



**AgEcon** SEARCH  
RESEARCH IN AGRICULTURAL & APPLIED ECONOMICS

*The World's Largest Open Access Agricultural & Applied Economics Digital Library*

**This document is discoverable and free to researchers across the globe due to the work of AgEcon Search.**

**Help ensure our sustainability.**

Give to AgEcon Search

AgEcon Search  
<http://ageconsearch.umn.edu>  
[aesearch@umn.edu](mailto:aesearch@umn.edu)

*Papers downloaded from **AgEcon Search** may be used for non-commercial purposes and personal study only. No other use, including posting to another Internet site, is permitted without permission from the copyright owner (not AgEcon Search), or as allowed under the provisions of Fair Use, U.S. Copyright Act, Title 17 U.S.C.*

# **A Spatial Analysis of the Role of Residential Real Estate Investment in the Economic Development of the Northeast Region of the United States**

## **Praveena Jayaraman**

PhD Candidate

Davis College of Agriculture, Natural Resources and Design

Morgantown, WV-26506

e-mail: [praveena.jayaraman@mail.wvu.edu](mailto:praveena.jayaraman@mail.wvu.edu)

## **Dr. Donald Lacombe**

Research Associate Professor

Regional Research Institute

Morgantown, WV-26506

e-mail: [donald.lacombe@mail.wvu.edu](mailto:donald.lacombe@mail.wvu.edu)

## **Dr. Tesfa Gebremedhin**

Professor

Davis College of Agriculture, Natural Resources and Design

Morgantown, WV-26506

e-mail: [tesfa.gebremedhin@mail.wvu.edu](mailto:tesfa.gebremedhin@mail.wvu.edu)

*Selected Paper prepared for presentation at the Agricultural and Applied Economics Association's 2013 AAEA & CAES Joint Annual Meeting, Washington, DC, August 4-6, 2013.*

Copyright 2013 by Praveena Jayaraman, Donald Lacombe and Tesfa Gebremedhin. All rights reserved. Readers may make verbatim copies of this document for non-commercial purposes by any means, provided that this copyright notice appears on all such copies.

# **A Spatial Analysis of the Role of Residential Real Estate Investment in the Economic Development of the Northeast Region of the United States**

**Praveena Jayaraman, Donald Lacombe, and Tesfa Gebremedhin<sup>1</sup>**

## **Abstract**

Spatial dependence is an important factor in regional economic growth analysis, especially in terms of population density, employment, and median income. This paper employs spatial econometric techniques and U.S. Census Bureau county-level data for the period of 1980-2010 to identify and estimate the impacts of residential real estate investment on the economic development of the Northeast region. A spatial panel method is used to analyze the spillover effect of county level economic development on neighboring counties.

**Key Words:** Residential Real Estate Investments, Spatial Econometrics, Spatial Panel

---

<sup>1</sup> Graduate Research Assistant. Research Associate Professor and Professor, Division of Resource Management, West Virginia University Selected Paper prepared for presentation at the Agricultural and Applied Economics Association's (AAEA) 2013 Annual Meeting, Washington DC, August 4-6, 2013

# **A Spatial Analysis of the Role of Residential Real Estate Investment in the Economic Development of the Northeast Region of the United States**

## **Introduction**

Residential real estate investment has been recognized as an agent of economic growth since the 1970s, because residential real estate investment was predicted as a major economic activity with large multiplier effects. Residential real estate improvement is also linked to many external social and economic benefits. Many studies have examined the role of residential real estate in economic development through various approaches, such as the effects of employment and income (Leung, 2004), household saving (Turner and Luea, 2009), labor productivity (Ofori and Han, 2003), health productivity and growth from real estate investment (Arku and Harris, 2005), as well as home ownership effects (Carruthers and Mulligan, 2005; Carruthers and Mulligan, 2008).

In the United States, real estate is an important investment for individual investors. In the Census Bureau's Survey of Income and Program Participation (SIPP), residential real estate was the largest class of assets held by individuals amounting to 78.7 percent of total household asset value (Census, 2000). Of this, homes represented 67.2 percent, rental properties 4.9 percent, and other real estate such as vacation homes and land holdings 6.6 percent of total investment portfolios. By contrast, the value of commercial real estate has decreased by 40 percent since 2007 in the United States. According to Deloitte LLP (2009), loss of jobs and reductions in consumer spending negatively affected all types of real estate investment in general and office and retail properties in particular. Rental rates and real estate prices decreased due to high vacancy rates of properties. However, in 2010 a potential recovery in economic growth of the country was leading to increases in property values again.

