Determinants of Liquidity on the German Housing Market

Dissertation zur Erlangung des Grades eines

Doktors der Wirtschaftswissenschaft

eingereicht an der Fakultät für Wirtschaftswissenschaften

der Universität Regensburg

Vorgelegt von:	Philipp Freudenreich MScRE, MScISIB
Berichterstatter:	Prof. Dr. Wolfgang Schäfers
	Prof. Dr. Tobias Just
Tag der Disputation:	28. November 2018

Contents

List of Figures

List of Tables

1. Introduction

1.1 General Motivation

With a net asset value of € 4.81 trillions in 2015, the German residential real estate stock far exceeded the values of other European countries including France and the UK.[1] At that time, the market consisted of about 41.4 million dwellings.[2] Despite strong economic performance over the last decades, Germany developed into a nation of tenants. Voigtlaender (2009), Bentzien et al. (2012), Lerbs and Oberst (2014), Kohl (2016), and Reisenbichler (2016), among others, describe reasons for this very distinctive market feature. Another peculiarity of the German market is the high proportion of private landlords. According to the GdW, roughly 37% of all dwellings are offered by about 3.9 million non-professional private landlords, while only 20% are offered by professional landlords.

Despite its tremendous size and widely dispersed ownership structure, the German residential market is relatively opaque and underresearched. This fact is mainly due to the lack of available data. Technological progress in terms of computational capacity, data gathering from internet sources, and the possibility to store large amounts of data have started to improve this situation. Private data providers now offer micro data on a few million transactions on the residential investment and rental market. Although the quality of the data is not comparable to e.g. the US, as only asking prices and asking rents are observable, the new and large scaled data sets enable researchers to explore the market with a higher degree of detail. The data provider Empirica, for example, publishes quarterly price indices and price-bubble indicators. The digital market place Immobilienscout24 publishes the proprietary IMX index developed by Bauer et al. (2013) and even the Federal Statistical Office and the statistical offices of the federal states have started to publish monthly price and rent indices.

Irrespective of the individual provider, these indices share one particular finding. All of them reveal an increase in prices on the investment market, which far exceeds the increase on the rental market. Of course, the historically low interest rates fuel the demand on the investment market, while new landlords have stated to accept lower rental yields due to missing investment alternatives. Nevertheless, landlords are expected to hand over the price increase on the investment market as far as rental protection laws allow them to. German media publishes a record amount of articles about rising rents, the lack of affordable housing, and the threat of gentrification. However, the moderately rising rental price indices show no sign of surplus demand on the rental market.

The sole consideration of price to explain the movements on the residential market ignores the second integral component, when marketing a dwelling. The process of marketing a dwelling

[1] Eurostat (2018)
[2] **IRE**|BS (2017)

starts with the introduction of the dwelling onto the market at a price (or monthly asking rent) determined by the seller (or landlord). With the introduction onto the market, the observation of the second integral component, which is the time it takes until a prospective buyer (or tenant) is willing to take the dwelling off the market and to pay the required price (or monthly amount of rent), starts. Contingent upon a matching of expectations, a market is able to function and the easier, thus faster this matching occurs, the shorter the timespan and consequently the higher the liquidity on the market. Typically, this matching will occur if the price (or asking rent) for the dwelling is supported by its particular location and building characteristics. Depending on the level of the demand in certain regions, buyers (or tenants) might start to tolerate higher prices (or rents). But as long as there is sufficient supply, the prospective buyer (or tenant) will continue to search the market and not rush into an undesired contractual agreement. In accordance with Fisher et al. (2003), the buyer (or tenants) is the provider of liquidity, as he has the financial resources to afford the dwelling and to convert it into cash (or a dividend yielding asset) for the owner. Only if it's up to "take what you can get", buyers (or tenants) will be accepting a price (or rent) which is exceeding their initial reservation price in no time.

Therefore, the aim of the dissertation is to emphasize the importance of a complementary detailed analysis of liquidity by underlining the significance of the component time on market, in order to get a more comprehensive understanding of the German residential real estate market. In this context, the dissertation examines liquidity solely with a time-based measure and does not include transaction cost, price, or volume measures.

Internationally, and in particular on the US residential real estate market, hedonic pricing is a very extensive field of research, going back to Kain and Quigley (1970) and Rosen (1974), among others. The combined analysis of price and time until sale goes back to Cubbin (1974), among others. Traditionally, the very transparent US market with its availability of high quality data gives room for inventions regarding modelling techniques. In the field of real estate hedonic pricing, a number of articles concludes that the inclusion of spatial variables improves the explanatory power of the model e.g. Goodman and Thibodeau (2007), Turnbull and Dombrow (2006), Pavlov (2000), Fik et al. (2003), and Bourassa et al. (2010), among others. Smith (2010) was the first to include district dummies and Cartesian coordinates as well as a distance variable in the context of liquidity analysis. The Quantile Regression (QR) approach, going back to Koenker and Bassett Jr. (1978), has been applied in the field of real estate pricing on the US market by Zietz et al. (2008), Farmer and Lipscomb (2010), Mak et al. (2010), and Liao and Wang (2012), among others. To use the model for the estimation of time on market, censoring in the data has to be taken into account. Due to insufficient computational capacity, this has not yet been possible. Thus, to the best of the author's knowledge, the Censored Quantile Regression (CQR) has not yet been introduced to academic literature regarding real estate liquidity analysis.

On the German residential market, only the pricing aspect of residential real estate has received some form of attention in academic literature see Maennig and Dust (2008), Bischoff (2012), Kajuth et al. (2013), Belke and Keil (2018), among others. Ahlfeldt and Maennig (2010), Brandt and Maennig (2011), and Brandt et al. (2014), among others, introduced spatial variables and spatial gravity variables to the field of hedonic residential real estate pricing on the German market. However, they only concentrate their studies on one specific city. In the field of hedonic pricing, contemporary econometric models like the Quantile Regression and the Generalized Additive Models for Location Scale and Shape (GAMLSS) were introduced to the German market by Thomschke (2015) and Cajias (2018), respectively. However, in the field of liquidity analysis on the German residential real estate market, a substantial research gap exists.

The first paper of the dissertation enables the combined analysis of price and liquidity, by the introduction of quality- and spatial-adjusted hedonic liquidity indices to the German residential investment and rental market. The liquidity index for the rental market is able to reveal hidden demand, which is not represented by the price development. A subsequent clustering based on the index values identifies “hot” and “cold” regional markets. The clustering allows the deduction of investment strategies and assists public institutions when deriving policy measures regarding the residential market. Throughout the dissertation, liquidity is consistently defined as the inverse of time on market, as by Wood and Wood (1985). The estimation of time with econometric methods entails some particular features, like e.g. the absence of negative values or the existence of right censoring in the data. For the analysis of time on market this means, that some dwellings remain available on the market until the end of the observation period. Therefore, the second paper of the dissertation identifies and incorporates those features by exploring the determinants of liquidity by means of survival analysis, more precisely by the application of the Cox (1972) Proportional Hazards Model (PHM), which is able to estimate right censored data. While Kluger and Miller (1990) initially used the model for real estate liquidity analysis, the present paper adapts the PHM to the German market and, to the best of the author’s knowledge, it introduces spatial gravity variables to the field of residential real estate liquidity analysis on the German market. The application of a large scaled datasets allows the identification of heterogeneity across the cities and the deduction of liquidity patterns within the cities. The last paper of the dissertation builds upon the findings of the Cox PHM model in order to introduce the Censored Quantile Regression to the field of real estate liquidity analysis. To the best of the author’s knowledge, it is the first time, the model has been applied to the field in an international perspective. The advanced econometric model allows the examination of the determinants of liquidity on a very granular basis, as it explores the impact of different covariates for individual dwelling offerings across the time on market distribution. While the results confirm the proportional hazards assumption underlying the Cox PHM, they also reveal significant differences in the magnitude and the direction of the impact of individual characteristics on the time on market quantiles.

1.2 Research Questions

This section provides an overview of the three papers comprising the dissertation and the research questions addressed within those papers.

Paper 1: Closing the liquidity gap: Why the consideration of time on market is inevitable for understanding the residential real estate market

- How did prices on the residential investment and rental market develop according to official statistics?
- What is the current state of research on liquidity analysis for the residential real estate market?
- Is it possible to introduce a quality- and spatial-adjusted hedonic liquidity index to the German residential market?
- How did price and liquidity on the residential investment and rental market measured by quality- and spatial-adjusted hedonic indices evolve over the last five years?
- Did the indices for the residential investment and residential rental market develop differently?
- Is the strong demand pressure on the rental market captured by the rental price index?
- In how far do price and liquidity move together?
- Is the clustering of residential markets into “hot” and “cold” market states possible?
- How are the regions of each market cluster characterized? What similarities and differences do these regions share?
- Which investment strategies can be derived with respect to the individual market clusters?

Paper 2: Exploring the determinants of liquidity with big data – market heterogeneity in German markets

- What is the current state of research on liquidity analysis for the residential real estate market using the Cox Proportional Hazard Model?
- How did the liquidity analysis using econometric survival models evolve?
- Is it possible to adapt the Cox Proportional Hazard Model to the German residential rental market using a large scaled data set?
- Is it possible to adapt the Cox Proportional Hazard Model in order to include spatial information next to hedonic and socioeconomic variables and various fixed-effects?
- Does the inclusion of spatial gravity variables significantly increase the explanatory power of the model?
- What additional information can be derived by including the “atypicality” measure introduced by Haurin (1988) and the “degree of overpricing” introduced by Anglin et al. (2003)?
- What are the determinants of liquidity of rental dwellings in the largest seven German cities?
- Are there differences or commonalities across the seven largest German cities?

- Is it possible to derive geographic liquidity patterns for the seven cities?

Paper 3: Exploring the determinants of real estate liquidity from an alternative perspective – Censored Quantile Regression in real estate research

- How did advanced econometric survival analysis evolve in other fields of research?
- Is it possible to introduce the Censored Quantile Regression to the field of real estate liquidity analysis using hedonic, socioeconomic, and spatial variables and various fixed-effects?
- Is it possible to adapt the Censored Quantile Regression to the German residential real estate market using a large scaled dataset?
- What additional information does the model display for the determinants of liquidity of rental dwellings in the largest seven German cities by estimating each quantile separately?
- Does the direction and magnitude of the effect individual covariates have on liquidity change along the time on market distribution?
- Is it possible to confirm the proportional hazards assumption underlying the Cox PHM model?
- What information about the marketability of their current and future dwellings can landlords infer from the CQR results and how can they improve the marketability?

1.3 Course of Analysis

This section provides an overview of the course of analysis with regard to the research purpose, the study design, the authors, the submission details and conference presentations.

Paper 1: **Closing the liquidity gap: Why the consideration of time on market is inevitable for understanding the residential real estate market**

The aim of this paper is to emphasize the importance of time on market when analyzing the residential real estate market. Using 3,055,343 observations, the paper is the first, to the best of the authors' knowledge, to introduce liquidity indices to the German residential investment and rental market. The paper is able to reveal the hidden demand on the rental market and creates clusters in order to summarize common market movements among the 161 observed regions and to facilitate the interpretation of the results. In addition, a higher tendency for spill over effects was found for the rental market.

Authors: Marcelo Cajias, Philipp Freudenreich, Anna Heller, and Wolfgang Schaefers

Submission to: Journal of Business Economics

Status: Under review

This paper was presented at the 2018 Annual Conference of the European Real Estate Society (ERES) in Reading, United Kingdom.

Paper 2: **Exploring the determinants of liquidity with big data – market heterogeneity in German markets**

This paper explores the determinants of liquidity on the residential rental market by examining 335,972 observations on the largest seven German real estate markets. The paper applies the Cox PHM in order to identify and measure those determinants. The model is adapted to include both absolute and relative spatial information in terms of coordinates and distance variables. To the best of the authors' knowledge, this is the first paper to include spatial gravity variables to the field of real estate liquidity analysis on the German residential market in order to increase the explanatory power. The model is able to identify heterogeneity across the cities as well as liquidity patterns within the cities.

Authors: Marcelo Cajias and Philipp Freudenreich

Submission to: Journal of Property Investment and Finance

Status: Published in Volume 31, Issue 1

This paper was presented at the 2017 Annual Conference of the American Real Estate Society (ARES) in San Diego, USA and at the 2017 Annual Conference of the ERES in Delft, Netherlands.

Paper 3: Exploring the determinants of real estate liquidity from an alternative perspective – Censored Quantile Regression in real estate research

Based upon the findings of the previous paper, this study introduces an advanced econometric model to the field of real estate liquidity analysis in order to explore the determinants of liquidity with a higher level of granularity. To the best of the authors' knowledge, this is the first paper to apply the Censored Quantile Regression in the field of real estate liquidity analysis. By using 482,196 observations on the seven largest German cities, CQR allows the identification and measurement of the impact an individual covariate has on time on market across the time on market distribution. The results reveal significant differences in the magnitude as well as the direction of the impact of individual characteristics on the time on market quantiles.

Authors: Marcelo Cajias, Philipp Freudenreich, and Anna Heller

Submission to: Housing Studies

Status: Under Review

This paper was presented at the 2018 Annual Conference of the ERES in Reading, United Kingdom.

1.4 References

Ahlfeldt, G. and Maennig, W. (2010), “Substitutability and Complementarity of Urban Amenities: External Effects of Built Heritage in Berlin”, *Real Estate Economics*, Vol. 38 (2), pp. 285-323.

Anglin, P.M., Rutherford, R.C., and Springer, T.M. (2003), “The Trade-off Between Selling Price of Residential Properties and Time-on-the-Market: The Impact of Price Setting”, *Journal of Real Estate Finance and Economics,* Vol. 26 (1), pp. 95-111.

Bauer, T. K., Feuerschuette, S., Kiefer, M., an de Meulen, P., Micheli, M., Schmidt, T., and Wilke, L. H. (2013), “Ein hedonischer Immobilienpreisindex auf Basis von Internetdaten: 2007–2011“, *AStA Wirtschafts-und Sozialstatistisches Archiv*, Vol. 7 (1-2), pp. 5-30.

Belke, A. and Keil, J. (2018), “Fundamental determinants of real estate prices: A panel study of German regions”, *International Advances in Economic Research*, Vol. 24 (1), pp. 25-45

Benefield, J., Cain, C., and Johnson, K. (2014), “A review of literature utilizing simultaneous modeling techniques for property price and time-on-market”, *Journal of Real Estate Literature*, Vol. 22 (2), pp. 149-175.

Bentzien, V., Rottke, N., and Zietz, J. (2012), “Affordability and Germany’s low homeownership rate”, *International Journal of Housing Markets and Analysis*, Vol. 5 (3), pp. 289-312.

Bischoff, O. (2012), “Explaining regional variation in equilibrium real estate prices and income”, *Journal of Housing Economics*, Vol 21 (1), pp. 1-15.

Bourassa, S., Cantoni, E., and Hoesli, M. (2010), “Predicting House Prices with Spatial Dependence: A Comparison of Alternative Methods”, *Journal of Real Estate Research*, Vol. 32 (2), pp. 139-159.

Brandt, S. and Maennig, W. (2011), “Road noise exposure and residential property prices: Evidence from Hamburg”, *Transportation Research Part D: Transport and Environment*, Vol. 16 (1), pp. 23-30

Brandt, S., Maennig, W., and Richter, F. (2014), “Do Houses of Worship Affect Housing Prices? Evidence from Germany”, *Growth and Change*, Vol 45 (4), pp. 549-570.

Cajias, M. (2018), “Is there room for another hedonic model? The advantages of the GAMLSS approach in real estate research”, *Journal of European Real Estate Research*, Vol. 11 (2), pp. 204-245.

Cox, D.R. (1972), “Regression models and life-tables”, *Journal of the Royal Statistical Society: Series B (Methodological)*, Vol. 34 (2), pp. 187-220.

Cubbin, J.S. (1974), “Price, Quality, and Selling Time in the Housing Market”, *Applied Economics*, Vol. 6 (3), pp. 171-187.

Eurostat (2018), “Balance sheets for non-financial assets (code: nama_10_nfa_bs)”, available at: https://ec.europa.eu/eurostat/de/home (accessed: September 18, 2018).

Farmer, M. and Lipscomb, C. (2010), “Using quantile regression in hedonic analysis to reveal submarket competition”, *Journal of Real Estate Research*, Vol. 32 (4), pp. 435-460.

Fik, T.J., Ling, D.C., and Mulligan, G.F. (2003), “Modeling Spatial Variation in Housing Prices: A Variable Interaction Approach”, *Real Estate Economics*, Vol. 31 (4), pp. 623-646.

Fisher, J., Gatzlaff, D., Geltner, D., and Haurin, D. (2003), “Controlling for the impact of variable liquidity in commercial real estate price indices”, *Real Estate Economics*, Vol. 31 (2), pp. 269-303.

GdW Bundesverband deutscher Wohnungs- und Immobilienunternehmen e.V. (2014), “Anbieterstruktur auf dem deutschen Wohnungsmarkt“, available at: https://web.gdw.de/uploads/pdf/infografiken/2016/Anbieterstruktur.pdf (accessed: September 18, 2018).

Goodman, A.C. and Thibodeau, T.G. (2007), “The Spatial Proximity of Metropolitan Area Housing Submarkets”, *Real Estate Economics*, Vol. 35 (2), pp. 209-232.

Haurin, D.R. (1988), “The Duration of Marketing Time of Residential Housing”, *Real Estate Economics*, Vol. 16 (4), pp. 396-410.

International Real Estate Business School (**IRE**|BS), “Immobilienbestand und -struktur”, in: Just, T., Voigtlaender, M., Eisfeld, R., Henger, R., Hesse, M., and Toschka, A. (2017) “Wirtschaftsfaktor Immobilien 2017”, Berlin, p. 19-48.

Kain, J.F. and Quigley J.M. (1970), “Measuring the value of housing quality*”, Journal of the American Statistical Association*, Vol. 65 (330), pp. 532-548.

Kajuth, F., Knetsch, T.A., and Pinkwart, N. (2013), “Assessing house prices in Germany: evidence from an estimated stock-flow model using regional data”, *Deutsche Bundesbank Working Paper*, No. 46/2013, Deutsche Bundesbank, Frankfurt am Main.

Kluger, B.D. and Miller, N.G. (1990), “Measuring residential real estate liquidity”, *Real Estate Economics*, Vol. 18 (2), pp. 145-159.

Kohl, S. (2016), “Urban history matters: Explaining the German-American homeownership gap”, *Housing Studies*, Vol. 31 (6), pp. 694-713.

Koenker, R. and Bassett Jr., G. (1978), “Regression quantiles”, *Econometrica,* Vol. 46 (1), pp. 33-50.

Lerbs, O. and Oberst, A. (2014), “Explaining the spatial variation in homeownership rates: Results for German regions”, *Regional Studies*, Vol. 48 (5), pp. 844-865.

Liao, W.C. and Wang, X. (2012), “Hedonic house prices and spatial quantile regression”, *Journal of Housing Economics*, Vol. 21 (1), pp. 16-27.

Maennig, W. and Dust, L. (2008), “Shrinking and growing metropolitan areas – asymmetric price reactions? The case of German single-family houses”, *Regional Science and Urban Economics*, Vol. 38 (1), pp. 63-69.

Mak, S., Choy, L., and Ho, W. (2010), “Quantile regression estimates of Hong Kong real estate prices”, *Urban Studies*, Vol. 47 (11), pp. 2461-2472.

Pavlov, A.D. (2000), “Space-Varying Regression Coefficients: A Semi-parametric Approach Applied to Real Estate Markets”, *Real Estate Economics*, Vol. 28 (2), pp. 249-283.

Reisenbichler, A. (2016), “A rocky path to homeownership: Why Germany eliminated large-scale subsidies for homeowners”, *Cityscape: Ajournal of policy development and research*, Vol. 18 (3), pp. 283-290.

Rosen, S. (1974), “Hedonic prices and implicit markets: product differentiation in pure competition”, *Journal of Political Economy*, Vol. 82 (1), pp. 34-55.

Smith, B.C. (2010), “Spatial Heterogeneity in Listing Duration: The Influence of Relative Location to Marketability”, *Journal of Housing Research*, Vol. 18 (2), pp. 151-171.

Thomschke, L. (2015), “Changes in the distribution of rental prices in Berlin”, *Regional Science and Urban Economics*, Vol. 51, pp. 88-100.

Turnbull, G.K. and Dombrow, J. (2006), “Spatial Competition and Shopping Externalities: Evidence from the Housing Market”, *Journal of Real Estate Finance and Economics*, Vol. 32 (4), pp. 391-408.

Voigtlaender, M (2009), “Why is the German homeownership rate so low?”, *Housing Studies*, Vol. 24 (3), pp. 335-372.

Wood, J.H. and Wood N.L. (1985), “Financial Markets”, San Diego: Harcourt Brace Jovanovich.

Zietz, J., Zietz, E.N., and Sirmans, G.S. (2008), “Determinants of house prices: a quantile regression approach”, *The Journal of Real Estate Finance and Economics*, Vol. 37 (4), pp. 317-333.

2. Closing the liquidity gap: Why the consideration of time on market is inevitable for understanding the residential real estate market

Abstract

This paper identifies the liquidity of dwellings, defined as the inverse of time on market, as an integral component when analyzing the residential market. Based on quality- and spatial-adjusted price and liquidity indices for about three million observations on the German residential investment and rental is clustered in order to summarize common market trends and to facilitate a regional interpretation. In that way, the study not only reveals the true demand on the rental market, but is able to identify "hot" and "cold" regions. This classification allows the deduction of investment strategies and enables a more precise drawing of policy implications. In addition, a stronger tendency for spill over effects was identified for the rental market.

Acknowledgement: The authors especially thank PATRIZIA Immobilien AG for contributing the dataset for this study. All statements of opinion are those of the authors and do not necessarily reflect the opinions of PATRIZIA Immobilien AG or its associated companies.

2.1. Introduction

On the residential market, the process of selling or renting out a dwelling comprises of two essential components. The first component is the introduction of the dwelling onto the market at a price (monthly asking rent) determined by the seller (landlord). The second component is the time it takes until a prospective tenant is willing to take the dwelling off the market and to pay the required price (monthly amount of rent). Contingent upon a matching of those expectations, a market is able to function and the easier thus faster this matching occurs, the higher the liquidity on the market. In the following, liquidity is defined as the inverse of the time on market (TOM) in accordance with Wood and Wood (1985). In this context, the study examines liquidity solely with a time-based measure and does not include transaction cost, price, or volume measures. Typically, this matching will occur if the price (asking rent) for the dwelling is supported by its particular location and building characteristics. Depending on the level of the demand in certain regions, buyers (tenants) might start to tolerate higher prices (rent). But as long as there is sufficient supply, the prospective buyer (tenant) will continue to search the market and not rush into an undesired contractual agreement. In accordance with Fisher et al. (2003), the buyer (renter) is the provider of liquidity, as he has the financial resources to afford the dwelling and to convert it into cash (a dividend yielding asset) for the owner. Only if it's up to "take what you can get", buyers (tenants) will be accepting a price (rent) which is exceeding their initial reservation price in no time. Without a combined analysis of price (rent) and time, these exceptional levels of demand cannot be captured. Nevertheless, it is mainly price which is at the center of attention of market participants and captured by a variety of indices worldwide. To improve the assessment of the residential market, this paper provides quality- and spatial-adjusted liquidity indices for the residential investment and rental market, as complementary demand indices.

Over the last decade, the German residential real estate market has experienced significant changes. The fundamental economic data exhibits a growing GDP accompanied by historically high levels of labor demand. The consistently favorable macroeconomic situation and geopolitical events triggered high migration from within the European Union as well as from many conflict zones. In addition, the number of households has been increasing due to the social trend towards smaller households. Furthermore, interest rates for mortgages have been extremely low, resulting in higher affordability of home ownership. Unsurprisingly, this economic development led to booming demand for residential real estate. Despite rising building permissions and construction activity, building completions have been insufficient to meet the demand in many regions. The Federal Institute for Research on Building, Urban Affairs and Spatial Development (BBSR) has identified an increased demand of 272,000 new dwellings per year for the years 2015-2020. According to the Federal Statistical office, dwelling completion was 216,727 in 2015, 235,658 in 2016, and 245,304 in 2017. The statistics show, that not in a single year since the BBSR published the study, enough new dwellings entered the market. As a consequence, vacancy rates fell below sustainable levels in many regions and house prices as well as rents have experienced upside pressure. The official national house price index of the Federal Statistical Office reveals a national price increase of 19.7 percentage points (pp) for the last 5 years. According to the exclusive sample used for this study, which includes 973,164 observations on the investment market, asking prices increased by 33.1 pp on average within same period. Only 161 regions with more than 100 offers per quarter are included in the sample in order to avoid a bias stemming from those inactive markets. The Federal Statistical Office, however, includes all regions, irrespective of the activity of the market. A decomposition of the consumer price index published by the Federal Statistical Office reveals an increase in rents of a mere 5.7 pp for the last five years. Again, the current sample of 2,082,179 rental offers reveals a higher average increase of 13.5 pp within the same period. The price appreciation on the rental market seems innocuous in comparison to the one experienced on the investment market and even internationally, with the UK and the US undergoing a nationwide five-year rental increase of 10.8 and 16.0 pp, respectively. With a homeownership rate of 43% as of 2013, the first year covered by the current sample, more than half of the German population rent their homes. While Voigtlaender (2009), Bentzien et al. (2012), Lerbs and Oberst (2014), and Reisenbichler (2016) discussed in detail the reasons for the extraordinarily low homeownership rate, research on the nationwide rental market is rather scarce, although not only the 2:1 ratio of available data for this study demonstrates the importance of the rental market. Simply based on the moderate appreciation in asking rents, it is hardly possible to make any inferences with regard to a tight rental market. Are the stories about property viewings with more than 50 competitors for the same flat only urban myths? Maybe the analysis of the liquidity on both residential markets reveals the somehow hidden demand. By only looking at the mean change in time on market, it seems as there is actually no difference between the investment and the rental market, as the liquidity improved by 52 and 46 pp respectively. An estimation of quality- and spatial-adjusted liquidity indices, however, exposes the real difference in market tightness.

As most other markets, the residential real estate market exhibits cyclical movements over time. According to the seminal work of Kluger and Miller (1990) who developed a liquidity measure by using the Cox (1972) Proportional Hazards Model, housing prices and liquidity exhibit a positive correlation. Thus, prices and liquidity should match along “hot” and “cold” market states. Krainer (1999) defines a market as “hot”" when prices are increasing, the time on market is short and transaction volume is above average. In contrast, decreasing prices, relatively long selling times and low transaction volumes point to a “cold” housing market. A positive correlation is found in Stein (1995), Berkovec and Goodman Jr. (1996), and Ortalo-Magné and Rady (2004) for instance. Follain and Velz (1995) and Hort (2000), among others, find a negative correlation. While Stein (1995) and Genesove and Mayer (1997) reason the correlation with sellers' equity constraints, i.e. with frictions on the credit market, Krainer (1999) shows that “hot” and “cold” real estate markets emerge due to search frictions and asymmetric information. Cauley and Pavlov

(2002) show evidence for the option value of homeowners and for nominal loss aversion. Substantial deviations from these two market states might indicate speculative expectations by investors and landlords, adjustment processes or supply and demand changes. To detect these deviations is essential for real estate market participants, as it is otherwise impossible to build decisions correctly.

Literature in this field focuses predominantly on the US residential investment market, while academic research concerning real estate market movements on the German market is rather scarce. The lack of research on the German residential real estate market might be traced back to the fact, that micro data and computational power have not been sufficiently available only a few years ago. While most of this literature strand focuses on "hot" and "cold" market phases along the residential cycle, this paper aims to detect "hot" and "cold" market spots on a regional basis. As one of the few papers on the German market, an de Meulen and Mitze (2014) identified "hot" and "cold" spots on the Berlin residential market. In order to detect those, the authors exclusively investigated the price aspect of dwellings. In general, the movements on the residential real estate market are described primarily with price indices. On the German market, there are hedonic price indices provided by the Federal Statistical Office as well as indices provided by private companies like e.g. bulwiengesa and Immobilienscout24 (IMX). The methodology and data behind the IMX are described in Bauer et al. (2013). A complementary liquidity index, however, is missing. Nevertheless, for central banks, policy makers, institutional investors, and private households it is essential to be aware of the liquidity momentum, as both indices might move in opposite directions, pointing to different market states. Thus, solely considering the price index for classifying a regional market might lead to incorrect investment strategies and policy implications. Therefore, this paper develops a quality- and spatial-adjusted price and a complementary liquidity indicator for the investment and rental market of 161 German regions.

According to the indices, the regional housing markets are then clustered in order to reassess the assumption that prices and liquidity move together or whether their dynamic behavior exhibits frictions. For more than 25 years, bulwiengesa has been providing a clustering of German cities according to their size, measured by the number of inhabitants, the size of the office market and the importance of the city for the national as well as international real estate market.[3] Heinrich and Just (2016) have noted, that those characteristics might not be entirely sufficient for analyzing the housing market and developed a solution which primarily includes information on the housing market and the population within those cities. While the approach of Heinrich and Just (2016) and the one presented in this paper both use a form of k-means clustering, the latter one does not directly cluster a variety of variables, but uses them for the preceding index calculation. In addition to the quality- and spatial-adjusted regional price and rent indices, the liquidity index is introduced as an additional layer, in order to cluster the regions. The indexing and two-

[3] RIWIS database, Bulwiengesa AG

dimensional clustering on a regional level yields a very granular analysis of the German residential investment and rental market and allows the identification of “hot” and “cold” spots as starting points for the deduction of investment strategies. The findings should also be of interest to consumers of living space and policy makers trying to steer the residential market.

The study aims to answer the following questions regarding the residential market:

- How did prices on the residential investment and rental market develop according to official statistics?
- Is it possible to introduce a quality- and spatial-adjusted hedonic liquidity index to the German residential market?
- How did price and liquidity on the residential investment and rental market measured by quality- and spatial-adjusted hedonic indices evolve over the last five years?
- Did the indices for the residential investment and residential rental market develop differently?
- Is the strong demand pressure on the rental market captured by the rental price index?
- In how far do price and liquidity move together?
- Is the clustering of residential markets into “hot” and “cold” market states possible?
- How are the regions of each market cluster characterized? What similarities and differences do these regions share?
- Which investment strategies can be derived with respect to the individual market clusters?

The remainder of this paper proceeds as follows: The next section describes the dataset and the descriptive statistics. Section 2.3 presents a description of the econometric model, including the derivation of the hedonic price and liquidity indices as well as the clustering algorithm. Estimation results are presented and discussed in section 2.4. Section 2.5 concludes.

2.2. Data and descriptive Statistics

The sample used to estimate the price and liquidity indices combines three merged data sets. It contains information of 3,055,343 observations on the rental (2,082,179) and investment market (973,164) in 161 German regions from the first quarter of 2013 to the fourth quarter of 2017. Information on the dwellings is gathered from various Multiple Listing Services (MLS) as collected from the Empirica Systems Database, containing real estate market data from the most important MLS providers. Hedonic characteristics of the dwellings contain the time on market as the number of weeks the dwelling was listed in the MLS calculated by the start and end date, the asking price in € for the investment market sample and the asking rent in € per month for the rental market sample due to a lack of transaction prices and contract rents. Nevertheless, a substantial bias is not to be expected, see e.g. Shimizu et al. (2012) and Lyons (2013), among others. Typical housing attributes included in the study are e.g. binary hedonic characteristics like "with balcony", being 1 if the dwelling exhibits a balcony and 0 otherwise. Since the data is georeferenced, two spatial gravity indicators measuring the Euclidian distance of each dwelling to the geographical centroid of the ZIP and NUTS3 polygon in kilometers have been calculated. NUTS3 regions correspond to the "Nomenclature of territorial units for statistics", which is a hierarchical system for dividing the economic territory in Europe. The NUTS3 regions cover small regions similar to counties or administrative districts. For the German case, those NUTS3 regions are either rather extensive counties containing many communities and smaller cities or urban districts. One example are the neighboring NUTS3 regions Munich city and Munich county. The spatial gravity variables control for the spatial distribution of dwellings within a region. Furthermore, the socioeconomic variables, purchasing power per household and the number of households at the ZIP code level, are extracted from the GfK-database. The population density per sqkm in a ZIP code area is then calculated in ArcGIS. The last source is Thomson Reuters Eikon, providing the 10-year interest rate for housing loans as a macro variable. The variables, their units and sources can be found in table 2.1.

Table 2.1: Variables and sources

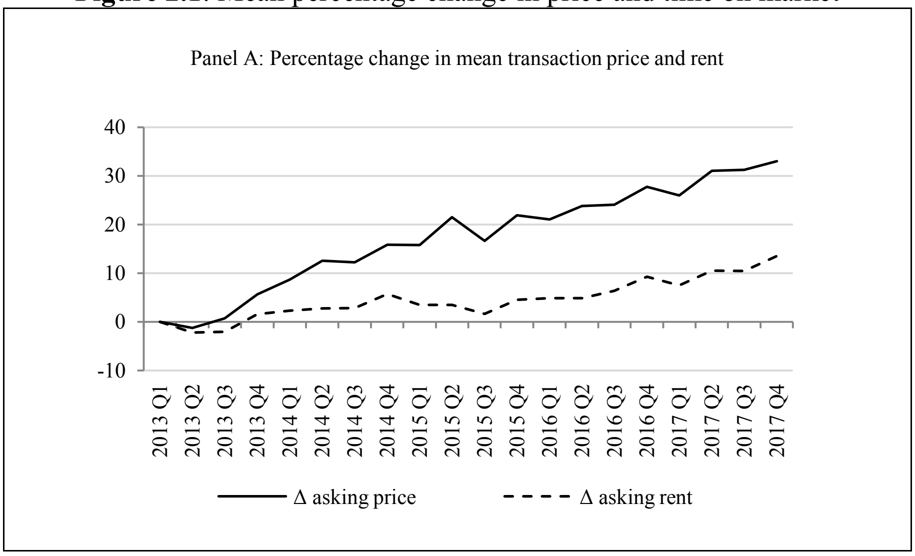
Variable	Unit	Effect in the survival equations			Source			
		Hedonic effects	Spatial effects	Socioeconomic effects	Empirica	GfK	ArcGIS	Reuters
Asking price	€	✓						
Asking rent	€ per month	✓			✓			
Time on market	Weeks	✓			✓			
Living area	M²	✓			✓			
Age	Years	✓			✓			
Rooms	Number	✓			✓			
With bathtub	Binary	✓			✓			
With built-in kitchen	Binary	✓			✓			
With car space	Binary	✓			✓			
With terrace	Binary	✓			✓			
With balcony	Binary	✓			✓			
With elevator	Binary	✓			✓			
Newly built dwelling	Binary	✓			✓			
Refurbished dwelling	Binary	✓			✓			
Gaussian longitude	Coordinate		✓				✓	
Gaussian latitude	Coordinate		✓				✓	

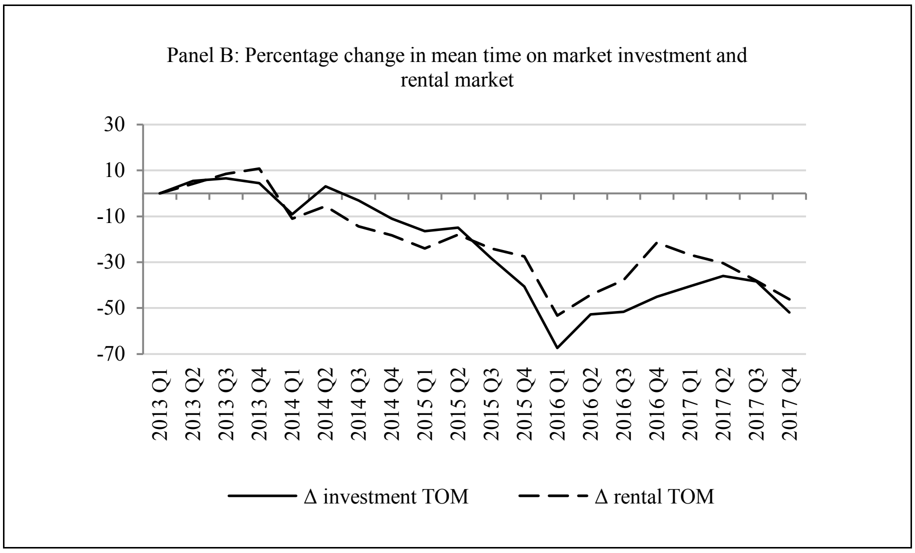
Distance to ZIP centroid	Km		✓				✓	
Distance to NUTS3 centroid	Km		✓				✓	
Households in ZIP	HHs/ZIP			✓		✓		
Purchasing power of HHs in ZIP	€/HH/p.a./ ZIP			✓		✓		
Population density in ZIP	Persons/km² /ZIP			✓		✓		
IR for housing loan 10 years	Effective interest rate in %			✓				✓
N						3,055,343		

Notes: This table reports the unit, the type of effect, and the source of all variables included in the hedonic price and liquidity index calculations as well as the number of observations.

Figure 2.1 shows the mean asking price and time on market development on the investment and rental market from the first quarter of 2013 until the end of the observation period. It is visible that prices have been increasing accompanied by a diminishing time on market on both markets. Hence, both indicators point to a boom phase on the German housing market, triggered by ongoing demand with supply lagging behind. Moreover, it is observable that transaction prices have been increasing considerably more than rents. While rents have been rising by a mere 13.5%, transaction prices have experienced a substantial growth of 33.1% over the last five years. Those figures indicate a particularly high demand on the investment market, probably triggered by constantly low mortgage rates on housing loans and a lack of interest bearing investment opportunities.

Figure 2.1: Mean percentage change in price and time on market



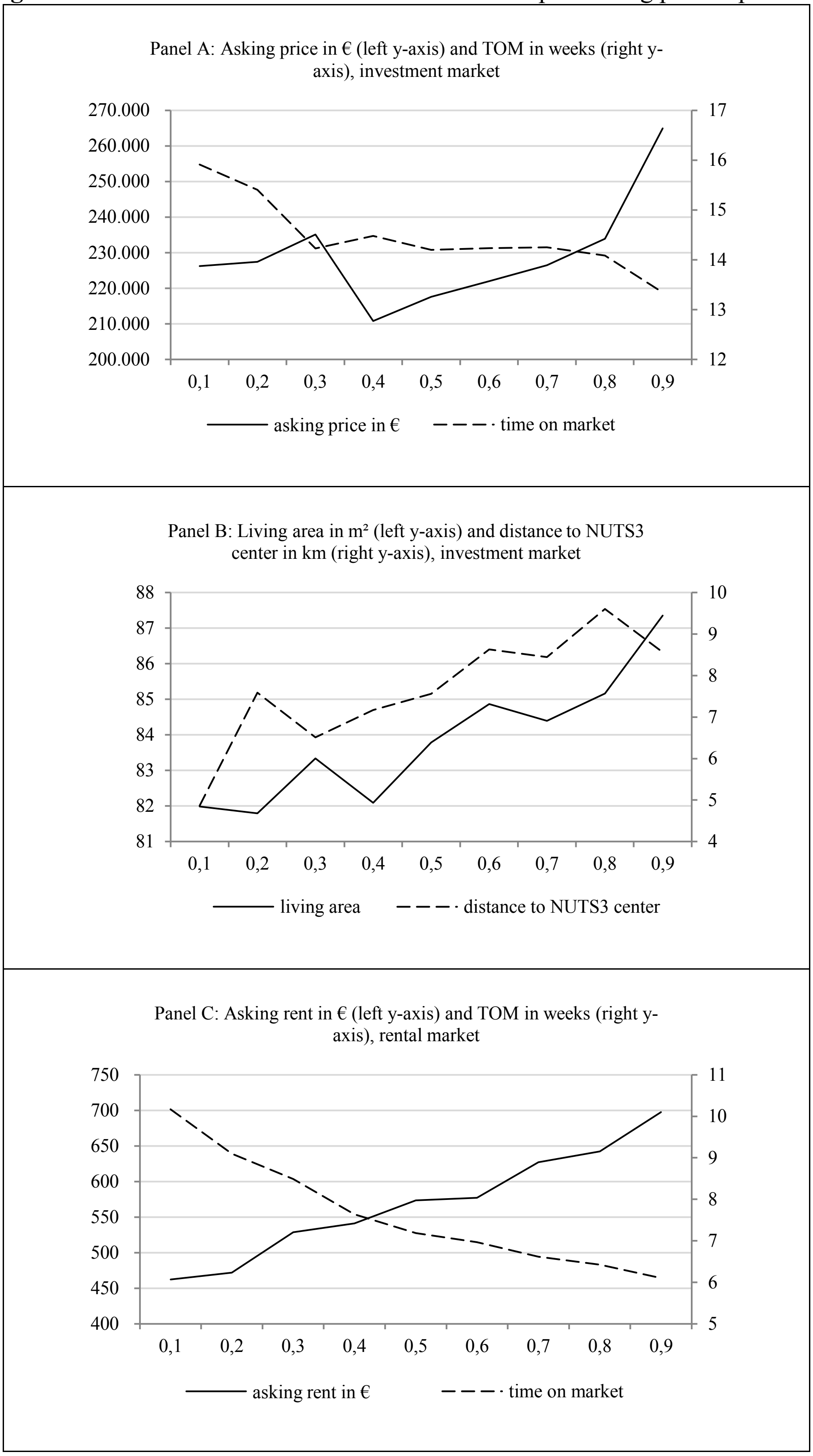


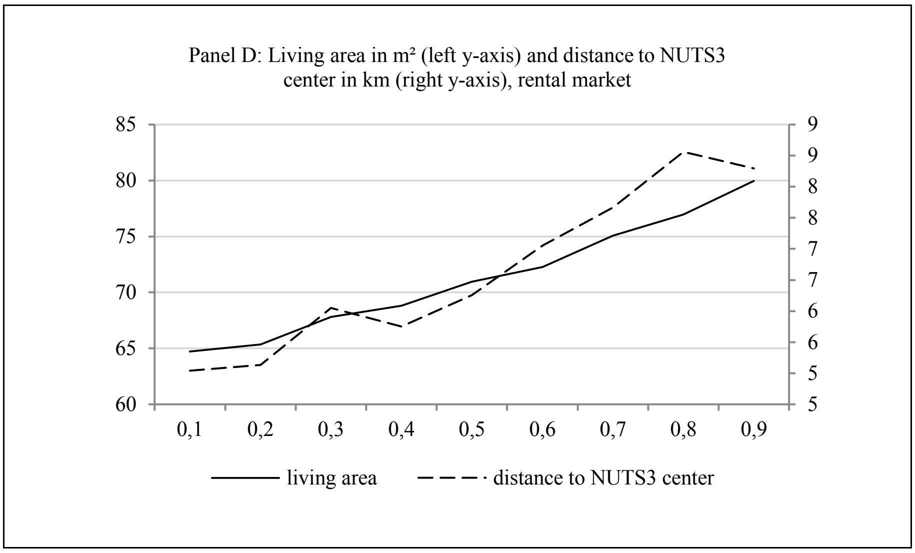
Notes: This figure plots the cumulative percentage change in mean transaction price and rent as well as the cumulative percentage change in time on market on the residential investment and rental market. The data consists of 973,164 observations on the residential investment market and 2,082,179 observations on the residential rental market. The sample period is 2013 Q1 to 2017 Q4.

However, it seems that the price development has not yet been fully absorbed by the rental market. The relatively moderate growth in rents seems to only reflect the natural demand, which obviously was higher in cities. As landlords will try to pass on the rising prices on the investment market to their tenants, it might indicate further rental growth in the near future. Of course, rental protection laws prohibit landlords to hand over the entire increase in transaction prices to tenants in order to meet their target return. Asking exorbitant rent has been prohibited on the German market for years, not only since the introduction of the "Mietpreisbremse" in 2015. Because of lacking investment alternatives, new landlords somehow became acquainted to shrinking rental yields. Nevertheless, time on market exhibits an enormous decrease of about 50% with an almost parallel development on both markets. Although price and rent have not experienced growth of equal magnitude, the similar development in time on market indicates considerably high demand on the rental market, which might also result in upward pressure on rents. To reason the similar drop in the time on market with relatively more supply on the rental market in relation to the investment market is rather less plausible, as newly built dwellings are usually offered on the investment market, before they appear on the rental market. Thus, this slightly controversial finding emphasizes the importance of focusing on both indicators – price and time on market – when analyzing the residential real estate market.

