

RENT ANALYSIS BEYOND CONDITIONAL EXPECTATION – HOW TO ESTIMATE THE SPATIAL DISTRIBUTION OF QUOTED RENTS BY USING GEOGRAPHICALLY WEIGHTED REGRESSION

Name (s) of author(s):

Anne Seehase

Magdeburg (GERMANY)

anne.seehase@stat.magdeburg.de

Organization: Department of Statistics, Magdeburg

Motivation

Living in a rented accommodation is the most common way of housing in Germany: 57 percent of all households are governed by a rental agreement (Statistisches Bundesamt, Wiesbaden 2015, effective 2013). Increasing rents lead to tense rental markets in most of the big German cities like Berlin, Hamburg or Munich. But also in other areas of high population density the gap between the amount of affordable apartments required and construction activity (in this particular field) is still high. It is expected that the increasing number of households caused by the demographic transition and urbanization is likely to deteriorate the situation. Even in cities like Magdeburg, which right now are characterized by quite a relaxed rental market, rising rents are expected (Rieß 2016).

Alternative Data-Sources

Since more and more rental property offers are published on online platforms, it is getting easier to get information on quoted rents immediately. Unlike traditional data collection of rental prices, using these platforms benefits from monitoring developments in the rental market almost in real-time (Hausmann et al. 2016).

Traditionally, the German regional rent-index-analysis (Mietspiegel) is based on the asset rents over the previous 4 years. They are specifically used as an instrument of market regulation and market orientation by defining a regional comparable rent. Admittedly, the focus on the asset rents for the analysis leads to insufficient conclusions for the market price by re-letting, especially in very dynamic markets (Held et. al. 2013). Furthermore, regional rent indexes have to be updated just every two (four) years. Their compilation is complex and expensive, therefore rental index for the most German cities are not available.

The analysis of rental offers data could be a cost-efficient alternative for the rental market monitoring. Although rental offers prices may differ from the real price by re-letting an apartment, they allow a differentiated view on the current market situation. This is important

for lodging allowance by the social welfare office for example. The analysis of the development in the course of time and space could show current trends. This could be an indicator of changes in urban districts, e.g. for social segmentation or popularity of certain areas.

The main advantage of this approach is the availability of high numbers of rental offers, e.g. via online platforms or the advertisement section in local newspapers. Since the data is public, data-collection is quite easy. By dint of web-crawling-algorithms or Web-APIs it is possible to collect data from the online platforms in real-time.

Alternatively, data may be purchased from commercial service providers. These companies collect rental data offers over a longer period from several sources and offer licenses for their data base. Data cleansing and double identification are usually part of the service.

Nevertheless, disadvantages of using rental offers data have to be considered. First, they are biased by the interest of the supplier. Second, the researcher has no bearing on the characteristics covered in the rent offer. A lot of special specifications about the apartment are missing. Third, very good housing supplies are likely to be re-hired without a publication on an online platform.

In summary, using data of rental offers could be a cost-efficient way to get an overview of the current development in the rental market. Regardless, the origin of the data and their original purpose should be kept in mind. Results should be interpreted within the context of the data.

Methodology

A rental apartment is a hedonic product, i.e. the rental price of a property is determined by the individual characteristics of the flat and the characteristics of the surrounding neighborhood. Because some of the characteristics of an accommodation are specific, e.g. the position or the condition at time x , it is difficult to compare one flat with another. The method of hedonic modeling estimates the effects of the individual characteristics.

Thus it will be possible to estimate the expected price of a hypothetical average flat and to compare this price depending on the location of the flat or the date of the offer.

Hedonic price models are often based on (multiple) linear regression which also enables the estimation of the conditional expectation by a special constellation of characteristics (Brachinger (2002)).

This research-project is based on the model of the quantile regression (QR). The quantile regression developed by Koenker and Basset (1978) enables the analysis of the conditional quantile to the quantile value $\tau \in (0,1)$ of a dependent variable in conjunction with a set of explanatory variables.

Referring to Zietz et al. (2008), Lia and Wang (2012), Su and Yang (2007) and Kostov (2009), it can be stated that QR-models are used in several studies within the context of hedonic

pricing. Most of them used an AR-extension to integrate the spatial effects. In contrast to our research, their subject of analysis is housing prices or land values instead of rental prices.

Quantile regression (QR) ¹

In contrast to the linear regression the model of the quantile regression enables the estimation of other parts of the conditional distribution, not just the middle tendency.

The model equation from the quantile regression is quite similar to the linear model. Given a random sample $(Y_i, X_{1i}, \dots, X_{(p-1)i})$, the relationship between Y_i , and the set of independent variables in the quantile regression model for the quantile-value $\tau \in (0,1)$ could be expressed by the formula:

$$Y_i = \beta_0(\tau) + \beta_1(\tau)X_{1i} + \dots + \beta_{(p-1)}(\tau)X_{(p-1)i} + U_i(\tau) = \mathbf{X}_i^T \boldsymbol{\beta}(\tau) + U_i(\tau)$$

$U_i(\tau)$ represents the perturbation in subject to the quantile-value of interest. Unlike the linear Model, the linear link $\boldsymbol{\beta}(\tau)$ and the perturbation $U_i(\tau)$ depend on the quantile-value τ of interest. The major difference to the linear regression is the assumption for the perturbation $U_i(\tau)$. Instead of independent and identical distributed errors with the expected value of zero, it is solely necessary to assume that $U_i(\tau)$ are independent distributed and $F_{U_i(\tau)}(0) = \tau$.

Therefore, it can be shown that $\mathbf{Q}_Y(\tau|\mathbf{X}) = \mathbf{x}_i^T \boldsymbol{\beta}(\tau)$ describes the conditional τ^{th} -quantile of Y .

