choosing hotel j obtains utility u_{ij}. There is an outside option, called u_{i0}, which can be normalized to zero when the consumer chooses not to choose a hotel that night and city. Therefore, in general, the utilities are modeled as, for each consumer i and hotel j:

$$u_{ij} = \alpha_j + \beta p_j + \epsilon_{ij} \quad \text{e} \quad u_{i0} = \xi_0 \tag{5}$$

The term α_j represents the utility of the hotel j, which is assumed to be the same for all consumers. The terms $\varepsilon_{i0}, \varepsilon_{i1}, \varepsilon_{i2}, \ldots, \varepsilon_{iN}$ represent the consumer's specific taste shock for all rooms in a given market, including the outside option ϵ_{i0}. As is standard, let's assume that the random vector ε is drawn i.i.d. from the type I extreme value distribution. With this assumption, the demand q_j for hotel j can be seen as follows:

$$\begin{aligned} q_j = \mathcal{M} s_j &= \mathcal{M} \cdot \mathbb{P}\left(\alpha_j + \beta p_j + \varepsilon_{ij} = \max_{0 \le k \le N} \{\alpha_k + \beta p_k + \varepsilon_{ik}\}\right), \\ &= \mathcal{M} \cdot \frac{\mathrm{e}^{\alpha_j + \beta p_j}}{1 + \sum_{k=1}^{N} \mathrm{e}^{\alpha_j + \beta p_j}}. \end{aligned} \tag{6}$$

If one wants to adapt the demand equation to be an equation that can be estimated, the Berry inversion, initially proposed in Berry et al. (1995), can be used,

$$\log s_j - \log s_0 = \log\left(s_j / s_0\right) = \alpha_j + \beta p_j. \tag{7}$$

With this, having market share data and a reasonable candidate for the potential market, it is possible to estimate this equation. Unfortunately, as we did not observe the market share data, we could not estimate demand assuming a multinomial logit. However, the intention is to show that the linear best response in rivals' prices is reasonable even in the context of multinomial logit demand. To do this, let's once again assume that the firms compete a la Bertrand, and combine Eq. 2 of profit maximization with the demand equation defined in Eq. 6:

$$\max_{p_j} \pi_j(p_j, p_{-j}) = \max_{p_j} (p_j - c_j)\, d_j(p_j, p_{-j}) = \max_{p_j} (p_j - c_j) \; \underbrace{\mathcal{M} \cdot \frac{e^{\alpha_j + \beta p_j}}{1 + \sum_{j=1}^{N} e^{\alpha_j + \beta p_j}}}_{d_j} \tag{8}$$

The market share s_j is defined as:

$$s_j = \frac{e^{\alpha_j + \beta p_j}}{1 + \sum_{k=1}^{N} e^{\alpha_k + \beta p_k}} \tag{9}$$

The derivative of s_j with respect to p_j is defined as:

$$\begin{aligned}\frac{\partial s_j}{\partial p_j} &= \frac{\partial}{\partial p_j}\left[\frac{e^{\alpha_j+\beta p_j}}{1+\sum_{k=1}^{N} e^{\alpha_k+\beta p_k}}\right] \\ &= \beta\frac{e^{\alpha_j+\beta p_j}}{1+\sum_{k=1}^{N} e^{\alpha_k+\beta p_k}} - \beta\frac{e^{\alpha_j+\beta p_j}\cdot e^{\alpha_j+\beta p_j}}{\left(1+\sum_{k=1}^{N} e^{\alpha_k+\beta p_k}\right)^2} \\ &= \beta s_j - \beta s_j^2\end{aligned} \tag{10}$$

The first-order condition for profit maximization with respect to price p_j is obtained by taking the derivative of the profit function and setting it to zero.

$$\begin{aligned}\frac{\partial \pi_j}{\partial p_j} &= \frac{\partial}{\partial p_j}\left[(p_j-c_j)\cdot\mathcal{M}\cdot\frac{e^{\alpha_j+\beta p_j}}{1+\sum_{k=1}^{N} e^{\alpha_k+\beta p_k}}\right] \\ &= \mathcal{M}\left[(p_j-c_j)\cdot\frac{\partial s_j}{\partial p_j}+s_j\right]\end{aligned} \tag{11}$$

Plugging in the derivative of s_j:

$$\begin{aligned}\frac{\partial \pi_j}{\partial p_j} &= \mathcal{M}\left[(p_j-c_j)\cdot\left(\beta s_j-\beta s_j^2\right)+s_j\right] \\ &= \mathcal{M}\left[(p_j-c_j)\cdot(\beta s_j(1-s_j))+s_j\right] \\ &= \frac{\mathcal{M}}{s_j}\left[(p_j-c_j)\beta(1-s_j)+1\right]\end{aligned} \tag{12}$$

Setting this derivative equal to zero and solving for p_j:

$$\begin{aligned}\frac{\mathcal{M}}{s_j}\left[(p_j-c_j)\beta(1-s_j)+1\right] = 0 \\ (p_j-c_j) &= \frac{-1}{\beta(1-s_j)} \\ p_j &= c_j - \frac{1}{\beta}\frac{1}{1-s_j}\end{aligned} \tag{13}$$

This results in the best-response function for hotel j in terms of its own price and the prices of its competitors, as s_j depends on both p_j and p_{-j}. Now, we are going to show that the best-response function is non linear in rivals' prices. For this, considering the chain rule:

$$\frac{\partial s_j}{\partial p_k} = \frac{\partial s_j}{\partial p_j}\cdot\frac{\partial p_j}{\partial p_k}+\frac{\partial s_j}{\partial p_k} \tag{14}$$

And it is straightforward to see that the following equations hold:

$$\begin{aligned}\frac{\partial s_j}{\partial p_j} &= \beta(1-s_j)s_j \\ \frac{\partial s_j}{\partial p_k} &= -\beta s_j s_k\end{aligned} \tag{15}$$

Now, consider $k \neq j$. The derivative $\frac{\partial p_j}{\partial p_k}$ is then equal to:

$$\begin{aligned}\frac{\partial p_j}{\partial p_k} &= -\frac{1}{\beta}\frac{\partial}{\partial p_k}\left[\frac{1}{1-s_j}\right] \\ \frac{\partial p_j}{\partial p_k} &= -\frac{s_j}{(1-s_j)}\frac{\partial p_j}{\partial p_k} + \frac{s_j s_k}{(1-s_j)^2} \\ \frac{\partial p_j}{\partial p_k}\left(\frac{1}{1-s_j}\right) &= \frac{s_j s_k}{(1-s_j)^2} \\ \frac{\partial p_j}{\partial p_k} &= \frac{s_j s_k}{(1-s_j)}\end{aligned} \tag{16}$$

Now, it is possible to take the last term, multiply by $\frac{-\beta}{-\beta}$ and use the first order condition:

$$\begin{aligned}\frac{\partial p_j}{\partial p_k} &= \frac{s_j s_k}{(1-s_j)} = s_j s_k\left(\frac{1}{1-s_j}\frac{-\beta}{-\beta}\right) \\ \frac{\partial p_j}{\partial p_k} &= -\beta s_j s_k \underbrace{-\frac{1}{\beta}\frac{1}{1-s_j}}_{\text{FOC}} \\ \frac{\partial p_j}{\partial p_k} &= -\beta s_j s_k (p_j - c_j)\end{aligned} \tag{17}$$

Finally, we arrive at the result of Li et al. (2018), which shows that since the term $\frac{\partial p_j}{\partial p_k}$ is not constant, p_j is not linear in the prices of rivals. However, the silver lining showed by Li et al. (2018) is that we can apply Taylor's linear approximation to the best-response function of p_j, thus obtaining:

$$p_j\left(p_{-j}\right) = p_j\left(p_{-j}^*\right) + \left.\frac{\partial p_j}{\partial p_1}\right|_{p_{-j}^*}(p_1 - p_1^*) + \ldots + \left.\frac{\partial p_j}{\partial p_N}\right|_{p_{-j}^*}(p_N - p_N^*) + \varepsilon \tag{18}$$

The first-order Taylor approximation ensures that the cross-price derivatives are constant, making it an approximation of the best-response function where rivals' prices enter additively. The term ε represents the approximation error, and Li et al. (2018) shows that the linear approximation is bounded by an error of approximately 0.08% - 0.6% of the equilibrium price when considering a range of $\pm\$50$ of the equilibrium price. With this, one can see that even assuming a more robust demand as opposed to simple linear demand, linear best-responses in rivals' prices are still an adequate approximation. The equation can be operationalized in the same way as linear demand.

we discuss the details of the methodology and estimation in the next chapter.

4 Methodology and Data

Our aim is to estimate best-response functions in the dataset of hotels in Germany, paying specific attention to the 9-Euro ticket event and whether it has changed the nature of the hotel competition. In this section, we will first present the methodology that will be used to achieve the objectives of this study. Subsequently, we shall present the hotel dataset, as well as discuss some preliminaries.

4.1 The Model

Consider N the number of hotels available in a given market (combination of city and night), from Section 3, we have the following equation:

$$p_j(p_{-j}) = \delta_j + \sum_{\substack{i=1 \\ i \neq j}}^{N} \beta_{ji} p_i \tag{19}$$

Recall that δ_j is the intercept of the reaction function specific to hotel j (which contains information about the marginal cost), p_i are the rivals' prices and β_{ij} are the estimated coefficients of the reaction slopes of each rival hotel. As presented in Rezvani and Rojas (2020), there is an empirical difficulty in estimating it the way it is. This is because if there are n hotels, we will have to estimate n^2 coefficients. We are going to use the same proposed solution as Rezvani and Rojas (2020) and Pinkse et al. (2002), which is to assume that the slopes of each best-response function are a function of the proximity between the hotels in some characteristic. As presented by Rezvani and Rojas (2020), this means that $\beta_{ij} = \theta x_{ij}$ where x_{ij} is the proximity between hotel j and hotel i in some dimension. This makes the previous equation look like:

$$p_j(p_{-j}) = \delta_j + \theta \sum_{\substack{i=1 \\ i \neq j}}^{N} x_{ij} p_i + \varepsilon_{ij} \tag{20}$$

Reducing the complexity of the estimation to a single coefficient, θ, at the expense of the assumption that the effect is linear in proximity and symmetrical between the different hotels. Incorporating the time component, and assuming that x_{ij} is an element of a weighting matrix $\mathbf{W}$, which has dimension NxN and diagonals being zero, the matrix version of the equation is:

$$\mathbf{p}_t = \boldsymbol{\delta} + \theta(\mathbf{W}_t \mathbf{p}_t) + \boldsymbol{\varepsilon}_t \tag{21}$$

Here, $\mathbf{p}_t$ is a vector of prices at time t, $\boldsymbol{\delta}$ is a vector of hotel-specific intercepts, and $\boldsymbol{\varepsilon}_t$ is a vector of error terms at time t. The coefficient θ is the common parameter representing the influence of prices of other hotels on firm j's price. In this study, as in Rezvani and Rojas (2020), We will look at three ways of assessing proximity between hotels: (i) geographical distance between hotels, in kilometers; (ii) quality, proxied by the number of stars each hotel has; and (iii) online reputation, proxied by a combination of the volume of ratings a hotel has and its 0 to 10 rating on the Booking.com platform. The online reputation of a hotel does not necessarily correlate with its quality, since in the descriptive analysis of our data it is possible to see that hotels with high ratings tend to be hotels considered good value for money, and they are often not the highest quality hotels. We therefore believe that these three dimensions are complementary and reasonable for the study of price competition between hotels.

From Section 3, it can be seen that β_{ij} is determined by the own and cross-price effects, i.e. the parameters γ_{jj} and γ_{ji}. Standard economic theory predicts that the sign of the cross-price effect is positive. Intuitively, the increase in the price of rivals, *ceteris paribus*, leads to an increase in demand. And the reverse is true for the own-price effect, i.e. an increase in one's own price tends to cause one's own demand to fall. For example, assuming linear demand, $\beta_{ij} = -\frac{\gamma_{ij}}{2}\gamma_{jj} > 0$. Therefore, economic theory predicts the sign of the estimated θ parameter to be positive.

In this estimation strategy, it is possible to identify the most relevant dimensions for hotel pricing decisions. Following the framework of Rezvani and Rojas (2020), we can empirically test the different dimensions of competition directly and see which one or ones are best suited to explain the variation in hotel data in Germany. If the $\mathbf{W}_t$ matrix is constructed on the basis of considering the geographical distance between hotels, and the θ estimate is positive and significant, then this provides us with evidence that the geographical dimension is relevant to competition between hotels. The same is true for the reputation and quality dimensions. To test the different dimensions, and also to differentiate between competition being local or global (regardless of the dimension), we made the same matrices as for Rezvani and Rojas (2020), it is worth noting that the diagonal is always zero:

- Global Competition
  - Benchmark - $\mathbf{W}^B$: entries in matrix are equal to 1.
  - Geographic - $\mathbf{W}^G$: entries are the inverse of the Euclidean distance (km) between the hotels.
  - Reputation - $\mathbf{W}^R$: entries are the inverse of the Euclidean distance of the number of reviews and rating average (normalized) between the hotels.
  - Geo-Reputation - $\mathbf{W}^{GR}$: entries are the inverse of the Euclidean distance of Reputation and distance (km), both normalized, between hotels.