Figure 2.2 exhibits, that heterogeneity is omnipresent on the housing market. Panel A to D show, that households within different purchasing power quantiles demand a different price, living area, and distance to the city center. Furthermore, it is shown that the sales and letting process with respect to the marketing time varies.

Figure 2.2: Distribution of selected variables across purchasing power quantiles





Notes: The figures plot the distribution of selected variables segmentd by nine purchasing power quantiles. The data consists of 973,164 observations on the residential investment market and 2,082,179 observations on the residential rental market. The sample period is 2013 Q1 to 2017 Q4.

Generally, relatively richer households can afford more expensive dwellings, prefer larger living areas, tend to live further away from the city center and spend less time on the search and matching process. These preferences are visible for the investment as well as the rental market. However, Panel C and D show, that the trend of the selected variables from the lowest to the highest purchasing power quantile is much smoother on the rental market. While the trend on the rental market is almost linear, it exhibits fluctuations on the investment market. Surprisingly, buyers living within zip codes with the lowest purchasing power (0.1-, 0.2-, 0.3-quantile) are asked to pay higher prices than the middle-income (0.4-, 0.5-, 0.6-quantile) groups. Another interesting fact is, that the range between the highest and lowest income group with respect to prices, living area and time on market is remarkably more pronounced on the rental market relative to the investment market. While on the investment market asking prices, living area and the time on market between the richest and poorest quantile vary by 17.1%, 6.5% and 16.2%, the differences on the rental market are much stronger with 50.9%, 23.6% and 40.0% respectively. This infers, that the participants on the rental market are much more diversified than those on the investment market.

2.3. Econometric Approach

The aim of a price index is to measure the price development over successive periods for the same quality-adjusted good. However, residential dwellings are not transacted periodically, but rather irregularly and even infrequently. Furthermore, residential real estate is extremely heterogeneous, both in terms of its physical characteristics and its location. Dwellings with different characteristics and in different locations might exhibit distinct price and liquidity dynamics in terms of volatility and cyclicality. Thus, idiosyncratic price and liquidity movements might be to observe in diverse markets, due to social, economic, and political circumstances in a particular region. In order to control for heterogeneity, hedonic indexing is applied in this paper. The hedonic approach dates back to Waugh (1928), Court (1939), Stone (1954), and Griliches (1961), who developed a method for generally indexing economic prices of goods affected by quality changes. Conceptually this approach goes back to Lancaster (1966) and Rosen (1974). Kain and Quigley (1970) were among the first to apply hedonic pricing to the real estate market. Given hedonic data, the hedonic model decomposes the price as well as the liquidity of residential real estate into individual characteristics. Hence, the computed index reveals constant characteristics and consequently points out the pure price and liquidity changes over time. The location of a dwelling is probably one of the most important determinants of prices and liquidity. Therefore, not only postcode identifiers as well as longitude and latitude data are considered, but the price and liquidity indices are estimated individually for each market $p \in \{1,\ldots,161\}$, defined by the NUTS3-regions. In this paper, the time-dummy method is applied. As the time-dummy price (liquidity) indices are defined as the marginal change in price (liquidity) with respect to time, a transformation of the estimated coefficients of the time fixed effects yields the price (liquidity) index, referring to the percentage marginal change in prices (liquidity) in period t_t relative to t_0. Hence, the indices can be computed directly from the estimated coefficients. Compared to the imputed hedonic index no "representative dwelling" has to be defined and it is less data intensive and therefore very well suited for the construction of regional price and liquidity indices. The standard model for the estimation of a time-dummy hedonic index is given as

$$\mathrm{y} = \mathrm{X}\beta + \mu\theta + \mathrm{u}. \quad (1)$$

As the semi-log functional form has proven appropriate and is used in most hedonic regression models according to Halvorsen and Pollakowski (1981), Pollakowski (1982), Diewert (2003), and Malpezzi (2003), among others. y is an $I \times 1$-vector consisting of the elements $y_i = ln(p_i)$. I denotes the number of dwellings in the sample. X is defined as an $I \times C$-matrix of covariates, with C being the number of covariates without the time dummies, β is a $C \times 1$-vector, describing the shadow price of each covariate. To generate an intercept as the first item of β, the first column of X solely consists of ones. μ is an $I \times T\text{-}1$-matrix of time dummies for each period, with T being the number of observation periods, θ is a $T\text{-}1 \times 1$-vector of period shadow prices relative to a fixed time period t_0, and u is an $I \times 1$-vector of error terms. As the purpose is to generate a price index, the coefficient of interest is the time dummy parameter θ. θ quantifies the time period-specific fixed effects, i.e. the impact of each time period, on the log price after controlling for quality and spatial characteristics of a dwelling. Exponentiating the estimated coefficient $\widehat{\theta_t}$, yields the time-dummy index as

$$\widehat{P_t} = \exp(\widehat{\theta_t}). \quad (2)$$

A transformation via $\left[\exp\left(\widehat{\theta_t}\right)-1\right]\cdot 100$ corresponds to the marginal change in prices in t_t relative to t_0. It is to note, that the time dummy index estimated above is not unbiased. According to Goldberger (1968), Teekens and Koerts (1972) and Kennedy (1981),

$$\widehat{P_t^*}=\exp\left(\widehat{\theta_t}+0.5\left(\widehat{se}\left(\widehat{\theta_t}\right)\right)^2\right), \tag{3}$$

yields a standard bias correction. $se(\theta_t)$ refers to the standard error of the time-dummy coefficient. However, according to Goldberger (1968), Kennedy (1981) and Syed et al. (2008), among others, the bias is in general very small. Syed et al. (2008) analyze house price indices in Australia from 2001 to 2006 and show that the difference in the indices appears only in the fourth decimal place. Thus, there is no need to correct for the bias according to Triplett (2004), Hill et al. (2009), and de Haan (2010), among others.

As this paper aims to investigate the dynamics of prices and liquidities, four models are estimated in order to obtain the price index for the investment market, the rental price index and the two liquidity indices for the residential investment and rental market. While for the price indices hedonic regressions are estimated, survival models are set up to obtain the liquidity indices. The four models are estimated individually for each market $p \in \{1, \ldots, 161\}$, defined by the NUTS3-regions, as independently pooled cross-sectional regressions.

2.3.1 The Residential Transaction and Rental Market Price Index

This section describes the derivation of the time-dummy price index for the residential investment as well as rental market. In the following, the term "price" refers to the transaction price as well as the rental market price. The hedonic equation (4) is estimated for both markets separately, see Cajias (2018). Estimation is conducted via a semiparametric Generalized Additive Model for Location, Scale and Shape (GAMLSS) introduced by Rigby and Stasinopoulos (2005) and parameterization is as follows:

$$\ln\left(P_{ijt}\right)=\ \mathrm{X}_i\beta+\mathrm{Z}_{jt}\alpha+\mathrm{R}_t\gamma+\mu_t\theta_t+\mu_j\rho_j+\mathrm{u}_{ijt}. \tag{4}$$

The hedonic regression decomposes the log price P of a dwelling i in ZIP-code j and in observation period t into dwelling-specific characteristics X_i, ZIP-code-specific covariates Z_{jt} and the interest rate for 10-year housing loans R_t. As location is undoubtedly a key determinant of the price of a dwelling, besides the longitude and lattitude of each dwelling, ZIP-code fixed effects μ_j are included. μ_t captures the time fixed effects, thus is the focus of the index calculation. The error term u_{ijt} describes the variation in prices that cannot be explained by the model. In this case independently and identically distributed (u ~ iid) robust standard errors are used for the regression. As the time dummy index is defined as the marginal change in price P_{ijt} with respect to μ_t, a transformation of the estimated coefficients $\widehat{\theta_t}$ according to

$$\widehat{PI}_t = \left[\exp(\hat{\theta}_t) - 1\right] * 100, \tag{5}$$

yields the price index PI_t, referring to the percentage marginal change in prices in period t_t relative to t_0.

2.3.2 The Residential Transaction and Rental Market Liquidity Index

Without any doubt the leading model for the analysis of survival data is the Cox (1972) Proportional Hazards Model (PHM). This model is used for exploring the determinants of the duration of an event or elapse of time, e.g. it determines the variables that accelerate or restrict the elapse of time that a response variable needs to change its state. In this case, the response variable is defined as a non-negative continuous variable, measuring the elapse of time that a dwelling requires for changing its status from being offered on the market into being out of the market in weeks, i.e. time on market. For understanding and estimating survival data, two main functions are essential: the survival function *S(t)* and the hazard rate function *λ(t)*. The survival function specifies the probability that an event has not occurred until a certain time *t* and is formally defined as

$$S(t) = P\left(T \geq t\right) = 1 - F(t) = \int_t^{\infty} f(x)dx, \tag{6}$$

with *f(x)* being the probability density function (p.d.f.) of the time until the event. The hazard function *λ(t)*, in contrast, describes the probability at *t* that an event occurs at time *T*, given that the event has not occurred before and is given by

$$\lambda(t) = \lim_{\Delta t \to 0} \frac{P(t \leq T < t + \Delta t | t \leq T)}{\Delta t}. \tag{7}$$

The relationship between those two functions is straightforward since the integrated hazard rate $\Lambda(t) = \int_0^t \lambda(x)dx$ can be expressed as the negative log of the survival rate *S(t)* as *λ(t) = -logS(t)*. In other words, the survival function expresses the probability of a dwelling for staying in the market while the hazard function measures the risk of the same dwelling for leaving the market.

The Cox PHM estimates the survival function, but coefficients can be transformed to hazard rates, giving the probability of “mortality” per unit of time, and hence describing a liquidity indicator. The semiparametric Cox proportional hazards regression is parameterized as

$$\lambda(\tilde{t}_{ijt}) = \tilde{X}_i\tilde{\beta} + Z_{jt}\tilde{\alpha} + R_t\tilde{\gamma} + \mu_t\tilde{\theta}_t + \mu_j\tilde{\rho}_j + e_{ijt}. \tag{8}$$

The time on market $\tilde{t}$ of a dwelling *i* in ZIP-code *j* and in observation period *t* is decomposed into dwelling-specific characteristics $\tilde{X}_i$, ZIP-code-specific covariates Z_{jt} and the interest rate for 10-year housing loans R_t. In addition to *X*, $\tilde{X}$ includes the log of asking prices as the data generating process (DGP) of the time on market $\tilde{t}$ is influenced by the initial asking price, as landlords set the asking price when offering the dwelling in the MLS. As in the hedonic regression, time fixed effects μ_t are included, μ_j captures the ZIP-code fixed effects and e_{ijt} describes the error term. With $\exp(\widetilde{\theta_t})$ being defined as the hazard rate, the estimated coefficients $\widehat{\widetilde{\theta_t}}$ can be transformed into the liquidity index LI_t according to

$$\widehat{LI}_t = [\exp(\widetilde{\theta_t}) - 1] * 100, \tag{9}$$

referring to the percentage marginal change in the hazard rate, i.e. in liquidity, in period t_t relative to t_0.

2.3.3 Cluster Analysis

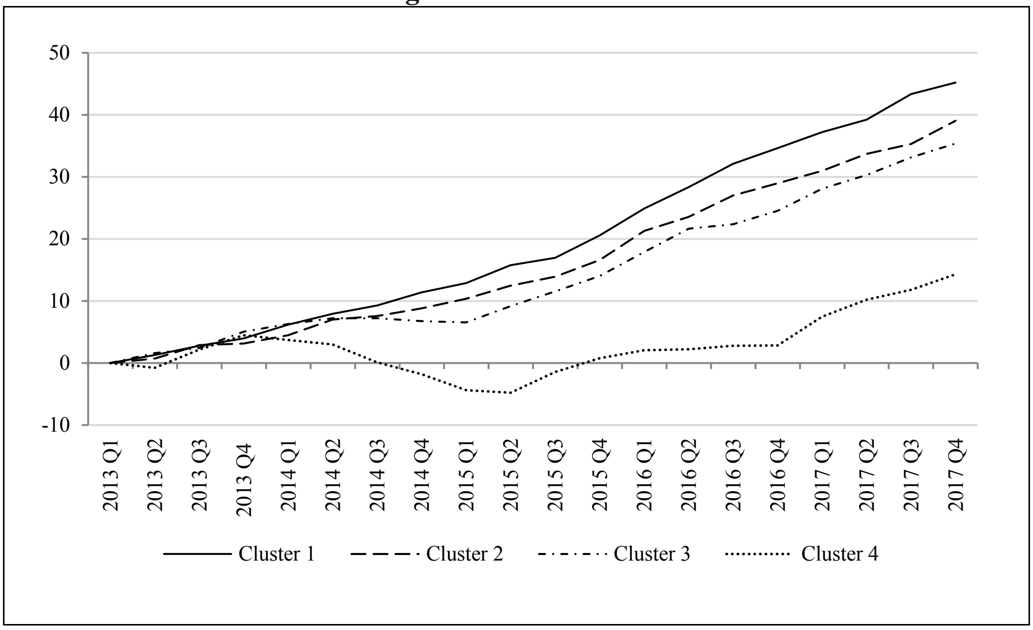
In order to determine regional markets that coincide according to their market movements, proceeding from the price and liquidity indices, the 161 regions are assigned to one of four clusters. The clustering is conducted separately for the price and liquidity indices on the transaction and rental market. The aim of the cluster analysis is to assign regions to the same cluster, so that the dissimilarity within a cluster is minimized and maximized between the clusters. Therefore, the “Partitioning Around Medoids (PAM)” clustering algorithm, going back to Kaufman and Rousseeuw (1987), is applied. The PAM clustering algorithm belongs to the *k*-medoids clustering procedure, that is closely related to the *k*-means procedure, however, is more robust to outliers and noise. While the *k*-means algorithm aims to minimize the sum of squared Euclidean distances, the *k*-medoids algorithm minimizes the average dissimilarity between the “representative” object, i.e. the medoid, and all other objects of the respective cluster. The algorithm consists of two major steps. At first, *k* initial objects are selected as medoids, i.e. these objects minimize the sum of the distances to all other objects. Second, the objective is to optimize the set of medoids. Therefore, each pair of medoid and remaining object is exchanged. If a swap indeed improves the cluster quality, the initial medoid and the other object change positions. This iteration is conducted until the quality of each cluster is optimal. The decisive variables underlying the clustering procedure are the estimated time-dummy coefficients θ_{Tp} and $\tilde{\theta}_{Tp}$ at each observation period *t*. *k* is chosen to be four, as the goal is to divide the regions in weak, rather weak, rather strong and strong developing markets in terms of price (rent) and liquidity. Based on these four categories and the actual progress of cluster means it should be possible to identify “hot” and “cold” regions.

2.4. Results

2.4.1 Investment Market

After estimating the price and liquidity indices for the 161 NUTS3 regions, each region is assigned to one of four clusters according to the methodology described in chapter 2.3.3. The regions with the highest increase in quality- and spatial-adjusted price are assigned to price cluster 1, whereas those with the highest increase in liquidity are allocated to liquidity cluster 1. Regions with the lowest increase in the attributes are allocated to cluster 4. Berlin for example, is assigned to cluster 1 for its price development and cluster 3 for its liquidity development. In the following, the city will be referred to as Berlin (1, 3).

Figure 2.3: Price index

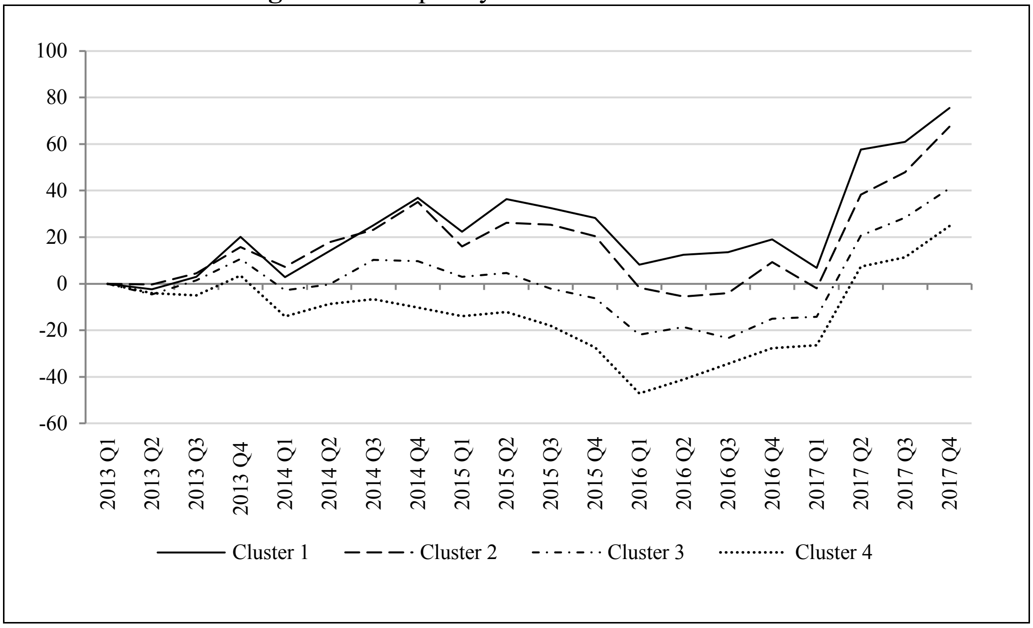


Notes: The figure plots the mean cumulative percentage price change for dwellings allocated to the individual clusters. The price changes are presented as the coefficients of the time dummy variable of a quality- and spatial-adjusted GAMLSS regression. To cluster the index values, the Partitioning Around Medoids (PAM) algorithm was used. The data consists of 973,164 observations on the residential investment market. The sample period is 2013 Q1 to 2017 Q4.

The trend of the quality- and spatial-adjusted cluster means is shown in figure 2.3 and reveals the cumulated average price change of all dwellings allocated to the particular clusters, indexed to zero in 2013 Q1. At the beginning of 2013, the clusters reveal an inconclusive development. During the second year, the clusters find their final rank order and maintain it until the end of the observation period. From then on, the clusters 1 to 3 display a quite similar progress, resulting in on average 45 to 35 pp higher prices in 2017 Q4 than in 2013 Q1. While cluster 1 and 2 show an almost linear and consistently positive trend, cluster 3 depicts falling prices for the quarters between 2014 Q3 and 2015 Q1. Price appreciation for dwellings in regions allocated to cluster 4 was surpassed by the other locations and experienced an increase of a mere 14 pp over the last five years. Regarding the remarkable development of real estate prices within the past years, "shelf warmer" locations might to be expected within this cluster. In fact, prices for dwellings allocated to cluster 4 even fell below the level of 2013 Q1 for more than a year between 2014 Q3 and 2015 Q4.

Assigning the regions to clusters by their liquidity development, which is based on the time it takes to sell a dwelling within the respective regions, displays a quite distinct pattern. While for cluster 2 only in five out of 20 quarters liquidity was worse than for the base quarter, this relation is found 10 times for cluster 3 and 15 out of 20 quarters for cluster 4. Nevertheless, all clusters finish with a liquidity level which is higher than for the base quarter.

Figure 2.4: Liquidity index investment market



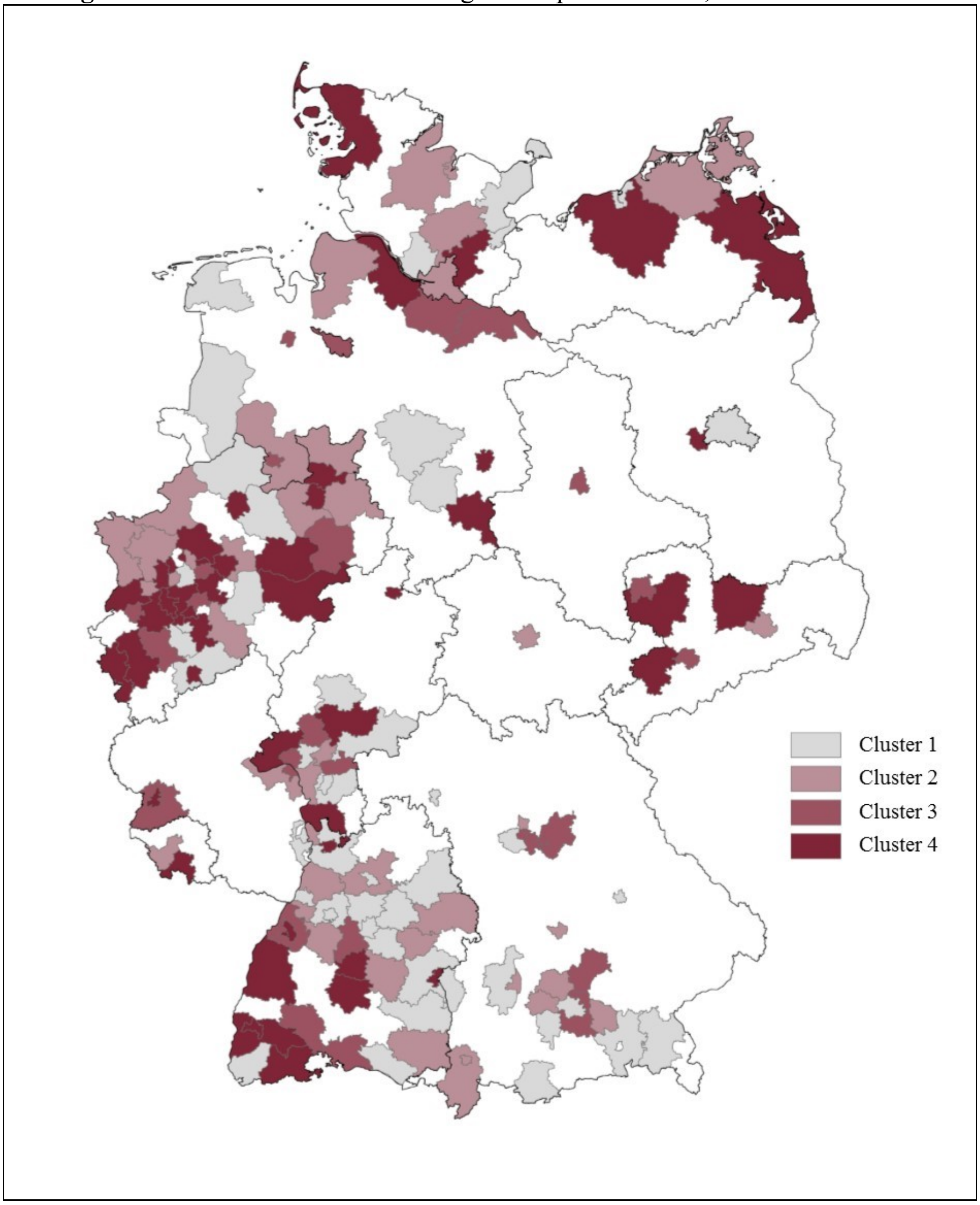
Notes: The figure plots the mean cumulative percentage change in liquidity for dwellings allocated to the individual clusters. The changes are presented as the coefficients of the time dummy variable of a quality- and spatial-adjusted Cox proportional hazards model. To cluster the index values, the Partitioning Around Medoids (PAM) algorithm was used. The data consists of 973,164 observations on the residential investment market. The sample period is 2013 Q1 to 2017 Q4.

As mentioned above, a combination of the cluster ranks derived from the price and liquidity development is used in order to classify the 161 regions. For this purpose, the price cluster rank is regarded as the primary determinant and the liquidity cluster rank as the complementary secondary determinant, which enables a higher granularity in the classification and a more precise market assessment. The full list of all 161 NUTS3 regions is found in the appendix.

2.4.1.1 Price Cluster 1

With Berlin (1, 3), Cologne (1, 2), Munich (1, 2), and Stuttgart (1, 1), four out of the German TOP 7 cities are assigned to price cluster 1. Although the regression controls for socioeconomic effects, cluster 1 consists to 85% of regions located in Baden-Wuerttemberg, Bavaria, Hesse, Lower Saxony, and North Rhine-Westphalia, the five economically strongest federal states measured by GDP. Six out of the 10 most populated counties in Baden-Wuerttemberg are assigned to cluster 1. As can be seen in figure 2.5, most of Baden-Wuerttemberg's regions which are assigned to cluster 1 are directly connected to the metropolis area around Stuttgart and expand the conurbation in all directions. As expected, the Bavarian capital Munich is allocated to cluster 1 together with other southern Bavarian growth regions. With Garmisch-Partenkirchen (1, 3), Rosenheim (1, 2) and Traunstein (1, 3), three of them are found in the alpine upland. In the northern half of Bavaria, the university cities Regensburg (1, 3) and Wuerzburg (1, 1) together with Fuerth county (1, 3), which is part of the metropolis area of Nuremberg, experienced a significant quality- and spatial-adjusted price increase for residential dwellings. Besides Kassel (4, 3), all regions in Hesse are located in the southern half of the state within the Rhine-Main metropolitan area around Frankfurt (2, 3). A price increase large enough to be allocated to cluster 1 was found outside the area's major economic center, indicating a strong demand for dwellings in surrounding regions for the last five years. Directly adjacent to Frankfurt are the regions Main-Taunus-Kreis (1, 2) and Offenbach (1, 1). Darmstadt (1, 1), Darmstadt county (1, 3), and Giessen (1, 2) are linked to the widespread public transport network of the Rhine-Main-Region, enabling frictionless commuting within the area and with that, the possibility to move out of the major conurbations. The most populated county of Lower Saxony, the region of Hanover (1, 3), is allocated to the cluster together with its adjacent region Hildesheim (1, 3). Two rather surprising representatives of Lower Saxony in cluster 1 are Aurich (1, 4) and Emsland (1, 4). The rural districts are in direct vicinity to the Netherlands. Aurich is famous for its touristic North Sea islands, which are mainly responsible for the high average price increase. Schaefer and Just (2018) have explored the importance of touristic attractiveness for real estate price development on the German market, among others. Emsland is one of the few rural regions that experience ongoing growth in its population. In addition, in both regions demand for residential real estate is fuelled by Dutch people moving to the area, while still working in the Netherlands. In North Rhine-Westfalia, the most populated city Cologne (1, 2) and its neighbouring districts Leverkusen (1, 1) and Rhein-Sieg-Kreis (1, 3) are allocated to cluster 1. Together with Bonn (4, 3), the districts form the second most populated region in North Rhine-Westfalia after the Ruhrgebiet. While the Ruhrgebiet is famous for outdated heavy industry, the area around Cologne is renowned for media and chemical industry. In the more northern part of the state, Essen (1, 3), Steinfurt, (1, 3), Warendorf (1, 3), and the Maerkische Kreis (1, 3) are assigned to the highest price cluster. Rostock (1, 1), the largest city of Mecklenburg, is the only eastern German district in the cluster besides Berlin (1, 3).

Figure 2.5: Allocation of NUTS3 regions to price clusters, investment market



Notes: The figure displays the geographic distribution of the individual clusters. The NUTS3 regions are allocated to a particular cluster by applying the Partitioning Around Medoids (PAM) algorithm on the price index values. The data consists of 973,164 observations on the residential investment market. The sample period is 2013 Q1 to 2017 Q4.

Out of the 48 regions allocated to cluster 1 by price, 15 are as well allocated to liquidity cluster 1. Those regions can be declared as absolute "hot" markets, where an extraordinary price development is supported by an equally strong liquidity development. About 60% of these absolute "hot" regions are located in Baden-Wuerttemberg, where besides Karlsruhe (1, 1), all of them are directly linked to the metropolis region of Stuttgart (1, 1), making this area the most performing German residential investment market for the last five years. This evaluation is reinforced by the fact, that of the 24 regions assigned to cluster 1 by liquidity, irrespective of the price cluster, 63% are located in Baden-Wuerttemberg. A similar relation is found for the regions allocated to liquidity cluster 2. As only in five out of 20 quarters quality- and spatial-adjusted liquidity was worse than in 2013 Q1, the regions are still characterized by continuous enhancement. In Baden-Wuerttemberg, the regions Loerrach (1, 2) and Bodenseekreis (1, 2), which are in direct vicinity to Switzerland, are found in this market segment together with Biberach (1, 2) and the Rhein-Neckar-Kreis (1, 2), both on the far ends of the Stuttgart metropolitan region. In Bavaria, Munich (1, 2) and Starnberg (1, 2), Germany's wealthiest county, are the most prominent representatives.

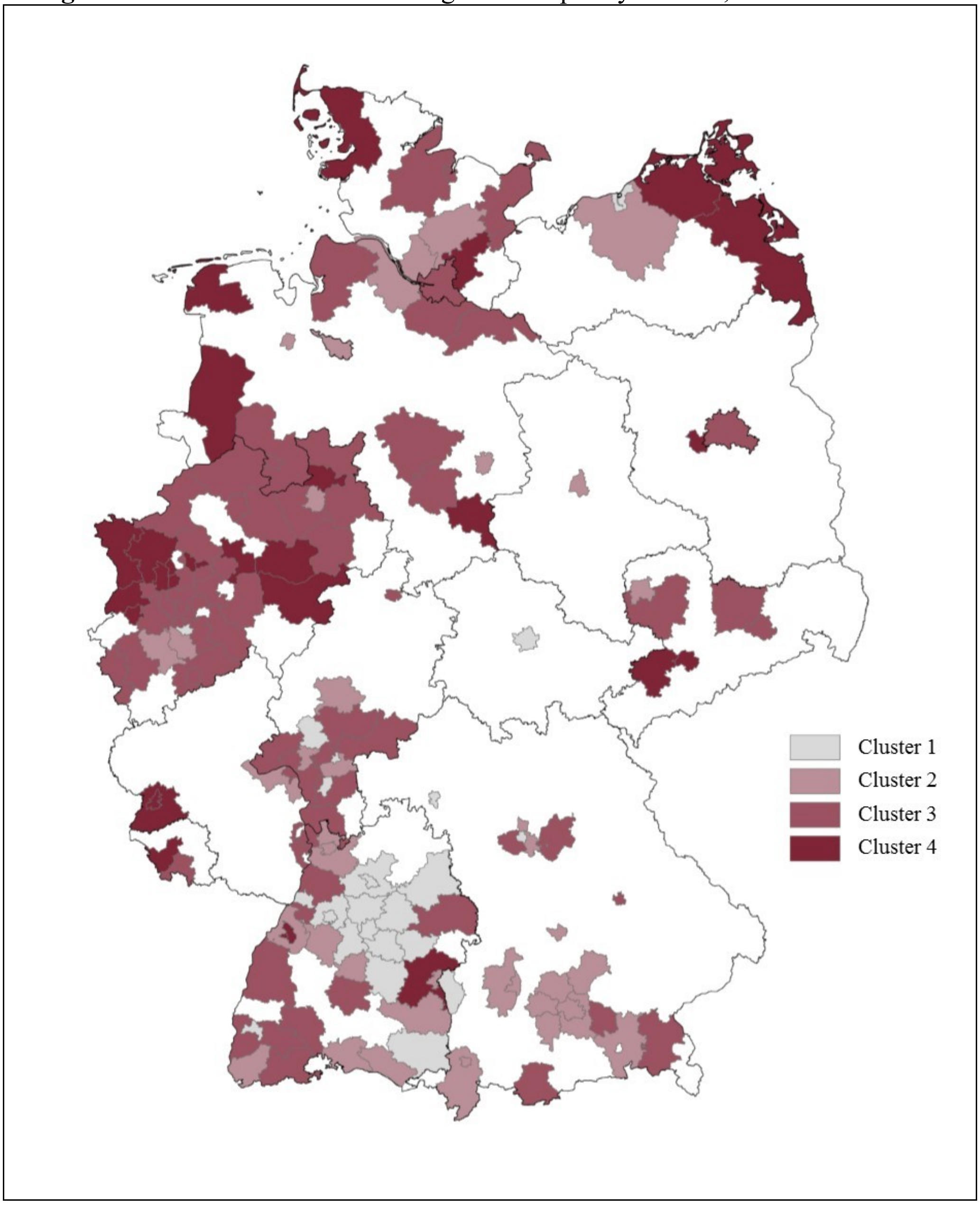
Observing the development of the NUTS3 regions over time, by clustering the cities for different time frames, shows a quite specific pattern. While most of the regions in liquidity cluster 1 and 2 either improve their price cluster or at least stay in the particular cluster, many regions allocated to liquidity cluster 3 and 4 descend in the price rank. One of the most decisive patterns is the persistence of almost all cities allocated to cluster 3, 3 and below.

As depicted in figure 2.6, liquidity cluster 1 and 2 are predominantly located in the southern half of the country, while cluster 3 and 4 are dominating the northern and western part. This finding is independent of the particular price cluster. Within price cluster 1 for example, 75% of the regions allocated to liquidity clusters 3 and 4 are in the northern half of the country. Among the regions in these clusters are the old industrial city Essen (1, 3) and the very remote regions Aurich (1, 4) and Emsland (1, 4). Although the quality- and spatial-adjusted price increase in those regions has been strong, they experienced negative liquidity development compared to 2013 Q1, indicating a rather relaxed market. According to Deschermeier et al. (2017), Emsland (1, 4) is one of the few German districts where more new dwellings are built than needed. At the same time, this assessment might underestimate the cross boarder demand from the Netherlands.

2.4.1.2 Price Cluster 2

With a consistent increase in prices, the appreciation of dwellings allocated to cluster 2 is almost as strong as for those allocated to cluster 1. Since 80% of all regions allocated to cluster 2, 1 and 2, 2 are located in Baden-Wuerttemberg and southern Bavaria, the tightness of the residential investment market is reinforced. The conurbation around Stuttgart (1, 1) again dominates the liquidity cluster 1, while counties in southern Bavaria dominate the subsequent liquidity clusters. Erfurt, (2, 1), the capital of Thuringia, might improve its price cluster with ongoing demand for investments. In the lower liquidity clusters, many large cities from northern and western states are found. The largest of them are the two TOP 7 cities Frankfurt am Main (2, 3) and Hamburg (2, 3). While for the residential market in Frankfurt further tightening is expected due to limited new supply and strong demographic inflow, Hamburg was able to manage the supply and demand gap by increasing building permits hence building completions, while at the same time fewer people moved to Hamburg. An agglomeration of regions in the lower liquidity clusters is found in the Ruhrgebiet, North Rhine-Westfalia's former mining and steel industry hub. Consistently negative liquidity development might result in a weakening price development and lower position in the price rank order.

Figure 2.6: Allocation of NUTS3 regions to liquidity clusters, investment market



Notes: The figure displays the geographic distribution of the individual clusters. The NUTS3 regions are allocated to a particular cluster by applying the Partitioning Around Medoids (PAM) algorithm on the liquidity index values. The data consists of 973,164 observations on the residential investment market. The sample period is 2013 Q1 to 2017 Q4.

2.4.1.3 Price Cluster 3

Price cluster 3 is the last cluster to show a price increase of more than 30%, however, it differentiates itself from cluster 1 and 2 by virtually constant prices from 2014 Q1 to 2015 Q1. The cluster displays the lowest number of regions among all price clusters. Because of strong quality- and spatial-adjusted liquidity performance, the Hochtaunuskreis (3, 1) just north of Frankfurt (2, 3) with its affluent municipalities Bad Homburg, Koenigstein, and Kronberg, might experience higher price pressure as do Fuerth (3, 1) and Boeblingen (3, 1). With slightly more than doubled liquidity, regions allocated to liquidity cluster two like e.g. Munich county (3, 2), Freising (3, 2) in direct vicinity to Munich, and Leipzig city (3, 2), might participate in a liquidity triggered price increase. As already mentioned, regions within the lower liquidity clusters, like e.g. Paderborn (3, 3) and Chemnitz (3, 4), are very likely to descend in price rank or to maintain it.

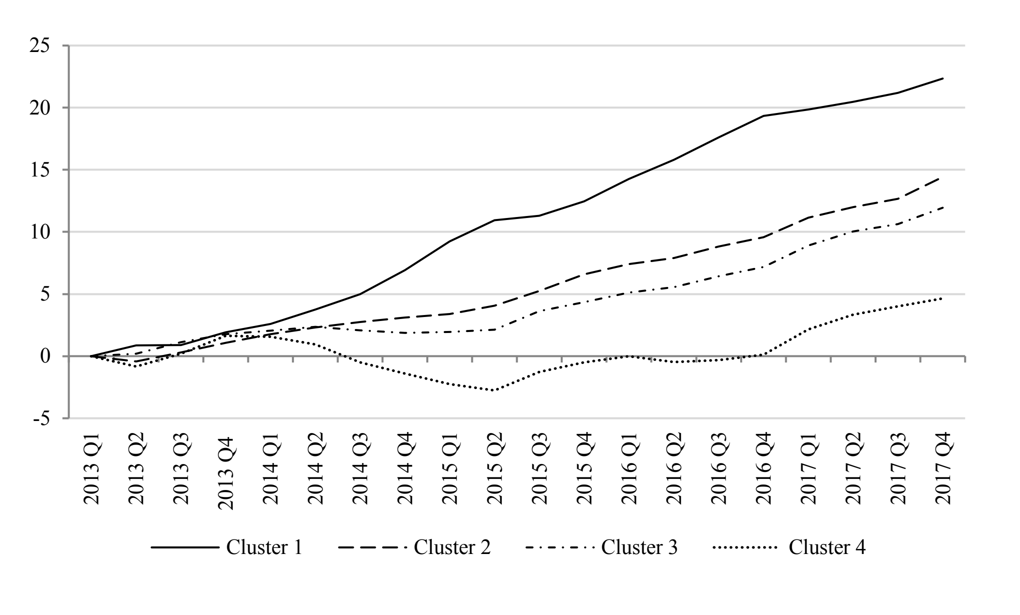
2.4.1.4. Price Cluster 4

Regions assigned to cluster 4 experienced the least strong price increase, while quality- and spatial-adjusted prices even fell for slightly more than a year during the observation period. Those regions display a geographic concentration in the southern part of Baden-Wuerttemberg, the Ruhrgebiet and eastern states. People familiar with the German market might be surprised to find the university cities Freiburg (4, 1), Heidelberg (4, 2), and Tuebingen (4, 2) in price cluster 4. Although high liquidity development points to a future price increase, these regions are known as strong performing markets and are not expected to be among the worst performers. Other regions which might not be expected among the worst performers are e.g. Bergstrasse (4, 3), Rheingau-Taunus-Kreis (4, 3), Duesseldorf city (4, 3), Bonn (4, 3), Muenster city (4, 3), Baden-Baden city (4, 4), and Potsdam city (4, 4). This classification might be due to the limited sample period covering only the five years from 2013 Q1 to 2017 Q4. Within the Ruhrgebiet, regions like for example Dortmund (4, 3), Recklinghausen (4, 3), Duisburg (4, 4), Herne (4, 4), and Oberhausen (4, 4) are assigned to price cluster 4. Slightly outside of the Ruhrgebiet, Soest (4, 4), Solingen (4, 3) and Mettman (4, 3) are allocated to cluster 4 Because of a combination of low price appreciation and persistently weak liquidity behaviour, the residential investment market in these "cold" regions is very unlikely to strengthen. While among the eastern regions, prices in Rostock county (4, 2) might benefit from the city's strong liquidity development on the investment market, Leipzig (4, 3), Meissen (4, 3), and Zwickau (4, 4) seem to lack liquidity for a price improvement.

2.4.2 Rental Market

Figure 2.7 shows the average cumulated change in quality- and spatial-adjusted rent for the dwellings allocated to the respective clusters. Those clusters do not necessarily contain the same regions as the price clusters. Cluster 1 sets itself apart by a very strong and consistently positive development. Nevertheless, the final index value of 22.3 is just half of the final index value for the respective quality- and spatial-adjusted price cluster, underlining the comparatively stronger increase in prices on the investment market than the rental market. Clusters 2 and 3 display a quite similar progress, while cluster 4 contains regions where rent even decreased in comparison to 2013 Q1 for 9 quarters between 2014 Q3 and 2016 Q3. Over the whole observation period, rents for dwellings allocated to cluster 4 increased a mere 4.6 pp, which is even less than the cumulated inflation rate. The analysis of the results of the quality- and spatial-adjusted hedonic price and rent indices is not only able to confirm the findings based on pure descriptive statistics but can assure, that this diverging progress is not attributable to regional effects or the heterogeneity of the dwellings themselves.

Figure 2.7: Rental index



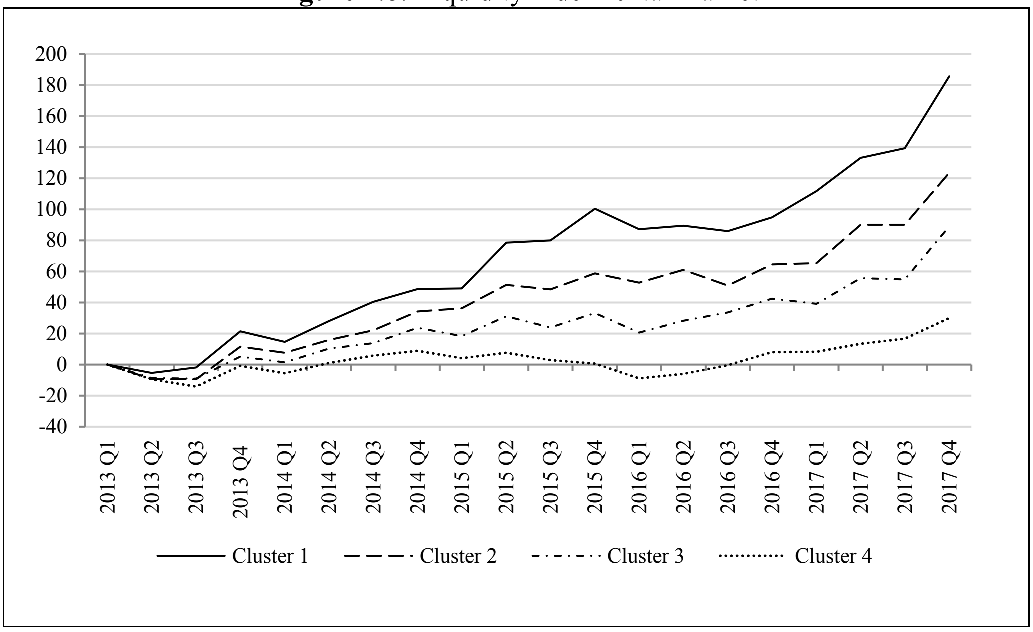
Notes: The figure plots the mean cumulative percentage price change for dwellings allocated to the individual clusters. The price changes are presented as the coefficients of the time dummy variable of a quality- and spatial-adjusted GAMLSS regression. To cluster the index values, the Partitioning Around Medoids (PAM) algorithm was used. The data consists of 2,082,179 observations on the residential rental market. The sample period is 2013 Q1 to 2017 Q4.

The liquidity development on the rental market depicted in figure 2.8, however, distinguishes itself from the investment market by a far stronger and almost consistently positive progress. After 2014 Q1 only in four quarters in cluster 4, liquidity was weaker than in 2013 Q1. The vastly rising demand pressure for living space is much better reflected in the liquidity on the rental market, resulting in an almost tripled liquidity for dwellings in regions allocated to cluster 1 and more than doubled liquidity for regions allocated to cluster 2 in 2017 Q4. Considering the country's comparatively low homeownership rate, this finding is not very surprising. It is of particular interest, however, that this positive development is observed across the whole country, while the only cluster on the investment market with consistently higher liquidity than in 2013 Q1 was mainly found in the southern parts of the country. In addition to that, the liquidity clusters will again be used as a complementary determinant to evolve a deeper understanding of the market and to detect "hot" and "cold" spots.

2.4.2.1 Rental Cluster 1

In opposite to the investigation of the investment market, the rental cluster 1 is the smallest cluster containing 30 regions. Nevertheless, as depicted in figure 2.9, strong quality- and spatial-adjusted rental development is found in eight different states. It is not surprising to find TOP 7 cities like Berlin [1, 2] and Munich [1, 2] and regions already allocated to cluster 1 on the investment market like e.g. Ludwigsburg [1, 1], Pforzheim [1, 1], Rosenheim [1, 2], and Wuerzburg [1, 3]. However, there are more surprising regions like e.g. Zollernalbkreis [1, 1] (4, 1 on the investment market), Oberhausen [1, 3] (4, 4), and Dueren [1, 3] (4, 3). Dwellings located in those cities experienced a tremendous rental growth, while they are among the worst performing on the investment market. Of course, it is also possible to observe the reverse behaviour. For example in the counties Starnberg [4, 2] (1, 2) and Main-Taunus-Kreis [4, 2] (1, 2), both well known for top-level purchasing power, the rental market lags far behind the investment market. Potential inhabitants of these particular regions, which are to a large extent consisting of very exclusive single-family homes, might have a higher preference for owning one of these properties. Therefore, the regional rental market is not as developed as the investment market.

