

Cajias, Marcelo; Freudenreich, Philipp; Freudenreich, Anna

**Article — Published Version**

## Exploring the determinants of real estate liquidity from an alternative perspective: censored quantile regression in real estate research

Journal of Business Economics

**Provided in Cooperation with:**

Springer Nature

*Suggested Citation:* Cajias, Marcelo; Freudenreich, Philipp; Freudenreich, Anna (2020) : Exploring the determinants of real estate liquidity from an alternative perspective: censored quantile regression in real estate research, Journal of Business Economics, ISSN 1861-8928, Springer, Berlin, Heidelberg, Vol. 90, Iss. 7, pp. 1057-1086,  
<https://doi.org/10.1007/s11573-020-00988-w>

This Version is available at:

<https://hdl.handle.net/10419/288609>

**Standard-Nutzungsbedingungen:**

Die Dokumente auf EconStor dürfen zu eigenen wissenschaftlichen Zwecken und zum Privatgebrauch gespeichert und kopiert werden.

Sie dürfen die Dokumente nicht für öffentliche oder kommerzielle Zwecke vervielfältigen, öffentlich ausstellen, öffentlich zugänglich machen, vertreiben oder anderweitig nutzen.

Sofern die Verfasser die Dokumente unter Open-Content-Lizenzen (insbesondere CC-Lizenzen) zur Verfügung gestellt haben sollten, gelten abweichend von diesen Nutzungsbedingungen die in der dort genannten Lizenz gewährten Nutzungsrechte.

**Terms of use:**

*Documents in EconStor may be saved and copied for your personal and scholarly purposes.*

*You are not to copy documents for public or commercial purposes, to exhibit the documents publicly, to make them publicly available on the internet, or to distribute or otherwise use the documents in public.*

*If the documents have been made available under an Open Content Licence (especially Creative Commons Licences), you may exercise further usage rights as specified in the indicated licence.*



<https://creativecommons.org/licenses/by/4.0/>



# Exploring the determinants of real estate liquidity from an alternative perspective: censored quantile regression in real estate research

Marcelo Cajias<sup>1</sup> · Philipp Freudenreich<sup>2</sup> · Anna Freudenreich<sup>2</sup>

Published online: 3 June 2020  
© The Author(s) 2020

## 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 seven largest 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 between cities but also within a city, each of the seven cities is considered individually. 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 direction of the regression coefficients for the different liquidity quantiles are found. Furthermore, both the magnitude and direction of the impact of an explanatory variable on the liquidity, differ between the cities. To the best of the authors' knowledge this is the first paper, to apply censored quantile regressions to liquidity analysis of 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.

**Keywords** Residential · Housing · Liquidity · Censored quantile regression · Time on market

---

**Electronic supplementary material** The online version of this article (<https://doi.org/10.1007/s11573-020-00988-w>) contains supplementary material, which is available to authorized users.

---

✉ Philipp Freudenreich  
[philipp.freudenreich@irebs.de](mailto:philipp.freudenreich@irebs.de)  
Marcelo Cajias  
[marcelo.cajias@patrizia.ag](mailto:marcelo.cajias@patrizia.ag)  
Anna Freudenreich  
[anna.freudenreich@irebs.de](mailto:anna.freudenreich@irebs.de)

<sup>1</sup> Patrizia Immobilien AG, Fuggerstrasse 26, 86150 Augsburg, Germany

<sup>2</sup> International Real Estate Business School, University of Regensburg, Universitaetsstrasse 31, 93053 Regensburg, Germany

**JEL Classification** C31 · C34 · R21 · R32

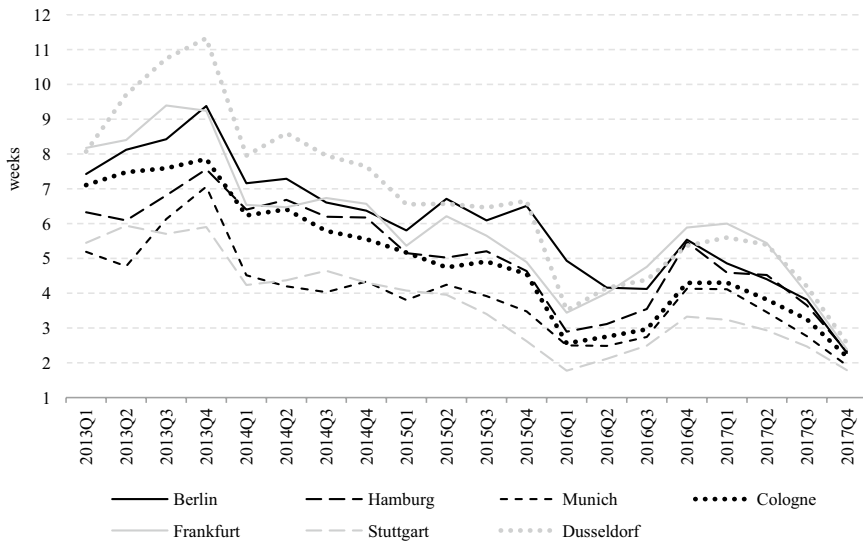
## 1 Introduction

The concept of asset liquidity in the residential rental market is somewhat blurred. In the investment market, asset liquidity traditionally measures the time it takes the owner to turn an asset into cash. In 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 to the market until the signing of the rental contract. In this paper, liquidity is defined as the inverse of the time on market (TOM) of rental dwellings, i.e. the higher the time on market, the lower the liquidity.<sup>1</sup> Whether the letting process is quick or slow, depends mainly, on the initial asking rent, the structural quality and location of the asset, demand for space, and the overall market conditions. A detailed understanding of these major drivers of marketing time of rental dwellings is the objective of this paper. To extract more insightful information from the available data, the study introduces a modelling technique which is new to the field and uses the results to validate the outcome of the most established approach to measure liquidity, which is the Cox (1972) proportional hazards model (PHM).

This paper focuses on the residential rental markets of the seven largest German cities (descending order by population): Berlin, Hamburg, Munich, Cologne, Frankfurt, Stuttgart, and Dusseldorf. While for the methodological aim of the paper, these markets are subsidiary, it is useful to understand the German rental regulatory framework and gain some insights into the housing markets in order to follow the contextual aim of this specific study. By law, tenants in Germany usually have a 3 months cancelation period, for which reason a dwelling is typically brought onto the market before the tenant leaves.<sup>2</sup> With a national ownership rate of roughly 43% as of 2013, the first year of the sample period, clearly more than half of German households rent their homes and hence the rental market deserves closer attention. Voigtländer (2009), Bentzien et al. (2012), Lerbs and Oberst (2014), and Reisenbichler (2016) explain in detail the reasons for this distinctive market feature. The considered cities exhibit ownership rates that are far lower than the German average, ranging from 33% in Stuttgart to 16% in Berlin. Therefore, especially in these cities, the rental market plays a significant role. While six of the cities can be generalized as long-established markets, Berlin shows stronger growth rates in population, purchasing power, employment, rents and purchase prices for real estate, however, starting from lower absolute values in those figures. For decades, the western German cities have developed into trans-regional hubs for education and employment. Berlin has “only” been able to benefit from the nationwide improvement after the reunification of Germany in 1990, however, since then its positive development was substantial.

<sup>1</sup> Based on the definition of Wood and Wood (1985).

<sup>2</sup> Landlords can only terminate the contract to use the dwelling for themselves or for other household members.



**Fig. 1** Average time on market in weeks in the seven largest German cities. This figure displays the average time on market in weeks in the seven largest German cities from 2013 Q1 to 2017 Q4. The data consists of 482,196 observations of the residential rental market

Figure 1 illustrates the average time on market for the seven largest German cities extracted from the underlying dataset. Despite some up and down movements, the graph clearly shows a continuous decline in time on market within the last years. This development points to an increasingly strong demand on the rental market, as prospective tenants have to shorten their decision-making process because of high competition for an insufficient supply. This excess demand was recognized by Held and Waltersbacher (2015) who claimed that this comprises 272,000 newly constructed dwellings per year on a national level for the years 2015–2020, while the actual completion did not satisfy this demand.<sup>3</sup> In addition, the use of online platforms might lead to better informed market participants, resulting in a narrower price assessment and hence less market uncertainty and increased market efficiency.

