The spatial impact of employment centres on housing markets.

Nikodem Szumilo, Edyta Laszkiewicz and Franz Fuerst

Abstract: The rapid growth of house prices in many fast-growing economic regions has proven difficult to explain by existing economic theories. Large scale models have proven especially difficult to apply to explain pricing differences between micro-locations. This research provides a model of spatial and temporal interactions between housing and employment markets and applies it to the example of Cambridge, UK. The results show that rapid growth of employment centres increases house prices in neighbouring locations even after adjusting for income, employment or macroeconomic conditions. It appears that by clustering together in large employment centres companies are able to create value to their employees at no additional direct cost to themselves. While it is unclear what exactly is the source of this value the effect appears to allow businesses to hire employees at below the wage rate which would be dictated by a parity of disable income.

1. Introduction

Over the last two decades house prices in most developed economies have increased at an unprecedented rate (Knoll et al 2014). It would seem that many centres of economic activity around the world struggle with house prices being increasingly unaffordable (Chakrabarti and Zhang 2010, Quigley and Raphael 2004). Researchers have been trying to explain this trend by analysing the determinants of supply and demand in housing markets. Some academics argued that increasing incomes and low interest rates have increased the ability of households and individuals to borrow and their ability to pay higher prices (McQuinn and O'reily 2008). Others pointed out that increasing capital values reduced the net user cost of housing and encouraged prices to grow faster than income (Himmelberg 2005). At the same time the supply has been reported not to be able to adjust accordingly. Limited availability of land due to physical restrictions or planning regulations has been used to argue that prices increase due to lack of adequate supply (Knoll et al. 2014).

While causes for the phenomenon are subject to an ongoing debate consequences of unaffordable house prices have also started to be recognized in academic literature. In addition to the most obvious social problems of inequality and spatial segregation (Massey and Kanaiaupuni 1993), economic problems of restricted labour supply have been reported (Duffy et al. 2005). High living costs at most desirable locations have also been linked to increased use of transportation which leads to considerable environmental concerns (Miles 2012).

However, the majority of studies that have considered the problem of increasing house prices has focused on national data and considered the phenomenon at a very high level. It is interesting to see that the attractiveness of particular locations has not been investigated in more detail. While many researchers have reported differences between house prices and price to income ratios between cities within the same country (Gathergood 2011, Hiebert and Roma 2010) few have attempted to explain this disparity. Rather superficially this phenomenon has been attributed to the economic success of those areas and considered as one of many an effects of agglomeration benefits (Girouard et al. 2005, Kiel and Zabel 2008). At the same time economic success stories of regions have been studied meticulously but appear to have offered very few universal conclusions (Christopherson et al. 2010, Hospers and Beugelsdijk 2002). It would appear that better economic performance leads to higher income. Assuming that its constant proportion is paid for houses and that dwellings do not change over time house prices would go up proportionally with income. This logic assumes that the attractiveness of location is determined by the average income within an area and a number of other factors that do not change significantly over time like characteristics of the housing stock.

However, even with the same average income prices in different micro-locations within the same region often experience dissimilar rates of growth in price. This would suggest that there is a time-variable factor that affects the amount people are willing to spend on housing by changing the proportion of income individuals are willing to sacrifice for this purpose. By using a spatial dynamic panel data modelling technique this research shows that certain developments of employment centres can have a dramatic effect on house prices within their immediate areas that is not explained by growing income. With all physical characteristics kept constant the only other explanation is an increase in the proportion of income being spend on housing within the area.

This research uses the example of Cambridge UK to show that an increase in economic activity in certain micro-locations may generate a considerable spatial effect that affects prices within its surrounding areas. It begins with an overview of the unique characteristics of the residential property market in Cambridge UK and an explanation of why it is an ideal example for this research. It then discusses the theory of how expanding housing supply can interact with local labour markets and how this is expected to manifest itself in Cambridge. Research methodology and data are explained next with special attention given to the problem of time-variable and spatial determinants of house prices. A presentation and a discussion of the results follow and a summary concludes.

2. The curious case of Cambridge UK

In many ways the market for residential property in Cambridge UK is unique. After the recession of 2007 house prices grew the fastest of all cities in the country and increased by 44.7%. This rate was even higher than in London which saw a 44.4% increase. As most markets with high price to income ratios Cambridge is a popular location for buy-to-rent investors. According to Savills as much as 70% of new houses were purchased by investors interested in letting their properties in 2014. This trend is also encouraged by the fact that a considerable proportion of the population is related to a large university and many of its students and junior staff require housing only temporarily. This results in 27.8% of houses in Cambridge being rented (2011 census) with tenants coming from all social groups. This creates a vibrant and relatively developed rental market where rents are rising quickly (52% from 2009 to 2014).

With growing population as well as house prices the supply side has attempted to respond. In years 2014 alone the housing stock expanded by around 2.5% as new houses were built and a similar rate of growth is projected until 2020. New completions are expected to come mainly from three major projects (North-West Cambridge, Southern Fringe and Cambridge East) rather than small independently built houses. This is due to the fact that supply of land in Cambridge is considerably limited by planning policies and a Green Belt. Until mid-1990s Cambridge was constrained by a very strict planning policy which focused on preserving it historic centre. Since then the policy has been changed and the city committed to supporting economic growth.

The city of Cambridge is most commonly associated with its ancient university. In fact, Oxford Economic report that education was the biggest employment sector in the region in 2014 and grew by 21% over the preceding decade. The university alone employed 9,500 people and is a considerable force driving the local economy. The cutting edge research performed at the university attracts not only the brightest and most ambitious students from around the world but also high technology companies. By locating in Cambridge tech businesses can take advantage of the spill-over effects of working with the best researchers but also have access to a highly skilled labour force produced by the university. In fact, the area has attracted so many companies of this kind that it has been named the “silicon fen” (in comparison to the silicon valley). In July of 2015 there were 2,100 high technology companies in the Cambridge area totalling the sales of around £14billion per annum. They vary in size from university based start-ups supported by one of many local business parks to giant multinationals such as ARM technologies, Microsoft or Samsung. In addition, the number also includes bio-tech companies that are based in the Addenbrooke’s area which is expected to become the biggest campus of this kind in the world with over 17,000 people working on the site.