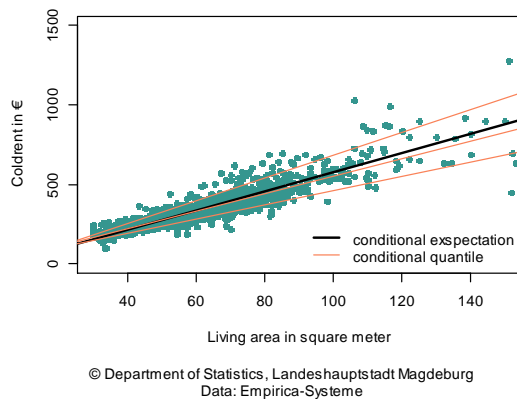


Figure 1: Cold-rent depending living area, first quarter 2016

Figure 1 shows the cold-rent prices of a flat in Magdeburg in depending on the living space. The conditional expectation of the cold-rent price is also evaluated as the conditional quantile for the quantile values $\tau \in \{0.1, 0.5, 0.9\}$. The different slopes indicate a variation of skewness and variance of the conditional distribution from Y . In that case, the conditional median is subject to the

conditional expectation of the cold-rent. This is caused by the right skewness of the conditional distribution. This point out the main-advantage of the quantile regression: it is robust against outliers.

In case of empirical analysis, the unknown parameter $\boldsymbol{\beta}(\tau)$ has to be estimated. One possible solution would be to use minimization. Unlike the linear regression, the sum of the quadratic difference will not be minimized. Instead, $\boldsymbol{\beta}(\tau)$ could be estimated by minimizing the sum of the weighted absolute deviation by the loss-function ρ_τ .

¹ E.g., for a further introduction to the methodology of quantile regression look at Koenker (2002) or Buchinsky (1998).

$$\hat{\beta}_\tau = \arg \min_{\beta \in \mathbb{R}^p} R_\tau(\beta) = \arg \min_{\beta \in \mathbb{R}^p} \sum_{i=1}^n \rho_\tau(y_i - \mathbf{x}_i^T \beta)$$

The loss-function ρ_τ is defined by:

$$\rho_\tau(u) = (u)(\tau - I(u < 0)) = \begin{cases} \tau \cdot u, & \text{for } u \geq 0 \\ u \cdot (\tau - 1), & \text{for } u < 0 \end{cases}$$

In contrast to the function of the sum of quadratic differences $R_\tau(\beta)$ is not differentiable. Hence the minimization-problem has to be solved by methods of linear programming.

Principal of the Geographic Weighted Regression

According to Tobler's first law of geography, "everything is related to everything else, but near things are more related than distant things" (Tobler (1970)), observations in a similar position share the same advantages and disadvantages according to the neighborhood. In

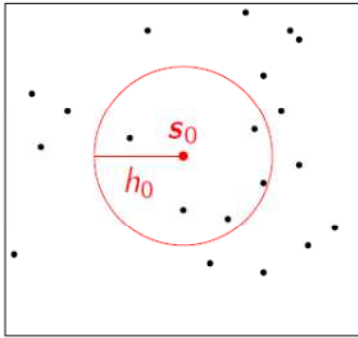


Figure 2: Scheme of spatial target-point and bandwidth

the classical rent index analysis the position effect will be added by predefined position-categories as dummy-variables. In contrast to this, our model will implement the position effect in that way, that we apply the geographical position of each observation.

Therefore, the classical quantile regression-model will be exceeded by the principal of the geographical weighting

regression (GWR) which is based on the ideas of Fotheringham et al. (2002). The extension of the OR-Modell is based on the description from Chen et al. (2014) and McMillen (2013).

The basic idea of the geographical weighting is a local estimation of the model at special target-points, which are denoted by their geographical position (lat_i, lon_i) .

The model of the geographical weighted quantile regression (GWQR) is denoted by:

$$Y_i = \beta_0(\tau; lat_i, lon_i) + \beta_1(\tau; lat_i, lon_i)X_{1i} + \dots + \beta_{(p-1); lat_i, lon_i}(\tau)X_{(p-1)i} + U_i(\tau) \\ = \mathbf{X}_i^T \beta(\tau; lat_i, lon_i) + U_i(\tau), \quad i = 1, \dots, n,$$

Where the coefficients $\{\beta_j(\tau; lat_i, lon_i)\}_{j=1, \dots, p-1}$ are the quantile regression coefficients at the location (lat_i, lon_i) at the quantile value τ . These parameters for the location $s_0 = (lat_0, lon_0)$ are estimated by the pointwise weighting of the loss-function ρ_τ by geographical proximity of the observation i to the target-point (lat_0, lon_0) .

$$\hat{\beta}_\tau(lat_0, lon_0) = \arg \min_{\beta \in \mathbb{R}^p} R_\tau(\beta(lat_0, lon_0)) = \arg \min_{\beta \in \mathbb{R}^p} \sum_{i=1}^n \rho_\tau(y_i - \mathbf{x}_i^T \beta(lat_0, lon_0)) \cdot K\left(\frac{d_{i0}}{h}\right)$$

According to Tobler's first law of geography, nearby observations will be integrated with a higher weight.

The weights are denoted by a kernel function $K\left(\frac{d_{i0}}{h}\right)_{i=1,\dots,n}$ of the scaled distance d_{i0} of the observation i to the target-point by the bandwidth d_{i0} . For the fixed kernel weighting routine the bandwidth h will be identical for every target-point. For example by using the tri-cube kernel-function, h denotes a radius of the circle around the target-point (cp. **figure 2**). All observation-points inside this circle will receive a positive weight; all observations outside the circle will receive the weight zero.

The fixed kernel weighting routine could be problematic in regions with a sparse number of observations, because just a small number of observations are receiving an adequate weight. When the observation-densities vary over space, an adaptive bandwidth could be an alternative. By using an adaptive kernel weighting routine by the window method, for each target-point an own bandwidth could compute depending on the proportion between the points inside an outside the circle.