Since the aim of this paper is to understand the drivers of demand for rental space by means of a well-established method and the introduction of a method new to the field of real estate, the first step is to explore time on market via the Cox PHM. Next, an advanced econometric approach is introduced. To the best of the authors' knowledge this study is the first, to investigate the determinants of time on market by applying a censored quantile regression (CQR) in real estate research. The CQR aims to explain the variation of time on market as a function of dwelling characteristics and other spatial and socioeconomic characteristics. The decisive feature

<sup>3</sup> According to the Federal Statistical Office (2018) dwelling completion was: 216,727 (2015), 235,658 (2016) and 245,304 (2017).

of the analysis is that CQRs 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 in terms of their means but might differ especially in their upper and lower tails. CQRs can quantify the impact of a covariate on the dependent variable for any quantile, compared to only the center of the population. The CQR, as an expansion of the survival regression analysis, is expected to yield more accurate estimations and to provide a more complete statistical analysis for understanding the factors driving liquidity and the underlying demand. The dataset consists of 482,196 observations on the rental markets of the seven largest German cities between 2013 Q1 and 2017 Q4. The study reveals that the magnitude and direction of impact of an explanatory variable on time on market differs between the cities. Hence, this implies the importance of analyzing each city individually and not the seven cities or even the German market as whole. Furthermore, the magnitude and direction of effect of an explanatory variable on the liquidity of a dwelling exhibits differences between time on market quantiles within a city. This implies that, the proportional hazards assumption, underlying the Cox PHM is violated for individual explanatory variables and cities and thus emphasizes the use of the CQR approach for the time on market analysis. 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. These findings should of course be of interest to current and future landlords, as they reveal both the characteristics of dwellings along the liquidity distribution (e.g. the existence of built-in kitchen), as well as the impact of a change in characteristics on the liquidity of dwellings (e.g. impact of installing a built-in kitchen). 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 increase the expected liquidity, e.g. install a built-in kitchen or change the floor plan to increase the number of rooms. Furthermore, the findings suggest that nationwide or even statewide policy measures might not be expedient to address the specific situation on the residential rental real estate market of a particular region, city or neighborhood.

The remainder of this paper proceeds as follows: The next Section provides a literature review. Section 3 describes the underlying econometric model, followed by a detailed description of the dataset and the descriptive statistics in Sect. 4. Estimation results are presented and discussed in Sect. 5. Section 6 concludes.

## 2 Literature review

Belkin et al. (1976) conducted one of the first empirical studies analyzing real estate liquidity for different market segments. They divide the market according to geographic areas, price segments and buyers' search space. By doing so, they analyze the relationship between time on market and the spread between listing price and selling price, using ordinary least squares (OLS) estimation techniques. They find essential differences between market segments. Especially in

high-price segments, deviations from the initial list price had a more pronounced effect on time on market. The determinants of time on market considering different price segments have comprehensively been analyzed in the literature. Kang and Gardner (1989) found that the impacts on time on market do vary in magnitude between 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 segment, whereas this effect is insignificant for houses in the low- and high-price segments. Allen et al. (2009) also use a multi-step procedure to analyze the relationship between asking rent and time on market in the rental market for single-family residential rental listings. Based on asking rents, the sample is divided into three price segments. They find that underpricing of asking rents and time on market move in the same direction in all price segments, although the effect is stronger in the medium- and high-price segment. In further studies real estate markets have been segmented according to the number of rooms, the number of units in a structure, the geographical region, the property type or by market cycle, respectively. The conclusion emerging from these findings is that market segmentation seems to make a valuable contribution to our understanding of liquidity patterns.

More closely related to the present study is the article of Turnbull and Dombrow (2006), given that they divide their sample into low-, medium- and high-liquidity segments. 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 spatial competition variables on time on market, vary between the different liquidity segments.

Based on these findings, the following hypotheses are deduced and investigated in this paper:

1. Difference between quantiles: The direction and magnitude of the effect of covariates on real estate liquidity are not equal for and vary across low, medium and high-liquidity segments. If this is the case, the assumption of proportional hazards underlying the Cox PHM would be violated, justifying the need for an approach able to deal with heterogeneous effects.
2. Difference between cities: It is hypothesized that the direction and magnitude of the impact of the covariates on time on market vary across the seven cities. If this is true, an approach jointly capturing the residential rental markets of the seven cities or even regarding a national market might lead to the loss of important information.

Nowadays, the most popular model for the estimation of duration data is the Cox PHM. Also commonly used is the accelerated failure time (AFT) model.

However, in econometric terms, this paper differs substantially from the preceding studies, as 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 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 in terms of their means but might differ especially in their upper and lower tails. 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.<sup>4</sup> In the real estate literature, more precisely in the area of hedonic pricing, QRs have been applied by Zietz et al. (2008) and Liao and Wang (2012), among others. Thom-schke (2015) for example used the method for 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) initially, as well as Lüdemann et al. (2006), Schmillen and Möller (2012) among others, have lately applied CQRs. Conceptually the present analysis is highly related to Lüdemann et al. (2006). In particular, the CQR method used in this paper goes back to Koenker and Biliás (2002).

### 3 Econometric approach

#### 3.1 Cox proportional hazards model

Without doubt, the leading model for the analysis of survival data is the Cox PHM. This model is used for exploring the determinants of the duration of an event or elapse of time. For example, 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 time elapse that a dwelling requires to change its status from being offered on the market to being out of the market in weeks, i.e. time on market. For understanding and estimating survival data, two main functions are necessary: the survival function

---

<sup>4</sup> For survival analysis, see e.g. Crowley and Hu (1977), Koenker and Geling (2001); for medical research, see e.g. Beyerlein et al. (2008), Wehby et al. (2009); for financial economics, see e.g. Bassett Jr. and Hsiu-Lang (2002); for environmental research, see e.g. Hendricks and Koenker (1992); for labour economics, see e.g. Buchinsky (1994, 1995).

$S(t)$  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

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

with  $f(x)$  being the probability density function (p.d.f.) of the time until the event. The hazard function  $\lambda(t)$ , by 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

$$\lambda(t) = \lim_{\Delta t \rightarrow 0} \frac{P(t \leq T \leq t + \Delta t | t \leq T)}{\Delta t}. \quad (2)$$

The survival function expresses the probability of a dwelling staying on the market, while the hazard function measures the risk of the same dwelling leaving the market. The Cox regression for a specific observation  $i$  is given as

$$\lambda_i(t|x_i) = \lambda_0(t) \exp(-x_i' \beta) \quad (3)$$

where  $x$  is a vector of covariates (without the constant),  $\beta$  is a vector of parameters and  $\lambda_0(t)$  is the non-negative baseline hazard. The Cox PHM requires no specification of the functional form of the baseline hazard  $\lambda_0(t)$ . However, it assumes proportional hazards, meaning that the hazard function is a constant function of time. The elapse of time that a dwelling is offered on the market corresponds to an event that might be right-censored. Right-censoring arises when the landlord does not change the status of the dwelling in the Multiple Listing Services (MLS) database, or the dwelling is still being offered on 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  $k=7$  cities according to

$$h(t) = \exp(X\beta) + u \quad (4)$$

The hazard function  $h$  of the time on market  $t$  depends on the covariate matrix  $X$  plus an iid error  $u$ .

### 3.2 Cox proportional hazards model versus censored quantile regression

While the Cox PHM is the most common tool for explaining time on market in social sciences, natural sciences and real estate studies, new techniques have been developed to account for conditional survival functions across different levels of 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. In contrast, the quantile approach estimates different coefficients for different quantiles of the population. A



traditional example when explaining quantile regression is 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 CQR takes the different quantiles of the response, i.e. long-term and short-term unemployment into consideration and estimates several equations with 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. For the analysis of duration data CQRs are estimated to overcome the issue of right-censoring. This paper is the first to introduce the CQR to the context of real estate liquidity, to the best of the authors' knowledge. The Cox PHM has an advantage over the CQR regarding the issues of competing risks, time-varying covariates and unobserved heterogeneity. However, the CQR does not come with the cost of the proportional hazards assumption that needs to be empirically proven. Hence, misspecification of the model is less likely. The model requires only very weak assumptions on the error term, for example homoskedasticity of the error terms is not required, and thus is more robust against non-normal errors, outliers and misspecification of the error term. Moreover, Powell (1986) demonstrates that under appropriate conditions for a certain value of  $\tau$ , the censored regression quantile estimator  $\beta^\tau$  is consistent and asymptotic normality is proven, if the appropriate assumptions hold for each  $\tau \in \{\tau_1, \dots, \tau_J\}$ , irrespective of the distribution of the error term. Furthermore, the interpretation of regression results is straight forward. In summary, the CQR yields a robust and more flexible alternative for the estimation of parametric and semiparametric duration models and provides a very comprehensive statistical analysis.

### 3.3 Quantile regression model

The origin of the QR model extends back to Koenker and Bassett Jr. (1978). It is a location model estimating the relationship between the covariate matrix  $X$  and the dependent variable  $y$ , conditional on the quantile  $\tau$  of  $y$ . The quantile  $\tau \in (0, 1)$  is defined as the value of  $y$  that separates the observations into the fraction  $\tau$  below and the fraction  $1 - \tau$  above.