If the large concentration of technology oriented companies can indeed be attributed to the world-wide success of the university then its expansion should be beneficial for the local economy and stimulate further growth. However, the strict planning policy in the area considerably limits the ability of the university to expand its facilities. As a result, the institution has been able to develop new research centres only on two sites: the Addenbrooke’s campus oriented on medical research and West Cambridge Site focused on other sciences and engineering. This makes Cambridge an excellent natural experiment for investigating the spatial impact of expanding size and capacity of employment centres that drive the local economy on the housing market.

3. Expanding inelastic supply and its impact on demand

By applying their model to US data from 1980-2000, Glaeser et al. (2006) find that the extent to which increases in productivity will create bigger cities or just higher paid workers and more expensive housing depends on the elasticity of housing supply (Glaeser, Gyourko and Saks 2006). We modify their model an apply it to Cambridge to find if expansion of inelastic supply of housing can lead to an increase in demand for it.

If the location has an elastic housing supply, new housing will be supplied in response to the increase in demand, resulting in only a modest increase in house prices but a significant increase in population. This works to dilute the marginal product of labour, thus, reducing the wage level and returning the utility provided by the location to equilibrium level (Glaeser, Gyourko and Saks 2006). On the other hand, if the location has an inelastic housing supply, new housing will not be supplied in response to the increase in demand, resulting in heightened competition for the existing housing stock. This works to bid up the price of housing, returning the utility provided by the location to equilibrium level, whilst the population remains relatively unchanged (Glaeser, Gyourko and Saks 2006).

The higher the demand for labour in a particular location is, the higher wages will be offered to workers. If a particular set of skills is required, the new workers can come from training the labour force or relocating individuals from areas where their skills are in lower demand. The longer the training time required for a job, the higher the likelihood of attracting new workers form outside of the region. In fact, national employment mobility polices encourage this process (Forrest 1986, Booth et al. 1999, Battu et al. 2008). This leads to a stark difference between national and local employment. It would seem that if housing supply is restricted, local unemployment can be decreased by pricing the population that struggles to find work out of the area.

Increasing the stock of dwellings in the centre of employment in the current period should lead to a temporary reduction in prices. This can be expected due to an outwards shift in supply. However, increasing supply of housing can be seen as an expansion of the labour pool. This normally leads to a reduction in wages and decreasing housing demand. However, in a developing local economy where demand for labour is very elastic and marginal gains in productivity form hiring new workers are positive, this may not translate into lower income. If the local economy expands in size and productivity to reflect the benefit of the new workers, new demand for labour is likely to be created. This would maintain the upwards pressure on income as well as housing demand. In fact, it would seem that in this situation an expansion of the labour force may increase productivity sufficiently to stimulate additional demand for it. If marginal returns from expanding the level of employment are increasing, then adding new housing stock will eventually lead to increases in labour demand, income and finally house prices. This increasing marginal return can only be expected form growing economies with restricted labour supply. The response of house prices with inelastic supply to a supply shock depends on elasticity of demand. In expanding economies with restricted labour supply it has been shown to be very high. It can be concluded that in markets with inelastic supply of housing and elastic demand for labour that is positively related to the current labour level housing supply shocks are unlikely to translate into lower property prices. In fact, it is likely that if the economy is expanding its productivity and size along with the supply shock house prices are more likely to increase than to fall. The more inelastic house prices and the more elastic labour demand the lower is the shift in productivity required to increase house prices in response to a supply shock.

4. Modelling house prices

These characteristics of the labour market have important implications for the market for housing. Assuming that the selling price of a property depends on the supply of and demand for dwellings in each location. In general it is possible to formulate a basic demand function of the following form:

$$q_{it} = f(I_{it}, p_{it}, Wp_{it}, \omega_{it}), \tag{1}$$

where q_{it} is the demand for housing in district *i* (*i* = 1,…,*N*) at time *t* (*t* = 2,…,*T*), I_{it} is income within commuting distance, p_{it} is the average house price, Wp_{it} is the spatial lag of house prices, and ω_{it} represents other factors (such as macroeconomic indicators).

The formula given by equation 1 can be adjusted to reflect the factors discussed above. Most importantly it needs to adjust for the growth of employment centres and the distance to the nearest one. This needs to be controlled for overall unemployment and the fact that the past increase in prices influences expectations of their future growth should also be captured. The following function allows those factors to be reflected:

$$q_{it} = f(p_{it}, p_{it-1}, I_{it}, U_{it}, wEC_{it}, Wp_{it}, \omega_{it}), \tag{2}$$

where wEC_{it} is spatially weighted growth in major employment centres and $U_{it}$is unemployment. It is assumed that q_{it} will be influenced by the mean selling price p_{it} and also the lagged price p_{it-1}, thus current demand is assumed to be a response to both contemporaneous and lagged price signals. The effect of the price increment is distributed over two periods (see Nerlove, 1958).

Demand is expected to be negatively affected by prices and unemployment but positively by income and growth in nearby research centres. On the supply side, the initial variables are the same, except that we substitute the stock of dwellings for income within commuting distance, indicators of labour demand are removed, and there are reverse assumptions about the signs of the coefficients. This is apparent from the equation below:

$$q_{it} = f(S_{it}, wS_{Ctr\,t}, p_{it}, p_{it-1}, Wp_{it}, \varsigma_{it}), \tag{3}$$

where S_{it} is the current housing stock, $wS_{Ctr\,t}$ is the current housing stock in Cambridge and ς_{it} represents other factors (such as macroeconomic indicators). Solving the supply function with respect to p_{it} (Eq. 1), and substituting for q_{it} (Eq. 3) using the demand function, it is possible to arrive at of the following reduced form equation:

$$p_{it} = \phi p_{it-1} + \rho \mathrm{W} p_{it} + \beta_1 U_{it} + \beta_2 wEC_{it} + \beta_3 S_{ctr\,t} + \beta_4 S_t + \beta_5 wp_{ctr\,t} + \beta_6 I_{it} + v_{it},$$
(4)

where $wp_{ctr\,t}$ is the average house price weighted by its geographical distance to Cambridge, $\mathrm{W}p_{it}$is the spatial lag of house prices in surrounding areas, the v_{it} error term is the sum of the usual error ε_{it} and the fixed effects μ_i for individuals which take into account the inter-location heterogeneity. Following specification testing our model does not assume that the disturbances comprise an autoregressive spatial dependence process. Instead we control for all unidentified differences between locations that do not change with time and include spatially weighted variables that reflect changes in the key determinants over time in the current and past periods. Since we found income to be closely correlated with lagged prices the two variables are reflected as a lagged value of the affordability ratio (price/income).

