The spatial impact of employment centres on housing markets.

Abstract: Local economic growth tends to affect neighbourhood house prices unevenly. It has been observed that prime locations experience price hikes far in excess of the surrounding local area. Yet, this phenomenon is not well captured by existing economic models. This research provides a model of spatial and temporal interactions between housing and employment markets. The results show that rapid growth of employment centres increases house prices in neighbouring locations even after adjusting for fundamentals. The study concludes that spatial clustering of companies creates an option value for existing and potential employees that goes beyond ease of access for commuting purposes.

Keywords: Housing markets, labour markets, employment centres, spatial regional economics

JEL codes: J21, J31, O18, R11

Authors: Nikodem Szumilo[1], Edyta Laszkiewicz[2], Franz Fuerst[1&3]

1. University of Cambridge

   Department of Land Economy

   Silver Street, CB3 9EP, Cambridge, UK

2. University of Lodz

   Faculty of Economics and Sociology

   Rewolucji 1905, 90-214 Lodz, Poland

3. University of Melbourne

   Faculty of Architecture, Building and Planning

   Masson Road, VIC 3010, Melbourne, Australia

1. Introduction

Over the past two decades, house prices in most developed economies have increased at an unprecedented rate (Knoll et al 2014). As a consequence, many urban centres around the world have become increasingly unaffordable (Chakrabarti and Zhang 2010, Quigley and Raphael 2004). Some studies argue that the combination of rising incomes and historically low interest rates have simultaneously increased the purchasing power and borrowing capabilities of households and individuals which in turn has fuelled house price growth (McQuinn and O'Reilly 2008). Others point 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, supply inelasticity due to limited availability of land and/or planning restrictions has been identified as a contributing factor (Knoll et al. 2014). 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 motorised transportation which leads to considerable environmental concerns (Miles 2012).

While many house price models been created, they mainly adopt very wide (Pain and Westaway, 1997; Brown et al., 1997) or detailed views (Sirmans et al. 2005; Mills and Simenauer, 1996). More recent literature indicates that incorporating local regional and national factors into house price modelling can be very important (Hwang and Quigley, 2006) but also that their spatial interactions can significantly affect the results (Case et al. 2004). 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) but few have attempted to explain this disparity (Faggio and Overman, 2014). Rather superficially, this phenomenon has been attributed to the economic success of those areas and considered as one of many 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). Better economic performance is associated with higher income. Assuming that the proportion of income paid for housing is fixed and the quality of the housing stock remains largely constant, house prices should grow in line with average local incomes. This logic assumes that the attractiveness of a location is determined by the average income of an area and a number of other factors, such as housing stock characteristics, that do not change significantly over time.

However, even locations with identical average income often experience vastly different rates of house price growth (Campbell and Cocco 2007, Case and Shiller 2003, Glaeser et al. 2001). This suggests that the housing proportion of income spending is variable over time and across space. Campbell and Cocco (2007) identify a number of factors that change over time such as credit availability or consumer confidence. However, spatial effects appear more difficult to measure (Sirmans et al. 2005). This study uses a dynamic panel data model to show that expansion of employment centres can have a sizable effect on house prices within their immediate areas that is not explained by growing income or better quality of housing. 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, increases labour costs and constrains local economic growth. Section 2 provides an overview of the residential property market in Cambridge, an explanation of why it is an ideal example for this research and a discussion of how its results can provide insights into other markets. Section 3 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 in sections 5 and 6 (respectively) with special attention given to the problem of time-variable and spatial determinants of house prices. Empirical results are presented and discussed in section 7 and conclusions in section 8 .

2. The curious case of Cambridge, UK

In some ways every market for residential property is unique while in others all locations are similar. Cambridge is characterised by rapid economic development and restricted housing supply. The rate of expansion and the severity of the constraints on housing development in Cambridge make it stand out but both processes occur in many other locations which can learn from this extreme case. After the recession of 2007 house prices in Cambridge grew the fastest of all cities in the country and increased by 44.7%. Similar to 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 in 2014 with the intention of renting them out. This trend is also encouraged by the fact that a considerable proportion of the population is affiliated with a large university and many of its students and junior staff require housing only temporarily (Jones 2007). This results in 27.8% of houses in Cambridge being rented accommodation (2011 census) with tenants coming from all socio-economic groups. It creates a vibrant and relatively developed rental market and dynamic rental growth patterns (52% growth from 2009 to 2014).