Applying the log-transformation of the time on market  $T_i$  of a particular dwelling  $i$ , according for example, to 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

$$\ln T_i = x_i' \beta^\tau + u_i^\tau. \quad (5)$$

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

$$\text{Quant}_\tau(\ln T_i | x_i) = x_i' \beta^\tau, \quad (6)$$

where  $Quant_{\tau}(lnT_i|x_i)$  represents the  $\tau$ th conditional quantile of  $lnT_i$  given a  $k \times 1$  vector of covariates  $x_i$ . The conditional quantile of the iid error term  $u$ ,  $Quant_{\tau}(u^{\tau}|x)$ , is 0.

### 3.4 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, which addresses fixed censoring, is applied. For CQRs 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 only 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  $\hat{\beta}^{\tau}$  is the value of  $\beta^{\tau}$  solving the minimization problem of the distance function

$$Q_N(\beta; \tau) \equiv \frac{1}{N} \sum_{i=1}^N \rho_{\tau}(lnT_i - \min(x_i' \beta^{\tau}, C_i)). \quad (7)$$

Thus,  $x_i' \beta^{\tau}$  is censored from above at the upper threshold  $C_i$ . The “check-function”  $\rho_{\tau}(u)$  is defined as:

$$\rho_{\tau}(u) = \begin{cases} \tau * |u| & u \geq 0 \\ (1 - \tau) * |u| & u < 0 \end{cases}. \quad (8)$$

$\tau * |u|$  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  $\hat{\beta}^{0.5}$ . 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 in an explanatory variable by a marginal unit, ceteris paribus, prolongs or shortens the time on market by  $\left[ \left| 1 - \exp(\hat{\beta}^{\tau}) \right| * 100 \right] \%$  keeping time on market in the same quantile. A prolongation of the time to event occurs if the respective coefficient  $\beta^{\tau}$  is positive, i.e. the hazard ratio  $\exp(\hat{\beta}^{\tau})$  is greater than 1, and a reduction of the time to event if  $\beta^{\tau}$  is negative, i.e. the hazard ratio is smaller than 1. For example, if  $\beta_{refurbished}^{0.2}$  was 0.077, the interpretation would be as follows: Dwell-

ings of the 0.2 quantile which are newly renovated, *ceteris paribus*, stay 8.1% longer on the market, than dwellings which are not newly renovated.

## 4 Data and descriptive statistics

The estimation sample is composed of three merged datasets, containing information from 482,196 observations of single- and multi-family rental dwellings in the seven largest German cities from the first quarter of 2013 to the fourth quarter of 2017. Observations regarding student housing, affordable housing, temporary housing and retirement housing have not been incorporated into the dataset. The variables, their units and sources can be found in the Online Appendix in Table A1. Information on the rental dwellings was provided from empirica systems (<http://www.empirica-systeme.de>) which collects georeferenced real estate data from more than 100 German Multiple Listing Systems (MLS) such as ImmoScout, Immonet or Immowelt but also regionally focused marketplaces and newspapers for the whole German market. As the market leader of real estate data for Germany, empirica has an own proprietary algorithm that identifies duplicates and harmonizes the sample of big data. The variable of major interest, time on market, is defined as the number of weeks a dwelling was listed in the MLS calculated by the start and end date according to Benefield and Hardin (2015) for example. The asking rent is defined in absolute terms measured in euros per month. The common issue of reverse causality between time on market and rent, discussed by Yavas and Yang (1995) among others, is not addressed in this paper due to a lack of contract rents on the German residential market. In this paper the asking rent is set by the landlord at the beginning of the data generating process (DGP). The time on market is established as the time the dwelling is offered in the MLS given the initially set asking rent. The asking rent operates as a “take it or leave it option” to the tenant and thus rent negotiations are not considered. On the residential rental market this assumption is plausible as negotiations about the monthly rental payments, especially in the overheated markets of the seven major cities considered in this paper, are rather an exception. Hence, this approach provides a descriptive analysis, however, is not able to draw causal conclusions. Since the data is georeferenced, two spatial gravity indicators, measuring the Euclidian distance of each dwelling to the geographical centroid of the ZIP and NUTS 3 polygon in kilometers, are included. NUTS 3 regions correspond to “the nomenclature of territorial units for statistics”, which is a hierarchical system for dividing the economic territory in Europe. The NUTS 3 regions cover small regions similar to counties or administrative districts. In the sample, each city represents one NUTS 3 region and therefore, the distance to the NUTS 3 centroid describes the distance to the geometric city center. The socioeconomic variables of purchasing power per household and the number of households at the ZIP code level, are obtained from the “Gesellschaft für Konsumforschung” (GfK). The population density per km<sup>2</sup> in a ZIP code area is calculated in ArcGIS. The final source is Thomson Reuters Eikon, providing the 10-year interest rate for housing loans as a macro variable. The analysis focusses on seven time

**Table 1** Descriptive statistics of selected variables

Variable	Mean	Median	Std. Dev.	Q1	Q99
Panel A: Berlin, ownership rate: 15.61%					
Asking rent	692.37	549.94	470.12	233.42	2500.35
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	11,997	3645.74	3890	20,434
Purchasing power of HHs in ZIP	35,272.76	34,352.80	4954.13	27,548.31	52,669.38
Population density	3899	3542	2398.80	61	7908
N = 180,858					
Panel B: Hamburg, ownership rate: 21.14%					
Asking rent	795.44	658.55	483.79	254.11	2624.01
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	11,051	10,677	3458.73	2143	17,979
Purchasing power of HHs in ZIP	43,559.36	42,779.87	7670.23	32,723.47	63,894.32
Population density	3800	3757	2342.76	45	7560
N = 101,008					
Panel C: Munich, ownership rate: 25.23%					
Asking rent	1209.59	1034.08	721.71	379.96	3960.00
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	11,459	12,074	3241.20	3573	16,896
Purchasing power of HHs in ZIP	55,942.79	54,728.80	6132.63	45,586.18	69,752.31
Population density	4173	4463	2184.26	253	7933
N = 47,394					
Panel D: Cologne, ownership rate: 27.42%					
Asking rent	720.55	639.6	369.88	250.00	2039.92
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	13,521	3594.13	6176	20,561
Purchasing power of HHs in ZIP	45,466.38	44,370.20	5748.89	34,685.48	58,827.02
Population density	3675	3390	2243.07	395	7598
N = 47,527					
Panel E: Frankfurt, ownership rate: 20.67%					
Asking rent	1012.81	850.20	634.55	299.98	3499.85
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	11,669	4351.39	1546	20,945

**Table 1** (continued)

Variable	Mean	Median	Std. Dev.	Q1	Q99
Purchasing power of HHs in ZIP	47,528.31	46,692.74	6663.02	37,419.27	76,088.03
Population density	3930	4194	2098.80	146	7785
N = 41,446					
Panel F: Stuttgart, ownership rate: 32.92%					
Asking rent	910.71	775.20	501.26	270.98	2749.82
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	10,927	3098.97	1104	15,899
Purchasing power of HHs in ZIP	47,058.05	46,374.77	4440.57	40,041.62	61,972.64
Population density	3507	3353	1766.72	254	7404
N = 17,967					
Panel G: Dusseldorf, ownership rate: 24.08%					
Asking rent	762.86	630.00	486.69	240.00	2579.00
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	9726	9703	3021.54	2721	15,045
Purchasing power of HHs in ZIP	47,869.27	46,140.99	5851.67	40,382.03	65,472.45
Population density	3999	3913	2370.53	23	7906
N = 45,996					

This table reports selected descriptive statistics, the ownership rate 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 2** Degree of censoring and average time on market in weeks per quantile

	Degree of censoring (%)	Q20	Q30	Q40	Median	Q60	Q70	Q80
Berlin	6.7	0.6	1.1	1.7	2.6	3.8	5.6	8.8
Hamburg	6.0	0.5	0.9	1.4	2.1	3.1	4.7	7.4
Munich	5.2	0.4	0.7	1.1	1.7	2.4	3.5	5.3
Cologne	4.9	0.6	1.1	1.6	2.4	3.4	4.8	7.4
Frankfurt	6.7	0.7	1.2	1.9	2.7	4.0	5.7	8.7
Stuttgart	3.7	0.3	0.7	1.1	1.6	2.3	3.5	5.4
Dusseldorf	7.0	0.9	1.5	2.2	3.2	4.6	6.4	10.0

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

on market quantiles from the 0.2 to the 0.8 quantile. The top and bottom end of the time on market distribution display extremely high and extremely low values respectively and hence are not the subject of this analysis.