The optimal bandwidth or window could be determined by a Cross-Validation-criteria (Abberger 1998). In this case we are using an adaptive kernel weighting method with a windows-size of 10 % for every quantile-value of interest. The kernel weights are computed by a Gaussian kernel function, whereas our model will be including dummy variables.

Target-points could be the location of the observation. As an alternative a different location in the region of interest could be used. For our analysis we will use the center points of a grid over the town area. The model is just evaluated over areas, which are covered with buildings. The indication of the covered grids is based on the information of open streets maps.

Estimation results

These analysis are based on the data base of *empirica systeme AG*, which include 24461 observations of rental offers, which were published between the 1st January 2012 and the 31th March 2016 for the city area of Magdeburg. Only geocoded offers with a location precision under 10 meters are used. Additionally a subset was built by the variable living area. Only the observations with a living area of 30 to 250 squares meters are included. On top of this observations whose start time of the offer was before the 1st January 2012 are excluded.

	Quantile value					OLS
	0.1	0.25	0.5	0.75	0.9	
Intercept	1.614 ***	1.62 ***	1.658 ***	1.657 ***	1.631 ***	1.672 ***
log(living aera)	-0.05 ***	-0.03 ***	-0.02 ***	0	0.02 ***	-0.002
elevator	0.024 ***	0.009 ***	0.006 *	0.021 ***	0.054 ***	0.029 ***
balcony/terrace	0.004	0.017 ***	0.023 ***	0.024 ***	0.029 ***	0.021 ***
parking	0.034 ***	0.049 ***	0.068 ***	0.068 ***	0.066 ***	0.060 ***
kitchen	0.052 ***	0.041 ***	0.046 ***	0.061 ***	0.082 ***	0.056 ***
social housing	-0.01 ***	-0.02 ***	-0.03 ***	-0.06 ***	-0.1 ***	-0.053 ***
good condition	0.03 ***	0.014 ***	0.003	-0	0.003	0.010 ***
bad condition	-0.1	-0.09 ***	-0.1 ***	-0.06	-0.02	-0.084 ***
time effect	0.028 ***	0.021 ***	0.017 ***	0.019 ***	0.021 ***	-0.023 ***
R, R ²	0.048	0.05	0.052	0.054	0.055	0.13

Table 1: Global QR- and OLS-estimation. Significance levels [0,001]***; [0,001,0,01]**; [0,01;0,05]*; [0,5,0,1]^o; [0,1,1]'

After some steps of data cleaning, 18442 offers were included into the research.

At first a global semi-log OLS-model and semi-log QR-models by the quantile values

$\tau = (0.1, 0.25, 0.5, 0.75, 0.9)$ were estimated. The response variable was the log-value of the price per square meter. Results show that additional characteristics like an elevator, a balcony or terrace, a parking lot or a kitchen increase the price per square meter for all models. If the flat condition was classified as bad by *empirica system AG*², it has a negative influence on the price. Likewise if someone needs a permission for a social housing, the price decreases. This effect grows within the higher priced segment. The variable "time effect" indicates the time difference between the oldest offer from the 1st January 2012 and the start time of the observation in years. The estimated effect indicates that the price increase depending on the time.

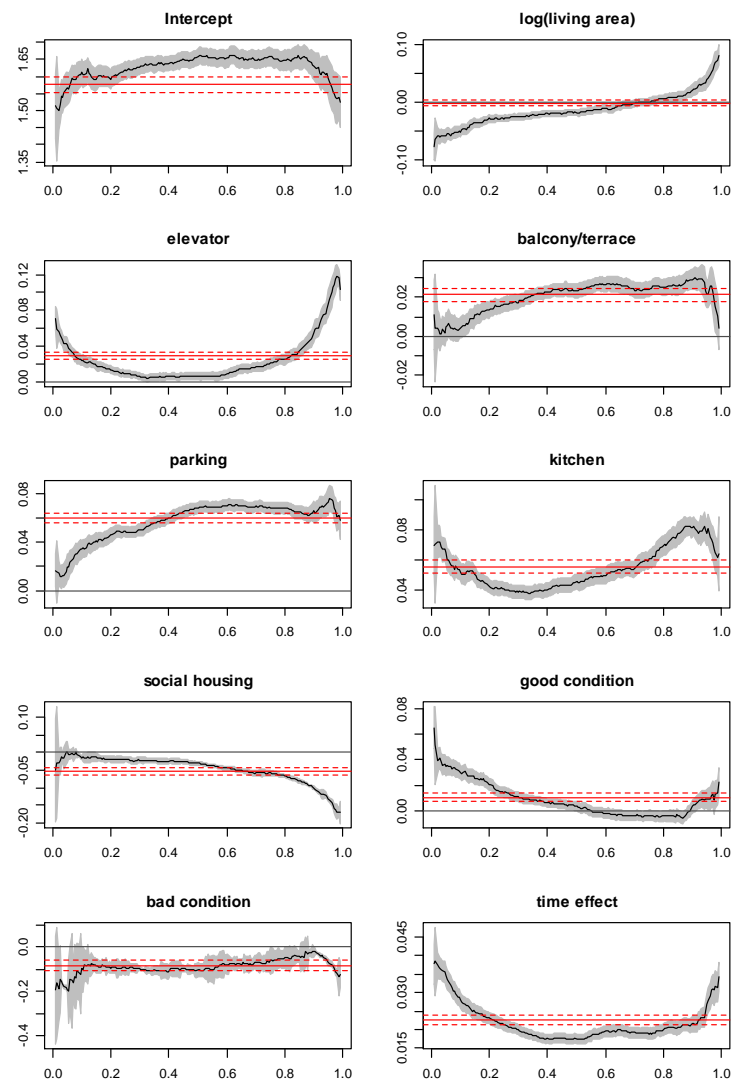


Figure 3: Estimated effects of the global QR-model by a closed meshed set of quantile values