- Local Competition
  - Geographic - $\mathbf{W}^{G^k}$: entry i,j is 1 if j is one of the k closest hotels geographically to i.
  - Quality - $\mathbf{W}^{Q}$: entries are 1 if hotels i and j have the same number of stars (0 to 5).
  - Reputation - $\mathbf{W}^{R^k}$: entry i,j is 1 if j is one of the k closest hotels reputationally to i.
  - Geo-Reputation - $\mathbf{W}^{GR^k}$: entry i,j is 1 if j is one of the k closest hotels (GR dimension) to i.

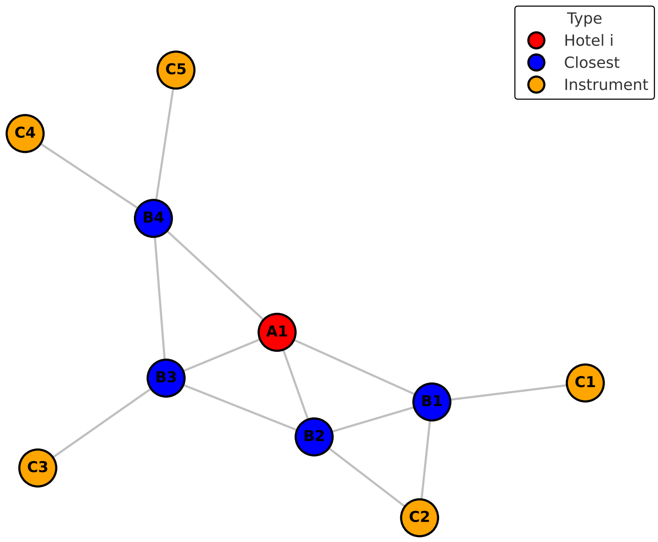
With these matrices, it is possible to test the nature of hotel competition in Germany. In addition, we calculated k to be equal to 4,7 and 10. We limited ourselves to these k because of the computational demand to calculate the matrices, and subsequently, the instruments needed to estimate them. The matrices that test competition are global because they take into account all the prices of the day in their estimation, but with varying weights. The same is true for reputation. Meanwhile, matrices that take localized competition into account completely ignore the prices of hotels that are not part of the matrix. For example, when considering a 3-star hotel, the quality matrix is made such that the right-hand-side price is an average of the prices of the other 3-star hotels, without taking into account distances (all same-quality hotels have equal weights). Of course, we assume that competition is only internal to the city (or region), and this is one of the limitations of our estimation[3].

For the local estimations that consider k hotels, we created a set of instrumental variables along the lines of Rezvani and Rojas (2020), which followed the idea behind Pinkse et al. (2002). The instrument could have been the sum of the characteristics of rival hotels; however, considering hotel fixed effects and the lack of extensive data on hotel characteristics, we chose not to pursue this approach. Instead, we used the same instrumental variable creation rule as Rezvani and Rojas (2020). For hotel i, for example, the right-hand side of the equation takes into account the k hotels closest to hotel i. It is reasonable to assume that these k prices are endogenous since they are subject to the same demand shocks as hotel i and are close to each other. Therefore, to instrument this, we use the prices of other hotels such that (i) they are one of the k closest hotels of i's neighbors, and (ii) are far from hotel i, which is not one of i closest k hotels, and (iii) is not i. Figure 1 illustrates an example where hotel i is A1 and each hotel has a maximum of 4 neighbors connected by a line. The instruments are the prices from yellow hotels, which, assuming that competition is local, the prices of the yellow

[3]This means that price variations in Munich, for example, are not considered in Berlin. We think this is reasonable for now, but it is something that could be explored in further research.

hotels are independent of the price of hotel i (exclusion) but is correlated with the price of the blue hotels (relevance).

Figure 1: Instrumental Variable Illustration



Finally, our aim is to investigate the effect of the 9-Euro Ticket event on the nature of the competition. To this end, we adjusted Equation 21, adding the interaction of the $\mathbf{W}_t\mathbf{p}_t$ term with the event dummy, and we used panel data with fixed-effects to be able to control for Hotel FE and Date FE. The equation that estimates the best-response and the variation of θ during the event is given by:

$$\mathbf{p}_t = \boldsymbol{\delta} + \theta(\mathbf{W}_t\mathbf{p}_t) + \theta'(\text{Event}_{9\text{euro},t} \times \mathbf{W}_t\mathbf{p}_t) + \boldsymbol{\varepsilon}_t \tag{22}$$

In order to estimate this equation, we have to construct a new set of instrumental variables. For this, we take the previously defined instrumental variable and interact it with the event dummy. In other words, for each estimation conducted via 2SLS, we have two sets of instruments for the two endogenous variables.

4.2 Data

The novel dataset was collected during the months of August and September 2022 to investigate the effect of the 9-Euro ticket on competition between hotels. The dataset was created by using a web scraper to collect information on the cheapest rooms in various hotels for a range of regions in Germany and other European cities. We set out to investigate the price 1 day before, 3 days before, and 7 days before the check-in date. The first day we have a date for is August 10, 2022, until September 30, 2022. Subsequently, the same web scraper was executed again in order to capture the price variations of the same hotels in Germany for the year 2023, where we have the first day as August 9, 2023, and the last day as October 1, 2023. Table 1 illustrates the information obtained during the web scraping process.

Table 1: Description of Fields

Field	**Description**
scrape day	The date the data was scraped
days until check in	The number of days until check-in
hotel id	A unique identifier for each hotel (based on the hotel link)
country	The country where the hotel is located
region	The region where the hotel is located
hotel location	The location of the hotel
hotel link	The URL of the hotel's page on Booking.com
hotel name	The name of the hotel
hotel rating	The rating of the hotel
hotel price	The price of the hotel for the night
hotel volume of reviews	The number of reviews of the hotel
room type	The type of room offered
other text	Additional information about the room
occupancy info	The occupancy information for the room
hotels in area	The number of hotels in the area
check in date	The check-in date
page number	The page number of the search results
hotel number on page	The position of the hotel on the search results page
html	The HTML code for the hotel's page on Booking.com

The cities we collected data from were: Berlin, Frankfurt, Hamburg, Köln, and München in Germany; Lyon and Paris in France; London and Manchester in the UK; and Copenhagen, Vienna, and Prague in other countries. We also collected data from hotels in 4 different regions: Ruhr and Rhein-Neckar Region in Germany; Katowice Region in Poland; and Lille Region in France. To enhance the dataset, we geocoded all the unique hotels in the dataset. This was achieved through the utilization of the

OpenCage and Google Maps API services, processing the addresses of the hotels to ascertain the geographical coordinates. Additionally, for the purpose of establishing a reference point within each city[4], coordinates denoting the center of each city were acquired from the city coordinate listings on the respective GeoHack site. In a final step towards enhancing the dataset, the OpenStreetMap API was used to calculate the distance from each hotel to the nearest public transport stop. Table 2 illustrates the number of hotels we have data on for each European region, and Figure 2 displays a map of the hotels we are investigating in Germany.

Table 2: Regional Distribution.

(a) Non-German Regions

Region	**Hotels**
Copenhagen	215
London	1236
Lyon	489
Vienna	467
Manchester	199
Paris	2451
Prague	646
Region: Katowice	185
Region: Lille	399
Total (Non-German)	6287

(b) German Regions

Region	**Hotels**
Berlin	624
Frankfurt	705
Munich	693
Hamburg	471
Cologne	546
Region: Rhein-Neckar	425
Region: Ruhr	438
Total (German)	3902

To facilitate a descriptive analysis, we filtered the data 7 days before check-in. We did this because it is the filtering of the database that is the most complete in terms of hotel information, since the closer to the check-in date, the fewer options are available. In addition, we made a selection to remove outliers of 1% of the price both upwards and downwards by region, outliers in distance to closest public transport stop (25 km), and we removed outliers of location, hotels that are the top 1% percentile of distance (43 km from the city center). This last selection means that we exclude hotels that, despite being listed as belonging to a city, are too far away to be considered relevant. We have maintained this filtering for the rest of the article, except at times when we explicitly inform that the filtering has been modified.

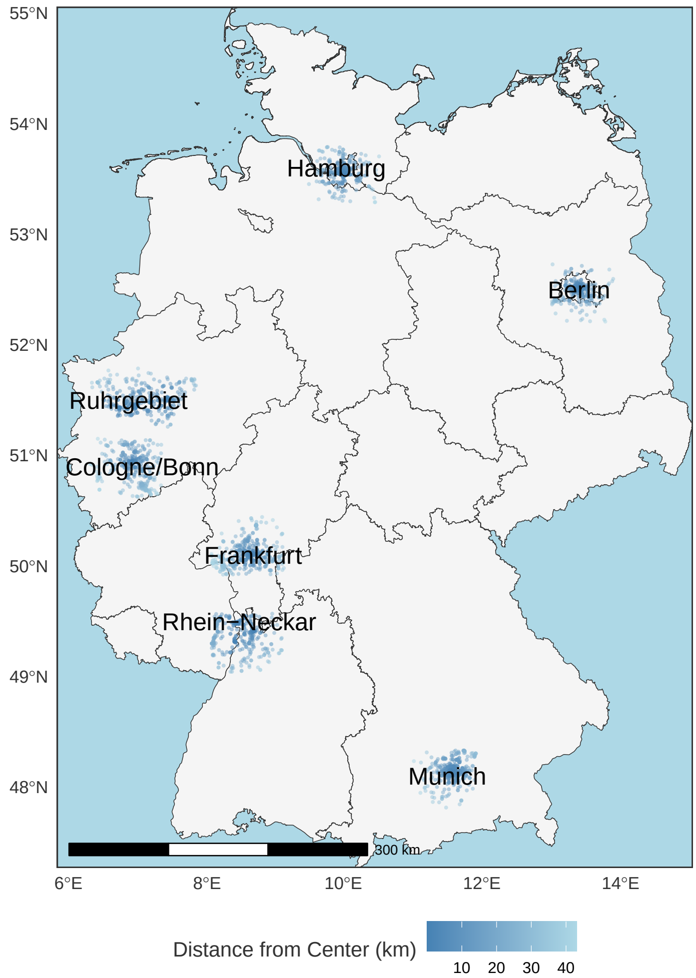
Here we present some descriptive statistics of our data. Figure 2 shows the spatial distribution of the hotels used in our database. With it, it is possible to see that we have price information from the main German hubs and the data also has a reasonable spatial variation. Table 3 shows descriptive statistics for the main variables used in our

[4]For the regions, we selected the largest cities in each region and did the same process of defining the center of each city.

analysis. The price has been capped to eliminate erroneous offers of around 1 dollar that are sometimes published on Booking.com, with a minimum of 26 euros and an average of 163 euros per night. In addition, we are deflating German hotel prices from 2023 to 2022 prices using the Consumer Price Index calculated by the Federal Statistical Office of Germany[5]. The rating captures user evaluations for each hotel, with a range from 0 to 10, with an average of 7.9, accompanied by the rating volume, which is the number of times a room has been evaluated. The number of stars is present in almost all the hotels, and for those where it was not, we considered a group consisting of 0 stars. The Daily Number of Hotel's variable is our proxy for the Capacity of the day, which adds up the number of rooms available on a given night. Finally, the variables distance from the city center and distance from the nearest transport stop capture time-constant information for each hotel. After removing 1% of the outliers, the study is limited to hotels up to 43 km from the respective city center and hotels no more than 25 km from a public transport stop. However, most hotels are relatively close to a transport stop, as the average is 2.5 km.

[5]We are using the August 2022 index as a reference, Part 11 - Restaurants and accommodation services. See more at `https://www.destatis.de/DE/Themen/Wirtschaft/Konjunkturindikatoren/Basisdaten/vpi003a.html?nn=213544`.