5. Methodology

5.1. Spatial dynamic panel data model with fixed effects

Housing choices follow a spatial and temporal diffusion process (Nanda and Yeh, 2014). On the one hand, changes in the average house price in a certain region affect this value in neighbouring locations. Hence, any local house price shocks are propagated to surrounding areas. On the other hand, anchoring effects observed in the real estate market result in an autoregressive dependence over time (Nanda and Yeh, 2014). This makes modelling longitudinal housing data relatively complex as both those processes need to be adjusted for. Ignoring correlation between spatial units over time or their spatial dependence on would lead to misspecification (Bouayad-Agha and Védrine, 2010).

While dynamic panel models are now relatively common (Arellano and Bond, 1991; Blundell and Bond, 1998) and spatial econometric models have been well documented (Elhorst, 2003), the analysis of spatial-dynamic processes is still under development. By assuming that house prices in spatial units are jointly determined by their regional characteristics, their past values and prices in neighbouring regions, we obtain a spatial autoregressive dynamic panel model with individual effects. It can be expressed as:

$$\mathbf{Y} = \phi\mathbf{Y}_{t-1} + \rho\mathbf{WY} + \boldsymbol{\beta}\mathbf{X} + (\boldsymbol{\mu} + \boldsymbol{\varepsilon}), \quad (5)$$

where $\mathbf{Y} = [p_{1,t},\ldots, p_{N,T}]'$ is a vector of house price for N regions and T time units, $\mathbf{Y}_{t\text{-}1}$ is a vector of lagged house prices, $\mathbf{X} = [E_{it}, wEC_{it}, S_{\mathrm{crt},t}, S_t, wp_{\mathrm{crt},t}]'$ is a matrix of exogenous variables which characterize supply and demand on real estate market, $\mathbf{W} = \mathbf{I}_T \otimes \boldsymbol{W}_N$ is a nonstochastic, time-invariant row-standardized spatial weight matrix, such that diag($\mathbf{W}$) = $\mathbf{0}$, $\boldsymbol{\beta}$ is a vector of structural parameters, $\boldsymbol{\mu} = [\mu_1,\ldots,\mu_N]'$ is a vector of individual fixed-effects, $\boldsymbol{\varepsilon}$ is a vector of error terms, ρ is an endogenous interaction effect (spatial autoregressive term) and ϕ is an autoregressive time effect.

According to the Eq. 5, we capture unobserved heterogeneity for regions by individual fixed-effects μ_i. They represent time-invariant regional characteristics and differences in real estate markets between spatial units. Moreover, the model accounts for spatial dependence by including a spatially autoregressive component $\rho\mathbf{WY}$ and explores housing market imperfections (like a non-instantaneous price reaction) by accounting for temporal dependence ($\phi\mathbf{Y}_{t-1}$). The stability condition for Eq. 5 is $|\phi|+|\rho|<1$. The stability condition is violated if the potential space-time covariance in the model is ignored. The spatial weight matrix has been set using an algorithm of k closest neighbours with k=25. This is discussed in more detail in chapter 5.3

Estimation strategies that allow obtaining consistent and efficient estimates for spatial dynamic panel data models seems to be widely discussed in academic literature (see e.g. Elhorst, 2012). Two popular estimation techniques which are used for such models are: the maximum likelihood method (MLE) and the generalized method of moments (GMM). As pointed by Kukenova and Monteiro (2008), if the endogenous part of the model consists only of spatial and temporal autoregression components, estimators such as MLE, quasi-MLE, C2SLSDV or MLE-GMM can be used. In the case of additional endogenous variables in the model, a system-GMM estimation is more appropriate.

In models with no additional endogeneity MLE-type estimators are more efficient than the corresponding GMM (Yang, 2015) and might be preferred for spatial dynamic panel data models like the one specified above. MLE-type estimators have been presented by both Elhorst (2005) and Yu et al. (2008). While Elhorst considered a panel model in which N was large and T was fixed, Yu et al. concentrated on a data structure in which both N and T were large. Two different ways of incorporating individual fixed-effects were proposed based on the differences in the data; Elhorst uses fist-differenced and Yu at al. demeaned variables. However, modelling initial values of dependent variable (initial differences in the case of fixed-effects models) is required for both methods. Obtaining inappropriate initial values results in biased and inconsistent MLE estimates. When T is large, initial values are easy to achieve. The process is more complicated for short panels (ones with relatively small T) and approximation procedures need to be used. It is possible to use the Bhargava and Sargan (1983) approximation, however, Su and Yang (2015) proposed that using a quasi-MLE estimator and to modelling initial differences with an adaptation of Hsiao et al. (2002) assumption would yield better results. More recently, Yang (2015) established an M-estimator which does not require initial values and is robust against non-normality of errors.

In this study parameters of Eq. 5 are estimated using the quasi-MLE method, proposed by Yu et al. (2008). Although for large N and relatively small T, estimators are consistent not more than with rate T, the bias correction used by Yu et al. eliminate the bias and yield a centred confidence intervals if T grows faster than $N^{1/3}$. The procedure was necessary as in the panel used for this study $T = 4$ and $N^{1/3} = 4.2$. In order, to evaluate the bias of the estimates resulting from an insufficient T, a Monte Carlo simulation has been used (see chapter 5.3 for the results).

Estimation software and procedure choices have been guided by the work presented by Belotti et al. (2014). In addition to the model expressed by equation 5 (dynamic SAR-FEM model), simpler models have been estimated: FEM (with ρ = 0 i ϕ = 0) and SAR-FEM (with ϕ = 0).

In spatial autorgressive models conclusions based solely on coefficients of explanatory variables are biased due to spatial spillover effects (LeSage i Pace, 2009). In order to detect and interpret existing relationships direct, indirect and total effects need to be calculated. In addition, for dynamic models they can be divided into short and long term effects. For the model presented in equation 5 those effects are calculated using the following method presented by Belotti et al. (2016)

- Short-term direct effect:

$$\left[(\mathbf{I} - \rho\mathbf{W})^{-1} \times (\beta_k \mathbf{I})\right]^{\bar{d}}, \qquad (6)$$

- Long-term direct effect:

$$\left[((1 - \phi)\mathbf{I} - \rho\mathbf{W})^{-1} \times (\beta_k \mathbf{I})\right]^{\bar{d}}, \qquad (7)$$

- Short-term indirect effect:

$$[(\mathbf{I} - \rho\mathbf{W})^{-1} \times (\beta_k \mathbf{I})]^{\overline{rsum}}, \qquad (8)$$

- Long-term indirect effect:

$$\left[((1 - \phi)\mathbf{I} - \rho\mathbf{W})^{-1} \times (\beta_k \mathbf{I})\right]^{\overline{rsum}}, \qquad (9)$$

where: $\bar{d}$ – mean diagonal element of a matrix, $\overline{rsum}$ – mean row sum of the non-diagonal elements.

