Spatial Quantile Analysis of Real Estate Prices in Germany

The aim of the paper is to identify and establish empirical facts on the determinants of the real estate prices by analyzing spatial regional data, considering the price level of the region. We provide empirical analysis on the panel data set of 401 German regions for the period 2004 – 2020 taking into account their relative geographical location and prices. The main contribution of our paper is the analysis of determinants and spatial effects in housing prices, taking into account whether the region belong to high-prices or low-prices clusters using quantile regression analysis.

The panel dimension of the analysis allows to account for regional heterogeneity, whereas spatial regional dimension catches the interaction of close-located regions: how shocks in real estate price determinants in neighboring regions affect the housing prices level and to what extent the shock in one region is expanded to other closely located regions. Finally, spatial quantile regression reveals the differences between high-prices and low-prices regions. Taken together they provide a unique opportunity to analyze the fundamental factors affecting real estate prices from the different perspectives.

First, we investigate whether real estate prices in German regions experience spatial correlation and whether this correlation is connected to the fundamental factors of the real estate prices. We find factors from the demand and supply side that affect housing prices. The prices may be influenced by socio-economic factors, such as the share of population with higher education, purchasing power, the quality of human capital, which affect the level of development of the region's economy and in turn affects the real estate prices. Second, we perform spatial econometric analysis, applying spatial autoregressive quantile model (QSAR)

QSAR with quantile τ (Zhang and Wang 2016) helps to record spatial effects.

$$Y = \lambda(\tau)WY + \alpha i_n + X\beta(\tau) + \varepsilon(\tau)$$

Here estimation $\beta(\tau)$ may be defined as:

$$\widehat{\beta(\tau)} = \arg min \frac{1}{n}\left(\sum_{y_i \geq x_i'b} \tau|y_i - x_i'b| + \sum_{y_i < x_i'b} (1-\tau)|y_i - x_i'b|\right)$$

The estimation of QSAR can have two ways: by two-stage quantile regression (2SQR) (Kim and Muller 2004) or quantile regression with instrumental variables (IVQR) (Chernozhukov and Hansen 2006).

For 2SQR the estimations of two serial quantile regressions foe each quantile are needed. It enhances the effectiveness of spatial lag model estimations. IVQR has better final sample results, but this method requires more time and efforts. (Garza and Ovalle 2019)

Thus, we estimate the model by two stages (McMillen 2012):

1. We estimate quantile regression with dependent variable WY:

$$WY = X\beta(\tau) + WX\gamma(\tau) + \varepsilon(\tau)$$

2. We estimate predicted value from the first stage and add this component to the quantile regression:

$$Y = \rho(\tau)\widehat{WY} + X\beta(\tau) + \varepsilon(\tau)$$

In this work we estimate QSAR using the method described above for our panel data. The specification is:

$$\ln{(Price)_{it}} = \alpha_{0i} + \rho(\tau)W\widehat{Price}_{it} + \alpha_1(\tau)\ln{(Price)_{it-1}} + \alpha_2(\tau)\ln{(Determinant_1)_{it}} + \cdots + \alpha_{2+k}(\tau)\ln{(Determinant_k)_{it}} + \varepsilon_t$$

where W is a $n\times n$ spatial weighting matrix. Non-zero elements of the matrix W indicate that the region j is a neighbor for the region i. Diagonal elements of the matrix are zeros. Matrices are row standardized so that the weights of all neighboring regions sum up to 1. We employ two specifications of weighting matrices in our analysis: a matrix based on inverse geographical distances between the regional centers (inverse distance matrix) and a matrix based on regional common borders (contiguity matrix). These types of matrices are often used in spatial regional analysis (see e.g. (Burgess and Profit 2001), (Niebuhr et al. 2009)), since they provide a good approximation for connectivity between regions.

The results of the analysis suggest that the demand-side factors increase the housing prices in most cases. Spatial allocation of regions plays an important role in real estate pricing: regions located close to regional centers benefit from it, which is easily accounted by spatial correlation. We also find spatial effects for the determinants: a demand change in a region affects the price also in the neighboring regions. From the spatial quantile regression analysis, we expect that regions with higher prices are more sensitive to infrastructural or policy changes, whereas low prices regions experience more sluggish reaction. More than that, the relative the positive spatial dependence is higher between regions with high housing prices and lower for less attractive regions.

References

Burgess, Simon, and Stefan Profit. 2001. “Externalities in the Matching of Workers and Firms in Britain.” *Labour Economics* 8(3): 313–33.

Chernozhukov, Victor, and Christian Hansen. 2006. “Instrumental Quantile Regression Inference for Structural and Treatment Effect Models.” *Journal of Econometrics* 132(2): 491–525.

Garza, Nestor, and Maria Carolina Ovalle. 2019. “Tourism and Housing Prices in Santa Marta, Colombia: Spatial Determinants and Interactions.” *Habitat International* 87: 36–43.

Kim, Tae-Hwan, and Christophe Muller. 2004. “Two-stage Quantile Regression When the First Stage Is Based on Quantile Regression.” *The Econometrics Journal* 7(1): 218–31.

McMillen, Daniel P. 2012. *Quantile Regression for Spatial Data*. Springer Science & Business Media.

Niebuhr, A., N. Granato, A. Haas, and S. Hamann. 2009. "Does Labour Mobility Reduce Disparities between Regional Labour Markets in Germany?" *Institut für Arbeitsmarkt- und Berufsforschung (IAB), Nürnberg [Institute for Employment Research, Nuremberg, Germany], IAB Discussion Paper* 46.

Zhang, Haiyong, and Xinyu Wang. 2016. "Effectiveness of Macro-Regulation Policies on Housing Prices: A Spatial Quantile Regression Approach." *Housing, Theory and Society* 33(1): 23–40.