Figure 2: Map - Distribution of Hotels in Germany



When we captured the data for the analysis in 2022, we also collected information on hotels in other European countries to serve as a control group. Here, we do some descriptive analysis and estimate some variables to test the validity of using other European hotels as a control group. Figure 3 shows information aggregated by date and country on the median price of each night. We aggregated between France, Germany, the UK, and others (Prague, Vienna, and Poland). The price trend between France, Germany, and Others is similar to each other, with obvious spikes, especially during weekends. In addition, it is clear to see that the median prices of German hotels are considerably lower than those of French or British hotels, and this is also true when taking the average rather than the median. The figure on the right shows the number of listings on Booking for each night and country. We can see a fairly high negative correlation between price and the number of listings. This is evidence that using this variable as a proxy for Capacity is reasonable[6]. In the Appendix, we present

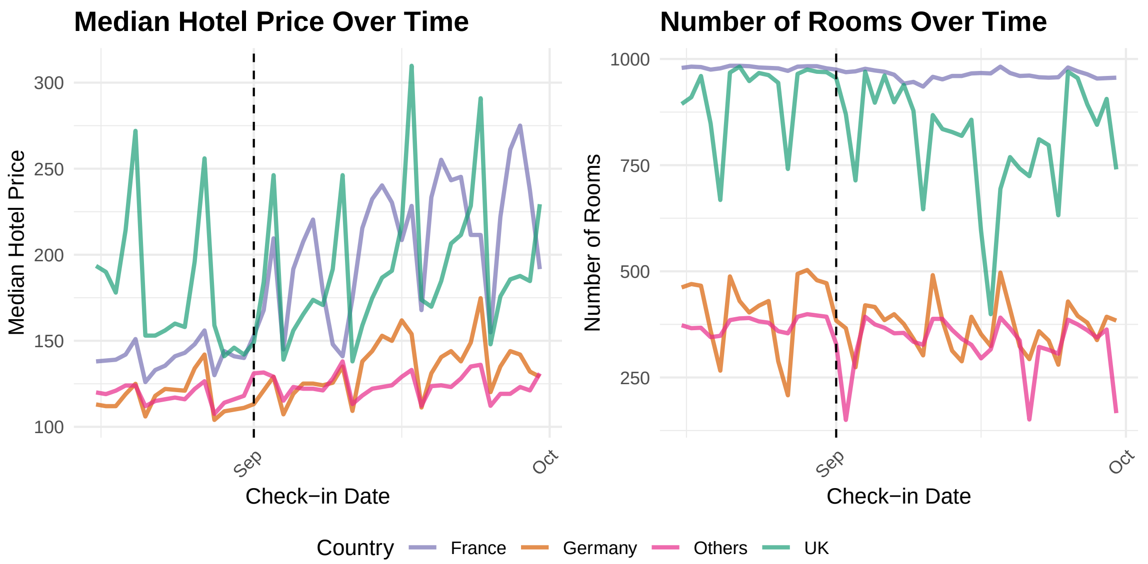
[6]The visual exception is in the case of France. Paris, for example, has thousands of hotels. However, Booking.com limits itself to less than 1000 hotels per night, hence we have not captured all the Parisian hotels. Therefore, the correlation between price and capacity is weaker and because of this and other

the correlations for all regions. They range from -0.091 in Rhein-Neckar to -0.45 for Berlin (with the exception of Lille, which has a slight positive correlation). German cities range from -0.30 to -0.45.

Table 3: Summary Statistics

Variable	N	Mean	Std. Dev.	Min	Pctl. 25	Pctl. 75	Max
Hotel Price	432967	163	107	26	94	194	830
Hotel Rating	432382	7.9	0.83	1.6	7.5	8.5	10
Ratings Volume	432967	1583	1928	1	434	1970	25114
Stars	432967	2.9	1.4	0	3	4	5
Daily Number of Hotels	432967	495	255	27	318	556	990
Distance to Center (km)	432967	9.2	9.4	0.047	2	14	43
Distance to Closest Stop (km)	432967	2.5	5.4	0	0.074	0.93	25

Figure 3: Median Price and Listings by Country

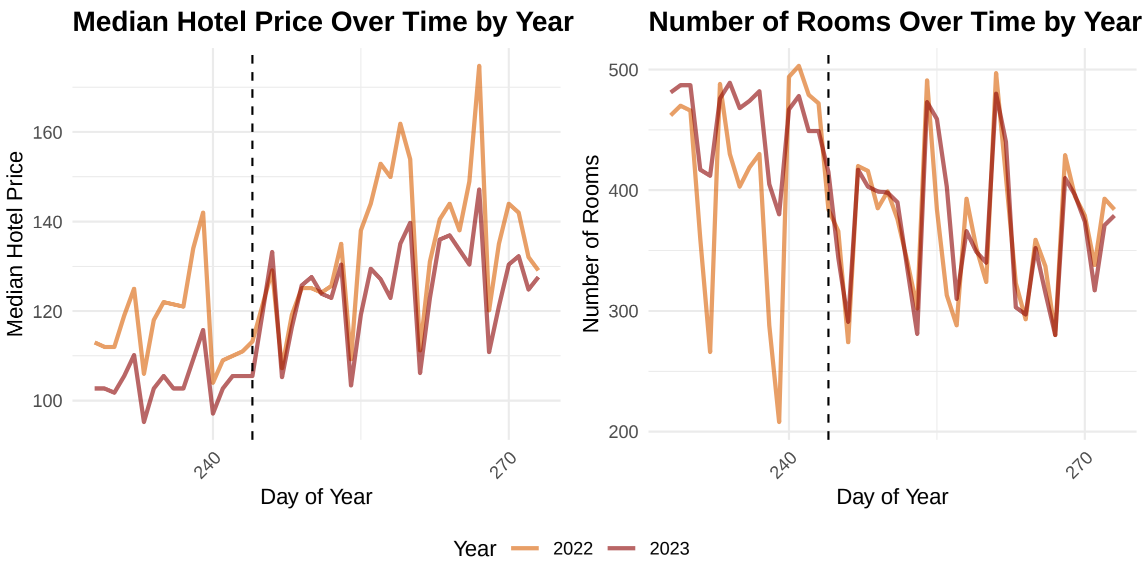


In Figure 4, we present the median prices and aggregate capacity for Germany over the same period between 2022 and 2023. It is worth noting that we have slightly offset the data. The comparison here is made using the Day of the Year, and we matched the days of the year 2023 one day earlier. This is to be able to compare, for example, September 1, 2023 (Friday) with September 2, 2022 (Friday). This offsetting served as a better comparison mainly because of the trend in the data. Adjusting for the day of the week to be the same, at the expense of slightly losing the exact comparison between the day of the year, is an adjustment that we believe is reasonable in order to seek a

factors, we later decided not to put too much emphasis on the other European countries as a control group.

better understanding of how hotels and consumers behave and to have a proper control group.

Figure 4: Median Real Price and Listings by Year - Germany



Figures 5 and 6 essentially split the previous figure by region. Specifically in Figure 5, it can be seen that using the offset strategy causes price peaks to fall on the same day of the year. Events in the cities (concerts, Olympics, sports games, etc.) explain the few spikes that occur for Frankfurt, Berlin, and Munich in 2022, where there is a price spike, causing the median price to exceed 500 euros in Munich[7]. However, these price peaks are captured by Figure 6, where exactly on these outlier days there is a huge variation in the number of rooms available. In other regions such as Rhein-Neckar, Hamburg, and Cologne, the behavior of median prices and capacity in the years 2022 and 2023 is extremely similar. This provides visual evidence that the overall effect of the 9-Euro Ticket on hotel pricing may not have been as strong. In the next subsection, we do a preliminary analysis to try to estimate the overall effect of the 9-Euro Ticket on hotel prices during the event.

[7]We also performed the analysis of this work considering a dummy that captures the days in which there is an outlier, i.e. peak days, and the conclusions of the results remained the same.

Figure 5: Median Real Price and Listings by Year - All German Regions

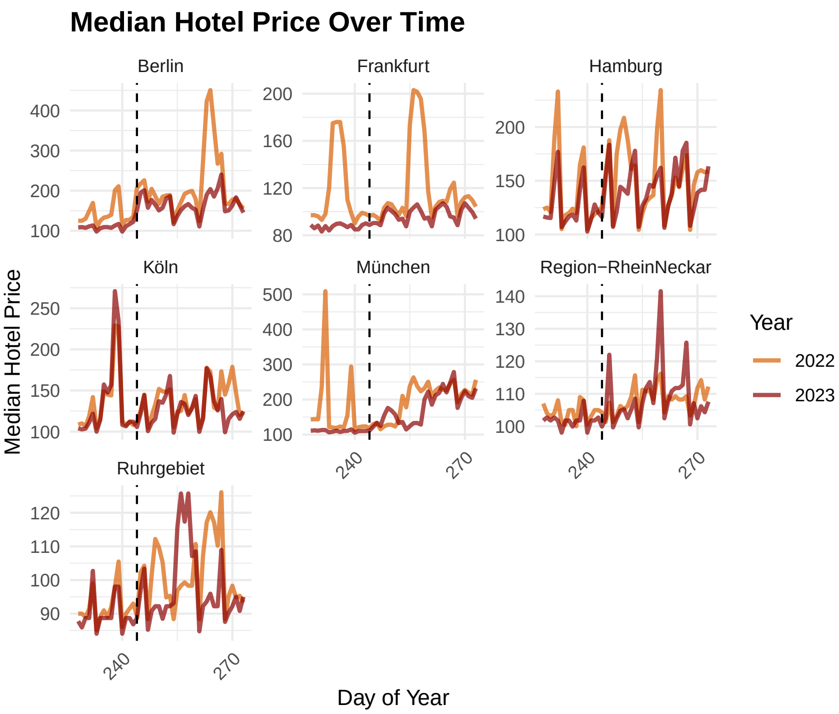
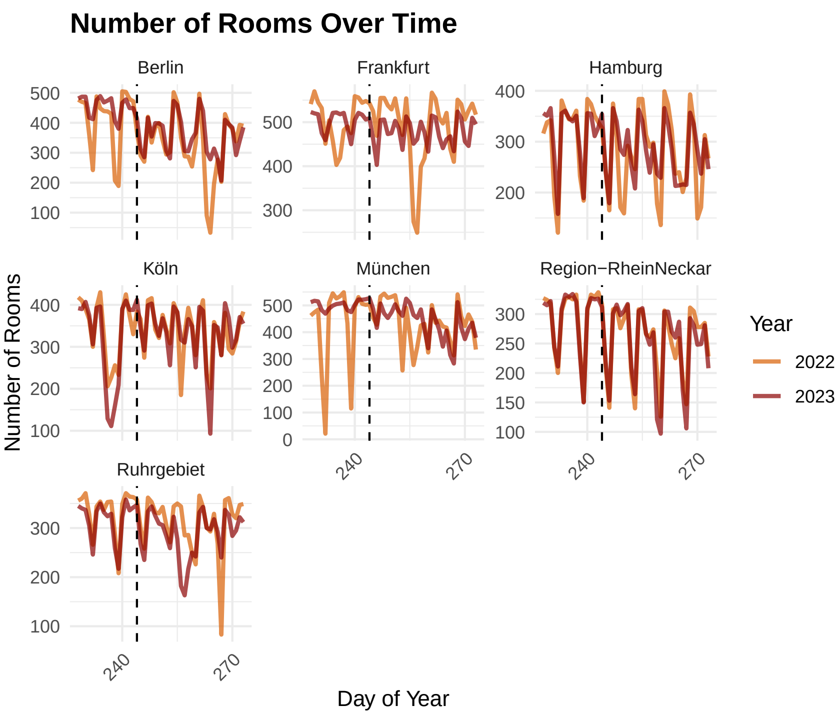


Figure 6: Available Rooms - All German Regions



4.2.1 Preliminary Effect of the 9-Euro Ticket

In this section, we do a simple Differences-in-Differences (DiD) exercise to estimate the overall effect of the 9-Euro ticket on the price of hotels in Germany in August 2022.

Our hypothesis is that, because of the discount on transportation in Germany, there will naturally be a greater demand for hotels in Germany during the month of August 2022. Consequently, it is direct to expect that hotel prices in Germany will be higher at that time. However, we believe that there is an additional effect: the hotels have a higher value, which means that, for the same level of demand, the hotel sector is able to sell at a higher price.

To measure the gains of the hotel sector in Germany, we estimate a DiD identification, where we use two different control groups: (i) hotels in August and September 2022 from other European cities; and (ii) the same German hotels in August and September 2023. Our estimator is as follows:

$$p_{it} = \delta_i + \delta_t + \beta_1 Q_{it} + \tau \mathbb{I}(i = \text{Ger})\mathbb{I}(t = \text{Aug}) + \epsilon_{it} \tag{23}$$

Where $i \in \{\textit{Ever Treated}, \textit{Never Treated}\}$, t indicates the time period, p_{it} is the lowest price for hotel i on day t, δ_i is the FE (which can be at city level, or even hotel level), δ_t is the FE time, which can be binary, such as either September or August, or daily, and Q_{it} is our proxy for capacity of the day, which aims to control for shocks in demand due to events or other things that could increase the demand for hotels and therefore the price on a given day. The variable is given by the sum of hotels being offered on Booking.com for the given combination of day and city. Finally, the categorization of i as *Ever Treated* is always for hotels in Germany during August 2022, and *Never Treated* otherwise. We believe that with this identification it is possible to separate the effect of a simple increase in demand from an increase in the value of German hotels as a result of the discount on transportation. To investigate this stylized fact more specifically in the break time (the end of the 9-Euro Ticket), we limited the analysis to capture the last two weeks of August and the first two weeks of September. We adjusted the variable *Day of the Year*, offsetting it according to the year, making the FE more meaningful[8]. To essentially capture the break between the end of the 9-Euro Ticket and the start of the period without the transport discount, we limited the analysis to $t = 14$ days before and after the event, which we defined as September 1, 2023 (the first day without the discount)[9].

Table 4 shows the DiD results for Panel A, which uses the data from the other European hotels in 2022 as a control group. We did three regressions, varying the fixed effects between them. The first column is the regression where we capture the variation within the month and region. The effect of the estimated 9-Euro Ticket (estimated τ) is approximately 6 euros. In the second column, we captured the variation between day (Date FE) and Region, obtaining a slightly lower estimated τ of 4.2 euros, and

[8] As previously mentioned, offsetting the *Day of the Year* of 2023 by 1 day makes the comparison more valid, as we are comparing the same time of year and the same day of the week.