With increasing population as well as house prices, levels of new supply have also risen (CCCRP 2013). In 2014 alone, housing stock expanded by around 2.5% and a similar annual growth rate 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 properties. This is due to the fact that supply of land in Cambridge is considerably limited by planning policies and the Green Belt. Until the mid-1990s, Cambridge was constrained by a very strict planning policy focussed on preserving its historic centre. Since then, the overall planning policy has shifted towards supporting sustainable economic growth.

The city of Cambridge is most commonly associated with its ancient university. In fact, Oxford Economics reports that education was the biggest employment sector in the Cambridge region in 2014 and had grown by 21% over the preceding decade. The university alone employed 9,500 people and is a considerable force driving the local economy. 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 employed by the university. In fact, the area has attracted so many companies of this kind that it has been dubbed the "Silicon Fen". There are around 2,100 high technology companies in the Cambridge area totalling revenues of around £14billion per annum (2015 figures). 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, this 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 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 science 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. The results will be applicable to all locations that experience both economic growth and restricted housing supply.

3. Supply shocks and their impact on demand

By applying their model to US data from 1980-2000, Glaeser et al. (2006) find that the extent to which productivity gains will create bigger cities or just higher paid workers and more expensive housing depends on the elasticity of housing supply. The present study uses a modified version of this model to establish if a localised housing supply shock can subsequently lead to increased demand. The first modification (discussed in this section) is that demand for labour depends on the local level of productivity which is a function of population size. The second modification allows for interactions between neighbouring areas (discussed in section 5).

If a location has flexible housing supply, new housing will only 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). If the location has inelastic housing supply, new housing will not be provided in response to an increase in demand, resulting in heightened competition for the existing housing stock. Glaeser, Gyourko and Saks (2006) specify that the sum of utility and cost of housing in a location equals its flow of amenities and wages. Therefore, the rising income works to bid up the price of housing, returning the utility provided by the location to equilibrium level, whilst the population remains relatively unchanged.

The higher the demand for labour in a particular location is, the higher the wages that will be offered to workers. If a particular set of skills is required, the new workers can come from training up the existing 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 policies 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. Glaeser, Gyourko and Saks (2006) define labour demand as:

$$Log(N_i) = \alpha A_i - \alpha W_i$$

where N is the level of employment, A is a measure of productivity and W are wages. Therefore, if employment level is constant and productivity increases, wages have to grow accordingly.

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 outward 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 rapidly expanding 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 wage. If the local economy expands in size and productivity to reflect the benefit of the new workers, new demand for labour is likely to occur. This maintains the upwards pressure on income as well as on housing demand. An expansion of the labour pool should entail productivity gains which in turn stimulates further economic progress. The labour demand equation would therefore have to reflect area productivity as a sum of productivity defined by A_i and a function of employment ∂N_i. This means that the final equations of housing labour demand and income from Glaeser, Gyourko and Saks (2006) would have to be modified to be:

$$H_i = \frac{\alpha\delta\rho(A_i + \partial N_i) + K_i + \alpha\delta\rho C_i - \delta Log(L_i) - \alpha\delta Ut}{1 + \alpha\delta\rho}$$

$$W_i = \frac{\alpha\delta\rho(A_i + \partial N_i) + \rho K_i - C_i - \delta Log(L_i) - Ut}{1 + \alpha\delta\rho}$$

$$Log(N_i) = \frac{\alpha\delta\rho(A_i + \partial N_i) - \alpha\rho K_i + \alpha C_i + \alpha\delta\rho Log(L_i) - \alpha Ut}{1 + \alpha\delta\rho}$$

where K is location-specific cost of new construction, C is a location-specific flow of amenities, Ut is reservation utility, H are house prices derived from rents by dividing them by a coefficient ρ, α is a coefficient of the labour demand equation, L is land area and δ is a coefficient of the influence of housing density on construction costs.