Table 1 shows the descriptive statistics for the variables of interest for each of the seven cities. Comparing the seven cities shows that the mean time on market is fairly low in Stuttgart and Munich and relatively high in Dusseldorf, Berlin and Frankfurt. With 3.73 weeks, the mean time on market is lowest in Stuttgart and highest in Dusseldorf with 6.6 weeks. Table 2 shows the time on market for each of the selected quantiles and for all seven cities. It confirms that the time on market is lowest in Stuttgart and Munich and highest in Dusseldorf, Frankfurt and Berlin along the entire distribution. Furthermore, Table 2 exhibits the degree of censoring present in the data for each city. The degree of censoring can be interpreted as the percentage of dwellings that have not been taken out of the MLS until the cut of date. For example, in Berlin 6.7% of all dwellings that appeared in the MLS from 2013 Q1 to 2017 Q4 still remain in the MLS after the last day of the observation period.

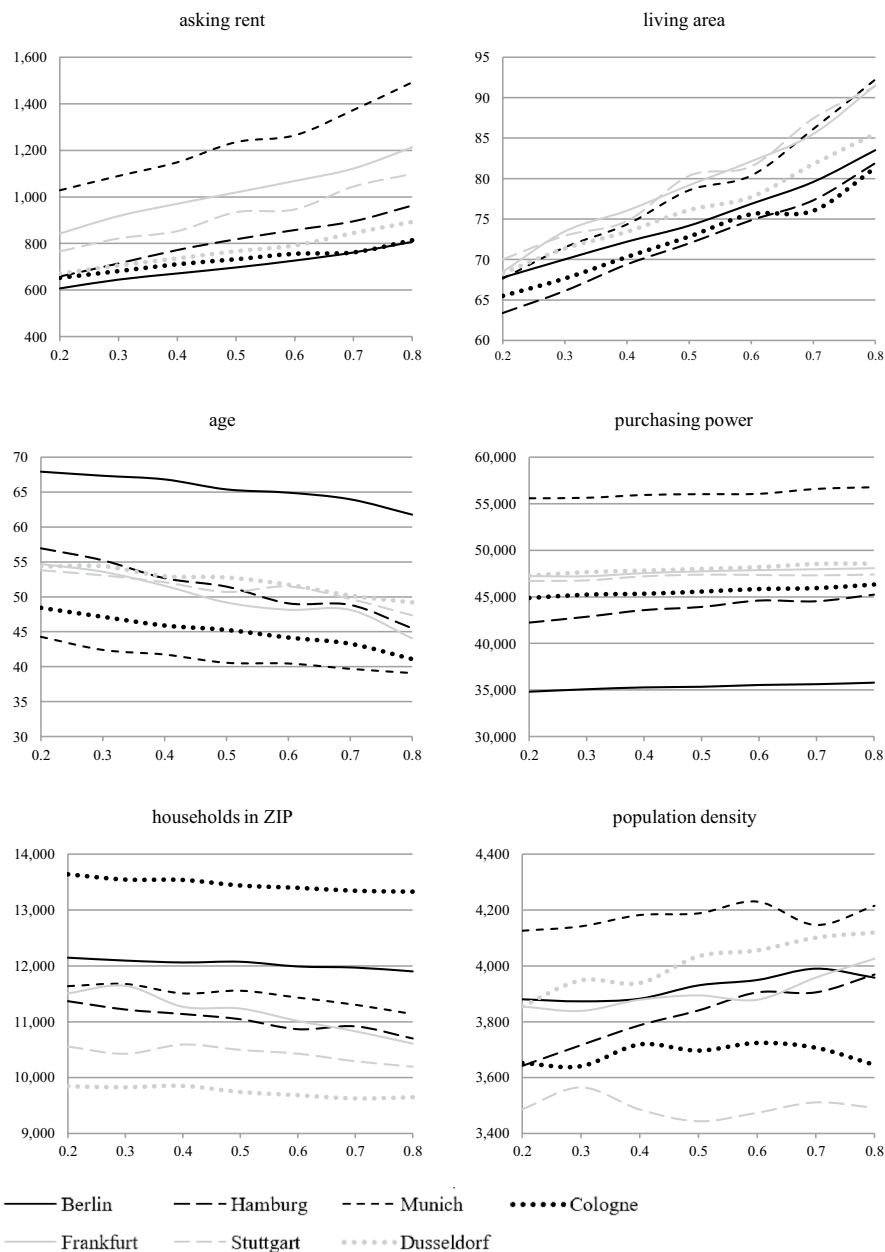
The covariate of asking rent is lowest in Berlin with a mean of 692.37 € and highest in Munich with a mean of 1209.59 €. Interestingly, these cities respectively exhibit the lowest and highest mean purchasing power of 35,272.76 € in Berlin and 55,942.79 € in Munich. Furthermore, they are characterized having on average the oldest and youngest buildings respectively. This indicates high development activity in the more recent past in Munich. These variables reveal interesting differences between the cities and give rise to an individual investigation of each city.

The aim of the study is to investigate the impact of changes in the explanatory variables on the time on market of rental dwellings, divided by their liquidity level. Hence, it is of special interest whether there are patterns in the dwelling characteristics, which might explain the affiliation to a respective liquidity quantile. Figure 2 shows that across all seven cities, the dwellings in the most liquid quantile, 0.2 (Q20), are on average the cheapest, the smallest and the oldest. Furthermore, they are located in ZIP code areas with the lowest purchasing power, but a relatively large number of households. In contrast to the most liquid quantile, dwellings assigned to the 0.8-quantile (Q80) display on average 33.71% higher rents, are 25.3% larger, 12.13% younger, 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. These insights based on pure analysis of descriptive statistics give rise to the assumption that the level of time on market strongly depends on the characteristics of a dwelling. Hence, this supports the CQR approach to get a deeper understanding of the determinants of time on market.

## 5 Results

### 5.1 Results of the cox survival regression

In a first step, covariates raising or lowering the time on market of rental dwellings of the seven largest German cities have been considered. The results of the Cox survival regressions are presented in Table 3. Since the results are displayed as hazard



**Fig. 2** Descriptive development of selected covariates for time on market quantiles. This figure displays descriptive statistics for selected variables across seven time on market quantiles for the seven largest German cities. The time on market quantiles are labelled on the x-axis and the respective diagram title is labelled on the y-axis. Altogether, the data consists of 482,196 observations on the residential rental market. The sample period is 2013 Q1 to 2017 Q4

**Table 3** Results of the Cox survival regression

Hazard rates and robust standard errors	Berlin	Hamburg	Munich	Cologne	Frankfurt	Stuttgart	Dusseldorf
Housing characteristics—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
Housing characteristics—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 elevator	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***



Table 3 (continued)

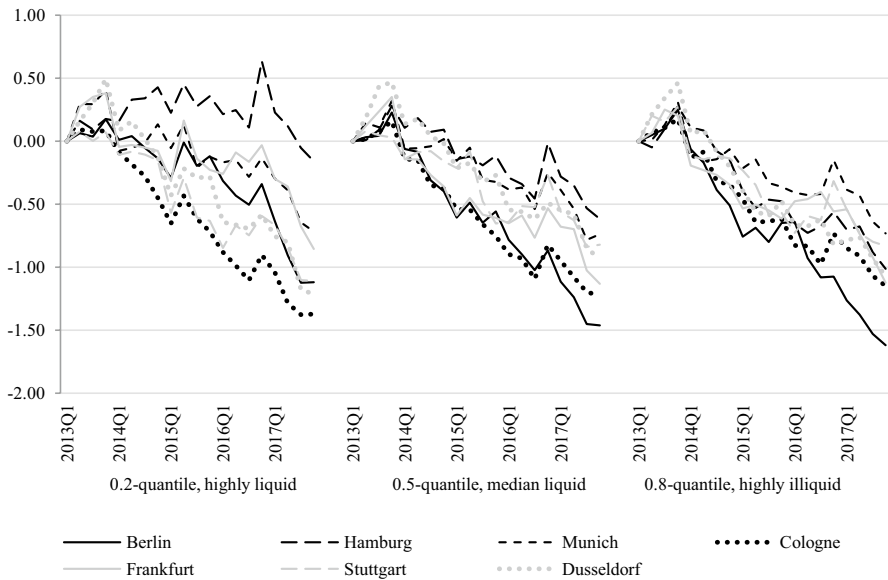
Hazard rates and robust standard errors	Berlin	Hamburg	Munich	Cologne	Frankfurt	Stuttgart	Dusseldorf
Refurbished	0.983 0.013	0.907 0.014***	0.967 0.014**	0.987 0.014	0.938 0.017***	0.960 0.021*	0.930 0.012***
Spatial variables							
Longitude	0.683 0.132***	0.597 0.236***	1.346 0.184	0.226 0.296***	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 NUTS 3 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

