MULTILEVEL AND SPILLOVER EFFECTS ESTIMATED FOR SPATIAL PANEL DATA, WITH APPLICATION TO ENGLISH HOUSE PRICES

Badi H. BALTAGI[a] **, Bernard FINGLETON**[b]* **and Alain PIROTTE**[cd]

Abstract - *This paper estimates a nested random effects spatial autoregressive panel data model to explain annual house price variation for 2000-2007 across 353 local authority districts in England. The paper applies a newly-proposed estimator based on the instrumental variable approach for the cross-sectional spatial autoregressive model.*

Keywords - HOUSE PRICES, PANEL DATA, SPATIAL LAG, NESTED RANDOM EFFECTS, INSTRUMENTAL VARIABLES, SPATIAL DEPENDENCE

JEL classification - C33

[a] Syracuse University, Department of Economics and Center for Policy Research.
[b] University of Cambridge, Department of Land Economy.
[c] CRED-TEPP (CNRS), University of Panthéon-Assas Paris II.
[d] IFSTTAR-DEST, French Institute of Science and Technology for Transport, Development and Networks.
* Corresponding author. Bernard Fingleton. E-mail: bf100@cam.ac.uk

1. INTRODUCTION

In this paper, we focus on the estimation of a theoretical house price model in which spatio-temporal variations in house prices are driven by supply and demand conditions, with spatial effects coming from two distinct sources. One is the direct dependence of house prices in a given locality on house prices in nearby localities. The other is *via* hierarchical error components typical of multilevel models. Direct dependence is the net effect of what we refer to as displaced demand and displaced supply. Displaced demand occurs where, ceteris paribus, high prices nearby cause demand to increase, because the negative relationship between prices and demand leads to purchasers switching away from high price locations nearby. Displaced supply occurs where high prices nearby cause supply to fall as a result of suppliers of housing switching to where higher prices give better investment returns. The supply and demand equations lead to a reduced form in which prices depend directly on prices nearby. We refer to this as a spillover effect.

The second source of spatial heterogeneity comes from the presence of hierarchical error components which represent the impact of local (district) effects embedded within wider (county) effects. Intuitively, local effects can be thought of as postcode effects, where particular postcodes are associated with more or less prestige. Likewise we envisage an independent county effect (a number of districts together are nested within a county). County A, which is a prestigious address, will tend to be associated with higher prices than less prestigious county B. The difference between these two sources of spatial heterogeneity in house prices is that the hierarchical district and county effects are constant within counties and districts, and terminate abruptly at county or district boundaries. In contrast, the spillover effects have an autoregressive specification, so that they extend across space with diminishing effect as distance increases. Thus spatial heterogeneity and autocorrelation is partly accounted for by the discrete non-overlapping effects of the components of the hierarchical level, and partly by spatial dependence operating, *via* the spatial autoregressive process, simultaneously across all areas.

Our solution to the problem of estimating the panel data model with spatial spillover effects and random hierachical effects, which we refer to as a nested random effects spatial autoregressive panel data model, is to propose a novel methodological approach. While the focus of the paper is squarely on the theory and application in relation to house prices, this has been introduced elsewhere in the literature. In contrast, our proposed estimator as set out intially in Baltagi et al (2014) is new. Recent developments in econometrics allow us to take into account cross-sectional correlation in a panel data context, as demonstrated for example by Elhorst (2003, 2010*a*), Anselin, Le Gallo and Jayet (2008), Anselin (2010) and Baltagi (2011). Early work on hierarchical panels was carried out by Fuller and Battese (1973), Montmarquette and Mahseredjian (1989), Antweiler (2001) and Baltagi, Song and Jung (2001, 2002), among others. The contribution here combines both hierarchical and spatial panel approaches, building on the sparse earlier literature which suggests combining nested models and spatial autoregressive processes in a cross-sectional context (see Corrado and Fingleton, 2011, 2012).

From the purely spatial perspective, a common way to proceed is to apply Maximum Likelihood methods. However, Maximum Likelihood procedures are often challenging when the sample size is large. Moreover, they call for explicit distributional assumptions. In fact, this is why Kelejian and Prucha (1998) and Lee (2003) proposed an Instrumental Variable (IV) procedure for the cross-sectional spatial autoregressive model, which is computationally simple and less restrictive regarding the distribution of the disturbances. In this paper, we extend this cross-sectional IV approach to the nested random effects spatial autoregressive panel data model[5].