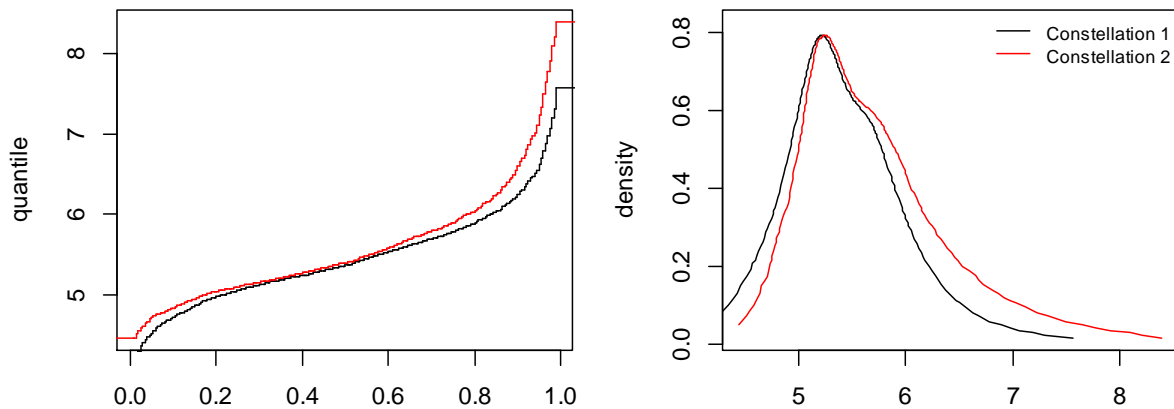
The estimation results illustrate that the effects could vary over the quantile values. Likewise, the effects do not have a significant influence on all parts of the distribution.

When we evaluate the quantile regression for a close meshed set of quantile values, it is possible to estimate the whole trend of the estimated effects as well as the conditional distribution by a special constellation of the characteristics, e. g. by an average of the characteristic in the data.

The difference in the estimation of the effects could be indicating a change in the variance and skews of the conditional distribution by different specification of one of the independent

² The variable flat condition is classified by *empirica systeme AG*. The classification is based upon a derivation from different variables according to the condition and the evaluation of the information written in the free text box of the offer. Their values are a "bad", "normal" and "good".

variables. **Figure 4** indicates that the variance of the conditional distribution increase, if the variable elevator changes from 0 (no elevator) to 1.



© Amt für Statistik, Landeshauptstadt Magdeburg
Data: Empirica-Systeme

Figure 4: Function of the conditional quantiles and conditional density by a flat of 61.8 m², with a balcony or terrace and in a good condition. The starttime of the offer was the 1st January 2015. (Constellation 1- no elevator; constellation 2 - evalator)

The coefficient of determination, denoted by R² at the global OLS-Estimation and R at the global QR-estimation, is very small. Therefore the model is extended by adding the geographical position of the observation by the principles of geographic weighting.

	Intercept							kitchen							R ² /R
	mean	min	p=0.25	p=0.5	p=0.75	max	sd	mean	min	p=0.25	p=0.5	p=0.75	max	sd	
GWR	1.627	1.493	1.584	1.613	1.652	1.896	0.067	0.048	0.005	0.041	0.051	0.058	0.075	0.014	0.46
GWQR=0.1	1.662	1.428	1.596	1.663	1.743	1.941	0.111	0.042	-0.01	0.028	0.043	0.055	0.093	0.023	0.95
GWQR=0.25	1.663	1.492	1.616	1.668	1.707	1.922	0.08	0.035	-0.01	0.025	0.034	0.046	0.079	0.018	0.90
GWQR=0.5	1.692	1.553	1.659	1.683	1.723	1.912	0.061	0.039	-0.01	0.03	0.044	0.05	0.067	0.015	0.88
GWQR=0.75	1.697	1.514	1.652	1.674	1.718	1.974	0.071	0.055	0.001	0.046	0.058	0.066	0.082	0.016	0.90
GWQR=0.9	1.686	1.519	1.632	1.662	1.714	2	0.083	0.075	0.029	0.065	0.079	0.086	0.098	0.016	0.94
	log(living area)							social housing							
GWR	-0.01	-0.07	-0.02	-0.01	-0	0.017	0.016	-0.06	-0.11	-0.07	-0.06	-0.05	-0.01	0.017	
GWQR=0.1	-0.06	-0.12	-0.08	-0.06	-0.05	-0.01	0.021	-0.03	-0.13	-0.03	-0.02	-0.01	0.017	0.023	
GWQR=0.25	-0.04	-0.09	-0.05	-0.04	-0.03	0.003	0.017	-0.04	-0.15	-0.04	-0.03	-0.02	0.006	0.028	
GWQR=0.5	-0.03	-0.07	-0.03	-0.02	-0.02	0.004	0.014	-0.04	-0.1	-0.04	-0.04	-0.03	-0	0.015	
GWQR=0.75	-0.01	-0.07	-0.02	-0	0.002	0.029	0.018	-0.06	-0.1	-0.07	-0.06	-0.05	0.007	0.016	
GWQR=0.9	0.007	-0.07	-0	0.013	0.021	0.044	0.02	-0.1	-0.14	-0.11	-0.1	-0.09	-0.01	0.024	
	elevator							good condition							
GWR	0.038	-0.01	0.019	0.029	0.057	0.106	0.027	0.007	-0.02	0.005	0.009	0.011	0.022	0.006	
GWQR=0.1	0.031	-0.02	0.021	0.031	0.042	0.085	0.017	0.025	-0.02	0.016	0.024	0.034	0.072	0.016	
GWQR=0.25	0.019	-0.03	0.003	0.015	0.033	0.074	0.021	0.012	-0.02	0.007	0.012	0.017	0.035	0.009	
GWQR=0.5	0.019	-0.05	-0	0.012	0.038	0.102	0.03	-0	-0.03	-0	0	0.005	0.011	0.007	
GWQR=0.75	0.033	-0.03	0.005	0.02	0.06	0.151	0.037	-0.01	-0.04	-0.01	-0	-0	0.016	0.008	
GWQR=0.9	0.064	-0.03	0.024	0.055	0.109	0.177	0.052	0.002	-0.04	-0	0.003	0.006	0.016	0.007	
	balcony/terrace							bad condition							
GWR	0.025	-0.01	0.012	0.028	0.04	0.06	0.017	-0.07	-0.14	-0.1	-0.07	-0.05	0.033	0.036	
GWQR=0.1	0.015	-0.04	-0	0.01	0.038	0.067	0.025	-0.1	-0.31	-0.13	-0.07	-0.04	0.059	0.089	
GWQR=0.25	0.02	-0.02	0.004	0.026	0.037	0.056	0.02	-0.09	-0.32	-0.11	-0.08	-0.06	0.054	0.044	
GWQR=0.5	0.026	-0.02	0.017	0.03	0.039	0.053	0.016	-0.08	-0.17	-0.11	-0.08	-0.05	0.022	0.041	
GWQR=0.75	0.027	-0.02	0.018	0.028	0.037	0.054	0.013	-0.06	-0.17	-0.08	-0.05	-0.04	0.115	0.047	
GWQR=0.9	0.03	-0.01	0.022	0.032	0.04	0.056	0.012	-0.04	-0.13	-0.08	-0.05	-0.02	0.104	0.047	
	parking							time							
GWR	0.06	0.028	0.047	0.06	0.07	0.113	0.017	0.022	0.019	0.022	0.022	0.023	0.027	0.002	
GWQR=0.1	0.041	-0.03	0.026	0.044	0.059	0.107	0.029	0.026	0.015	0.02	0.024	0.032	0.042	0.007	
GWQR=0.25	0.047	-0.01	0.035	0.047	0.06	0.107	0.022	0.021	0.015	0.018	0.02	0.023	0.029	0.003	
GWQR=0.5	0.064	0.015	0.051	0.064	0.076	0.113	0.019	0.018	0.014	0.016	0.018	0.02	0.025	0.002	
GWQR=0.75	0.066	0.023	0.054	0.066	0.075	0.132	0.018	0.019	0.011	0.018	0.019	0.02	0.028	0.002	
GWQR=0.9	0.064	0.02	0.053	0.064	0.072	0.131	0.017	0.022	0.014	0.02	0.022	0.023	0.029	0.002	

