

# Online Exchange Markets and Empirical Analysis of Ads: Willhaben, a Primarily C2C Marketplace

COURSE ESSAY  
VU ONLINE MARKET PLACES

MAGDALENA SCHINDL (K1455323)  
FELIX REICHEL (K12008176)

AUGUST 2024

## **Contributional Acknowledgements**

### **Chapter 1 Introduction:**

Main contributor: Magdalena Schindl, Editor: Felix Reichel

### **Chapter 2 Theoretical Part:**

Main contributor: Magdalena Schindl, Editor: Felix Reichel

### **Chapter 3 Empirical Part:**

Main contributor: Felix Reichel, Editor: Magdalena Schindl

### **Chapter 4 Conclusion and Implications:**

Main contributor: Felix Reichel, Editor: Magdalena Schindl

### **Chapter 5 Appendix: Tables and Figures:**

Main contributor: Felix Reichel

### **Submission of Supplementary Materials:**

Code: Felix Reichel

# Contents

<b>1</b>	<b>Introduction</b>	<b>5</b>
<b>2</b>	<b>Theoretical Part: Online Marketplaces - The Case of willhaben.at, a primarily C2C Platform in Austria</b>	<b>6</b>
<b>3</b>	<b>Empirical Part: An Analysis of Willhaben Ads for a Single Product</b>	<b>11</b>
3.1	Introduction . . . . .	11
3.2	Data Sources . . . . .	11
3.3	Models . . . . .	13
3.3.1	Bivariate Simple Linear Regression (SLR) Models . . . . .	13
3.3.2	Stepwise AIC-Best Multiple Linear Regression (MLR) Model	13
3.3.3	Multiple Linear Regression (MLR) Models . . . . .	13
3.3.4	MLR Model with Fixed Effects (FE) for Size, Condition, and Color . . . . .	14
3.3.5	Moran's I Test for Spatial Autocorrelation on Psychological Pricing (9, 90, 99 Prices) . . . . .	14
<b>4</b>	<b>Conclusion and Implications</b>	<b>16</b>
<b>5</b>	<b>Appendix: Tables &amp; Figures</b>	<b>20</b>

# List of Tables

5.1	Multiple Linear Regression Models for Listing Prices on Willhaben. Snapshot sample size: $n = 254$ .	21
5.2	Stepwise AIC-Best Regression Model. Sample size: $n = 826$ .	22
5.3	Multiple Linear Regression Model. Sample size: $n = 826$ .	23
5.4	Multiple Linear Regression Model. Sample size: $n = 826$ .	24
5.5	Fixed Effects Model for <i>hasPsychologicalPricing_i</i>	25
5.6	Moran's I Test Results for Psychological Pricing. Two snapshots: $n$ $= 826$ .	25
5.7	Moran's I Test Results for Log Price. Two snapshots: $n = 826$ .	25

# List of Figures

5.1	Histograms of Key Variables. This figure provides a view of the distribution of most variables, including categorical variables across the dataset. . . . .	26
5.2	Correlation Heatmap. This heatmap visualizes the pairwise correlations between numeric variables in the dataset. . . . .	27
5.3	Spatial Distribution of Psychological Pricing. This figure illustrates the geographical distribution of the psychological pricing binary dummy variable across Austria. . . . .	28
5.4	This figure depicts the nearest neighbor relationships among bike data points using K-Nearest Neighbors (KNN) with $K = 1$ . . . . .	29
5.5	Residual Analysis/Q-Q Plot. . . . .	30

# Chapter 1

## Introduction

According to Kadir et al. (2023) online marketplaces are used to buy and sell products and services, as well as to exchange money and data between users or the platform. Due to the large product selection, low costs and the ease of shopping without physical restrictions as well as the technical possibilities, online marketplaces have grown rapidly Kadir et al. (2023). Online marketplaces are also used in the consumer-to-consumer (C2C) sector and thus offer a broad user group a marketplace, for example for used products. This article focuses on Willhaben.at (2024), a leading C2C marketplace in Austria, as stated by Obersteiner, Schmied, and Pamperl (2023).

The empirical analysis in this course essay centers around the offer ads of **Woom Bikes**, a standardised product which is sold on Willhaben. Through web scraping, a dataset of approximately 826 observations was created, focusing on mid-to-high price segment bicycles, which are characterized by price stability and uniformity as we claim. This analysis aims to create analyse ad listing prices through predictive models using willhaben product listing attributes and using the spatial distribution of one of the product attributes.

Our course essay thus tackles with following two main **research questions**:

1. How does information asymmetry between buyers and sellers affect market efficiency and transaction?
2. How much of the variance of woom bike prices can we explain using a multiple linear regression (MLR) model of observed product characteristics through web-scraping (e.g. size, color, condition and some feature engineering/encoded variables)?

## Chapter 2

# Theoretical Part: Online Marketplaces - The Case of willhaben.at, a primarily C2C Platform in Austria

Willhaben Internet Service GmbH & Co KG was founded in Austria in 2006. Willhaben is a virtual marketplace focused on the exchange between  $n$  buyers and  $m$  sellers. In October 2021, willhaben.at had 82,127,327 visits, 1,510,387,470 page views and 93% of them from Austria according to Österreichische Wirtschaftskammer (2024). Willhaben primarily serves as a classified platform and does not use a specific allocation mechanism such as an auction format for transactions. The platform offers users the opportunity to offer items for sale. Buyers can contact sellers directly to obtain further information and (possibly) negotiate prices. With the help of "PayLivery", willhaben offers a service with a small fee to ensure (secure) transactions. Therefore, this marketplace is more of a direct exchange or matching mechanism than an auction-based system. On the platform, buyers and sellers can be both private and business. The following companies or marketplaces can be named as main competitors: eBay, Vinted, kleinanzeigen.de, offline shops and flea markets as well as private garage sales. The products offered on willhaben are mostly used products, and can be either standardized products or handmade products. Obersteiner et al. (2023) have also taken a closer look at willhaben and describe in their work that willhaben also offers a marketplace for used cars and apartments as well as houses for rent and purchase. These product areas are no longer taken into account in our work. According to Heinemann (2019), regional

comparison platforms such as Check24 and Geizhals can be named as competitors in the e-commerce landscape, offering companies new opportunities for information and purchasing processes about customers. These platforms offer insight into search trends and differ regionally and offer valuable insights for companies to choose the right channels and understand geographical and temporal differences in consumer behavior Heinemann (2019). Hu et al. (2023) describe that information on platforms can have both positive and negative effects. The authors differentiate between seller and buyer markets. The seller's market is defined by the fact that there are more buyers than sellers. According to Hu et al., disclosing product information is an advantage for the platform. The product information helps buyers to evaluate individual products, and this also varies the customers' willingness to pay. In a buyer's market, there are more sellers than buyers. However, information can also be a disadvantage. According to Hu et al. (2023), withholding product information can be an advantage. The article by Hu et al. (2023) describes that it leads to less variation in product ratings and increases the likelihood that buyers have a sufficient willingness to pay. This can also lead to higher overall sales for the platform Hu, Wei, and Cui (2023).