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. Our model assumes that this increasing marginal return can only be expected from growing local economies with restricted labour supply where the major limitation to growth is unavailability of labour. The response of house prices in markets with constrained population to a supply shock depends on the value of ∂. If expansion of an economy depends on availability of the right workers, ∂ can be expected to be high. It can be concluded that in markets with restricted supply of housing and elastic demand for labour housing supply shocks are unlikely to translate into lower property prices. In fact, it is likely that if the economy is increasing its productivity due to the increased population, income would grow and house prices are more likely to increase than fall.

4. Modelling house prices

While it is clear from the previous section that house prices depend on productivity, employment, amenities and construction costs in an area, it also needs to be noted that there may be interactions between areas. Assuming that the selling price of a property depends on the supply of and demand for dwellings in a location as well as within a commutable distance, it is possible to formulate a basic demand function of the following form:

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

where q_{it} is the demand for housing in district *i* (*i* = 1,…,*N*) at time *t* (*t* = 2,…,*T*), $\mathrm{I}_{\mathrm{it}-1}$ 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 productivity or 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 as (after controlling for unemployment) this reflects labour demand within the area as a function of its productivity. The equation also needs to adjust for the fact that the past increase in prices influences expectations of their future growth as well as macroeconomic conditions driving house prices. The following function allows those factors to be reflected:

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

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 as even over a four year period serial correlation of prices may persist in small areas with fewer transactions (see Nerlove, 1958 and Capozza et al., 2004).

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:

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

where S_{it} is the current housing stock, $wS_{Ctr\,t}$ is the distance-weighted 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. 2), and substituting for q_{it} (Eq. 3) using the demand function, it is possible to arrive at of the following reduced form equation:

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

where $wp_{ctr\,t}$ (denoted as ω_{it} in equation 2 and ς_{it} in equation 3) is the average house price weighted by its geographical distance to Cambridge that proxies for macroeconomic conditions (Bitter et al., 2007; Huang et al., 2010), the $\mathrm{v_{it}}$ 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.

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). Changing house prices in a region affect transactions occurring in neighbouring locations. Hence, any local house price shocks are propagated to surrounding areas. Furthermore, anchoring effects observed in the real estate market result in 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 might 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:

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

where $\mathbf{Y} = [p_{1,t},\dots, p_{N,T}]'$ is a vector of house price for N regions and T time units, $\mathbf{Y}_{t-1}$ 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 reduced 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 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 ($\phi\mathbf{Y}_{t-1}$). Hence, Eq. 5 implies global spatial effects (Hassan, 2017). 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 are widely discussed in the 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 out 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, 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 considers a panel model in which N is large and T is fixed, Yu et al. concentrate on a data structure in which both N and T are large. Two different ways of incorporating individual fixed-effects are proposed based on the differences in the data; Elhorst uses firstdifferenced and Yu at al. demean 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 (small T), therefore, approximation procedures need to be used. It is possible to use the Bhargava and Sargan (1983) approximation, however, Su and Yang (2015) propose that using a quasi-MLE estimator and modelling initial differences with an adaptation of Hsiao's 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, the 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 is 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 are guided 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 $\rho = 0$ and $\phi = 0$) and SAR-FEM (with $\phi = 0$).

In spatial autoregressive models, conclusions based solely on coefficients of explanatory variables are biased due to spatial spillover effects (LeSage and 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:

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

- Short-term indirect effect:

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

- Long-term indirect effect:

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

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

5.2. Spatial weights matrix

One of the key elements of a spatial econometric model is the spatial structure of relationships between entities represented by a matrix **W**. This paper follows the work of Ezcurra and Rios (2015) and compares a number of spatial expression methods. Maximising the value of the Log-Likelihood function or minimising the residual variance of the estimated model can be adopted as selection criteria. Alternatively, the value of the Bayesian posterior model probability can be used. This research focused on minimising 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 ---

The weighting matrix **W** offering the best results is 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 these results, the spatial weight matrix W used in the remainder of this paper assumes *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 the T in this empirical analysis is small, a Monte Carlo experiment is 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) are 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 $\mathbf{X}$, $\boldsymbol{\varepsilon}$ and $\boldsymbol{\mu}$ are generated independently form a normal distribution while elements of matrix W is calculated using an algorithm of k=25 closest neighbours (see section 5.2). Target parameter values are ρ = 0.3, ϕ = 0.2, $\boldsymbol{\beta}$ = [-10, -5, 10, 20]' for scenario 1 and ρ = 0.6, ϕ = 0.1, $\boldsymbol{\beta}$ = [-5, -0.2, 0.3, 1.5]' for scenario 2. The accuracy of the estimates is indicated by two common measures: 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 is 10,000 in each case. Table 2 presents the obtained results.