Table 2: Estimation of the GWR and QR

By evaluating the local GWR-models and GWQR-models at the position of each observation the coefficient of determination could be improved. The local effects of the coefficients take into account the unobserved influences of the neighborhood.

As a result, for every target-point a single model has to be estimated at the assumption that the effects could vary over space.

In **table 3** you can find descriptive statistics, which gives an overview about the general tendency. Every model evaluation is punctual. Additional color coded maps of the coefficients could help to analyze the influence of the effects in respect to their position. To increase the readability it will be evaluated at the center points of the grid. The whole cell will be marked according to the estimated values for the center.

For example in **figure 5** the time effect for the quantile values $\tau \in \{0.1, 0.9\}$ was mapped. Because the frequencies of the offers vary over space, their evaluation will be concentrated at the target-points with an adaptive bandwidth under 1 kilometer.

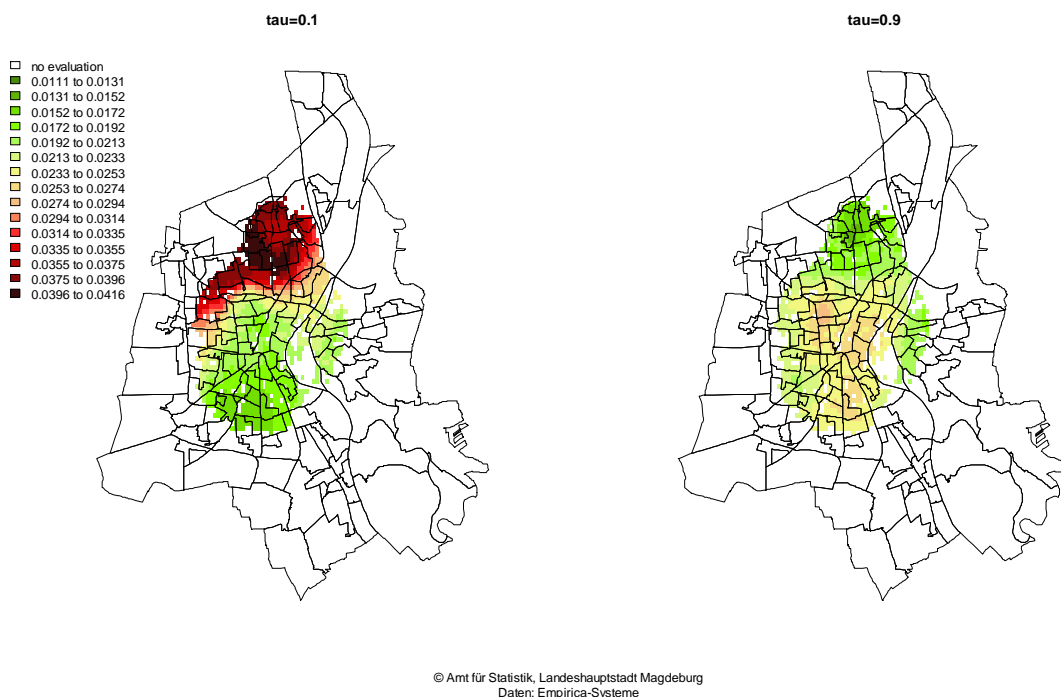


Figure 5: Spatial time-effect for the 0.1- and 0.9-quantile. The coefficients indicate a relative chance of the price per square meter, when the start time of the offer shifts for one year.