Willhaben.at 2024 provides the seller the following possibilities to describe their products and inform the buyer about our in the empirical part investigated product:

- |                                     |  |
|-------------------------------------|--|
| • title                             | • frame size (explained by the version number) |
| • point of sale, paylivery possible |  |
| • fotos, pictures, images           | • color  |
| • price                             | • condition (used, new, almost like new)       |
| • description                       | • transfer                                     |
| • type of brake                     | • seller: name, location, private or business  |
| • wheel size                        |  |

However, we looked at this possible information from the seller's perspective using the example of a bicycle, as we dealt with this product in more detail in the empirical part. The possible information (product-specific features) also varies depending on the product. Pictures of products, which could serve marketing purposes, are excluded since they would be needed to classify post-scraping, they weren't scraped either.

To register as a buyer or as a private seller, all you need is an email address for



registration and a device with internet access, otherwise there are no restrictions. During registration, the "Terms and Conditions" and a "Captcha" must also be confirmed. Nevertheless, willhaben now offers many functions that distinguish willhaben from the previously mentioned marketplaces. The platform is highly localized and, as already mentioned, focuses mainly on the Austrian market and on used items. This regional focus simplifies transactions and logistics and often allows buyers and sellers to meet in person to pick up the item and thus view the item before buying it. The seller can initially choose how the item is to be sold, such as by pickup, shipping, PayLivery or one of the multiple options. The PayLivery option means that the payment process and also the shipping label are taken over by willhaben and should therefore offer a secure transaction for buyers and sellers. Willhaben charges a service fee for this. In addition, willhaben offers users the opportunity to list items for free, which attracts a large number of private sellers. One of willhaben's business models is to list or highlight products better. The seller pays the fee for this. Because the platform offers a wide range of product categories, from real estate and vehicles to electronics, fashion, sporting goods, books and even jobs, willhaben is a marketplace for many buyers and sellers. Willhaben also allows users to remain anonymous in their transactions, which can be attractive for those who value privacy. In addition, willhaben offers another service through a "Trusted Seller" program that highlights sellers with positive reviews and thus gives buyers more trust. In summary, this service has many advantages for buyers and sellers, which leads to willhaben's good position on the market, which can be deduced from the statistics from Österreichische Wirtschaftskammer [2024](#).

Although information about the products is very important, according to Hu, Wei, and Cui ([2023](#)), there are also other factors that influence the buyer's decision Moriuchi and Takahashi ([2023](#)). Moriuchi and Takahashi (2022) investigated consumer behavior in the consumer-to-consumer (C2C) e-commerce marketplace in Japan, focusing on the role of value, trust, and engagement in buyer satisfaction. The research of Moriuchi and Takahashi (2022) extends the Means-End Chain (MEC) theory by integrating trust, engagement and satisfaction as critical factors. The quality of products affects the value that buyers perceive, the shopping pleasure and also the emotional value. The emotional connection is important for the brand experience. In addition, Moriuchi and Takahashi (2022) found that functional value has a stronger influence on trust in both the intermediary and the seller than emotional value. The authors conclude that buyers primarily look for products that meet their needs, although emotional engagement can also increase satisfaction. Trust in the intermediary, such as a platform like willhaben, is crucial

for customer loyalty and yet it has no direct influence on satisfaction. In contrast, according to Moriuchi and Takahashi (2022), trust in the seller has a direct impact on satisfaction. The authors conclude that this suggests that active interaction between buyers and sellers leads to higher satisfaction, as stated in Moriuchi and Takahashi (2023).

Market liquidity refers to the ease with which assets can be bought or sold in a marketplace without significantly affecting their price. In the context of lateral exchange marketplaces, such as those enabled by platforms like Airbnb or eBay, liquidity is influenced by the ability of these platforms to efficiently match supply and demand. The concept of liquid ownership, which involves temporary access to goods rather than full ownership, is central to understanding these marketplaces. Gleim, Stevens, and Johnson (2023) Platforms like willhaben can benefit from recognizing the distinctions between different types of marketplace models, such as access-based consumption (e.g., renting through platforms) and collaborative consumption (e.g., sharing economy platforms). For willhaben, enhancing market liquidity means ensuring that users can easily access and exchange goods, thereby improving the overall efficiency and attractiveness of the platform. Understanding these nuances can help willhaben better position itself, optimize its platform characteristics, and effectively compete in the evolving digital marketplace landscape. In addition, newer technologies could be adopted by online second-hand marketplaces. According to Kadir et al. (2023), the second-hand market in particular, especially for high-value items such as luxury goods, faces the challenge of establishing trust, transparency and product authenticity. For this reason, the authors propose blockchain technology to address these problems with its secure and immutable ledger system by ensuring the traceability of ownership and the authenticity of products. The OWNTRAD platform, for example, uses blockchain to improve transparency and trust in vintage e-commerce by allowing users to track ownership transactions and ensure the quality and legality of the items sold. According to Kadir et al. (2023), this system not only increases customer satisfaction but also improves the company's profitability by strengthening consumer trust. Willhaben could use similar blockchain-based solutions to expand on the findings of Kadir et al. (2023) regarding the trust and transparency issues on its second-hand marketplace.

Based on our research, we outline the following advantages and disadvantages for buyers and sellers using the willhaben marketplace:

**Advantages for Buyers:** Product range and local focus: Willhaben offers sellers the opportunity to list products with a wide range of properties and places few

restrictions on this. With the local focus on the Austrian market, willhaben ensures that buyers can find items that are relevant to their location, which often leads to lower shipping costs and the opportunity to view items in person or pick them up before purchasing. Ease of use and accessibility: Based on the number of users, we claim that the platform is user-friendly. Only an email address is required for registration, making it accessible to a wide range of users. Price negotiation: Buyers can negotiate prices directly with sellers via an integrated chat solution, which offers the opportunity for a direct exchange. Buyers can also suggest a different price via PayLivery.

**Disadvantages for Buyers:** Information asymmetry: One of the biggest challenges for buyers on willhaben is the possibility of information asymmetry. Since the platform relies heavily on information provided by sellers, product descriptions may contain discrepancies or inaccuracies, leading to uncertainty about the actual condition or value of the item. Buyer protection programs: Unlike some other e-commerce platforms, willhaben does not always offer buyer protection programs. This can make transactions riskier, especially for higher-value items where the risk of fraud or receiving a fake product is higher.

**Advantages for Sellers:** Large user base: With millions of page views and visits, Willhaben is one of Austria's most popular online marketplaces. This large target group increases the visibility of the listed items and improves the chances of a successful sale. Cost-effective selling: For private sellers, listing items on willhaben is generally free. Flexibility in pricing: Sellers have the freedom to set their own prices and descriptions.

**Disadvantages for Sellers:** Competition: In a C2C market, the same product may be offered by different sellers at different prices, which may lead to (local) competition. Sellers may find it difficult to price their items competitively and still achieve a desirable profit margin. Length of the sales process: Private sellers in particular often want to get rid of the products on offer quickly due to a lack of storage space. The sales process can vary greatly in terms of lead time, from the initial inquiry to price negotiation to pickup. Buyer trust: Since buyers are often cautious when purchasing used goods, sellers may have difficulty convincing potential buyers of the quality and authenticity of their products. This can be particularly problematic for high-value items where trust is important.