--- Table 2 ---

As expected, the majority of estimated parameters have RB values higher than those obtained by Yu et al (2008) while CR values are lower. 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 and 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. Economies in university towns are traditionally highly dependent on the skills of the local university graduates and researchers. Companies often locate close to centres of education to take advantage of research collaboration opportunities, spillover effects and to 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 recognised as such for centuries and have strict planning regulations (for historical and cultural reasons) governing the development of new housing. This limits opportunities for re-development of land (Barker 2008) and reduces elasticity of housing supply within those locations. In addition, many UK cities have introduced "green belts". This prohibits outward urban growth and further constrains the supply of new dwellings. Some UK university towns may be ideal examples of markets that have restricted 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 or 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 restricted housing markets, those locations appear to be highly interesting for research of the effect of such action. This is particularly true for Cambridge which has experienced unprecedented growth in both economic activity and house prices while the university has developed its facilities and harnessed its reputation as one of the leading research universities in the world.

--- Table 3 ---

The majority of the data collected for this study is publicly available from the UK government (see the appendix for a list of sources). The Land Registry provides data on all house transactions in Cambridgeshire. This information is then supplemented with the data from the Office for National Statistics on Small Area Model-Based Income Estimates[1]. 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 are grouped at middle layer super output area levels[2] using a normalised model-based index of prices controlling for the type of building and identifying new structures and leasehold transactions. This approach also allows converting a spatio-temporal dataset into a spatial panel (see Thanos et al. 2016 for a discussion of the difference). Since income data is only available for certain years, the study is 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

[1] ONS produces four measures of mean weekly household income: total; net; Equalised before house costs; and equalized after housing costs. The method for producing the estimates involves combining data from the Family Resources Survey (FRS) with relevant administrative data sources (including benefit claimant counts, council tax bandings and tax credit claims). ONS produces a model which describes the relationship between the survey and administrative data. It then applies this relationship to the administrative sources at the small area level to produce estimates of weekly household income. ONS constrains the regional income estimates to the equivalent FRS regional statistics. (UK Statistics Authority 2011)

[2] Super Output Areas are a geography for the collection and publication of small area statistics. They are used on the Neighbourhood Statistics site and across National Statistics. Middle Layer SOAs are generated automatically by zone-design software using census data from groups of LSOAs. They have a minimum size of 5,000 residents and 2,000 households with an average population size of 7,500. They fit within local authority boundaries.

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 is around 70%. While this introduces a measurement bias into the data, its effect on market estimates is limited as dwellings not captured by this procedure are usually not available for purchase by the general public and included mostly built-to-rent student properties designated for temporary residents. Nevertheless, changes in the size of the housing stock are not tracked with perfect accuracy. Prices in past periods and their growth used for estimation are 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 appears 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 are relatively high.

7. Results and analysis

At first, a simple OLS method is used to estimate a FEM model without time and spatial lags (column 2 table 5). The results indicate that increases in income in the previous period, average prices in Cambridgeshire, employment and size of employment centres significantly influence house prices in all studied areas. Both stock variables are insignificant. Due to an expectation of spatial effects based on the logic outlined above in section 5, residuals of the FEM model are 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 are also performed using robust estimates. Test results presented in table 4 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 effects 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 4 ---

Estimation results for this model are presented in Column 3 of Table 5. It allows spatial but not temporal autoregression. Including spatial effects has no effect on the signs of estimated parameters or 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. In addition, it needs to be noted that the estimates of spatial dependence may actually be biased downwards due to the high density of the spatial weight matrix (Smith, 2009).