It has to be highlighted, that the north areas of the town have the highest time effect at the 0.1-quantile. With higher quantile values this will be inverted. This indicates that the range of conditional distribution will get smaller for the newer offers. On the other side the lower part of the conditional distribution at the south is not changing as much with time. However the upper part of the conditional distribution will be moved to higher prices. The range of the distribution is increasing for newer offers.

It is helpful to predict the (retransformed) conditional expectation and quantiles of a flat with special characteristics to analyze the spatial conditional distribution. Due do this, we take a look at a flat with average characteristics (61.8 m² living area, with a balcony and a good condition) at two starting points (July 1, 2012 vs. July 1, 2015).

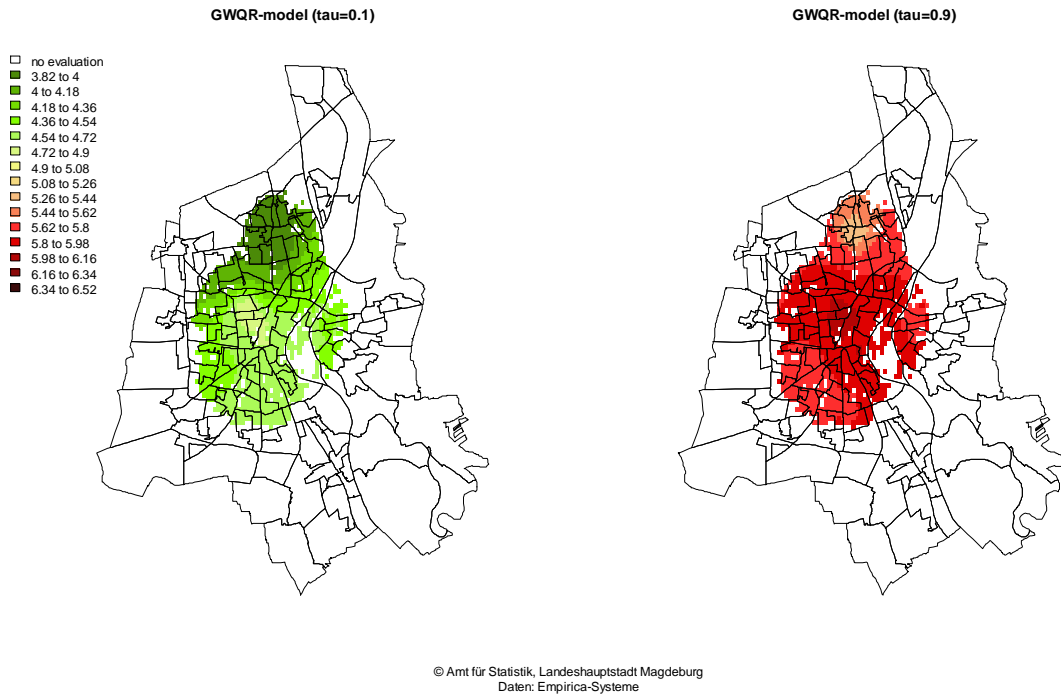


Figure 6: Predicted cold-rent prices per square meter (in €) at the quantile values 0.1 and 0.9 for an average flat (2012)

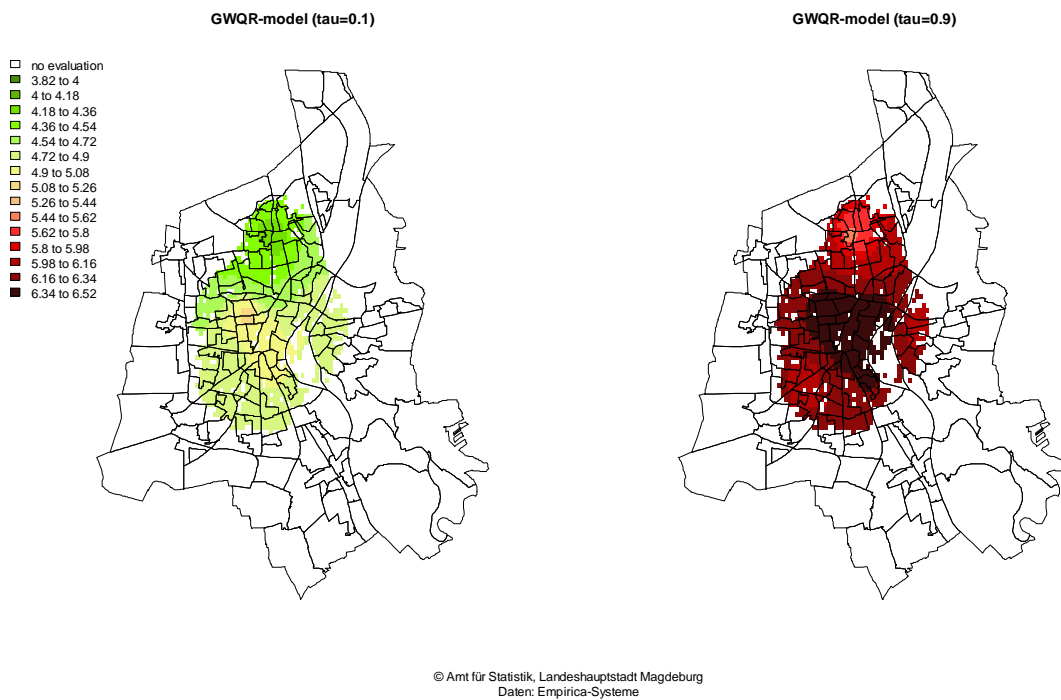


Figure 7: Predicted cold-rent prices per square meter (in €) at the quantile values 0.1 and 0.9 for an average flat (2015)

For the 0.1-quantile the cold-rent price per m² varies 2012 between 3.83 € (4.28 € 2015) in the north and 4.85 € (5.14 € 2015) in the center. According to the spatial distribution of the time effect the prices of the lower part of the distribution are converging. For the year 2012 the difference is 1.02 € between the highest and lowest priced region. For the year 2015 the estimated difference is just 0.86 €. For the 0.9-quantile the cold-rent price per m² varies 2012 between 5.31 € (6.02 € 2015) in the north and 5.59 € (6.51 € 2015) in the center. The difference between the highest and lowest priced region increased (2012: 0.71€, 2015:0.92 €).