Willhaben offers significant advantages to both buyers and sellers, such as a large user base and the flexibility to negotiate prices. In addition, people can act as both buyers and sellers. However, willhaben also has to address the challenges of information asymmetry and provide a platform for buyers and sellers.

## Chapter 3

# Empirical Part: An Analysis of Willhaben Ads for a Single Product

### 3.1 Introduction

For the empirical analysis of our online marketplace, we focus on a single product: "Woom Bikes," categorized under Marketplace → Sport / Sports Equipment → Bicycles / Cycling → Bicycles. Initially, data were collected through two web scraper executions, yielding approximately 3,000 observations. After addressing duplicates, missing values, and merging the observations with other datasets, approximately 900 observations remained. Further refinement to remove pricing outliers outside the 1.5 Interquartile-Range (IQR) resulted in a final sample size of 826 observational units. We claim that our product is a suitable observational unit for analyzing ads in the mid-to-high-level price segment, characterized by price stability and product uniformity, varying only in size and color (mostly).

### 3.2 Data Sources

In addition to the web-scraped data of Woom Bikes<sup>1</sup>, we utilized other data sources to augment the initial dataset. Zip codes were provided by the Austrian parcel de-

---

<sup>1</sup>Woom Bikes on a Primarily C2C Online Exchange Marketplace: <https://www.willhaben.at/iad/kaufen-und-verkaufen/marktplatz/fahrraeder-radsport/fahrraeder-4552?sfId=b8725e40-07af-41a5-bb6d-6d32deed8220&rows=30&isNavigation=true&keyword=woom+4>

livery service Post.at<sup>2</sup>, and geocoded data was used to link these zip codes to longitude and latitude coordinates<sup>3</sup>. Additionally, shapefiles of Austria with various grid sizes (1 km, 10 km, and 100 km) were obtained from the European Environment Agency (EEA)<sup>4</sup>. Currently not included in our analysis but potentially valuable are data sets such as the official locations of Woom Bike resellers who sell new bikes to customers, obtainable from a dealer locator website<sup>5</sup>. Other useful data sets might include the official prices of new products from the bicycle manufacturer itself<sup>6</sup> and additional Austria-specific shapefiles, such as those depicting zip code level borders, which would be beneficial for further empirical analysis or investigating spatial and/or spatiotemporal dependencies among observational units whilst controlling for underlying spatial dependencies on a zip code level.

## Descriptive Statistics

Figure 5.1 shows descriptive statistics in the form of histograms for all numeric variables and categorical variables: size (WoomCategory), color, and condition. Price outliers have already been removed, resulting in  $n = 826$  observations. Many variable values are quite imbalanced. For example, note the almost nonexistent observations of 'Orange' bikes. For condition, most bikes are encoded with the category 'as good as new' or 'used,' while only very few observations are marked as 'good.' Furthermore, observe the imbalance in the observed sizes; most bikes fall into size categories 3 or 4. The histograms of zip code and the difference count variables for bikes of the same size and condition might suggest some clustering hinting towards cities, urban areas, or unbalanced assignment of zip codes regarding different population density of municipalities. However, bike ads (offers) seem to be fairly geographically distributed. To provide an overview of descriptive statistics, Figure 5.2 presents a heatmap of all numerical, categorical (color), and ordinal (size and condition) variables. The heatmap reveals no surprisingly strong correlations ( $> 0.7$ ) between any two variables other than expected correlations, such as re-coded variables.

---

<sup>2</sup>Austrian Zip Codes Data: According to ISO standard: <https://www.post.at/en/g/c/postal-encyclopedia>

<sup>3</sup>Geocoded Zip Codes: <https://gist.github.com/PeterTheOne/7135a370b37406e6801d36827e0316cf>

<sup>4</sup>Shapefiles of Austria (Grid 1, 10, and 100 kilometres): <https://www.eea.europa.eu/data-and-maps/data/eea-reference-grids-2/gis-files/austria-shapefile>

<sup>5</sup>All Woom B2C Resellers with Zip Codes and Country: <https://intl-checkout.woom.com/apps/dealerlocator>

<sup>6</sup>Woom Official Website: [https://woom.com/de\\_AT/](https://woom.com/de_AT/)

## Non-Imputed Computed Variables

Note that the numerical variable for modeling logistic costs is currently not calculated using a feasible imputation strategy, which introduces possible biases due to some observations (where the value equals 0, because the unit has not any comparable bike units within a 60-kilometre radius, as noted in Table 5.3).

## 3.3 Models

### 3.3.1 Bivariate Simple Linear Regression (SLR) Models

No bivariate simple linear regression (SLR) model reached statistical significance in predicting the Willhaben ad listing price (*Price*) of the bike.

### 3.3.2 Stepwise AIC-Best Multiple Linear Regression (MLR) Model

Using the *MASS* package in R, the *stepAIC* function was used to run an algorithm that yields a multiple linear regression (MLR) model with the lowest (best) Akaike Information Criterion (AIC). The resulting model is presented in Table 5.2. The Ordinary Least Squares (OLS) coefficients for the intercept and categorical variables (*Size*, *Condition*, and *Color*) are highly statistically significant at the 1 percent level. However, the regression coefficient for the dummy variable *Dealer* remains insignificant. Table 5.2 also displays the F-statistic and various other statistics. This model also reached the highest  $R^2$  of approximately 0.71.

**Residual Analysis:** Figure 5.5 presents the residual analysis graphically, including plots of residuals vs. fitted values, Q-Q plot of residuals, scale-location plot, and residuals vs. leverage.

### 3.3.3 Multiple Linear Regression (MLR) Models

Table 5.1 shows the initial run of MLR models based on a single snapshot ( $n = 254$ ) of ads. One model includes the calculated variable *Logistic Costs*, defined as being inversely proportional to the sum of similar products weighted by distance. This variable might be highly biased as it lacks an imputation strategy, however, it achieved statistical significance in the single snapshot. Including this additional variable in the regression model also increased the proportion of explained variance (Adjusted  $R^2$ ).

### 3.3.4 MLR Model with Fixed Effects (FE) for Size, Condition, and Color

Table 5.5 presents a Fixed Effects (FE) model on statistically significant categorical variables (*Size*, *Color*, *Condition*) for the variable *hasPsychologicalPricing*, which is defined as one if the Willhaben ad listing price contains a 9, 90, or 99 in the price. The OLS regression coefficient for *hasPsychologicalPricing* is statistically insignificant at the 10% level but has the expected sign, as the literature suggests that such pricing strategies lead to increased consumer spending. According to 5.1 the dummy variable for *hasPsychologicalPricing* shifts the predicted prices upwards by cet. par. approximately 3.4%.

### 3.3.5 Moran's I Test for Spatial Autocorrelation on Psychological Pricing (9, 90, 99 Prices)

To test for spatial autocorrelation on the binary dummy *hasPsychologicalPricing*, Table 5 reports the results of Moran's I Test.

The definition of Moran's I, as used by Moran in Moran (1950), is:

$$I = \frac{N \sum_{i=1}^N \sum_{j=1}^N w_{ij} (x_i - \bar{x})(x_j - \bar{x})}{\left( \sum_{i=1}^N \sum_{j=1}^N w_{ij} \right) \sum_{i=1}^N (x_i - \bar{x})^2}$$

where

$$w_{ij}$$

are the spatial weights.