[9] We also did the analysis with variations in t (e.g. 10 and 20) and the conclusions remain the same.

finally, in the last column, in the more granular variation with Hotel FE and Date FE, the effect was not considered statistically significant. For the three models estimated, the proxy variable for Capacity is statistically significant at the 1% level, with the expected sign. Marginal increases in the number of rooms in a given night and region cause hotel prices to fall by an average of 0.3 euro cents. We estimated the regression controlling for the interaction between Capacity×Region, which aims to control more for heterogeneity between regions. These preliminary results indicate that there may have been a significant effect of the 9-Euro Ticket, but we investigated our results with some caution.

Table 4: Effect of 9-Euro Ticket - Panel A: 2022 Hotels

Dependent Variable:	Daily Hotel Price		
Model:	(1)	(2)	(3)
Variables			
τ: 9-Euro Event	5.685***	4.034***	0.0530
	(0.9053)	(0.8989)	(0.7093)
Capacity	-0.3234***	-0.3315***	-0.2980***
	(0.0106)	(0.0112)	(0.0092)
Fixed-effects			
Region	Yes	Yes	
Month	Yes		
Date		Yes	Yes
Hotel ID			Yes
Fit statistics			
Observations	161,565	161,565	161,565
R^2	0.31557	0.32090	0.86749
Within R^2	0.07043	0.05392	0.29559

Clustered (Hotel ID) standard-errors in parentheses.
Capacity×Region as controls.
*Signif. Codes: ***: 0.01, **: 0.05, *: 0.1*

Figure 7 is our way of testing for the assumption of parallel trends, necessary for a valid estimation of DiD. In it, we have estimated Event Studies, where the regression is the same as column (3) of the panels, but controlling for lead and lag. The graphs have to be analyzed as follows: the event period runs from -10 to -1, where -1 is the baseline[10]. The 0 shows the end of the treatment day (end of the 9-Euro Ticket). From that day onwards, the ideal and expected is to have statistically indifferent coefficients from zero. For figure (a), which is the case for the other European hotels in 2022, there is no clear differentiation between the 9-Euro Ticket period and the period without it at the level of each hotel. Furthermore, the coefficients are mostly non-zero for the period after the end of the 9-Euro Ticket.

Table 5: Effect of 9-Euro Ticket - Panel B: 2022-2023 German Hotels

Dependent Variable:	Daily Hotel Price		
Model:	(1)	(2)	(3)
Variables			
τ: 9-Euro Event	2.325***	2.349***	1.785***
	(0.5756)	(0.5566)	(0.4597)
Capacity	-0.2501***	-0.4108***	-0.3863***
	(0.0073)	(0.0111)	(0.0091)
Fixed-effects			
Year	Yes	Yes	Yes
Month	Yes		
Region	Yes	Yes	
Day of Year		Yes	Yes
Hotel ID			Yes
Fit statistics			
Observations	142,998	142,998	142,998
R^2	0.20670	0.25216	0.70164
Within R^2	0.18107	0.11150	0.23409

Clustered (Hotel ID) standard-errors in parentheses.
Capacity×Region as controls.
*Signif. Codes: ***: 0.01, **: 0.05, *: 0.1*

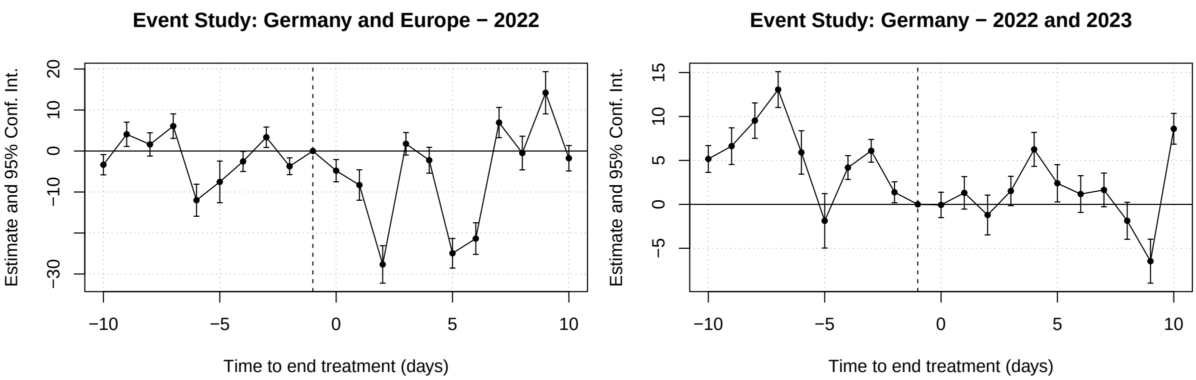
In our second panel, Figure 5 shows Panel B, where we used data from the same hotels in Germany, but one year later as a control group. We performed the same three variations of the estimation. The difference here is that we control for Year FE to capture the difference between treatment and control. The first column is from the variation within month and region, with an estimated value of 2.32 euros. The second column is

[10]In order to have a more concise analysis around the end of the event, -10 means dates less than or equal to 10 days prior to the event, and 10 shows dates greater than or equal to 10 days after the event.

the variation within day and region, with a similar coefficient of 2.35 euros. Finally, the variation within hotel and daily rate, the most granular, is statistically significant at 1% and has a point estimate of 1.78 euros. The Event Study graph in Figure 7 (b) shows Panel B. The estimation is done in a similar way to Panel A, therefore the event takes place before the dotted line. The way to interpret the graph is to check whether there is a significant difference between the event period and the post-event period, and for the post-event, the lag estimate should be zero. For the 9 estimated periods before the end of the 9-Euro Ticket, 8 (7) are positive and significant at the 5% (1%) level. For the 11 periods estimated after the end of the 9-Euro Ticket, 7 (8) are statistically indifferent from zero at 5% (1%) significance.

Figure 7: Event Study

(a) Panel A: 2022 - 5% (b) Panel B: 2022 and 2023 for Germany - 5%



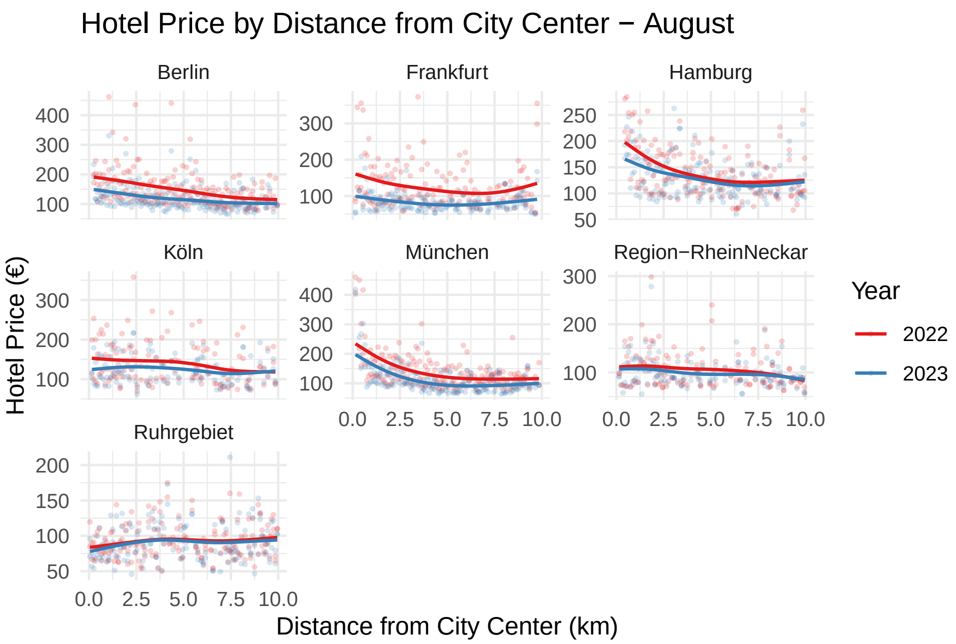
From this preliminary analysis, we aimed to estimate the overall effect of the 9-Euro ticket on hotel prices during the discount event, as well as present the control groups and test which one fits best. From the estimations carried out here in this part, added to the visual analysis of the graphs, we conclude that Panel B is the most suitable to serve as the input for our main analysis. With this panel, we were able to estimate some effect, albeit small, of the 9-Euro ticket on hotel prices in 2022. Controlling for the year, day, and heterogeneity of each hotel, it was possible to see that prices were marginally higher in 2022. It is worth noting that this is an estimate, as we recognize the limitations of our results. Other confounding effects may occur in 2023 which may make our assumption of parallel trends weak, and our control variables may not capture all the demand shocks that occurred during the period. But despite this, we believe that our analysis can still be considered robust and aims to do the best with the data acquired. As a result, from now on our analysis will only use the data from Panel B. This means that we will no longer use the other European hotels in 2022 as a control group, as we believe that they are not the ideal control group for our research.

4.2.2 Distance from City Center and Hotel Pricing

In this section, we aim to investigate the relationship between city centers and hotel prices. Our expectation is that the introduction of the 9-Euro ticket in Germany might influence the behavior of hotel prices in relation to their distance from city centers. Traditionally, hotels closer to city centers charge more due to the higher demand from tourists and business travelers who value proximity to the main attractions and amenities. With an event that causes the cost of transportation to substantially fall, distant hotels can become more attractive, potentially narrowing the price gap and the slope of the price-distance curve. In other words, the price difference between central and more peripheral hotels could decrease, making distant hotels relatively more attractive and possibly leading to a more even distribution of prices across distances.

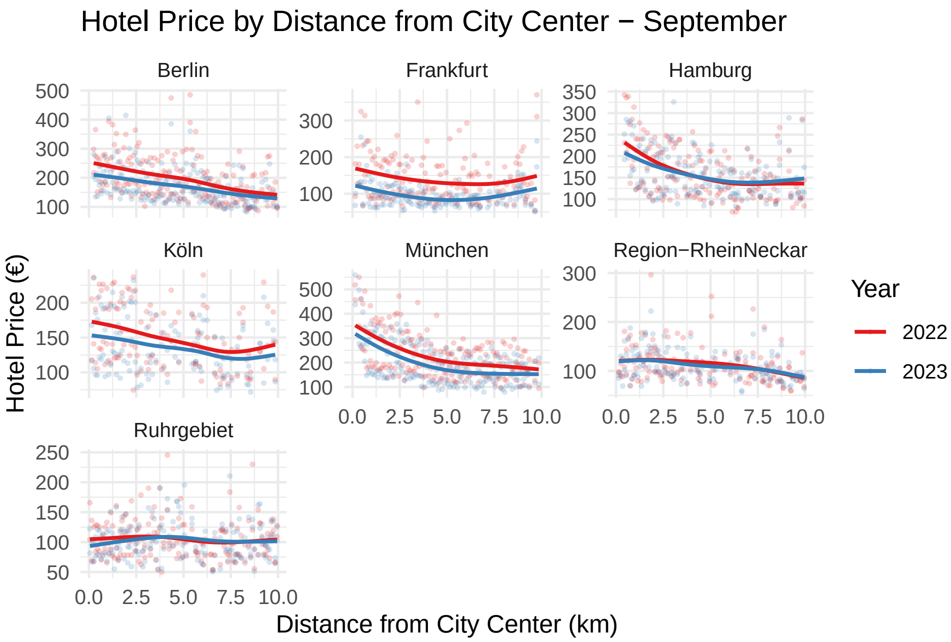
To empirically test this hypothesis, we grouped the data by the interquartile range of prices. We believe that grouping in this way allows us to get a real sense of the price trend by excluding outlier groups. In addition, we grouped the data for hotels within 100 meters of the city center and focused on hotels with a public transport stop within 920 meters (75th percentile). We restricted it to this group of hotels because we believe that the closer a hotel is to a public transport stop, the more pronounced the effect of the discount on transport should be. To make it easier to see the trend, we added a smooth line made by interpolating splines. Figure 8 shows the distribution of prices by distance for the 7 regions for the same period of August 2022 and August 2023, and Figure 9 shows the same, but for the month of September 2022 and September 2023.

Figure 8: Interquartile Range of Real Prices - August



However, the graphs in Figures 8 and 9 suggest that this effect was not as pronounced as might have been predicted. The expectation was for less price sensitivity in August 2022, indicated by a flatter spline curve. However, for some regions, prices in 2023, without the ticket event, decrease less sharply, potentially indicating less price sensitivity to distance. Cities like Frankfurt, Munich, and Berlin have a 2022 curve very similar to the 2023 curve, only vertically shifted. This can be explained by the demand shocks that occurred on specific days in these cities. Despite this, the price x distance curve is very similar between 2022 and 2023 for all cities. Looking at the curves for September, we can see that the trend that happened with Frankfurt, Munich, and Berlin was carried over to September, so the 9-Euro Ticket is not the reason for the divergence between the years. For regions such as Ruhrgebiet and Rhein-Neckar, the curve is practically flat. This can be explained by the fact that these are regions where there are many small towns and more differentiated hotels, and this may mean that there is no pronounced effect of the event[11]. Overall, this descriptive analysis of the data provides some evidence that the effect of the 9-Euro ticket was not as strong in changing the general price behavior based on distance from the city center.