The plan of the paper is as follows: Section 2 describes the spatial autoregressive model with nested random effects, Section 3 briefly summarises an IV procedure to estimate this model (details of which are given in Baltagi et al, 2014) and Section 4 describes the house price data. Section 5 discusses the empirical results and Section 6 gives our concluding remarks.

2. THE SPATIAL LAG MODEL INCLUDING NESTED RANDOM EFFECTS

Based on a theoretical framework that assumes the existence of equilibrium housing supply and demand functions, details of which are given in Baltagi et al(2014), we arrive at an empirical specification that allows for a exogenous effects on house prices, spatial spillover of price levels across districts, and nested random effects attributable to districts nested within counties, thus

$$y_{ijt} = \rho \sum_{g=1}^{N} \sum_{h=1}^{M_g} w_{ij,gh} y_{ght} + \mathbf{x}_{ijt}\beta + \varepsilon_{ijt}, \qquad (1)$$

where $i = 1, \dots, N$, $j = 1, \dots, M_i$ and $t = 1, \dots, T$. The dependent variable y_{ijt} denotes the average house price of district j in county i at time period t. $\mathbf{x}_{ijt}$ is a $(1 \times K)$ vector of explanatory (exogenous) variables, namely income within commuting distance (Y_{ijt}^c) and the available stock of dwellings (Λ_{ijt}). ρ is a scalar and β is a $(K \times 1)$ vector of parameters to be estimated. The weight $w_{ij,gh} = w_{k,l}$ is the $(k = ij, l = gh)$ element of the matrix $\mathbf{W}_S$ with ij denoting district j within county $i, i = 1, \dots, N, j = 1, \dots, M_i$ and similarely for gh.Thus $k, l = 1, \dots, S$ where $S = \sum_{i=1}^{N} M_i$ and $\mathbf{W}_S$ is an $(S$ x $S)$ matrix of known spatial weights denoting inter-district connectivity, which has zeros on the main diagonal and is row-normalised so that for row $k, \sum_{g=1}^{N} \sum_{h=1}^{M_g} w_{k,gh}$ =1. We assume that ρ is bounded numerically to ensure spatial stationarity, i.e., $e_{\min}^{-1} < \rho < 1$ where $e_{\min}$ is the minimum real characteristic root of $\mathbf{W}_S$. The nested random effects are introduced via the remainder term ε_{ijt} which follows an error component structure :

[5] In a similar spirit, Baltagi and Liu (2011) derived an IV estimator in the context of spatial autoregressive random effects panel data model. However, this estimator does not deal with the unbalancedness or the nested structure of the panel data.

$$\varepsilon_{ijt} = \alpha_i + \mu_{ij} + v_{ijt}, \tag{2}$$

where α_i denotes an unobservable *county* specific time-invariant effect which is assumed to be i.i.d.$(0, \sigma_\alpha^2)$. μ_{ij} is the nested effect of *district j* within the ith county which is assumed to be i.i.d.$\left(0, \sigma_\mu^2\right)$, and v_{ijt} is a remainder disturbance which is also assumed to be i.i.d. $(0, \sigma_v^2)$. The α_i 's, μ_{ij} 's and v_{ijt} 's are independent of each other and among themselves. In contrast to the classical literature on panel data, grouping the data by periods rather than units is more convenient when we consider the spatial dependence. In matrix notation, for a cross-section, the model (1) corresponds to

$$\mathbf{y}_t = \rho \mathbf{W}_S \mathbf{y}_t + \mathbf{X}_t \beta + \varepsilon_t, \tag{3}$$

where $\mathbf{y}_t$ is of dimension $(S \times 1)$, $\mathbf{X}_t$ is an $(S \times K)$ matrix of explanatory variables.

The assumption that the nested error components are independent of each other and among themselves is a standard assumption in the literature. While it would be interesting to attempt to take account of possible interdependencies in a nested random effects context, this is something very much on the research horizon and beyond the scope of the present paper. Additionally, as is typical in random effects specifications, we are assuming the independence of the compound errors and the explanatory variables leading to consistent estimation, but as pointed out by Baltagi (2008), there is no entirely satisfactory test of this assumption. Typically, following Hausman (1978), in a general panel setting one would compare estimates from a fixed effects specification with the estimates from a random effects specification. Unfortunately, although this test is very popular in the literature, it is conditional on an assumption that the fixed effects estimates are consistent, something which cannot be guaranteed.