Figure 2.8: Liquidity index rental market

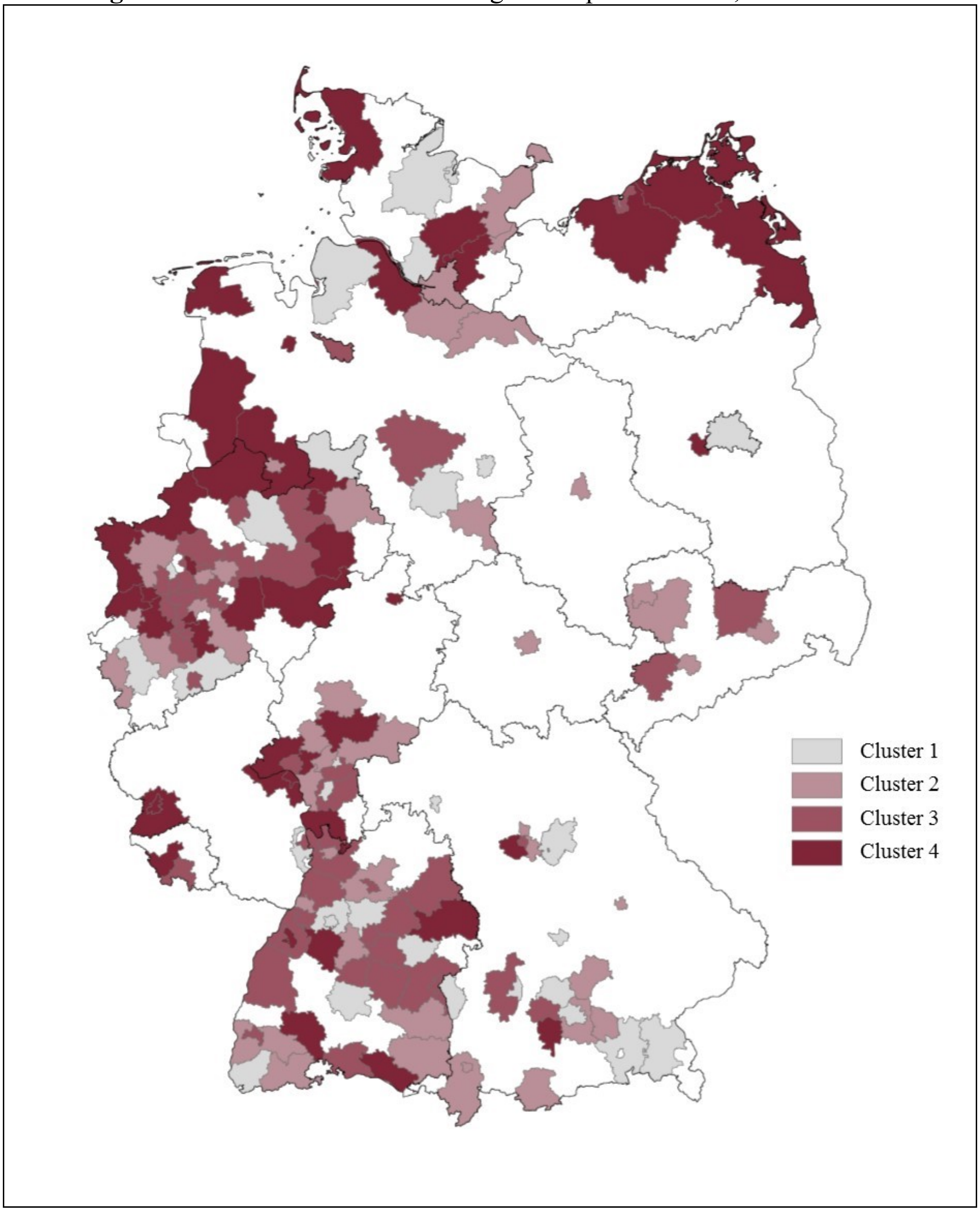


Notes: The figure plots the mean cumulative percentage change in liquidity for dwellings allocated to the individual clusters. The changes are presented as the coefficients of the time dummy variable of a quality- and spatial-adjusted Cox proportional hazards model. To cluster the index values, the Partitioning Around Medoids (PAM) algorithm was used. The data consists of 2,082,179 observations on the residential rental market. The sample period is 2013 Q1 to 2017 Q4.

Of those 30 regions, seven are as well allocated to liquidity cluster 1. Thus, Goeppingen [1, 1] and Zollernalbkreis [1, 1], both in Baden-Wuerttemberg, Dachau [1, 1] in direct vicinity to Munich [1, 2] and Hildesheim [1, 1] are among the “hottest” German rental markets. The regions are characterized by the strongest possible rental increase and tremendous liquidity development. Pforzheim [1, 1], which is surrounded by Enzkreis [1, 1], thus directly adjacent to Ludwigsburg

[1, 1] just north and north-west of Stuttgart [3, 1], are among the "hottest" markets on both the investment and the rental market. Therewith, these three regions are the most sought after, thus best performing residential markets in Germany. Only the simultaneous classification by quality- and spatial-adjusted price and rent development in combination with the corresponding liquidity, enables a classification with such a high level of detail.

Figure 2.9: Allocation of NUTS3 regions to price clusters, rental market



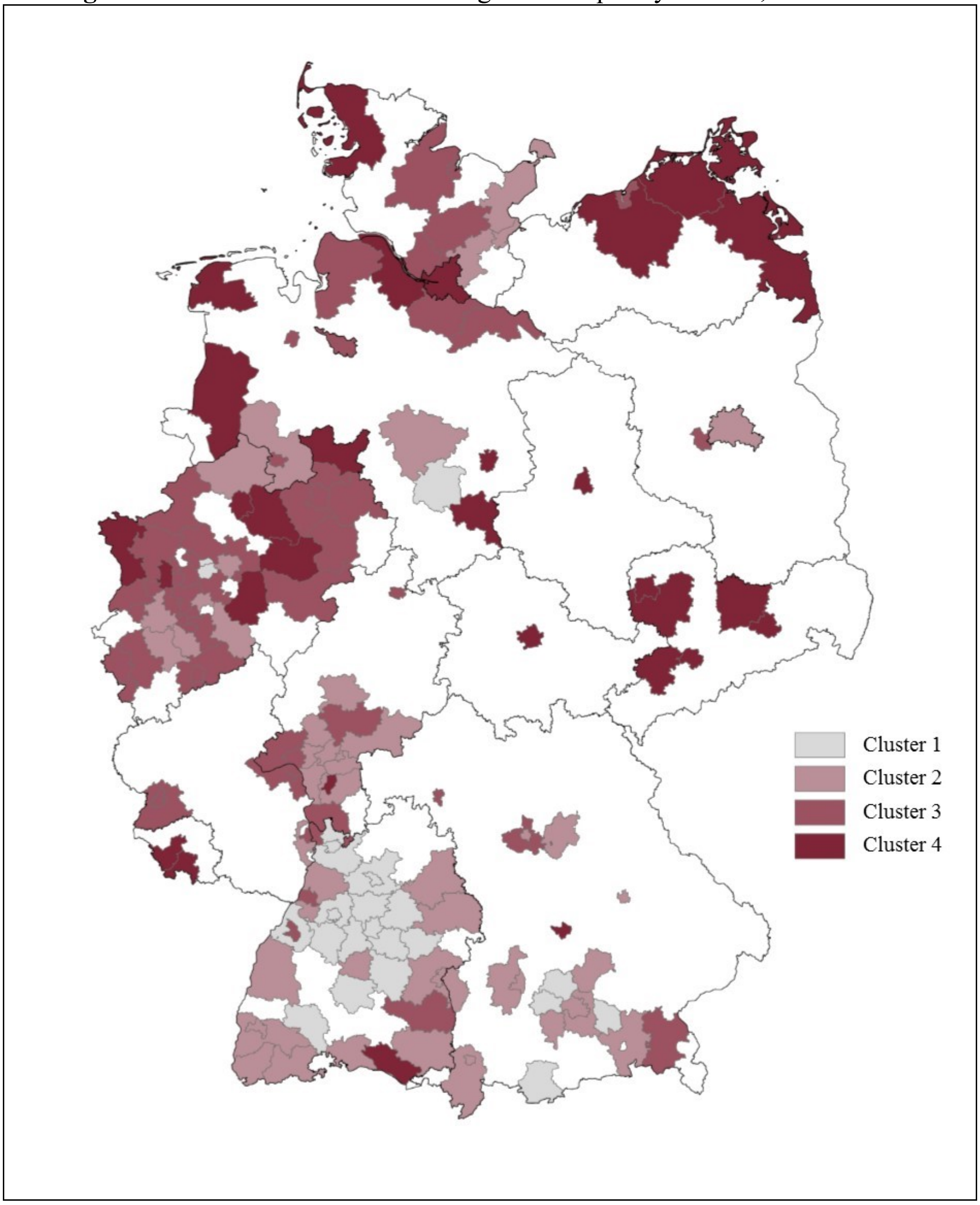
Notes: The figure displays the geographic distribution of the individual clusters. The NUTS3 regions are allocated to a particular cluster by applying the Partitioning Around Medoids (PAM) algorithm on the price index values. The data consists of 2,082,179 observations on the residential rental market. The sample period is 2013 Q1 to 2017 Q4.

The fact, that three out of the seven "hottest" rental markets are in direct vicinity of each other, is a great example for one of the rental markets distinctive characteristics. While on the investment market, the metropolis region Stuttgart and the Ruhrgebiet, are two very extensive conurbations which reveal a consistent development in both price and liquidity behaviour, there is a considerably higher number of neighbouring regions being identically clustered on the rental market. This characteristic is found across all price and liquidity clusters and indicates, that spillover effects are more likely on the rental market. A possible explanation might be the higher preference for a specific location when buying a home in comparison to renting it. When it comes to renting a dwelling, potential tenants seem to be willing to accept a location within a wider area, thus allowing the rental markets more space for joint development. By that, the price and liquidity development of e.g. Pforzheim [1, 1] was able to expand to the surrounding Enzkreis [1, 1]. In addition to equally behaving markets in direct vicinity, there is a high concentration of regions in cluster 2, 4 in the eastern states. Although only Leipzig city [2, 4] and Leipzig county [2, 4] are direct adjacent, the spillover effects seem to stretch to Dresden [2, 4] and Chemnitz [2, 4], both Saxon cites, and Magdeburg [2, 4] and Erfurt [2, 4] located in Saxony-Anhalt and Thuringia respectively.

2.4.2.2 Rental Clusters 2 and 3

With 44 regions each, rent cluster 2 and 3 not only behave almost identically, but also contain the same number of regions. With 43.2% of all regions allocated to rental cluster 2, the highest density of university cities on the rental market is found for cluster 2, followed by cluster 3 with 34.1%. On the investment market on the other hand, the highest density of university cities is found for cluster 4, the worst performing regions in terms of price appreciation. While rental development according to cluster 2 and 3 is found across the whole country, regions allocated to the corresponding liquidity cluster 1 are exclusively located in Bavaria, Baden-Wuerttemberg and North Rhine-Westfalia, indicating a strong dependence on economically solid regions. The corresponding liquidity cluster 4 on the other hand shows a significant proportion of regions located in eastern states.

Figure 2.10: Allocation of NUTS3 regions to liquidity clusters, rental market



Notes: The figure displays the geographic distribution of the individual clusters. The NUTS3 regions are allocated to a particular cluster by applying the Partitioning Around Medoids (PAM) algorithm on the liquidity index values. The data consists of 2,082,179 observations on the residential rental market. The sample period is 2013 Q1 to 2017 Q4.

2.4.2.3 Rental Cluster 4

A similar but weaker relationship is found for cluster 4. While for example the Bavarian county Starnberg [4, 2] and Main-Tauber-Kreis [4, 2] (Baden-Wuerttemberg) are assigned to the higher liquidity clusters, regions in the structurally weaker northern and western states are found for the subsequent liquidity clusters. As for the investment market, regions in the Ruhrgebiet are allocated to liquidity cluster 3 together with more rural regions of Hesse. With Rostock county [4, 4], Vorpommern-Ruegen [4, 4], and Vorpommern-Greifswald [4, 4] the cluster containing the "coldest" rental markets, however, is dominated by regions in Mecklenburg and Lower-Saxony.

Among the regions in Lower-Saxony are Aurich [4, 4] and Emsland [4, 4], which have been allocated to price cluster 1 for the investment market. As the least demanded regions on the entire German residential market, Vorpommern-Greifswald [4, 4] (4, 4) and Nordfriesland [4, 4] (4, 4) are identified.

When analysing both the investment and the rental market, it is possible to detect 23 regions, which are identically clustered on the investment and the rental market. Among them are for example Ludwigsburg [1, 1] (1, 1), Heilbronn [2, 1] (2, 1), Wiesbaden [3, 3] (3, 3) and Solingen [4, 3] (4, 3). Obviously, the absolute magnitude of the change in price and liquidity is not the same on both markets but apparently, those regions behave equally in proportion to the other regions on both markets. Heilbronn for example is in the second best performing cluster on the investment and the rental market, while the respective liquidity is both times in the best performing cluster. This might indicate a “fair” relation of investment market prices to rental prices.

2.4.3 Investment Strategies and Policy Implications

Based on a very granular analysis of the 161 regions and the subsequent classification of the regions by price, rent, and the corresponding market liquidity, it is now possible to derive ideas for investment strategies. In the following, the four most differentiated cases will be discussed.

2.4.3.1 High Price and High Rent Cluster

Apparently, these regions already experienced strong price development on both the investment and rental market and seem to be fairly priced, as the markets move along. Nevertheless, the liquidity clustering can be used as a guideline for ongoing demand in order to speculate on further positive development. In case liquidity is high on both markets and the analysis of the socioeconomic outlook is positive, it should still be possible to participate in rising prices and rents. A higher liquidity on the rental market than on the investment market seems like a more appealing combination for the investor, as consumers of living space face a tighter market with less options. Therefore, they might be more willing to accept higher monthly rent payments, which will in turn result in price appreciation. In addition to that, a slightly more relaxed investment market facilitates the acquisition for the investor. Of course an identification of bargains is less likely in the strongest performing regions like e.g. Ludwigsburg [1, 1] (1, 1), Munich city [1, 2] (1, 2), Berlin [1, 3] (1, 2), and Hildesheim [1, 1] (1, 3). More risk friendly investors might have a look at the subsequent price clusters containing worse performing markets, in order to identify a liquidity mismatch in otherwise fairly priced markets like e.g. Ebersberg [2, 1] (2, 3), and Frankfurt [2, 2] (2, 3).

2.4.3.2 Low Price and Low Rent Cluster

If despite the current market cycle, a region displays a low price and rent clustering, the region seems not to be within the top spot for investment opportunities. If however, liquidity is high for both or at least one market, this might be an initial signal for rising prices. By all means, a very detailed further analysis of fundamentals and socioeconomic outlooks has to be conducted for regions like e.g. Rostock county [4, 4] (4, 2), Rhein-Kreis Neuss [4, 2] (4, 3), and Stormann [4, 2] (4, 4).

2.4.3.3 Lower Price than Rent Cluster

This is the most appealing combination for investors, as it enables the acquisition of dwellings at a high net initial yield thus indicating an “underpriced” region. Besides relatively high rental payments, the investor might furthermore benefit from rising prices, as far as it is possible to rule out that the very positive rental development did arise as a consequence of a price appreciation previous to the observation period. A high liquidity cluster on the investment market might eliminate this concern and simultaneously hint to a future increase in prices. As these regions appear to be the best investment opportunities, the very obvious ones are of course rather scarce. Potential target regions are e.g. Braunschweig [1, 4] (4, 2), Zollernalbkreis [1, 1] (4, 3), Dueren [1, 3] (4, 3), Oberhausen [1, 3] (4, 4), Heidelberg [2, 1] (4, 2), Nuremberg county [1, 2] (3, 3), and Ingolstadt [1, 4] (2, 2).

2.4.3.4 Higher Price than Rent Cluster

While the previous investment idea is more promising the higher the spread between the investment and the rental market, the contrary situation is desirable for a region allocated to a higher price than rent cluster. Besides a large price to rent ratio, indicating a rather “overpriced” market, a higher spread might result in the rental market not being able to catch up to the appreciation on the investment market. Hence, an investor would be more dependent on a further yield compression. The higher the liquidity on both markets, the more attractive this scenario is for investors du to apparently strong demand. Possible regions to invest based on this scenario are e.g. Garmisch-Partenkirchen [2, 1] (1, 3), Regensburg [2, 2] (1, 3), Giessen [2, 2] (1, 2), Reutlingen [3, 1] (2, 1), and Fuerstenfeldbruck [3, 1] (2, 2).

2.4.3.5 Policy Implications

The affordability of living space is a very topical issue on the German market, as especially in cities and metropolitan regions, large parts of the population are not able to create ownership or have to accept a monthly rent at or even above 50% of the monthly household income. Whereas this study is not able to derive strategies for providing affordable living space or easing tight residential markets, it might assist policy makers in detecting overheating markets and facilitate a categorization of regions according to their market tightness. Herewith, policy makers should have a better understanding of which regions they should focus on and where market interventions are most urgent. The separate clustering enables policy makers to derive regionally varying strategies based on the pricing of markets, the tightness of markets measured by liquidity, and a combination of both.

2.5 Conclusion

It is the aim of this paper to build quality- and spatial-adjusted price indices for the major German residential investment and rental markets on a regional basis and to complement each with a liquidity index, in order to obtain a very detailed assessment of the German residential real estate market. In the context of this study, liquidity is solely examined with a time-based measure and does not include transaction cost, price, or volume measures.

While the mere analysis of descriptive statistics on the investment and rental market already indicated a trend of rising prices and declining time on market, the analysis of quality- and spatial-adjusted price and rent indices eliminates the possibility that these trends are due to changes in the housing stock. The regions are then assigned to one of four clusters based on the partitioning around medoids (PAM) clustering algorithm, in order to identify common market movements and to facilitate the interpretation of the results for the 161 regions. Of the 402 NUTS3 regions, only active residential markets with more than 100 offers per quarter are included in the study. Over the observation period from 2013 Q1 to 2017 Q4 and based on more than 3 million observations, the increase in prices on the investment market was far stronger than the increase on the rental market. The divergence of the markets is emphasized by a much stronger liquidity development on the rental market in comparison to the investment market. Prices for dwellings assigned to cluster 1 rose about twice as much as rents for dwellings assigned to the same cluster. On the other hand, liquidity on the rental market for dwellings assigned to cluster 1 almost tripled, while liquidity on the investment market for dwellings assigned to the same cluster not even doubled. This finding reveals, that the strong demand on the rental market is much better reflected by the liquidity index in comparison to the moderate development of rental prices.

The regional analysis of the price and rent clusters yields a diversified pattern of strong investment and rental markets. While a slight concentration on southern states might be indicated, high performing markets in terms of price are found across the whole country. Only the combined classification of price and liquidity clusters reveals a strong focus on Baden-Wuerttemberg. For the investment market, the metropolis region of Stuttgart is clearly the strongest performing region with eight districts assigned to the highest cluster for price and liquidity. For the rental market, parts of this region are among the "hottest" markets. Many of the least performing regions on both markets are found in structurally weak regions in North Rhine-Westfalia and eastern states. In addition, the regional analysis suggest stronger spillover tendencies on the rental market, as a larger number of adjacent regions which experienced an identical development was found.

Those findings are of course limited by the rather short sample period of only five years. Price and liquidity development happening before 2013 Q1 could not be incorporated because of the absence of data.

Based on the joint classification, it is possible to derive investment strategies for different combinations of price, rent and the respective market liquidity. The classification might also assist policy makers on the identification of tight markets and a prioritization of subsequent actions.

A peculiarity of the analysis described in the paper is, that it is easily applicable to other residential markets. While it is not possible to use transaction prices and contracted rent on the German market, it might be one of the most interesting extensions of the model together with an analysis of the intertemporal relationship of the investment and the rental market. Furthermore, this classification can be used to extend the price versus rent literature.

2.6 References

an de Meulen, P. and Mitze, T. (2014), “Exploring the spatial variation in quality-adjusted rental prices and identifying hot spots in Berlin‘s residential property market” *Regional Studies, Regional Science*, Vol. 1 (1), pp. 310-328.

Bauer, T. K., Feuerschuette, S., Kiefer, M., an de Meulen, P., Micheli, M., Schmidt, T., and Wilke, L. H. (2013), “Ein hedonischer Immobilienpreisindex auf Basis von Internetdaten: 2007–2011“, *AStA Wirtschafts-und Sozialstatistisches Archiv*, Vol. 7 (1-2), pp. 5-30.

Bentzien, V., Rottke, N., and Zietz, J. (2012), “Affordability and Germany’s low homeownership rate”, *International Journal of Housing Markets and Analysis*, Vol. 5 (3), pp. 289-312.

Berkovec, J.A. and Goodman, J.L. Jr. (1996), “Turnover as a measure of demand for existing homes”, *Real Estate Economics*, Vol. 24 (4), pp. 421-440.

Cajias, M. (2018), “Is there room for another hedonic model? The advantages of the GAMLSS approach in real estate research”, *Journal of European Real Estate Research*, Vol. 11 (2), pp. 204-245.

Cauley, S.D. and Pavlov, A.D. (2002), “Rational delays: The case of real estate”, *The Journal of Real Estate Finance and Economics*, Vol. 24 (1-2), pp. 143-165.

Court, A.T. (1939), “Hedonic price indexes with automotive examples”, in: *The Dynamics of Automobile Demand*, pp. 99-117, New York; The General Motors Corporation.

Cox, D.R. (1972), “Regression models and life-tables”, *Journal of the Royal Statistical Society: Series B (Methodological)*, Vol. 34 (2), pp. 187-220.

Deschermeier, P., Henger, R., Seipelt, B., and Voigtlaender, M. (2017), “Wohnungsmangel in den Staedten, Leerstand auf dem Land”, IW-Kurzberichte, 44.2017.

Diewert, E.W. (2003), “Hedonic regressions: A review of some unresolved issues”, in *7th Meeting of the Ottawa Group*, Paris, May, pp. 27-29.

Federal Statistical Office, “Preisindizes fuer Wohnimmobilien (Code 61262-0001)”, available at: https://www-genesis.destatis.de/genesis/online (accessed: May, 24 2018)

Federal Statistical Office, “Verbraucherpreisindex fuer Deutschland (Code 61111-0003)”, available at: https://www-genesis.destatis.de/genesis/online (accessed: May, 24 2018)

Fisher, J., Gatzlaff, D., Geltner, D., and Haurin, D. (2003), “Controlling for the impact of variable liquidity in commercial real estate price indices”, *Real Estate Economics*, Vol. 31 (2), pp. 269-303.

Follain, J.R. and Velz, O.T. (1995), “Incorporating the number of existing home sales into a structural model of the market for owner-occupied housing”, *Journal of Housing Economics*, Vol. 4 (2), pp. 93-117.

Genesove, D. and Mayer, C.J. (1997), “Equity and time to sale in the real estate market”, *American Economic Review*, Vol. 87, pp. 255-269.

Goldberger, A.S. (1968), “The interpretation and estimation of Cobb-Douglas functions”, *Econometrica: Journal of the Econometric Society*, Vol. 36 (3-4), pp. 464-472.

Griliches, Z. (1961), “Hedonic price indexes for automobiles: An econometric of quality change”, in: *The Price Statistics of the Federal Government*, Report of the Price Statistics Review Committee, National Bureau of Economic Research, pp. 173-196.

de Haan, J. (2010), “Hedonic price indexes: A comparison of imputation, time dummy and ‘re-pricing’ methods”, Jahrbuecher fuer Nationaloekonomie und Statistik, Vol. 230 (6), pp. 772-791.

Halvorsen, R. and Pollakowski, H.O. (1981), “Choice of functional form for hedonic price equations”, *Journal of Urban Economics*, Vol. 10 (1), pp. 37-49.

Heinrich M. and Just, T. (2016), “Wie heterogen sind deutsche Staedte? Zur Segmentierung deutscher Wohnungsmaerkte“, Beitraege zur Immobilienwirtschaft, Vol. 13.

Hill, R.J., Melser, D., and Syed, I. (2009), “Measuring a boom and bust: The Sydney housing market 2001-2006”, *Journal of Housing Economics*, Vol. 18 (3), pp. 193-205.

Hort, K. (2000), “Prices and turnover in the market for owner-occupied homes”, *Regional Science and Urban Economics*, Vol. 30 (1), pp. 99-119.

Kain, J.F. and Quigley J.M. (1970), “Measuring the value of housing quality”, *Journal of the American Statistical Association*, Vol. 65 (330), pp. 532-548.

Kaufman, L. and Rousseeuw, P.J. (1987), “Clustering by means of medoids”, in: *Statistical data analysis based on the L1-norm and related methods*, North-Holland, New York, pp. 405-416.

Kennedy, P.E. (1981), “Estimation with correctly interpreted dummy variables in semilogarithmic equations”, *American Economic Review*, Vol. 71 (4), pp. 801.

Kluger, B.D. and Miller, N.G. (1990), “Measuring residential real estate liquidity”, *Real Estate Economics*, Vol. 18 (2), pp. 145-159.

Krainer, J. (1999), “Real estate liquidity”, *Economic Review Federal Reserve Bank of San Francisco*, Vol. 3 (1), pp. 14-26.

Lancaster, K.J. (1966), “A new approach to consumer theory”, *Journal of political economy*, Vol. 74 (2), pp. 132-157.

Lerbs, O. and Oberst, A. (2014), “Explaining the spatial variation in homeownership rates: Results for German regions”, *Regional Studies*, Vol. 48 (5), pp. 844-865.

Lyons, R.C. (2013), “Price signals in illiquid markets – The case of residential properties in Ireland, 2006-2012”, Trinity Economics Papers tep0616, Trinity College Dublin, Department of Economics.

Malpezzi, S. (2003), “Hedonic pricing models: A selective and applied review”, in: *Housing Economics and Public Policy: Essays in honor of Duncan Maclennan,* Oxford, Blackwell, pp. 67-89.

Ortalo-Magné, F. and Rady, S. (2004), “Housing transactions and macroeconomic fluctuations: A case study of England and Wales”, *Journal of Housing Economics*, Vol. 13 (4), pp. 287-303.

Pollakowski, H.O. (1982), “Urban housing markets and residential location”, Lexington, MA: D. C. Heath and Company.

Reisenbichler, A. (2016), “A rocky path to homeownership: Why Germany eliminated large-scale subsidies for homeowners”, *Cityscape: Ajournal of policy development and research*, Vol. 18 (3), pp. 283-290.

Rigby, R.A. and Stasinopoulos, D.M. (2005), “Generalized additive models for location, scale and shape”, *Journal of the Royal Statistical Society: Series C (Applied Statistics)*, Vol. 54 (3), pp. 507-554.

Rosen, S. (1974), “Hedonic prices and implicit markets: product differentiation in pure competition”, *Journal of Political Economy*, Vol. 82 (1), pp. 34-55.

Schaefer, P. and Just, T. (2018), “Does urban tourism attractiveness affect young adult migration in Germany?”, *Journal of Property Investment and Finance*, Vol. 36 (1), pp. 68-90.

Shimizu, C., Nishimura, K.G., and Watanabe, T. (2012), “House prices from magazines, realtors, and the land registry”, in: *Property Market and Financial Stability*, BIS Papers No. 64, Bank of International Settlements, pp. 29-38.

Stein, J.C. (1995), “Prices and trading volume in the housing market: A model with down-payment effects”, *The Quarterly Journal of Economics*, Vol. 110 (2), pp. 379-406.

Stone, R. (1954), “The measurement of consumer behaviour and expenditure in the United Kingdom, 1920-1938”, *Studies in the National Income and Expenditure of the United Kingdom (assisted by DR Rowe, WJ Corlett, R. Hursteld, and M. Potter)*, Cambridge: University Press, Vol. 1.

Syed, I., Hill, R.J., and Melser, D. (2008), “Flexible spatial and temporal hedonic price indexes for housing in the presence of missing data”, discussion paper 2008/14, School of Economics, The University of New South Wales, Sydney, Australia.

Teekens, R. and Koerts, J. (1972), “Some statistical implications of the log transformation of multiplicative models”, *Econometrica: Journal of the Econometric Society*, Vol. 40 (5), pp. 793-819.

Triplett, J. (2004), “Handbook on hedonic indexes and quality adjustments in price indexes: Special application to information technology products”, OECD Science, Technology and Industry Working Papers, No. 2004/09, OECD Publishing, Paris.

Voigtlaender, M (2009), “Why is the German homeownership rate so low?”, *Housing Studies*, Vol. 24 (3), pp. 335-372.

Waugh, F.V. (1928), “Quality factors in influencing vegetable prices”, *Journal of farm economics*, Vol. 10 (2), pp. 185-196.

Wood, J.H. and Wood N.L. (1985), “Financial markets”, San Diego: Harcourt Brace Jovanovich.

2.7 Appendix

Table 2.2: Overview of classification of the 161 NUTS3 regions

#	NUTS3	State	Region	Cluster Price	Cluster Liquidity	NUTS3	State	Region	Cluster Rent	Cluster Liquidity
1	DE111	Baden-Wuerttemberg	Stuttgart city	1	1	DE114	Baden-Wuerttemberg	Goeppingen	1	1
2	DE113	Baden-Wuerttemberg	Esslingen	1	1	DE115	Baden-Wuerttemberg	Ludwigsburg	1	1
3	DE115	Baden-Wuerttemberg	Ludwigsburg	1	1	DE129	Baden-Wuerttemberg	Pforzheim city	1	1
4	DE116	Baden-Wuerttemberg	Rems-Murr-Kreis	1	1	DE12B	Baden-Wuerttemberg	Enzkreis	1	1
5	DE117	Baden-Wuerttemberg	Heilbronn city	1	1	DE143	Baden-Wuerttemberg	Zollernalbkreis	1	1
6	DE11A	Baden-Wuerttemberg	Schwaebisch Hall	1	1	DE217	Bavaria	Dachau	1	1
7	DE122	Baden-Wuerttemberg	Karlsruhe city	1	1	DE925	Lower Saxony	Hildesheim	1	1
8	DE129	Baden-Wuerttemberg	Pforzheim city	1	1	DE139	Baden-Wuerttemberg	Loerrach	1	2
9	DE12B	Baden-Wuerttemberg	Enzkreis	1	1	DE212	Bavaria	Munich city	1	2
10	DE263	Bavaria	Wuerzburg city	1	1	DE21K	Bavaria	Rosenheim county	1	2
11	DE279	Bavaria	Neu-Ulm	1	1	DE259	Bavaria	Nuremberg county	1	2
12	DE711	Hesse	Darmstadt city	1	1	DE271	Bavaria	Augsburg city	1	2
13	DE713	Hesse	Offenbach am Main city	1	1	DE279	Bavaria	Neu-Ulm	1	2
14	DE803	Mecklenburg-Western P.	Rostock city	1	1	DE300	Berlin	Berlin	1	2
15	DEA24	North Rhine-Westfalia	Leverkusen city	1	1	DE713	Hesse	Offenbach am Main city	1	2
16	DE128	Baden-Wuerttemberg	Rhein-Neckar-Kreis	1	2	DEB3I	Rhineland	Rhein-Pfalz-Kreis	1	2
17	DE139	Baden-Wuerttemberg	Loerrach	1	2	DE21M	Bavaria	Traunstein	1	3
18	DE146	Baden-Wuerttemberg	Biberach	1	2	DE263	Bavaria	Wuerzburg city	1	3
19	DE147	Baden-Wuerttemberg	Bodenseekreis	1	2	DE932	Lower Saxony	Cuxhaven	1	3

20	DE212	Bavaria	Munich city	1	2	DEA17	North Rhine-Westfalia	Oberhausen city	1	3
21	DE21K	Bavaria	Rosenheim county	1	2	DEA26	North Rhine-Westfalia	Dueren	1	3
22	DE21L	Bavaria	Starnberg	1	2	DEA2C	North Rhine-Westfalia	Rhein-Sieg-Kreis	1	3
23	DE276	Bavaria	Augsburg county	1	2	DEF02	Schleswig-Holstein	Kiel city	1	3
24	DE71A	Hesse	Main-Taunus-Kreis	1	2	DEF09	Schleswig-Holstein	Pinneberg	1	3
25	DE721	Hesse	Giessen	1	2	DEF0B	Schleswig-Holstein	Rendsburg-Eckernfoerde	1	3
26	DEA23	North Rhine-Westfalia	Cologne city	1	2	DE211	Bavaria	Ingolstadt city	1	4
27	DEF09	Schleswig-Holstein	Pinneberg	1	2	DE711	Hesse	Darmstadt city	1	4
28	DE21D	Bavaria	Garmisch-Partenkirchen	1	3	DE911	Lower Saxony	Braunschweig city	1	4
29	DE21M	Bavaria	Traunstein	1	3	DEA38	North Rhine-Westfalia	Warendorf	1	4
30	DE232	Bavaria	Regensburg	1	3	DEA46	North Rhine-Westfalia	Minden-Luebbecke	1	4
31	DE258	Bavaria	Fuerth county	1	3	DE112	Baden-Wuerttemberg	Boeblingen	2	1
32	DE300	Berlin	Berlin	1	3	DE118	Baden-Wuerttemberg	Heilbronn county	2	1
33	DE716	Hesse	Darmstadt-Dieburg	1	3	DE125	Baden-Wuerttemberg	Heidelberg city	2	1
34	DE719	Hesse	Main-Kinzig-Kreis	1	3	DE218	Bavaria	Ebersberg	2	1
35	DE925	Lower Saxony	Hildesheim	1	3	DE21D	Bavaria	Garmisch-Partenkirchen	2	1
36	DE929	Lower Saxony	Region Hanover	1	3	DEA51	North Rhine-Westfalia	Bochum city	2	1
37	DEA13	Nordrhein-Westfalen	Essen city	1	3	DE132	Baden-Wuerttemberg	Breisgau-Hochschwarzwald	2	2
38	DEA2C	North Rhine-Westfalia	Rhein-Sieg-Kreis	1	3	DE13A	Baden-Wuerttemberg	Waldshut	2	2
39	DEA37	North Rhine-Westfalia	Steinfurt	1	3	DE148	Baden-Wuerttemberg	Ravensburg	2	2
40	DEA38	North Rhine-Westfalia	Warendorf	1	3	DE21B	Bavaria	Freising	2	2
41	DEA58	North Rhine-Westfalia	Maerkischer Kreis	1	3	DE21H	Bavaria	Munich county	2	2

42	DEB34	Rhineland	Ludwigshafen am Rhein city	1	3	DE232	Bavaria	Regensburg	2	2
43	DEB3I	Rhineland	Rhein-Pfalz-Kreis	1	3	DE273	Bavaria	Kempten (Allgaeu) city	2	2
44	DEF03	Schleswig-Holstein	Luebeck city	1	3	DE27E	Bavaria	Oberallgaeu	2	2
45	DEF08	Schleswig-Holstein	Ostholstein	1	3	DE712	Hesse	Frankfurt am Main city	2	2
46	DE145	Baden-Wuerttemberg	Alb-Donau-Kreis	1	4	DE717	Hesse	Gross-Gerau	2	2
47	DE947	Lower Saxony	Aurich	1	4	DE718	Hesse	Hochtaunuskreis	2	2
48	DE949	Lower Saxony	Emsland	1	4	DE719	Hesse	Main-Kinzig-Kreis	2	2
49	DE114	Baden-Wuerttemberg	Goeppingen	2	1	DE721	Hesse	Giessen	2	2
50	DE118	Baden-Wuerttemberg	Heilbronn county	2	1	DEA1A	North Rhine-Westfalia	Wuppertal city	2	2
51	DE141	Baden-Wuerttemberg	Reutlingen	2	1	DEA27	North Rhine-Westfalia	Rhein-Erft-Kreis	2	2
52	DE148	Baden-Wuerttemberg	Ravensburg	2	1	DEA2A	North Rhine-Westfalia	Oberbergischer Kreis	2	2
53	DEG01	Thuringia	Erfurt city	2	1	DEA52	North Rhine-Westfalia	Dortmund city	2	2
54	DE12A	Baden-Wuerttemberg	Calw	2	2	DEF03	Schleswig-Holstein	Luebeck city	2	2
55	DE211	Bavaria	Ingolstadt city	2	2	DEF08	Schleswig-Holstein	Ostholstein	2	2
56	DE217	Bavaria	Dachau	2	2	DE122	Baden-Wuerttemberg	Karlsruhe city	2	3
57	DE21C	Bavaria	Fuerstenfeldbruck	2	2	DE146	Baden-Wuerttemberg	Biberach	2	3
58	DE252	Bavaria	Erlangen city	2	2	DE252	Bavaria	Erlangen city	2	3
59	DE271	Bavaria	Augsburg city	2	2	DE254	Bavaria	Nuremberg city	2	3
60	DE273	Bavaria	Kempten (Allgaeu) city	2	2	DE933	Lower Saxony	Harburg	2	3
61	DE27E	Bavaria	Oberallgaeu	2	2	DE935	Lower Saxony	Lueneburg	2	3
62	DEB3J	Rhineland	Mainz-Bingen	2	2	DE944	Lower Saxony	Osnabrueck city	2	3
63	DEF0D	Schleswig-Holstein	Segenberg	2	2	DEA15	North Rhine-Westfalia	Moenchenglad-bach city	2	3

64	DE11D	Baden-Wuerttemberg	Main-Tauber-Kreis	2	3	DEA1F	North Rhine-Westfalia	Wesel	2	3
65	DE123	Baden-Wuerttemberg	Karlsruhe county	2	3	DEA2D	North Rhine-Westfalia	Staedteregion Aachen	2	3
66	DE126	Baden-Wuerttemberg	Mannheim city	2	3	DEA45	North Rhine-Westfalia	Lippe	2	3
67	DE218	Bavaria	Ebersberg	2	3	DE600	Hamburg	Hamburg	2	4
68	DE600	Hamburg	Hamburg	2	3	DE916	Lower Saxony	Goslar	2	4
69	DE712	Hesse	Frankfurt am Main city	2	3	DED21	Saxony	Dresden city	2	4
70	DE717	Hesse	Gross-Gerau	2	3	DED41	Saxony	Chemnitz city	2	4
71	DE932	Lower Saxony	Cuxhaven	2	3	DED51	Saxony	Leipzig city	2	4
72	DE94E	Lower Saxony	Osnabrueck county	2	3	DED52	Saxony	Leipzig county	2	4
73	DEA14	North Rhine-Westfalia	Krefeld city	2	3	DEE03	Saxony-Anhalt	Magdeburg city	2	4
74	DEA2A	North Rhine-Westfalia	Oberbergischer Kreis	2	3	DEG01	Thuringia	Erfurt city	2	4
75	DEA34	North Rhine-Westfalia	Borken	2	3	DE111	Baden-Wuerttemberg	Stuttgart city	3	1
76	DEA42	North Rhine-Westfalia	Guetersloh	2	3	DE113	Baden-Wuerttemberg	Esslingen	3	1
77	DEA45	North Rhine-Westfalia	Lippe	2	3	DE116	Baden-Wuerttemberg	Rems-Murr-Kreis	3	1
78	DEA46	North Rhine-Westfalia	Minden-Luebbecke	2	3	DE117	Baden-Wuerttemberg	Heilbronn city	3	1
79	DED21	Saxony	Dresden city	2	3	DE124	Baden-Wuerttemberg	Rastatt	3	1
80	DEF02	Schleswig-Holstein	Kiel city	2	3	DE128	Baden-Wuerttemberg	Rhein-Neckar-Kreis	3	1
81	DEF0B	Schleswig-Holstein	Rendsburg-Eckernfoerde	2	3	DE141	Baden-Wuerttemberg	Reutlingen	3	1
82	DE80L	Mecklenburg-Western P.	Vorpommern-Ruegen	2	4	DE21C	Bavaria	Fuerstenfeldbruck	3	1
83	DEA16	North Rhine-Westfalia	Muelheim a.d. Ruhr city	2	4	DEA55	North Rhine-Westfalia	Herne city	3	1
84	DEA1B	North Rhine-Westfalia	Kleve	2	4	DE11A	Baden-Wuerttemberg	Schwaebisch Hall	3	2
85	DEA1F	North Rhine-Westfalia	Wesel	2	4	DE123	Baden-Wuerttemberg	Karlsruhe county	3	2

86	DEA32	North Rhine-Westfalia	Gelsenkirchen city	2	4	DE131	Baden-Wuerttemberg	Freiburg city	3	2
87	DEA5C	North Rhine-Westfalia	Unna	2	4	DE134	Baden-Wuerttemberg	Ortenaukreis	3	2
88	DEC04	Saarland	Saarlouis	2	4	DE138	Baden-Wuerttemberg	Konstanz	3	2
89	DE112	Baden-Wuerttemberg	Boeblingen	3	1	DE142	Baden-Wuerttemberg	Tuebingen	3	2
90	DE253	Bavaria	Fuerth city	3	1	DE144	Baden-Wuerttemberg	Ulm city	3	2
91	DE718	Hesse	Hochtaunuskreis	3	1	DE145	Baden-Wuerttemberg	Alb-Donau-Kreis	3	2
92	DE124	Baden-Wuerttemberg	Rastatt	3	2	DE253	Bavaria	Fuerth city	3	2
93	DE138	Baden-Wuerttemberg	Konstanz	3	2	DE276	Bavaria	Augsburg county	3	2
94	DE21B	Bavaria	Freising	3	2	DE716	Hesse	Darmstadt-Dieburg	3	2
95	DE21H	Bavaria	Munich county	3	2	DE71C	Hesse	Offenbach county	3	2
96	DE254	Bavaria	Nuremberg city	3	2	DE929	Lower Saxony	Region Hanover	3	2
97	DE501	Bremen	Bremen city	3	2	DEA23	North Rhine-Westfalia	Cologne city	3	2
98	DE71C	Hesse	Offenbach county	3	2	DE126	Baden-Wuerttemberg	Mannheim city	3	3
99	DE943	Lower Saxony	Oldenburg city	3	2	DE501	Bremen	Bremen city	3	3
100	DEA27	North Rhine-Westfalia	Rhein-Erft-Kreis	3	2	DE714	Hesse	Wiesbaden city	3	3
101	DED51	Saxony	Leipzig city	3	2	DE803	Mecklenburg-Western P.	Rostock city	3	3
102	DEE03	Saxony-Anhalt	Magdeburg city	3	2	DEA11	North Rhine-Westfalia	Dusseldorf city	3	3
103	DE136	Baden-Wuerttemberg	Schwarzwald-Baar-Kreis	3	3	DEA13	North Rhine-Westfalia	Essen city	3	3
104	DE259	Bavaria	Nuremberg county	3	3	DEA16	North Rhine-Westfalia	Muelheim a.d. Ruhr city	3	3
105	DE714	Hesse	Wiesbaden city	3	3	DEA1C	North Rhine-Westfalia	Mettmann	3	3
106	DE933	Lower Saxony	Harburg	3	3	DEA22	North Rhine-Westfalia	Bonn city	3	3
107	DE935	Lower Saxony	Lueneburg	3	3	DEA24	North Rhine-Westfalia	Leverkusen city	3	3

108	DE944	Lower Saxony	Osnabrueck city	3	3	DEA36	North Rhine-Westfalia	Recklinghausen	3	3
109	DEA1A	North Rhine-Westfalia	Wuppertal city	3	3	DEA42	North Rhine-Westfalia	Guetersloh	3	3
110	DEA47	North Rhine-Westfalia	Paderborn	3	3	DEA56	North Rhine-Westfalia	Ennepe-Ruhr-Kreis	3	3
111	DEA51	North Rhine-Westfalia	Bochum city	3	3	DEA5C	North Rhine-Westfalia	Unna	3	3
112	DEB35	Rhineland	Mainz city	3	3	DEB34	Rhineland	Ludwigshafen (Rhine) city	3	3
113	DEA15	North Rhine-Westfalia	Moenchenglad-bach city	3	4	DEA12	North Rhine-Westfalia	Duisburg city	3	4
114	DEB25	Rhineland	Trier-Saarburg	3	4	DEA33	North Rhine-Westfalia	Muenster city	3	4
115	DED41	Saxony	Chemnitz city	3	4	DEA5B	North Rhine-Westfalia	Soest	3	4
116	DE131	Baden-Wuerttemberg	Freiburg city	4	1	DEC01	Saarland	Regionalverband Saarbruecken	3	4
117	DE125	Baden-Wuerttemberg	Heidelberg city	4	2	DED2E	Saxony	Meissen	3	4
118	DE142	Baden-Wuerttemberg	Tuebingen	4	2	DED45	Saxony	Zwickau	3	4
119	DE144	Baden-Wuerttemberg	Ulm city	4	2	DE12A	Baden-Wuerttemberg	Calw	4	1
120	DE80K	Mecklenburg-Western P.	Rostock county	4	2	DE136	Baden-Wuerttemberg	Schwarzwald-Baar-Kreis	4	1
121	DE911	Lower Saxony	Braunschweig city	4	2	DE11D	Baden-Wuerttemberg	Main-Tauber-Kreis	4	2
122	DE939	Lower Saxony	Stade	4	2	DE21L	Bavaria	Starnberg	4	2
123	DEA41	North Rhine-Westfalia	Bielefeld city	4	2	DE71A	Hesse	Main-Taunus-Kreis	4	2
124	DE132	Baden-Wuerttemberg	Breisgau-Hochschwarzwald	4	3	DE94E	Lower Saxony	Osnabrueck county	4	2
125	DE134	Baden-Wuerttemberg	Ortenaukreis	4	3	DEA1D	North Rhine-Westfalia	Rhein-Kreis Neuss	4	2
126	DE13A	Baden-Wuerttemberg	Waldshut	4	3	DEA37	North Rhine-Westfalia	Steinfurt	4	2
127	DE143	Baden-Wuerttemberg	Zollernalbkreis	4	3	DEF0F	Schleswig-Holstein	Stormann	4	2
128	DE715	Hesse	Bergstrasse	4	3	DE121	Baden-Wuerttemberg	Baden-Baden city	4	3
129	DE71D	Hesse	Rheingau-Taunus-Kreis	4	3	DE258	Bavaria	Fuerth county	4	3

130	DE71E	Hesse	Wetteraukreis	4	3	DE404	Brandenburg	Potsdam city	4	3
131	DE731	Hesse	Kassel city	4	3	DE715	Hesse	Bergstrasse	4	3
132	DEA11	North Rhine-Westfalia	Duesseldorf city	4	3	DE71D	Hesse	Rheingau-Taunus-Kreis	4	3
133	DEA19	North Rhine-Westfalia	Solingen city	4	3	DE71E	Hesse	Wetteraukreis	4	3
134	DEA1C	North Rhine-Westfalia	Mettmann	4	3	DE731	Hesse	Kassel city	4	3
135	DEA1D	North Rhine-Westfalia	Rhein-Kreis Neuss	4	3	DE943	Lower Saxony	Oldenburg city	4	3
136	DEA22	North Rhine-Westfalia	Bonn city	4	3	DEA14	North Rhine-Westfalia	Krefeld city	4	3
137	DEA26	North Rhine-Westfalia	Dueren	4	3	DEA19	North Rhine-Westfalia	Solingen city	4	3
138	DEA2B	North Rhine-Westfalia	Rheinisch-Bergischer Kreis	4	3	DEA1E	North Rhine-Westfalia	Viersen	4	3
139	DEA2D	North Rhine-Westfalia	Staedteregion Aachen	4	3	DEA2B	North Rhine-Westfalia	Rheinisch-Bergischer Kreis	4	3
140	DEA33	North Rhine-Westfalia	Muenster city	4	3	DEA32	North Rhine-Westfalia	Gelsenkirchen city	4	3
141	DEA36	North Rhine-Westfalia	Recklinghausen	4	3	DEA34	North Rhine-Westfalia	Borken	4	3
142	DEA52	North Rhine-Westfalia	Dortmund city	4	3	DEA41	North Rhine-Westfalia	Bielefeld city	4	3
143	DEA56	North Rhine-Westfalia	Ennepe-Ruhr-Kreis	4	3	DEA43	North Rhine-Westfalia	Herford	4	3
144	DEC01	Saarland	Regionalverband Saarbruecken	4	3	DEA47	North Rhine-Westfalia	Paderborn	4	3
145	DED2E	Saxony	Meissen	4	3	DEA57	North Rhine-Westfalia	Hochsauerland-kreis	4	3
146	DED52	Saxony	Leipzig county	4	3	DEB21	Rhineland	Trier city	4	3
147	DE121	Baden-Wuerttemberg	Baden-Baden city	4	4	DEB25	Rhineland	Trier-Saarburg	4	3
148	DE404	Brandenburg	Potsdam city	4	4	DEB35	Rhineland	Mainz city	4	3
149	DE80N	Mecklenburg-Western P.	Vorpommern-Greifswald	4	4	DEB3J	Rhineland	Mainz-Bingen	4	3
150	DE916	Lower Saxony	Goslar	4	4	DEF0D	Schleswig-Holstein	Segenberg	4	3
151	DEA12	North Rhine-Westfalia	Duisburg city	4	4	DE147	Baden-Wuerttemberg	Bodenseekreis	4	4

152	DEA17	North Rhine-Westfalia	Oberhausen city	4	4	DE80K	Mecklenburg-Western P.	Rostock county	4	4
153	DEA1E	North Rhine-Westfalia	Viersen	4	4	DE80L	Mecklenburg-Western P.	Vorpommern-Ruegen	4	4
154	DEA43	North Rhine-Westfalia	Herford	4	4	DE80N	Mecklenburg-Western P.	Vorpommern-Greifswald	4	4
155	DEA55	North Rhine-Westfalia	Herne city	4	4	DE939	Lower Saxony	Stade	4	4
156	DEA57	North Rhine-Westfalia	Hochsauerland-kreis	4	4	DE947	Lower Saxony	Aurich	4	4
157	DEA5B	North Rhine-Westfalia	Soest	4	4	DE949	Lower Saxony	Emsland	4	4
158	DEB21	Rhineland	Trier city	4	4	DEA1B	North Rhine-Westfalia	Kleve	4	4
159	DED45	Saxony	Zwickau	4	4	DEA58	North Rhine-Westfalia	Maerkischer Kreis	4	4
160	DEF07	Schleswig-Holstein	Nordfriesland	4	4	DEC04	Saarland	Saarlouis	4	4
161	DEF0F	Schleswig-Holstein	Stormann	4	4	DEF07	Schleswig-Holstein	Nordfriesland	4	4

3. Exploring the Determinants of Liquidity with Big Data – Market Heterogeneity in German Markets

Abstract

Purpose – The purpose of this paper is to examine the market liquidity (time on market) and its determinants, for rental dwellings in the largest seven German cities, with big data.