--- Table 5 ---

The next step involves estimating the dynamic SAR-FEM model allowing temporal autoregression. The results presented in table 5 show that the lagged value of the dependant variable is positive and significant. Interestingly, including the temporally lagged parameter does not yield estimates different to the ones obtained by other models with the exception of $wS_{ctr\ t-1}$, which becomes 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 confirms that the dynamic approach is superior.

Since spatial lags are 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 6.

--- Table 6 ---

Most explanatory variables (I_{it-1}, U_{it}, ω_{it}, wEC_{it}) are significant over both short and long terms. Direct effects for these variables are close to the parameters estimated in Table 5. This is indicative of a small feedback effect of these factors on prices in area *i* through affecting prices in neighbouring locations. 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. This is not unexpected since the variable reflects the spatial distribution of macroeconomic variables.

7.1 Discussion

The results presented above clearly show that both spatial and temporal effects are significant in modelling house prices. Signs and significance levels of both effects are consistent with findings of similar studies which used hedonic models to examine individual assets (Dubé and Legros, 2014; Nappi-Choulet and Maury, 2011) but the fact that area price indices were used as the dependant variable means that their magnitudes cannot be directly compared.

Most importantly however, the results support the theory that growing employment centres attract an additional premium if housing supply is constrained. The claim that the Cambridgeshire market has restricted supply finds support in the finding of no impact of new supply on price levels. With growing demand, prices appear to be determined mainly by income, employment and macroeconomic factors. Furthermore, only supply in the city of Cambridge appears to reduce prices in the examined region showing the importance of this location relative to all other areas. The evidence indicates a strong preference for living in Cambridge with prices around it affected considerably by a spatial spill-over. However, spatial interactions do not end there as 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. 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 determined both by the size of the hub and distance from it. This is consistent with the findings of Ciccone (2002) who shows that agglomeration benefits in Europe (including the UK) had a significant spatial effect on regional productivity. Increasing employment density of a micro-location may affect all surrounding areas. Ciccone and Hall (1993) find a similar spatial effect on productivity in US cities.

Assuming that physical characteristics of individual houses do not change, a larger employment centre must generate additional economic value. This is consistent with the work of Fik et al. (2003) who find that attributes of locations change over time and affect house prices while asset characteristics may remain constant. The results of this paper appear to suggest that the feature being priced differently is proximity to an employment centre. As centres grow in size, being located close to them is becoming more and more valuable. This result is consistent with the work of Baumont (2004) who shows a positive and significant effect of improving accessibility to the CBD on house prices in Dijon. Rather than using the growth of employment (as this article does) as an indicator of influence of a centre, the author uses its accessibility as a proxy of its influence but reaches a similar conclusion.

The ramifications of this conclusion are pertinent to our present analysis 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 through an increased value effect of distance to locations.

8. Conclusions

This study finds that restricted housing supply may limit population size may create labour shortages when labour demand increases. The example of Cambridge (UK) with its rapidly growing local economy and its severely constrained housing supply yields results that are applicable to any location that suffers from similar problems with housing supply. Irrespective of their causes, housing shortages appear to have a profound impact on regional economic development. In Cambridge this manifested itself by the spatial impact of research centres on house prices but the process is likely to be similar in other locations where employment is expanding but housing stock fails to adopt.

Newcomers may be able to afford to price locals out of the most expensive areas but in order to continue to develop further, the regional economy will require both current and 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. 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 jobs appear to be in very high demand as house prices around employment centres rise faster than the average income. This leads to the conclusion that growing regional economies create value to households that has not been previously captured. The most productive workers appear to be willing to sacrifice a greater proportion of their income in order to locate 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 to 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 the source of this value is unclear, the effect appears to allow businesses to hire employees below the wage rate necessary to obtain the same levels of disposable income elsewhere.

Acknowledgment: The authors would like to thank the Higher Education Funding Council for England for providing funding for a project which inspired this study.

9. References

Barker, K. (2008). Planning policy, planning practice, and housing supply. Oxford Review of Economic Policy, 24(1), 34-49.

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

Baumont, C. (2004). Spatial effects in housing price models: do housing prices capitalize urban development policies in the agglomeration of Dijon (1999)?.

Baumont, C., Ertur, C., & Gallo, J. (2004). Spatial analysis of employment and population density: the case of the agglomeration of Dijon 1999. Geographical analysis, 36(2), 146-176.