The test statistic for Moran's I is calculated as:

$$z(I) = \frac{I - \mathbb{E}(I)}{\sqrt{\text{Var}(I)}}$$

where  $\mathbb{E}(I)$  is the expected value of Moran's I under the null hypothesis, and  $\text{Var}(I)$  is its variance.

The test hypotheses are:

$$H_0 : \text{No spatial autocorrelation } (I = \mathbb{E}(I))$$

$$H_1 : \text{Presence of spatial autocorrelation } (I \neq \mathbb{E}(I))$$

Furthermore, Cliff and Ord (1970) provides a historical overview of spatial autocorrelation measures, which stem from the field of economic geography and trace

back to Tobler’s first law of geography.

The Moran’s I test using  $k$ -nearest neighbors ( $k = 1$ ) for constructing a spatial lag weight matrix is statistically significant at the 10% level, with a very low positive value, indicating that the observations, on average, slightly fall in the high-high quadrant of a Moran’s I scatterplot. Given that only 120 out of 826 units have a positive value for the *hasPsychologicalPricing* dummy, the slight spatial autocorrelation is not surprising and could be attributed to urban areas (e.g., Vienna). At the more typical 5% significance level, the null hypothesis is not rejected, suggesting there is no significant spatial autocorrelation for the psychological pricing dummy. Ideally, the Moran’s I test should be conducted for each comparable unit, typically at the same or neighboring sizes (assuming no buyer preferences for color or condition); more observations would then be needed for robust statistical inference.

Figure 5.3 illustrates the spatial distribution of bikes with *hasPsychologicalPricing* using different color schemes. Figure 5.4 outlines the potential construction of a spatial weight matrix using  $k$ -nearest neighbors ( $k = 1$ ) for testing potential spatial hypotheses regarding competition and *hasPsychologicalPricing*. However, this approach might be unrealistic if the buyer has specific preferences for the color of the bike.



## Chapter 4

# Conclusion and Implications

According to Hu et al. (2023), information asymmetry can affect market efficiency on platforms such as willhaben. Especially in a used goods market, the seller has much more information about the item, such as the quality, signs of wear, condition or even the previous use and maintenance of the items. This information asymmetry can lead to uncertainty among buyers. Research by Moriuchi and Takahashi (2022) has shown that trust plays a crucial role. It can also lead to higher satisfaction and customer loyalty among buyers on a C2C marketplace. Willhaben is also trying to take other factors into account for example with the "Trusted Seller" function. If the buyer feels that there is an information asymmetry, then the customer is less likely to buy via the platform. Hu et al. (2023) describe in their paper that in a market where there are more sellers than buyers, it can be a disadvantage if sellers share a lot of information with their customers. According to Hu et al. (2023) customers can no longer compare the products. Hu et al. (2023) also point out that information can increase the probability of a sale because the sellers can differ from other sellers and thus leads to differentiation. We conclude from these findings that willhaben is faced with the challenge of providing or being able to obtain the right depth of information in the interests of buyers and sellers. The models used in our empirical analysis reveal significant predictors of Willhaben ad listing prices for one product, capturing much of the variation. One potential area of investigation is the role of spatial competition in psychological pricing strategies, which are also present in modern e-commerce (see Hackl, Kummer, and Winter-Ebmer (2014), for example). We attempted to use a psychological pricing dummy as a proxy variable for spatial competition. However, we did not find statistical significance using Moran's I Test of Spatial Autocorrelation for products, unconditional on their characteristics. To test spatial competition theories

more specifically, one could examine whether the Moran’s I Test of Spatial Autocorrelation yields more statistical significance when applied to other marketing characteristics or to psychological pricing conditional on product characteristics (e.g., size, condition), ideally using a larger dataset with more observations.

Additionally, since the coefficient on logistic cost estimates lost statistical significance when using two snapshots compared to one, implementing a feasible imputation strategy for that variable could improve the robustness of the analysis (see Table 5.1). It is noteworthy that we did not obtain scraped information on whether PayLivery is available for an item, which is a newer feature on the platform. Such a PayLivery variable would set the logistic costs of self-pickup to the same as shipping costs via the PayLivery feature and should therefore be modeled accordingly.

A potential research hypothesis could explore spatial competition in psychological pricing. For existing research questions whether psychological pricing is myth, see Ortega and Tabares (2023). A more specific hypothesis would be: If two sellers of the same bike size of Woom Bikes are close to each other and one uses a psychological price, then their nearest neighbors are also more likely to use a psychological price. A suitable spatial regression model would then also control for underlying spatial differences (e.g. a Spatial Durbin Model) and ideally incorporate temporal aspects, e.g., through spatiotemporal models and more snapshots at equal time intervals. However, quantities cannot be observed directly, as a private seller can simply delist an item, or the listing period on the platform may have ended.

One of the key weaknesses of our paper is the lack of precise data regarding the final purchase price paid by the buyer and whether the product was ultimately bought. This gap in information limits our ability to fully analyze market behavior and pricing strategies. Furthermore, sellers may have differing levels of knowledge about the product’s value or condition, leading to inconsistent pricing. This asymmetry can result in pricing inefficiencies, where buyers either overpay or miss out on potential deals due to a lack of reliable information on the platform.

# Bibliography

- Cliff, A. D. and J. K. Ord (1970). “Spatial Autocorrelation: A Review of Existing and New Measures with Applications”. In: *Economic Geography* 46, pp. 269–292.
- Gleim, M., J. Stevens, and C. Johnson (2023). “Platform Marketplaces: Unifying Our Understanding of Lateral Exchange Markets”. In: *European Journal of Marketing* 57.1, pp. 1–28.
- Hackl, F., M. E. Kummer, and R. Winter-Ebmer (2014). “99 Cent: Price Points in E-commerce”. In: *Information Economics and Policy* 26, pp. 12–27.
- Heinemann, S. (2019). *Werbegeschichte(n)*. 1st ed. Springer Fachmedien Wiesbaden GmbH, ein Teil von Springer Nature.
- Hu, S., M. Wei, and S. Cui (2023). “The Role of Product and Market Information in an Online Marketplace”. In: *Production and Operations Management* 32.10, pp. 3100–3118.
- Kadir, K. et al. (2023). “OWNTRAD: A Blockchain-Based Decentralized Application for Vintage E-commerce Marketplaces”. In: *2023 IEEE Symposium on Wireless Technology and Applications (ISWTA)*, pp. 12–17. DOI: [10.1109/ISWTA58588.2023.10250080](https://doi.org/10.1109/ISWTA58588.2023.10250080).
- Moran, P. A. P. (1950). “Notes on Continuous Stochastic Phenomena”. In: *Biometrika* 37.1/2, pp. 17–23. DOI: [10.2307/2332142](https://doi.org/10.2307/2332142).
- Moriuchi, E. and I. Takahashi (2023). “An Empirical Study on Repeat Consumer’s Shopping Satisfaction on C2C E-commerce in Japan: The Role of Value, Trust and Engagement”. In: *Asia Pacific Journal of Marketing and Logistics* 35.3, pp. 560–581.