Against this there are some advantages of a random effects specification. For example, with 353 districts, there would be a considerable loss of degrees of freedom by invoking fixed effects. In addition, random effects allows one to obtain estimates taking account of permanent cross-section or between variation. In comparison fixed effects focuses on short term variation (Partridge (2005), Baltagi (2008), Elhorst (2010*b*)). So while correlation of the random effects with exogenous regressors may induce inconsistency, there will be compensation in the form of enhanced precision of the estimates. Higgins, Levy and Young (2006) use county-level data to analyze growth and convergence across the US, and with 3000 counties and only 3 time periods, the within variation is small and uninteresting compared with the between. Likewise Barro (1997) criticizes fixed effects panel data methods that rely purely on time series information, arguing that the conditioning variables often vary slowly over time, so that the most important information is lost.

3. ESTIMATION

The key to understanding our estimation procedure is the covariance matrix of ε, which is denoted by $\mathbf{\Omega}$. This captures the nested random effects attributable to the influence of counties, districts nested within counties, and the remainder. The paper by Baltagi et al (2014) gives the technical detail of how this is accomplished.

Stacking the T periods, the model (3) becomes

$$\mathbf{y} = \rho \mathbf{W}\mathbf{y} + \mathbf{X}\beta + \varepsilon \tag{4}$$

$$= \mathbf{B}^{-1}(\mathbf{X}\beta + \varepsilon), \tag{5}$$

where $\mathbf{B} = \mathbf{I}_T \otimes \mathbf{B}_S$, $\mathbf{B}_S = \left(diag\left(\mathbf{I}_{M_i}\right) - \rho \mathbf{W}_S\right)$ and $\mathbf{W} = (\mathbf{I}_T \otimes \mathbf{W}_S)$.

It is well known that the spatially lagged dependent variable $\mathbf{Wy}$ is correlated with the disturbances ε and therefore, the Ordinary Least Squares estimator will be inconsistent. Let $\mathbf{Z} = (\mathbf{X}, \mathbf{Wy})$ and $\delta = (\beta, \rho)^T$, the model (5) can be written as

$$\mathbf{y} = \mathbf{Z}\delta + \varepsilon. \tag{6}$$

In the cross-section, Kelejian and Prucha (1998) suggested a Two-Stage Least Squares spatial estimator (S2SLS) for the spatial lag model model, advising that the instrument set should be kept to a low order in order to avoid linear dependence and retain full column rank for the matrix of instruments. We therefore use the recommended instrument set $[\mathbf{X}, \mathbf{WX}, \mathbf{W}^2\mathbf{X}]$.

Premultiplying by $\mathbf{\Omega}^{-1/2}$ gives

$$\mathbf{y}^* = \mathbf{Z}^*\delta + \varepsilon^*, \tag{7}$$

where $\mathbf{y}^* = \mathbf{\Omega}^{-1/2}\mathbf{y}, \mathbf{Z}^* = \mathbf{\Omega}^{-1/2}\mathbf{Z} = \mathbf{\Omega}^{-1/2}(\mathbf{X}, \mathbf{Wy})$ and $\varepsilon^* = \mathbf{\Omega}^{-1/2}\varepsilon$. Applying the Kelejian and Prucha (1998) 2SLS procedure to (7), we obtain our nested random effects spatial Two-Stage Least Squares estimator (NRE-S2SLS) of δ, given by

$$\hat{\delta}_{NRE-S2SLS} = \left(\mathbf{Z}^{*^T}\mathbf{P}_{\mathbf{H}^{**}}\mathbf{Z}^*\right)^{-1}\mathbf{Z}^{*^T}\mathbf{P}_{\mathbf{H}^{**}}\mathbf{y}^*, \tag{8}$$

where $\mathbf{H}^{**} = \mathbf{\Omega}^{-1/2}\mathbf{H}$ and $\mathbf{P}_{\mathbf{H}^{**}} = \mathbf{H}^{**}(\mathbf{H}^{**T}\mathbf{H}^{**})^{-1}\mathbf{H}^{**T}$.