Looking at the middle tendency given by the conditional median and expectation, we can confirm, that the north areas are cheaper than the central areas of the town. In contrast to the predicted values 2015 by the GWR-model we see, that the GWR-model overestimates the middle tendency in most of the central and southern area of town. The reason for this could be a higher frequency from outliers caused by unobserved higher quality of new built flats in the higher price segments.

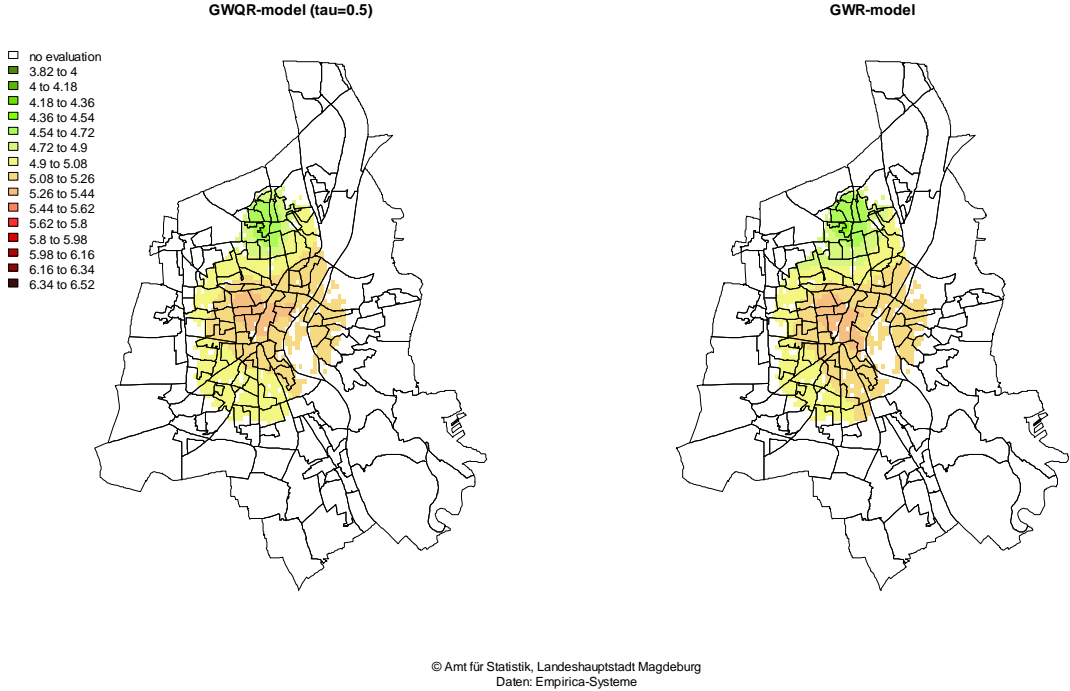
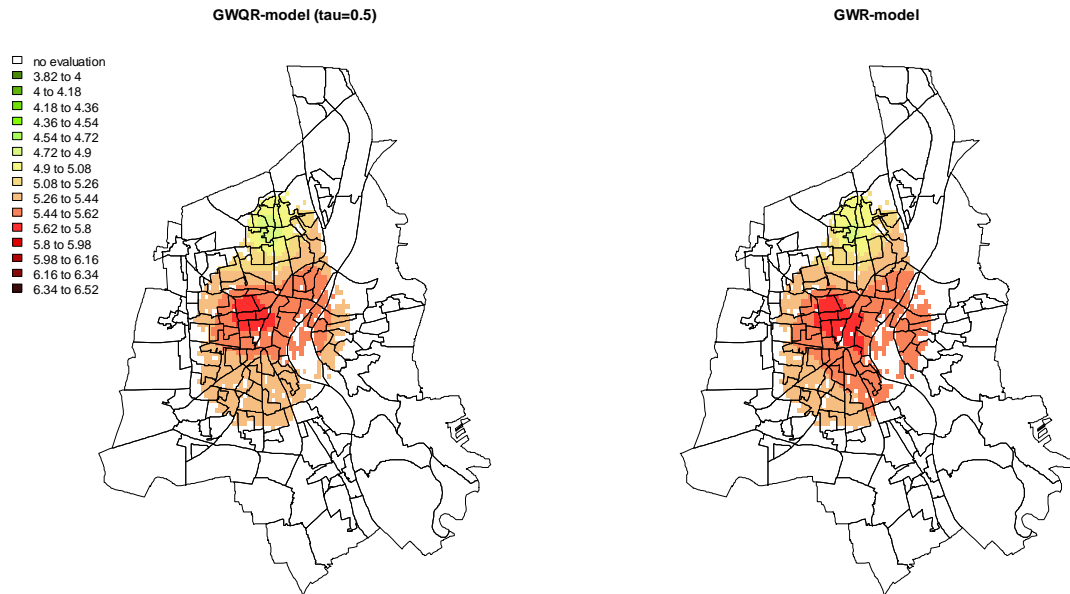