Figure 9: Interquartile Range of Real Prices - September



In terms of numbers, Table 6 shows the price-distance correlations grouped by month. The expectation is to have a negative correlation for all regions, which is true for all cases except the Ruhrgebiet in August. With the 9-Euro Ticket event, the correlation in August 2022 is expected to be less negative. This would mean that the price would tend to be less sensitive to distance. According to the table, when comparing the column

[11]Consumers could be more inclined to go driving to these hotels, making the effect of the 9-Euro Ticket less pronounced.

for August 2022 with August 2023, this only happens for Frankfurt (-0.12 in Aug. 2022 and -0.13 in Aug. 2023) and Munich (-0.31 in Aug. 22 and -0.39 in Aug. 23). For cities like Berlin, Hamburg and Cologne, the correlation increased even more. This potentially indicates a reverse effect: hotels in the city center raised their prices more to capture the additional demand coming from the 9-Euro Ticket event. Looking at September, all the changes in the comparison between 2022 and 2023 were of the same direction and similar magnitude to August, with the exception of Munich (which had the opposite direction in September) and Ruhr-Gebiet (which had a different magnitude and sign).

Table 6: Correlation Table: Hotel Real Price and Distance from City Center

Region	Aug. 2022	Aug. 2023	Sept. 2022	Sept. 2023
Berlin	-0.27	-0.23	-0.27	-0.27
Frankfurt	-0.12	-0.13	-0.09	-0.13
Hamburg	-0.28	-0.20	-0.31	-0.21
Cologne	-0.17	-0.09	-0.25	-0.17
Munich	-0.31	-0.39	-0.36	-0.35
Rhein-Neckar	-0.18	-0.18	-0.20	-0.22
Ruhrgebiet	0.08	0.09	-0.08	-0.04

This section sought to present the methodology, the dataset, and some preliminary analysis regarding the general effect of the 9-Euro ticket on hotel prices, using a Differences in Differences exercise and price-distance correlation analysis to test some hypotheses. It was observed that the increase in demand due to the transport discount in August 2022 could be a reason for the rise in hotel prices. However, the analyses indicate that this effect was not as disruptive as expected on consumers' sensitivity to the distance of hotels from the center. This result suggests the potential greater demand from tourists during the 9-Euro event had a moderate effect on prices, without significantly altering price behavior in relation to distance. The estimates also show that the increase of hotel prices during the event, when controlling our proxy to hotel capacity, is consistent and significant, indicating marginal increases in hotel prices that are not explained by changes in quantity could be attributed to the 9-Euro Ticket. Overall, this analysis serves as a basis for more detailed investigations carried out in the next section, where we aim to study the nature of hotel competition in relation to their rivals, study which type of dimension is more important for this competition, and whether it is localized or global.

5 Empirical Analysis

In this section, we estimate Equation 22. Our objective is multifaceted: we want to identify which dimensions of competition between hotels best explain the data during the period of August and September 2022 and 2023, as well as to identify how the 9-Euro Ticket event potentially affected this behavior. To this end, we estimate best-response functions derived in Section 3.1. The estimated θ parameter can be considered as a metric of how strong the strategic interactions between firms are, and a positive θ would mean that hotels have a price as a strategic complement variable. This means that a hotel's optimal strategy given the price increases of its rivals would be to also increase prices.

To understand the potential effect of strategic price complements in the case of hotels, take as an example a situation in a market (combination of city and day) where there is a tourist event or popular attraction (such as a soccer match, concert, etc.) that attracts a significant increase in the number of additional visitors. Hotels close to popular attractions or transportation hubs, such as airports and train stations, may find that their competitors, also in advantageous locations, raise their prices due to the increased demand. If θ is positive and significant, this suggests that these hotels see each other's price increases as a cue to also raise their prices, with the aim of maximizing profits while demand is high. This potential strategic complementarity in pricing means that hotels are not only responding to general market demand but are also sensitive to how their competitors are setting prices, leading to a collective upward adjustment in the market. We will also use this intuition in the way we think about the instructional variables used.

The intuition behind the instrumental variables chosen involves taking advantage of the price dynamics of hotels that are influential but not directly competitive in relation to hotel i. The approach assumes that competition between hotels is localized and that price dynamics given demand shocks occur in a domino effect manner. When selecting prices from hotels that are among the closest options for a hotel's neighbors, but not for the hotel itself, and which are relatively far from hotel i, the aim is to capture exogenous price influences. This strategy seeks to minimize the direct impact of competition and the possible simultaneity bias in price setting. We tested whether the instruments are relevant and have the expected sign using a t-test and F-test, and the relevance condition was met in all cases estimated.

For all the regressions estimated here, we used Hotel ID, Day of the Year, and Year Fixed-Effects. In addition, we chose not to estimate the proxy variable of the number of rooms on a given day. This is because the variable is potentially endogenous and we do not have a good instrument for it. Prices vary at the room/hotel level, but the total number of rooms is the same for each region, and the strategy used for the instrumental variable would not work. Furthermore, the rivals' price variable would already be capturing the potential effects of demand shocks that the capacity variable could be controlling for, thus adding the variable would bring potential problems to our estimation[12].

In Table 7 we have estimated the core matrices calculated via OLS. The matrix $\mathbf{W}^B$ is our baseline, where for all markets, rivals' prices have equal weight in the matrix. To compare the results and investigate which dimensions of competition best explain the variation in prices over the period investigated, we compare the R^2 of each model. Our baseline has a R^2 of 0.74, which is higher than in the case of $\mathbf{W}^R$. Then comes $\mathbf{W}^Q$ with a value of 0.76, $\mathbf{W}^{GR}$ with 0.77 and, finally, $\mathbf{W}^G$ with 0.8. In all models, the coefficient of θ is positive and significant at 1% level. Our 9-Euro Ticket dummy is negative and significant at the 5% level in the $\mathbf{W}^G$ and $\mathbf{W}^Q$ matrices. In general, although it is OLS, the initial results indicate that geographical distance tends to explain more of the variation in prices, followed by quality and then reputation. Furthermore, the θ' (9-Euro Event Dummy) is negative in all cases, potentially indicating the lower price sensitivity of rival hotels during the event.

Table 7: Core Matrices Regressions (OLS)

	Daily Hotel Price				
	$\mathbf{W}^B$	$\mathbf{W}^G$	$\mathbf{W}^R$	$\mathbf{W}^Q$	$\mathbf{W}^{GR}$
	(1)	(2)	(3)	(4)	(5)
Estimated θ	0.9315***	0.8565***	0.8659***	0.9225***	0.9063***
	(0.0137)	(0.0076)	(0.0352)	(0.0139)	(0.0111)
θ': Estimated θ $\times$ 9-Euro Event	-0.0038	-0.0068**	-0.0053	-0.0087**	-0.0027
	(0.0037)	(0.0034)	(0.0060)	(0.0042)	(0.0036)
Observations	233,004	233,004	232,904	233,004	232,904
R^2	0.74405	0.80265	0.73953	0.75775	0.76847
Within R^2	0.35951	0.50615	0.34832	0.39377	0.42071

Clustered (Hotel ID) standard-errors in parentheses.

Fixed Effects: Hotel ID, Day of Year and Year.

In Table 8 we made the estimations considering the k closest hotels, where k is 4, 7, or 10. The table is structured in parts. The first part considers geographical distance,

[12] We did estimations considering the Capacity variable and also Capacity $\times$ Region, but the results is that the Capacity coefficient is practically positive in all scenarios, indicating potential endogeneity problems with this regression.

the second considers online reputation, and the third considers geographic-reputation. Regarding the columns, (1) is $k = 4$, (2) $k = 7$, (3) $k = 10$ and (4) is the element-wise multiplication of the core matrix with $k = 10$[13]. The estimated θ's tends to be slightly lower than in the core matrices' from Table 7, and in all cases, the θ' is also negative, with most of them being significant at the 5% level. It is difficult to compare the θ's in different regressions because they are constructed differently. However, one potential explanation for the difference in magnitude when considering only nearby hotels versus all hotels is that those nearby tend to have more similar prices, which makes individual price variations less sensitive than when considering all prices in a market.

In all three parts of Table 8, we see a progressive improvement in R^2 as k increases. In the first part of the table, where the geographic dimension regressions were estimated, the coefficient of θ' was significant and negative in all the cases estimated. In addition, the R^2 in column (3) explained most of the variation in the data, with approximately 81% of the variation in prices being explained by it. In the second part of the table, reputation is not as strong as geographic distance in explaining price variation, since all the models show R^2 below the benchmark, but θ' remained negative in all four cases, and significant in (2) and (3). Finally, the third part of the table shows the combination of geographic-reputation. The R^2 are above the benchmark, but always below its corresponding estimate in the first part of the table, the one that only takes geographical distance into account. From this table, we conclude that geography explains the variation in the data better, and adding the interaction with reputation tends to reduce the quality of the explanation of the variation. Also, increasing k tends to increase R^2, but combining it with the core matrix causes the coefficient to fall.

In Table 9, we consider the role of quality interacting with other dimensions. The table is divided into three parts and three columns. Column (1) considers the core matrix multiplied element-wise with the quality matrix, (2) considers the $k = 10$ matrix multiplied with quality, and (3) considers the $k = 10$ matrix multiplied by the core and also by the quality matrix. In the first and last parts of the table, the tendency is for the coefficient of determination in column (2) to be the highest, and column (1) to be the lowest. For the second part of the table, R^2 worsened with each incremental change. Also, when compared to Table 7, the interaction of the core matrices with quality causes the R^2 to fall in all three cases. In nearly all estimations, the estimated coefficient for θ' is negative, but only in the first part is it significant in all columns. This leads us to consider that based on the OLS results, geographical distance is the best explainer of price variation, and the one that most captured a significant difference for the 9-Euro Ticket period.

[13] We chose to show the case of $k = 10$ because it is the one with the best R^2 in all cases among the estimated options. Also, the interaction of the core matrix with the k-closest matrix yields a weighted average of the prices of the k-closest hotels, where the weight is given by the core matrix.

Table 8: K-Closest Hotels Regressions (OLS)

	Daily Hotel Price			
	$\mathbf{W}^{G^4}$	$\mathbf{W}^{G^7}$	$\mathbf{W}^{G^{10}}$	$\mathbf{W}^{G \odot G^{10}}$
	(1)	(2)	(3)	(4)
Estimated θ	0.6946***	0.7623***	0.7859***	0.7461***
	(0.0065)	(0.0061)	(0.0061)	(0.0060)
θ': Estimated θ $\times$ 9-Euro Event	-0.0084**	-0.0108***	-0.0124***	-0.0140***
	(0.0040)	(0.0038)	(0.0037)	(0.0038)
Observations	233,004	233,004	233,004	233,004
R^2	0.78257	0.80000	0.80629	0.79595
Within R^2	0.45589	0.49951	0.51524	0.48938
	$\mathbf{W}^{R^4}$	$\mathbf{W}^{R^7}$	$\mathbf{W}^{R^{10}}$	$\mathbf{W}^{R \odot R^{10}}$
	(1)	(2)	(3)	(4)
Estimated θ	0.5920***	0.7054***	0.7647***	0.7198***
	(0.0095)	(0.0103)	(0.0108)	(0.0109)
θ': Estimated θ $\times$ 9-Euro Event	-0.0073*	-0.0092**	-0.0079**	-0.0028
	(0.0040)	(0.0040)	(0.0040)	(0.0041)
Observations	233,004	233,004	233,004	232,904
R^2	0.70638	0.72759	0.73816	0.73110
Within R^2	0.26523	0.31831	0.34475	0.32722
	$\mathbf{W}^{GR^4}$	$\mathbf{W}^{GR^7}$	$\mathbf{W}^{GR^{10}}$	$\mathbf{W}^{GR \odot GR^{10}}$
	(1)	(2)	(3)	(4)
Estimated θ	0.6586***	0.7310***	0.7599***	0.7385***
	(0.0074)	(0.0067)	(0.0069)	(0.0069)
θ': Estimated θ $\times$ 9-Euro Event	-0.0157***	-0.0138***	-0.0123***	-0.0071*
	(0.0040)	(0.0038)	(0.0037)	(0.0037)
Observations	233,004	233,004	233,004	232,904
R^2	0.76735	0.78730	0.79359	0.78985
Within R^2	0.41781	0.46773	0.48346	0.47421

Clustered (Hotel ID) standard-errors in parentheses.

Fixed Effects: Hotel ID, Day of Year and Year.

Table 9: Quality Augmented Regressions (OLS)

	Daily Hotel Price		
	$\mathbf{W}^{G\odot Q}$	$\mathbf{W}^{G^{10}\odot Q}$	$\mathbf{W}^{G\odot G^{10}\odot Q}$
	(1)	(2)	(3)
Estimated θ	0.6820***	0.6529***	0.6195***
	(0.0149)	(0.0133)	(0.0126)
θ': Estimated θ $\times$ 9-Euro Event	-0.0134***	-0.0113**	-0.0114**
	(0.0033)	(0.0051)	(0.0050)
Observations	233,004	213,146	212,946
R^2	0.70813	0.77149	0.76177
Within R^2	0.26960	0.43725	0.41334
	$\mathbf{W}^{R\odot Q}$	$\mathbf{W}^{R^{10}\odot Q}$	$\mathbf{W}^{R\odot R^{10}\odot Q}$
	(1)	(2)	(3)
Estimated θ	0.7715***	0.5671***	0.5337***
	(0.0146)	(0.0108)	(0.0103)
θ': Estimated θ $\times$ 9-Euro Event	0.0003	-0.0075	-0.0053
	(0.0033)	(0.0053)	(0.0052)
Observations	232,904	215,005	214,318
R^2	0.71542	0.69591	0.69053
Within R^2	0.28799	0.26518	0.25250
	$\mathbf{W}^{GR\odot Q}$	$\mathbf{W}^{GR^{10}\odot Q}$	$\mathbf{W}^{GR\odot GR^{10}\odot Q}$
	(1)	(2)	(3)
Estimated θ	0.8234***	0.6128***	0.5987***
	(0.0127)	(0.0156)	(0.0146)
θ': Estimated θ $\times$ 9-Euro Event	-0.0045	-0.0092*	-0.0081
	(0.0033)	(0.0055)	(0.0055)
Observations	232,904	215,050	214,329
R^2	0.72668	0.75180	0.74939
Within R^2	0.31616	0.39207	0.38608

Clustered (Hotel ID) standard-errors in parentheses.