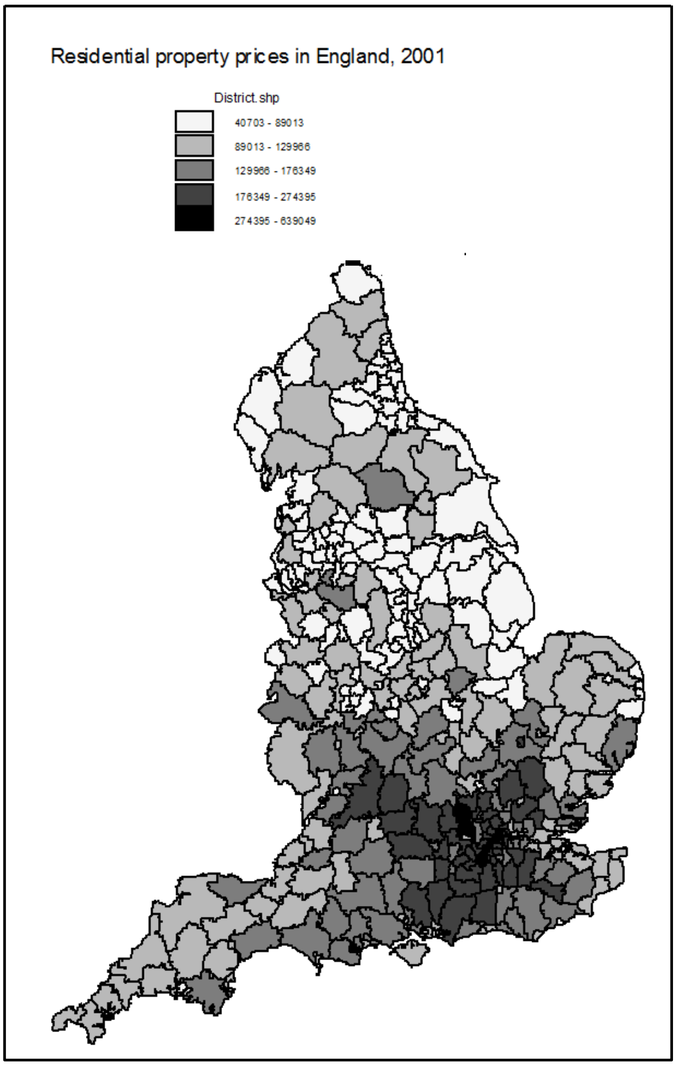
Although we have derived a NRE-S2SLS estimator, in Baltagi et al (2014) we also go one stage further and obtain a (very closely related) nested best spatial 2SLS estimator (NB-S2SLS) estimator, following Lee (2003). This is omitted here to save space.

4. DATA DESCRIPTION

The house price data give p_{ijt}, $i = 1, \dots, N; j = 1, \dots, M_i; t = 1, \dots, T$, which is the average selling prices by local authority district (of which there are 353 covering England). Our modeling is based on $T = 8$ years of data observed over the period 2000 to 2007. While these data may appear fixed (at least until boundary changes occur), we do not regard them as a population, but rather as a realization taken from an infinite number of posisble realizations or super-population. We can think of the remainder disturbance in our model capturing unpredictable unmodelled variance across space and time as a driver of different realizations. We are interested in the underlying process that could have generated different realizations and which has generated the data we observe, and use the 'sample' data to make inferences about this process.