**Table 3** (continued)

Hazard rates and robust standard errors	Berlin	Hamburg	Munich	Cologne	Frankfurt	Stuttgart	Dusseldorf
Fixed effects							
Construction dummies	Included						
Quarterly dummies	Included						
Intercept	Included						
Spatial adjusted (Lin-Wei) standard-errors	Considered						
$R^2$ -concordance	65.4%	66.7%	67.5%	64.3%	65.4%	66.7%	64.4%

This table displays the hazard ratios and robust standard errors of a Cox PHM of the time on market of a dwelling in weeks for housing, spatial and socioeconomic variables, as well as the effective 10-year interest rate and various fixed effects. In order to obtain robust standard errors, the Lin-Wei (Lin and Wei 1989) procedure, similar to the sandwich covariance matrix of the OLS estimator, is applied and extended by a spatial component, i.e. the ZIP codes. 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

rates, a rate larger than one decreases the time on market and 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). The seller faces the trade-off between the time the dwelling is on the market and the price he might receive, as mentioned by Anglin et al. (2003) for the real estate transaction market. According to that, the result is not surprising, that a higher asking rent leads to lower liquidity of dwellings. Since the densely populated cities regularly display an 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. Miller (1978) postulates that housing attributes, defining its attractiveness, are important determinants of time on market. An increase in living area significantly increases liquidity in two of seven cities, namely Berlin and Hamburg. As Hamburg is the city with the smallest average living area in this sample, people might desire larger apartments and hence take them off the market faster. Cologne is the only city where living area has a prolonging impact on time on market, however, is not significant. A segmentation into liquidity quantiles might be useful in order to consider the impact and its significance of, for instance, the living area for each quantile, as it might be the heterogeneity within the cities that leads to insignificant effects. The number of rooms in a dwelling, in contrast, has a statistically significant positive effect on liquidity for all seven cities. Hence, it seems that in most of these seven highly demanded regions, a larger dwelling in general is not favorable. A dwelling with a larger number of rooms, however, can provide more living space for more people and thus is highly demanded and consequently off the market more quickly. Surprisingly, the marketing time is *ceteris paribus* shorter, the older the dwelling, and is longer for newly built and refurbished ones. In Berlin the marketing time is longer the older the building, however, the effect is not statistically significant. This might be due to the fact, that in this sample Berlin exhibits by far the highest average building age. An increase in the distance to the NUTS 3 centroid, which is used as a proxy for the city center, is associated with a longer time on market. This seems obvious, as regions farer away from the city center are less demanded. The distance to the ZIP code centroid shows almost no statistically significant impact. The coefficients of each socioeconomic factor are not significant for most cities, again emphasizing the importance of considering each time on market quantile separately. A higher purchasing power results on average 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. Among the seven cities, Berlin and Hamburg are the cities with the lowest average purchasing power, hence this result points to the fact that in Berlin and Hamburg people would spend additional purchasing power on the search and matching process that would shorten marketing time. While an increase in the number of households significantly reduces time on market in Frankfurt and Dusseldorf, the population density has a significant prolonging effect on time on market in Hamburg, Munich, and Frankfurt. It should be noted that neither socioeconomic variable has an impact on time on market in Cologne and Stuttgart. The Pseudo- $R^2$  based on Kendall's tau, which measures the concordance between estimated survival time and the observed survival time for the



**Fig. 3** Estimated quantile regression coefficients of the quarterly dummy variables. The figure displays the development of the coefficients  $\beta_k^\tau$  of the quarterly dummy covariate for the 0.2-, 0.5- and 0.8-quantile for each of the seven cities  $k$  over time. Altogether, the data consists of 482,196 observations on the residential rental market. The sample period is 2013 Q1 to 2017 Q4

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 be interpreted directly or compared to the usual  $R^2$  calculated for OLS regressions.

## 5.2 Results of the censored quantile regression

In a second step, the same regressors as for the Cox survival regressions are used to estimate CQRs, in order to gain deeper insight into the time on market quantiles. Therefore, for each city, the sample of dwellings was divided into time on market quantiles from the 0.2- to the 0.8-quantile. The results for the covariates of interest are shown in Figs. 3, 4, 5 and 6. In Figs. 4, 5 and 6 the black lines plot the estimated coefficients, the grey lines display the standard deviations. Each plot shows the development of a coefficient  $\beta_k^\tau$  over the liquidity quantiles  $\tau$  for each of the seven cities  $k$ . A positive statistically significant coefficient increases time on market, thus decreasing liquidity, while a negative statistically significant coefficient has the opposite effect. The main effects, divided into quarterly factors, housing characteristics, spatial gravity variables, and socioeconomic characteristics, are reported below. The crucial point is that, contrary to the traditional regression models, the effect of a change in the covariate holds for the same quantile  $\tau$ , rather than across quantiles.<sup>5</sup>

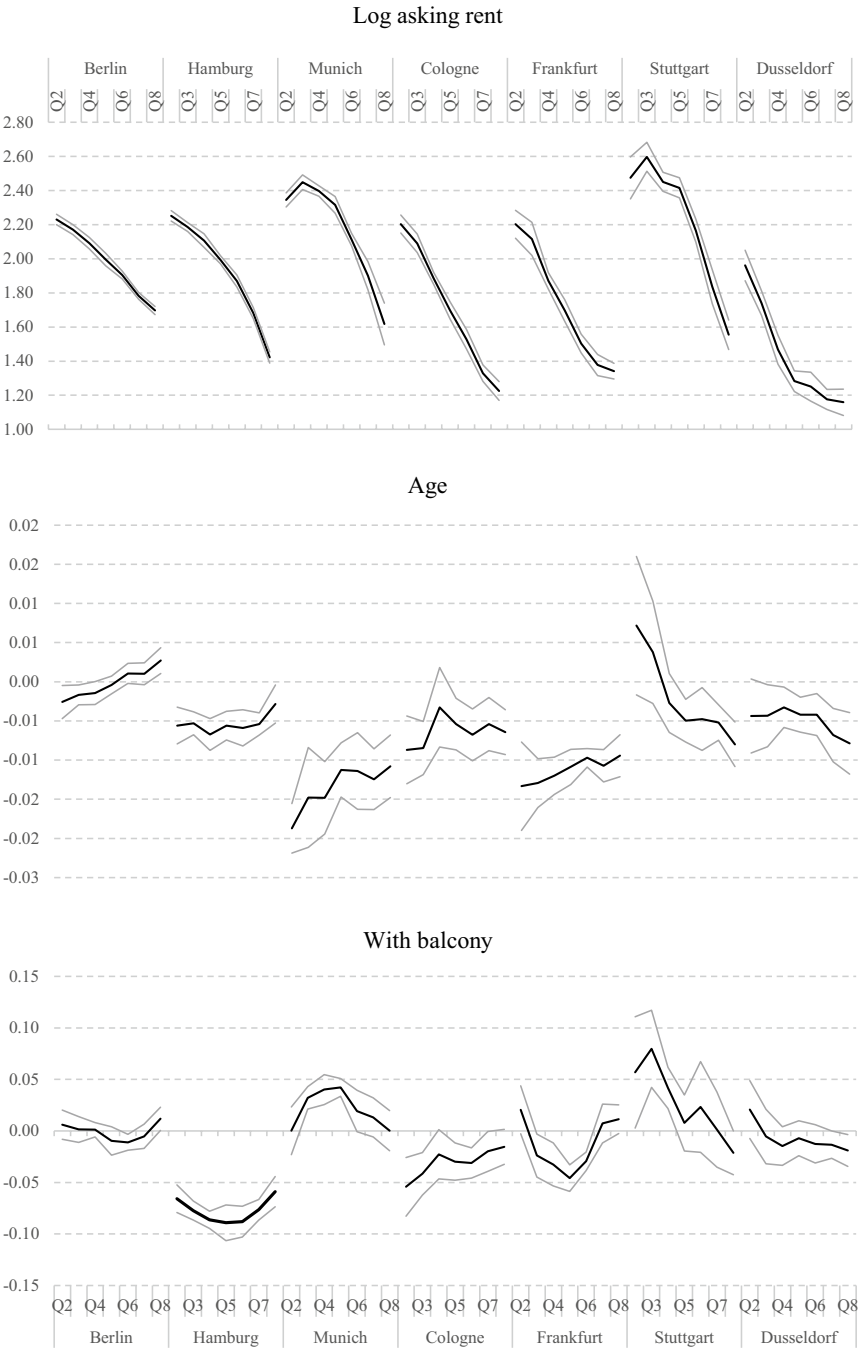
<sup>5</sup> Table A2 in the Online Appendix shows the estimated coefficients for all cities and quantiles.