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.

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

Bitter, C., Mulligan, G. F., & Dall'erba, S. (2007). Incorporating spatial variation in housing attribute prices: a comparison of geographically weighted regression and the spatial expansion method. Journal of Geographical Systems, 9(1), 7-27.

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.

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.

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),

Campbell, J. Y., & Cocco, J. F. (2007). How do house prices affect consumption? Evidence from micro data. Journal of monetary Economics, 54(3), 591-621.

Capozza, D. R., Hendershott, P. H., & Mack, C. (2004). An anatomy of price dynamics in illiquid markets: analysis and evidence from local housing markets. Real Estate Economics, 32(1), 1-32.

Case, K. E., & Shiller, R. J. (2003). Is there a bubble in the housing market?. Brookings Papers on Economic Activity, 2003(2), 299-342.

CCCRP (2013), Population, Housing and Employment Forecasts Technical Report 2013, Cambridgeshire County Council, Avilable at: http://cambridgeshireinsight.org.uk/housing/current-version/PopHseEmp_TechReport2013

Chakrabarti, R., & Zhang, J. (2010). *Unaffordable housing and local employment growth: Evidence from California municipalities*. *Urban Studies* (Vol. 10). 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*.

Ciccone, A. (2002). Agglomeration effects in Europe. European Economic Review, 46(2), 213-227.

Ciccone, A., & Hall, R. E. (1993). Productivity and the density of economic activity (No. w4313). National Bureau of Economic Research.

Combes, P., Duranton, G. (2006). "Labour pooling, labour poaching and spatial clustering." Regional Science and Urban Economics, 36 (1), 1-28.

D'Este, P., & Patel, P. (2007). University–industry linkages in the UK: What are the factors underlying the variety of interactions with industry?. Research policy, 36(9), 1295-1313.

Dubé, J., Legros, D., 2014. Spatial econometrics and the hedonic pricing model: what about the temporal dimension?, Journal of Property Research 31(4), 333-359

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

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

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

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

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

Englund, P., & Ioannides, Y. M. (1997). House price dynamics: an international empirical perspective. Journal of Housing Economics, 6(2), 119-136.

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

Faggio, G., & Overman, H. (2014). The effect of public sector employment on local labour markets. Journal of urban economics, 79, 91-107.

Fik, T. J., Ling, D. C., & Mulligan, G. F. (2003). Modeling spatial variation in housing prices: a variable interaction approach. Real Estate Economics, 31(4), 623-646.

Forrest, R. (1987). Spatial mobility, tenure mobility, and emerging social divisions in the UK housing market. Environment and Planning A, 19(12), 1611-1630.

Gathergood, J. (2011). Unemployment risk, house price risk and the transition into home ownership in the United Kingdom. Journal of Housing Economics, 20(3), 200-209.

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

Glaeser, E. L., & Nathanson, C. G. (2015). An extrapolative model of house price dynamics (No. w21037). National Bureau of Economic Research.

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., & Saks, R. E. (2006). Urban growth and housing supply. Journal of Economic Geography, 6(1), 71-89.

Glaeser, E. L., Kolko, J., & Saiz, A. (2001). Consumer city. Journal of economic geography, 1(1), 27-50.

Guerrero, M., & Urbano, D. (2014). Academics' start-up intentions and knowledge filters: an individual perspective of the knowledge spillover theory of entrepreneurship. Small Business Economics, 43(1), 57-74.

Guren, A. (2014). The causes and consequences of house price momentum. Cambridge, MA.

Harris, R. D. F., Tzavalis, E., 1999. Inference for unit roots in dynamic panels where the time dimension is fixed. Journal of Econometrics 91, 201–226.

Hassan, A. R. (2017). The interplay between the Bayesian and frequentist approaches: a general nesting spatial panel data model. Spatial Economic Analysis, 12(1), 92-112.

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).

Hoogland, J. J., & Boomsma, A. (1998). Robustness studies in covariance structure modeling An overview and a meta-analysis. Sociological Methods & Research, 26(3), 329-367.