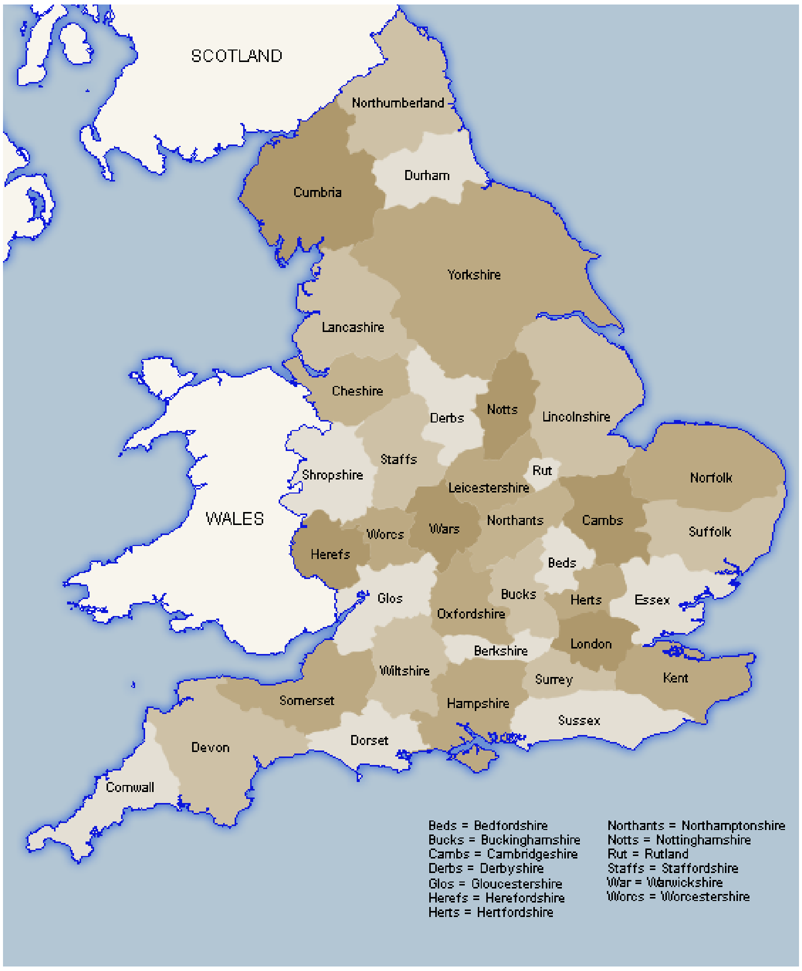
Figure 1 maps the data for a single snapshot of time, the year 2001, clearly showing a heterogeneous and spatially correlated mean house price distribution across districts (the other 7 years show similar patterns).

Figure 1. Mean residential property price (all types) in English districts 2001



The data are organised by districts nested within counties**.** Figure 1 gives the distribution of districts, while Figure 2 gives the counties within which districts are nested. The district and county definitions are not immutable, having changed over time typically for administrative reasons. Likewise we choose to change the definition of counties, thus increasing county-level variance, by merging all counties lying outside the South East of England. The rest of England is therefore treated as a single super-county, while the South Eastern counties remain as separate sources of variance. The 16 counties which are 'not merged', with the brackets containing the number of nested districts, are Bedfordshire (4), Berkshire (6), Buckinghamshire (5), East Sussex (6), Essex (14), Hampshire and the Isle of Wight (14), Hertfordshire (10), Kent (13), Oxfordshire (5), Surrey (11), West Sussex (7), Cambridgeshire (6), Norfolk (7), Suffolk (7), Inner London (14), Outer London (19). The remaining 30 counties are merged as one 'super county' nesting the remaining 205 districts.

Figure 2. English counties



Income by district for each year (Y_{ijt}) is based on data from the UK's Office of National Statistics.

The Office of National Statistics also provides a matrix of inter- and intra-district commuting frequencies (d^*) from the UK 2001 census.

Using these commuting frequencies, we calculate income within commuting distance Y_{ijt}^c, applying less weight to districts with relatively few commuters and most weight is applied to local income with large within-district commuting flows. Thus each district's income is weighted by its row normalized commuting frequency.

The spatial lag matrix $\mathbf{W}_S$ is also based on commuting (subsequently row-normalized to give $\mathbf{W}_S$), but rather than d^* we commence with δ,which is d^* with the main diagonal cells set to zero. Thus the vector $\mathbf{W}_S\mathbf{y}_t$ comprises weighted averages of 'nearby' districts. The rationale here is that inter-district commuting is an appropriate indicator of inter-district proximity or connectivity. The weights matrix $\mathbf{W}_S$ a function of the larger values (> 50) in the cells of δ.

On the supply side, Λ_{ijt} is equal to the annual dwelling stock estimates by district divided by the annual population estimates for each district, data which are also available from the UK's Office of National Statistics. The number of dwellings (per 100 persons) respresents the available supply.