Some important issues associated with real estate; population, income, cost, quality, and affordability of real estate all influence residential real estate prices. According to the U.S. Census Bureau (2011), the population of the Northeast region is approximately 73 million, which is equal to 23.4 percent of the U.S. population. Increasing population in urban areas is a burden on the residential real estate market. The population in urban areas in the region increased by 18 percent from 1980 to 2000. At the same time population in rural areas decreased by 18 percent. One possible reason for an increase in urban population and a decrease in rural population at the same time and by the same rate could be due to the migration of rural population to urban areas for employment. Neighborhood quality of life also has significant consequences on the health and well-being of children, and often plays a role when people move from one region to another.

After the national recession in the early 1980s, the Northeast region recovered rapidly according to the FDIC (2010). In this region, commercial and residential real estate markets grew quickly due to strong regional employment and economic growth between 1982 and 1988. However in the late 1980s economic growth in the region declined due to a decrease in employment and slow personal income growth, regional economic growth declined and overbuilt real estate markets intensified the effects. Residential real estate costs and quality of life issues continued to be a serious problem for low-income populations, especially in rural areas. More than 42 percent of unassisted low-income renters had severe residential problems in the region during the same years. In spite of the problem of acquiring affordable quality housing and available credit in rural areas, ownership of real estate has been one of the best methods of asset accumulation for low-income rural households.

Carlino and Mills, (1987) using simultaneous equations model estimated employment and population changes in U.S. counties and explained the migration patterns in the U.S. In

addition, a simultaneous equations approach was used to investigate whether “people follow jobs” or “jobs follow people”. Areas with high family incomes have relatively higher demand for goods and services, leading to higher levels of service and commercial employment, but lower levels of manufacturing employment. Lower levels of manufacturing employment were likely influenced by relatively higher land prices in areas with high family income and potentially more expensive residential real estate.

This research further attempts to delve into the question of whether “people follow jobs” or “jobs follow people” (Carlino and Mills, 1987). Policy makers frequently face choices of which types of regional economic development policies to support. They can support policies which influence business’ location decisions, or policies which influence peoples’ location decisions. This research attempts to identify a relationship between residential real estate as a measure of the value people place on a location and employment as a measure of how firms value a location.

Regional geographic variation within the U.S. influences population location decisions (Carlino and Mills, 1987). Counties in the relatively warmer Sunbelt region were more attractive than counties in the relatively colder regions in the U.S. Interregional differences within the U.S. were important, but intraregional differences were less important. Accordingly, there were large differences between counties in separate regions, but only small differences between counties in the same region. Counties in the Northeast were statistically different from counties in the South and other regions, but counties within the Northeast region were relatively similar to each other. If counties outside the Northeast are compared to counties in the Northeast, bias reflecting amenities like climate would affect the results. This study does directly consider the effects of amenities on housing location decisions. In order to limit the introduction of bias reflecting

amenities into this study, only counties in the Northeast are utilized, as they have been shown to be statistically similar to each other.

## **Literature Review**

Investing in residential real estate is both a consumption decision and a major investment (Plaut, 1987). Henderson and Ioannides (1979) designed a model of tenure choice where housing serves two purposes, a consumption good and an investment holding in a portfolio. A consumer simultaneously chooses the optimal level of housing consumption and optimal portfolio holdings. When consumption demand is less than investment demand, the consumer owner-occupies that portion of investment equaling consumption demand and rents out the remainder. If consumption demand is greater than investment demand, the consumer cannot own part of his consumption and must rent (Henderson and Ioannides, 1979).

For middle-income households in particular, home ownership is a very one-sided investment portfolio, with substantial financial risks (Forrest and Murie, 1989). Housing remains the largest asset investment of most American families. Some studies have shown that owning a house often put households in a favorable position relative to those who remained tenants for life (Kendig, 1984; Badcock, 1989; Hamnett, 1991). Various authors argue that accumulation of wealth through home ownership may well have developed into a critical social divide. Government subsidies encourage home ownership, but not everyone is in a financial position to take advantage.