Obersteiner, G., E. Schmied, and M. Pamperl (2023). *Grundlagenstudie zu Möglichkeiten und Potenzialen von Re-Use im Möbelsegment: Endbericht im Auftrag des BMK*. 1st ed. Universität für Bodenkultur Wien.

Ortega, A. M. and F. A. Tabares (2023). “Psychological Pricing: Myth or Reality? The Impact of Nine-Ending Prices on Purchasing Attitudes and Brand Revenue”. In: *Journal of Retailing and Consumer Services* 71, p. 103206.

Österreichische Wirtschaftskammer (2024). *Willhaben Dachangebot*. URL: <https://report.oewa.at/basic/online-angebote> (visited on 07/22/2024).

Willhaben.at (2024). *Das größte Anzeigenportal Österreichs*. URL: <https://willhaben.at> (visited on 08/10/2024).

## **Chapter 5**

## **Appendix: Tables & Figures**

Table 5.1: Multiple Linear Regression Models for Listing Prices on Willhaben. Snapshot sample size:  $n = 254$ .

Dependent Variable	Willhaben Ad Listing Price		
	logarithm (log)	Euros (€)	Euros (€)
Size Category 2	0.463*** (0.053)	122.719*** (10.196)	122.633*** (10.700)
Size Category 3	0.540*** (0.049)	147.148*** (8.411)	148.611*** (8.939)
Size Category 4	0.739*** (0.051)	222.415*** (11.200)	223.786*** (11.562)
Size Category 5	0.881*** (0.059)	284.006*** (16.447)	285.935*** (16.956)
Size Category 6	0.816*** (0.066)	250.610*** (13.771)	252.040*** (13.706)
Size Category 7	-0.024 (0.126)	-3.934 (8.863)	-3.934 (9.351)
Size Category 8	0.491*** (0.095)	134.110*** (49.283)	133.110*** (48.851)
Good Condition	0.102* (0.052)	36.055*** (11.513)	36.684*** (11.620)
Used Condition	-0.080*** (0.018)	-26.981*** (5.291)	-26.152*** (5.344)
Dealer $i$	-0.018 (0.034)	-22.672* (13.094)	-22.672* (12.979)
Last 48 Hours $i$	-0.045* (0.023)	-15.366* (8.057)	-17.010** (8.168)
Psychological Pricing $i$	0.034* (0.020)	13.648* (7.747)	13.648* (7.836)
Logistic Costs $i$	/	/	0.651* (0.585)
Constant	5.335*** (0.049)	210.243*** (9.022)	206.266*** (9.702)
Observations	254	254	254
R-squared	0.646	0.647	0.651
Adjusted R-squared	0.628	0.628	0.632

Notes: t-statistics are calculated using Huber-White robust standard errors. Significance levels: \*  $p \leq 0.1$ ; \*\*  $p \leq 0.05$ ; \*\*\*  $p \leq 0.01$ .

Sample size  $n = 254$  out of approximately 1,500 observations due to missing data. This snapshot was obtained by web-scraping the following

URL: <https://www.willhaben.at/iad/kaufen-und-verkaufen/marktplatz/fahrraeder/kinderfahrraeder-4558?keyword=woom>

Variable calculation:

$$\text{logistic\_costs} = \begin{cases} 0 & \text{if total\_count} = 0 \text{ (Note: potential bias due to missing imputation for this variable).} \\ \frac{1}{\text{weighted\_sum}} & \text{otherwise} \end{cases}$$

where  $\text{total\_count} = \text{Number of Products within 0-10 km (i)} + \text{Number of Products within 10-30 km (i)} + \text{Number of Products within 30-60 km (i)}$  and  $\text{weighted\_sum} = \text{Number of Products within 0-10 km (i)} \cdot \frac{1}{10} + \text{Number of Products within 10-30 km (i)} \cdot \frac{1}{30} + \text{Number of Products within 30-60 km (i)} \cdot \frac{1}{60}$ .

Table 5.2: Stepwise AIC-Best Regression Model. Sample size:  $n = 826$ .

	log(price)
Size Category 2	0.472*** (0.018)
Size Category 3	0.543*** (0.016)
Size Category 4	0.745*** (0.017)
Size Category 5	0.850*** (0.027)
Size Category 6	0.837*** (0.020)
Size Category 7	-0.056* (0.029)
Good Condition	0.142*** (0.030)
Used Condition	-0.098*** (0.009)
Color: Blue	-0.080*** (0.026)
Color: Yellow	-0.072*** (0.027)
Color: Green	-0.065** (0.026)
Color: Orange	-0.161*** (0.045)
Color: Red	-0.074*** (0.026)
Color: Violet	-0.033 (0.026)
Dealer $i$	0.012 (0.008)
Constant	5.400*** (0.030)
Observations	826
R-squared	0.718
Adjusted R-squared	0.713
Residual Std. Error	0.106 (df = 810)
F Statistic	137.695*** (df = 15; 810)

*Notes:* t-statistics are based on Huber-White standard errors.  
Significance levels: \* $p \leq 0.1$ ; \*\* $p \leq 0.05$ ; \*\*\* $p \leq 0.01$ .

Table 5.3: Multiple Linear Regression Model. Sample size:  $n = 826$ .

	Price Parsed
Size Category 2	121.165*** (4.809)
Size Category 3	144.205*** (3.767)
Size Category 4	219.943*** (4.680)
Size Category 5	277.963*** (8.130)
Size Category 6	261.302*** (8.164)
Size Category 7	5.253 (5.037)
Good Condition	53.938*** (16.337)
Used Condition	-35.406*** (3.414)
Dealer $i$	5.066 (3.407)
Last48Hours $i$	-0.337 (5.065)
HasPsychologicalPricing $i$	2.865 (4.616)
Number of Same Size in 0-10 km Radius $i$	0.044 (0.044)
Number of Same Size in 10-30 km Radius $i$	-0.027 (0.025)
Number of Same Size in 30-60 km Radius $i$	-0.001 (0.015)
Constant	217.381*** (4.857)
Observations	826
R-squared	0.680
Adjusted R-squared	0.675

*Notes:* t-statistics are based on Huber-White robust standard errors.  
Significance levels:  $*p \leq 0.1$ ;  $**p \leq 0.05$ ;  $***p \leq 0.01$ .



Table 5.4: Multiple Linear Regression Model. Sample size:  $n = 826$ .

	Price Parsed
Size Category 2	120.533*** (4.880)
Size Category 3	143.698*** (3.814)
Size Category 4	219.195*** (4.700)
Size Category 5	276.562*** (8.106)
Size Category 6	260.235*** (8.200)
Size Category 7	4.943 (4.784)
Good Condition	54.776*** (16.184)
Used Condition	-35.241*** (3.412)
Dealer $i$	4.538* (2.742)
Last48Hours $i$	-0.505 (5.089)
HasPsychologicalPricing $i$	2.895 (4.633)
Logistic Costs	-0.504 (1.949)
Constant	217.430*** (4.509)
Observations	826
R-squared	0.679
Adjusted R-squared	0.675

*Notes:* t-statistics are based on Huber-White robust standard errors.  
Significance levels: \* $p \leq 0.1$ ; \*\* $p \leq 0.05$ ; \*\*\* $p \leq 0.01$ .