5. MODEL ESTIMATES

Table 1 shows the estimates obtained for our house price model using various model specifications and estimators. Our preferred estimates are the NRE-S2SLS. These estimators provide theoretically acceptable and appropriately signed parameter estimates[6]. OLS estimation, which omits both the spatial lag and the nested error structure, clearly introduces bias. Comparing the OLS estimates with other estimation outcomes which include the endogenous spatial lag $\mathbf{Wy}$ highlights its omission as an important source of bias. Although it controls for the spatial lag, S2SLS estimation produces a positive coefficient on Λ_{ijt} which is contrary to the negative estimate antici-pated from theory. We attribute this to a lack of control for nested district and county effects.

The nested random effects (ANOVA) estimates do include district and county effects and the anticipated negative sign on the Λ_{ijt} coefficient, but here the model set up excludes the endogenous spatial lag. Compared with NRE-S2SLS, the ANOVA estimates of β_1, β_2, σ_α^2, σ_μ^2, and σ_v^2 are very different, indicating omitted variable bias. This illustrates that it is insufficient to pick up spatial heterogeneity and correlation simply via nested error components and to ignore spillover effects. However, including both sources of spatial correlation, that due to the spatial lag and that due to the nested error components, elimi-

[6] Appendices 2 and 3 in Baltagi et al (2014) gives the small sample performance of these estimators using Monte Carlo simulations.

nates this source of bias and gives us the preferred NRE-S2SLS, which gives outcomes that one would anticipate from theory, with a highly significant estimate of ρ pointing to the importance of the simultaneous spatial interaction of price levels. Also in line with theory is the significant positive coefficient for income within commuting distance Y_{ijt}^{c}, and significant negative coefficient for the number of dwellings (per 100 persons) Λ_{ijt}.

Table 1. Various parameter estimates for the house price model

	OLS	S2SLS	SML	RE-S2SLS	RE-SML	NRE	NRE-S2SLS	NRE-SML	NB-S2SLS
Income within Comm. Distance (Y_{ijt}^{c})	**1.1318**	**0.3631**	**0.2536**	**0.5057**	**0.4641**	**6.4756**	**0.4869**	**0.4649**	**0.4975**
s.e.	(0.0355)	(0.0368)	(0.0363)	(0.0732)	(0.0810)	(0.0917)	(0.0724)	(0.0894)	(0.0641)
t-ratio	(31.882)	(9.845)	(6.975)	(6.907)	(5.727)	(70.617)	(6.722)	(5.201)	(7.760)
Supply (Λ_{ijt})	**1391.90**	**2125.303**	**2100.05**	**-730.4573**	**-724.167**	**-2747.29**	**-741.3797**	**-733.5315**	**-744.557**
s.e.	(432.4)	(299.405)	(370.92)	(105.442)	(568.42)	(230.86)	(105.502)	(528.706)	(105.534)
t-ratio	(3.220)	(7.098)	(5.66)	(-6.927)	(-1.274)	(-11.90)	(-7.027)	(-1.387)	(-7.055)
Spatial Lag of House Prices ($p_{ijt}^{W_S}$)	-	**0.5781**	**0.6107**	**0.8489**	**0.7961**	-	**0.8732**	**0.8078**	**0.8713**
s.e.	-	(0.0207)	(0.0150)	(0.0124)	(0.0171)	-	(0.0112)	(0.0301)	(0.0111)
t-ratio	-	(27.8744)	(40.508)	(68.2956)	(46.531)	-	(77.7798)	(26.8258)	(78.7695)
$\hat{\sigma}_{\alpha}^{2}$	-	-	-	-	-	0.2533	0.0562	0.1066	(*)
$\hat{\sigma}_{\mu}^{2}$	-	-	-	0.2114	0.2482	0.3337	0.1738	0.0973	(*)
$\hat{\sigma}_{v}^{2}$	0.4927	0.2343	0.3615	0.0132	0.4892	0.0647	0.0132	0.4707	(*)

(*) This estimator uses the same consistent estimates of σ_{α}^{2}, σ_{μ}^{2} and σ_{v}^{2} as those used for NRE-S2SLS.

With regard to the county effects operating via the error components, we anticipate that there will be discernable county-level impacts in the South and East of England especially, resulting from the inter-county status differences that are perceived by some, even if they are stereotypical rather than being objective reality. Some supposed reputational differences between Essex and the Royal county of Berkshire are caricatured by the entries in Table 2.