Engelhardt (1995) examines how changes housing values cause homeowners to alter their savings and consumption behavior. If residential real estate is an investment, than an increase in the value of that investment should cause owners to alter their consumption and saving behavior, likewise with a decrease. Using PSID data from 1984 and 1989, Engelhardt (1995) showed that

owners who experienced gains in the value of their residential real estate did not tend to save more over time, while owners whose residential real estate investments lost value tended to save more income in attempt to compensate. This behavior demonstrates that owners realize that residential real estate investment is a primary savings vehicle for a large section of the U.S. population, and that when that investment decreases in value owners tend to increase savings. An increase in savings may lead to reduced consumption, as more income is saved instead of spent during the present time period.

An alternative view in economics is that housing prices are driven primarily by construction costs. For example, this view was neatly laid out in 1956 by Grebler, Blank and Winnick (1956). This model considers that people did not view housing as a speculative asset: almost all of the value of houses has been value of structure, a manufactured product. From this view, there would be no reason to think that one can make money by buying houses and holding them for resale than that one can make money by buying tables and chairs and holding them for resale (Schiller, 2007).

Gyourko et al. (2006) argues that great cities will indefinitely outperform the economy in general. They found that some “superstar cities” have shown long-term, that is 50-year, appreciation above national averages (Gyourko, Mayer et al., 2006). They use Census decadal owners’ evaluations of the value of their homes, but found only relatively small excess returns to homes in those cities. They report much smaller differences across cities than people expect. Their paper found that Los Angeles grew at 2.46% a year real 1950-2000, but this is far below the kind of expectations we have seen recently (Gyourko, Mayer et al., 2006). Moreover, in the decadal Census data there is no correction for quality change, and yet homes have been getting



larger in the superstar cities, so the actual appreciation of existing homes was likely less (Schiller, 2007).

DiPasquale, Forslid and Glaeser (2000) have found that homeowners tend to be more involved in local government, are more informed about their political leaders and join more organizations than renters do, even after controlling for other factors. This view has led to widespread political support for policies that encourage homeownership over much of the world, including the mortgage interest deduction in the U.S.

Alternately, there are many sensible reasons for people to rent rather than own. Some people who cannot currently bear the responsibilities of household management, who are likely to move soon or who have other plans for their time, should rent rather than own. Renting rather than owning encourages a better diversification of investments; many homeowners have very undiversified investment portfolios, and these investments are often highly leveraged (Schiller, 2007). Moreover, creating too much attention to housing as investments may encourage speculative thinking, and therefore, excessive volatility in the market for homes. Encouraging people into risky investments in housing may have bad outcomes (Schiller, 2007).

The physical nature of land and houses as forms of capital requires a different treatment than other forms. This leads Mayer and Somerville (2000) to examine the effect of housing construction in the general economy of a region and how new housing construction often leads both recessions and recoveries. Land and housing is physical capital, and is not as mobile as other forms of capital. When sold, the capital still occupies the same location. Also, housing capital physically depreciates and can be removed from the market.

The effect of spatial autocorrelation and spatial heterogeneity in housing markets in Dijon, France was investigated by Baumont (2007). Using a hedonic price function to account

for spatial effects, neighborhood attributes, and accessibility, a spatial error model showed that lower income areas within Dijon affected surrounding areas. Housing prices are autocorrelated due to their nature as a durable good in a fixed location. In cities and suburbs, houses within a neighborhood are often built at the same time and with similar structural features (Baumont, 2007). Neighborhoods draw from similar labor markets and amenities like schools and parks as well.

Employment, income, net migration, and government expenditure in Appalachia have been shown to be spatially correlated (Gebremariam, Gebremedhin et al., 2007). Utilizing panel data and a spatial approach to generalized Three-Stage Least Squares, the authors show the existence of feedback simultaneities between employment, income, migration, and government expenditures. Additionally, growth rates in one county were affected by the growth rates of the neighboring counties (Gebremariam, Gebremedhin et al., 2007).