5.2. Spatial weights matrix

One of the key elements of spatial econometrics is the choice of a spatial structure of relationships between entities represented by a matrix **W.** This paper followed the work of Ezcurra and Rios (2015) and compared a number of spatial expression methods. Maximizing value of the Log-Likelihood function or minimizing residual variance of the estimated model can be adopted as a selection criteria. Alternatively, value of the Bayesian posterior model probability can be used. This research focused on minimizing residual variance and AIC for the dynamic SAR-FEM model presented in equation 5. In addition, changes in the spatial interaction parameter ρ have been considered. The results are presented in table 1.

Table 1.Spatial weight matrix selection.

Spatial Matrix W_N	**AIC**	**Residual variance**	**Parameter ρ**
KNN = 1 closest neighbour	4842.738	2.96e+08	0.068
KNN = 2 closest neighbours	4843.280	2.97e+08	0.065
KNN = 5 closest neighbours	4843.154	2.97e+08	0.099
KNN = 10 closest neighbours	4840.638	2.93e+08	0.209*
KNN = 15 closest neighbours	4838.828	2.90e+08	0.264**
KNN = 20 closest neighbours	4839.526	2.91e+08	0.271**
KNN = 25 closest neighbours	**4838.220**	**2.89e+08**	**0.332****
KNN = 30 closest neighbours	4838.877	2.90e+08	0.379**
$1/d$	4843.141	2.96e+08	0.279
$1/d^2$	4840.640	2.91e+08	0.214*
$1/d^3$	4839.912	2.90e+08	0.170**
$\exp(-d/\bar{d})$	4844.536	2.99e+08	0.054

d – geographical distance. Obliczono na podstawie modelu dynamicznego SAR-FEM. Significant at the * 10% level, ** 5% level, *** 1% level.

The weighting matrix **W** offering the best results has been determined using an algorithm based on k-closest neighbours, assuming that each region relates to its 25 closest areas and the strength of the interaction is universal across them. Interestingly, the value of the spatial parameter ρ for this matrix is close to its maximum value obtained in the calculations. Based on those results the spatial weight matrix W used in the remainder of this paper uses *k=25.*

5.3. Monte Carlo simulation results

As indicated in section 5.1 estimating a dynamic SAR-FEM model using a QMLE method requires a relatively large number of entities (N) and observations (T). Since in this research T is small a Monte Carlo experiment has been conducted to estimate the bias of parameter estimators. For the simulation N=73 and T=4 (identical to the sample used in this study) and a data-generating process expressed by the equation presented below (as in equation 5) have been adopted:

$$\mathbf{Y} = (\mathbf{I} - \rho\mathbf{W})^{-1}(\phi\mathbf{Y}_{t-1} + \boldsymbol{\beta}\mathbf{X} + \boldsymbol{\mu} + \boldsymbol{\varepsilon}), \tag{10}$$

The equation uses a vector of starting-values $\mathbf{Y}_{t0}$ ~ N(0, $\mathbf{I}_N$). Values **X**, **ε** i **μ** were generated independently form a normal distribution while elements of matrix **W** were calculated using an algorithm of k=25 closest neighbours (see section 5.2). Target parameter values were ρ = 0.3, ϕ = 0.2, **β** = [-10, -5, 10, 20]' for scenario 1 and ρ = 0.6, ϕ = 0.1, **β** = [-5, -0.2, 0.3, 1.5]' for scenario 2. The accuracy of the estimates was indicated using two measures commonly used in the evaluation of the simulation results: relative bias of an estimator $\hat{\beta}$ for parameter β (RB) and rate of the coverage (based on 95% confidence interval) (CR). Number of iterations was 10,000 in each case. Table 2 presents the obtained results.

Table 2. Performance of quasi-MLE estimator in the case of panel data with small *T* and medium *N*.

Scenario	Parameter	RB (%)	CR (%)
1	β_1	-0.006	0.8785
	β_2	-0.089	0.8797
	β_3	-0.003	0.8813
	β_4	-0.004	0.8779
	ρ	-0.127	0.8834
	ϕ	-0.006	0.8832
2	β_1	-0.418	0.8658
	β_2	-0.019	0.8776
	β_3	-0.033	0.8772
	β_4	-0.048	0.8791
	ρ	-0.997	0.8791
	ϕ	-0.261	0.8767

As expected, the majority of estimated parameters have RB values higher while CR values are somewhat higher than those obtained by Yu et al. (2008). This is especially evident when the results are compared to the values reported by the authors for a small T (relative to N). For $T = 10$ i $N = 196$ they reported RB values ranging from -0.0250 to 0.0003 with coverage probability of 0.9020-0.9390 (not considering σ^2). Results presented in this paper appear to suggest that reducing T with a relatively large N increases the bias of estimates.

The question critical for this research is if the QMLE approach for the dynamic SAR-FEM model will allow drawing correct conclusions from the available panel data. The bias of estimates does not exceed 1% and appears acceptable since Hoogland and Boomsma (1998) argue that unbiased estimates are those for which the relative bias is less than 5%. Standard errors appear to be biased more but the value of the CR indicator is close to 90% in all cases. In this context it can be concluded that both the model and the estimation method are appropriate for the data.

6. Data description

This article examines labour markets with restricted supply of housing and developing local economies based on highly skilled labour. University towns appear to offer some unique characteristics that make them excellent research subjects for this research. Economies in university towns are traditionally highly dependent on the skills of the local university graduates and researchers. Companies often locate themselves close to reputable education centres in order to be able to take advantage of research collaboration opportunities, spillover effects and gain access to a skilled labour force (Combs and Durnaton 2006, Guerrero and Urbano 2014). This is especially true for towns of moderate size. Furthermore many university towns in the UK have been recognized as such for long time periods and for historical and cultural reasons have strict planning regulations governing the development of new housing. This limits opportunities for re-development of land (Barker 2008, Kim et al. 2012) and reduces elasticity of housing supply within those locations. In addition, many UK cities have introduced “green belts” which are green strips of land around town borders on which development is not permitted. This prohibits outward urban growth and further constrains the supply of new dwellings. It would appear that certain UK university towns may be ideal examples of markets that have inelastic housing supply and benefit from developing local economies that rely on highly skilled labour.