Fixed Effects: Hotel ID, Day of Year and Year.

In Table 10, we estimate the regressions via 2SLS with two sets of instrumental variables for each estimation, explained in Section 4.1. The table is again divided into three parts, where the first part takes into account geographical distance, the second online reputation, and the third the interaction of the two. In each part, columns (3) and (4) show the interaction with quality. We fixed $k = 10$ as it is the one that best explains the variation in prices[14]. When interacting with quality, the coefficient of determination decreases and the p-value increases compared to the model without quality. In general, the results of the IV regression are similar to those of the OLS. The difference is a slightly higher estimated θ, which is to be expected, given that the instrument aims to isolate variations for only the price variations of rivals' neighbors, which by the domino effect hypothesis is exogenous. This, therefore, caused the estimated coefficient to be marginally higher in absolute terms, an indication that the instrument may have acted to separate out the bias in the estimates. Furthermore, an θ as generally high is a good indicator that the competition between the hotels has strong strategic complementarity.

In the results of Table 10, all the estimated θ's are significant. Furthermore, following the trend of the last few tables, the regressions when competition by geographical distance is considered are the ones that best explain the variation in prices. In the first part of the table, both columns (1) and (2) have a significant θ' coefficient, which captures the effect of the 9-Euro Ticket. This means that both considering the 10 closest hotels and giving different weights to the 10 closest hotels make the θ' significant and negative. In other words, the results suggest that hotels were slightly less sensitive to their competitors' pricing strategies when the ticket was in effect. In the second part of the table, only in (1) is θ' significant. In the third part of the table, we have (1) and interestingly (3) as significant, which is the model that considers the three dimensions of competition together. Overall, the results are coherent when considering the sign and magnitude of the coefficients, providing evidence for the conclusions. Table A3 reproduces the regressions conducted but using the log of the price on both sides of the equation. The results are similar and the table provides robustness to the conclusions. When considering the log, we see a slightly higher R^2 and more cases where the θ' is significant.

[14]We also conducted the analysis with other k's, and obtained the same conclusions regarding the magnitude, sign, and significance of the estimates, but with a slightly lower R^2 when compared with $k = 10$.

Table 10: 10-Closest Hotels and Quality Regressions (IV)

	Daily Hotel Price			
	$\mathbf{W}^{G^{10}}$	$\mathbf{W}^{G\odot G^{10}}$	$\mathbf{W}^{G^{10}\odot Q}$	$\mathbf{W}^{G\odot G^{10}\odot Q}$
	(1)	(2)	(3)	(4)
Estimated θ	0.8846***	0.8754***	0.8914***	0.8862***
	(0.0068)	(0.0069)	(0.0081)	(0.0083)
θ': Estimated θ $\times$ 9-Euro Event	-0.0120***	-0.0138***	-0.0056	-0.0060
	(0.0037)	(0.0038)	(0.0048)	(0.0050)
Observations	232,246	232,263	173,926	174,056
R^2	0.80320	0.79029	0.76346	0.74679
Within R^2	0.50747	0.47516	0.42291	0.38227
	$\mathbf{W}^{R^{10}}$	$\mathbf{W}^{R\odot R^{10}}$	$\mathbf{W}^{R^{10}\odot Q}$	$\mathbf{W}^{R\odot R^{10}\odot Q}$
	(1)	(2)	(3)	(4)
Estimated θ	0.9542***	0.9562***	0.9660***	0.9635***
	(0.0134)	(0.0139)	(0.0171)	(0.0174)
θ': Estimated θ $\times$ 9-Euro Event	-0.0074*	-0.0020	0.0002	0.0042
	(0.0040)	(0.0041)	(0.0052)	(0.0054)
Observations	232,994	232,887	190,274	189,793
R^2	0.72967	0.71695	0.65670	0.63865
Within R^2	0.32352	0.29184	0.19274	0.15161
	$\mathbf{W}^{GR^{10}}$	$\mathbf{W}^{GR\odot GR^{10}}$	$\mathbf{W}^{GR^{10}\odot Q}$	$\mathbf{W}^{GR\odot GR^{10}\odot Q}$
	(1)	(2)	(3)	(4)
Estimated θ	0.8399***	0.8331***	0.8409***	0.8372***
	(0.0075)	(0.0077)	(0.0092)	(0.0093)
θ': Estimated θ $\times$ 9-Euro Event	-0.0103***	-0.0054	-0.0103**	-0.0074
	(0.0036)	(0.0037)	(0.0051)	(0.0050)
Observations	233,004	232,904	196,317	195,884
R^2	0.79142	0.78671	0.73825	0.73217
Within R^2	0.47804	0.46637	0.36578	0.35155

Clustered (Hotel ID) standard-errors in parentheses.

Fixed Effects: Hotel ID, Day of Year and Year.

The analysis presented in this section aimed to explore the strategic interactions between hotels through price competition, particularly in the period around the 9-Euro Ticket event. In addition, the analysis aimed to investigate which of the different dimensions of competition, namely geographical proximity, reputation, and quality, most explain the variation in prices for the period. The estimates indicate a positive relationship in all the models for θ. This is evidence that there is a high strategic complementarity in the determination of prices between hotels. Furthermore, based on the estimations, the dimension that most explained the variation in the data was geography. Based on the OLS results, it was followed by quality, and finally, online reputation had its R^2 marginally below the benchmark.

According to the results of both the OLS and IV models, geographical proximity was the factor that most explained the variation in the data. Limiting the matrix to the 10 closest hotels increased the R^2, suggesting that hotels tend to focus on local rather than global competition. The dummy for the 9-Euro Ticket event was mostly negative. During the event, hotels were less sensitive to rivals' prices. This may indicate that the competition may have been relaxed during the event that potentially increased demand for hotels, making pricing relatively more independent in price competition, as captured by the model estimated.

5.1 Robustness

In this section we present our robustness tests. The tables are in the Appendix A. Table A2 shows the estimates considering more than one matrix in the same regression. Each column shows a core matrix among geography, reputation, and geo-reputation. We have not considered quality alone, as it is being estimated together with the other matrices. The first row of results is the estimated coefficient for the $\mathbf{W}$ core matrix. The second line is for the quality matrix. Finally, we have the interaction of the core matrix with quality, and then we present the coefficients for the three interactions with the 9-Euro dummy. Table A9 shows a variation of multiple matrices that were estimated via 2SLS. In this second table, columns (1) and (2) show the results without considering quality, while (3) and (4) show the interaction with quality. The results are in the same direction as our main results. This means that the importance of geography is relatively greater than that of other matrices. Furthermore, when the model that also estimates geo-reputation or quality, is considered, the R^2 drops. Finally, the 9-Euro Ticket dummy is mostly negative, but sometimes not significant.

Table A3 presents the results estimating the log of prices. The results are in the same direction as our results previously described. Table A4 shows the F-test of the IV estimates for both the level and log models. The relevance condition is met in all cases. Subsequently, Tables A5, A6 and A7 present the estimations via 2SLS where instead of considering Hotel ID fixed-effect, we estimate using hotel characteristics. The coefficient of determination is lower in all cases, but θ remains positive and significant, and θ' remains mostly negative. Furthermore, the results with geography presented in Table A5 are the ones with the best R^2, corroborating our results. Another way of weighting geographical proximity was considered: inverse distance squared. The results of the IV regression are shown in Table A8. The results remain robust to change.

Finally, the last robustness test presented here was to consider other days until check-in. The data set is almost entirely different when considering 3 or 1 day until check-in. This is because hotels can appear or disappear from the database the closer to the check-in date it is considered. We believe that this test is a good way of checking the results, because the closer to the check-in date, the stronger the price effects tend to be. Tables A10 and A11 replicate Table 10 for 3 days and 1 day before check-in, respectively. The general conclusion when comparing the tables is twofold: the closer to the check-in date, (i) the estimated θ's are lower in all cases, and; (ii) the coefficient θ' tends to be more negative and significant. The two effects are in the same direction. That is, hotels tend to be less responsive to rivals' prices the closer they get to the check-in date, and the 9-Euro Ticket exacerbated this effect. This leads to best-response functions being less dependent on rivals, and hotels being relatively more independent.

6 Conclusion

The introduction of the 9-Euro ticket in the summer of 2022 by the German government has served as an interesting opportunity to explore the direct and indirect impacts of policies and incentives that go hand in hand with environmental and sustainability goals and their effects. The specific aim of this study was to investigate the demand shock caused by the 9-Euro and its impact on the German hotel sector. In particular, on competition between hotels in Germany. The thesis set out to explore whether the event had an impact on the nature of competition between hotels, which was measured using three dimensions: geographical distance, online reputation, and quality. It also aimed to measure the impact of the event on the perceived value of hotels.

The preliminary analysis of this work concluded that the 9-Euro Ticket caused a slight increase in hotel prices during the month of August 2022. Furthermore, the empirical analysis revealed that when analyzing hotels through price competition, it concluded that competition has price as strategic complements. Our results suggest that geographical proximity continued to be a significant determinant of price adjustments, corroborating the hypothesis that hotels compete more strongly among their closest rivals and competition is more localized. In addition, hotels adjusted their pricing strategies during the period in which the 9-Euro ticket was in force. Specific to the 9-

Euro ticket period, the best-response function estimations with the proposed instrument strategy revealed that the price elasticity of hotels in relation to rivals' prices decreased during the period. This suggests that the hotels are priced more independently. Furthermore, the study revealed that although quality and reputation also influenced pricing strategies, their impact was less pronounced than that of geographical factors. This highlights the fundamental role of location in the competitive dynamics of the hotel sector, especially in scenarios of greater public mobility.

This study, however comprehensive, has its limitations. The research was based on publicly available data scraped from the Booking.com platform, which does not capture all dimensions of hotel service quality, customer preferences, and financial data. Empirical models and strategies were limited due to data limitations and the absence of transactional information (e.g. number of rooms sold per night). Another limitation is that the study focused on short-term impacts, making generalizations of trends and behaviors for the medium and long-term limited, both in hotel operations and in consumer choices. In addition, external factors and other potential demand shocks may not have been fully captured by the estimation strategies, and the control groups may not be perfect, but are based on authentic data from an uncontrolled event.

The results of this study have policy implications. Policies such as the 9-Euro Ticket, even if they are not aimed directly at the hotel industry, can in fact have a profound influence on economic sectors such as hospitality. A potentially more robust investigation that could be done is to analyze aggregate sales data per night and per hotel region during the event to quantify how much extra sales the hotels received because of the transport subsidy, in order to measure the overall impact of this indirect support on the German hotel sector. Another future analysis is to investigate the extent to which regions compete with each other for tourists.

A Appendix

Table A1: Correlation Between Hotel Capacity and Prices by Region in 2022

Region	Correlation
Berlin	-0.449
Frankfurt	-0.420
Hamburg	-0.372
Köln	-0.348
Kopenhagen	-0.199
London	-0.242
Lyon	-0.090
Manchester	-0.363
München	-0.427
Paris	-0.279
Region-Katowice	-0.118
Region-Lille	0.052
Region-RheinNeckar	-0.088
Ruhrgebiet	-0.162
Wien	-0.135

Table A2: Multiple Matrices Regressions (OLS)

	Daily Hotel Price		
Core Matrix:	$\mathbf{W}^G$	$\mathbf{W}^R$	$\mathbf{W}^{GR}$
	(1)	(2)	(3)
Core Matrix P	0.8747***	0.2854***	0.8569***
	(0.0151)	(0.0478)	(0.0290)
$\mathbf{W}^Q$ P	0.1186***	0.6824***	0.3156***
	(0.0188)	(0.0357)	(0.0289)
Core $\odot$ $\mathbf{W}^Q$ P	-0.1472***	-0.0064	-0.2652***
	(0.0102)	(0.0176)	(0.0139)
Core P $\times$ 9-Euro Dummy	0.0386***	0.0034	0.1440***
	(0.0139)	(0.0309)	(0.0222)
$\mathbf{W}^Q$ P $\times$ 9-Euro Dummy	-0.0078	-0.0399**	-0.0635***
	(0.0127)	(0.0166)	(0.0143)
Core $\odot$ $\mathbf{W}^Q$ P	-0.0352***	0.0291	-0.0767***
	(0.0091)	(0.0200)	(0.0162)
Observations	233,004	232,904	232,904
R^2	0.80477	0.76172	0.77323
Within R^2	0.51144	0.40384	0.43264

Clustered (Hotel ID) standard-errors in parentheses.

Fixed Effects: Hotel ID, Day of Year and Year.