Design/methodology/approach – The determinants of time on market are estimated with the Cox Proportional Hazards Model. Hedonic characteristics, as well as socioeconomic and spatial variables, are combined with different fixed effects and controls for non-linearity, so as to maximise the explanatory power of the model.

Findings – Higher asking rent and larger living space decrease the liquidity in all seven markets, while the age of a dwelling, the number of rooms and proximity to the city centre accelerate the letting process. For the other hedonic characteristics heterogeneous implications emerge.

Practical implications – The findings are of interest for institutional and private landlords, as well as governmental organizations in charge of housing and urban development.

Originality/value – This is the first paper to deal with the liquidity of rental dwellings in the seven most populated cities of Europe's second largest rental market, by applying the Cox Proportional Hazards Model with spatial gravity variables. Furthermore, the German rental market is of particular interest, as approximately 60% of all rental dwellings are owned by private landlords and the German market is organized polycentrically.

Keywords: Liquidity/ Time on market; Housing real estate; Big data; Cox Proportional Hazards model; Non-linearity

Acknowledgement: The authors especially thank PATRIZIA Immobilien AG for contributing the dataset for this study. All statements of opinion are those of the authors and do not necessarily reflect the opinions of PATRIZIA Immobilien AG or its associated companies.

3.1 Introduction

Financial assets such as stocks and bonds are traded in tremendous volumes, turning over billions of dollars within seconds and with no spatial constraints. By contrast, the transaction process of direct real estate is more complex, often consuming several months. When it comes to residential real estate, a match may be even more difficult to achieve, as this is strongly determined by the individual preferences of homebuyers and the expectations of homesellers. A general understanding of liquidity in direct real estate is therefore essential for market players, whether private, institutional or governmental, not only in order to derive investment strategies, but also to assess market fundamentals and cyclical movements, as well as political implications. Moreover, the instruments needed to efficiently capture the factors both boosting and constraining liquidity, are crucial and far from trivial, as liquidity in terms of "time" requires advanced econometric modelling. To capture the uncertainty of finding a match, as well as the time a property is advertised on the market, liquidity in the residential real estate literature is widely proxied by time on market (TOM). In this context, this paper explores the liquidity of direct real estate, focussing on the seven largest German rental housing markets by means of advanced semiparametric survival techniques. The aim of the study is to explore liquidity concepts and examine the factors that determine liquidity, such as linear, binary, spatial as well as possible non-linear effects with big data, in order to derive both similarities and divergences between the cities. The paper may serve as a guide for market players and policy makers conducting liquidity analysis on and understanding future developments in rental housing markets in Germany. Especially for first-time buyers, an overview on the largest seven real estate markets and an indication of the factors affecting the letting process is of considerable importance, as during the marketing time, redemption and interest have to be borne by other sources of income.

The following brief literature review only covers the articles directly relevant for this study, and thus only a small fraction of the literature on time on market. Since their establishment within the real estate literature, survival models have been adopted by various researchers to estimate the determinants of time on market. Kluger and Miller (1990) introduced the semi-parametric Cox Proportional Hazards Model based on Cox (1972) to real estate studies, which allows a particularly flexible application, without any a priori assumptions regarding the distribution of the baseline hazard, in contrast to the widely used Weibull model. Studies using this approach include Krainer (1999), Smith (2010), Hoeberichts et al. (2013), Cirman et al. (2015), among others.

In searching for an instrument capturing "user taste" for dwellings and its effect on liquidity, Haurin (1988) developed an atypicality index and shows that for more atypical dwellings, the distribution of offers is prone to wider variation. A dwelling is more atypical when its hedonic properties deviate substantially from the mean hedonic market characteristics, e.g. a dwelling with 150 m², 1 room, located on the 10th floor without an elevator. Nowadays, atypicality is a widely recognised factor in hedonic survival regressions, as seen in Krainer (1999), Anglin et al. (2003), Bourassa et al. (2009), Haurin et al. (2010; 2013) and Hoeberichts et al. (2013), among others.

The signalling effect of setting the initial list price is also a widely researched area. Glower et al. (1998) for example, began to investigate the impact of the percentage difference in the observed list price from the expected list price. Anglin et al. (2003) extended this approach and introduced a new explanatory variable in the context of liquidity, called the degree of overpricing (DOP). They defined the variable as the percentage deviation of an individual property's list price from the empirically estimated market list price. While they found that abnormal list prices, i.e. overpricing, increase the marketing time of houses, further applications can also be found in Hoeberichts et al. (2013) and Cirman et al. (2015).

Over the last years, more and more emphasis has been placed on spatial effects when modelling the price and time on market of residential properties. Many articles have tested the theory of market segmentation in residential real estate markets, concluding that the inclusion of spatial variables improves the explanatory power of real estate pricing models e.g. Goodman and Thibodeau (2007), Turnbull and Dombrow (2006), Pavlov (2000), Fik et al. (2003), Bourassa et al. (2010) and Cirman et al. (2015) among others. Smith (2010) was the first to specify a Cox-model containing school districts and Cartesian coordinates. He found that, while the school district dummies and the coordinates are by themselves statistically significant and demonstrate a large impact on the liquidity, the combination of both, yields the largest explanatory power.

To the best of the authors' knowledge, the first study to estimate time on market in residential rental markets was conducted by Allen et al. (2009). The authors focus on the Dallas-Fort Worth area with a sample of over 20,000 listings and more than 11,000 corresponding letting contracts. Using a Weibull hazard model, the authors conclude that after resetting asking rent initially overpriced by 15%, the landlords face 9.5 days longer time on market on average. Due to the initial overpricing and thus longer TOM, these landlords also have to accept a contract rent which is on average 5.2% below the hedonically estimated level. Cajias et al. (2016), in contrast, used a similar approach to estimate the effect of energy consumption on time on market for the German rental market. Using a Cox model, the authors calculated the odds of a dwelling being let, dependent on energy consumption, and show that dwellings with higher energy consumption relative to the most energy efficient dwellings, stayed on the market for longer.

This paper investigates the determinants of liquidity on one of the largest rental markets in Europe. It should be noted, that the German housing market has the second lowest ownership rate among all European countries (45%) after Switzerland (43%) and compared to other industrialized countries. In Germany, home-ownership is at such a low level, because of a large stock of high-quality subsidized social housing built after World War II, low tax benefits for owners and a rather liberal rental market (Voigtlaender, 2009). A profound understanding of liquidity is therefore not only relevant for institutional landlords, but for millions of private providers of living space, as of those roughly 14.5 million rental properties, approximately 60% are owned and let by individuals. At the same time, Germany provides a unique research field, because of extraordinarily low homeownership rates in its largest seven real estate markets, ranging from about 15% in Berlin to ca. 32% in Stuttgart and because of the polycentric market organization. In comparison to other European countries such as England or France, the German market is not dominated by a single megacity, but consists of seven major real estate markets. Each of those seven cities has developed its own field of specialization, for example Frankfurt as the financial capital, Stuttgart as an automobile city and Munich a hybrid between new technology and Bavarian tradition. Therefore, the dataset consisting of Germany's largest seven cities, represents a socially, culturally and economically well diversified overview of major urban areas all over Germany and the study is thus able to yield comprehensive results explaining liquidity for the German rental market in an urban context.

Beyond exploring the hedonic characteristics, the authors wish to contribute to the existing literature by enhancing the modelling quality and introducing spatial gravity variables to time on market modelling, using the unique German rental market as an example.

The paper is organized as follows. The second section describes the methodology employed for the study, whereas the third section describes the data. The fourth section presents the results.

3.2 Econometric Approach

Prior to deriving the model, some statistical elements in the estimation of survival regression need to be defined. The time period (T) for which a flat is offered on the market corresponds to a continuous positive response variable without zeros and is interpreted as the duration of an event (t), in this case, the time in weeks before the signing of the letting agreement. Two main measures are important for understanding and estimating survival models: the survival function (S) and the hazard rate function (h). Formally, they are expressed as:

$$S(t) = P(T > t) = 1 - \int_t^{\infty} f(x)dx \tag{10}$$

$$h(t) = \frac{P(t < T \leq t + \Delta t | T > t)}{\Delta t} \tag{11}$$

While the survival function yields the probability that a dwelling "survives" until a certain time t, the hazard specifies the rate of failure at $T = t$ given that the flat survived up to time t. Since the numerator in equation (11) corresponds to a conditional probability and the denominator is time elapsed, the hazard function gives the probability or rate of "mortality" per units of time. Very important in survival analysis is the fact that some observations or dwellings do not change their event status, either because they remain available on the market or because the landlord does not change the status in the Multiple Listing Service (MLS) database. In this case, the response variable is said to be right-censored. Cox hazard models do account for censoring in the response variable, as they transform the response into a count variable per unit of time.

The Cox hazard model explains the factors that boost or restrict the letting process of a dwelling, as a probability function after controlling for dwelling- and market-specific characteristics. More specifically, the multivariate Cox hazard model expresses the elasticities as “odds”, e.g. a coefficient of 1.2 means a 1.2 times quicker “death” as the reference or baseline. In a first step, the equation is parametrized as a semiparametric proportional hazard model:

$$h(t_{ijp}) = \exp(\mathbf{X}_{ip}\beta + \mathbf{Z}_{j}\alpha + \boldsymbol{\mu}_{ip}\delta_{p} + \boldsymbol{\mu}_{ij}\rho_{j}) + e_{ijp} \; \forall \; m; \; m \in 1, \dots, 7, \tag{12}$$

where h corresponds to the hazard function of time on market $t$in market m, the $\mathbf{X}$ matrix contains the specific characteristics of dwelling i at observation period p, $\mathbf{Z}$ includes socioeconomic data on ZIP-area j and $\boldsymbol{\mu}_{ip}$ and $\boldsymbol{\mu}_{ij}$ account for p time- and j spatial effects respectively. The results of equation (12) are expected to provide information on the covariates boosting or limiting the marketing time of dwellings in the observed housing markets.

A second step captures the response of liquidity in space. While the covariates in $\mathbf{X}$ and $\mathbf{Z}$ are either continuous or binary, the p time- and j spatial effects in the matrices $\boldsymbol{\mu}_{ip}$ and $\boldsymbol{\mu}_{ij}$ are defined as follows:

$$\mu_{ip} = \{1 \; \Leftrightarrow \text{i in p}; 0 \Leftrightarrow \; \text{else}\} \tag{13}$$

$$\mu_{ij} = \{1 \; \Leftrightarrow \text{i in j}; 0 \Leftrightarrow \; \text{else}\} \tag{14}$$

For each m, the vector of $\hat{\rho}_j$ coefficients captures the ZIP-specific relative changes in liquidity over the entire observation period with respect to a certain reference category. The reference category in each market is defined as the ZIP-area with the highest asking rent R adjusted for sample size. Afterwards, the results of the $\hat{\rho}_j$ coefficients are presented in maps, so as to explore liquidity graphically within a spatial context.

In a third step, equation (12) is expanded by non-linear effects. This improves the estimates of two continuous hedonic covariates: dwelling rent and age. This is accomplished by applying a non-parametric smoothing estimator, which corresponds to a penalized approach comprising k knots. Simply expressed, for $k = 2$, the smoothing estimator minimizes the sum of squares of a “line” with one turning point or local minima, similar to quadratic terms. The knots are chosen iteratively by minimizing the sum of squares at different values of k (Heckman and Ramsay, 2000). The expanded Cox hazard model is as follows:

$$h(t_{ijp}) = \exp(\mathbf{X}_{ip}\beta + \mathbf{Z}_{j}\alpha + \boldsymbol{\mu}_{ip}\delta_{p} + \boldsymbol{\mu}_{ij}\rho_{j} + f(x_{ip}^{a}) + f(x_{ip}^{b})) + e_{ijp} \; \forall \; m; \; m \in 1, \dots, 7, \tag{15}$$

where $f(x_{ip}^{a})$ and $f(x_{ip}^{b})$ correspond to the smoothing function of dwelling rent and age respectively. The coefficients are interpreted graphically, for each covariate in each market.

3.3 Data and descriptive Statistics

The estimation sample comprises two merged databases. On the one hand, 335,972 observations of rental flats are gathered from multiple listing services (MLS) in Germany from 2013-Q1 until 2016-Q3, as collected by the Empirica Systems database, which contains the most important multiple listing service (MLS) providers. On the other hand, two socioeconomic variables, purchasing power per household and the number of households at the ZIP-code level, are extracted from the GfK-database. Since the data is georeferenced, two spatial gravity indicators measuring the Euclidian distance of each dwelling to the geographical centroid of the ZIP and NUTS3 polygon in kilometres, are calculated. Both variables are assumed to control for spatial distribution of dwellings within an urban area. NUTS3 regions correspond to the "Nomenclature of territorial units for statistics", which is a hierarchical system for dividing up the economic territory in Europe. While the NUTS1 consists of major socio-economic regions, the NUTS3 regions cover small regions similar to counties or administrative districts. Finally, the relevant variables in the context of hedonic survival regressions, dwelling atypicality and the degree of overpricing, are derived.

In table 3.1, Munich displays the highest density of ZIP-code areas with one ZIP-code for each 4.14 km². Although the density of postal areas for the Dusseldorf subsample is higher than for Berlin, dwellings are on average located closer to the ZIP-code area centroid. The highest construction activity seems to have taken place in Munich, as the dwellings are on average only 36.16 years old and 20.9% are listed for first occupancy. A very low degree of atypicality – also the lowest standard deviation in atypicality – together with the highest ratios for the amenities parking space and elevator are probably signs of a large share of professional housing construction, meeting the demand of the wealthiest households among the sample. Households in Munich have on average 57% more purchasing power than households in Berlin, but pay 73.1% more rent. Stuttgart and Munich have the most liquid markets for residential leasehold property, as in both cities, a dwelling is advertised for about 4.3 weeks on average, while in Stuttgart, the duration is slightly shorter and displays a lower standard deviation.

Table 3.1: Descriptive statistics

Variable Mean / (St. Deviation)	Berlin	Frankfurt	Munich	Stuttgart	Cologne	Dusseldorf	Hamburg
Asking rent €/m²/p.m.	8.567 (2.481)	12.226 (2.873)	14.831 (3.252)	11.06 (2.505)	9.788 (2.147)	9.599 (2.125)	10.567 (2.754)
Time on market weeks	6.71 (10.899)	6.564 (10.292)	4.331 (8.751)	4.316 (7.678)	5.996 (10.368)	7.717 (11.888)	5.734 (9.721)
Area m²	72.192 (27.939)	76.179 (31.489)	74.774 (31.475)	78.473 (31.244)	71.684 (27.746)	74.453 (29.78)	70.545 (26.988)
Age	56.139 (35.432)	40.549 (33.277)	36.159 (27.586)	44.776 (31.448)	41.803 (26.501)	49.55 (27.811)	46.308 (30.787)
Euclidean distance to ZIP centroid in Km.	0.687 (0.416)	0.748 (0.492)	0.601 (0.374)	0.717 (0.469)	0.915 (0.579)	0.652 (0.483)	0.808 (0.577)

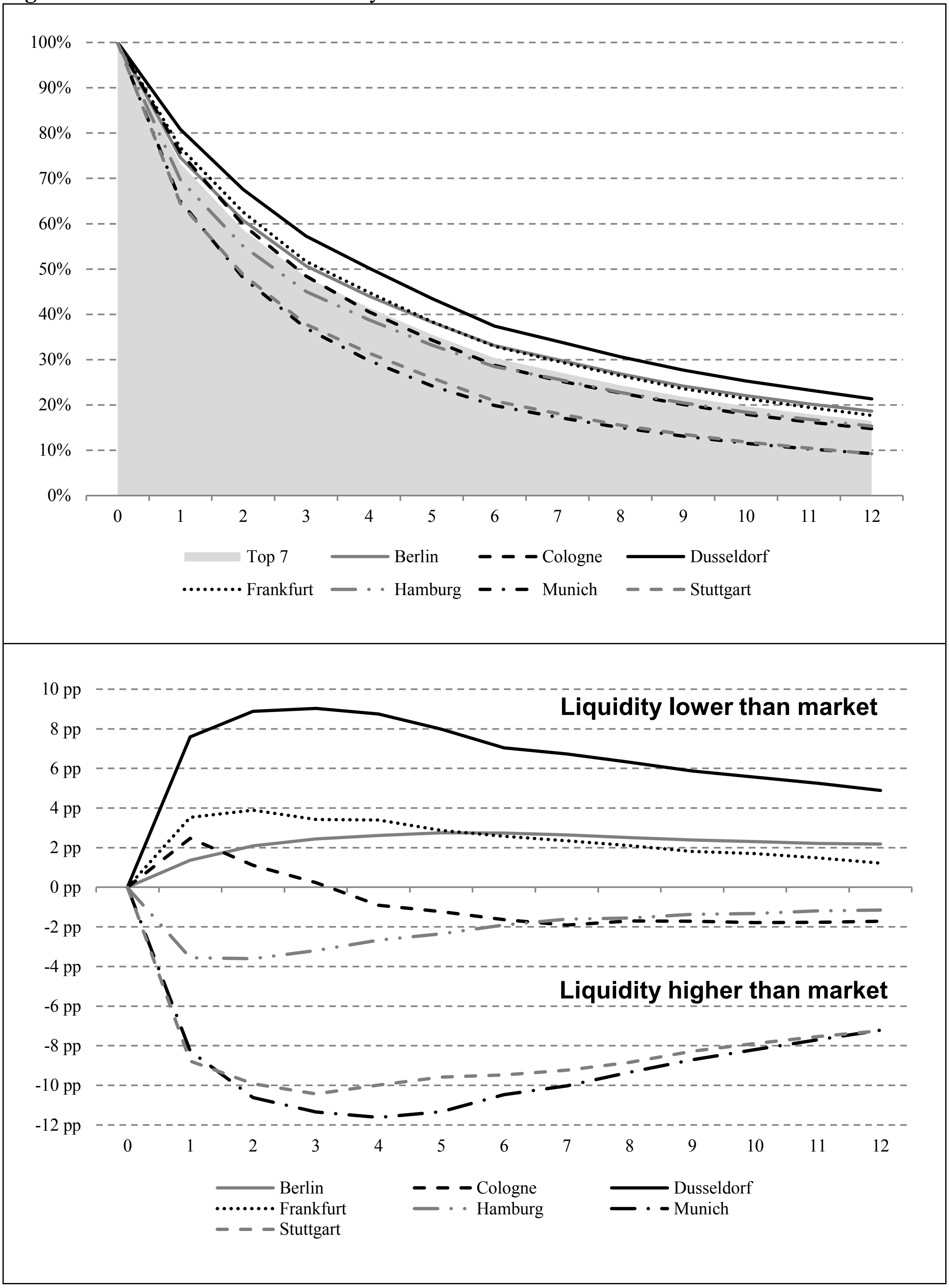
Euclidean distance to NUTS3 centroid in Km.	8.466 (4.035)	3.677 (1.942)	4.733 (2.069)	4.437 (2.33)	5.462 (2.841)	3.958 (2.609)	7.974 (3.772)
Degree of atypicality	3.906 (3.251)	0.219 (0.165)	0.106 (0.1)	1.063 (0.375)	0.375 (0.106)	-0.101 (0.153)	0.12 (0.285)
Degree of overpricing	0 (1)	0 (1)	0 (1)	0 (1)	0 (1)	0 (1)	0 (1)
Number of households in ZIP	11,856 (3,565)	10,963 (4,201)	11,402 (3,088)	10,271 (3,157)	13,066 (3,393)	9,599 (2,946)	10,960 (3,360)
Purchasing power per household in €	34,791 (4,924)	46,941 (6,320)	54,618 (6,072)	46,901 (4,428)	45,589 (5,776)	47,658 (5,881)	42,804 (7,644)
With bathtub	0.607 (0.488)	0.579 (0.494)	0.538 (0.499)	0.49 (0.5)	0.482 (0.5)	0.531 (0.499)	0.627 (0.484)
With built-in-kitchen	0.474 (0.499)	0.662 (0.473)	0.667 (0.471)	0.685 (0.464)	0.297 (0.457)	0.359 (0.48)	0.767 (0.423)
With parking space	0.258 (0.438)	0.524 (0.499)	0.65 (0.477)	0.575 (0.494)	0.424 (0.494)	0.394 (0.489)	0.408 (0.491)
With terrace	0.156 (0.363)	0.152 (0.359)	0.19 (0.392)	0.21 (0.408)	0.171 (0.376)	0.163 (0.369)	0.165 (0.371)
With balcony	0.668 (0.471)	0.642 (0.479)	0.638 (0.481)	0.616 (0.486)	0.627 (0.484)	0.637 (0.481)	0.659 (0.474)
With elevator	0.425 (0.494)	0.458 (0.498)	0.55 (0.497)	0.314 (0.464)	0.355 (0.479)	0.36 (0.48)	0.286 (0.452)
First occupancy	0.179 (0.384)	0.24 (0.427)	0.209 (0.406)	0.188 (0.391)	0.152 (0.359)	0.168 (0.374)	0.164 (0.37)
N	119,481	28,641	32,216	12,755	36,940	35,814	70,125

Notes: The sample contains 335,972 observations of dwellings advertised on multiple listing services (MLS). The sample covers 3.75 years from Q1 2013 until Q3 2016. While the asking rent is expressed in €/m²/month, area is expressed as m² and age as a number of years, the means of the characteristics bathtub, built-in kitchen etc. can be interpreted as ratios. The purchasing power per household and the number of households per ZIP-code area are extracted from the market research database of GfK. Spatial gravity indicators measure the Euclidian distance of each dwelling to the geographical centroid of the ZIP and NUTS3 polygon in kilometers. The degree of atypicality is calculated according to the definition by (Haurin, 1988), while the Degree of overpricing is constructed according to (Anglin et al., 2003).

In order to conceptualize the main idea of survival methods and liquidity per se, a survival regression with city dummies is presented. Figure 3.1 shows the survival function proceedings from regressing the time on market in weeks on city dummies over the entire sample period. The survival functions depict the mortality rate of an average dwelling as a function of time. When looking at the top 7 markets combined, the results show that the probability of letting a dwelling after four weeks is roughly 60%. Whereas the probability of letting an average dwelling in Munich and Stuttgart after four weeks is about 70%, the probability of finding a new tenant in Dusseldorf within the first month is only 50%. The lower panel shows that the sharpest increase in the probability of letting an apartment in Munich, Stuttgart and Hamburg, compared to the top 7 markets, takes place during the first three weeks, revealing the huge demand pressure which results in above-average liquidity for rental units in those cities. The inverse pattern appears for

Dusseldorf and Frankfurt. Landlords in Cologne face lower than average liquidity within the first month, before the probability of letting a dwelling increases above the market average. As the survival functions evidently show distinctions between the cities, this present paper aims at exploring the factors that boost or dampen the survival function in a multivariate approach.

Figure 3.1: Mean survival function by market



Notes: The upper panel shows the survival function from a Cox regression of dwellings' time on market in weeks on seven dummies (stratas) and the entire sample. The survival functions illustrate the mortality rate of an average dwelling as a function of time. The lower exhibit presents the survival functions relative the overall market survival.

3.4 Results

3.4.1 Main Liquidity Drivers

Table 3.2 presents the results of the three parameterizations from equation (3) for each city. The coefficients of the Cox Proportional Hazards Model are displayed, together with their respective standard deviations, whereas positive coefficients increase the hazard (shorten the survival time) and therefore increase liquidity (dwelling letting process). Since hazard models estimate event probabilities per unit of time, a coefficient of determination precisely as in the OLS, is difficult to obtain. As a substitute, the Pseudo-R^2 based on Kendall's Tau measures the concordance between estimated survival time and the observed survival time for only the non-censored response sample. Values between 80% and 60% are common in survival studies. Model I includes only hedonic covariates, whereas model II includes the atypicality and overpricing indices, as well as gravity variables and Cartesian coordinates. Model III presents the full model including all control variables.

The full models show that liquidity responds negatively to rent and size. Thus, a dwelling's letting process is longer, the higher the asking rent and the larger the dwelling, whereas the rent effect is insignificant in Hamburg. The age factor shortens the letting process in Berlin and Frankfurt, as the coefficients are significant and positive. In contrast, the design of the flats in terms of the number of rooms shows the expected effect, as the higher the dwellings' usability, the shorter the average letting process.

When focussing on the hedonic dummy variables, the coefficient interpretation is more tangible. Dwellings with a bathtub, a parking space, a balcony and an elevator are in general difficult to let, as the coefficients are in most cases negative. In Hamburg, the city with the highest ratio of flats with bathtubs, the feature seems to be accepted as standard, because it is the only city revealing no significant impact. A similar effect is observed for built-in kitchens in Stuttgart. In Dusseldorf and Cologne on the other hand, where the least dwellings have built-in kitchens, the presence of a kitchen has a rather strong positive impact on liquidity. A different picture emerges for Munich, where across the different cities on average, most dwellings include a parking space. Despite the high ratio of 65%, the feature cannot be declared as standard, as it yields a significant negative impact on liquidity. For terraces in Stuttgart and elevators in Munich, it can be said that the supply satisfies the demand. The highest ratio for terraces is found in Stuttgart, while across the cities, the most dwellings with elevators are found in Munich. For both cities, the presence of these features decreases the marketing time. In each city besides Frankfurt, where the highest ratio of dwellings offered for first occupancy are located, the feature increases the liquidity.

The degree of atypicality and the degree of overpricing show consistent results, as they confirm a restricted liquidity for highly atypical and overpriced dwellings with the exceptions of Munich and Stuttgart. Only there, does the degree of overpricing have a positive impact on liquidity, probably attributable to the strong demand over the last years, accompanied by insufficient housing construction. These factors are reflected in the lowest average time on market across the seven cities, and force households to let properties irrespective of the hedonic particularities. In addition to the strong overall demand for living space in Munich, there is a substantial demand for dwellings in the heart of the city centre, as the next section will demonstrate.

The spatial gravity variables included in the model show that for dwellings in Frankfurt, Cologne and Hamburg, the marketing time decreases with distance to the ZIP-code centroid. When looking at the spatial influence with larger granularity, it becomes clear that in six of the seven largest German cities, proximity to the city centre is significant when marketing a dwelling. The coefficients display prolonged marketing time for more decentralised flats.

The different model parameterizations reveal a substantial change in the hazard rates for the asking rents vector after controlling for atypicality, overpricing and gravity variables and especially when including time, spatial and socioeconomic control variables. The increase in the Pseudo R^2 between model I and III confirms however, that liquidity is captured more accurately when controlling for the latter variables, leading to a less pronounced bias from omitted variables. The assumption of proportionality is a central aspect of the Cox model. To test whether this assumption can be confirmed, the Therneau and Grambsch non-proportionality test, which measures the correlation between the covariate-specific residuals (so-called Schoenfeld residuals) and the event times, is applied. The test is Chi-squared distributed with one degree of freedom, and verifies that the assumption of proportionality in the models is not violated.

Table 3.2: Results Cox Proportional Hazards Model

Exp(coefficient) Mean / (P-Values)	Berlin			Frankfurt			Munich		
Model	I	II	III	I	II	III	I	II	III
Log rent	-0.056 0.002***	0.016 0.003***	-0.045 0.007***	-0.053 0.001***	-0.014 0.003***	-0.035 0.006***	-0.048 0.002***	0.002 0.005	-0.110 0.009***
Log area	-0.008 0.000***	-0.010 0.000***	-0.014 0.000***	-0.011 0.000***	-0.012 0.000***	-0.015 0.001***	-0.011 0.000***	-0.011 0.000***	-0.017 0.001***
Age	0.002 0.000***	0.000 0.000***	0.012 0.001***	0.001 0.000***	0.001 0.000***	0.014 0.003***	0.001 0.000***	0.001 0.000***	0.003 0.002
Number of rooms	0.123 0.006***	0.171 0.006***	0.199 0.007***	0.159 0.012***	0.174 0.013***	0.185 0.013***	0.144 0.012***	0.140 0.012***	0.203 0.015***
With bathtub (yes=1)	-0.159 0.007***	-0.177 0.007***	-0.167 0.007***	-0.110 0.014***	-0.094 0.014***	-0.082 0.014***	-0.056 0.013***	-0.070 0.013***	-0.024 0.013*
With built-in kitchen (yes=1)	0.042 0.007***	-0.082 0.008***	0.044 0.013***	0.099 0.016***	-0.001 0.018	0.043 0.024*	0.018 0.014	-0.037 0.015**	0.071 0.019***
With parking space (yes=1)	-0.112 0.008***	-0.159 0.008***	-0.039 0.011***	-0.073 0.018***	-0.114 0.018***	-0.057 0.020***	-0.085 0.016***	-0.101 0.017***	-0.050 0.018***
With terrace (yes=1)	-0.116 0.009***	-0.131 0.009***	0.011 0.012	-0.075 0.017***	-0.093 0.018***	-0.021 0.020	-0.113 0.016***	-0.147 0.016***	-0.011 0.021
With balcony (yes=1)	-0.053 0.007***	-0.042 0.007***	-0.030 0.008***	-0.034 0.015**	-0.051 0.015***	-0.040 0.016**	-0.060 0.014***	-0.061 0.014***	-0.022 0.014
With elevator (yes=1)	0.019 0.007***	-0.050 0.008***	-0.063 0.009***	0.001 0.017	-0.059 0.018***	-0.035 0.019*	0.077 0.014***	0.011 0.015	0.035 0.016**
First occupancy (yes=1)	-0.208 0.009***	-0.360 0,011***	0.097 0.030***	-0.144 0.019***	-0.215 0.019***	0.047 0.036	-0.236 0.017***	-0.374 0.019***	0.096 0.044**
Degree of atypicality		-0.266 0.007***	-3.571 0.29***		-0.230 0.015***	-1.904 0.493***		-0.279 0.017***	-1.955 0.622***
Degree of overpricing		-0.042 0.008***	-0.104 0.019***		0.024 0.016	-0.058 0.032*		0.054 0.018***	0.177 0.036***
Centroid to ZIP		-0.019 0.001***	-0.011 0.012		0.003 0.005	0.067 0.024***		-0.021 0.004***	-0.006 0.022
Centroid to NUTS3			-0.056 0.007***			-0.070 0.018***			-0.046 0.016***
Coordinates		✓	✓		✓	✓		✓	✓
Socioec. variables			✓			✓			✓
Age fixed effects			✓			✓			✓
ZIP fixed effects			✓			✓			✓
Time fixed effects			✓			✓			✓
Smoothing functions									
Pseudo R^2	61.9%	64.4%	65.9%	63.3%	64.3%	65.8%	65.3%	66.7%	68.1%
χ^2 cox prop test			11,100***			3,050***			5,640***
N	119,481			28,641			32,216		

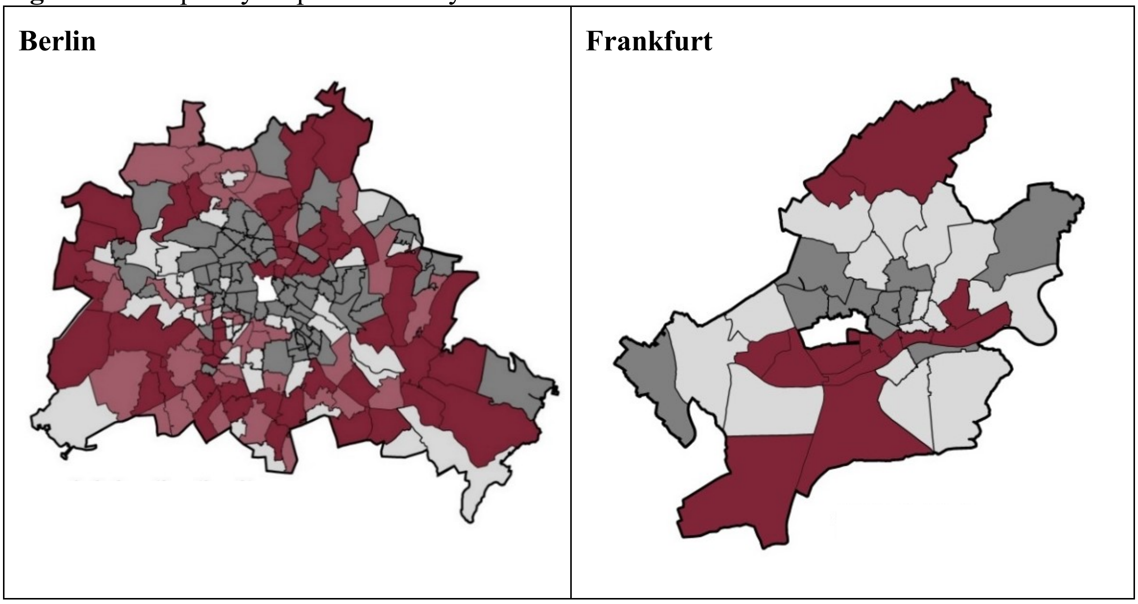
Notes:*Significant at the 10%-level; ** significant at the 5%-level; *** significant at the 1%-level. The exhibit shows the regression results of a semiparametric cox regression of dwellings time on market in weeks on hedonic, spatial, socioeconomic and smoothing covariates. The results are presented as coefficients. While significant positive values shorten the survival and thus increase the liquidity, significant negative coefficients decrease the hazard rate and lengthen the survival time. The three different model parameterizations control for different fixed effects. The Pseudo-R^2 based on Kendall's Tau measures the concordance between estimated survival time and the observed survival time for only the non-censored response sample. Proportionality test using the Therneau and Grambsch procedure with Schoenfeld-adjusted residuals under the null of non-proportionality.

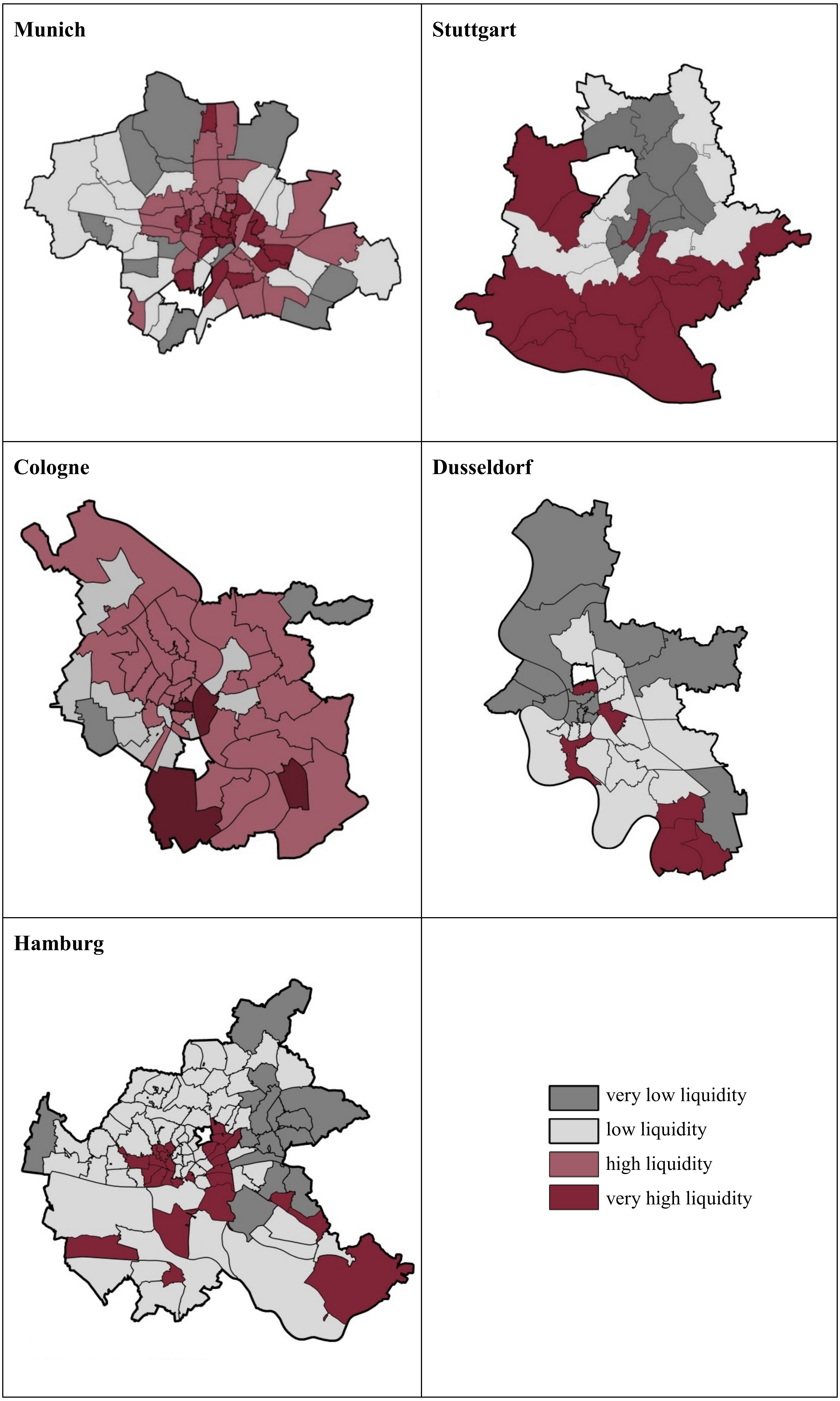
Stuttgart			**Cologne**			**Dusseldorf**			**Hamburg**		
I	II	III	I	II	III	I	II	III	I	II	III
-0.054 0.005***	0.088 0.012***	-0.093 0.017***	-0.019 0.003***	0.090 0.007***	-0.085 0.014***	-0.052 0.003***	0.051 0.007***	-0.035 0.011***	-0.045 0.002***	0.036 0.005***	0.003 0.008
-0.010 0.001***	-0.010 0.001***	-0.016 0.001***	-0.009 0.000***	-0.009 0.000***	-0.015 0.001***	-0.008 0.000***	-0.01 0.000***	-0.011 0.000***	-0.009 0.000***	-0.011 0.000***	-0.014 0.000***
0.001 0.000	0.001 0.000	-0.005 0.004	0.001 0.000***	0.000 0.000	-0.005 0.002**	0.000 0.000	0.000 0.000	0.000 0.002	0.001 0.000***	0.000 0.000	0.002 0.002
0.098 0.016***	0.099 0.017***	0.172 0.018***	0.116 0.011***	0.131 0.012***	0.175 0.013***	0.119 0.011***	0.140 0.012***	0.160 0.012***	0.126 0.009***	0.163 0.009***	0.172 0.010***
-0.129 0.020***	-0.136 0.020***	-0.053 0.021**	-0.071 0.011***	-0.071 0.012***	-0.042 0.012***	-0.118 0.012***	-0.122 0.012***	-0.102 0.012***	-0.022 0.009**	-0.026 0.010***	-0.012 0.010
-0.011 0.022	-0.191 0.028***	0.062 0.043	0.096 0.013***	-0.022 0.014	0.162 0.019***	0.150 0.013***	0.037 0.015**	0.228 0.019***	-0.112 0.012***	-0.227 0.014***	-0.246 0.028***
-0.045 0.025*	-0.097 0.026***	0.013 0.028	-0.063 0.013***	-0.109 0.014***	-0.021 0.016	-0.040 0.015***	-0.066 0.015***	0.026 0.017	-0.174 0.01***	-0.197 0.011***	-0.061 0.013***
-0.054 0.023**	-0.111 0.024***	0.088 0.032***	-0.074 0.015***	-0.144 0.015***	0.053 0.021**	-0.064 0.016***	-0.111 0.017***	0.109 0.024***	-0.049 0.011***	-0.042 0.012***	0.076 0.014***
-0.044 0.022**	-0.086 0.023***	-0.009 0.025	-0.035 0.012***	-0.044 0.013***	-0.012 0.014	-0.028 0.013**	-0.061 0.014***	-0.030 0.015**	0.023 0.010**	-0.003 0.010	0.021 0.011*
-0.055 0.027**	-0.129 0.028***	-0.066 0.030**	-0.083 0.013***	-0.132 0.014***	-0.151 0.014***	-0.046 0.013***	-0.117 0.014***	-0.040 0.016**	-0.064 0.011***	-0.094 0.012***	-0.162 0.013***
-0.117 0.029***	-0.343 0.034***	0.297 0.077***	-0.160 0.019***	-0.336 0020***	0.139 0.041***	-0.113 0.018***	-0.245 0.02***	0.235 0.042***	-0.171 0.013***	-0.234 0.014***	0.155 0.025***
	-0.407 0.031***	-2.765 0.821***		-0.329 0.014***	-3.221 0.492***		-0.236 0.015***	-4.246 0.500***		-0.243 0.012***	-3.749 0.351***
	0.049 0.024**	0.164 0.050***		0.077 0.011***	-0.058 0.030*		-0.034 0.015**	-0.019 0.022		-0.019 0.009**	-0.002 0.025
	0.022 0.007***	-0.040 0.029		-0.003 0.003	0.065 0.015***		-0.014 0.003***	0.016 0.019		-0.015 0.002***	0.027 0.013**
		-0.076 0.021***			0.004 0.008			-0.130 0.015***			-0.086 0.007***
	✓	✓		✓	✓		✓	✓		✓	✓
		✓			✓			✓			✓
		✓			✓			✓			✓
		✓			✓			✓			✓
		✓			✓			✓			✓
64.3%	66.2%	67.8%	60.9%	63.6%	65.3%	61.3%	62.6%	64.5%	64.9%	67.0%	68.2%
		1,900***			3,130***			2,050***			10,200***
12,755			36,940			35,814			70,125		

3.4.2 Liquidity in an Urban Spatial Context

When looking at the distribution of Berlin's most liquid ZIP-regions, one can clearly detect a more or less circular form surrounding the inner city. As it is impossible to infer a distinct pattern of spatial preferences, the cause of the strong demand for more decentralised dwellings might simply be the rental aspect. As central Berlin is becoming prohibitedly expensive, lower-income households often have to relocate to the more affordable outskirts. A very distinct constellation appears when looking at the most liquid regions of Hamburg. Since almost the entire southern half of Hamburg is enclosed by widespread natural reserves and Europe's third largest harbour, only solitary settlements are found within that area. The ZIP-regions with very high liquidity cluster themselves west- and eastwards of the old town. Particularly some of the western districts are among the most densely populated areas of Germany. Cologne yields a constant liquidity pattern for almost all of its ZIP-code regions. In Munich, it is notably the more expensive inner-city regions, that are highly liquid and let faster than the reference district. Those central districts contain on average 14% fewer parking spaces than the city average, thus explaining the negative effect on liquidity, when considering the entire city. The expansion of Stuttgart's inner-city is naturally bound by its geology. The districts displaying the highest market liquidity are found south of the circular valley containing the inner-city. The geology renders Stuttgart prone to traffic congestions and particle pollution. As large employers move to the outskirts, people do not hesitate to follow them, decreasing both commuting time and the health risk.