However, in order to study the effect of housing supply in restricted markets a change in the stock of dwellings and an expansion of employment are required. Over the last two decades many UK universities faced a growing demand from both students and local employers to expand their facilities. In response to this demand many institutions took extraordinary measures to create both new housing units for their staff and students and research facilities shared with local businesses (D'este and Patel 2007). This involved working with local authorities to obtain special permissions to either relax the restrictions on redevelopment of existing buildings or develop land classified as part of the green belt. Since this resulted in increased employment and supply of dwellings in otherwise inelastic housing markets those locations appear to be highly interesting for research of the effect of such action. As explained in preceding chapters this is especially true for Cambridge as the town has experienced an unprecedented growth in both economic activity and house prices while the university has developed its world-leading facilities.

Table 2. Summary statistics of the data

Location	Variable	min	mean	max
2001	Index value	975970.5	1053217	1221205
	housing stock	1518	3117.397	5305
	Income	486	699.3288	920
	Unemployment	2.4	3.210959	4.4
	Index value in Cambridge	1079007	6501119	6576427
	New supply	0	22.05479	367
	Research centre*	39.47318	202.6517	1141.594
2004	Index value	989387.5	1113750	1244288
	housing stock	2030	3187.945	5347
	Income	480	705.7534	890
	Unemployment	2.4	3.305479	4.6
	Index value in Cambridge	6960606	6960606	6960606
	New supply	0	21.26027	116
	Research centre*	47.6012	241.4511	1286.977
2007	Index value	982140.8	1161392	1320089
	housing stock	2058	3257.808	5467
	Income	500	764.7945	1020
	Unemployment	3.1	3.872603	5.3
	Index value in Cambridge	7221957	7221957	7221957
	New supply	0	20.71233	221
	Research centre*	51.54852	261.7162	1401.136
2011	Index value	1028906	1154866	1358617
	housing stock	2082	3325.315	5517
	Income	434	618.137	752
	Unemployment	3.9	5.668493	9.2
	Index value in Cambridge	7119444	7119444	7119444
	New supply	0	17.49315	198
	Research centre*	69.66153	350.313	1790.449

*the research centre variable is the spatially adjusted number of research centres at West Cambridge and Addenbrooke's sites

The majority of the data collected for this study is publicly available from the UK government. The Land Registry provided data on all house transactions in Cambridgeshire. This information was supplemented with the data from the Office for National Statistics on Small Area Model-Based Income Estimates. However, the information on income is not available at the same level of geographical detail as the transactional data. In order to match the two datasets all sales transactions have been grouped at a middle layer output area level (see the website of the UK office for national statistics for details) using a normalized model-based index of prices controlling for the type of building and identifying new structures and leasehold transactions. In addition, the fact that the income data is only available for certain years the study had been limited to years 2000, 2004, 2008 and 2011. Information on the geographical location of major employment centres has been obtained from local council reports. Expansion of research facilities was approximated by the number of university research centres opened at a particular location. Reliable data on the total dwelling stock in the period of interest proved difficult to obtain. It was estimated by adjusting the total stock reported by the 2011 population census for any new additions. New supply was estimated based on the number of newly built houses sold in a particular location in a particular year reported by the Land Registry database. Although this may not be a perfect approximation, it has been found that the correlation of the data obtained through this process with numbers reported by local authorities was around 70%. Prices in past periods and their growth used for estimation have been taken from intermediate periods between years of income measurements. Overall the study investigates 73 locations in Cambridgeshire over 4 time periods.

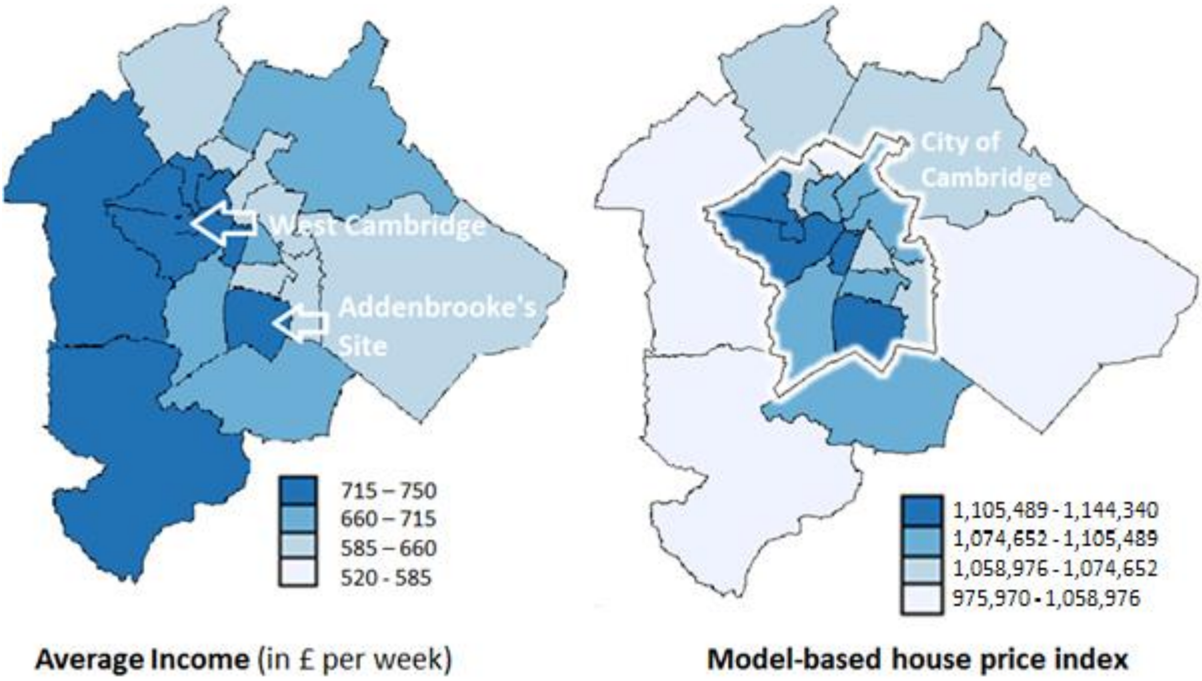

Figure 1. Levels of income and price index in middle layer super output areas in Cambridge in 2011.