District effects can also be envisaged, and are at their most potent in the so-called 'London postcode wars'. For example, people living in parts of the leafy and affluent district of Richmond-Upon-Thames are angry that their postcode is the one used in the more deprived borough of Hounslow. According to the property valuation website Zoopla, in 2013 the average house in Hounslow was worth £ 294,020, while the average house in Richmond was £ 423,982. to According to one newspaper report, a resident said " I don't know anyone who would turn down the offer of a Richmond postcode instead of a Hounslow one", given that it would suddenly increase the value of their home by such a large amount.

Table 2. County comparisons (stereotypes !)

Essex	Berkshire
• one of the country's most distinct reputations - and probably the worst • Simon Heffer, wrote about "Essex Man". Described as a bit "brutish", "culturally barren" and "breathtakingly right-wing", Essex Man quickly became shorthand for the successes - and excesses - of Thatcher's Britain. • big hair, big cars and big wads of cash • tanning salons, nail bars and nightclubs • Numerous private car number plates, such as MON EY1 or UR2 FAT or OOOO7 • TV show The Only Way Is Essex. Like Footballers Wives, this TV series stereotypes those from Essex	• known as the Royal County of Berkshire • Windsor castle • Eton (England's poshest school) • Thames Valley or M4 corridor full of IT firms such as Cable and Wireless, Hewlett-Packard, Microsoft, Sharp Telecommunications, Oracle Corporation, Sun Microsystem, Vodaphone, Cisco, Xerox, accounting firms Deloitte, KPMG, Ernst and Young, and PricewaterhouseCoopers • Ascot Racecourse : thoroughbred horse racing. The course is closely associated with the British Royal Family, being approximately six miles from Windsor Castle • Suppliers of silk top hats are saying demand is out-stripping supply as the Queen gently suggests that she likes to see men wearing them in the Royal Enclosure

The estimates of σ_α^2, σ_μ^2 and σ_v^2 which partition the error into components representing county, district and remainder variance, highlight the relative importance of the district component. With regard to the county-level variance $\hat{\sigma}_\alpha^2$, which is the outcome of dividing counties into two groups, there is discernable contrast between the South and East of England and elsewhere, with inter-county variation within the South East reflecting the (relatively minor) role of county-level factors operating to distinguish, for example, Essex from Berkshire. In the rest of England, we see much less inter-county variation. If we undo the bloc construction so as to introduce inter-county variation also within the rest of England, the NRE-S2SLS estimator gives $\hat{\rho} = 0.855105$ ($t = 70.63$), $\hat{\beta}_1 = 0.545864$ ($t = 7.20$) and $\hat{\beta}_2 = -752.1513$ ($t = -7.11$), and $\hat{\sigma}_\alpha^2$ is equal to 0.013932 ($\hat{\sigma}_\mu^2 = 0.17992$, $\hat{\sigma}_v^2 = 0.013185$) as a consequence of the influence on the estimates of the comparatively smaller inter-county variation in the rest of England.

6. CONCLUSION

The paper analyzes house price data observed in 353 English districts over 8 years. We show that, in line with theoretical expectation, income within commuting distance has a positive effect on prices, and the stock of housing has a negative effect, and that there is a significant spatial lag term indicating postive correlation between prices locally and prices in 'nearby' districts. Also we model additional price heterogeneity using nested error components attributable to district and county effects, showing that these and the spatial lag are both necessary elements of our house price model. Our Monte Carlo analysis detailed in Baltagi et al. (2014) indicates that for realizations of an artificial data generating process with both spatial lag effects and hierarchical error components, our estimators (NRE-S2SLS and NB-S2SLS) are superior to a number of alternatives.

To sum up, the spatial econometrics literature is almost totally devoid of hierarchical models, as pointed out by Corrado and Fingleton (2012). Moreover, the spatial panel literature makes no reference to spatial interaction effects in a nested context. In the presence of both sources of spatial dependence, omission of one or both from the estimator can lead to incorrect inference and an improper understanding of true causal mechanisms.

REFERENCES

Anselin, L., 2010, Thirty years of spatial econometrics, *Papers in Regional Science* 89, 3-25.

Anselin, L., J. Le Gallo and H. Jayet, 2008, *Spatial Panel Econometrics*, Chapter 19 in L. Mátyás and P. Sevestre, eds., The Econometrics of Panel Data: Fundamentals and Recent Developments in Theory and Practice, Springer-Verlag, Berlin, 625-660.