Figure 3.2: Liquidity dispersion at city level

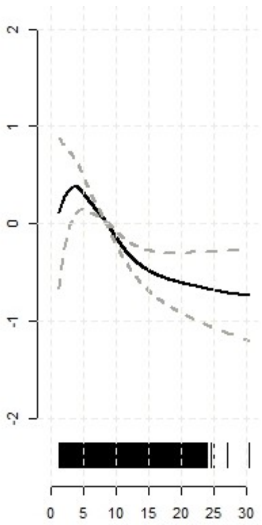
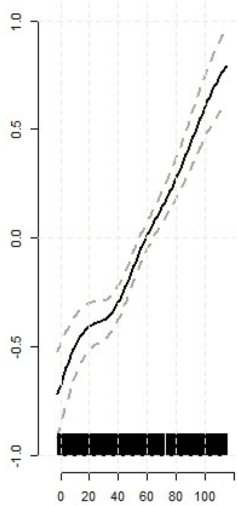



Munich
Stuttgart
Cologne
Dusseldorf
Hamburg
very low liquidity
low liquidity
high liquidity
very high liquidity

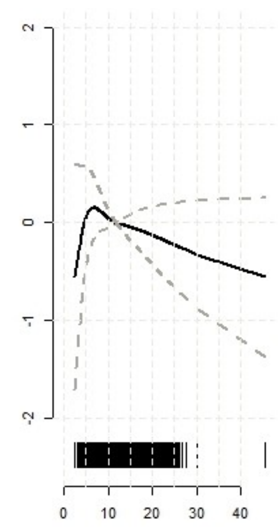
3.4.3 Non-linearity and its Impact on Liquidity

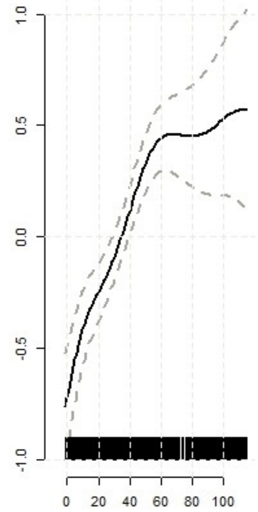
In this section, the effects of non-linear covariates on liquidity are presented. More specifically, the basic equation (12) is expanded by smoothing functions of rent and age, as described in equation (15). Based on the results from Exhibit 3, the non-linear effects of rent and age are presented in figure 3.3 as coefficients. In this context, values above zero increase the hazard, i.e. shorten the survival time, and therefore increase liquidity, i.e. dwelling's letting process. The graphs show that liquidity responds non-linearly to the dwelling asking rent and age. Besides Munich, where less expensive dwellings are always let faster, the functions display a kink, showing a short increase in liquidity with rent, before the market liquidity declines with higher rent. Each city exhibits its individual threshold at which liquidity is maximised. The non-linear results of age show a pronounced liquidity discount for dwellings smaller than ca. 60 m² in Berlin and 40 m² in Frankfurt. In Stuttgart and Cologne, on the other hand, there are liquidity discounts for flats larger than 80m² and 50m² respectively.

Figure 3.3: Non-linear effect of rent and age on coefficients

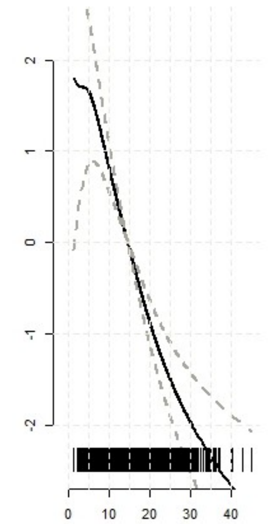
	Asking rent in €/m²	Dwelling age
Berlin		

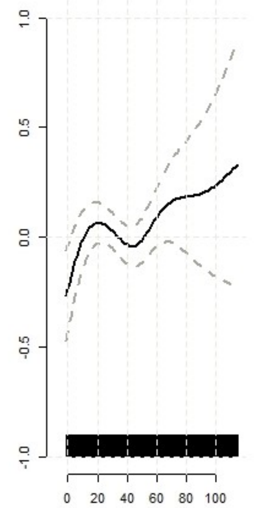
Frankfurt



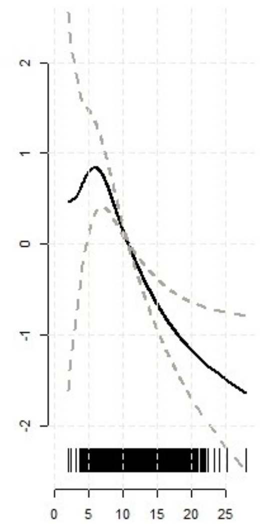


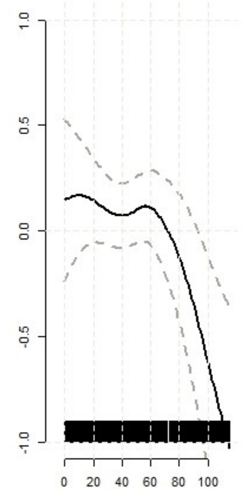
Munich

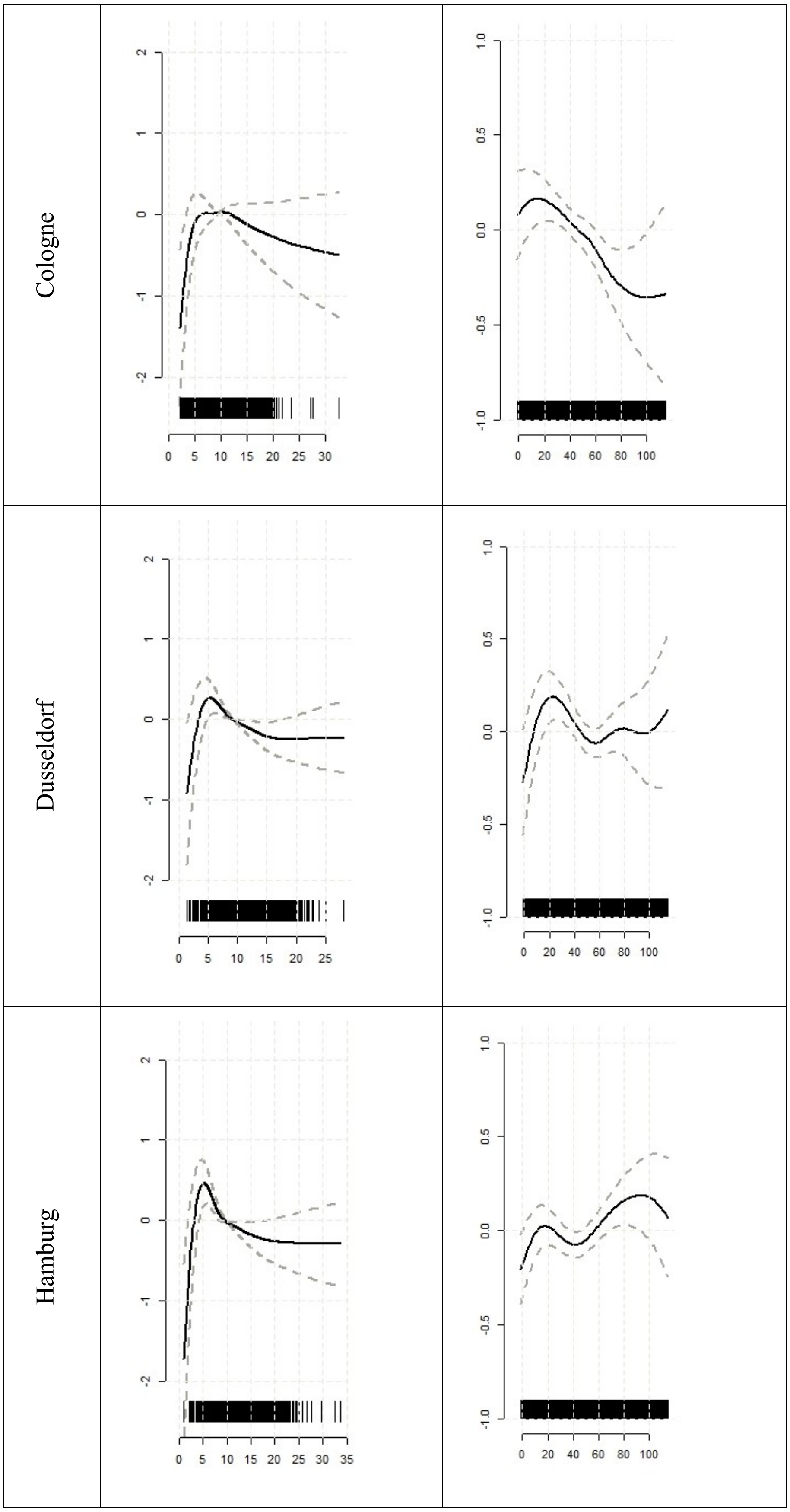




Stuttgart







Notes: The semiparametric Cox survival regression can be expanded to control for non-linear or smoothing effects of metric covariates. The results show the response of the log hazard to non-linear changes in asking rents and dwelling age. Values above zero increase the hazard and consequently liquidity. The vertical bars above the x axis display the density of the rent-level and age respectively. Confidence intervals are shown as dashed lines.

3.5 Conclusion

A broader knowledge of liquidity and the underlying factors is essential for assessing market movements with respect to the buying and selling of property by market players. This paper has explored several concepts of liquidity in residential rental markets and introduced a profound foundation of econometric tools that are necessary for capturing liquidity. The results based on big data can assist both private and institutional landlords in assessing the marketability of rental property within the observed markets and help governmental organisations in charge of housing and urban planning to derive policy implications. Especially for the abovementioned first time buyers, who are new to the housing market, this article contributes a condensed overview of the most liquid areas in the observed cities and provides an indication of which characteristics further increase the marketability of rental dwellings. Governmental organisations on the other hand, can obtain guidance on which trends are currently dominating the market and infer which regulatory or supportive actions might be necessary.

The paper contributes to a better understanding of liquidity in rental markets, an underexplored topic in traditional real estate housing research, using the German market with its extraordinarily low homeownership rate of about 45% as an example. Across the seven largest German real estate markets, the semiparametric Cox hazard models, controlling for hedonic, socioeconomic, spatial and various fixed effects displayed similarities as well as differences in the liquidity of rental dwellings and its determinants. While for each city, the asking rent, living area, dwelling age and distance to the NUTS3 centroid show consistent effects, the hedonic characteristics and the degree of overpricing display a market-specific impact on liquidity. Based on these results, geographic liquidity patterns are derived for the observed cities, and individual non-linear effects of asking rent and dwelling age are shown graphically. By means of this approach, the article contributes to the literature on time on market modelling, by enhancing the quality of the econometric approach, and by introducing spatial gravity variables to an extensive dataset covering a truly unique market in terms of homeownership rate.

3.6 References

Allen, M.T., Rutherford, R.C., and Thomson, T.A. (2009), "Residential asking rents and time on the market", *Journal of Real Estate Finance and Economics*, Vol. 38 (4), pp. 351-365.

Anglin, P.M., Rutherford, R.C., and Springer, T.M. (2003), "The trade-off between selling price of residential properties and time-on-the-market: The impact of price setting", *Journal of Real Estate Finance and Economics*, Vol. 26 (1), pp. 95-111.

Bourassa, S., Cantoni, E., and Hoesli, M. (2010), "Predicting house prices with spatial dependence: A comparison of alternative methods", *Journal of Real Estate Research*, Vol. 32 (2), pp. 139-159.

Bourassa, S., Haurin, D.R., Haurin, J.L., Hoesli, M., and Sun, J. (2009), "House price changes and idiosyncratic risk: The impact of property characteristics", *Real Estate Economics*, Vol. 37 (2), pp. 259-278.

Cajias, M., Fuerst, F., and Bienert, S. (2016), "Are energy efficiency ratings ignored in the German housing market? – Evidence from a large-sample hedonic study", Available at: https://papers.ssrn.com/sol3/papers.cfm?abstract_id=2799206 (accessed 22 December 2016).

Cirman, A., Pahor, M., and Verbic, M. (2015), "Determinants of time on the market in a thin real estate market", *Inzinerine Ekonomika-Engineering Economics*, Vol. 26 (1), pp. 4-11.

Cox, D.R. (1972), "Regression models and life-tables", *Journal of the Royal Statistical Society: Series B (Methodological)*, Vol. 34 (2), pp. 187-220.

Fik, T.J., Ling, D.C., and Mulligan, G.F. (2003), "Modelling spatial variation in housing prices: A variable interaction approach", *Real Estate Economics*, Vol. 31 (4), pp. 623-646.

Glower, M., Haurin, D., and Hendershott, P.H. (1998), "Selling time and selling price: The influence of seller motivation", *Real Estate Economics*, Vol. 26 (4), pp. 719-740.

Goodman, A.C. and Thibodeau, T.G. (2007), "The spatial proximity of metropolitan area housing submarkets", *Real Estate Economics*, Vol. 35 No. 2, pp. 209-232.

Haurin, D.R. (1988), "The duration of marketing time of residential housing", *Real Estate Economics*, Vol. 16 (4), pp. 396-410.

Haurin, D.R., Haurin, J.L., Nadauld, T., and Sanders, A. (2010), "List prices, sale prices and marketing time: An application to U.S. housing markets", *Real Estate Economics*, Vol. 38 (4), pp. 659-685.

Haurin, D.R., McGreal, S., Adair, A., Brown, L., and Webb, J.R. (2013), "List price and sales prices of residential properties during booms and busts", *Journal of Housing Economics*, Vol. 22 (1), pp. 1-10.

Heckman, N.E. and Ramsay, J. (2000), “Penalized regression with model-based penalties”, *Canadian Journal of Statistics*, Vol. 28 (2), pp. 241-258.

Hoeberichts, M., van Rooij, M., and Siegmann, A. (2013), “House list prices and durations in boom and bust“ available at: https://ssrn.com/abstract=2321521 (accessed 16 December 2016).

Kluger, B.D. and Miller, N.G. (1990), “Measuring real estate liquidity”, *Real Estate Economics*, Vol. 18 (2), pp. 145-159.

Krainer, J. (1999) “Real estate liquidity”, *Economic Review*, Vol. 3 (1), pp. 14-26.

Pavlov, A.D. (2000), “Space-varying regression coefficients: A semi-parametric approach applied to real estate markets”, *Real Estate Economics*, Vol. 28 (2), pp. 249-283.

Smith, B.C. (2010), “Spatial heterogeneity in listing duration: The influence of relative location to marketability”, *Journal of Housing Research*, Vol. 18 (2), pp. 151-171.

Turnbull, G.K. and Dombrow, J. (2006), “Spatial competition and shopping externalities: Evidence from the housing market”, *Journal of Real Estate Finance and Economics*, Vol. 32 (4), pp. 391-408.

Voigtlaender, M. (2009), “Why is the German homeownership rate so low?”, *Housing Studies*, Vol. 24 (3), pp. 355-372.

4. Exploring the determinants of real estate liquidity from an alternative perspective – Censored Quantile Regression in real estate research

Abstract

In this paper, the liquidity (inverse of time on market) of rental dwellings and its determinants for different liquidity quantiles are examined for the largest seven German cities. The determinants are estimated using Censored Quantile Regressions in order to investigate the impact on very liquid to very illiquid dwellings. As market heterogeneity is not only observed within one market but also between the cities, each of the seven cities is considered separately. Micro data for almost 500,000 observations from 2013 to 2017 is used to examine the time on market. Substantial differences in the magnitude and the significance of the regression coefficients for the different liquidity quantiles are found. Furthermore, the magnitude as well as the direction of the impact an explanatory variable has on the liquidity, differ between the cities. This is the first paper, to the best of the authors' knowledge, to apply censored quantile regressions to liquidity analysis on the real estate rental market. The model reveals, that the proportionality assumption underlying the Cox Proportional Hazards Model cannot be confirmed for all variables across all cities, but for most of them.

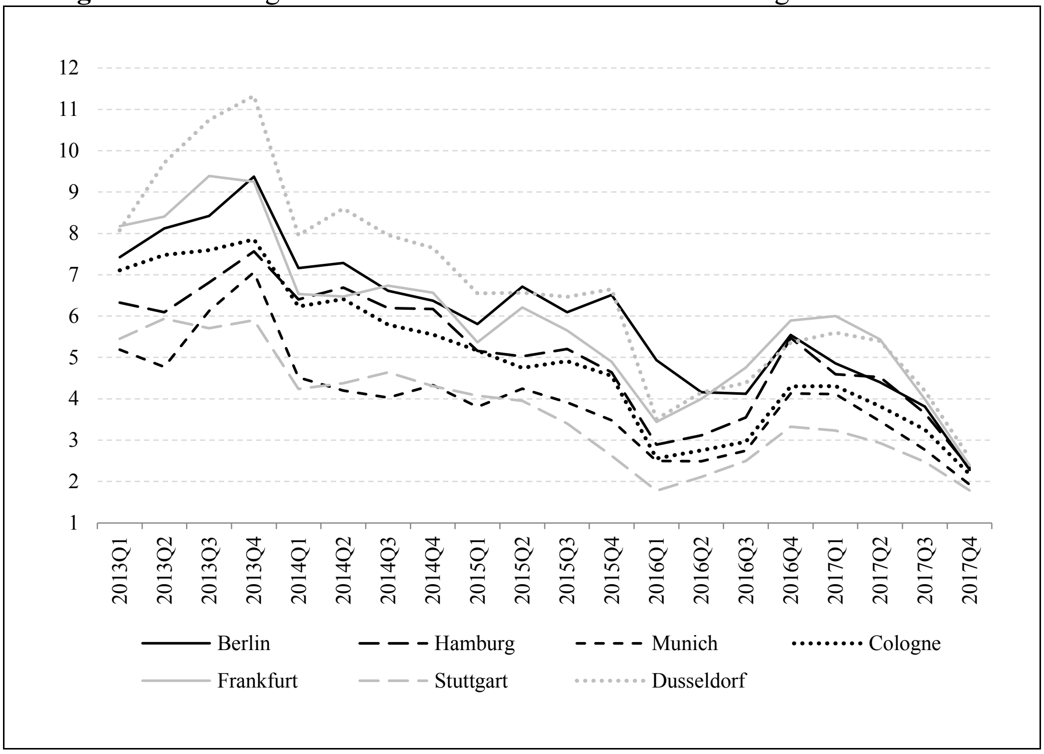
Acknowledgement: The authors especially thank PATRIZIA Immobilien AG for contributing the dataset for this study. All statements of opinion are those of the authors and do not necessarily reflect the opinions of PATRIZIA Immobilien AG or its associated companies.

4.1 Introduction

The concept of asset liquidity on the rental market is somewhat fuzzy. On the investment market, asset liquidity traditionally measures the time it takes the owner to turn the asset into cash. On the rental market on the other hand, asset liquidity measures the time it takes the landlord to find a new tenant, i.e. from introducing the dwelling onto the market until the signing of the rental contract. In this context, the study examines liquidity solely with a time-based measure and does not include transaction cost, price, or volume measures. Whether the letting process is quick or slow, depends among other things on the amount of the initial asking rent, the structural quality and location of the asset, the demand within the segment and the overall market conditions. A detailed examination of the factors that determine the time on market (TOM) on the rental market is the objective of this study. In a first step, TOM is explored from a traditional research perspective via the Cox (1972) Proportional Hazard Model (PHM). The survival regression is estimated for a unique data set consisting of 482,196 observations on the German rental market. In a next step, a new econometric approach is introduced to the field of real estate liquidity analysis. The Censored Quantile Regression (CQR) aims at explaining the variation of liquidity as a function regarding the dwelling characteristics and other spatial and socioeconomic characteristics. In other words, the CQR controls for the heterogeneity of the assets, assuming that highly liquid dwellings respond differently to certain covariates than very illiquid dwellings. The CQR, as an expansion of the survival regression analysis, is expected to yield more accurate estimations and to provide a much more solid basis for drawing conclusions about the factors affecting liquidity.

In the study, the inverse of the time on market of rental dwellings is used to construct a liquidity measure based on the liquidity definition by Wood and Wood (1985). The goal is to identify dwelling characteristics which shorten the time it takes to rent out a property or in other words, characteristics which make the rental dwellings more liquid. From the regulatory perspective, tenants in Germany have a three months cancelation period, for which reason a dwelling is usually brought onto the market before the tenant leaves. But why is it important to investigate the time on market or liquidity for rental dwellings on the German market? With about 43% ownership rate as of 2013, the first year of the sample period, obviously more than half of the German households rent their homes. Voigtlaender (2009), Bentzien et al. (2012), Lerbs and Oberst (2014), and Reisenbichler (2016) very well describe reasons for this distinctive market feature. The resulting tremendous amount of cash-flow generating rental dwellings is strongly demanded by a global base of institutional investors in order to benefit from the stable nationwide macroeconomic environment, the diversification benefits resulting from the high polycentricity of the German market, and the availability of large tradable portfolios. The largest seven cities by population are at the same time the main investment markets. Therefore, the analysis focuses on the top 7 markets Berlin, Cologne, Dusseldorf, Frankfurt, Hamburg, Munich, and Stuttgart. In these cities, the ownership rates are far lower than the German average, ranging from 33% in Stuttgart to 16% in Berlin. Therefore, the examination of the rental market should be better suited to allow conclusions regarding the entire residential real estate market within those cities. Figure 4.1 illustrates the average time on market for the top 7 rental markets. Despite some up and down movements, the graph clearly shows a continuous decline in time on market within the last five years. This development points to an increasingly strong demand on the rental market, as the prospective tenants have to shorten their decision making process because of high competition for insufficient supply. This excess demand was recognized by Held and Waltersbacher (2015) and declared to comprise 272,000 new dwellings per year for the years 2015-2020, while the actual completion did not meet that goal.[4]

Figure 4.1: Average time on market in weeks in the seven largest German cities



Notes*:* This figure displays the average time on market in weeks in the largest seven German cities. The data consists of 482,196 observations on the residential rental market. The sample period is 2013 Q1 to 2017 Q4.

The CQR splits the city subsamples into quantiles based on their liquidity level and measures the impact of each covariate on the particular quantile. This study is the first, to the best of the authors' knowledge, to investigate the determinants of time on market by applying a Censored Quantile Regression. Using an extensive data set, the study is able to identify patterns of impact the explanatory variables have on time on market. These patterns vary based on the location of the dwelling as well as the liquidity level of the dwelling. The study finds, that the magnitude of the impact an explanatory variable has on the liquidity of a dwelling differs between the cities as well

[4] According to the Federal Statistical Office (2018) dwelling completion was: 216,727 (2015), 235,658 (2016), 245,304 (2017)

as with the level of liquidity. In other words, the size of a dwelling for example exerts a different effect on the level of demand for certain dwellings between and within the cities, thus shows a varying effect on TOM. Furthermore, the direction of the effect an explanatory variable has on the liquidity of a dwelling exhibits statistically significant differences between and within the cities. Hence, the proportional hazard assumption, underlying the Cox (1972) Proportional Hazards Model is violated for individual explanatory variables and cities. The study concludes, that the heterogeneity across the liquidity quantiles as well as the heterogeneity between the cities are accountable for the distinguishable impacts of changes in the covariates on time on market. Those findings should of course be of interest to current and future landlords, as they might draw implications about their existing and future properties. Moreover, the study shows both the characteristics of dwellings along the liquidity distribution as well as the impact of a change in the characteristics on the liquidity of the dwellings. Therefore, landlords should be able to infer whether a dwelling displays the characteristics of a highly liquid thus highly demanded dwelling or what actions they could take in order to shorten the expected liquidity, e.g. install a built-in kitchen or change the floor plan to increase the number of rooms. Furthermore, the revealed necessity to analyze and interpret markets on a very granular basis yields implications for policy makers. The findings suggest, that nationwide or even statewide policy measures might not be sufficient to capture the individual needs of regions, cities or neighborhoods.

Since the seminal work by Rosen (1974), real estate hedonic pricing models are the most widely used methods for capturing the heterogeneity of residential units. Guntermann and Norbin (1987), Sirmans et al. (1989), Sirmans and Benjamin (1991), Valente et al. (2005), and Allen et al. (2009) among others, investigated the impact of individual characteristics on the rent level of residential property. Rental property on the German market was examined e.g. by an de Meulen and Mitze (2014) and Thomschke (2015), with both studies focusing on the Berlin market. As for the pricing of residential real estate liquidity estimates are as well likely to be biased and market knowledge might be distorted if heterogeneity is not taken into account. Deeper insights into specific market segments do not only provide reduced search costs to households, see e.g. Malpezzi (2003), Goodman and Thibodeau (2007), but also improve the financial appraisal of private and institutional lenders and investors. Therefore, the segmentation into submarkets leads to a more profound understanding of liquidity patterns in the residential real estate market.

Belkin et al. (1976) conducted one of the first empirical studies analyzing real estate liquidity for different market segments. They define submarkets according to geographic areas, price segments and buyers' search space and analyze the relationship between TOM and the spread between listing price and selling price using ordinary least squares (OLS) estimation techniques. They find essential differences by market segments. Especially in high-price submarkets, deviations from the initial list price had a more pronounced effect on time on market. The determinants of TOM considering different price segments are further analyzed by Kang and Gardner (1989), Kalra and Chan (1994), Yavas and Yang (1995) and Allen et al. (2009) among others. Kang and Gardner

(1989) found that the impacts on TOM do vary in magnitude between the low-, medium- and high-price segments. While Kang and Gardner (1989) did not identify the simultaneity problem between time on market and the selling price, Yavas and Yang (1995) applied a two-stage least squares (2SLS) estimation to deal with the fact that time on market and price mutually influence each other. They exhibit a significant positive impact of price on time on market in the medium-price subgroups, whereas this effect is insignificant for houses in the low- and high-price segments. Allen et al. (2009) analyze the relationship between asking rent and time on market on the rental market of single family residential rental listings using a multi-step procedure. Based on asking rents the sample is divided into three price subgroups. They find that underpricing of asking rents and time on market move in the same direction in every price segment, however the effect being stronger in the medium- and high-price submarket, compared to low-price houses. Further studies include Guasch and Marshall (1985), Ong and Koh (2000), Turnbull et al. (2006) and McGreal et al. (2009) among others, who segment markets according to the number of rooms, the number of units in a structure, the geographical region, the property type or by market cycle, respectively. As a conclusion from these findings, market segmentation seems to be a valuable contribution for understanding liquidity patterns.

More closely related to the present study is the article of Turnbull and Dombrow (2006), as they divide their sample into low-, medium- and high- liquidity submarkets. They explore the impact of listing density on time on market for a pooled sample, for different market cycles and for different market cycles combined with different liquidity segments. Applying a three-stage least squares (3SLS) estimation, they find that the significance as well as the magnitude and directions of the impact of the spatial competition variables on time on market vary between the different liquidity submarkets. Based on these findings, the following hypotheses are deduced

1. the direction and magnitude of the effect of covariates on real estate liquidity is not equal and varies across low, medium and high liquidity segments and
2. if the latter holds, the Cox proportional hazard assumption would be violated, justifying the usage of an approach able to deal with heterogeneous effects.

As it is possible to investigate and compare the impact of factors on the time on market for each of the seven largest German cities, the additionally hypothesis is, that

3. the direction and magnitude of the impact of the covariates vary across these cities.

Nowadays, the most popular model for the estimation of duration data is the Cox (1972) PHM. Also commonly used is the accelerated failure time (AFT) model. However, in terms of the econometric model, the paper strongly differs from the preceding studies, as censored quantile regressions (CQRs) are applied to real estate liquidity analysis. Quantile Regression (QR) has been formally introduced by Koenker and Bassett Jr. (1978). Compared to the accelerated failure time model or the Cox (1972) PHM, QR is a more flexible estimation method as it allows for consistent estimation of the regression model without restrictions on the variation of estimated coefficients over the quantiles. The decisive feature of the analysis, however, is that QRs are used to model any quantile of the distribution of the dependent variable. Chaudhuri et al. (1997) stress this feature as a great advantage compared to mean regressions as distributions might not only be different by their means but might especially differ in their upper and lower parts. Thus, QRs can quantify the impact of a covariate on the dependent variable for any quantile compared to only the center of the population. In contrast to linear regression, QR coefficients are computed via minimizing the sum of weighted absolute deviations.

Since its introduction, the QR approach has received increasing attention, theoretically as well as empirically, and has been applied to many different research areas.[5] In the real estate literature, more precisely in the area of hedonic pricing, QRs have been applied by Zietz et al. (2008), Farmer and Lipscomb (2010), Mak et al. (2010), Liao and Wang (2012), among others. An de Meulen and Mitze (2014) and Tomschke (2015) used the method on the German market. However, when it comes to real estate liquidity, this is the first paper, to the best of the authors' knowledge, to use QRs with censoring for duration analysis on the real estate market. For the closely related analysis of (un)employment durations Horowitz and Neumann (1987) have initially, as well as Luedemann et al. (2006), Fitzenberger and Wilke (2010), Schmillen and Moeller (2010) among others, have lately applied CQRs. Conceptually the present analysis is highly related to Luedemann et al. (2006). In particular, the CQR method used in this paper goes back to Koenker and Bilias (2002). A comprising description on the implementation of censoring into the R package "quantreg" can be found in Koenker (2008).

The remainder of this paper proceeds as follows: The next section describes the underlying econometric model, followed by a detailed description of the dataset and the descriptive statistics in section 4.3. Estimation results are presented and discussed in section 4.4, Section 4.5 concludes.

4.2 Econometric Approach

4.2.1 Cox Proportional Hazards Model and Quantile Regression Model

Without any doubt the leading model for the analysis of survival data is the Cox (1972) PHM. This model is used for exploring the determinants of the duration of an event or elapse of time, e.g. it determines the variables that accelerate or restrict the elapse of time that a response variable needs to change its state. In this case, the response variable is defined as a non-negative continuous

[5] for survival analysis see e.g. Crowley and Hu (1977), Yang (1999), Koenker and Geling (2001)
for medical research see e.g. Cole and Green (1992), Royston and Altman (1994), Harder et al. (2005), Owen et al. (2005), Wei et al. (2006), Beyerlein et al. (2008), Wehby et al. (2009)
for financial economics see e.g. Taylor (1999), Bassett Jr. and Hsiu-Lang (2002)
for environmental research see e.g. Hendricks and Koenker (1992), Pandey and Nguyen (1999)
for labour economics see e.g. Rose (1992), Buchinsky (1994, 1995)
for a review on the application fields of QRs see e.g. Yu et al. (2003)

variable, measuring the elapse of time that a dwelling requires for changing its status from being offered on the market into being out of the market in weeks, i.e. time on market. For understanding and estimating survival data, two main functions are essential: the survival function *S*(*t*) and the hazard rate function $\lambda(t)$. The survival function specifies the probability that an event has not occurred until a certain time t and is formally defined as

$$S(t) = P(T \geq t) = 1 - F(t) = \int_t^{\infty} f(x)dx, \tag{16}$$

with *f*(x) being the probability density function (p.d.f.) of the time until the event. The hazard function $\lambda(t)$, in contrast, describes the probability at *t* that an event occurs at time *T*, given that the event has not occurred before and is given by

$$\lambda(t) = \lim_{\Delta t \to 0} \frac{P(t \leq T < t + \Delta t | t \leq T)}{\Delta t}. \tag{17}$$

The relationship between those two functions is straightforward since the integrated hazard rate $\Lambda(\mathrm{t}) = \int_0^t \lambda(x)dx$ can be expressed as the negative log of the survival rate *S*(*t*) as $\Lambda(t) = -\log S(t)$. The survival function expresses the probability of a dwelling for staying on the market while the hazard function measures the risk of the same dwelling for leaving the market. The Cox PHM estimates the survival function, but most importantly focusses on the estimation of the hazard function by transforming the response variable in units of time into a discrete variable, so called conditional odds. The Cox regression for a specific observation *i* is given as

$$\lambda_i(t|x_i) = \lambda_0(t) \exp(-x_i'\beta), \tag{18}$$

where *x* is a vector of covariates (without the constant), β is a vector of parameters and λ t is the non-negative baseline hazard. The Cox PHM requires no specification of the functional form of the baseline hazard $\lambda_0(t)$. It assumes however, proportional hazards, meaning that the hazard function is a constant function of time. Taking logs results in a simple additive model for the log of the hazard rate

$$log\lambda(t|x) = log\lambda_0(t) - x'\beta, \tag{19}$$

And thus, the conditional survival function *S*(*t*|*x*) can be described as

$$\log(-logS(t|x)) = log\Lambda_0(t) - x'\beta. \tag{20}$$

Consequently, the model can be written as

$$log\Lambda_0(T) = x'\beta + u, \tag{21}$$

What is equivalent to the transformation model

$$h(T) = x'\beta + u, \tag{22}$$

with *h*(*T*) being a monotone transformation of the observed survival time *T* and *u* being iid with extreme value distribution $F(u) = 1 - \exp(-\exp(u))$. Hence, the Cox (1972) PHM can be written as a monotone transformation of the observed survival time *T* linearly depending on the covariates x plus an iid error u. The elapse of time that a dwelling is offered on the market corresponds to an event that might be censored on the right. Censoring refers to incomplete event cases in which the beginning or the end of an event is unknown. Right censoring is more common and arises

when the landlord doesn't change the status of the dwelling in the Multiple Listing Services (MLS) database or the dwelling is still being offered in the market. The Cox regression framework allows the censored events of the sample to contribute to the model until the end of the observation period. Therefore, a semiparametric PHM is estimated for each of the k=seven cities according to

$$h(t_{ijp}) = \exp(X_i\beta + Z_j\alpha + R_p\,\lambda + \mu_{ip}\delta_p + \mu_{ij}\rho_j) + e_{ijp}. \tag{23}$$

Building upon Cajias and Freudenreich (2018), a specification with slightly different regressors has been estimated. The hazard function h of the time on market t depends on the matrix X, containing the hedonic characteristics of dwelling i, Z including time-invariant socioeconomic data on ZIP-code level j, R representing the time-varying effective 10-year interest rate for housing loans and μ_{ip} and μ_{ij} accounting for p time- and j spatial effects respectively. Time fixed effects as quarterly dummies ranging from 1 to 20 have been included to absorb seasonal fluctuations and to account for structural breaks. The construction dummies describe the period of time, a dwelling was built in, on a ten year range from 1910 until today.

While the Cox (1972) PHM is the most common tool for explaining time on market in social sciences, natural sciences and also real estate studies, new techniques have been developed in order to account for conditional survival functions across different levels of the response. The traditional Cox regression estimates the conditional survival function for the entire sample based on the assumption of homoscedasticity within the sample. The covariates are expected to exert the same impact on the response regardless of the distribution of the response, e.g. highly liquid and poorly liquid dwellings. Thus, the approach ignores conditional elasticities, implying that for example highly liquid dwellings respond differently to certain covariates than very illiquid dwellings. In other words, the coefficients resulting from a Cox PHM are valid for the entire population, while the quantile approach estimates different coefficients for different segments of the population. A traditional example when explaining quantile regression in the duration of unemployment. When using a Cox PHM, the effect of the covariate experience will not distinguish between long-term and short-term unemployed persons. In contrast, the censored quantile regression takes the different segments of the response, i.e. long-term and short-term unemployment into consideration and estimates several equations with of course different elasticities.

In this context, quantile regressions have arisen as a method for estimating conditional regressions within the sample as a function of the quantile distribution of the response. In this paper, a unique technique corresponding to the survival quantile regression is employed, which has not been introduced to the context of real estate liquidity, to the best of the authors' knowledge. This method yields a robust and more flexible alternative for the estimation of parametric and semiparametric duration models by imposing less distributional assumptions. Moreover, there are no imposed modelling assumptions to be empirically proven true and thus, misspecification of the model is less likely.

4.2.2 Quantile Regression Model

The origin of the QR model goes back to Koenker and Bassett Jr. (1978). It is a location model estimating the relationship of the covariates x with the dependent variable y, conditional on the quantile τ of y. The quantile $\tau \in (0, 1)$ is defined as the value of y that separates the observations into the fraction τ below and the fraction 1-τ above. Thus, the quantile τ of a random variable Y is defined as the minimum value $q\tau$, so that

$$q_{\tau} = F^{-1}(\tau) = inf\big(y{:}\, F(y) \geq \tau\big), \tag{24}$$

where $F(y) = P\big(Y \leq y\big)$ denotes the cumulative distribution function (c.d.f.) of Y. Hence, the median for example is described by $\tau = 0.5$.

Following Doksum and Gasko (1990), several survival analysis models, such as the PHM, the proportional odds model or the accelerated failure time model, can be linked to the general transformation model

$$h(y_i) = x_i'\beta + u_i. \tag{25}$$

h(y_i) denotes a monotone transformation of the observed dependent variable y_i, linearly depending on a k x 1 vector of covariates x_i with $x_{1i} \equiv 1$ and an iid error u_i. With the error term u being defined as $u^{\tau} \equiv \mathrm{h(y)} - x'\beta^{\tau}$ it follows that $Quant_{\tau}(u^{\tau}|x) = 0$. Thus, the conditional quantile function of the transformed dependent variable y_i can be denoted as

$$Quant_{\tau}(h(y_i)|x_i) = x_i'\beta^{\tau}. \tag{26}$$

It describes the family of QR models. β^{τ} denotes a k x 1 vector of regression parameters dependent on the quantile τ.

Applying the "log"-transformation of T_i, $h(T_i)$ = lnT_i according to e.g. Chaudhuri et al. (1997), yields the accelerated failure time model as basis for the relationship between time on market and the covariates dependent on the conditional quantile. The underlying model can be described as

$$lnT_i = x_i'\beta^{\tau} + u_i^{\tau}. \tag{27}$$

The conditional quantile functions of the logarithm of the time on market can be written as

$$Quant_{\tau}(lnT_i|x_i) = x_i'\beta^{\tau}, \tag{28}$$

where $Quant_\tau(lnT_i|x_i)$ represents the τth conditional quantile of $\ln T_i$ given x_i. With the application of QR, it is possible to investigate changes in the relation between the covariates and the time on market depending on the liquidity segment. The quantile approach seems furthermore plausible, given the large datasets of almost 500,000 observations and due to the spatial heterogeneity in the data. The QR approach is expected to provide deeper insights into the underlying determinants of time on the market on the German rental housing market.

4.2.3 Censored Quantile Regression Model

An important feature of survival analysis is, that some observations do not change their event status throughout the observation period, as some dwellings remain available in the MLS database until the end of the observation period. If this is the case, the response variable, time on market T_i, is right-censored. To deal with censoring within the QR framework, three main approaches have been introduced by Powell (1984, 1986), Portnoy (2003) and Peng and Huang (2008). For the present dataset, Powell's (1984, 1986) approach is best suited as it addresses fixed censoring. For QRs with fixed censoring, it is necessary to know the observation specific censoring value C_i for all observations. If an observation *i* is censored, it is not possible to observe the actual survival time T_i, but to observe the observation specific censoring value C_i instead. Thus, in a right-censored dataset T_i is given by $T_i = \min\{T_i^*, C_i\}$. $C_i = lnT_i$, if an observation is censored and $C_i = +\infty$, if an observation is not censored. The CQR estimator $\widehat{\beta^\tau}$ is the value of β^τ solving the minimization problem of the distance function

$$Q_N(\beta;\tau) \equiv \frac{1}{N}\sum_{i=1}^{N} \rho_\tau(ln\,T_i - \min(x_i'\beta^\tau, C_i)). \tag{29}$$

The minimization term becomes $x_i'\beta^\tau$ if $x_i'\beta^\tau < C_i$ and is C_i otherwise. Thus, $x_i'\beta^\tau$ is censored from above at the upper threshold C_i. The "check-function" $\rho_\tau(u)$ is defined as

$$\rho_\tau(u) = \begin{cases} \tau * |u| & u \geq 0 \\ (1-\tau) * |u| & u < 0 \end{cases}. \tag{30}$$

$\tau * |u|$ denotes the penalty for underprediction and $(1-\tau) * |u|$ for overprediction. The estimator $\hat{\beta}$ that minimizes the distance function $Q_N(\beta; 0.5)$, i.e. at the median $\tau = 0.5$, describes a special case yielding the censored least absolute deviations (LAD) estimator $\widehat{\beta^{0.5}}$. The coefficients can be interpreted as the change in the dependent variable that, ceteris paribus, arises from a marginal change in the respective regressor while keeping the dependent variable in the same quantile according to Machado and Mata (2000). An increase of an explanatory variable e.g. "price" by a marginal unit, ceteris paribus, prolongs or shortens the time on market by $\left[\left|1-\exp\left(\widehat{\beta_{price}^{\tau}}\right)\right| * 100\right]\%$ in the same quantile τ . A prolongation of the time to event occurs if the hazard ratio $\exp(\widehat{\beta_{price}^{\tau}})$ is greater than 1 and a reduction of the time to event if the hazard ratio is smaller than 1.