Interestingly figure 1 appear to suggest that on average the correlation between income and value of the price index is quite low in Cambridge. However, around the main research areas of Cambridge University (West Cambridge and Addenbrooke’s Site) both incomes and house prices appear to be relatively high.

7. Results and analysis

At first a simple OLS method was used to estimate a FEM model without time and spatial lags (column 2 table 4). The results indicate that increases in income in the previous period, average prices in Cambridgeshire, employment and size of employment centres significantly influence house price in all studied areas. Both stock variables are insignificant. Due to an expectation of spatial effects based on the logic outlined above in chapter 5 residuals of the FEM model were tested for spatial autocorrelation using the Lagrange multiplier test for the lagged dependant variable (LM-LAG) and spatial autocorrelation of residuals (LM-ERR). Following the work of Elhorst (2014) both tests were also performed using robust estimates. Testing results presented in table 3 confirm that a null hypothesis of no spatial autoregression can be rejected while one of no spatial autocorrelation cannot. This is confirmed by robust estimates and leads to a conclusion that spatial effect take the form of autoregression but not autocorrelation. This confirms that a panel model with spatial autoregression (SAR-FEM) is appropriate for this research.

Table 3. Spatial correlation tests.

Test	Statistic	p-value
LAG-LM	6.986	0.008
ERR-LM	2.207	0.137
Robust LAG-LM	4.789	0.029
Robust ERR-LM	0.010	0.920

Estimation results for this model are presented in column 3 of table 4. It allows spatial but not temporal autoregression. Including spatial effects had no effect on the signs of estimated parameters nor on their significance. The spatial parameter ρ is also significant and positive which confirms the expectation of spatial effects. Using the Akaike criterion (AIC) as an indicator the SAR-FM model proved superior to the FEM alternative.

Table 4. Estimation results for FEM, SAR-FEM and dynamic SAR-FEM models

Variable	FEM	SAR-FEM	Dynamic SAR-FEM
I_{it-1}	75.014*** (22.204)	67.495*** (17.663)	59.125*** (17.850)
S_{it}	-18.769 (18.094)	-12.129 (14.425)	-14.892 (14.404)
$wS_{ctr\ t-1}$	-1.072 (10.622)	-14.429 (9.721)	-21.555** (9.995)
U_{it}	-3661.221** (1533.359)	-2504.848** (1278.990)	-4244.261*** (1411.642)
ω_{it} (apc)	0.180*** (0.015)	0.119*** (0.025)	0.101*** (0.026)
wEC_{it}	186.599*** (44.091)	183.880*** (34.646)	163.691*** (35.262)
ϕ			0.168*** (0.059)
ρ		0.368*** (0.137)	0.332** (0.136)
AIC	4846	4840	4838

Standard error in parentheses. Significance at the * 10% level, ** 5% level, *** 1% level.

The next step was estimating the dynamic SAR-FEM model allowing temporal autorgerssion. The results presented in table 4 show that the lagged value of the dependant variable is positive and significant. Interestingly, including the temporally lagged parameter did not appear to yield estimates different to ones obtained by other models with the exception of $wS_{ctr\ t-1}$ which became much lower and statistically significant and unemployment which doubled in its negative influence reported by the static SAR-FEM model. Comparing the two models using AIC values confirmed that the dynamic approach was superior.

Since spatial lags have been included in the model interpreting the results requires estimating direct, indirect and total effects. For the preferred dynamic SAR-FM model these have to be repeated for short and long terms. These are presented in table 5.

Table 5. Direct, indirect and total impacts for the dynamic SAR-FEM model.

Variable	Type of effect	Direct effects	Indirect effects	Total effects
I_{it-1}	Short-run	60.754*** (19.441)	30.303 (23.804)	91.056*** (35.251125)
S_{it}	Short-run	-15.077 (13.775)	-6.746 (8.880)	-21.823 (21.178)
$wS_{ctr\ t-1}$	Short-run	-19.685* (10.325)	-12.039 (12.324)	-31.724 (21.388)
U_{it}	Short-run	-4471.712*** (1396.031)	-2115.562 (1619.708)	-6587.274*** (2320.532)
ω_{it} (apc)	Short-run	0.104*** (0.029)	0.045** (0.022)	0.148*** (0.022)
wEC_{it}	Short-run	162.749*** (34.443)	82.782 (59.878)	245.530*** (78.429)
I_{it-1}	Long-run	4877.375* (2807.506)	3398.648 (3674.478)	8276.022 (5439.253)
S_{it}	Long-run	498.113 (600.978)	317.280 (519.416)	815.393 (1044.076)
$wS_{ctr\ t-1}$	Long-run	591.157 (537.129)	591.736 (907.181)	1182.893 (1386.407)
U_{it}	Long-run	2.63e+07* (1.51e+07)	1.68e+07 (1.95e+07)	4.31e+07 (2.87e+07)
ω_{it} (apc)	Long-run	0.014* (0.008)	0.007** (0.003)	0.021*** (0.007)
wEC_{it}	Long-run	33181.940** (13855.760)	24062.15 (23263.45)	57244.1* (32493.13)

Significant at the * 10% level, ** 5% level, *** 1% level.

Most explanatory variables (I_{it-1}, U_{it}, ω_{it}, wEC_{it}) are significant over both short and long terms. Direct effect for these variables are close to the parameters estimated in table 4. This can be interpreted as an indicator that the feedback effect of these factors influencing prices in area *i* through affecting prices in neighbouring locations is very small. For example the parameter for lagged income $I_{it-1}$is 59 while the short term direct effect is less than 61 suggesting that the feedback effect is practically almost negligible. The only variable that shows significant indirect effects is the spatially weighted average price in Cambridgeshire which can be expected since the variable reflects the spatial distribution of macroeconomic variables.

8. Discussion

The results presented above clearly show that both spatial and temporal effects are significant in modelling house prices. Most importantly however, they also appear to give support to the theory outlined earlier in this paper. The claim that the Cambridgeshire market has restricted supply finds support in the finding of no impact of supply on prices. With perfectly elastic demand prices appear to be determined mainly by income, employment and macroeconomic factors. Furthermore, only supply in Cambridge appears to be able to reduce prices in the examined region showing the importance of this location relative to all other areas. It would appear that there is a strong preference for living in Cambridge with prices around it affected considerably by a spatial spill-over. However, even spatial interactions do not end there. Even within the city there are locations that attract much higher prices.