Hospers, G., & Beugelsdijk, S. (2002). Regional Cluster Policies: Learning by Comparing? *Kyklos*, *55*(3), 381–402.

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.

Huang, B., Wu, B., & Barry, M. (2010). Geographically and temporally weighted regression for modeling spatio-temporal variation in house prices. International Journal of Geographical Information Science, 24(3), 383-401.

Jones, M. (2007) University of Cambridge Housing Requirements. Cambridge Centre for Housing and Planning Research, Available at: https://www.cambridge.gov.uk/public/ldf/coredocs/University%20of%20Cambridge%20Housing%20Needs%20Study%202008.pdf

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.

Knoll, K., Schularick, M., & Steger, T. M. (2014). No price like home: Global house prices, 1870-2012.

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

Massey, D.S., & Kanaiaupuni, S.M. (1993). Public housing and the concentration of poverty. Social Science Quarterly, 74, 109-122.

McQuinn K. and O'Reilly G., "Assessing the role of income and interest rates in determining house prices", *Economic Modeling*, vol. 25, pp. 377-390, 2008.

Miles, D. (2012). Population Density, House Prices and Mortgage Design. *Scottish Journal of Political Economy*, *59*(5), 444–466.

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

Nappi-Choulet, I., Maury, T.-P., 2011. A spatial and temporal autoregressive local estimation for the Paris housing market, Journal of Regional Science 51(4), 732–750

Nerlove, M. (1958). Adaptive expectations and cobweb phenomena. The Quarterly Journal of Economics, 227-240.

Pace, R. K., & LeSage, J. P. (2009). A sampling approach to estimate the log determinant used in spatial likelihood problems. Journal of Geographical Systems, 11(3), 209-225.

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

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

Sargan, J. D. (1983). Identification and lack of identification. Econometrica: Journal of the Econometric Society, 1605-1633.

Sirmans, S., Macpherson, D., & Zietz, E. (2005). The composition of hedonic pricing models. Journal of real estate literature, 13(1), 1-44.

Smith, T.E., 2009. Estimation Bias in Spatial Models with Strongly Connected Weight Matrices, Geographical Analysis 41, 307–332

Su, L., & Yang, Z. (2015). QML estimation of dynamic panel data models with spatial errors. Journal of Econometrics, 185(1), 230–258.

Thanos, S., Dubé, J., Legros, D., 2016. Putting Time into Space: The Temporal Coherence of Spatial Applications in the Housing market. Regional Science and Urban Economics, 58, 78–88

UK Statistics Authority (2011), Small Area Income Estimates for England and Wales., Available at: https://www.statisticsauthority.gov.uk/archive/assessment/assessment/assessment-reports/assessment-report-160---small-area-income-estimates-for-england-and-wales.pdf

Yang, Z. (2015). Unified M-Estimation of Fixed-Effects Spatial Dynamic Models with Short Panels. Research Collection School Of Economics.

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.

Zabel, J. E. (1999). Controlling for Quality in House Price Indices. Journal of Real Estate Finance and Economics, 19; Jg. 19(3), 223–41.

10. Appendix

Table A1. Data sources and management procedures

Variable	**Index value**	**Housing stock**
Source	Land Registry	2001 and 2011 Census data and Land Registry
Geographical Granularity	Street address	MSOA
Processing	Index of average prices at MSOA level using a regression controlling for the type of building, new structures and leasehold/freehold transactions	Computed for 2008 by adjusting the 2011 stock levels for new buildings sold between 2008 and 2011 and for 2004 levels by adjusting 2001 stock levels for properties sold until 2004
Variable	**Index value in Cambridge**	**New supply**
Source	Land Registry	Land Registry
Geographical Granularity	Street address	Street address
Processing	The average value of the index calculated for all areas within the city of Cambridge	Based on the number of new properties sold over the last 4 years
Variable	**Income**	**Unemployment**
Source	Small Area Model-Based Income Estimates	UK Neighbourhood Statistics
Geographical Granularity	MSOA	District
Processing	Unprocessed	Average district value assigned to each MSOA located within that district.
Variable	**Research centre**	
Source	Cambridgeshire County Council	
Geographical Granularity	Geographical coordinates	
Processing	Based on the size and distance of the nearest employment centre	