Table 5.5: Fixed Effects Model for *hasPsychologicalPricing\_i*

<i>Dependent Variable:</i>	
<i>log_price</i>	
<i>hasPsychologicalPricing_i</i>	3.841 (0.115)
<i>Fixed Effects:</i>	
Size	✓
Color	✓
Condition	✓
Observations	826
RMSE	37.3
Adjusted R-squared	0.689

Note: p-Value in parentheses. \*\*\* $p < 0.001$ ; \*\* $p < 0.01$ ; \* $p < 0.05$

Moran's I Test under Randomisation for Psychological Pricing	
Statistic	Value
Moran I statistic standard deviate	1.4564
p-value	0.07264
Expectation	-0.001212121
Variance	0.001543322

Table 5.6: Moran's I Test Results for Psychological Pricing. Two snapshots: n = 826.

Moran's I Test under Randomisation for Log Price	
Statistic	Value
Moran I statistic standard deviate	-0.62634
p-value	0.7345
Expectation	-0.001212121
Variance	0.001544525

Table 5.7: Moran's I Test Results for Log Price. Two snapshots: n = 826.

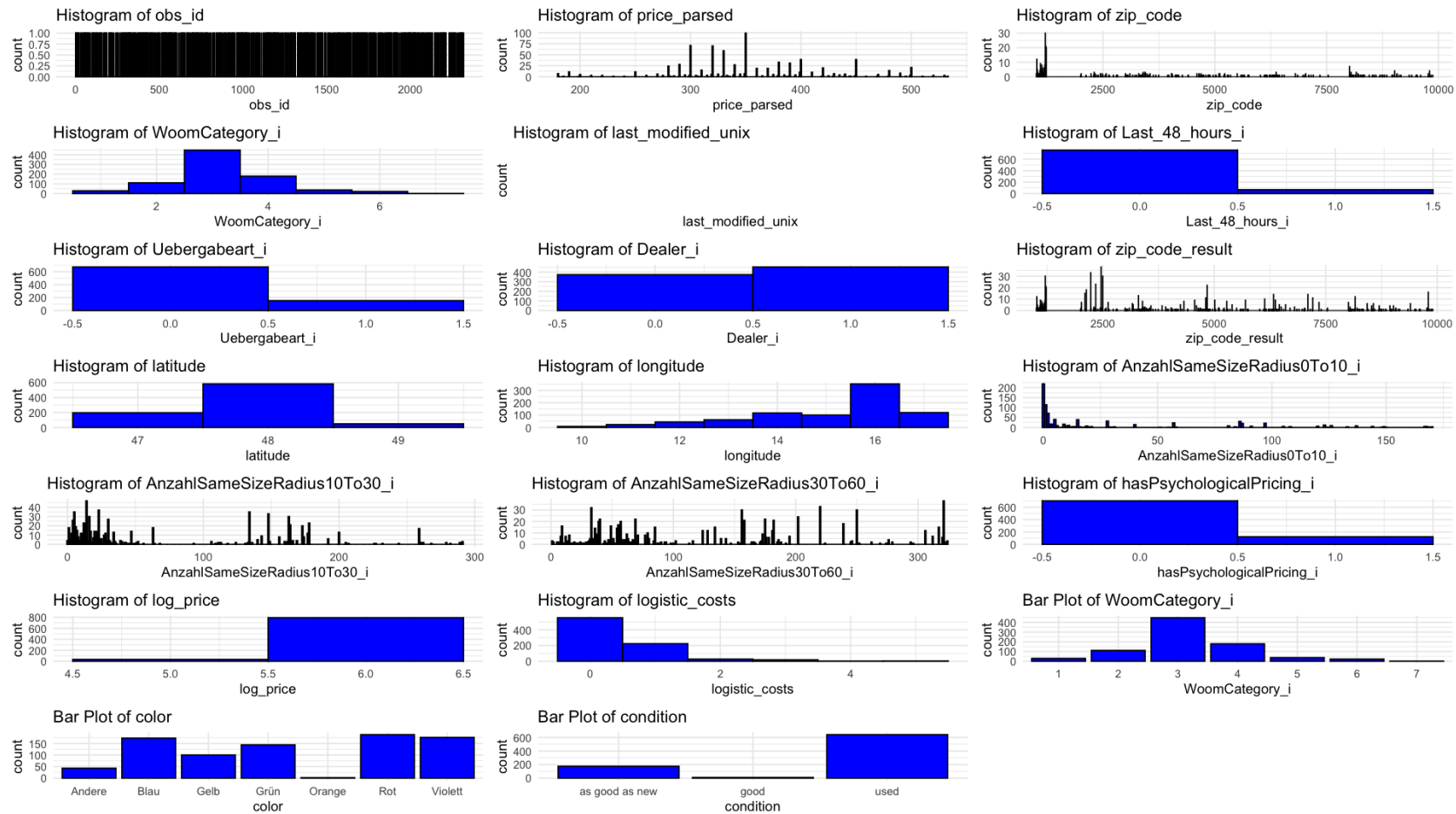


Figure 5.1: Histograms of Key Variables. This figure provides a view of the distribution of most variables, including categorical variables across the dataset.

*Note:* Notably, price outliers that fall outside the 1.5 IQR (Interquartile Range) have already been removed. Two snapshots:  $n = 826$ .