Growth of major employment centres appears to significantly increase the attractiveness of the neighbouring areas to house buyers. It is important to note that this effect is registered in addition to controls for income growth and overall economic growth. It appears that the proportion of income dedicated to purchasing a house changes with the size of the nearest employment centre. This suggests that the value derived from living close to a larger centre of employment is determine both by the size of the hub and distance from it.

Assuming that physical characteristics of individual hoses do not change there has to be an additional value generated by a larger employment centre. The results appear to suggest that the feature being priced differently is proximity to the desired location. As centres grow in size being located close to them is becoming more and more valuable.

This is interesting as it signals that in growing regional economies spatial effects of employment centres on housing markets are transferred not only through income and macroeconomic variables but also an increased value effect of distance to locations. The finding has some very important economic implications.

9. Conclusions

With restricted housing supply leading to limited population this may create labour shortages. Newcomers may be able to afford to price locals out of the most expensive areas but the in order to continue to develop the regional economy will require both current end additional workers. With house prices growing at a faster rate than income and clustering clearly around centres of employment those hubs may face upwards pressure on wages both from existing and new staff. This is likely to erode the marginal profit from increasing productivity of labour by making it more expensive not only to hire new workers but also retain existing ones. In this context, it appears that regional economies with restricted supply of housing are more likely to concentrate on increasing output through strategies relating to increasing labour productivity rather than on expanding employment (although the latter is also possible). This emphasizes the importance of skills and education of workers in such an economy. Only the most productive employees will be able to get and keep jobs under those circumstances.

Nevertheless, those positions appear to be in very high demand as houses around employments centres rise faster than average income. This leads to the conclusion that growing regional economies create value to households that has not been previously captured. The most productive individuals appear to be willing to sacrifice a greater proportion of their income in order to locate themselves closer to employment centres.

In conclusion, it appears that companies located in employment centres with restricted housing supply are forced to hire and retain only the best employees due to a limited population size. However, the wage they have to pay in order to attract those employees is not proportional the increase in living costs in the area. Consequently, it appears that by clustering together in large employment centres companies are able to create value to their employees at no additional direct cost to themselves. While it is unclear what exactly is the source of this value the effect appears to allow businesses to hire employees at below the wage rate which would be dictated by a parity of disable income.

10. References

Basu, S., & Thibodeau, T. (1998). Analysis of Spatial Autocorrelation in House Prices. *Journal of Real Estate Finance and Economics*, *17*(1), 61–85. http://doi.org/10.1023/A:1007703229507

Battu, H., Ma, A., & Phimister, E. (2008). Housing tenure, job mobility and unemployment in the UK. In *Economic Journal* (Vol. 118, pp. 311–328).

Booth, A. L., Francesconi, M., & Garcia-Serrano, C. (1999). Job tenure and job mobility in Britain. *Industrial and Labor Relations Review*, *53*(1), 43–70.

Bresnahan, T., Gambardella, A., & Saxenian, A. (2001). “Old Economy” Inputs for “New Economy” Outcomes: Cluster Formation in the New Silicon Valleys. *Industrial and Corporate Change*, *10*(4), 835–860. http://doi.org/10.1093/icc/10.4.835

Chakrabarti, R., & Zhang, J. (2010). *Unaffordable housing and local employment growth: Evidence from California municipalities*. *Urban Studies* (Vol. 10). Retrieved from http://usj.sagepub.com/content/early/2014/05/09/0042098014532854?papetoc\nhttp://usj.sagepub.com/cgi/doi/10.1177/0042098014532854

Cho, M. (1996). House Price Dynamics : A Survey of Theoretical and Empirical Issues. *Journal of Housing Research*, *7*(1995), 145–172.

Christopherson, S., Michie, J., & Tyler, P. (2010). Regional resilience: Theoretical and empirical perspectives. *Cambridge Journal of Regions, Economy and Society*. http://doi.org/10.1093/cjres/rsq004

Di Addario, S., & Patacchini, E. (2008). Wages and the City. Evidence from Italy. *Labour Economics*, *15*(5), 1040–1061. http://doi.org/10.1016/j.labeco.2007.09.003

Duffy, D., Gerald, J. F., & Kearney, I. (2005). Rising house prices in an open labour market. *Economic and Social Review*, *36*(3), 251–272.

Etzkowitz, H. (2012). Silicon Valley: The Sustainability of an Innovative Region. *Social Science Information*, *52*(3), 1–25. Retrieved from http://www.triplehelixassociation.org/wp-content/uploads/2012/07/Silicon-Valley-The-Sustainability-of-an-Innovative-Region.pdf

FITZPATRICK, T., & MCQUINN, K. (2007). HOUSE PRICES AND MORTGAGE CREDIT: EMPIRICAL EVIDENCE FOR IRELAND. *Manchester School (14636786)*, *75*(1), 82–103.

Gallin, J. (2006). The long-run relationship between house prices and income: Evidence from local housing markets. *Real Estate Economics*, *34*(3), 417–438.

Geerlings, H., & Stead, D. (2003). The integration of land use planning, transport and environment in European policy and research. *Transport Policy*, *10*(3), 187–196. http://doi.org/10.1016/S0967-070X(03)00020-9

Girouard, N., Kennedy, M., & Noord, P. Van Den. (2005). Recent house price developments: The role of fundamentals. *OECD Economic Outlook*, (2005), 123–154. http://doi.org/http://dx.doi.org/10.1787/864035447847

Glaeser, E. L. (2010). Agglomeration Economics. *Nber*, *I*(February), 0–21. http://doi.org/10.7208/chicago/9780226297927.001.0001

Glaeser, E. L., Gyourko, J., & Saiz, A. (2008). Housing supply and housing bubbles. *Journal of Urban Economics*, *64*(2), 198–217.

Glaeser, E. L., Gyourko, J., & Saiz, A. (2008). Housing supply and housing bubbles. *Journal of Urban Economics*, *64*(2), 198–217. http://doi.org/10.1016/j.jue.2008.07.007

Goetz, S. J., & Rupasingha, A. (2002). High-tech firm clustering: Implications for rural areas. *American Journal of Agricultural Economics*, *84*(5), 1229–1236. http://doi.org/10.1111/1467-8276.00383

Granziera, E., & Kozicki, S. (2015). House price dynamics: Fundamentals and expectations. *Journal of Economic Dynamics and Control*, *60*, 152–165.

Haynes, K. E. (1997). Labor markets and regional transportation improvements: the case of high-speed trains. *The Annals of Regional Science*, *31*(1), 57–76.