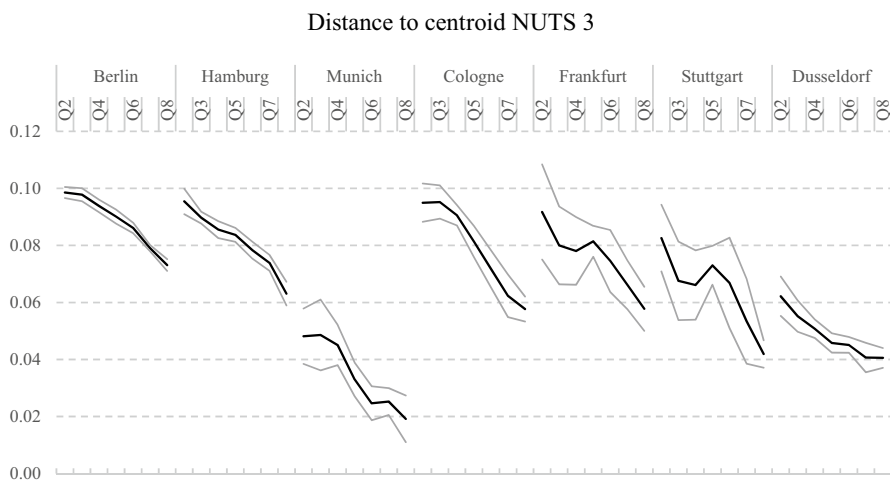
**Fig. 4** Estimated quantile regression coefficients of selected housing variables. The figure displays the development of the coefficients  $\beta_k^\tau$  of the housing covariates across the liquidity quantiles  $\tau$  for each of the seven cities and the respective confidence intervals. The impact 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

### 5.3 Quarterly time effects

The considered period is characterized by low interest rates, high migration to Germany, specially to the metropolises, and additionally far too low housing supply in these cities. Consequently, vacancy has mostly been diminishing, and real estate prices as well as rents 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 in the estimation equation to capture these developments over time and its impact on the time on market.<sup>6</sup> This time trend can be observed in Fig. 3 for each city and for quantiles representing high, medium and low liquidity. The base quarter is the first quarter of 2013, so that all changes are with respect to this basis. All in all, time on market has been decreasing compared to the base quarter for all time on market quantiles and for each city. However, the magnitude and direction of change differs between the cities and different time on market quantiles. In the first quarter of the year, time on the market is usually relatively low for rental dwellings. Hence, in 2013 and partly in the beginning of 2014 time on market has been increasing relative to the base quarter in all cities and all time on market quantiles. Over the course of 2014 time on market started to decrease relative to 2013 Q1 in six of the seven cities and all time on market quantiles. The exception is Hamburg, where time on market started to decrease relative to 2013 Q1 later than in all other cities and it is striking that the date of decline is documented later for more liquid quantiles. This supports the importance of analyzing all cities separately. Furthermore, it emphasizes the hypothesis that the time on market development within a city might be different for distinct time on market quantiles and thus the quantiles should be analyzed separately. By far the most pronounced decline in time on market experienced Berlin in the 0.5 and 0.8 quantile. In the 0.2 quantile the development in Berlin is rather average compared to the other cities as a sharp decline only starts in 2017. Hence, it seems that in Berlin relative to 2013 Q1 the demand for rental dwellings was higher in the 0.5 and 0.8 quantile. As described in Sect. 4. Data and Descriptive Statistics, the living area of dwellings in these quantiles is relatively larger than in the 0.2 quantile. A reason for this result might be the development of household members during the recent years. According to Vonovia and CBRE (2016), in 2014, the year when time on market started to decline compared to the base quarter, Berlin was amongst the cities with the highest percentage of single households and even had the highest percentage of one- and two-person-households together. Hence, the demand for larger dwellings was probably relatively small compared to the other cities. According to

<sup>6</sup> Of those quarterly effects, 82.0% are significant at the 10% significance level, while 74.5% are significant at the 5% significance level.



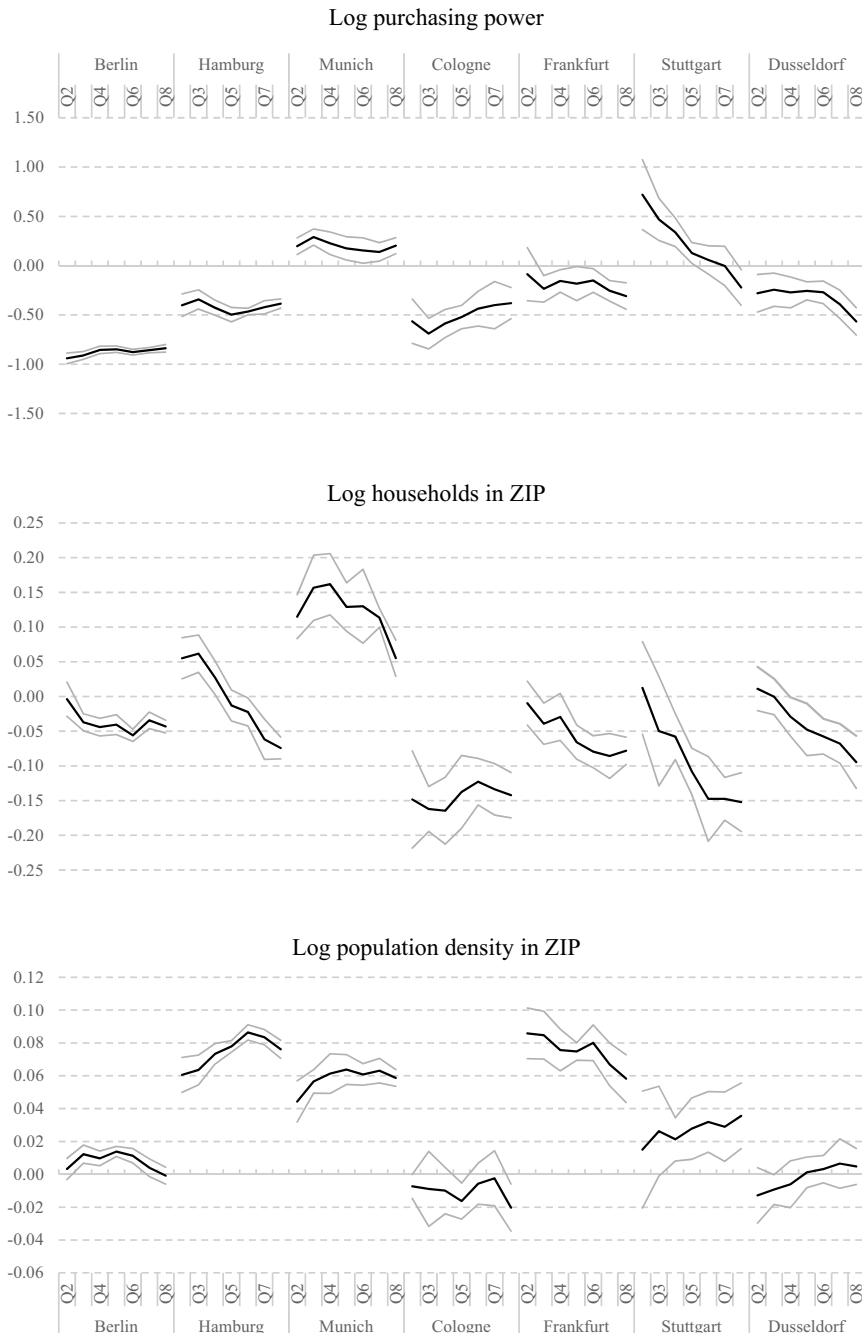


**Fig. 5** Estimated quantile regression coefficients of the spatial gravity variable. The figure displays the development of the coefficients  $\beta_k^\tau$  of the spatial gravity covariate distance to the NUTS 3 center across the liquidity quantiles  $\tau$  for each of the seven cities and the respective 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

data provided from Oxford Economics, over the observation period from 2013 to 2018 Berlin experienced an increase in population of 6.83%. Working population, however, has only increased by 4.94%. This is by far the most pronounced divergence of population and working population among the analyzed cities. It might be that relative to the other cities especially families with members younger than the working age, who demand relatively more living space, have moved to Berlin. Consequently, the demand and thus liquidity in Berlin have especially increased in the upper liquidity quantiles. Furthermore, during the observation period Berlin has experienced high migration inflow, a massive increase in real GDP and household disposable income combined with a decreasing unemployment rate. It seems that in Berlin, this favorable economic development resulted in higher demand for larger and more expensive dwellings, hence led to a strong liquidity decline at the upper end of the distribution. This again highlights the importance of the CQR approach.