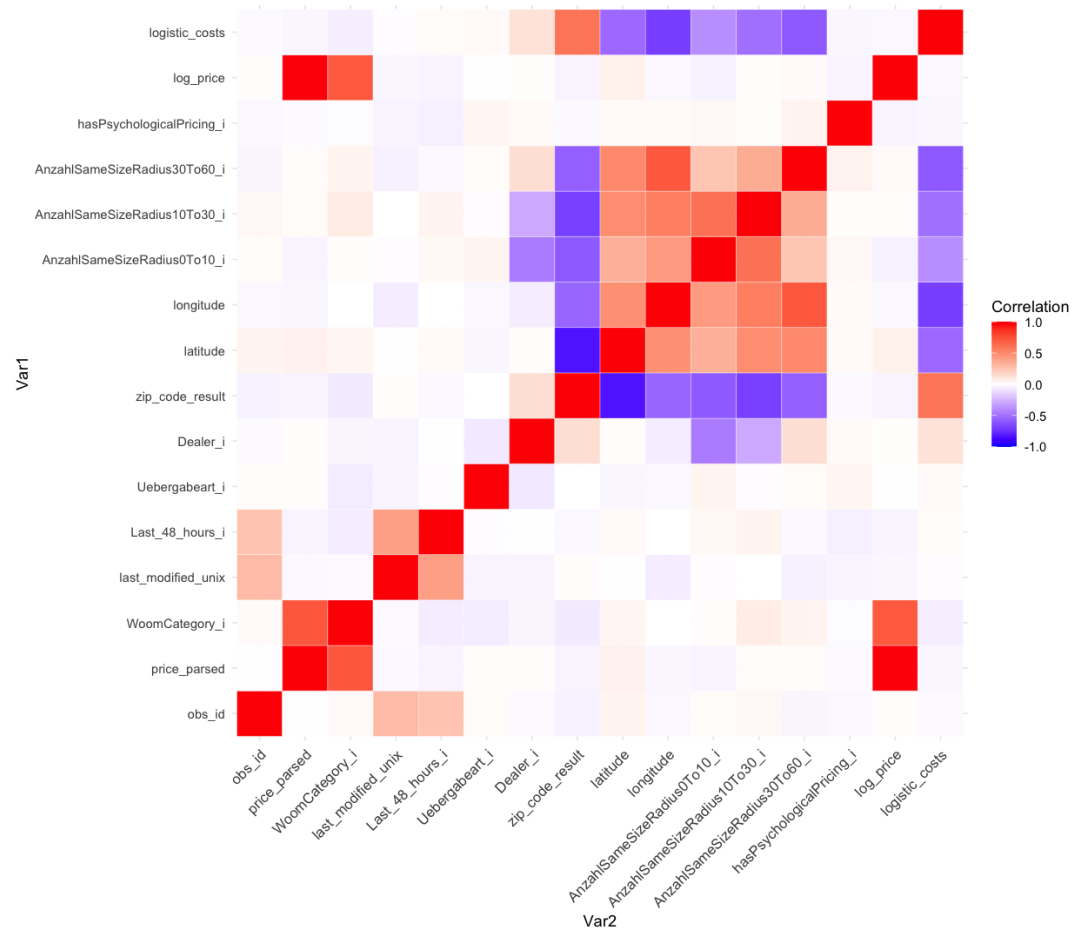


Figure 5.2: Correlation Heatmap. This heatmap visualizes the pairwise correlations between numeric variables in the dataset. *Note:* The heatmap helps in identifying patterns of multicollinearity and relationships between variables. Strong correlations (both positive and negative) are visible as colored blocks, while weak or no correlations are depicted in white. Two snapshots:  $n = 826$ .

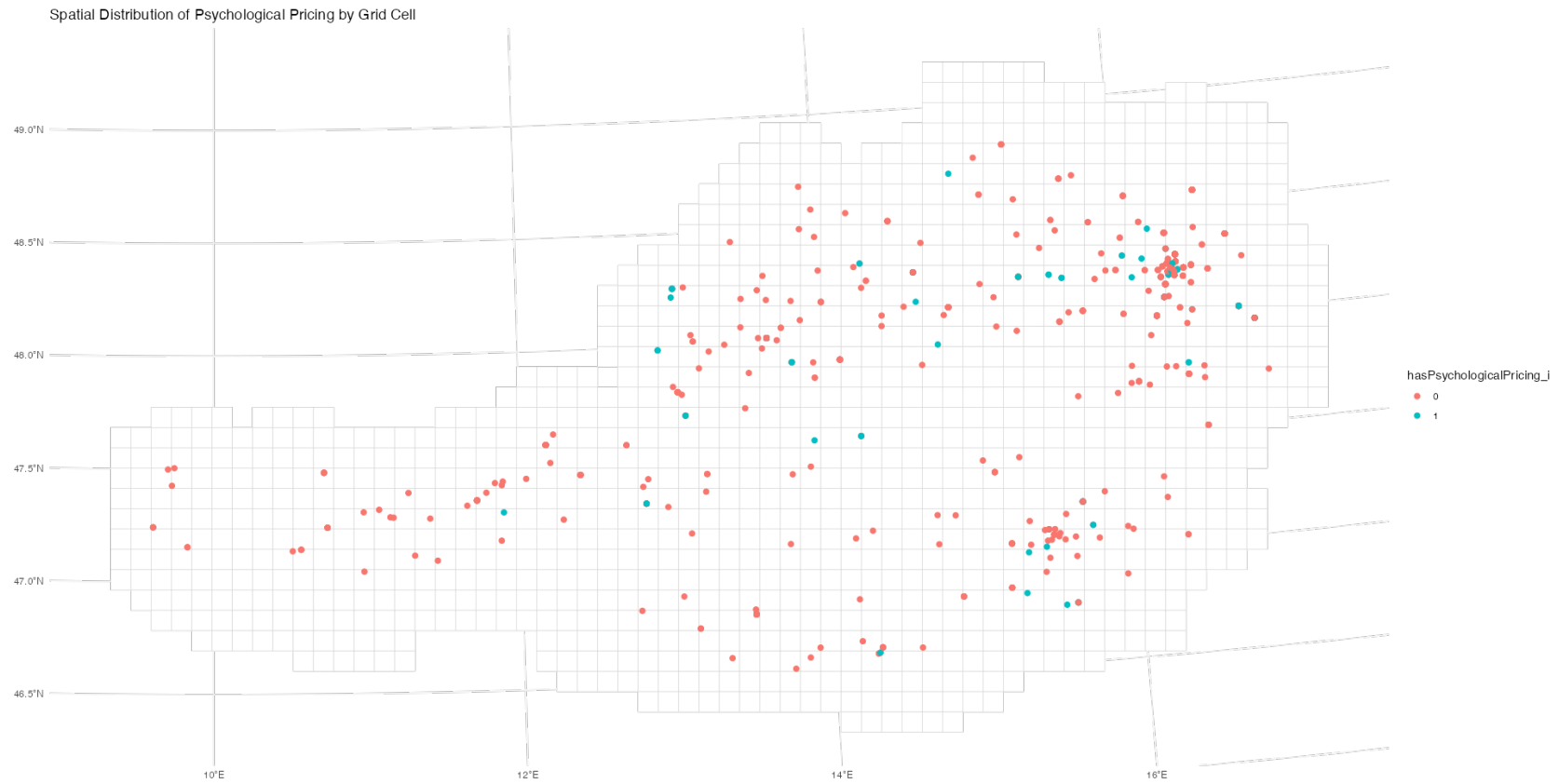


Figure 5.3: Spatial Distribution of Psychological Pricing. This figure illustrates the geographical distribution of the psychological pricing binary dummy variable across Austria.

*Note:* The plot overlays the spatial distribution of bike data points onto a shapefile of Austria, with different colors representing the presence or absence of psychological pricing. Two snapshots:  $n = 826$ .