Table A3: 10-Closest Hotels and Quality Log Regressions (IV)

	Daily Hotel Log(Price)			
	$\mathbf{W}^{G^{10}}$	$\mathbf{W}^{G\odot G^{10}}$	$\mathbf{W}^{G^{10}\odot Q}$	$\mathbf{W}^{G\odot G^{10}\odot Q}$
	(1)	(2)	(3)	(4)
Estimated θ	0.9552***	0.9522***	0.9523***	0.9497***
	(0.0061)	(0.0063)	(0.0086)	(0.0087)
θ': Estimated θ $\times$ 9-Euro Event	-0.0013***	-0.0016***	-0.0013**	-0.0014**
	(0.0005)	(0.0005)	(0.0006)	(0.0007)
Observations	232,246	232,263	173,926	173,926
R^2	0.81781	0.80280	0.76416	0.74549
Within R^2	0.44656	0.40099	0.29721	0.24159
	$\mathbf{W}^{R^{10}}$	$\mathbf{W}^{R\odot R^{10}}$	$\mathbf{W}^{R^{10}\odot Q}$	$\mathbf{W}^{R\odot R^{10}\odot Q}$
	(1)	(2)	(3)	(4)
Estimated θ	1.011***	1.024***	1.033***	1.036***
	(0.0103)	(0.0104)	(0.0132)	(0.0135)
θ': Estimated θ $\times$ 9-Euro Event	-0.0018***	-0.0012**	-0.0018***	-0.0012*
	(0.0005)	(0.0005)	(0.0006)	(0.0007)
Observations	232,994	232,887	190,274	189,793
R^2	0.77241	0.75996	0.70259	0.68525
Within R^2	0.30876	0.27127	0.13722	0.08863
	$\mathbf{W}^{GR^{10}}$	$\mathbf{W}^{GR\odot GR^{10}}$	$\mathbf{W}^{GR^{10}\odot Q}$	$\mathbf{W}^{GR\odot GR^{10}\odot Q}$
	(1)	(2)	(3)	(4)
Estimated θ	0.9288***	0.9267***	0.9021***	0.8954***
	(0.0063)	(0.0064)	(0.0142)	(0.0143)
θ': Estimated θ $\times$ 9-Euro Event	-0.0014***	-0.0007	-0.0009	-3.95×10^{-6}
	(0.0005)	(0.0005)	(0.0006)	(0.0006)
Observations	233,004	232,904	196,317	195,884
R^2	0.81531	0.81063	0.76189	0.75688
Within R^2	0.43903	0.42507	0.29653	0.28241

Clustered (Hotel ID) standard-errors in parentheses.

Fixed Effects: Hotel ID, Day of Year and Year.

Table A4: F-test from First Stage IV Regressions

	Daily Hotel Price			
	$\mathbf{W}^{G^{10}}$	$\mathbf{W}^{G \odot G^{10}}$	$\mathbf{W}^{G^{10} \odot Q}$	$\mathbf{W}^{G \odot G^{10} \odot Q}$
	(1)	(2)	(3)	(4)
Level Models				
F-test (1st stage), $\hat{\theta}$	282,202.1	254,352.5	83,080.2	144,512.3
F-test (1st stage), $\hat{\theta}'$	1,139,260.0	1,135,729.5	507,945.7	581,571.5
Log Models				
F-test (1st stage), $\hat{\theta}$	185,901.9	165,212.7	50,157.5	46,491.7
F-test (1st stage), $\hat{\theta}'$	31,625,767.6	31,279,746.9	10,940,830.4	10,589,447.4
	$\mathbf{W}^{R^{10}}$	$\mathbf{W}^{R \odot R^{10}}$	$\mathbf{W}^{R^{10} \odot Q}$	$\mathbf{W}^{R \odot R^{10} \odot Q}$
	(1)	(2)	(3)	(4)
Level Models				
F-test (1st stage), $\hat{\theta}$	207,736.0	167,979.5	51,814.0	46,392.5
F-test (1st stage), $\hat{\theta}'$	1,452,366.1	1,332,710.1	469,863.4	446,987.9
Log Models				
F-test (1st stage), $\hat{\theta}$	171,716.0	140,833.7	39,339.0	35,401.2
F-test (1st stage), $\hat{\theta}'$	46,591,705.48	42,856,324.3	13,556,986.	12,757,985.6
	$\mathbf{W}^{GR^{10}}$	$\mathbf{W}^{GR \odot GR^{10}}$	$\mathbf{W}^{GR^{10} \odot Q}$	$\mathbf{W}^{GR \odot GR^{10} \odot Q}$
	(1)	(2)	(3)	(4)
Level Models				
F-test (1st stage), $\hat{\theta}$	371,560.5	368,546.9	101,572.0	99,017.5
F-test (1st stage), $\hat{\theta}'$	1,745,821.6	1,759,743.0	639,467.2	625,744.0
Log Models				
F-test (1st stage), $\hat{\theta}$	282,893.7	274,283.6	72,418.3	71,571.3
F-test (1st stage), $\hat{\theta}'$	49,328,918.2	49,913,900.4	15,603,903.0	15,267,640.8

Clustered (Hotel ID). Fixed Effects: Hotel ID, Day of Year and Year.

Table A5: Geographical Distance - Hotel's Characteristics IV Regression

	Daily Hotel Price			
	$\mathbf{W}^{G^{10}}$	$\mathbf{W}^{G\odot G^{10}}$	$\mathbf{W}^{G^{10}\odot Q}$	$\mathbf{W}^{G\odot G^{10}\odot Q}$
	(1)	(2)	(3)	(4)
Estimated θ	0.8681***	0.8498***	0.8672***	0.8605***
	(0.0099)	(0.0102)	(0.0108)	(0.0107)
θ': Estimated θ $\times$ 9-Euro Event	-0.0139***	-0.0163***	-0.0053	-0.0059
	(0.0044)	(0.0046)	(0.0065)	(0.0066)
Hotel Rating	19.05***	19.48***	16.00***	16.32***
	(0.8208)	(0.8312)	(0.8583)	(0.8586)
Rating Volume	-0.0023***	-0.0023***	-0.0015***	-0.0014***
	(0.0004)	(0.0004)	(0.0004)	(0.0004)
Center Distance	0.0095	0.0040	-0.2726***	-0.2985***
	(0.0819)	(0.0854)	(0.0867)	(0.0901)
Closest Stop	-0.2167	-0.1791	0.0064	0.0284
	(0.1401)	(0.1451)	(0.1419)	(0.1456)
Stars	7.033***	7.226***	-0.6908	-0.8608*
	(0.4454)	(0.4609)	(0.4462)	(0.4604)
Observations	232,074	232,091	173,791	173,921
R^2	0.62204	0.59836	0.61349	0.59258
Within R^2	0.59102	0.56540	0.57911	0.55634

Clustered (Hotel ID) standard-errors in parentheses.

Fixed Effects: Day of Year and Year.

Table A6: Reputation - Hotel's Characteristics IV Regression

	Daily Hotel Price			
	$\mathbf{W}^{R^{10}}$	$\mathbf{W}^{R\odot R^{10}}$	$\mathbf{W}^{R^{10}\odot Q}$	$\mathbf{W}^{R\odot R^{10}\odot Q}$
	(1)	(2)	(3)	(4)
Estimated θ	0.9298***	0.9152***	0.9311***	0.9209***
	(0.0167)	(0.0175)	(0.0200)	(0.0199)
θ': Estimated θ $\times$ 9-Euro Event	-0.0042	0.0021	0.0047	0.0095
	(0.0049)	(0.0050)	(0.0070)	(0.0070)
Hotel Rating	-0.0102	-0.0411	3.029***	2.950***
	(0.7535)	(0.7569)	(0.7280)	(0.7309)
Rating Volume	-0.0027***	-0.0027***	-0.0011***	-0.0010**
	(0.0004)	(0.0005)	(0.0004)	(0.0004)
Center Distance	-1.207***	-1.223***	-1.132***	-1.150***
	(0.1070)	(0.1080)	(0.0997)	(0.0999)
Closest Stop	0.4174**	0.4628***	0.4476***	0.5019***
	(0.1734)	(0.1757)	(0.1639)	(0.1661)
Stars	7.445***	7.447***	-1.620***	-1.518***
	(0.5160)	(0.5206)	(0.4673)	(0.4677)
Observations	232,822	232,734	190,145	189,680
R^2	0.46577	0.45018	0.43606	0.41846
Within R^2	0.42193	0.40507	0.38426	0.36539

Clustered (Hotel ID) standard-errors in parentheses.

Fixed Effects: Day of Year and Year.

Table A7: Geo-Reputation - Hotel's Characteristics IV Regression

	Daily Hotel Price			
	$\mathbf{W}^{GR^{10}}$	$\mathbf{W}^{GR\odot GR^{10}}$	$\mathbf{W}^{GR^{10}\odot Q}$	$\mathbf{W}^{GR\odot GR^{10}\odot Q}$
	(1)	(2)	(3)	(4)
Estimated θ	0.8293***	0.8173***	0.8360***	0.8318***
	(0.0108)	(0.0111)	(0.0127)	(0.0129)
θ': Estimated θ $\times$ 9-Euro Event	-0.0086*	-0.0026	-0.0064	-0.0018
	(0.0044)	(0.0045)	(0.0064)	(0.0067)
Hotel Rating	3.050***	3.174***	5.653***	5.726***
	(0.6754)	(0.6784)	(0.6605)	(0.6693)
Rating Volume	-0.0010**	-0.0009**	-0.0002	-9.09×10^{-5}
	(0.0004)	(0.0004)	(0.0003)	(0.0004)
Center Distance	-0.2457***	-0.2863***	-0.4605***	-0.4919***
	(0.0858)	(0.0868)	(0.0861)	(0.0872)
Closest Stop	0.3595**	0.4480***	0.4651***	0.5407***
	(0.1450)	(0.1460)	(0.1402)	(0.1415)
Stars	7.219***	7.231***	0.1346	0.2053
	(0.4662)	(0.4676)	(0.4253)	(0.4283)
Observations	232,832	232,751	196,183	195,765
R^2	0.58266	0.57444	0.56770	0.55831
Within R^2	0.54842	0.53953	0.52925	0.51936

Clustered (Hotel ID) standard-errors in parentheses.

Fixed Effects: Day of Year and Year.

Table A8: Geographical (Inverse Distance Squared) IV Regression

	Price		Log(Price)	
	$\mathbf{W}^{G_2^{10}}$	$\mathbf{W}^{G_2^{10} \odot Q}$	$\mathbf{W}^{G_2^{10}}$	$\mathbf{W}^{G_2^{10} \odot Q}$
	(1)	(2)	(3)	(4)
Estimated θ	0.8846***	0.8914***	0.9552***	0.9523***
	(0.0068)	(0.0081)	(0.0061)	(0.0086)
θ': Estimated θ $\times$ 9-Euro Event	-0.0120***	-0.0056	-0.0013***	-0.0013**
	(0.0037)	(0.0048)	(0.0005)	(0.0006)
Fit statistics				
Observations	232,246	173,926	232,246	173,926
R^2	0.80320	0.76346	0.81781	0.76416
Within R^2	0.50747	0.42291	0.44656	0.29721

Clustered (Hotel ID) standard-errors in parentheses.

Fixed Effects: Day of Year and Year.

Table A9: Geographical and Reputation Multiple Matrices (IV)

	Daily Hotel Price			
	(1)	(2)	(3)	(4)
$\mathbf{W}^{G^{10}}$P	0.8121***	0.9018***		
	(0.0135)	(0.1238)		
$\mathbf{W}^{R^{10}}$P	0.1052***	0.1602**		
	(0.0144)	(0.0809)		
$\mathbf{W}^{G^{10}}$P$\times$9-Euro	-0.0249*	0.0206		
	(0.0137)	(0.0290)		
$\mathbf{W}^{R^{10}}$P$\times$9-Euro	0.0146	0.0495*		
	(0.0135)	(0.0295)		
$\mathbf{W}^{GR^{10}}$P		-0.1269		
		(0.1779)		
$\mathbf{W}^{GR^{10}}$P$\times$9-Euro		-0.0759*		
		(0.0456)		
$\mathbf{W}^{G^{10}Q}$P			0.8151***	-0.0049
			(0.0243)	(0.4831)
$\mathbf{W}^{R^{10}Q}$P			0.1044***	-0.6106
			(0.0292)	(0.3947)
$\mathbf{W}^{G^{10}Q}$P$\times$9-Euro			-0.0176	-0.1385
			(0.0280)	(0.2530)
$\mathbf{W}^{R^{10}Q}$P$\times$9-Euro			0.0137	0.1358
			(0.0284)	(0.1144)
$\mathbf{W}^{GR^{10}Q}$P				1.333*
				(0.7606)
$\mathbf{W}^{GR^{10}Q}$P$\times$9-Euro				-0.0008
				(0.3261)
Observations	232,237	232,237	154,706	148,971
R^2	0.80821	0.80103	0.77435	0.66434
Within R^2	0.52002	0.50206	0.46903	0.20867

Clustered (Hotel ID) standard-errors in parentheses.

Fixed Effects: Day of Year and Year. All relevance conditions are fulfilled.