Powell (1986) demonstrates, that under appropriate conditions for a certain value of τ, the censored regression quantile estimator β^{τ} is $\sqrt{N}$-consistent and asymptotic normality is proven true, if the appropriate assumptions hold for each $\tau \epsilon \{\tau_1, \ldots, \tau_J\}$. While in the uncensored QR, the objective function to be minimized $Q_N(\beta; \tau)$ is convex, this nice property is not given for the censored case, leading to some strong computational difficulties.

4.3 Data and descriptive Statistics

The estimation sample is composed of three merged data sets, containing information of 482,196 observations on single- and multi-family rental dwellings in the largest seven German cities from the first quarter of 2013 to the fourth quarter of 2017. Information on the rental dwellings are gathered from various Multiple Listing Services (MLS) as collected from the Empirica Systems Database. The database contains real estate market data from more than 100 sources, among them the most important MLS providers. Characteristics of the rental dwellings contain the time on market as the number of weeks the flat was listed in the MLS calculated by the start and end date according to e.g. Benefield and Hardin (2015) and the asking rent in absolute terms measured in € per month. A significant bias stemming from a possible deviation from contract rents is not to be expected as according to Shimizu et al. (2012) and Lyons (2013), among others. Other typical housing attributes and hedonic characteristics like "with balcony" are included as binary variables being 1 if the flat exhibits the characteristic and 0 otherwise. Since the data is georeferenced, two spatial gravity indicators, measuring the Euclidian distance of each dwelling to the geographical centroid of the ZIP and NUTS3 polygon in kilometers, are incorporated. NUTS3 regions correspond to "the nomenclature of territorial units for statistics", which is a hierarchical system for dividing the economic territory in Europe. The NUTS3 regions cover small regions similar to counties or administrative districts. In the sample, every city represents one NUTS3 region and therefore, the distance to the NUTS3 centroid describes the distance to the geometric city center. The socioeconomic variables purchasing power per household and the number of households at the ZIP code level, are extracted from the GfK-database. The population density per km² in a ZIP code area is calculated in ArcGIS. The last source is Thomson Reuters Eikon, providing the 10-year interest rate for housing loans as a macro variable. The variables, their units and sources can be found in table 4.1.

Table 4.1: Variables and sources

Variable	**Unit**	**Effect in the survival equations**			**Source**			
		Hedonic effects	Spatial effects	Socio-economic effects	Empirica	GfK	ArcGIS	Reuters
Asking rent	€ per month	✓			✓			
Time on market	Weeks2	✓			✓			
Living area	M²	✓			✓			
Age	Years	✓			✓			
Rooms	Number	✓			✓			
With bathtub	Binary	✓			✓			
With built-in kitchen	Binary	✓			✓			
With car space	Binary	✓			✓			
With terrace	Binary	✓			✓			
With balcony	Binary	✓			✓			
With elevator	Binary	✓			✓			
Newly built dwelling	Binary	✓			✓			
Refurbished dwelling	Binary	✓			✓			
Gaussian longitude	Coordinate		✓				✓	
Gaussian latitude	Coordinate		✓				✓	
Distance to ZIP centroid	Km		✓				✓	
Distance to NUTS3 centroid	Km		✓				✓	
Households in ZIP	HHs/ZIP			✓		✓		
Purchasing power of HHs in ZIP	€/HH/p.a./ ZIP			✓		✓		
Population density in ZIP	Persons/km²/ZIP			✓		✓		
IR for housing loan 10 years	Effective interest rate in %			✓				✓
N					482,196			

Notes: This table reports the unit, the type of effect, and the source of all variables included in the hedonic price and liquidity index calculations as well as the number of observations.

Table 4.2 shows the descriptive statistics for the variables of interest for each of the seven cities. It indicates the heterogeneity within each city by displaying the variation of the variables. The heterogeneity between the cities emerges by comparing the variation of the individual variables across the cities. The appeal of using the quantile regression to explain time on market becomes apparent by investigating the variation in the variable of interest. A relative standard deviation ranging from 1.35 to 1.63 implies, that on average, the time a dwelling is advertised on the market deviates from the mean by 1.35 to 1.63 times the mean. In absolute values, the time on market within the cities deviates on average between 5.93 to 8.91 weeks from the mean. Across the cities, the variation in time on market rises along the distribution curve as shown in table 4.3. While the average time a dwelling is advertised on the market stretches from 3.73 to 6.6 weeks, the variation becomes more apparent along the distribution, with a spread from 5.3 weeks in Munich to 10 weeks in Dusseldorf for the 80th percentile. The presence of strong variation is not only true for the dependent variable, but also for many covariates, some of which are displayed in figure 4.3. The monthly rent ranges from a mean of € 692.37 (80th-percentile: € 898.55) in Berlin to € 1,209.59 (80th-percentile: € 1,559.91) in Munich, which are at the same time the cities with the lowest and highest mean in purchasing power. With respect to the purchasing power, a huge variation between the cities can be found. While Berlin exhibits the lowest average value of € 35,272.76 per household, the average purchasing power in Munich is € 55,942.79. Dwellings offered in Berlin are on average the oldest dwellings within the sample. The lowest mean and median for building age is found for Munich, indicating high development activity in the more recent past. Living area spans from an average of 71.49 sqm in Hamburg to 79.19 sqm in Stuttgart. Within the cities, the standard deviation ranges from 30.13 sqm in Dusseldorf to 37.23 sqm in Frankfurt.

The aim of the study is to investigate the impact of changes in the explanatory variables on the time on market of rental dwellings segmented by their liquidity level. Hence, it is of particular interest, whether there are patterns in the dwelling characteristics, which might explain the affiliation to the respective liquidity quantile. The descriptive statistics show, that across all seven cities, the dwellings in the most liquid quantile, the 0.2-quantile (Q20), are on average the least expensive, the smallest, have the least number of rooms, are the oldest and least renovated ones. Furthermore, they are located in ZIP codes with the least amount of purchasing power but a relatively large number of households. The dwellings are also closest to the city center. However, the distribution of population density in a ZIP code area along the time on market quantiles exhibits variations across the cities. In contrast to the most liquid quantile, the dwellings assigned to the 0.8-quantile (Q80) display on average 33.71% higher rents, are 25.3% larger, are 12.13% younger, are located in ZIP codes with 2.67% higher purchasing power, have 3.46% less households in a ZIP code area and are 3.05% more densely populated.

Table 4.2: Descriptive statistics

Panel A: Berlin, ownership rate: 15.61%					
Variable	**Mean**	**Median**	**Std. Dev.**	**Q1**	**Q99**
Asking rent	692.37	549.94	470.12	233.42	2,500.35
Time on market	5.95	2.60	8.71	0.10	55.50
Living area	74.27	67.00	34.08	29.50	200.01
Age	65.57	59.00	39.20	0	117.00
Households in ZIP	12,008.03	11,997.00	3,645.74	3,890.00	20,434.00
Purchasing power of HHs in ZIP	35,272.76	34,352.80	4,954.13	27,548.31	52,669.38
Population density	3,898.94	3,542.00	2,398.80	61.00	7908.00
N=180,858					
Panel B: Hamburg, ownership rate: 21.14%					
Variable	**Mean**	**Median**	**Std. Dev.**	**Q1**	**Q99**
Asking rent	795.44	658.55	483.79	254.11	2,624.01
Time on market	4.98	2.10	7.51	0.10	40.20
Living area	71.49	65.65	30.30	26.00	180.00
Age	52.12	52.00	34.56	0	117.00
Households in ZIP	11,050.85	10,677.00	3,458.73	2,143.00	17,979.00
Purchasing power of HHs in ZIP	43,559.36	42,779.87	7,670.23	32,723.47	63,894.32
Population density	3,799.52	3,757.00	2,342.76	45.00	7,560.00
N=101,008					
Panel C: Munich, ownership rate: 25.23%					
Variable	**Mean**	**Median**	**Std. Dev.**	**Q1**	**Q99**
Asking rent	1,209.59	1,034.08	721.71	379.96	3,960.00
Time on market	3.84	1.70	6.30	0.10	33.60
Living area	77.66	72.00	37.20	23.00	209.33
Age	42.22	41.00	33.69	0	117.00
Households in ZIP	11,458.58	12,074.00	3,241.20	3,573.00	16,896.00
Purchasing power of HHs in ZIP	55,942.79	54,728.80	6,132.63	45,586.18	69,752.31
Population density	4,172.87	4,463.00	2,184.26	253.00	7,933.00
N=47,394					
Panel D: Cologne, ownership rate: 27.42%					
Variable	**Mean**	**Median**	**Std. Dev.**	**Q1**	**Q99**
Asking rent	720.55	639.6	369.88	250.00	2,039.92
Time on market	5.04	2.40	7.28	0.10	39.47
Living area	72.03	68.00	30.13	23.00	168.97
Age	45.24	46.00	29.60	1.00	117.00
Households in ZIP	13,452.47	13,521.00	3,594.13	6,176.00	20,561.00
Purchasing power of HHs in ZIP	45,466.38	44,370.20	5,748.89	34,685.48	58,827.02

Population density	3,675.18	3,390.00	2,243.07	395.00	7,598.00
N=47,527					
Panel E: Frankfurt, ownership rate: 20.67%					
Variable	**Mean**	**Median**	**Std. Dev.**	**Q1**	**Q99**
Asking rent	1,012.81	850.20	634.55	299.98	3,499.85
Time on market	5.89	2.70	8.29	0.10	45.90
Living area	79.03	72.00	37.23	23.00	208.00
Age	49.63	47.00	39.57	0	117.00
Households in ZIP	11,147.15	11,669.00	4,351.39	1,546.00	20,945.00
Purchasing power of HHs in ZIP	47,528.31	46,692.74	6,663.02	37,419.27	76,088.03
Population density	3,929.72	4,194.00	2,098.80	146.00	7,785.00
N=41,446					
Panel F: Stuttgart, ownership rate: 32.92%					
Variable	**Mean**	**Median**	**Std. Dev.**	**Q1**	**Q99**
Asking rent	910.71	775.20	501.26	270.98	2,749.82
Time on market	3.73	1.60	5.93	0.10	30.40
Living area	79.19	74.00	34.29	23.00	193.01
Age	50.25	48.00	34.85	0	117.00
Households in ZIP	10,385.10	10,927.00	3,098.97	1,104.00	15,899.00
Purchasing power of HHs in ZIP	47,058.05	46,374.77	4,440.57	40,041.62	61,972.64
Population density	3,506.76	3,353.00	1,766.72	254.00	7,404.00
N=17,967					
Panel G: Dusseldorf, ownership rate: 24.08%					
Variable	**Mean**	**Median**	**Std. Dev.**	**Q1**	**Q99**
Asking rent	762.86	630.00	486.69	240.00	2,579.00
Time on market	6.60	3.20	8.91	0.10	50.01
Living area	75.65	70.00	33.85	25.00	190.00
Age	52.18	54.00	29.78	1.00	117.00
Households in ZIP	9,725.86	9,703.00	3,021.54	2,721.00	15,045.00
Purchasing power of HHs in ZIP	47,869.27	46,140.99	5,851.67	40,382.03	65,472.45
Population density	3,999.35	3,913.00	2,370.53	23.00	7,906.00
N=45,996					

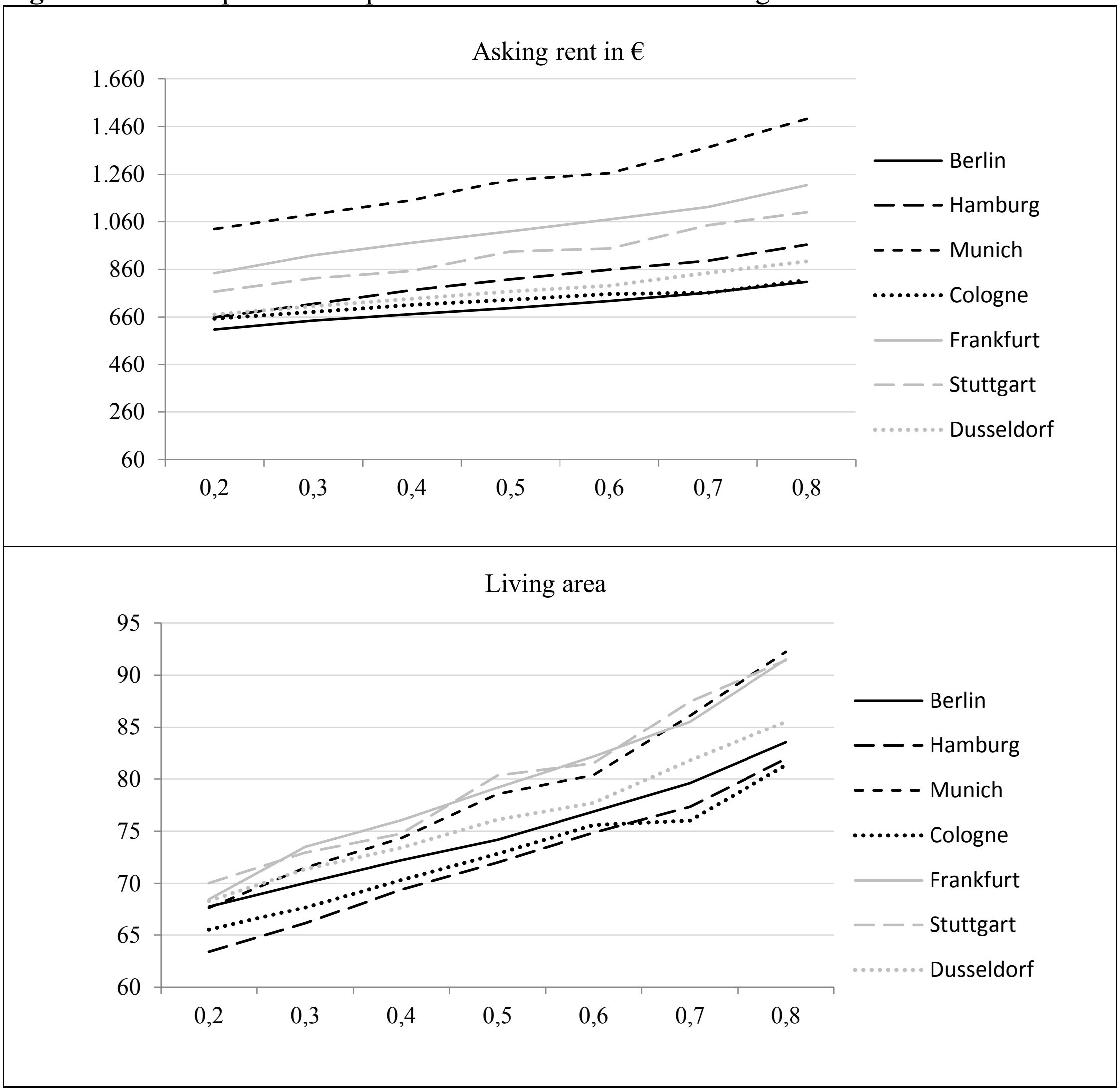
Notes*:* This table reports selected descriptive statistics and the number of observations for each of the seven cities. The ownership rates are as of 2013. The data consists of 482,196 observations on the residential rental market. The sample period is 2013 Q1 to 2017 Q4.

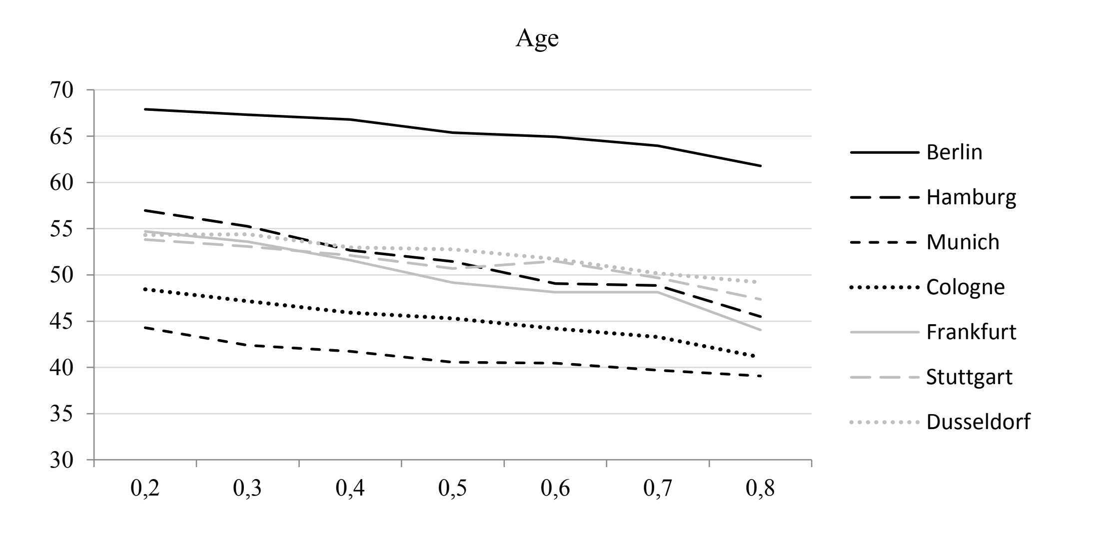
Table 4.3: Average time on market in weeks per quantile

	Q20	Q30	Q40	Median	Q60	Q70	Q80
Berlin	0.6	1.1	1.7	2.6	3.8	5.6	8.8
Hamburg	0.5	0.9	1.4	2.1	3.1	4.7	7.4
Munich	0.4	0.7	1.1	1.7	2.4	3.5	5.3
Cologne	0.6	1.1	1.6	2.4	3.4	4.8	7.4
Frankfurt	0.7	1.2	1.9	2.7	4.0	5.7	8.7
Stuttgart	0.3	0.7	1.1	1.6	2.3	3.5	5.4
Dusseldorf	0.9	1.5	2.2	3.2	4.6	6.4	10.0

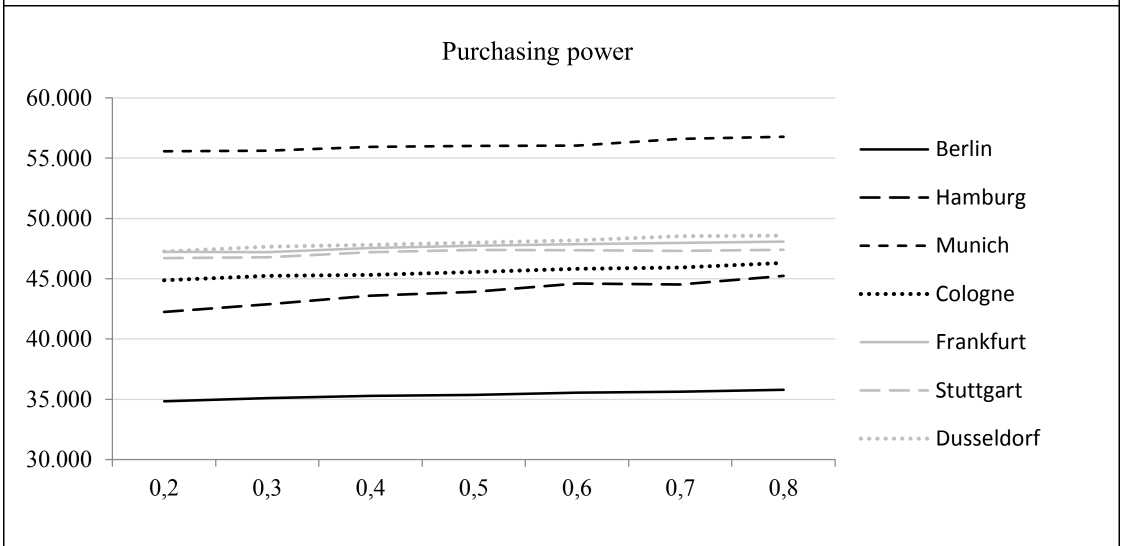
Notes*:* This table reports the average time on market in weeks for the seven cities in weeks. The data consists of 482,196 observations on the residential rental market. The sample period is 2013 Q1 to 2017 Q4.

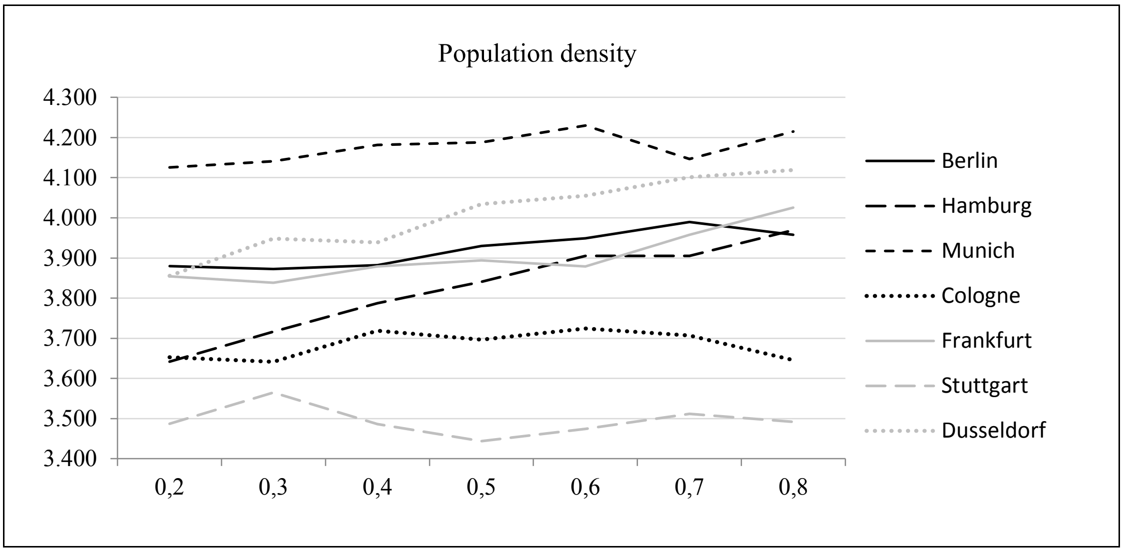
Figure 4.2: Descriptive development of selected covariates along the TOM distribution




Age
70
65
60
55
50
45
40
35
30
0,2
0,3
0,4
0,5
0,6
0,7
0,8
Berlin
Hamburg
Munich
Cologne
Frankfurt
Stuttgart
Dusseldorf


Households in ZIP
14.000
13.500
13.000
12.500
12.000
11.500
11.000
10.500
10.000
9.500
9.000
0,2
0,3
0,4
0,5
0,6
0,7
0,8
Berlin
Hamburg
Munich
Cologne
Frankfurt
Stuttgart
Dusseldorf


Purchasing power
60.000
55.000
50.000
45.000
40.000
35.000
30.000
0,2
0,3
0,4
0,5
0,6
0,7
0,8
Berlin
Hamburg
Munich
Cologne
Frankfurt
Stuttgart
Dusseldorf



Notes*:* These figures display descriptive statistics for selected variables across seven time on market quantiles for the seven largest German cities. The data consists of 482,196 observations on the residential rental market. The sample period is 2013 Q1 to 2017 Q4.

4.4 Results

4.4.1 Results of the Cox Survival Regression

In a first step, covariates boosting or limiting the time on market of rental dwellings on the housing markets of the seven largest German cities have been considered.

Table 4.4: Results of the Cox survival regression

Hazard rates and robust standard errors	Berlin	Hamburg	Munich	Cologne	Frankfurt	Stuttgart	Dusseldorf
Hedonic covariates – metric							
Log asking rent in €	0.285 0.052***	0.332 0.059***	0.396 0.058***	0.429 0.059***	0.378 0.095***	0.379 0.095***	0.410 0.058***
Log living area	1.329 0.053***	1.185 0.060***	1.035 0.061	0.994 0.062	1.026 0.079	1.151 0.102	1.039 0.070
Number of rooms	1.172 0.012***	1.165 0.012***	1.122 0.012***	1.134 0.015***	1.158 0.014***	1.096 0.019***	1.160 0.013***
Age	0.998 0.002	1.003 0.002	1.008 0.003***	1.007 0.003**	1.007 0.002***	1.002 0.003	1.003 0.003
Hedonic covariates – binary							
With bathtub	0.873 0.016***	0.964 0.013***	0.961 0.012***	0.970 0.018*	0.906 0.016***	0.939 0.020***	0.903 0.018***
With built-in kitchen	1.102 0.015***	0.964 0.018**	1.047 0.016***	1.120 0.015***	1.138 0.014***	1.093 0.033***	1.178 0.013***
With parking slot	0.917 0.020***	0.892 0.017***	0.936 0.016***	0.982 0.024	0.919 0.018***	0.988 0.030	0.989 0.020
With terrace	0.938 0.014***	0.989 0.017	0.933 0.016***	0.940 0.021***	0.922 0.025***	0.983 0.026	0.984 0.026
With balcony	1.002 0.014	1.047 0.017***	1.003 0.014	1.009 0.021	1.005 0.017	0.989 0.020	1.016 0.023
With lift	0.974 0.022	0.916 0.019***	1.035 0.018**	0.852 0.023***	0.955 0.022***	0.913 0.025***	0.942 0.020***
Newly built	0.830 0.025***	0.922 0.020***	0.839 0.021***	0.940 0.028**	0.921 0.021***	0.917 0.030***	0.922 0.026***
Refurbished	0.983 0.013	0.907 0.014***	0.967 0.014**	0.987 0.014	0.938 0.017***	0.960 0.021*	0.930 0.012***
Spatial variables							
Longitude	0.683 0.132***	0.597 0.236***	1.346 0.184	0.226 0.296***	1.310 0.426*	0.443 0.308***	0.075 0.404***
Latitude	0.679 0.262***	1.447 0.314	1.300 0.412	0.218 0.477***	1.824 0.499**	0.806 0.521	0.106 0.700***
Distance to ZIP centroid	0.961 0.030*	0.922 0.027***	0.973 0.029	0.976 0.027	0.903 0.046	0.976 0.033*	0.945 0.037
Distance to NUTS3 centroid	0.942 0.004***	0.952 0.006***	0.982 0.005***	0.949 0.010***	0.964 0.009***	0.975 0.008***	0.971 0.008***
Socioeconomic variables at ZIP level							
Log purchasing power	1.781 0.126***	1.281 0.126**	0.854 0.142	1.166 0.216	1.135 0.181	0.811 0.156	1.526 0.187**

Log number of households	1.009 0.038	1.016 0.059	0.994 0.040	0.944 0.074	1.057 0.027**	1.045 0.033	1.071 0.037*
Log population density	1.001 0.014	0.945 0.016***	0.969 0.014**	0.965 0.023	0.959 0.018**	0.983 0.018	0.991 0.014
Financial conditions at day of release							
Effective 10Y IR for housing loan	1.118 0.101	1.039 0.075	0.943 0.129	1.083 0.100	0.954 0.121	0.829 0.174	1.164 0.112
Fixed effects							
Construction dummies	Included						
Quarterly dummies	Included						
Intercept	Included						
Spatial adjusted (Win-Lei) standard-errors	Considered						
R^2-concordance	65.4%	66.7%	67.5%	64.3%	65.4%	66.7%	64.4%

Notes: This table displays the hazard ratios and robust standard errors of a Cox PHM of dwellings' time on market in weeks on hedonic, spatial and socioeconomic variables as well as the effective ten-year interest rate and various fixed effects. The Pseudo-R^2 based on Kendall's Tau measures the concordance between estimated survival time and the observed survival time for only the non-censored response sample. The data consists of 482,196 observations on the residential rental market. The sample period is 2013 Q1 to 2017 Q4. *Significant at the 10%-level; ** significant at the 5%-level; *** significant at the 1%-level.

The results of the Cox survival regressions are presented in table 4.4. Since the results are displayed as hazard rates, a rate larger than one decreases the time on market thus increases liquidity while a rate smaller than 1 decreases liquidity. The results show that an increase in the asking rent ceteris paribus leads to a longer time on market across all seven cities. The same direction of the effect was found by Cajias and Freudenreich (2018). This result is not very surprising, as an increase in asking rent is expected to have a negative implication on the liquidity of dwellings. Since the densely populated cities regularly show excess demand for housing, it is of particular interest to investigate the rental effect for individual liquidity quantiles, as the magnitude of these effects is supposed to differ widely along the distribution curve. An increase in living area increases liquidity in two out of seven cities, which are Berlin and Hamburg. Hence, a segmentation into liquidity submarkets might be useful to consider the impact and its significance of e.g. living area for each quantile, as it might be the heterogeneity within the cities leading to insignificant effects. The number of rooms in a dwelling has the same statistically significant positive effect on liquidity for all seven cities. Surprisingly, the marketing time is ceteris paribus shorter, the older the dwelling and is longer for newly built and refurbished ones. An increase in the distance to the NUTS3 centroid, which is used as a proxy for the city center, is associated with a longer time on market. The distance to the ZIP code centroid almost shows no statistically significant impact. The coefficients of the socioeconomic factors exhibit great variation between the cities, again emphasizing the heterogeneity within Germany and thus the importance of considering the seven cities separately. As would be expected, a higher purchasing power results in a shorter time on market in Berlin, Hamburg, and Dusseldorf, all else equal, whereas for the remaining cities, the effect is not statistically significant. While an increase in the number of households significantly reduces time on market in Frankfurt and Dusseldorf, the population density has a significant time on market prolonging effect in Hamburg, Munich, and Frankfurt. It is to notice, that neither socioeconomic variable has an impact on time on market in Cologne and Stuttgart. The effective 10-year interest rate for housing loans was included to account for market interactions between the property and the rental market. As for none of the seven cities a significant impact on time on market was found, it seems like interactions between the markets are negligible for the non-segmented sample. The Pseudo-R^2 based on Kendall's Tau which measures the concordance between estimated survival time and the observed survival time for the non-censored response sample, ranges from 64.3% to 67.5%. Those values are common in survival studies. At this point it is necessary to note, that these values cannot directly be interpreted as or compared to the usual R^2 calculated for OLS regressions.

4.4.2 Results of the Censored Quantile Regression

In a second step, the same regressors as for the Cox survival regressions are used to estimate censored quantile regressions, in order to get deeper insights into the liquidity segments. Therefore, for each city, the rental market was divided into seven time on market quantiles. The results for the covariates of interest are shown in figure 4.3 to 4.6. Each plot shows the development of a coefficient β_k^τ over the liquidity quantiles τ for each of the seven cities k. A positive and statistically significant coefficient increases time on market, thus decreases liquidity while a negative statistically significant coefficient has the opposite effect. The main effects divided into quarterly factors, hedonic characteristics, spatial gravity variables, and socioeconomic characteristics are reported in the following. The crucial point is, that contrary to the traditional regression models, the effect of a change in the covariate holds for the same quantile τ rather than across quantiles.

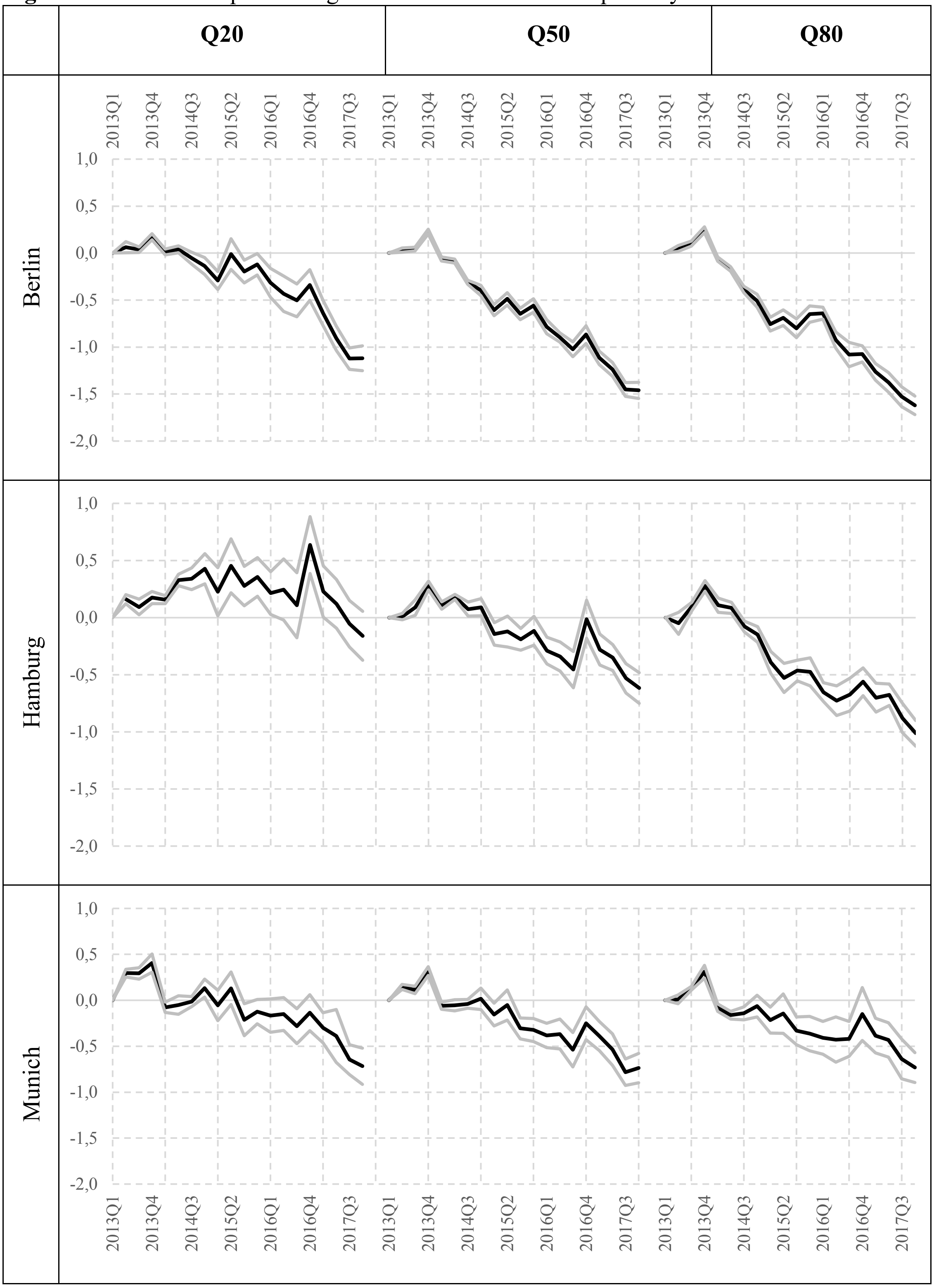
4.4.2.1 Quarterly Time Effects

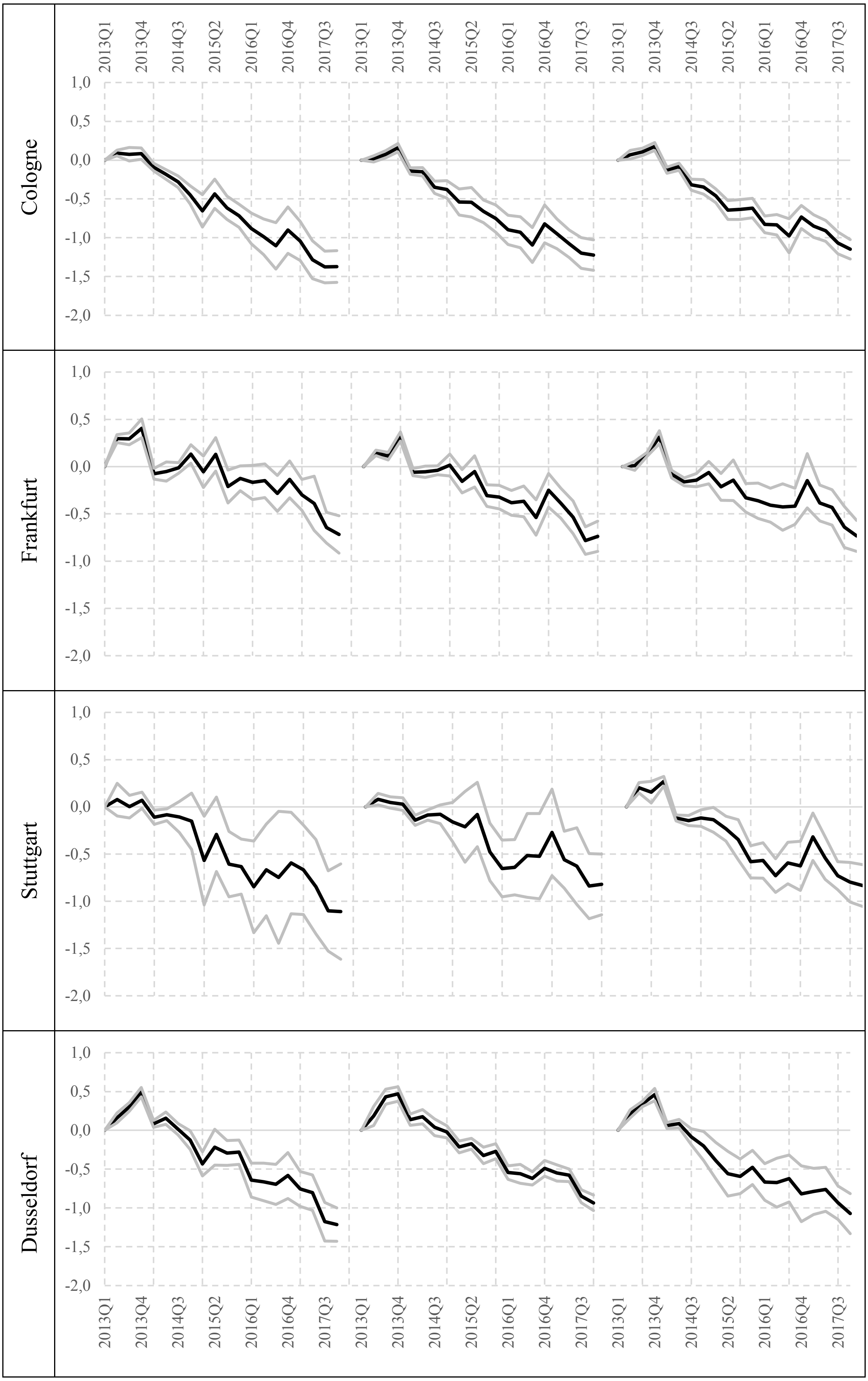
The considered period, the period between the first quarter of 2013 and the fourth quarter of 2017, is characterized by low interest rates, high migration to Germany, especially to the metropolises and additionally way too little housing supply in these cities. As a consequence, vacancy has mostly been diminishing and real estate prices as well as rents have been increasing. Despite rising construction activity in most cities, building completion was insufficient to meet demand, leading to excess demand. Time fixed effects as quarterly dummies have been included to capture the time trend of the time on market.[6] This time trend can be observed in figure 4.3 for quantiles

[6] Of those quarterly effects, 82.0% are significant at the 10% significance level while 74.5% are significant at the 5% significance level.

representing high, medium and low liquidity. The black line plots the coefficients of the individual quarterly dummies, the grey lines display the standard deviations. The base quarter is the first quarter of 2013, thus all changes are with respect to this basis. Comparing the highly liquid, average liquid and highly illiquid segments, the magnitude of the change in the time on market is in general very similar. However, the development of each city differs for the different liquidity quantiles. In the first quarter of the year, time on the market is usually relatively low for rental dwellings. At the 0.8-quantile, the time on market started to decrease in 2014 and kept this direction until the end of the observation period across all seven cities. This is also true at the median, with Hamburg as an exception, starting to decrease with the first quarter of 2015. For highly liquid dwellings, all cities besides Hamburg exhibit a declining development of the time on market starting in 2014. For highly liquid dwellings in Hamburg, the time on market was increasing relative to the base quarter until mid of 2017 and has only started to decrease in the last two observation periods. The increasing time on market compared to the base quarter at the beginning of the period under consideration was relatively strong in Dusseldorf across all quantiles. This might be due to the strong increase in construction completions in Dusseldorf from 2013 to 2014 and a rather high vacancy rate compared to the other top cities. Across all quantiles, the decline in time on market from about 2014 onwards, has been particularly strong in Cologne. A possible explanation might be the strong population growth, which the city experienced. At the median, Frankfurt exhibits a similar decrease. A reason for that might be the prospering economic condition in combination with surging employment rates and corresponding demand for space. Furthermore, building permits and building completions have been strongly diverging so that newly built living space is still scarce. In Berlin, the city with the lowest ownership rate, a huge and for some quarters even the strongest decrease in the time on market can be observed at the median and for very illiquid dwellings. After years of high unemployment and debt, Berlin is now flourishing, coming along with strong economic growth, a positive labor market dynamic and an ongoing population inflow. The strong increase in demand faces a massive shortage of dwellings due to insufficient construction completions and scarce availability of building land. For highly liquid dwellings in Berlin, time on market has been decreasing only moderately compared to the other cities, however with a relatively strong decrease starting in 2017. The reduction in time on market for highly and median liquid dwellings is weakest in Hamburg and is average for very illiquid dwellings. This moderate development can be explained by a very dynamic construction activity in the previous years and relatively small population growth. As a consequence, rents have only been increasing slightly. A city with marginal changes in the time on market across all quantiles is Munich. This city exhibits also the largest rental increase. In Munich, the number of inhabitants has been vastly rising, so that an enormous excess demand for rental dwellings has been emerging along with vacancy rates close to zero. Furthermore, purchasing power per person is highest in Munich and unemployment is lowest. Thus, the time on market development might seem somehow surprising. However, with a glance at the descriptive statistics in table 4.3, it might to some extent be due to the anyway relatively low levels of time on market in each of the respective quantiles compared to the other cities.

Figure 4.3: Estimated quantile regression coefficients of the quarterly dummies





Notes: The figures display the development of the coefficients β_k^τ of the quarterly dummy covariate for each of the seven cities and the respective confidence intervals over time. The impact of an individual coefficient is insignificant if the confidence interval includes zero. The data consists of 482,196 observations on the residential rental market. The sample period is 2013 Q1 to 2017 Q4.

4.4.2.2 Hedonic Characteristics

It is possible to cluster the hedonic variables in three groups based on their impact on time on market. Asking rent unsurprisingly has a consistently positive impact on time on market, suggesting that an increase in rent increases the time a dwelling is advertised on the market. The statistically significant coefficients of first occupancy and renovation as well as the equipment variables with bathtub and with car space also show positive effects. Living area, number of rooms and building age, besides the 0.8-quantile in Berlin, show the opposite impact for all statistically significant quantiles across the seven cities. For the other equipment features, the statistically significant coefficients show distinct effects, depending on the location of the dwelling.

For the covariate asking rent, the magnitude of the impact on time on market decreases with the level of illiquidity for all cities. This is probably due to the fact that the more illiquid quantiles are characterized by larger living areas, more rooms and are farther from the city center. With the population in the seven metropolises consisting to a great extent of young single households, these dwellings are less demanded and so the reaction to a rental price increase is less pronounced. In Munich and Stuttgart, the effect increases from the 0.2- to the 0.3-quantile but decreases afterwards. The impact of an increase in asking rent is weakest for dwellings located in Dusseldorf. The variation between the cities is smallest for very liquid quantiles, increases for average liquid quantiles and converges again towards the end of the distribution. While a ten percent increase in asking rent within the most liquid quantile results in a 8.5% to 10.8% higher time on market in Dusseldorf and Stuttgart. The spread decreases to 4.9% to 7.3% higher time on market in Dusseldorf and Berlin for the most illiquid quantile. The strongest impact for the more liquid quantiles is found for Stuttgart, but switches to Munich with the 70% percentile and to Berlin for the 80% percentile. This is not surprising, as rents in Munich as well as Stuttgart are on a very high level. Accordingly, further rental price increases strongly affect time on market. The effect in Berlin, however, is considerably weaker for all quantiles and only exceeds the impact in Munich and Stuttgart for the most illiquid dwellings.