Teen employment and wages are investigated using a spatial panel approach by Kalenkoski and Lacombe (2011). Since minimum wages laws vary among the states in the U.S. There exist neighboring counties with different minimum wage rates. The authors study the economic effects of this disparity on teens who earn these wages by looking at their incomes and employment. Since employment is also correlated across states, a spatial panel approach is the correct tool for this analysis. Decreases in teen employment caused by increases in the minimum wage are greater when accounting for spatial dependence than in previous studies (Kalenkoski and Lacombe, 2011)

Mayer and Somerville (2000) discuss the spatial role of housing in a metropolitan area which must be accounted for in the analysis. These results indicate that as housing prices rise, the

boundary of a city may increase and houses at the fringe should have the same value as houses that were at the fringe before the expansion of the city boundary. Thus, the general value of housing itself has not changed, but instead the boundary of the city has expanded.

By including a Bayesian component, Deller, Lledo, and Marcouiller, (2008) create an objective method to choose variables for their study of the effects of amenities on regional economic growth. The traditional empirical growth literature has been criticized as being ad hoc in the selection of right-hand-side control variables (Deller, Lledo et al., 2008). After the Bayesian step, the authors then use Spatial Error, Spatial Autoregressive, and Spatial Durbin Models for estimation.

Jeanty, Partridge, and Irwin (2010) use Michigan Census tract-level data, to estimate a spatial simultaneous equations model that jointly models population change and housing values. The model explicitly considers the spatial interactions between housing price and population change within and across neighborhoods, while also controlling for spurious correlations (Jeanty, Partridge et al., 2010). The model is estimated using a generalized spatial two-stage (GS2SLS) procedure based on the work of Kelejian and Prucha (2004). Their results demonstrate the significance of substantive spatial interactions with neighboring census tracts. Average tract-level housing values are positively associated with average housing values of neighboring tracts. Population gains in neighboring tracts have a positive influence on population gains in the original tract. In addition, the coefficient estimates provide evidence of feedback simultaneity: neighborhoods are likely to experience an increase in their housing values if they gain population and they are more likely to lose population if they experience an increase in housing values. In decomposing the interactions of population change and housing values into direct (the “own” effect), indirect (spatial spillover effect) and total impacts (Jeanty, Partridge et al., 2010) find that

housing value in a census tract is affected by a change in population growth in both the own and neighboring tracts, while population growth in a tract is only affected by changes in housing values in the same tract. Strong evidence of spatial spillovers in both population and housing values indicate that research models should account for spatial interaction of these variables.

### **Theoretical Model**

Firms and individuals simultaneously choose where to locate based upon the costs and benefits of the qualities of each location (Roback, 1982). The basic assumptions for a simplified model include: both capital and labor are completely mobile across cities; land is fixed in quantity for cities but allowed to be mobile for uses; workers ignore leisure, have homogeneous preferences, and supply one unit of labor independent of the wage rate. If every location has a vector of characteristics  $s$  and each worker can only produce and consume a good  $x$ , then each worker will attempt to maximize their utility with respect to their consumption of  $x$ , and the amount of residential land consumed  $l^c$  with respect to their budget constraint. That is each worker will attempt to:

$$(1) \max U(x, l^c; s) \text{ subject to } w + r l^c = x + I$$

Where wage and rental payments are  $w$  and  $r$ , respectively, and non-labor income is  $I$  ( $I$  does not depend on location).

Indirect utility function  $V$  is associated with equation (1) is:

$$(2) V(w, r; s) = k$$

Wages and rents across location must be equal, or else workers will move to locations where they can have a higher utility.

Firms also face a similar location choice. Firms use land,  $l^p$ , and the number of workers in a city,  $N$ , to produce good  $x$ . Firms face constant returns to scale, and locate so as to minimize

the cost of producing good  $x$ . By assumption, the firms' equilibrium condition is when cost equals price, assumed to be 1. Therefore, firms face cost function:

$$(3) C(w,r;s) = I$$

If not, firms can choose to move to other cities where profits are higher. In order to fully describe the problem, the wages and rents must be at equilibrium within both markets. Roback (1982) uses a simplified model to show that wages, the number of employees per city, land rents, and amenities all combine in equilibrium to describe cities. This research will attempt to utilize this theoretical relationship among these variables.

For the empirical analysis both non-spatial and spatial models will be used. A brief description of the models is given below:

### ***Spatial Model***

The focus of this study is to analyze the relationship between residential real estate investment and economic growth represented by changes in population, employment, and median income. Spatial dependence is an important factor in regional economic growth analysis, especially in terms of population, employment, and per capita income (LeSage and Fischer, 2009). In cases of simultaneous equations, spatial dependencies appear due to two reasons (Kelejian and Prucha, 1999, 2004). The first reason is that error terms are not only assumed to be spatially correlated but also correlated across equations. The second reason is that the value of the endogenous variables in a given equation is assumed to depend upon a weighted sum of those endogenous variables over neighboring regions.