Table A10: 10-Closest Hotels and Quality Regressions - 3 Days Before (IV)

	Daily Hotel Price			
	$\mathbf{W}^{G^{10}}$	$\mathbf{W}^{G\odot G^{10}}$	$\mathbf{W}^{G^{10}\odot Q}$	$\mathbf{W}^{G\odot G^{10}\odot Q}$
	(1)	(2)	(3)	(4)
Estimated θ	0.8748***	0.8658***	0.8809***	0.8774***
	(0.0068)	(0.0068)	(0.0080)	(0.0081)
θ': Estimated θ $\times$ 9-Euro Event	-0.0106***	-0.0111***	0.0030	0.0033
	(0.0034)	(0.0035)	(0.0044)	(0.0045)
Observations	242,213	242,234	181,133	181,262
R^2	0.79600	0.78275	0.75198	0.73542
Within R^2	0.46792	0.43336	0.37319	0.33138
	$\mathbf{W}^{R^{10}}$	$\mathbf{W}^{R\odot R^{10}}$	$\mathbf{W}^{R^{10}\odot Q}$	$\mathbf{W}^{R\odot R^{10}\odot Q}$
	(1)	(2)	(3)	(4)
Estimated θ	0.9449***	0.9453***	0.9639***	0.9581***
	(0.0120)	(0.0123)	(0.0157)	(0.0160)
θ': Estimated θ $\times$ 9-Euro Event	-0.0094**	-0.0023	-0.0042	0.0020
	(0.0037)	(0.0038)	(0.0049)	(0.0050)
Observations	243,018	242,921	198,617	198,051
R^2	0.73262	0.71958	0.65362	0.63594
Within R^2	0.30263	0.26866	0.15680	0.11486
	$\mathbf{W}^{GR^{10}}$	$\mathbf{W}^{GR\odot GR^{10}}$	$\mathbf{W}^{GR^{10}\odot Q}$	$\mathbf{W}^{GR\odot GR^{10}\odot Q}$
	(1)	(2)	(3)	(4)
Estimated θ	0.8395***	0.8302***	0.8483***	0.8406***
	(0.0074)	(0.0075)	(0.0098)	(0.0097)
θ': Estimated θ $\times$ 9-Euro Event	-0.0114***	-0.0050	-0.0040	0.0014
	(0.0034)	(0.0034)	(0.0046)	(0.0046)
Observations	243,038	242,937	204,769	204,252
R^2	0.78612	0.78168	0.72905	0.72315
Within R^2	0.44214	0.43060	0.32051	0.30596

Clustered (Hotel ID) standard-errors in parentheses.

Fixed Effects: Hotel ID, Day of Year and Year.

Table A11: 10-Closest Hotels and Quality Regressions - 1 Day Before (IV)

	Daily Hotel Price			
	$\mathbf{W}^{G^{10}}$	$\mathbf{W}^{G\odot G^{10}}$	$\mathbf{W}^{G^{10}\odot Q}$	$\mathbf{W}^{G\odot G^{10}\odot Q}$
	(1)	(2)	(3)	(4)
Estimated θ	0.8622***	0.8533***	0.8837***	0.8781***
	(0.0070)	(0.0071)	(0.0093)	(0.0096)
θ': Estimated θ $\times$ 9-Euro Event	-0.0136***	-0.0145***	-5.3×10^{-5}	-0.0020
	(0.0032)	(0.0033)	(0.0043)	(0.0044)
Observations	242,181	242,193	181,400	181,526
R^2	0.79482	0.78137	0.74830	0.73297
Within R^2	0.44619	0.40988	0.34228	0.30226
	$\mathbf{W}^{R^{10}}$	$\mathbf{W}^{R\odot R^{10}}$	$\mathbf{W}^{R^{10}\odot Q}$	$\mathbf{W}^{R\odot R^{10}\odot Q}$
	(1)	(2)	(3)	(4)
Estimated θ	0.9296***	0.9307***	0.9511***	0.9511***
	(0.0115)	(0.0118)	(0.0153)	(0.0157)
θ': Estimated θ $\times$ 9-Euro Event	-0.0120***	-0.0040	-0.0062	-0.0005
	(0.0035)	(0.0036)	(0.0044)	(0.0045)
Observations	242,961	242,857	198,475	197,949
R^2	0.73637	0.72367	0.65894	0.63780
Within R^2	0.28873	0.25450	0.14297	0.09127
	$\mathbf{W}^{GR^{10}}$	$\mathbf{W}^{GR\odot GR^{10}}$	$\mathbf{W}^{GR^{10}\odot Q}$	$\mathbf{W}^{GR\odot GR^{10}\odot Q}$
	(1)	(2)	(3)	(4)
Estimated θ	0.8259***	0.8198***	0.8368***	0.8303***
	(0.0072)	(0.0074)	(0.0098)	(0.0097)
θ': Estimated θ $\times$ 9-Euro Event	-0.0128***	-0.0063*	-0.0090*	-0.0033
	(0.0032)	(0.0032)	(0.0052)	(0.0052)
Observations	242,971	242,865	204,742	204,227
R^2	0.78417	0.77952	0.73023	0.72454
Within R^2	0.41769	0.40518	0.30020	0.28593

Clustered (Hotel ID) standard-errors in parentheses.

Fixed Effects: Hotel ID, Day of Year and Year.

Bibliography

Anderson, S. P., J. K. Goeree, and R. Ramer (1997, November). Location, Location, Location. *Journal of Economic Theory 77*(1), 102–127.

Anderson, S. P. and D. J. Neven (1991). Cournot Competition Yields Spatial Agglomeration. *International Economic Review 32*(4), 793–808. Publisher: [Economics Department of the University of Pennsylvania, Wiley, Institute of Social and Economic Research, Osaka University].

Armona, L., G. Lewis, and G. Zervas (2021, June). Learning Product Characteristics and Consumer Preferences from Search Data.

Aydin, E. and K. Kürschner Rauck (2023, July). Public Transport Subsidization and Air Pollution: Evidence from the 9-Euro-Ticket in Germany.

Baum, J. and T. Lant (2003, September). Hits and misses: Managers'(mis) categorization of competitors in the Manhattan hotel industry. *Advances in Strategic Management 20*, 119–156.

Baum, J. A. C. and S. J. Mezias (1992). Localized Competition and Organizational Failure in the Manhattan Hotel Industry, 1898-1990. *Administrative Science Quarterly 37*(4), 580–604. Publisher: [Sage Publications, Inc., Johnson Graduate School of Management, Cornell University].

Becerra, M., J. Santaló, and R. Silva (2013, February). Being better vs. being different: Differentiation, competition, and pricing strategies in the Spanish hotel industry. *Tourism Management 34*, 71–79.

Berry, S., J. Levinsohn, and A. Pakes (1995). Automobile Prices in Market Equilibrium. *Econometrica 63*(4), 841–890. Publisher: [Wiley, Econometric Society].

Biscaia, R. and I. Mota (2013). Models of spatial competition: A critical review*. *Papers in Regional Science 92*(4), 851–871. _eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1111/j.1435-5957.2012.00441.x.

Bissel, M. (2023, August). A Public Transport Ticket that Moved a Country: Assessing the Value of the German 9-Euro-Ticket as a Socio-Technical Experiment. *Findings*. Publisher: Findings Press.

Caplin, A. and B. Nalebuff (1991). Aggregation and Imperfect Competition: On the Existence of Equilibrium. *Econometrica 59*(1), 25–59. Publisher: [Wiley, Econometric Society].

Combes, P.-P. (1997). Industrial Agglomeration under Cournot Competition. *Annales d'Économie et de Statistique* (45), 161–182. Publisher: [GENES, ADRES].

d'Aspremont, C., J. J. Gabszewicz, and J.-F. Thisse (1979). On Hotelling's "Stability in Competition". *Econometrica 47*(5), 1145–1150. Publisher: [Wiley, Econometric Society].

d'Aspremont, C., J. Jaskold Gabszewicz, and J. F. Thisse (1983, January). Product differences and prices. *Economics Letters 11*(1), 19–23.

Farronato, C. and A. Fradkin (2022, June). The Welfare Effects of Peer Entry: The Case of Airbnb and the Accommodation Industry. *American Economic Review 112*(6), 1782–1817.

Feenstra, R. C. and J. A. Levinsohn (1995). Estimating Markups and Market Conduct with Multidimensional Product Attributes. *The Review of Economic Studies 62*(1), 19–52. Publisher: [Oxford University Press, Review of Economic Studies, Ltd.].

Gohl, N. and P. Schrauth (2022). Ticket to Paradise? : The Effect of a Public Transport Subsidy on Air Quality. (50).

Greenhut, J. G. and M. L. Greenhut (1975). Spatial Price Discrimination, Competition and Locational Effects. *Economica 42*(168), 401–419. Publisher: [London School of Economics, Wiley, London School of Economics and Political Science, Suntory and Toyota International Centres for Economics and Related Disciplines].

Hamilton, J. H., J.-F. Thisse, and A. Weskamp (1989, February). Spatial discrimination: Bertrand vs. Cournot in a model of location choice. *Regional Science and Urban Economics 19*(1), 87–102.

Hotelling, H. (1929, March). Stability in Competition. *The Economic Journal 39*(153), 41–57.

Kalnins, A. (2016). Beyond Manhattan: Localized competition and organizational failure in urban hotel markets throughout the United States, 2000–2014. *Strategic Management Journal 37*(11), 2235–2253. _eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1002/smj.2578.

Kalnins, A., L. Froeb, and S. Tschantz (2017). Can mergers increase output? Evidence from the lodging industry. *The RAND Journal of Economics 48*(1), 178–202. Publisher: [RAND Corporation, Wiley].

Kosová, R., F. Lafontaine, and R. Perrigot (2013). Organizational Form and Performance: Evidence from the Hotel Industry. *The Review of Economics and Statistics 95*(4), 1303–1323. Publisher: The MIT Press.

Lee, S. K. (2015, February). Quality differentiation and conditional spatial price competition among hotels. *Tourism Management 46*, 114–122.

Lewis, G. and G. Zervas (2016). The Welfare Impact of Consumer Reviews: A Case Study of the Hotel Industry.

Li, J., S. Netessine, and S. Koulayev (2018, September). Price to Compete ... with Many: How to Identify Price Competition in High-Dimensional Space. *Management Science 64*(9), 4118–4136. Publisher: INFORMS.

Mayer, T. (2000, May). Spatial Cournot competition and heterogeneous production costs across locations. *Regional Science and Urban Economics 30*(3), 325–352.

Mazzeo, M. J. (2002a, November). Competitive Outcomes in Product-Differentiated Oligopoly. *The Review of Economics and Statistics 84*(4), 716–728.

Mazzeo, M. J. (2002b). Product Choice and Oligopoly Market Structure. *The RAND Journal of Economics 33*(2), 221–242. Publisher: [RAND Corporation, Wiley].

McFadden, D. (1978). Modelling the Choice of Residential Location. Cowles Foundation Discussion Paper 477, Cowles Foundation for Research in Economics, Yale University.

Neven, D. J. (1986, January). On Hotelling's competition with non-uniform customer distributions. *Economics Letters 21*(2), 121–126.

Osborne, M. J. and C. Pitchik (1987). Equilibrium in Hotelling's Model of Spatial Competition. *Econometrica 55*(4), 911–922. Publisher: [Wiley, Econometric Society].

Pinkse, J. and M. Slade (2004). Mergers, brand competition, and the price of a pint. *European Economic Review 48*(3), 617–643. Publisher: Elsevier.

Pinkse, J., M. E. Slade, and C. Brett (2002). Spatial Price Competition: A Semiparametric Approach. *Econometrica 70*(3), 1111–1153. Publisher: [Wiley, Econometric Society].

Rezvani, E. and C. Rojas (2020). Spatial price competition in the Manhattan hotel market: The role of location, quality, and online reputation. *Managerial and Decision Economics 41*(1), 49–63. _eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1002/mde.3092.

Salop, S. C. (1979). Monopolistic Competition with Outside Goods. *The Bell Journal of Economics 10*(1), 141–156. Publisher: [RAND Corporation, Wiley].

Tabuchi, T. and J.-F. Thisse (1995, June). Asymmetric equilibria in spatial competition. *International Journal of Industrial Organization 13*(2), 213–227.

Affidavit – Ehrenwörtliche Erklärung

Ich versichere, dass ich die vorliegende Arbeit ohne Hilfe Dritter und ohne Benutzung anderer als der angegebenen Quellen und Hilfsmittel angefertigt und die den benutzten Quellen wörtlich oder inhaltlich entnommenen Stellen als solche kenntlich gemacht habe. Diese Arbeit hat in gleicher oder ähnlicher Form noch keiner Prüfungsbehörde vorgelegen. Ich bin damit einverstanden, dass meine Arbeit zum Zwecke eines Plagiatsabgleichs in elektronischer Form anonymisiert versendet und gespeichert werden kann.

I affirm that this Master's thesis was written by myself without any unauthorised third-party support. All used references and resources are clearly indicated. All quotes and citations are properly referenced. This thesis was never presented in the past in the same or similar form to any examination board. I agree that my thesis may be subject to electronic plagiarism check. For this purpose an anonymous copy may be distributed and uploaded to servers within and outside the University of Mannheim.

Mannheim, den 04.07.2024

`figures/signatur3.png`

Guilherme Schultz Lindenmeyer