## 5.4 Housing characteristics

It is possible to cluster the housing characteristics into three groups based on their impact on time on market. Unsurprisingly, asking rent 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 coefficients of first occupancy and renovation (besides the statistically insignificant 0.8 quantile in Cologne), as well as the equipment variable with bathtub, also show positive effects. Living area (besides the statistically insignificant 0.7 quantile in Dusseldorf) as well as number of rooms



**Fig. 6** Estimated quantile regression coefficients of the socioeconomic variables. The figure displays the development of the coefficients  $\beta_k^\tau$  of the socioeconomic covariates across the liquidity quantiles  $\tau$  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



show the opposite impact for all quantiles across the seven cities. For the other dwelling characteristics, the coefficients yield distinct effects, depending on the respective quantile and location of the dwelling. Coefficients of variables of interest, supporting the hypotheses stated in Sect. 2, are shown in Fig. 4.

Amentioned above, the covariate of asking rent is positive and statistically significant for each quantile of all cities. The magnitude of impact on time on market decreases with increasing time on market for all cities. This might be due to the fact that the more illiquid quantiles are on average characterized by larger living areas, more rooms and are further from the city center. With the population in the seven metropolises consisting to a great extent of young single households, larger dwellings are less in demand, so that the reaction of time on market to a rental price increase is less pronounced. Furthermore, higher time on market quantiles are also characterized by higher asking rents. Hence, an additional reason might be that a further rental increase starting from a relatively high rental level does not affect demand and liquidity that strong. The impact of an increase in asking rent is weakest for dwellings located in Dusseldorf. A reason for that might be that the overall socioeconomic and economic development has been relatively weak in Dusseldorf compared to the other cities. According to data provided from Oxford Economics, Dusseldorf experienced a relatively low increase in population, working population and household disposable income during the observation period. Though the vacancy rate has been decreasing during the last years, vacancy is still high in comparison to the other metropolises. Additionally, building completion has been increasing and is at a relatively high level, while the city exhibits the second largest purchasing power per household in this sample. These factors influence real estate prices and rents. Consequently, asking rents in Dusseldorf have been rising below-average relative to the other considered cities. Hence, these socioeconomic and economic factors dampen the impact on time on market. The strongest impact for the more liquid quantiles is found for Stuttgart and Munich and switches to Berlin for the 0.8-quantile. This is not surprising, as rents in Munich as well as Stuttgart are at a very high level compared to the other cities. Accordingly, further rental price increases substantially 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. While a ten percent increase in asking rent within the most liquid quantile results in an 8.5–10.8% higher time on market in Dusseldorf and Stuttgart, the impact decreases to between 4.9 and 7.3% higher time on market in Dusseldorf and Berlin for the most illiquid quantile.

The impact of an increase in living area as well as the number of rooms shows 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. 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. 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 analyzing each city individually.

In contrast to the uniform direction of impact of asking rent, 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. All statistically significant coefficients are negative, hence an increase in age reduces time on market, meaning that older dwellings are rented out faster. However, regarding the statistically insignificant coefficients as well, exhibits that within Berlin and Stuttgart the direction of effect changes along the time on market distribution. While in Berlin an increase in age leads to a longer time on market at the upper tail of the distribution, in Stuttgart this direction of effect is true at the lower tail. Consequently, the proportional hazards assumption underlying the Cox PHM cannot be verified for the covariate of age for all cities. This supports the CQR approach, where the effect of a covariate is analyzed for each quantile individually.

The direction of impact a change in certain equipment characteristics has on time on market does not only differ between the cities but also between the time on market quantiles within a city. For example, a balcony negatively influences time on market in Hamburg but positively in Munich in each quantile. In Berlin, Frankfurt and Dusseldorf the impact of a balcony changes its sign along the time on market quantiles, giving rise to the need of censored quantile regressions. The effect of an elevator on time on market is positive in Berlin and negative in Munich. Moreover, in Frankfurt the effect is significantly negative at the 0.2 quantile, however, significantly positive at the 0.6, 0.7, and 0.8 quantile. A similar picture is visible for the effect of a built-in kitchen. The impact on time on market is negative in Berlin, Munich, Cologne, Frankfurt, and Dusseldorf but positive in Hamburg for each quantile. In Stuttgart in contrast the impact of a built-in kitchen is significantly positive at the 0.2 quantile and significantly negative at the 0.6, 0.7, and 0.8 quantile. These binary variables strongly show the need for regarding each city separately as well as the importance of the CQR approach.

## 5.5 Spatial gravity variables

Since the study analyzes the seven largest German cities geographically each restricted by its NUTS 3 boundaries, the distance of a dwelling to the NUTS 3 center is of particular interest as it can be interpreted as its distance to the geographical city center.

Figure 5 shows that the effect of the distance to the geographical city center on time on market is positive and statistically significant for each city and each quantile. Hence, a higher distance to the center extends the time a rental dwelling is listed on the market. The effect is strongest for the highly liquid quantiles and becomes weaker for more illiquid ones. A possible explanation might be the fact that the most liquid dwellings are 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 1 km closer to or further from the center makes little difference. Across the whole distribution, the effect is highest in Berlin and lowest in Munich.

With on average 8.44 km, the rental dwellings in Berlin display on average the highest distance to the approximated city center. People in sprawling Berlin evidently have higher preferences for living closer to the city center compared to the other cities, especially Munich. A reason for this might be the location of popular residential areas right across the metropolitan area of Berlin, as well as the still relatively low asking rents in the city center compared to other cities. The distance to the ZIP code center is of less interest and has no explanatory power.

## 5.6 Socioeconomic characteristics

Each of the socioeconomic variables exhibits a quite different impact on time on market depending on the time on market quantile as well as the city.

An increase in purchasing power leads *ceteris paribus* to a shorter time on market in five of the seven cities. In these cities purchasing power drives the demand for residential real estate on the rental market and thus results in a lower marketing time. The impact is strongest in Berlin across all quantiles. Berlin shows by far the lowest average purchasing power. The substantial 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. In Munich, however, the effect of purchasing power is time on market-prolonging. Hence, this points to the importance of considering the cities separately. Furthermore, the effect in Stuttgart exhibits a different sign depending on the respective quantile. Consequently, this supports the use of CQRs. 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, reveals the huge heterogeneity within each city.

The effect of the number of households in a ZIP code area on time on market reveals a similar picture. While Berlin, Cologne, and Frankfurt show the expected demand effect, meaning that an additional household increases the demand for space and thus time on market decreases, in Munich the impact is again time on market-prolonging. In the three remaining cities the impact on time on market even switches from positive to negative, implying a violation of the proportional hazards assumption. The direction of impact of population density is also not consistent for all time on market quantiles in two of the seven cities. These findings strongly emphasize to analyze the seven cities separately and to permit the direction of the effect to change with time on market quantiles and hence support the preference of the CQR approach over the Cox PHM.

## 6 Conclusion

Academic research on the factors affecting the liquidity of dwellings on the rental market is scarce and for most countries non-existent. Over the last years, the improvement in computational power and the ability to gather and store large amounts of data opens new research fields, especially when it comes to the

application of advanced econometric methods. 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 dataset, comprising almost 500,000 observations across the seven largest German residential markets.

The main 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 between the liquidity quantiles. Hence, the proportionality assumption underlying the Cox proportional hazards model (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 proportional hazards regression, the CQR accurately detects these differences as it distinguishes between differential effects across quantiles. Furthermore, it is evident that the magnitude, significance, and direction of the impact of the covariates on time on market is quite different across the cities. These findings emphasize the importance of considering each city individually for a more detailed analysis and understanding of the respective rental real estate market. To summarize, the CQR applied on each city individually provides a more complete statistical analysis than commonly used approaches. The results of the individual city might depend on its economic and socioeconomic development during the last years. This overall development might include factors like population, working population, real GDP, household disposable income, unemployment rate, vacancy rate and construction activity. Additionally, the socio-demographic structure of the population regarding the age distribution, the distribution of the number of household members and the income distribution, among others, might play an important role.

With this detailed market assessment, landlords 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 their 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. While the study uses the rental market of the seven largest 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 time on market quantiles.