Antweiler, W., 2001, Nested random effects estimation in unbalanced panel data, *Journal of Econometrics* 101, 295-313.

Baltagi, B.H., 2008, *Econometric Analysis of Panel Data* (Wiley, Chichester).

Baltagi, B.H., 2011, *Spatial Panels*, Chapter 15 in *The Handbook of Empirical Economics and Finance*, Aman Ullah and David E.A. Giles, editors, Chapman and Hall, 435-454.

Baltagi, B.H. and L. Liu, 2011, Instrumental variable estimation of a spatial autoregressive panel model with random effects, *Economics Letters* 111, 135-137.

Baltagi, B.H., S.H. Song and B.C. Jung, 2001, The unbalanced nested error component regression model, *Journal of Econometrics* 101, 357-381.

Baltagi, B.H., S.H. Song and B.C. Jung, 2002, Simple LM tests for the unbalanced nested error component regression model, *Econometric Reviews* 21, 167–187.

Baltagi, B.H., B. Fingleton and A. Pirotte, 2014, Spatial Lag Models with Nested Random Effects: an IV procedure with application to English house prices, *Journal of Urban Economics* 80, 76-86.

Barro, R.J., 1997, *Determinants of Economic Growth: A Cross-Country Empirical Study* (Cambridge, MA: MIT Press).

Corrado, L. and B. Fingleton, 2011, Multilevel modelling with spatial effects, *Discussion paper n°11-05*, University of Strathclyde.

Corrado, L. and B. Fingleton, 2012, Where is the economics in spatial econometrics?, *Journal of Regional Science* 52, 210-239.

Elhorst, J.P., 2003, Specification and estimation of spatial panel data models, *International Regional Science Review* 26, 244-268.

Elhorst, J.P., 2010*a*, *Spatial Panel Data Models*, Chapter 2 in M.M. Fischer and A. Getis, eds., Handbook of Applied Spatial Analysis, Berlin Springer-Verlag, 377-405.

Elhorst, J.P., 2010*b*, Applied spatial econometrics: raising the bar, *Spatial Economic Analysis* 5, 9-28.

Felsenstein, D., M. Beenstock and D. Feldman, 2012, Testing for unit roots and cointegration on spatial cross section data, *Spatial Economic Analysis* 7, 203-222.

Fuller, W.A. and G.E. Battese, 1973, Transformations for estimation of linear models with nested error structures, *Journal of the American Statistical Association* 68, 626-632.

Hausman, J., 1978, Specification tests in econometrics, *Econometrica* 46, 1251-1271.

Higgins, M.J., D. Levy and A.T. Young, 2006, Growth and convergence across the United States: Evidence from county-level data, *Review of Economics and Statistics* 88, 671–681.

Kelejian, H.H. and I.R. Prucha, 1998, A generalized spatial two-stage least squares procedure for estimating a spatial autoregressive model with autoregressive disturbances, *Journal of Real Estate Finance and Economics* 17, 99-121.

Lee, L.-F., 2003, Best spatial two-stage least square estimators for a spatial autoregressive model with autoregressive disturbances, *Econometric Reviews* 22, 307-335.

Montmarquette, C. and S. Mahseredjian, 1989, Does school matter for educational achievement? A two-way nested-error components analysis, *Journal of Applied Econometrics* 4, 181-193.

Partridge, M.D., 2005, Does income distribution affect U.S. state economic growth?, *Journal of Regional Science* 45, 363-394.

UNE ESTIMATION DES EFFETS DE DÉBORDEMENT MULTI-NIVEAUX POUR DES MODÈLES SPATIAUX AVEC DES DONNÉES DE PANEL - UNE APPLICATION AUX PRIX IMMOBILIERS EN ANGLETERRE

Résumé *- Le travail présenté dans cet article s'appuie sur un modèle autorégressif spatial qui utilise des données de panel aléatoirement relevées, afin d'étudier la variation annuelle des prix immobiliers durant la période 2000-2007 dans 353 districts locaux en Angleterre. L'originalité de l'article repose sur l'utilisation d'un nouvel estimateur basé sur l'approche des variables instrumentales dans les modèles autorégressifs spatiaux.*

Mots-clés - PRIX IMMOBILIERS, DONNÉES DE PANEL, DÉCALAGE SPATIAL, VARIABLES INSTRUMENTALES, DÉPENDANCE SPATIALE