The study is derived from a spatial autoregressive model with autoregressive disturbances for the simultaneous equations of Kelejian and Prucha (2004). They built this model by extending Cliff-Ord's single equation model (1973, 1981). Kelejian and Prucha (2004) extended

it into a simultaneous system to estimate a simultaneous system of cross-sectional equations with spatial dependencies. To examine the feedback simultaneities among the endogenous variables of the model, the existence of spatial autoregressive lag effects and spatial cross-regressive lag effects with respect to the endogenous variables of the model, we will use a system of spatial simultaneous equation of population, employment, income, and residential real estate investment.

Anselin (1988) argued that in the presence of spillover effects, estimation of the econometric model will be biased or inefficient if spatial dependencies are ignored in the model. Anselin (1988) also showed that OLS estimation results are inconsistent.

This means that the non-spatial simultaneous equations should be estimated by incorporating spatial dependency. Two approaches, which incorporate spatial dependencies, are the Spatial Durbin Model (SDM) and the Spatial Autoregressive (SAR) model. LeSage and Pace (2009) explained that SDM incorporates not only spatial lag of dependent variables but also independent variables. LeSage and Fischer (2009) indicated that SDM also deals with omitted variable bias.

A simple pooled linear regression model with spatial specific effects is considered, but without spatial interaction effects.

$$(4) \quad y_{it} = x_{it}\beta + \mu_i + \varepsilon_{it}$$

where  $i$  is an index for the cross-sectional dimension (spatial units), with  $i=1, \dots, N$ , and  $t$  is an index for the time dimension (time periods), with  $t=1, \dots, T$ .  $y_{it}$  is an observation on the dependent variable at  $i$  and  $t$ ,  $x_{it}$  an  $(1, K)$  row vector of observations on the independent variables, and  $\beta$  a matching  $(K, 1)$  vector of fixed but unknown parameters.  $\varepsilon_{it}$  is an independently and identically distributed error term for  $i$  and  $t$  with zero mean and variance  $\sigma^2$ , while  $\mu_i$  denotes a spatial

specific effect. The standard reasoning behind spatial specific effects is that they control for all space-specific time-invariant variables whose omission could bias the estimates in a typical cross sectional study (Elhorst, 2010).

When specifying interaction between spatial units, the model may contain a spatially lagged dependent variable or a spatial autoregressive process in the error term, known as the spatial lag and the spatial error model, respectively. The spatial lag model posits that the dependent variable depends on the dependent variable observed in neighboring units and on a set of observed local characteristics

$$(5) \quad y_{it} = \delta \sum_{j=1}^N w_{ij} y_{jt} + x_{it} \beta + \mu_i + \varepsilon_{it}$$

where  $\delta$  is called the spatial autoregressive coefficient and  $w_{ij}$  is an element of a spatial weights matrix  $\mathbf{W}$  describing the spatial arrangement of the units in the sample (Elhorst, 2010). It is assumed that  $\mathbf{W}$  is a pre-specified non-negative matrix of order  $N$ .

The spatial error model, on the other hand, posits that the dependent variable depends on a set of observed local characteristics and that the error terms are correlated across space

$$(6) \quad y_{it} = x_{it} \beta + \mu_i + \phi_{it}$$

$$(7) \quad \phi_{it} = \rho \sum_{j=1}^N w_{ij} \phi_{jt} + \varepsilon_{it}$$

where  $\phi_{it}$  reflects the spatially autocorrelated error term and  $\rho$  is called the spatial autocorrelation coefficient.

In both the spatial lag and the spatial error model, stationarity requires that  $1/\omega_{min} < \delta < 1/\omega_{max}$  and  $1/\omega_{min} < \rho < 1/\omega_{max}$ , where  $\omega_{min}$  and  $\omega_{max}$  denote the smallest (i.e., most negative) and largest characteristic roots of the matrix  $\mathbf{W}$ . While it is often suggested in the literature to constraint  $\delta$  or  $\rho$  to the interval  $(-1,+1)$ , this may be unnecessarily restrictive (Elhorst,2010).