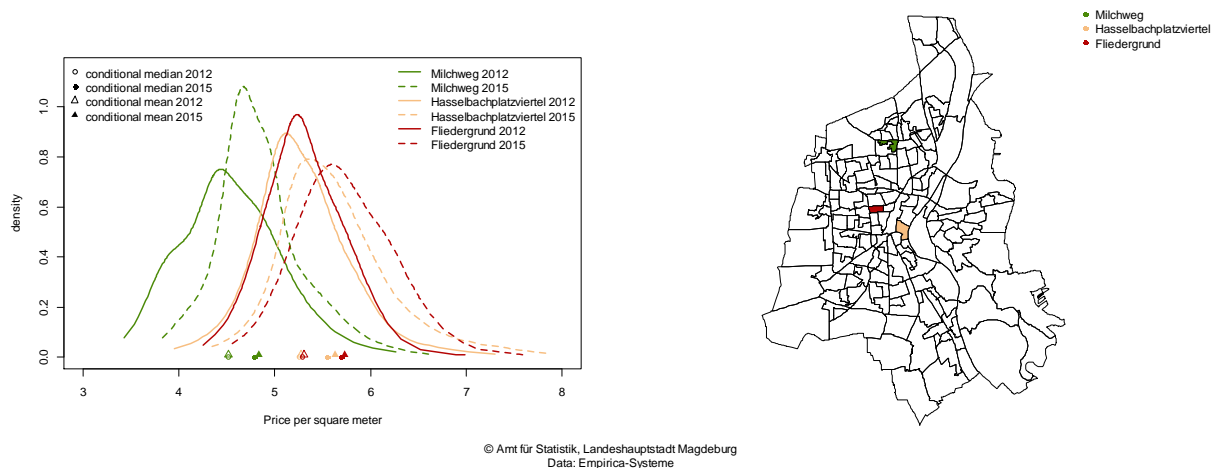
Figure 8: Predicted cold-rent prices per square meter (in €) at the quantile values 0.5 and the GWR-model of an average flat (2012)



© Amt für Statistik, Landeshauptstadt Magdeburg
 Daten: Empirica-Systeme

Figure 9: Predicted cold-rent prices per square meter at the quantile values 0.5 and the GWR-model of an average flat (2015)

This is confirmed by the estimation of the whole conditional distribution at 3 example target-points. Especially for the right skew conditional distribution from the area “Hasselbachplatz”, the difference between conditional median and mean increases from 2012 to 2015. This area has a lot of old buildings and is stamped by studentical lifestyle.



© Amt für Statistik, Landeshauptstadt Magdeburg
 Data: Empirica-Systeme

Figure 10: Conditional distribution at 3 target-points of an average flat at 2012 and 2015

Conclusion

The first results of our research are showing the potential of the geographic weighted quantile regression. This enables a detailed view at the conditional distribution considering the location of interest. Time trends and other effects can be analyzed. Also, geographical autocorrelation is handled by the estimation algorithm. Nevertheless, the local estimation is quite computationally intensive. Also, the QR-algorithm is just a local approach for the

quantile value of interest. The determination of the conditional quantile values could lead to quantile crossing. This means, the conditional quantile to the quantile value τ_1 is less than the conditional quantile to the value τ_2 ($\tau_1 \leq \tau_2$), unless $Q(\tau_1|x)$ is bigger than $Q(\tau_2|x)$. It is in disagreement with the definition of the τ^{th} -quantile.

References

- Abberger, Klaus: Cross-validation in nonparametric quantile regression. In: Journal of German Statistical Society 82, Nr.2, 1998, pp. 149-161
- Brachinger, Hans W.: Statistical Theory of Hedonic Price Indices. Working Paper, University of Fribourg 2002
- Buchinsky, Moshe: Recent Advances in Quantile Regression Models: A Practical Guideline for Empirical Research. In: Journal of Human Resources, vol. 33, issue 1, 1998, pp. 88-126
- Chen, Vivian Y.; Deng, Wen-Shuenn; Yang, Tse-Chuan; Matthews, Stephan: Geographically Weighted Quantile Regression (GWQR): An Application to U.S. Mortality Data. In: Geographical Analysis 44(2) 2012, pp. 134-150
- Fotheringham, A. Stewart ; Brunson, Chris ; Charlton, Martin: Geographically Weighted Regression: The Analysis of Spatially Varying Relationships. Wiley 2002
- Haussmann, Patrick; von Bodelschwingh, Arnt: Wohnungsmarktanalysen mit Daten von "ImmobilienScout24" In: Stadtforschung und Statistik Jg. 29 (1), 2016, pp. 43-52
- Held, Tobias; Nielsen, Jörg; Schürt, Alexander: Aktuelle Mietentwicklung und ortsübliche Vergleichsmiete: Liegen die erzielbaren Mietpreise mittlerweile deutlich über dem örtlichen Bestandsniveau? Hintergrundpapier BBSR 2013
- Liao, When Chi; Wang, Xizhu 2012: Hedonic house prices and spatial quantile regression. In: Journal of Housing Economics, Volume 21, Issue 1, 2012, pp. 16-27
- Koenker, Roger: Quantile regression, Cambridge university press 2005
- Koenker, Roger; Basset Jr., Gilbert: Regression quantiles. In: Econometrica: journal of the Econometric Society, 1978, pp. 33-50
- Koenker, Roger; Machado, José A.F.: Goodness of fit and related inference processes for quantile regression. In: Journal of the American statistical association 94 (448), 1999, pp. 1296-1310
- Kostov, Philip: A Spatial Quantile Regression Hedonic Model of Agricultural Land Prices. In: Spatial Economic Analysis, vol. 4, issue 1, 2009, pp. 53-72
- McMillen, Daniel P.: Quantile regression for spatial data. Berlin, Springer, 2013
- Rieß, Martin: Stadt braucht 700 neue Wohnungen im Jahr. In: Volksstimme May 5 2016
- Su, Liangjun; Yang, Zhenlin: Instrumental Variable Quantile Estimation of Spatial Autoregressive Models. Paper No. 05-2007, School of Economics, Singapore Management University 2007
- Tobler, Waldo R.: A Computer Movie Simulating Urban Growth in the Detroit Region. In: Economic Geography, Vol. 46, 1999, pp. 234-240
- Zietz, Joachim; Zietz, Emily N.; Sirmans, Stacy G. 2008: Determinants of House Price: A Quantile Regression Approach. In: The Journal of Real Estate Finance and Economics, Volume 37, Issue 4, 2008, pp. 317-333