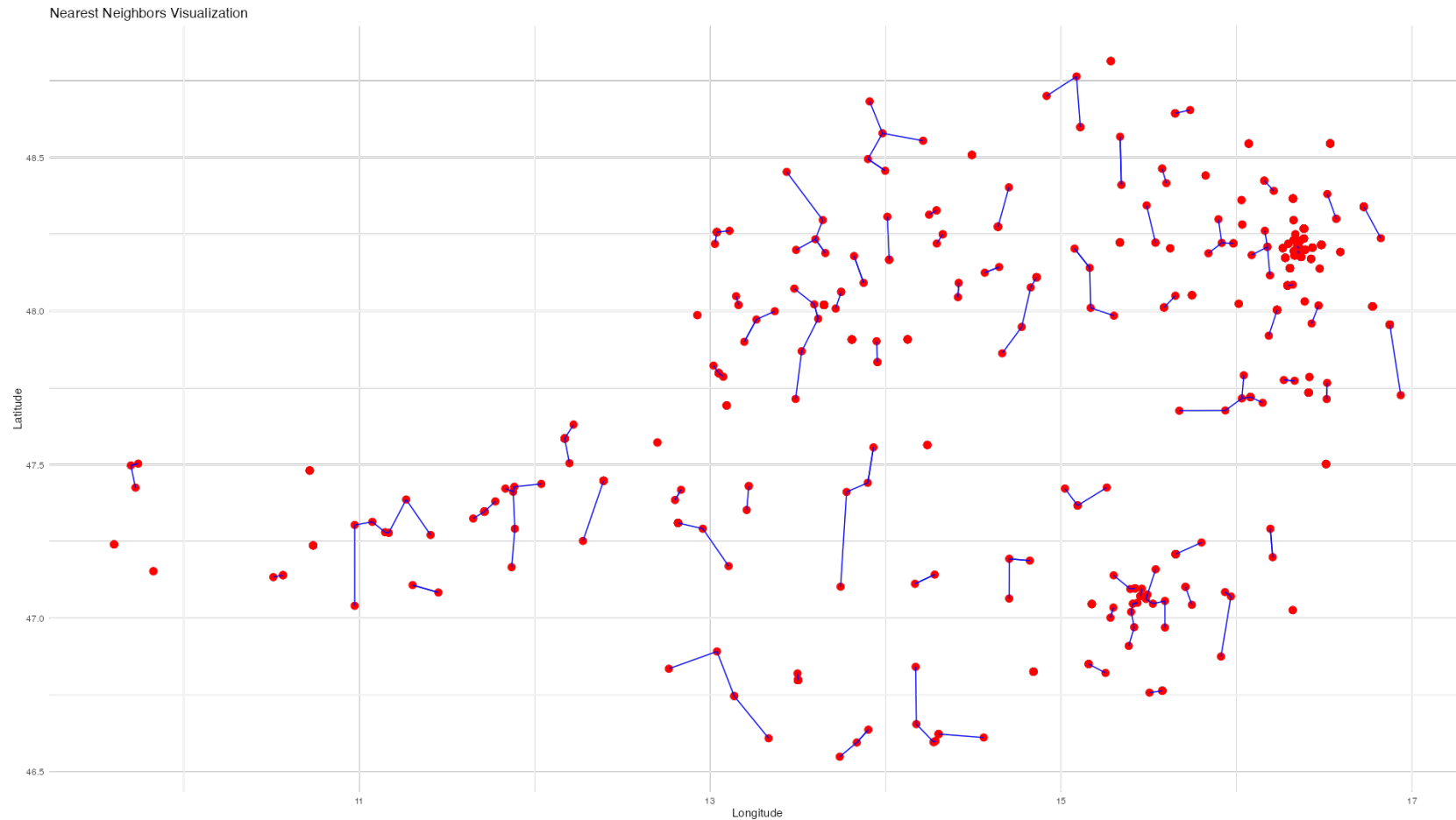


Figure 5.4: This figure depicts the nearest neighbor relationships among bike data points using K-Nearest Neighbors (KNN) with  $K = 1$ .

*Note:* Each data point is connected to its nearest neighbor, highlighting the spatial connections and proximity between observations.

The red points represent the bike data locations, while the blue lines indicate the connections to the nearest neighbors.

Two snapshots:  $n = 826$ .

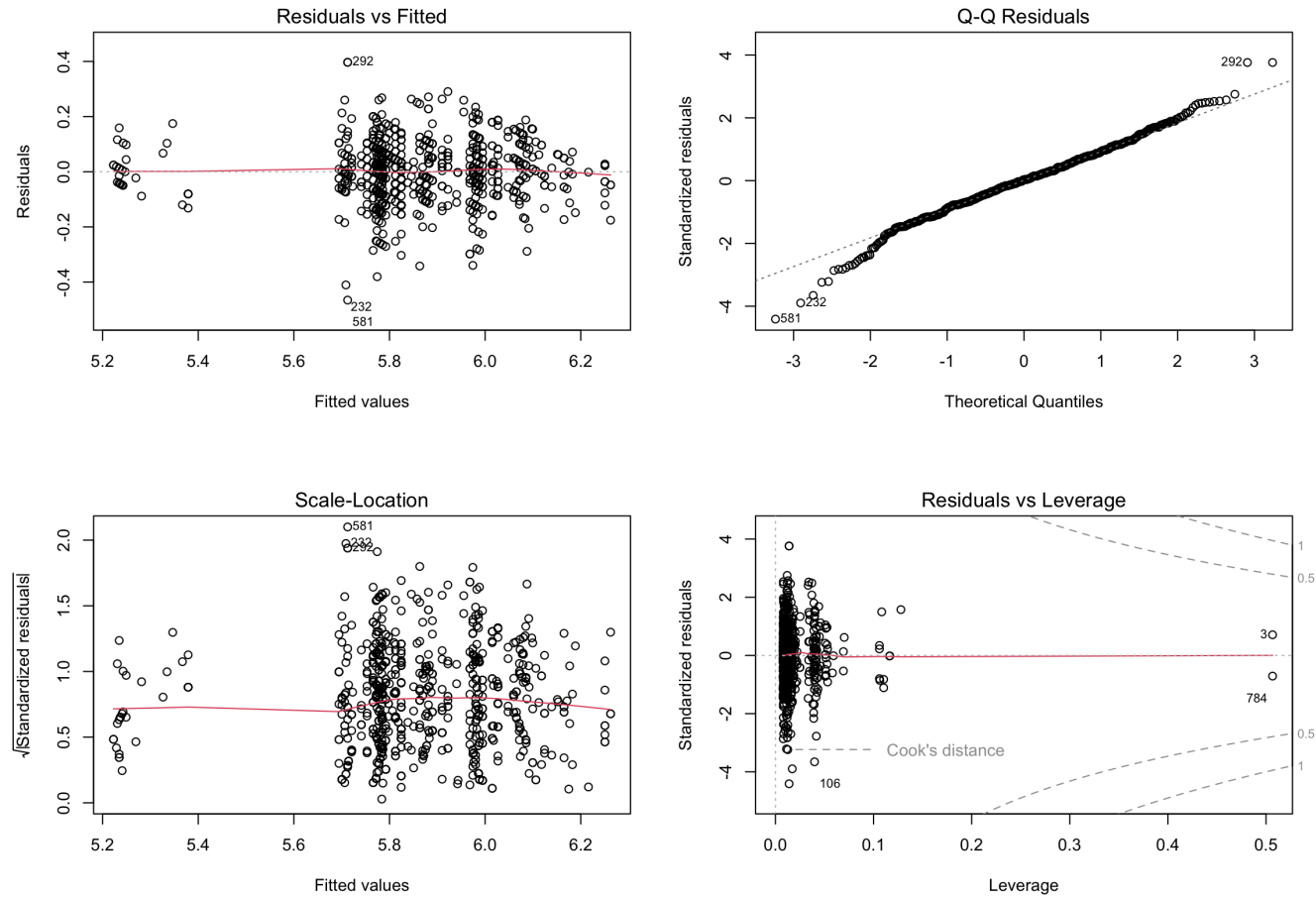


Figure 5.5: Residual Analysis/Q-Q Plot.

*Note:* The Q-Q plot is a diagnostic tool for checking the normality assumption in regression analysis. Points deviating from the 45-degree line may indicate issues with the residuals. Two snapshots:  $n = 826$ .