As addressed in Sect. 4. Data and descriptive statistics the covariate asking rent is defined as the rent initially set by the landlord. Rent negotiations as well as contract rents are not available for the German residential rental market. On the residential rental market, the assumption that asking rents operate as a "take it or leave it option" to the tenant seems to be plausible as rent negotiations, especially in the overheated markets of the seven largest cities considered in this paper, are rather an exception. Nevertheless, the availability of contract rents would further improve the findings of this paper and moreover would expand the range of possible research questions. In general, using asking rents might come at the cost of a potential upward bias of rents. A lower contract rent than the initially asked rent might be possible. This scenario seems most likely for dwellings with a long marketing time. Consequently, in this paper the results in the time on market quantiles at the

upper end of the distribution might be biased with respect to rents. The availability of contract rents would allow to estimate the impact of initial overpricing on time on market. Specifically, does initial mispricing significantly lead to longer marketing times? Furthermore, it would be interesting to investigate in which time on market quantiles the contract rent deviates from the asking rent. Another important covariate might be transaction costs. It could be investigated whether transaction costs (e.g. broker fees) lead to shorter or longer marketing times. It could be assumed, that landlords with higher transaction costs prefer to let a dwelling more quickly. That in turn might influence the initial price setting. Furthermore, incorporating further building characteristics, like the type of building (e.g. multi-family building, old building, detached building), the number of floors, and the floor level of the respective dwelling might alter the results. Especially, some of the binary housing attributes like with balcony or with elevator might yield more distinct results. Further research should also concentrate on a deeper analysis of an individual city. The city center might be specified not only geographically but rather in terms of an economic and sociodemographic city center. Furthermore, the proximity to local public transport and other amenities, tourism hotspots, commuting flows as well as the distance to adjacent urban areas should be included in the analysis. The type of neighborhood, school quality and noise, e.g. from a highway, should be considered as well. Including this mass of spatial variables, would lead to a more precise and detailed spatial mapping of a city. These variables would probably change the impact of the current variable “distance to the NUTS 3 center”. Moreover, the spatial variables are assumed to have a different effect on different dwelling and inhabitant types and thus in different time on market quantiles. Moreover, the impact of competition on the rental market, for example modelled as spatial concentration of simultaneously offered similar dwellings in a neighborhood, on time on market might be of further interest. The turnover of similar rental dwellings and the market tightness in a neighborhood might be an indicator of the bargaining power of landlords and tenants. As these competition variables are enormously associated with the rental price setting, they might influence the current results of the variable asking rent. This might depend on the respective location and type of dwelling. Hence, these results might be different for different time on market quantiles. 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 characteristics.

**Acknowledgements** Open Access funding provided by Projekt DEAL. 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.

**Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission

directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

## References

- Allen MT, Rutherford RC, Thomson TA (2009) Residential asking rents and time on the market. *J Real Estate Financ Econ* 38:351–365
- Anglin PM, Rutherford RC, Springer TM (2003) The trade-off between the selling price of residential properties and time-on-the-market: the impact of price setting. *J Real Estate Financ Econ* 26:95–111
- Bassett GW Jr, Hsiu-Lang C (2002) Portfolio style: return-based attribution using quantile regression. In: Fitzenberger B, Koenker R, Machado J (eds) *Economic applications of quantile regression, studies in empirical economics, physica*. Springer, Heidelberg, pp 293–305
- Belkin J, Hempel DJ, McLeavey DW (1976) An empirical study of time on market using multidimensional segmentation of housing markets. *Real Estate Econ* 4:57–75
- Benefield JD, Hardin WG (2015) Does time-on-market measurement matter? *J Real Estate Financ Econ* 50:52–73
- Bentzien V, Rottke N, Zietz J (2012) Affordability and Germany's low homeownership rate. *Int J Housing Mark Anal* 5:289–312
- Beyerlein A, Toschke AM, von Kries R (2008) Breastfeeding and childhood obesity: shift of the entire BMI distribution or only the upper parts? *Obesity* 16:2730–2733
- Buchinsky M (1994) Changes in the US wage structure 1963–1987: application of quantile regression. *Econometrica* 62:405–458
- Buchinsky M (1995) Quantile regression, Box–Cox transformation model, and the US wage structure, 1963–1987. *J Econom* 65:109–154
- Cajias M, Freudenreich P (2018) Exploring the determinants of liquidity with big data—market heterogeneity in German markets. *J Prop Invest Financ* 36:3–18
- Chaudhuri P, Doksum K, Samarov A (1997) On average derivative quantile regression. *Ann Stat* 25:715–744
- Cox DR (1972) Regression models and life-tables. *J R Stat Soc Ser B (Methodol)* 34:187–220
- Crowley J, Hu M (1977) Covariance analysis of heart transplant survival data. *J Am Stat Assoc* 72:27–36
- Federal Statistical Office (2018) Baufertigstellungen von Wohn- und Nichtwohngebäuden unter anderem nach Bauherren-Lange Reihen. [https://www.destatis.de/DE/Publikationen/Thematisch/Bauen/Bautatigkeit/Wohnungsbau/BaufertigstellungenBauherrenPDF\\_5311201.pdf?\\_\\_blob=publicationFile](https://www.destatis.de/DE/Publikationen/Thematisch/Bauen/Bautatigkeit/Wohnungsbau/BaufertigstellungenBauherrenPDF_5311201.pdf?__blob=publicationFile). Accessed 8 Aug 2018
- Held T, Waltersbacher M (2015) Wohnungsmarktprognose 2030. BBSR-Analysen KOMPAKT, Nr. 7, Bonn. [https://www.bbsr.bund.de/BBSR/DE/Veroeffentlichungen/AnalysenKompakt/2015/DL\\_07\\_2015.pdf?\\_\\_blob=publicationFile&v=5](https://www.bbsr.bund.de/BBSR/DE/Veroeffentlichungen/AnalysenKompakt/2015/DL_07_2015.pdf?__blob=publicationFile&v=5). Accessed 28 Feb 2020
- Hendricks W, Koenker R (1992) Hierarchical spline models for conditional quantiles and the demand for electricity. *J Am Stat Assoc* 87:58–68
- Horowitz JL, Neumann GR (1987) Semiparametric estimation of employment duration models. *Econom Rev* 6:5–40
- Kang H, Gardner M (1989) Selling price and marketing time in the residential real estate market. *J Real Estate Res* 4:21–35
- Koenker R, Bassett GW Jr (1978) Regression quantiles. *Econometrica* 46:33–50
- Koenker R, Biliyas Y (2002) Quantile regression for duration data: a reappraisal of the Pennsylvania reemployment bonus experiments. In: Fitzenberger B, Koenker R, Machado J (eds) *Economic applications of quantile regression, studies in empirical economics, physica*. Springer, Heidelberg, pp 199–220
- Koenker R, Geling O (2001) Reappraising medfly longevity: a quantile regression survival analysis. *J Am Stat Assoc* 96:458–468
- Lerbs O, Oberst A (2014) Explaining the spatial variation in homeownership rates: results for German regions. *Region Stud* 48:844–865
- Liao WC, Wang X (2012) Hedonic house prices and spatial quantile regression. *J Hous Econ* 21:16–27
- Lin DY, Wei LJ (1989) The robust inference for the Cox proportional hazards model. *J Am Stat Assoc* 84:1074–1078

- Lüdemann E, Wilke RA, Zhang X (2006) Censored quantile regressions and the length of unemployment periods in West Germany. *Empirical Econ* 31:1003–1024
- Machado J, Mata J (2000) Box-Cox quantile regression and the distribution of firm sizes. *J Appl Econ* 15:253–274
- Miller NG (1978) Time on the market and the selling price. *Real Estate Econ* 6:164–174
- Peng L, Huang Y (2008) Survival analysis with quantile regression models. *J Am Stat Assoc* 103:637–649
- Portnoy S (2003) Censored regression quantiles. *J Am Stat Assoc* 98:1001–1012
- Powell JL (1984) Least absolute deviations estimation for the censored regression model. *J Econom* 25:303–325
- Powell JL (1986) Censored regression quantiles. *J Econom* 32:143–155
- Reisenbichler A (2016) A rocky path to homeownership: why Germany eliminated large-scale subsidies for homeowners. *Cityscape* 18:283–290
- Schmillen A, Möller J (2012) Distribution and determinants of lifetime unemployment. *Labour Econ* 19:33–47
- Thomschke L (2015) Changes in the distribution of rental prices in Berlin. *Region Sci Urban Econ* 51:88–100
- Turnbull GK, Dombrow J (2006) Spatial competition and shopping externalities: evidence from the housing market. *J Real Estate Financ Econ* 32:391–408
- Voigtländer M (2009) Why is the German homeownership rate so low? *Hous Studi* 24:335–372
- Vonovia SE, CBRE GmbH (2016) Wohnmarktreport Deutschland 2016
- Wehby GL, Murray JC, Castilla E, Lopez-Camelo JS, Ohsfeldt RL (2009) Prenatal care effectiveness and utilization in Brazil. *Health Policy Plan* 24:175–188
- Wood JH, Wood NL (1985) Financial markets. Harcourt Brace Jovanovich, San Diego
- Yavas A, Yang S (1995) The strategic role of listing price in marketing real estate: theory and evidence. *Real Estate Econ* 23:347–368
- Zietz J, Zietz EN, Sirmans GS (2008) Determinants of house prices: a quantile regression approach. *J Real Estate Financ Econ* 37:317–333

**Publisher's Note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.