The impact of an increase in living area as well as the number of rooms show the expected opposite pattern, as more space and a higher usability are positive factors for the marketability of a dwelling, all else equal. With growing illiquidity, the effect of living area, as well as of the number of rooms on the marketing time diminishes. Again, this might be due to the fact that in all seven cities small dwellings for single households are the most demanded ones. In consequence, the marketing time of larger, thus more illiquid dwellings, reacts less to changes in the size or the number of rooms. For the most liquid dwellings, a ten percent increase in living area decreases time on market by 2.2% in Hamburg to 3.9% in Stuttgart. The spread converges for highly illiquid dwellings and reveals a 0.1% increase in Cologne and a 1.6% increase in Stuttgart. For all quantiles, the strongest impact of an increase in living area is found for dwellings in Stuttgart. However, the impact of the number of rooms on time on market is lowest in Stuttgart.

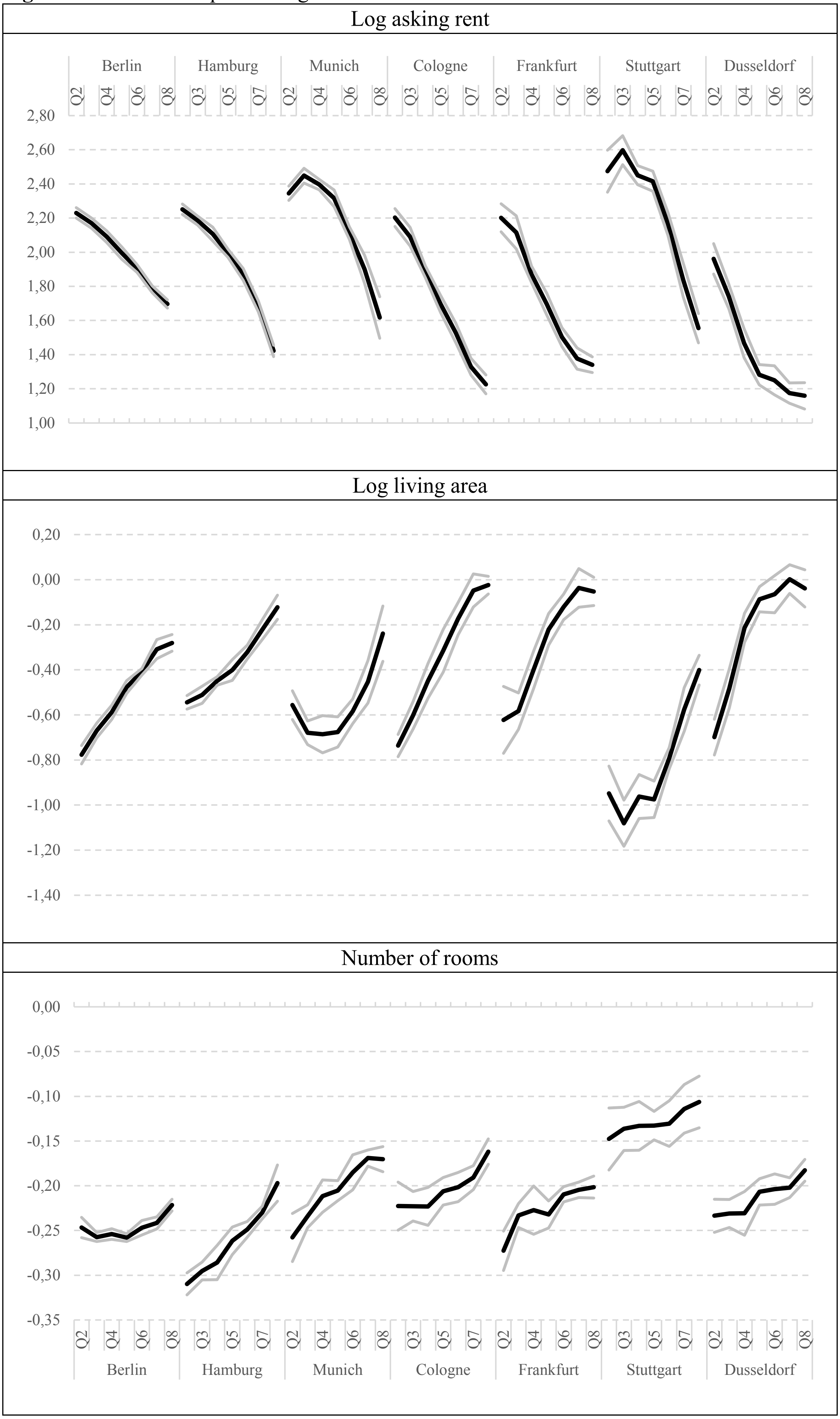
Thus relative to the other cities, people in Stuttgart prefer larger apartments with less rooms. Munich exhibits the second weakest effect for the most liquid dwellings but the second strongest effect for all other quantiles. As in Stuttgart, the number of rooms has a relatively low impact on the marketing time of dwellings in Munich. The weakest impact of an additional sqm of living space on marketing time is found for highly liquid dwellings in Hamburg but changes to Dusseldorf with the 0.3-quantile. For the most illiquid dwellings, the effects strongly converge for Dusseldorf, Cologne and Frankfurt. Hamburg and Berlin exhibit the strongest impact of a change in the number of rooms on time on market. When comparing the effect of an additional sqm with the effect of an additional room, the marketing time of dwellings in Stuttgart and Munich is more strongly affected by the living area whereas the influence of an additional room is low. However, in Hamburg the importance of an additional room is highest among the seven cities, whereas the effect of the living area is lowest for highly liquid dwellings. A similar picture holds true for Berlin. These findings emphasize the presence of heterogeneity between the cities and the importance of market segmentation.

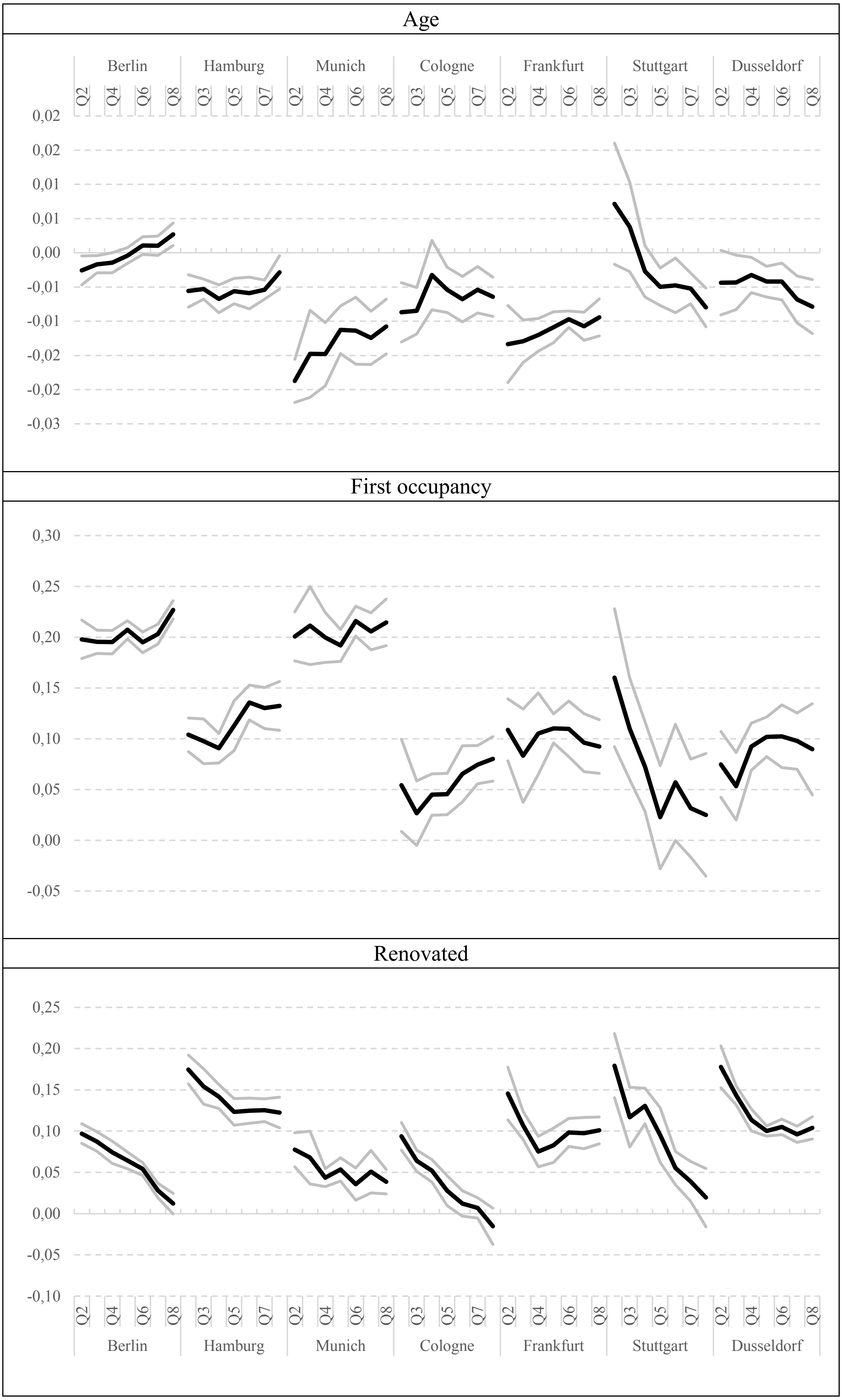
Unlike the impact of a change in other hedonic variables, the effect of a change in age does not have a consistent effect on time on market across all quantiles and cities, nor is it statistically significant for all of them. The impact is only consistently statistically significant and thus conclusive for Munich, Frankfurt, and Hamburg besides the 0.8-quantile. The corresponding dwellings exhibit a negative impact of an increase in age on time on market, meaning that older dwellings are rented out faster. Among these three cities, the impact is strongest for Munich. These findings show, that the proportionality assumption underlying the Cox Proportional Hazards Model cannot be verified for the covariate age for all cities.

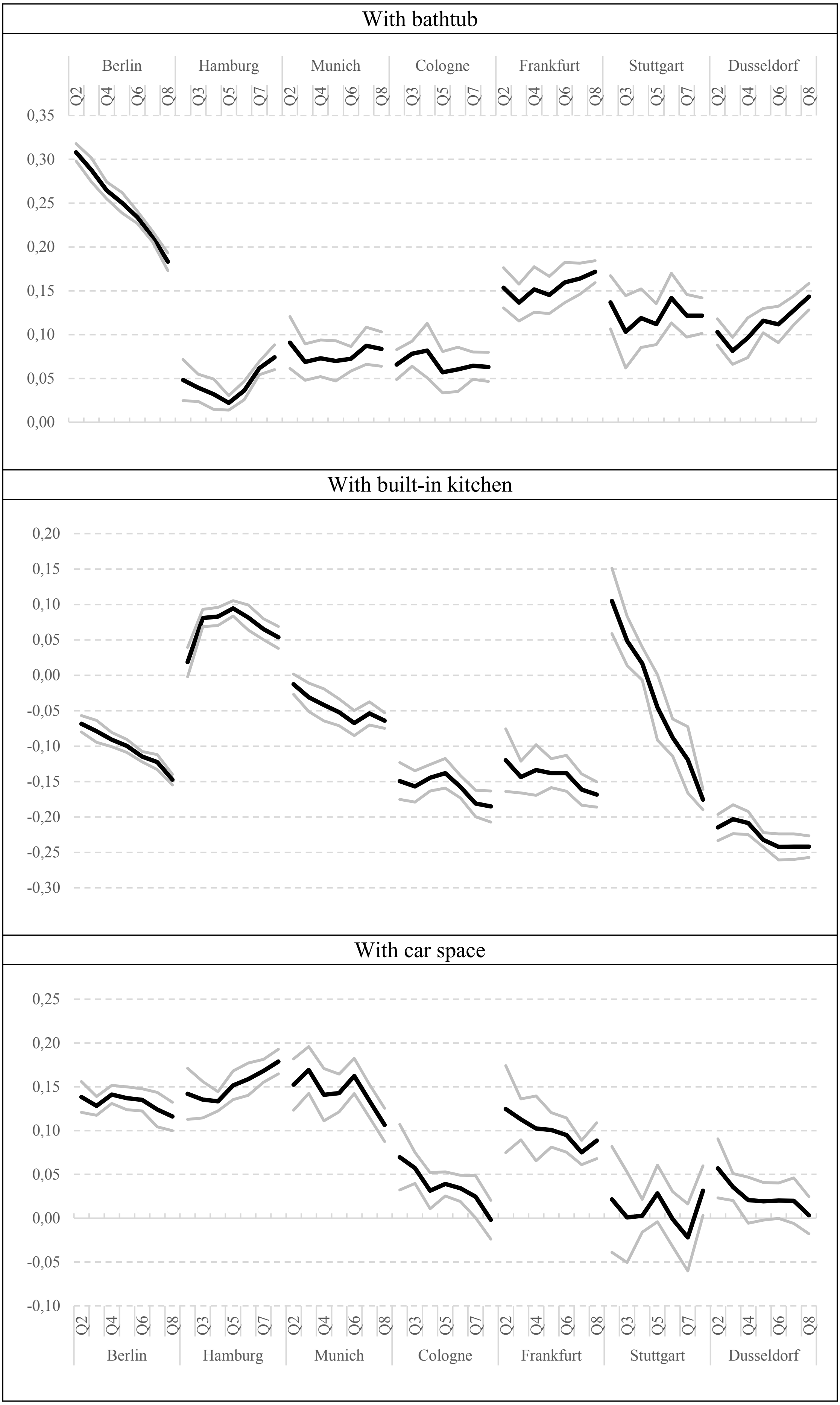
Whether a dwelling was offered for first occupancy or has been renovated has a prolonging impact on the marketing time for all cities across all quantiles. For renovated dwellings, the strongest effect on the time on market is again observed for the highly liquid quantiles. For first-time occupancy however, the effect is quite distinguishable across the cities. While the trend is increasing relatively flat for Berlin and Munich, Hamburg and Cologne exhibit a stronger increase. The strongly decreasing impact in Stuttgart is only significant for the first two quantiles. The effect of first-time occupancy is by far strongest in Berlin and Munich, renovation, however, plays a minor role in these cities. For Berlin, the positive, thus time on market prolonging effect of whether the dwelling is offered for fist occupancy, ranges from a 21.9% higher time on market for the most liquid quantile to a 25.5% higher time on market for the most illiquid quantile.

The importance of considering the seven cities separately is emphasized by the varying effects, a change in certain equipment characteristics has on time on market. A built-in kitchen reduces the time on market in six out of seven cities. It however has a time on market prolonging effect in Hamburg. An elevator significantly increases the marketing time in all cities besides Munich, where an elevator increases liquidity. Whether the dwelling offers a balcony is only statistically significant for all quantiles in Hamburg and consistently reduces the time on market. A time on market reducing effect is also visible for certain quantiles in Cologne and Frankfurt. In Munich and Stuttgart, a balcony prolongs the marketing time for the 0.3- to 0.5-quantile and the 0.3- and 0.4-quantile, respectively. These findings underscore the importance of segmenting the market into different quantiles when analyzing the rental market in more detail, as for all covariates and cities with a changing direction of the impact, the proportionality assumption underlying the Cox PHM cannot be verified.

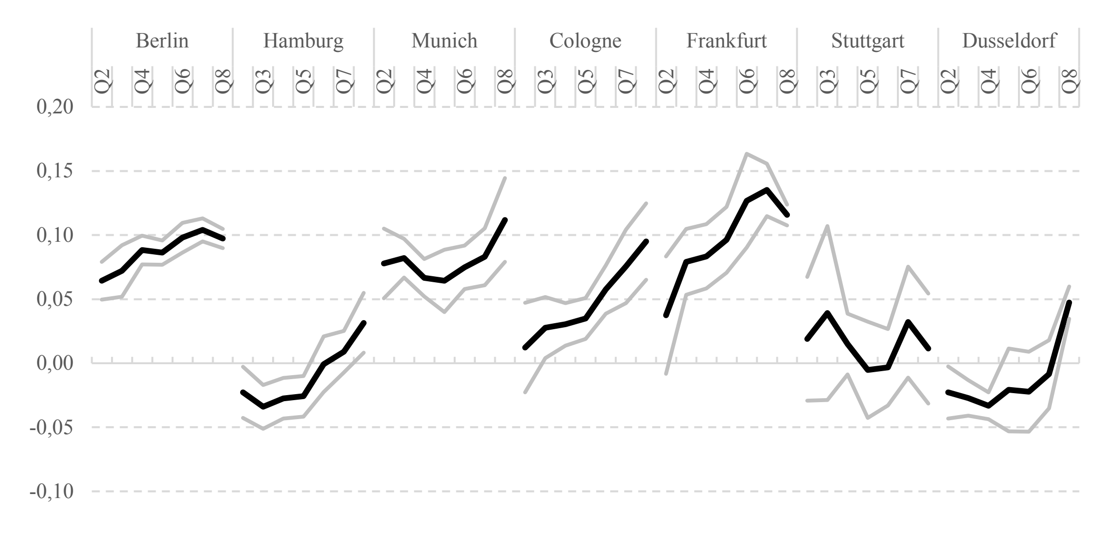
Figure 4.4: Estimated quantile regression coefficient of the hedonic variables



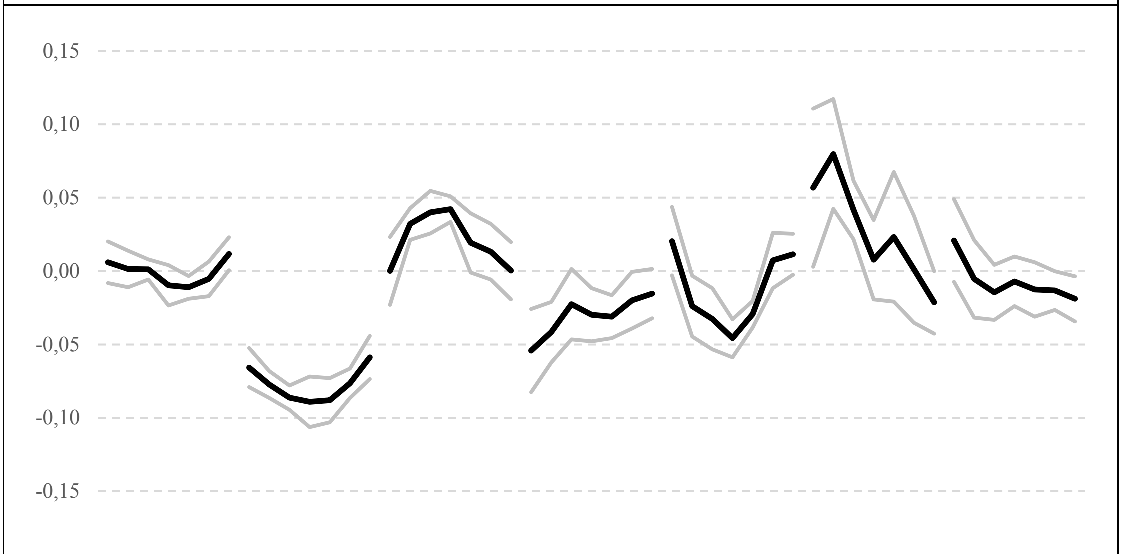

Age
Berlin
Hamburg
Munich
Cologne
Frankfurt
Stuttgart
Dusseldorf
Q2
Q4
Q6
Q8
Q3
Q5
Q7
0,02
0,02
0,01
0,01
0,00
-0,01
-0,01
-0,02
-0,02
-0,03
First occupancy
0,30
0,25
0,20
0,15
0,10
0,05
0,00
-0,05
Renovated
0,25
0,20
0,15
0,10
0,05
0,00
-0,05
-0,10


With bathtub
Berlin
Hamburg
Munich
Cologne
Frankfurt
Stuttgart
Dusseldorf
Q2
Q4
Q6
Q8
Q3
Q5
Q7
0,35
0,30
0,25
0,20
0,15
0,10
0,05
0,00
With built-in kitchen
0,20
0,15
0,10
0,05
0,00
-0,05
-0,10
-0,15
-0,20
-0,25
-0,30
With car space
0,25
0,20
0,15
0,10
0,05
0,00
-0,05
-0,10

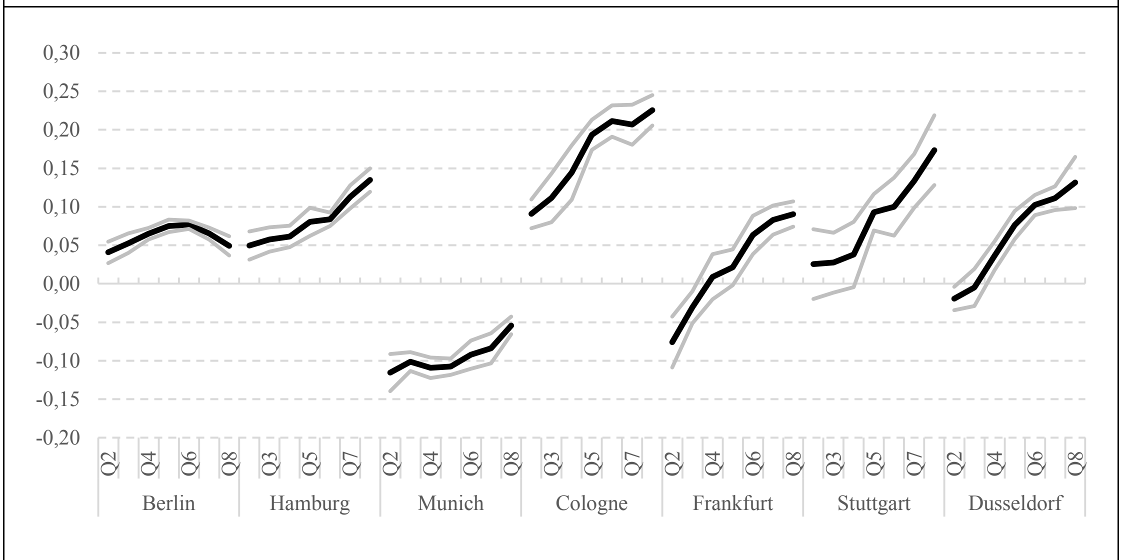
With terrace



With balcony



With elevator



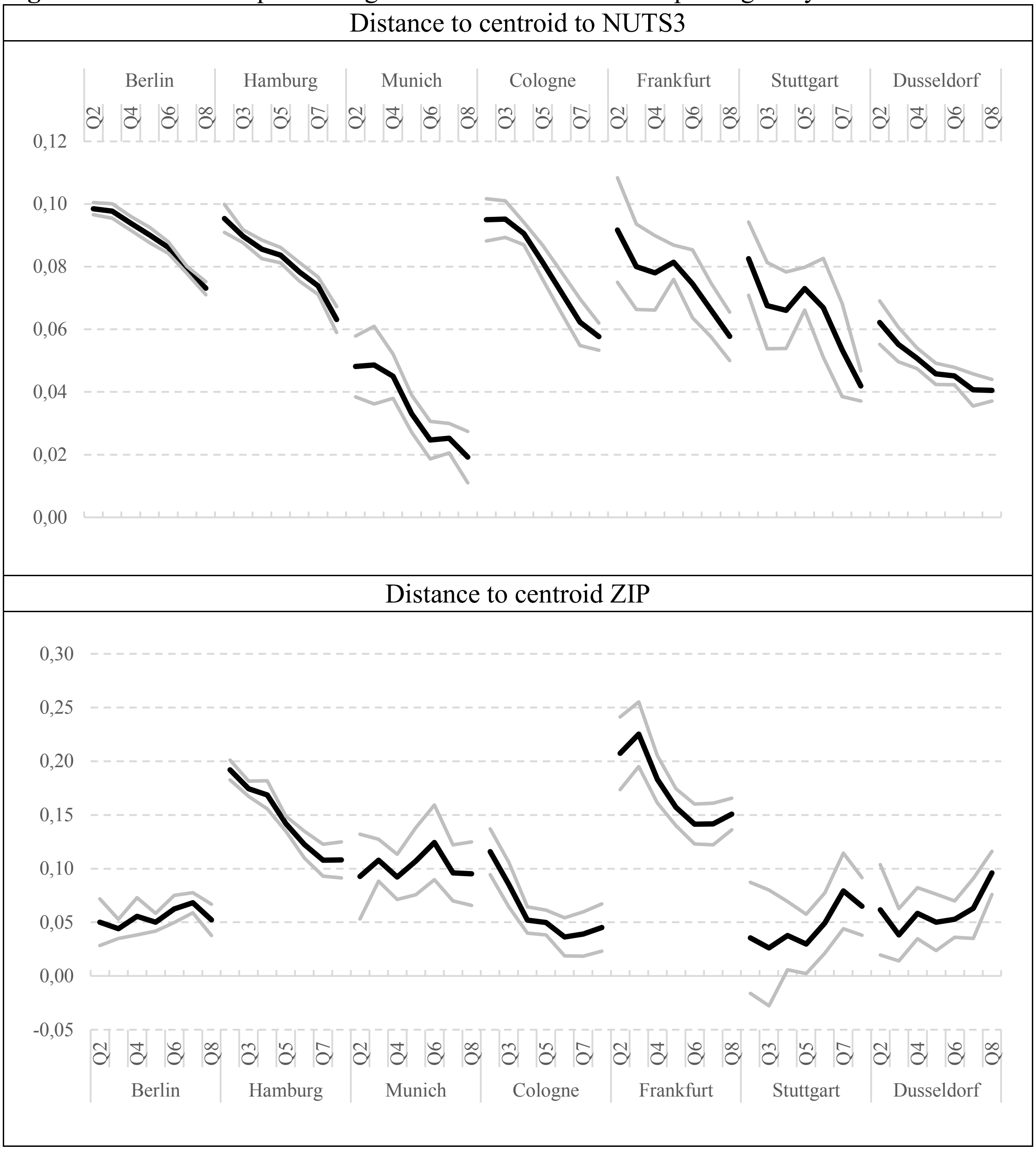
Notes*:* The figures display the development of the coefficients β_k^τ of the hedonic covariates across the liquidity quantiles τ for each of the seven cities and the respective confidence intervals. The impact of an individual coefficient is insignificant if the confidence interval includes zero. The data consists of 482,196 observations on the residential rental market. The sample period is 2013 Q1 to 2017 Q4.

4.4.2.3 Spatial Gravity Variables

Considering the spatial variables, the distance to the NUTS3 center is of particular interest, since the study analyzes the largest seven German cities geographically restricted by its NUTS3 boundaries. Therefore, the distance from an individual dwelling to the center of the NUTS3 region can be interpreted as its distance to the geographical city center.

Not surprisingly, a higher distance to the center is extending the time a rental dwelling is listed on the market. The effect is strongest for the highly liquid quantiles and is getting weaker for more illiquid quantiles. A possible explanation might be the fact, that the most liquid dwellings are mostly located in the city centers, thus moving away from the center has the strongest effect. In contrast, the more illiquid dwellings are located at the outskirts where one kilometer closer to or further from the center does merely play a role. Across the whole distribution, the effect is highest in Berlin and lowest in Munich. While for the more liquid half of the quantiles, distance exhibits a meaningful effect on time on market in Cologne, it fades below the impact in Hamburg and Frankfurt for the more illiquid quantiles. In Berlin, distance to the city center seems to be of high importance with respect to the marketing time. With 8.44 kilometers to the centroid of the NUTS3 region, the rental dwellings in Berlin display on average the highest distance to the approximated city center. The dwellings assigned to the most illiquid quantile are on average about 7% further from the city center as the dwellings in the 0.2-quantile. Apparently, people in widespread Berlin have higher preferences living closer to the city center compared to the other cities, especially Munich. A reason for that might be the allocation of popular residential areas all over the metropolitan area as well as staggering asking rents in the city center. The distance to the ZIP code center is of less interest and has no explanatory power.

Figure 4.5: Estimated quantile regression coefficients of the spatial gravity variables



Notes: The figures display the development of the coefficients β_k^τ of the spatial gravity covariates across the liquidity quantiles τ for each of the seven cities and the respective confidence intervals. The impact of an individual coefficient is insignificant if the confidence interval includes zero. The data consists of 482,196 observations on the residential rental market. The sample period is 2013 Q1 to 2017 Q4.

4.4.2.4 Socioeconomic Characteristics

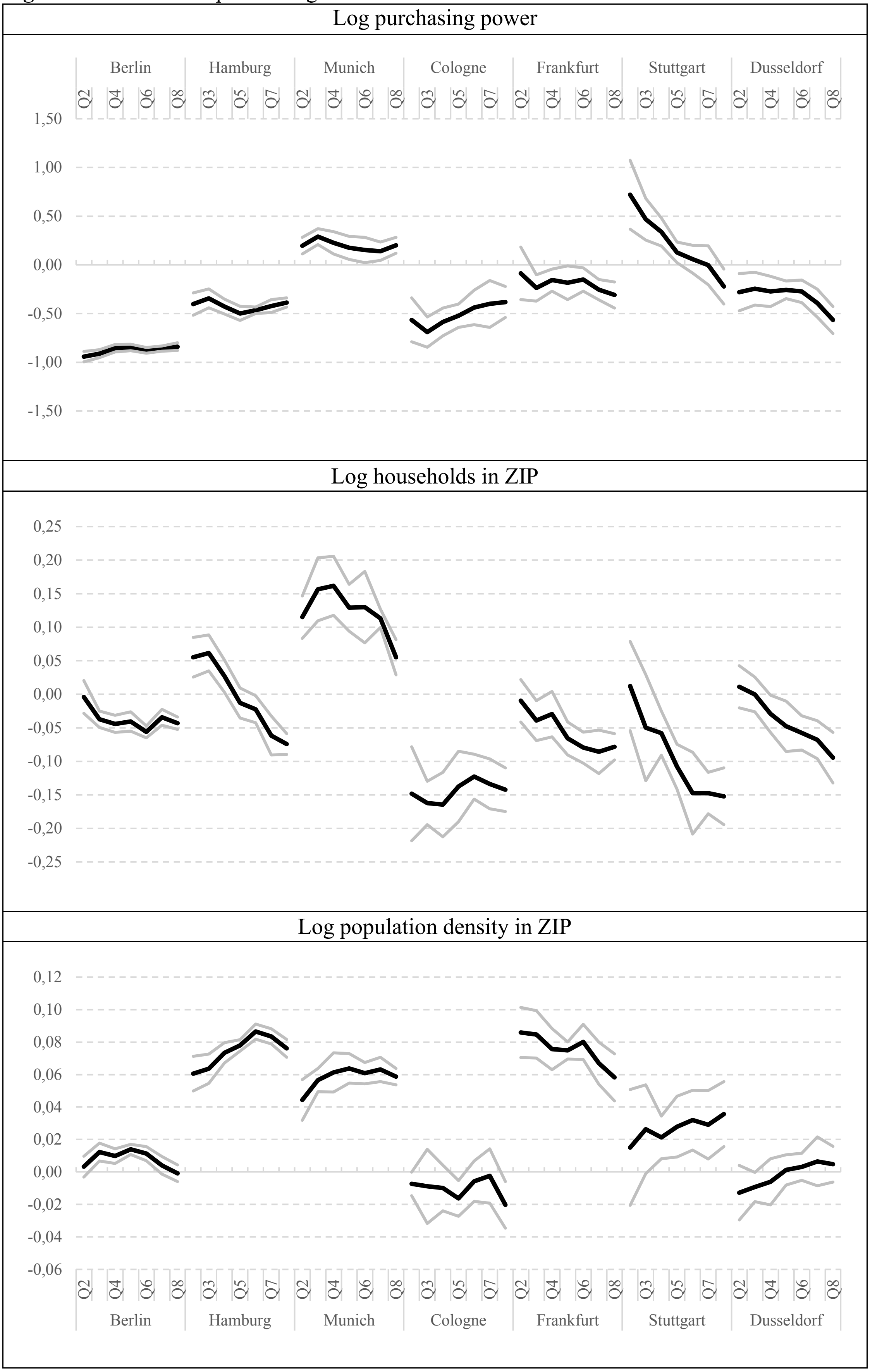
The socioeconomic factors purchasing power and number of households per ZIP code area exhibit a negative impact on the time on market in most cities. This result is generally common across all cities and all liquidity quantiles. The significant effects of the population density are positive, thus time on market prolonging.

From the three analyzed socioeconomic variables, purchasing power has the strongest impact on liquidity. In five of the seven cities, an increase in purchasing power leads to more demand on the real estate market, thus decreases the time a dwelling stays on the market. This development is relatively flat across all quantiles. The impact of purchasing power is strongest in Berlin across all quantiles. Berlin is by far the city with the lowest average purchasing power. The strong reduction in marketing time caused by an increase in purchasing power might indicate, that people would be willing to spend more of the additional income on housing than in other cities. The contrary might be true for Munich and Stuttgart, where a higher purchasing power leads to longer time on the market. However, the coefficients are only statistically significant for the very liquid half of the dwellings and for the 0.8-quantile in Munich. While in Berlin, Hamburg and Cologne the impact of purchasing power decreases with growing illiquidity, the effect on time on market increases in Frankfurt and Dusseldorf. As can be seen in figure 4.6, high illiquidity can be attributed to higher levels of purchasing power. This might indicate that in Frankfurt and Dusseldorf, richer households spend more time for the search and matching process, whereas in Berlin, Hamburg and Cologne, richer households match faster when their income increases. Comparing the results of the CQR to the results of the Cox survival regressions, where the coefficient of purchasing power was only significant for three of the seven cities, highlights the huge heterogeneity within each city. Hence, it again emphasizes the importance of segmenting the market, as the dwellings exhibit substantial differences depending on their level of liquidity. In addition, the proportionality assumption cannot be verified for Stuttgart and Frankfurt, again highlighting the importance of a detailed analysis.

The effect of the number of households in a ZIP code area on time on market shows a similar picture. The impact of an additional household has a quite distinguishable effect on the time on market across the cities. Berlin, Cologne, Frankfurt, Stuttgart and Dusseldorf show the expected demand effect, whereas the impact is time on market prolonging in Munich. In Hamburg, the impact on time on market even switches from positive to negative, but is only significant for the two most liquid and the two most illiquid quantiles. Only for dwellings in Munich and Cologne, the direction of the impact is consistent. Hence, the proportionality assumption is violated for the other cities.

On the other side, an increase in population density significantly leads to a longer marketing time for Hamburg, Munich, Frankfurt and Stuttgart. While the impact increases with growing illiquidity for Hamburg and Munich, a decreasing trend is visible for Frankfurt. For the other cities, the effect is mainly insignificant.

Figure 4.6: Estimated quantile regression coefficients of the socioeconomic variables



Notes*:* The figures display the development of the coefficients β_k^τ of the socioeconomic covariates across the liquidity quantiles τ for each of the seven cities and the respective confidence intervals. The impact of an individual coefficient is insignificant if the confidence interval includes zero. The data consists of 482,196 observations on the residential rental market. The sample period is 2013 Q1 to 2017 Q4.

4.5 Conclusion

Academic research on the factors affecting the liquidity of dwellings on the rental market is difficult to find and for most countries not even existent. On the other hand, the enormous improvement in computational power and the possibility to gather and store large amounts of data allows the application of advanced econometric methods in the field. To the best of the authors' knowledge, this study is the first to apply a Censored Quantile Regression (CQR) in order to explore the determinants of liquidity with an extensive data set, comprising almost 500,000 observations across the top 7 German residential markets. As the censored quantile regression is able to identify the factors that drive time on market for dwellings allocated to low-, medium- and high-liquidity quantiles, the study is able to confirm that liquidity behaves differently across those segments, as measured by the elasticities of covariates within the liquidity quantiles. This distinct behavior implies, that the utility functions of landlords and tenants of residential dwellings behave differently when it comes to letting "stars-dwellings" and "poor-dogs-dwellings".

Apparently, the hedonic characteristics and in particular asking rent, the number of rooms, and whether the dwelling is offered for first occupancy or newly renovated affect liquidity the most. Another important factor is the distance to the respective geographical city center. Although the socioeconomic variables play only a minor role, among them purchasing power exhibits the largest impact. Across the cities, asking rent and living area have the strongest effect on time on market for the highly liquid dwellings in Munich and Stuttgart. The lowest impact of asking rent is found for dwellings in Dusseldorf. Living area shows the weakest effect for the most liquid dwellings in Hamburg. The number of rooms on the other hand, exhibits the highest effect for dwellings in Hamburg until the 0.6-quantile. The weakest impact of an additional room on liquidity is found for Stuttgart. Age and first occupancy have the strongest impact on time on market for dwellings in Munich.

The major findings can be summarized as follows. For many covariates consistent signs of the regression coefficients were found across the quantiles of the time on market distribution. However, for some covariates in individual cities, the impact of a change in the explanatory variables differs in direction across the liquidity quantiles. Hence, the proportionality assumption underlying the Cox PHM is violated for those covariates in the associated cities. In addition, the impact of a change in the explanatory variables differs in magnitude and significance across the liquidity quantiles. In contrast to the traditional Cox hazard regression, the Censored Quantile Regression properly detects these differences. Thus, the model is better suited to explore the heterogeneity of individual dwellings within a certain market. Furthermore it is found, that the magnitude, the significance, and the direction of the impact of the covariates on time on market is quite different across the cities. These findings emphasize the importance of market segmentation for a more detailed analysis and understanding of the rental real estate market or the real estate market as a whole.

With this detailed market assessment, buyers of dwellings should be able to infer how fast they will be able to let them, or which actions to take in order to increase the marketability. The variation in the impact of individual covariates on time on market across the liquidity quantiles and across the cities reveals the very distinct market characteristics in terms of marketability and location. This finding points to the need for very granular policy measures in order to control the market in an effective manner.

While the study uses the rental market of the largest seven German cities, it is of course possible to adapt the methodology to more cities or conurbations and other international real estate markets in order to examine the individual liquidity or time on market quantiles. Further research might consider the expansion of the censored quantile regression by segmenting the spatial component or including the spatial lags in the error component. In addition, a counterfactual decomposition could reveal whether the impact on time on market is attributable to a pure change in the characteristics of the dwellings or a shift in the assessment of the characteristics.

4.6 References

an de Meulen, P. and Mitze, T. (2014), "Exploring the spatial variation in quality-adjusted rental prices and identifying hot spots in Berlin's residential property market", *Regional Studies, Regional Science*, Vol. 1 (1), pp. 310-328.

Allen, M.T., Rutherford, R.C., and Thomson, T.A. (2009), "Residential asking rents and time on the market", *The Journal of Real Estate Finance and Economics*, Vol. 38 (4), pp. 351-365.

Bassett Jr., G.W. and Hsiu-Lang, C. (2002), "Portfolio style: Return- based attribution using quantile regression", In: *Economic Applications of Quantile* Regression, Studies in Empirical Economics, Fitzenberger, B., Koenker, R., and Machado, J.A.F. (eds.). Heidelberg: Physica, pp. 293-305.

Belkin, J., Hempel D.J., and McLeavey, D.W. (1976), "An empirical study of time on market using multidimensional segmentation of housing markets", *Real Estate Economics*, Vol. 4 (2), pp. 57-75.

Benefield, J.D. and Hardin, W.G. (2015), "Does time-on-market measurement matter?", *The Journal of Real Estate Finance and Economics*, Vol. 50 (1), pp. 52-73.

Bentzien, V., Rottke N., and Zietz, J. (2012), "Affordability and Germany's low homeownership rate", *International Journal of Housing Markets and Analysis*, Vol. 5 (3), pp. 289-312.

Beyerlein, A., Toschke, A.M., and von Kries, R. (2008), "Breastfeeding and childhood obesity: shift of the entire BMI distribution or only the upper parts?", *Obesity*, Vol. 16 (12), pp. 2730-2733.

Buchinsky, M. (1994), "Changes in the US wage structure 1963-1987: Application of quantile regression", *Econometrica*, Vol. 62 (2), pp. 405-458.

Buchinsky, M. (1995), "Quantile regression, Box-Cox transformation model, and the US wage structure, 1963–1987", *Journal of Econometrics*, Vol. 65 (1), pp. 109-154.

Cajias, M. and Freudenreich, P. (2018), "Exploring the determinants of liquidity with big data - market heterogeneity in German markets", *Journal of Property Investment and Finance*, Vol. 36 (1), pp. 3-18.

Chaudhuri, P., Doksum, K., and Samarov, A. (1997), "On average derivative quantile regression", *The Annals of Statistics*, Vol. 25 (2), pp. 715-744.

Cole, T.J. and Green, P.J. (1992), "Smoothing reference centile curves: the LMS method and penalized likelihood", *Statistics in Medicine*, Vol. 11 (10), pp. 1305-1319.

Cox, D.R. (1972), "Regression models and life-tables", *Journal of the Royal Statistical Society: Series B (Methodological)*, Vol. 34 (2), pp. 187-220.

Crowley, J. and Hu, M. (1977), "Covariance analysis of heart transplant survival data", *Journal of the American Statistical Association*, Vol. 72 (357), pp. 27-36.

Doksum, K.A. and Gasko, M. (1990), "On a correspondence between models in binary regression analysis and in survival analysis", *International Statistical Review/Revue Internationale de Statistique*, Vol. 58 (3), pp. 243-252.

Farmer, M. and Lipscomb, C. (2010), "Using quantile regression in hedonic analysis to reveal submarket competition", *Journal of Real Estate Research*, Vol. 32 (4), pp. 435-460.

Federal Statistical Office (2018), "Baufertigstellungen von Wohn- und Nichtwohngebaeuden unter anderem nach Bauherren – Lange Reihen", available at: https://www.destatis.de/DE/Publikationen/Thematisch/Bauen/BautaetigkeitWohnungsbau/BaufertigstellungenBauherrenPDF_5311201.pdf?__blob=publicationFile (accessed: August 8, 2018).

Fitzenberger, B. and Wilke, R.A. (2010), "New insights on unemployment duration and post unemployment earnings in Germany", *Oxford Bulletin of Economics and Statistics*, Vol. 72 (6), pp. 794-826.

Goodman, A.C. and Thibodeau, T.G. (2007), "The spatial proximity of metropolitan area housing submarkets", *Real Estate Economics*, Vol. 35 (2), pp. 209-232.

Guasch, J.L. and Marshall R.C. (1985), "An analysis of vacancy patterns in the rental housing market", *Journal of Urban Economics*, Vol. 17 (2), pp. 208-229.

Guntermann, K. and Norbin, S. (1987), "Explaining the variability of apartment rents", *Real Estate Economics*, Vol. 15 (4), pp. 321-340.

Harder, T., Bergmann, R., Kallischnigg, G., and Plagemann, A. (2005), "Duration of breastfeeding and risk of overweight: a meta-analysis", *American Journal of Epidemiology*, Vol. 162 (5), pp. 397-403.

Held, T. and Waltersbacher, M. (2015), "Wohnungsmarktprognose 2030", BBSR-Analysen KOMPAKT, Nr. 7, Bonn.

Hendricks, W. and Koenker, R. (1992), "Hierarchical spline models for conditional quantiles and the demand for electricity", *Journal of the American Statistical Association*, Vol. 87 (417), pp. 58-68.

Horowitz, J. L. and Neumann, G. R. (1987), "Semiparametric estimation of employment duration models", *Econometric Reviews*, Vol. 6 (1), pp. 5-40.

Kalra, R. and Chan, K.C. (1994), "Censored sample bias, macroeconomic factors, and time on market of residential housing", *Journal of Real Estate Research*, Vol. 9 (2), pp. 253-262.

Kang, H. and Gardner, M. (1989), "Selling price and marketing time in the residential real estate market", *Journal of Real Estate Research*, Vol. 4 (1), pp. 21-35.

Koenker, R. (2008), "Censored quantile regression redux", *Journal of Statistical Software*, Vol. 27 (6), pp. 1-25.

Koenker, R. and Bassett Jr., G. (1978), "Regression quantiles", *Econometrica,* Vol. 46 (1), pp. 33-50.

Koenker, R. and Bilias, Y. (2002), "Quantile regression for duration data: A reappraisal of the Pennsylvania reemployment bonus experiments", In: *Economic Applications of Quantile* Regression, Studies in Empirical Economics, Fitzenberger, B., Koenker, R., and Machado, J. A. F. (eds.). Heidelberg: Physica, pp. 199-220.

Koenker, R. and Geling, O. (2001), "Reappraising medfly longevity: a quantile regression survival analysis", *Journal of the American Statistical Association*, Vol. 96 (454), pp. 458-468.