Hiebert, P., & Roma, M. (2010). Relative house price dynamics across euro area and US cities Convergence or divergence ? *ECB Working Paper Series*, *1206*(June 2010). Retrieved from http://ecomod.net/sites/default/files/document-conference/ecomod2010/1226.pdf

Himmelberg, C., Mayer, C., & Sinai, T. (2005). Assessing High House Prices: Bubbles, Fundamentals and Misperceptions. *Journal of Economic Perspectives*, *19*(4), 67–92.

Hospers, G., & Beugelsdijk, S. (2002). Regional Cluster Policies: Learning by Comparing? *Kyklos*, *55*(3), 381–402. http://doi.org/10.1111/1467-6435.00192

Hsiehchen, D., Espinoza, M., & Hsieh, A. (2015). Multinational teams and diseconomies of scale in collaborative research. *Science Advances*, *1*(8), e1500211–e1500211. http://doi.org/10.1126/sciadv.1500211

Iacoviello, M., & Neri, S. (2010). Housing Market Spillovers: Evidence from an Estimated DSGE Model. *American Economic Journal: Macroeconomics*, *2*(2), 125–164.

Johansson, B., & Quigley, J. M. (2003). Agglomeration and networks in spatial economies. *Papers in Regional Science*, *83*(1), 165–176.

Kiel, K. A., & Zabel, J. E. (2008). Location, location, location: The 3L Approach to house price determination. *Journal of Housing Economics*, *17*(2), 175–190. http://doi.org/10.1016/j.jhe.2007.12.002

Klepper, S. (2001). Employee Startups in High-Tech Industries. *Industrial and Corporate Change*, *10*(3), 639–674. http://doi.org/10.1093/icc/10.3.639

Lambson, V. E., McQueen, G. R., & Slade, B. A. (2004). Do Out-of-State Buyers Pay More for Real Estate? An Examination of Anchoring-Induced Bias and Search Costs. *Real Estate Economics*.

Miles, D. (2012). Population Density, House Prices and Mortgage Design. *Scottish Journal of Political Economy*, *59*(5), 444–466. http://doi.org/10.1111/j.1467-9485.2012.00589.x

Mueser, P. R., & Graves, P. E. (1995). Examining the Role of Economic Opportunity and Amenities in Explaining Population Redistribution. *Journal of Urban Economics*, *37*(2), 176–200. Retrieved from http://www.sciencedirect.com/science/article/pii/S0094119085710108

Nunes, P. M., Serrasqueiro, Z., & Leitão, J. (2012). Is there a linear relationship between R&D intensity and growth? Empirical evidence of non-high-tech vs. high-tech SMEs. *Research Policy*, *41*(1), 36–53. http://doi.org/10.1016/j.respol.2011.08.011

Ortalo-Magné, F., & Sven, R. (2006). Housing market dynamics: On the contribution of income shocks and credit constraints. *Review of Economic Studies*, *73*(2), 459–485.

Quigley, J. M., & Raphael, S. (2004). Is Housing Unaffordable? Why Isn't It More Affordable? *Journal of Economic Perspectives*, *18*(1), 191–214. http://doi.org/10.1257/089533004773563494

Richardson, H. W. (1995). Economies and Diseconomies of Agglomeration. In *Urban Agglomeration and Economic Growth* (pp. 123–155). http://doi.org/10.1007/978-3-642-79397-4

Wachs, M., & Kumagai, T. G. (1973). Physical accessibility as a social indicator. *Socio-Economic Planning Sciences*, *7*(5), 437–456.

Zabel, J. E. (1999). Controlling for Quality in House Price Indices. *Journal of Real Estate Finance and Economics*, *19; Jg. 19*(3), 223–41. http://doi.org/10.1023/A:1007832527244

Bouayad-Agha, S., & Védrine, L. (2010). Estimation strategies for a spatial dynamic panel using GMM. A new approach to the convergence issue of European regions. *Spatial Economic Analysis*, 5(2), 205-227. http://doi.org/10.1080/17421771003730711

Arellano, M., & Bond, S. (1991). Some tests of specification for panel data: Monte Carlo evidence and an application to employment equations. The review of economic studies, 58(2), 277-297.

Blundell, R., & Bond, S. (1998). Initial conditions and moment restrictions in dynamic panel data models. Journal of econometrics, 87(1), 115-143.

Elhorst, J. P. (2003). Specification and estimation of spatial panel data models. International regional science review, 26(3), 244-268.

Kukenova, M., & Monteiro, J. A. (2008). Spatial dynamic panel model and system GMM: a Monte Carlo investigation. Available at SSRN 1300871.

Elhorst, J. P. (2012). Dynamic spatial panels: models, methods, and inferences. Journal of geographical systems, 14(1), 5-28.

Yang, Z. (2015). Unified M-Estimation of Fixed-Effects Spatial Dynamic Models with Short Panels. Research Collection School Of Economics. Retrieved from http://ink.library.smu.edu.sg.

Yu, J., De Jong, R., & Lee, L. F. (2008). Quasi-maximum likelihood estimators for spatial dynamic panel data with fixed effects when both n and T are large. Journal of Econometrics, 146(1), 118-134.

Bhargava A, Sargan JD (1983) Estimating dynamic random effects models from panel data covering short time periods. Econometrica 51(6):1635–1659.

Elhorst, J. P. (2005). Unconditional Maximum Likelihood Estimation of Linear and Log-Linear Dynamic Models for Spatial Panels. Geographical analysis, 37(1), 85-106.

Hsiao, C., Pesaran, M. H., & Tahmiscioglu, A. K. (2002). Maximum likelihood estimation of fixed effects dynamic panel data models covering short time periods. Journal of econometrics, 109(1), 107-150.

Nanda, A., & Yeh, J. H. (2014). Spatio-temporal diffusion of residential land prices across Taipei regions. SpringerPlus, 3(1), 1.

Belotti, F., Hughes, G., & Mortari, A. P. (2014). XSMLE: Stata module for spatial panel data models estimation. Statistical Software Components.

Belotti, F., Hughes, G., & Piano Mortari, A. (2016). Spatial panel data models using Stata. CEIS Research Papers, Vol. 14, Issue 5, No. 373.

Ezcurra, R., & Rios, V. (2015). Volatility and regional growth in Europe: Does space matter?. Spatial Economic Analysis, 10(3), 344-368.

Elhorst, J. P. (2014). Spatial panel data models. In Spatial Econometrics (pp. 37-93). Springer Berlin Heidelberg.