Lerbs, O. and Oberst, A. (2014), "Explaining the spatial variation in homeownership rates: Results for German regions", *Regional Studies*, Vol. 48 (5), pp. 844-865.

Liao, W.C. and Wang, X. (2012), "Hedonic house prices and spatial quantile regression", *Journal of Housing Economics*, Vol. 21 (1), pp. 16-27.

Luedemann, E., Wilke, R.A., and Zhang, X. (2006), "Censored quantile regressions and the length of unemployment periods in West Germany", *Empirical Economics*, Vol. 31 (4), pp. 1003-1024.

Lyons, R.C. (2013), "Price signals in illiquid markets – The case of residential properties in Ireland, 2006-2012", Trinity Economics Papers tep0616, Trinity College Dublin, Department of Economics.

Machado, J. and Mata J. (2000), "Box-Cox quantile regression and the distribution of firm sizes", *Journal of Applied Econometrics*, Vol. 15 (3), pp. 253-274.

Mak, S., Choy, L., and Ho, W. (2010), "Quantile regression estimates of Hong Kong real estate prices", *Urban Studies*, Vol. 47 (11), pp. 2461-2472.

Malpezzi, S. (2003), "Hedonic pricing models: A selective and applied review", In: *Housing Economics and Public Policy: Essays in honor of Duncan Maclennan*, O'Sullivan, T. and Gibb, K. (eds.). Oxford: Blackwell, pp. 67-89.

McGreal, S., Adair, A., Brown, L., and Webb, J.R. (2009), "Pricing and time on the market for residential properties in a major UK city", *Journal of Real Estate Research*, Vol. 31 (2), pp. 201-234.

Ong, S.E. and Koh, Y.C. (2000), "Time on-market and price trade-offs in high-rise housing sub-markets", *Urban Studies*, Vol. 37 (11), pp. 2057-2071.

Owen, C.G., Martin, R.M., Whincup, P.H., Smith, G.D., and Cook, D.G. (2005), "Effect of infant feeding on the risk of obesity across the life course: a quantitative review of published evidence", *Pediatrics*, Vol. 115 (5), pp. 1367-1377.

Pandey, G.R. and Nguyen, V.T.V. (1999), “A comparative study of regression based methods in regional flood frequency analysis”, *Journal of Hydrology*, Vol. 225 (1-2), pp. 92-101.

Peng, L. and Huang, Y. (2008), “Survival analysis with quantile regression models”, *Journal of the American Statistical Association*, Vol. 103 (482), pp. 637- 649.

Portnoy, S. (2003), “Censored regression quantiles”, *Journal of the American Statistical Association*, Vol. 98 (464), pp. 1001-1012.

Powell, J.L. (1984), “Least absolute deviations estimation for the censored regression model”, *Journal of Econometrics*, Vol. 25 (3), pp. 303-325.

Powell, J.L. (1986), “Censored regression quantiles”, *Journal of Econometrics*, Vol. 32 (1), pp. 143-155.

Reisenbichler, A. (2016), “A rocky path to homeownership: Why Germany eliminated large-scale subsidies for homeowners”, *Cityscape: Ajournal of policy development and research*, Vol. 18 (3), pp. 283-290.

Rose, S.J. (1992), “Social Stratification in the United States: The American Profile Poster”, New York: New Press.

Rosen, S. (1974), “Hedonic prices and implicit markets; product differentiation in pure competition”, *Journal of Political Economy*, Vol. 82, pp. 34-55.

Royston, P. and Altman, D.G. (1994), “Regression using fractional polynomials of continuous covariates: parsimonious parametric modelling”, *Applied Statistics*, Vol. 43 (3), pp. 429-467.

Shimizu, C., Nishimura, K.G., and Watanabe, T. (2012), “House prices from magazines, realtors, and the land registry”, In: *Property Market and Financial Stability*, BIS Papers No. 64, Bank of International Settlements, pp. 29-38.

Sirmans, G.S., and Benjamin, J.D. (1991), “Determinants of market rent”, *Journal of Real Estate Research*, Vol. 6 (3), pp. 357-379.

Sirmans, G.S., Sirmans C.F., and Benjamin J.D. (1989), “Determining apartment rent: The valuation of amenities, services, and external factors”, *Journal of Real Estate Research*, Vol. 4 (2), pp. 33-44.

Schmillen, A. and Moeller, J. (2010), “Determinants of lifetime unemployment: a micro data analysis with censored quantile regressions”, In: *IAB-Discussion Paper, Institut fuer Arbeitsmarkt- und Berufsforschung (IAB)*, 3.

Taylor, J.W. (1999), “A quantile regression approach to estimating the distribution of multiperiod returns”, *The Journal of Derivatives*, Vol. 7 (1), pp. 64-78.

Thomschke, L. (2015), “Changes in the distribution of rental prices in Berlin”, *Regional Science and Urban Economics*, Vol. 51, pp. 88-100.

Turnbull, G.K. and Dombrow, J. (2006), "Spatial competition and shopping externalities: Evidence from the housing market", *The Journal of Real Estate Finance and Economics*, Vol. 32 (4), pp. 391-408.

Turnbull, G.K., Dombrow, J., and Sirmans, C.F. (2006), "Big house, little house: relative size and value", *Real Estate Economics*, Vol. 34 (3), pp. 439-456.

Valente, J., Wu, S.S., Gelfand, A., and Sirmans C.F. (2005), "Apartment Rent Prediction Using Spatial Modeling", *Real Estate Research*, Vol. 27 (1), pp. 105-136.

Voigtlaender, M. (2009), "Why is the German homeownership rate so low?", *Housing Studies*, Vol. 24 (3), pp. 335-372.

Wehby, G.L., Murray, J.C., Castilla, E., Lopez-Camelo, J.S., and Ohsfeldt, R.L. (2009), "Prenatal care effectiveness and utilization in Brazil", *Health Policy and Planning*, Vol. 24 (3), pp. 175-188.

Wei, Y., Pere, A., Koenker, R., and He, X. (2006), "Quantile regression methods for reference growth charts", *Statistics in Medicine*, Vol. 25 (8), pp. 1369-1382.

Wood, J.H. and Wood, N.L. (1985), "Financial Markets", San Diego: Harcourt Brace Jovanovich.

Yang, S. (1999), Censored median regression using weighted empirical survival and hazard functions, *Journal of the American Statistical Association*, Vol. 94 (445), pp. 137-145.

Yavas, A. and Yang, S. (1995), The strategic role of listing price in marketing real estate: theory and evidence, *Real Estate Economics*, Vol. 23 (3), pp. 347-368.

Yu, K., Lu, Z., and Stander, J. (2003), Quantile regression: applications and current research areas, *Journal of the Royal Statistical Society: Series D (The Statistician)*, Vol. 52 (3), pp. 331-350.

Zietz, J., Zietz, E.N., and Sirmans, G.S. (2008), Determinants of house prices: a quantile regression approach, *The Journal of Real Estate Finance and Economics*, Vol. 37 (4), pp. 317-333.

4.7 Appendix

Table 4.5: Results of the Censored Quantile Regression

CQR Coefficients and P-Values per quantile	Q20	Q30	Q40	Q50	Q60	Q70	Q80
Effect of log of asking rent on TOM							
Berlin	2.230 0.000***	2.171 0.000***	2.091 0.000***	1.995 0.000***	1.905 0.000***	1.783 0.000***	1.697 0.000***
Hamburg	2.252 0.000***	2.186 0.000***	2.106 0.000***	1.994 0.000***	1.870 0.000***	1.682 0.000***	1.422 0.000***
Munich	2.345 0.000***	2.448 0.000***	2.396 0.000***	2.316 0.000***	2.113 0.000***	1.900 0.000***	1.618 0.000***
Cologne	2.203 0.000***	2.090 0.000***	1.885 0.000***	1.697 0.000***	1.528 0.000***	1.330 0.000***	1.226 0.000***
Frankfurt	2.202 0.000***	2.116 0.000***	1.871 0.000***	1.697 0.000***	1.503 0.000***	1.377 0.000***	1.341 0.000***
Stuttgart	2.474 0.000***	2.597 0.000***	2.451 0.000***	2.416 0.000***	2.165 0.000***	1.837 0.000***	1.555 0.000***
Dusseldorf	1.961 0.000***	1.740 0.000***	1.467 0.000***	1.283 0.000***	1.251 0.000***	1.176 0.000***	1.159 0.000***
Effect of log of living area on TOM							
Berlin	-0.776 0.000***	-0.669 0.000***	-0.589 0.000***	-0.476 0.000***	-0.409 0.000***	-0.308 0.000***	-0.280 0.000***
Hamburg	-0.544 0.000***	-0.511 0.000***	-0.449 0.000***	-0.401 0.000***	-0.321 0.000***	-0.222 0.000***	-0.122 0.023**
Munich	-0.556 0.000***	-0.679 0.000***	-0.686 0.000***	-0.676 0.000***	-0.584 0.000***	-0.453 0.000***	-0.240 0.052*
Cologne	-0.736 0.000***	-0.602 0.000***	-0.449 0.000***	-0.316 0.001***	-0.172 0.016**	-0.048 0.512	-0.024 0.541
Frankfurt	-0.622 0.000***	-0.583 0.000***	-0.399 0.000***	-0.220 0.002***	-0.121 0.032**	-0.037 0.667	-0.052 0.404
Stuttgart	-0.949 0.000***	-1.080 0.000***	-0.962 0.000***	-0.974 0.000***	-0.791 0.000***	-0.575 0.000***	-0.401 0.000***
Dusseldorf	-0.698 0.000***	-0.482 0.000***	-0.213 0.001***	-0.087 0.116	-0.065 0.433	0.002 0.974	-0.039 0.641
Effect of age on TOM							
Berlin	-0.003 0.222	-0.002 0.180	-0.001 0.322	-0.000 0.719	0.001 0.409	0.001 0.470	0.003 0.100*
Hamburg	-0.006 0.018**	-0.005 0.000***	-0.007 0.001***	-0.006 0.002***	-0.009 0.011**	-0.005 0.000***	-0.003 0.244
Munich	-0.019 0.000***	-0.015 0.020**	-0.015 0.001***	-0.011 0.001***	-0.011 0.020**	-0.012 0.001***	-0.011 0.007***
Cologne	-0.009 0.045**	-0.008 0.013**	-0.003 0.521	-0.005 0.101	-0.007 0.040**	-0.005 0.110	-0.006 0.025**
Frankfurt	-0.013 0.018**	-0.013 0.000***	-0.012 0.000***	-0.011 0.000***	-0.010 0.000***	-0.011 0.000***	-0.009 0.000***
Stuttgart	0.007 0.418	0.004 0.563	-0.003 0.471	-0.005 0.069*	-0.005 0.234	-0.005 0.021**	-0.008 0.005***
Dusseldorf	-0.004 0.3555	-0.004 0.275	-0.003 0.207	-0.004 0.060*	-0.004 0.117	-0.007 0.046*	-0.008 0.046*

Effect of number of rooms on TOM							
Berlin	-0.247 0.000***	-0.257 0.000***	-0.254 0.000***	-0.258 0.000***	-0.247 0.000***	-0.242 0.000***	-0.222 0.000***
Hamburg	-0.310 0.000***	-0.295 0.000***	-0.286 0.000***	-0.262 0.000***	-0.248 0.000***	-0.230 0.000***	-0.197 0.000***
Munich	-0.258 0.000***	-0.234 0.000***	-0.212 0.000***	-0.205 0.000***	-0.185 0.000***	-0.169 0.000***	-0.170 0.000***
Cologne	-0.223 0.000***	-0.223 0.000***	-0.223 0.000***	-0.206 0.000***	-0.201 0.000***	-0.191 0.000***	-0.162 0.000***
Frankfurt	-0.273 0.000***	-0.233 0.000***	-0.227 0.000***	-0.232 0.000***	-0.210 0.000***	-0.205 0.000***	-0.202 0.000***
Stuttgart	-0.148 0.000***	-0.136 0.000***	-0.133 0.000***	-0.133 0.000***	-0.130 0.000***	-0.114 0.000***	-0.106 0.000***
Dusseldorf	-0.233 0.000***	-0.231 0.000***	-0.231 0.000***	-0.207 0.000***	-0.204 0.000***	-0.202 0.000***	-0.183 0.000***
Effect of first occupancy on TOM							
Berlin	0.178 0.000***	0.196 0.000***	0.195 0.000***	0.207 0.000***	0.195 0.000***	0.203 0.000***	0.227 0.000***
Hamburg	0.104 0.000***	0.097 0.000***	0.091 0.000***	0.113 0.000***	0.136 0.000***	0.130 0.000***	0.133 0.000***
Munich	0.201 0.000***	0.211 0.000***	0.200 0.000***	0.192 0.000***	0.216 0.000***	0.206 0.000***	0.215 0.000***
Cologne	0.054 0.232	0.027 0.399	0.045 0.028**	0.045 0.025**	0.065 0.020**	0.074 0.000***	0.082 0.000***
Frankfurt	0.109 0.000***	0.083 0.070**	0.105 0.009***	0.110 0.000***	0.110 0.000***	0.096 0.001***	0.092 0.000***
Stuttgart	0.160 0.018**	0.110 0.028**	0.073 0.101	0.023 0.655	0.057 0.320	0.032 0.512	0.025 0.678
Dusseldorf	0.075 0.021**	0.053 0.110	0.092 0.000***	0.102 0.000***	0.102 0.000***	0.098 0.000***	0.090 0.046**
Effect of renovated on TOM							
Berlin	0.097 0.000***	0.087 0.000***	0.074 0.000***	0.065 0.000***	0.054 0.000***	0.028 0.001**	0.012 0.329
Hamburg	0.175 0.000***	0.154 0.000***	0.142 0.000***	0.123 0.000***	0.125 0.000***	0.125 0.000***	0.123 0.000***
Munich	0.077 0.000***	0.068 0.033**	0.044 0.000***	0.054 0.000***	0.036 0.067*	0.051 0.048**	0.039 0.009***
Cologne	0.094 0.000***	0.064 0.000***	0.052 0.000***	0.028 0.127	0.012 0.421	0.007 0.567	-0.015 0.486
Frankfurt	0.146 0.000***	0.107 0.000***	0.075 0.000***	0.083 0.000***	0.098 0.000***	0.098 0.000***	0.101 0.000***
Stuttgart	0.179 0.000***	0.117 0.001***	0.130 0.000***	0.095 0.004***	0.055 0.006***	0.039 0.108	0.019 0.583
Dusseldorf	0.178 0.000***	0.143 0.000***	0.113 0.000***	0.100 0.000***	0.105 0.000***	0.096 0.000***	0.104 0.000***
Effect of with bathtub on TOM							
Berlin	0.308 0.000***	0.288 0.000***	0.265 0.000***	0.250 0.000***	0.234 0.000***	0.212 0.000***	0.183 0.000***
Hamburg	0.048 0.041**	0.039 0.011**	0.032 0.065*	0.022 0.008***	0.036 0.001***	0.062 0.000***	0.074 0.000***
Munich	0.091 0.002***	0.069 0.001***	0.073 0.000***	0.070 0.002***	0.072 0.000***	0.0872 0.000***	0.084 0.000***

Cologne	0.066 0.000***	0.078 0.000***	0.082 0.009***	0.057 0.015**	0.060 0.017**	0.065 0.000***	0.063 0.000***
Frankfurt	0.153 0.000***	0.137 0.000***	0.152 0.000***	0.145 0.000***	0.160 0.000***	0.164 0.000***	0.172 0.000***
Stuttgart	0.137 0.000***	0.103 0.012**	0.119 0.000***	0.112 0.000***	0.142 0.000***	0.122 0.000***	0.122 0.000***
Dusseldorf	0.103 0.000***	0.081 0.000***	0.096 0.000***	0.116 0.000***	0.112 0.000***	0.127 0.000***	0.143 0.000***
Effect of with built-in kitchen on TOM							
Berlin	-0.068 0.000***	-0.079 0.000***	-0.091 0.000***	-0.100 0.000***	-0.115 0.000***	-0.123 0.000***	-0.147 0.000***
Hamburg	0.019 0.373	0.081 0.000***	0.083 0.000***	0.094 0.000***	0.082 0.000***	0.065 0.000***	0.053 0.001***
Munich	-0.013 0.379	-0.031 0.124	-0.042 0.066*	-0.052 0.006***	-0.067 0.000***	-0.054 0.001***	-0.064 0.000***
Cologne	-0.149 0.000***	-0.157 0.000***	-0.145 0.000***	-0.138 0.000***	-0.157 0.000***	-0.181 0.000***	-0.185 0.000***
Frankfurt	-0.120 0.007***	-0.144 0.000***	-0.134 0.000***	-0.138 0.000***	-0.138 0.000***	-0.161 0.000***	-0.168 0.000***
Stuttgart	0.105 0.023**	0.049 0.163	0.016 0.479	-0.045 0.327	-0.088 0.001***	-0.119 0.011**	-0.175 0.000***
Dusseldorf	-0.215 0.000***	-0.203 0.000***	-0.209 0.000***	-0.232 0.000***	-0.242 0.000***	-0.242 0.000***	-0.242 0.000***
Effect of with car space on TOM							
Berlin	0.138 0.000***	0.128 0.000***	0.141 0.000***	0.137 0.000***	0.135 0.000***	0.124 0.000***	0.116 0.000***
Hamburg	0.142 0.000***	0.135 0.000***	0.133 0.000***	0.152 0.000***	0.159 0.000***	0.168 0.000***	0.179 0.000***
Munich	0.152 0.000***	0.169 0.000***	0.141 0.000***	0.143 0.000***	0.162 0.000***	0.134 0.000***	0.107 0.000***
Cologne	0.070 0.062*	0.057 0.001***	0.032 0.126	0.039 0.004***	0.034 0.023**	0.024 0.311	-0.002 0.937
Frankfurt	0.125 0.012**	0.113 0.000***	0.102 0.006***	0.101 0.000***	0.095 0.000***	0.075 0.000***	0.088 0.000***
Stuttgart	0.021 0.722	0.001 0.985	0.003 0.883	0.028 0.382	-0.001 0.974	-0.022 0.566	0.031 0.269
Dusseldorf	0.057 0.092*	0.0355 0.021**	0.021 0.433	0.019 0.365	0.020 0.323	0.020 0.442	0.003 0.869
Effect of with terrace on TOM							
Berlin	0.064 0.000***	0.072 0.000***	0.088 0.000***	0.086 0.000***	0.098 0.000***	0.104 0.000***	0.097 0.000***
Hamburg	-0.023 0.257	-0.034 0.047**	-0.027 0.083*	-0.026 0.104	-0.001 0.971	0.009 0.578	0.032 0.177
Munich	0.078 0.004***	0.082 0.000***	0.067 0.000***	0.064 0.008***	0.075 0.000***	0.083 0.000***	0.112 0.001***
Cologne	0.012 0.728	0.028 0.246	0.030 0.067*	0.035 0.029**	0.058 0.002***	0.076 0.008***	0.095 0.001***
Frankfurt	0.038 0.412	0.079 0.002***	0.083 0.001***	0.096 0.000***	0.127 0.001***	0.135 0.000*	0.116 0.000***
Stuttgart	0.019 0.695	0.039 0.564	0.015 0.527	-0.005 0.890	-0.003 0.914	0.032 0.460	0.011 0.789

Dusseldorf	-0.023 0.259	-0.028 0.050*	-0.033 0.002***	-0.021 0.520	-0.022 0.477	-0.009 0.746	0.047 0.000***
Effect of with balcony on TOM							
Berlin	0.006 0.669	0.001 0.908	0.001 0.865	-0.010 0.480	-0.011 0.154	-0.005 0.653	0.012 0.297
Hamburg	-0.066 0.000***	-0.077 0.000***	-0.086 0.000***	-0.089 0.000***	-0.088 0.000***	-0.076 0.000***	-0.059 0.000***
Munich	0.000 0.995	0.032 0.003***	0.040 0.006***	0.042 0.000***	0.019 0.339	0.013 0.488	0.000 0.990
Cologne	-0.054 0.056*	-0.042 0.044**	-0.023 0.345	-0.028 0.099*	-0.031 0.034**	-0.020 0.306	-0.015 0.360
Frankfurt	0.020 0.381	-0.024 0.251	-0.032 0.119	-0.046 0.000***	-0.0294 0.001***	0.007 0.700	0.011 0.413
Stuttgart	0.057 0.292	0.080 0.033**	0.042 0.037**	0.008 0.775	0.023 0.598	0.001 0.974	-0.021 0.316
Dusseldorf	0.021 0.459	-0.005 0.840	-0.015 0.437	-0.007 0.680	-0.013 0.499	-0.013 0.316	-0.019 0.220
Effect of with elevator on TOM							
Berlin	0.041 0.004***	0.053 0.000***	0.065 0.000***	0.075 0.000***	0.077 0.000***	0.065 0.000***	0.049 0.000***
Hamburg	0.050 0.007***	0.058 0.000***	0.061 0.000***	0.080 0.000***	0.084 0.000***	0.113 0.000***	0.135 0.000***
Munich	-0.112 0.000***	-0.101 0.000***	-0.109 0.000***	-0.108 0.000***	-0.092 0.000***	-0.084 0.000***	-0.054 0.000***
Cologne	0.091 0.000***	0.112 0.000***	0.145 0.000***	0.194 0.000***	0.211 0.000***	0.207 0.000***	0.225 0.000***
Frankfurt	-0.076 0.021**	-0.031 0.140	0.009 0.759	0.021 0.362	0.063 0.011**	0.083 0.000***	0.091 0.000***
Stuttgart	0.025 0.575	0.027 0.481	0.038 0.370	0.093 0.000***	0.100 0.008***	0.133 0.000***	0.174 0.000***
Dusseldorf	-0.019 0.201	-0.005 0.842	0.036 0.052*	0.076 0.000***	0.102 0.000***	0.111 0.000***	0.132 0.000***
Effect of log of purchasing power per household in ZIP code area on TOM							
Berlin	-0.941 0.000***	-0.911 0.000***	-0.855 0.000***	-0.848 0.000***	-0.878 0.000***	-0.858 0.000***	-0.838 0.000***
Hamburg	-0.402 0.000***	-0.343 0.000***	-0.427 0.000***	-0.498 0.000***	-0.466 0.000***	-0.423 0.000***	-0.386 0.000***
Munich	0.197 0.020**	0.290 0.000***	0.227 0.047**	0.175 0.140	0.153 0.235	0.139 0.138	0.202 0.013**
Cologne	-0.565 0.012**	-0.690 0.000***	-0.587 0.000***	-0.522 0.000***	-0.437 0.013**	-0.401 0.095*	-0.381 0.016**
Frankfurt	-0.086 0.748	-0.235 0.081*	-0.156 0.171	-0.183 0.290	-0.151 0.208	-0.256 0.014**	-0.308 0.022**
Stuttgart	0.720 0.042**	0.470 0.028**	0.338 0.019**	0.128 0.223	0.058 0.685	-0.003 0.986	-0.222 0.217
Dusseldorf	-0.280 0.141	-0.244 0.147	-0.272 0.081*	-0.256 0.005***	-0.272 0.019**	-0.393 0.006***	-0.566 0.000***
Effect of log of households in ZIP code area on TOM							
Berlin	-0.004 0.873	-0.037 0.002***	-0.044 0.001***	-0.041 0.005***	-0.056 0.000***	-0.034 0.004***	-0.043 0.000***
Hamburg	0.055 0.062*	0.062 0.022**	0.027 0.259	-0.013 0.563	-0.022 0.269	-0.062 0.034**	-0.074 0.000***

Munich	0.115 0.000***	0.157 0.001***	0.162 0.000***	0.129 0.000***	0.130 0.015**	0.113 0.000***	0.055 0.035**
Cologne	-0.148 0.034**	-0.162 0.000***	-0.165 0.000***	-0.137 0.009***	-0.123 0.000***	-0.134 0.000***	-0.142 0.000***
Frankfurt	-0.010 0.762	-0.039 0.186	-0.029 0.384	-0.066 0.007***	-0.080 0.001***	-0.086 0.008***	-0.078 0.000***
Stuttgart	0.012 0.853	-0.050 0.528	-0.058 0.079*	-0.108 0.001***	-0.148 0.016**	-0.147 0.000***	-0.152 0.000***
Dusseldorf	0.011 0.725	-0.000 0.995	-0.029 0.296	-0.048 0.205	-0.057 0.024**	-0.068 0.017**	-0.095 0.012**
Effect of log of population density in ZIP code area on TOM							
Berlin	0.003 0.607	0.012 0.026**	0.010 0.028**	0.014 0.000***	0.011 0.010**	0.004 0.453	-0.001 0.876
Hamburg	0.061 0.000***	0.064 0.000***	0.073 0.000***	0.078 0.000***	0.086 0.000***	0.084 0.000***	0.076 0.000***
Munich	0.044 0.000***	0.057 0.000***	0.061 0.000***	0.064 0.000***	0.061 0.000***	0.063 0.000***	0.059 0.000***
Cologne	-0.007 0.313	-0.009 0.698	-0.010 0.481	-0.016 0.137	-0.006 0.647	-0.002 0.885	-0.020 0.158
Frankfurt	0.086 0.000***	0.085 0.000***	0.076 0.000***	0.075 0.000***	0.080 0.000***	0.067 0.000***	0.058 0.000***
Stuttgart	0.015 0.672	0.026 0.337	0.021 0.105	0.028 0.136	0.032 0.084*	0.029 0.169	0.036 0.075*
Dusseldorf	-0.013 0.447	-0.009 0.308	-0.006 0.672	0.001 0.896	0.003 0.703	0.007 0.665	0.005 0.662
Effect of distance to ZIP code area centroid on TOM							
Berlin	0.050 0.021**	0.044 0.000***	0.056 0.001***	0.050 0.000***	0.062 0.000***	0.068 0.000***	0.052 0.000***
Hamburg	0.192 0.000***	0.174 0.000***	0.169 0.000***	0.142 0.000***	0.122 0.000***	0.108 0.000***	0.108 0.000***
Munich	0.093 0.019**	0.108 0.000***	0.092 0.000***	0.107 0.001***	0.124 0.000***	0.096 0.000***	0.095 0.001***
Cologne	0.116 0.000***	0.085 0.000***	0.052 0.000***	0.050 0.000***	0.036 0.039**	0.039 0.059*	0.045 0.039**
Frankfurt	0.207 0.000***	0.225 0.000***	0.183 0.000***	0.157 0.000***	0.141 0.000***	0.142 0.000***	0.151 0.000***
Stuttgart	0.036 0.491	0.026 0.627	0.038 0.235	0.030 0.281	0.049 0.079*	0.079 0.024**	0.065 0.015**
Dusseldorf	0.062 0.142	0.038 0.113	0.058 0.014**	0.050 0.056*	0.053 0.002***	0.063 0.024**	0.096 0.000***
Effect of distance to NUTS3 region centroid on TOM							
Berlin	0.099 0.000***	0.098 0.000***	0.094 0.000***	0.090 0.000***	0.086 0.000***	0.079 0.000***	0.073 0.000***
Hamburg	0.095 0.000***	0.090 0.000***	0.086 0.000***	0.084 0.000***	0.078 0.000***	0.074 0.000***	0.063 0.000***
Munich	0.048 0.000***	0.049 0.000***	0.045 0.000***	0.033 0.000***	0.025 0.000***	0.025 0.000***	0.019 0.019**
Cologne	0.095 0.000***	0.095 0.000***	0.091 0.000***	0.081 0.000***	0.072 0.000***	0.062 0.000***	0.058 0.000***
Frankfurt	0.092 0.000***	0.080 0.000***	0.078 0.000***	0.081 0.000***	0.075 0.000***	0.066 0.000***	0.058 0.000***

Stuttgart	0.083 0.000***	0.068 0.000***	0.066 0.000***	0.073 0.000***	0.067 0.000***	0.053 0.000***	0.042 0.000***
Dusseldorf	0.062 0.000***	0.055 0.000***	0.051 0.000***	0.046 0.000***	0.045 0.000***	0.041 0.000***	0.041 0.000***
Effect of effective interest rate on TOM							
Berlin	0.369 0.017**	0.113 0.341	0.056 0.516	-0.078 0.174	-0.194 0.003***	-0.305 0.000***	-0.349 0.000***
Hamburg	0.540 0.001***	0.297 0.005***	0.168 0.170	0.042 0.684	-0.060 0.659	-0.222 0.171	-0.325 0.007***
Munich	0.595 0.001***	0.349 0.113	0.419 0.030**	0.294 0.037**	0.413 0.000*	0.371 0.000***	0.228 0.259
Cologne	0.124 0.515	0.112 0.552	0.020 0.937	-0.029 0.879	-0.097 0.402	-0.052 0.572	-0.127 0.195
Frankfurt	0.480 0.129	0.250 0.400	0.143 0.556	-0.028 0.875	0.010 0.967	-0.004 0.985	-0.077 0.795
Stuttgart	0.425 0.321	0.333 0.318	0.426 0.200	0.415 0.215	0.583 0.005***	0.0645 0.847	0.139 0.431
Dusseldorf	0.068 0.731	0.130 0.381	0.155 0.369	0.010 0.910	-0.113 0.463	-0.163 0.194	-0.289 0.197

Notes*:* This table displays the coefficients and p-values of the individual censored quantile regressions of dwellings' time on market in weeks on hedonic, spatial and socioeconomic variables as well as the effective ten-year interest rate. The data consists of 482,196 observations on the residential rental market. The sample period is 2013 Q1 to 2017 Q4. *Significant at the 10%-level; ** significant at the 5%-level; *** significant at the 1%-level.

5. Conclusion

The following section provides a summary of the three papers comprising this dissertation including the research purpose, the study design, key findings, and practical implications. The dissertation concludes with some final remarks and an outlook on further research on the liquidity of residential real estate.

5.1 Executive Summaries

Paper 1: Closing the liquidity gap: Why the consideration of time on market is inevitable for understanding the residential market

The aim of this paper is to emphasize the importance of time on market when analyzing the residential real estate market. With the increasing availability of internet based data on the German residential market, some authors and institutions have started to investigate market movements with the exclusive consideration of price, see e.g. Bauer et al. (2013) and an de Meulen and Mitze (2014), among others. The second integral component when marketing a dwelling, which is the time it takes until the transaction is completed, is not taken into account. Nevertheless, the consideration of time on market seems to be of crucial importance for detecting tight market conditions, as especially for the rental market, the price development appears not to suggest surplus demand. Those clustered conditions are then used to derive investment strategies.

A large scaled data set consisting of 973,164 observations on the German residential investment market and 2,082,179 observations on the German rental market of 161 NURS3 regions over a five-year sample period from 2013 Q1 to 2017 Q4 is applied. The data contains information on the hedonic characteristics of the dwellings, extended with absolute and relative spatial information and socioeconomic information. The time dummy approach is used in order to generate price and liquidity indices for the German residential investment and rental market. The coefficients of the time dummy variable are extracted from a quality- and spatial-adjusted hedonic regression. While the General Additive Model for Location, Scale and Shape (GAMLSS) introduced by Rigby and Stasinopoulos (2005) is applied to generate the price indices, a Cox (1972) Proportional Hazards Model is applied to generate the liquidity indices. Subsequently, the index values are clustered using the Partitioning Around Medoids (PAM) algorithm going back to Kaufman and Rousseeuw (1987), in order to summarize similar markets and to facilitate the interpretation of the results regarding the 161 regions.

The study is able to proof that the consideration of time on market is essential to detect the true demand on the German residential rental market. While the official German rental index displays a rental increase of a mere 5.7% over the last five years, the liquidity for dwellings in locations assigned to the highest liquidity cluster almost tripled during the same period. Overall, the regional analysis of the price and rent clusters yields a diversified pattern of strong investment and rental markets. While a slight concentration on southern states might be indicated, high performing markets in terms of price are found across the whole country. Only the combined classification of price and liquidity clusters reveals a strong focus on Baden-Wuerttemberg. In addition, the regional analysis suggest stronger spillover tendencies on the rental market, as a larger number of adjacent regions which experienced an identical development was found.

To sum up, the findings of this paper contribute to a better understanding of the German residential real estate market by the introduction of a liquidity index and are able to emphasize the importance of a combined analysis of price and liquidity. The results allow the deduction of investment strategies and assist policy makers on the identification of tight markets and a prioritization of subsequent actions.

Paper 2: Exploring the determinants of liquidity with big data – market heterogeneity in German markets

While the first paper was able to underline the importance of liquidity, especially when analyzing the residential rental market, the second paper aims at exploring the determinants of liquidity within the largest seven German cities, which are at the same time the largest residential real estate markets. While Germany is well known for its low home ownership rate, the proportion of the population which is renting their homes, is far exceeding the national average within these seven cities. Therefore, the examination of the rental market within those cities should yield a robust assessment of the entire residential market within those cities.

A dataset consisting of 335,972 observations on the rental market within the largest seven German cities Berlin, Cologne, Dusseldorf, Frankfurt, Hamburg, Munich, and Stuttgart over a 45 months sample period starting in 2013 Q1 is applied. The data contains information on the hedonic characteristics of the dwellings, extended with absolute and relative spatial information and socioeconomic information. While spatial information was used in the context of pricing by e.g. Goodman and Thibodeau (2007), Turnbull and Dombrow (2006), Bourassa et al. (2010) and Cirman et al. (2015), among others, Smith (2010) was the first to include district dummies and Cartesian coordinates as well as a distance variable in the context of liquidity analysis. To the best of the author’s knowledge, the present paper is the first to employ relative spatial information in terms of distance variables in the context of real estate liquidity analysis on the German residential market. A Cox Proportional Hazards Model is used in order to reveal the direction and the magnitude of the impact the various covariates have on time on market. The study also examines, which additional information can be derived by including the variables “degree of atypicality” introduced by Haurin (1988) and the “degree of overpricing” introduced by Anglin et al. (2003) into the analysis.

The paper contributes to a better understanding regarding the determinants of liquidity on the largest seven German real estate markets and is able to identify communalities and differences across those markets. While for each city, asking rent, living area, dwelling age, and distance to the NUTS3 centroid show consistent effects, the other hedonic characteristics and the degree of overpricing display a market-specific impact on liquidity. Based on these results, geographic liquidity patterns are derived for the observed cities.

By means of this approach, the article contributes to the literature on time on market modelling on the German market, by enhancing the quality of the econometric approach, and by introducing spatial gravity variables and other fixed effects.

Paper 3: Exploring the determinants of real estate liquidity from an alternative perspective – Censored Quantile Regression in real estate research

Building upon the previous findings, the final paper of the dissertation aims at exploring the determinants of liquidity with a higher level of granularity by introducing the advanced Censored Quantile Regression (CQR) model to the field of real estate liquidity analysis. The heterogeneity in the impact of various covariates on time on market across the cities, which was identified by the application of the Cox PHM, indicates the need for a more segmented, thus detailed analysis of the individual cities.

A dataset consisting of 482,196 observations on the rental market within the largest seven German cities Berlin, Cologne, Dusseldorf, Frankfurt, Hamburg, Munich, and Stuttgart from 2013 Q1 to 2017 Q4 is applied. The data contains information on the hedonic characteristics of the dwellings, extended with absolute and relative spatial information and socioeconomic information. The Quantile Regression (QR) approach, going back to Koenker and Bassett Jr. (1978), allows the estimation of the relationship a covariate has on the independent variable, conditional on an individual quantile of the independent variable. The QR method has been applied in real estate pricing by e.g. Zietz et al. (2008), Farmer and Lipscomb (2010), Mak et al. (2010), Liao and Wang (2012), among others. An de Meulen and Mitze (2014) and Tomschke (2015) used the method on the German market. An important feature time on market analysis is, that some observations do not change their event status throughout the observation period. This means, that some dwellings remain available on the market until the end of the observation period. If this is the case, the response variable, time on market is right-censored. To deal with censoring within the QR framework, three main approaches have been introduced by Powell (1984, 1986), Portnoy (2003) and Peng and Huang (2008). For the present dataset, Powell's (1984, 1986) approach is best suited as it addresses fixed censoring.

For almost all regression coefficients, consistent signs across the quantiles of the time on market distribution are found. Thus, the proportional hazard assumption, underlying the Cox PHM, is not violated for most covariates. However, the impact of a change in the explanatory variables differs in magnitude and significance across the liquidity quantiles. In contrast to the traditional Cox Proportional Hazards Model, the Censored Quantile Regression properly detects these differences. Thus, the model is better suited to explore the heterogeneity of individual dwellings. Furthermore, the magnitude, the significance and the direction of the impact of the covariates on time on market is quite different across the cities.

These findings emphasize the importance of market segmentation for a more detailed analysis and understanding of the rental real estate market or the real estate market as a whole. The paper extends the existing literature on residential liquidity analysis by the initial application of the advanced Censored Quantile Regression (CQR). In addition, the paper enables landlords to infer whether an individual dwelling displays the characteristics of a highly liquid thus highly demanded dwelling and what actions they could take in order to shorten the expected liquidity.

5.2 Final Remarks and Further Research

In times of an ongoing discussion about urbanization, the associated rural exodus and a significant shortage of living space in conurbations, this dissertation lays the foundation for residential real estate liquidity analysis on the German residential market. The first paper demonstrates the importance of liquidity analysis, in order to reveal the real demand on the rental market and identifies “hot” and “cold” markets on a regional basis. Therefore, it is the first paper, to the best of the author’s knowledge, to introduce a liquidity index to the field of residential real estate research. The second paper focuses on a detailed examination of the largest seven German real estate markets Berlin, Cologne, Dusseldorf, Frankfurt, Hamburg, Munich, and Stuttgart, by extending an established econometric model for real estate liquidity analysis with the initial integration of spatial gravity variables to real estate liquidity analysis on the German market. The approach is able to increase the explanatory power of the model, while at the same time identifying heterogeneity across the cities and liquidity patterns within the top seven markets. The third paper aims at embracing this heterogeneity in order to get an even deeper understanding of the variables driving liquidity within the largest seven German cities. The paper is the first, to the best of the author’s knowledge, to apply the advanced Censored Quantile Regression to the field of real estate research. The model allows the identification of the impact of individual covariates on the time on market of dwellings, depending on the individual time on market quantile. The findings of this dissertation should be of interest to both consumers and providers of living space as well as policy makers, as they provide the second integral component for a better understanding of residential real estate markets by complementing the exclusive consideration of price.

Since the aim of this dissertation is to emphasize the importance of liquidity analysis on the German residential real estate market, there are still plenty of research questions which might to be addressed with further research. Although the increasing availability of data on the German residential market enabled this study, it is mainly insufficient data, which sets limits to future research. Up to now, only asking price and asking rent are observable. The consideration of transaction prices and contracted rent, however, would allow the researcher to infer in how far liquidity on the German market is influenced by pricing behavior of sellers and landlords. With a longer sample period it should also be possible to include more granular socioeconomic data, as the Federal Institute for Research on Building, Urban Affairs and Spatial Development (BBSR) only publishes this data with a more than two year time-lag. In addition, a longer sample period might shed light on the inter-temporal causality between price or rent and liquidity. Based on this ex-post analysis, a forecasting model for price and liquidity might be constructed. The inclusion of sentiment indicators such as Google trends data or sentiment indicators created by textual analysis, might improve the explanatory power of those forecasting models for the German residential market, see e.g. Hohenstatt et al. (2011) or Beracha and Wintoki (2013), among others. Furthermore, a decomposition of the impact of individual variables might yield further clarification. For a decomposition in the field of pricing on the German residential market see e.g. Thomschke (2015).

With residential real estate being an essential part of day-to-day life, the dissertation might not only attract the attention of researchers, institutional investors, and policy makers, but also be of interest to private individual landlords, property owners, and tenants on the German market.

5.3 References

an de Meulen, P. and Mitze, T. (2014), “Exploring the spatial variation in quality-adjusted rental prices and identifying hot spots in Berlin‘s residential property market” *Regional Studies, Regional Science*, Vol. 1 (1), pp. 310-328.

Anglin, P.M., Rutherford, R.C. and Springer, T.M. (2003), “The Trade-off Between Selling Price of Residential Properties and Time-on-the-Market: The Impact of Price Setting”, *Journal of Real Estate Finance and Economics*, Vol. 26 (1), pp. 95-111.

Bauer, T. K., Feuerschuette, S., Kiefer, M., an de Meulen, P., Micheli, M., Schmidt, T., and Wilke, L. H. (2013), “Ein hedonischer Immobilienpreisindex auf Basis von Internetdaten: 2007–2011“, *AStA Wirtschafts-und Sozialstatistisches Archiv*, Vol. 7 (1-2), pp. 5-30.

Beracha, E. and Wintoki, B. (2013), “Forecasting residential real estate price changes from online search activity”, *Journal of Real Estate Research*, Vol. 35 (3), pp. 283-312.

Bourassa, S., Cantoni, E. and Hoesli, M. (2010), “Predicting House Prices with Spatial Dependence: A Comparison of Alternative Methods”, *Journal of Real Estate Research*, Vol. 32 (2), pp. 139-159.

Cirman, A., Pahor, M. and Verbic, M. (2015), “Determinants of Time on the Market in a Thin Real Estate Market”, *Inzinerine Ekonomika-Engineering Economics*, Vol. 26 (1), pp. 4-11.

Cox, D.R. (1972), “Regression models and life-tables”, *Journal of the Royal Statistical Society: Series B (Methodological)*, Vol. 34 (2), pp. 187-220.

Farmer, M. and Lipscomb, C. (2010), “Using quantile regression in hedonic analysis to reveal submarket competition”, *Journal of Real Estate Research*, Vol. 32 (4), pp. 435-460.

Goodman, A.C. and Thibodeau, T.G. (2007), “The Spatial Proximity of Metropolitan Area Housing Submarkets”, *Real Estate Economics*, Vol. 35 (2), pp. 209-232.

Haurin, D.R. (1988), “The Duration of Marketing Time of Residential Housing”, *Real Estate Economics*, Vol. 16 (4), pp. 396-410.

Hohenstatt, R., Kaesbauer, M., and Schaefers, W. (2011), “’Geco’ and its Potential for Real Estate Research: Evidence from the U.S. Housing Market”, *Journal of Real Estate Research*, Vol. 33 (4), pp. 471-506.

Kaufman, L. and Rousseeuw, P.J. (1987), “Clustering by means of medoids”, in: *Statistical data analysis based on the L1-norm and related methods*, North-Holland, New York, pp. 405-416.

Koenker, R. and Bassett Jr., G. (1978), “Regression quantiles”, *Econometrica,* Vol. 46 (1), pp. 33-50.

Liao, W.C. and Wang, X. (2012), “Hedonic house prices and spatial quantile regression”, *Journal of Housing Economics*, Vol. 21 (1), pp. 16-27.

Mak, S., Choy, L., and Ho, W. (2010), "Quantile regression estimates of Hong Kong real estate prices", *Urban Studies*, Vol. 47 (11), pp. 2461-2472.

Peng, L. and Huang, Y. (2008), "Survival analysis with quantile regression models", *Journal of the American Statistical Association*, Vol. 103(482), pp. 637-649.

Portnoy, S. (2003), "Censored regression quantiles", *Journal of the American Statistical Association*, Vol. 98 (464), pp. 1001-1012.

Powell, J.L. (1984), "Least absolute deviations estimation for the censored regression model", *Journal of Econometrics*, Vol. 25 (3), pp. 303-325.

Powell, J.L. (1986), "Censored regression quantiles", *Journal of Econometrics*, Vol. 32 (1), pp. 143-155.

Rigby, R.A. and Stasinopoulos, D.M. (2005), "Generalized additive models for location, scale and shape", *Journal of the Royal Statistical Society: Series C (Applied Statistics)*, Vol. 54 (3), pp. 507-554.

Smith, B.C. (2010), "Spatial Heterogeneity in Listing Duration: The Influence of Relative Location to Marketability", *Journal of Housing Research*, Vol. 18 (2), pp. 151-171.

Thomschke, L. (2015), "Changes in the distribution of rental prices in Berlin", *Regional Science and Urban Economics*, Vol. 51, pp. 88-100.

Turnbull, G.K. and Dombrow, J. (2006), "Spatial Competition and Shopping Externalities: Evidence from the Housing Market", *Journal of Real Estate Finance and Economics*, Vol. 32 (4), pp. 391-408.

Zietz, J., Zietz, E.N., and Sirmans, G.S. (2008), "Determinants of house prices: a quantile regression approach", *The Journal of Real Estate Finance and Economics*, Vol. 37 (4), pp. 317–333.