An unconstrained spatial Durbin model with spatial fixed effects looks like

$$(8) \quad y_{it} = \delta \sum_{j=1}^N w_{ij} y_{jt} + x_{it} \beta + \sum_{j=1}^N w_{ij} x_{ijt} \gamma + \mu_i + \varepsilon_{it}$$

where  $\gamma$ , just as  $\beta$ , is an  $(K,1)$  vector of fixed but unknown parameters. The hypothesis  $H_0: \gamma=0$  can be tested to investigate whether this model can be simplified to the spatial lag model and the hypothesis  $H_0: \gamma+\delta\beta=0$  whether it can be simplified to the spatial error model.

To test for spatial interaction effects in a cross-sectional setting, Anselin et al. (1996) developed Lagrange Multiplier (LM) tests for a spatially lagged dependent variable, for spatial error correlation, and their counterparts robustified against the alternative of the other form. These tests have become very popular in empirical research. Recently, Anselin et al. (2006) also specified the first two LM tests for a spatial panel

$$(9) \quad LM_{\delta} = \frac{[e'(I_T \otimes W)Y / \hat{\sigma}^2]^2}{J} \quad \text{and} \quad LM_{\rho} = \frac{[e'(I_T \otimes W)e / \hat{\sigma}^2]^2}{T * T_w}$$

where the symbol  $\otimes$  denotes the Kronecker product,  $I_T$  denotes the identity matrix and its subscript the order of this matrix, and  $e$  denotes the residual vector of a pooled regression model without any spatial or time specific effects or of a panel data model with spatial and/or time period fixed effects. Finally,  $J$  and  $T_w$  are defined by

$$(10) \quad J = \frac{1}{\hat{\sigma}^2} [((I_T \otimes W)X \hat{\beta})'(I_{NT} - X(X'X)^{-1}X')(I_T \otimes W)X \hat{\beta} + TT_w \hat{\sigma}^2]$$

$$(11) \quad T_w = \text{trace}(WW + W'W)$$

The robust LM tests for a spatial panel are

$$(12) \quad \text{robust LM}_{\delta} = \frac{[e'(I_T \otimes W)Y / \hat{\sigma}^2 - e'(I_T \otimes W)e / \hat{\sigma}^2]^2}{J - TT_w}$$



$$(13) \text{ robust LM}_{\rho} = \frac{[e'(I_T \otimes W)e / \hat{\sigma}^2 - TT_w / J * e'(I_T \otimes W)Y / \hat{\sigma}^2]^2}{TT_w[1 - TT_w / J]}$$

The most appropriate spatial model will be selected for analysis based on the statistical significance of the  $\rho$  and  $\delta$  variable results from the two spatial models, and by the statistical significance of the LM tests for errors and lags.

### **Specification of Variables**

The empirical models are used to analyze the effect of residential real estate investment in regional economic growth using changes in population, employment, and median income. The model will be explained as a spatial panel with residential real estate as a function of human capital, economic, and demographic variables.

The residential real estate dependent variable is the median housing value per county in dollars (HVM). The HVM values are inflated to their 2010 value using the Housing Price Index (HPI). Explanatory variables for residential real estate investment include the number of banks per county (BNK), the number of new building permits per county (PER), the number of vacant housing units per county (VAC), and the number of occupied housing units per county (OCC). Number of banks and permits are related to the availability of financing and new housing constructions, while the number of vacancy and households are related to the supply of vacant housing and the potential demand for housing.

The population density per county (DEN) as an independent variable is directly related to HVM, as areas with an increase in population density during the study period should have an increase in the demand for residential real estate. Government expenditures per county (GOV) is also an explanatory variable for the population growth.

The employment per county (EMP) as an independent variable is directly related to HVM, as areas with an increase in employment during the study period should have an increase

in the demand for residential real estate. The explanatory variable for employment is number of businesses per county (BUS), as areas with more businesses, may have increased demand for housing for workers, increasing residential housing values.

The change in median income per county (INC) is another independent variable, inflated to its 2010 value using the Consumer Price Index (CPI). Median income is directly related to HVM, as an increase in the median income per county means an increase in the demand for residential real estate, leading to a higher value for residential real estate. The number of people living in poverty per county (POV) is an explanatory variable for income. Poverty is a reflection of actual income situation and the surrounding area.

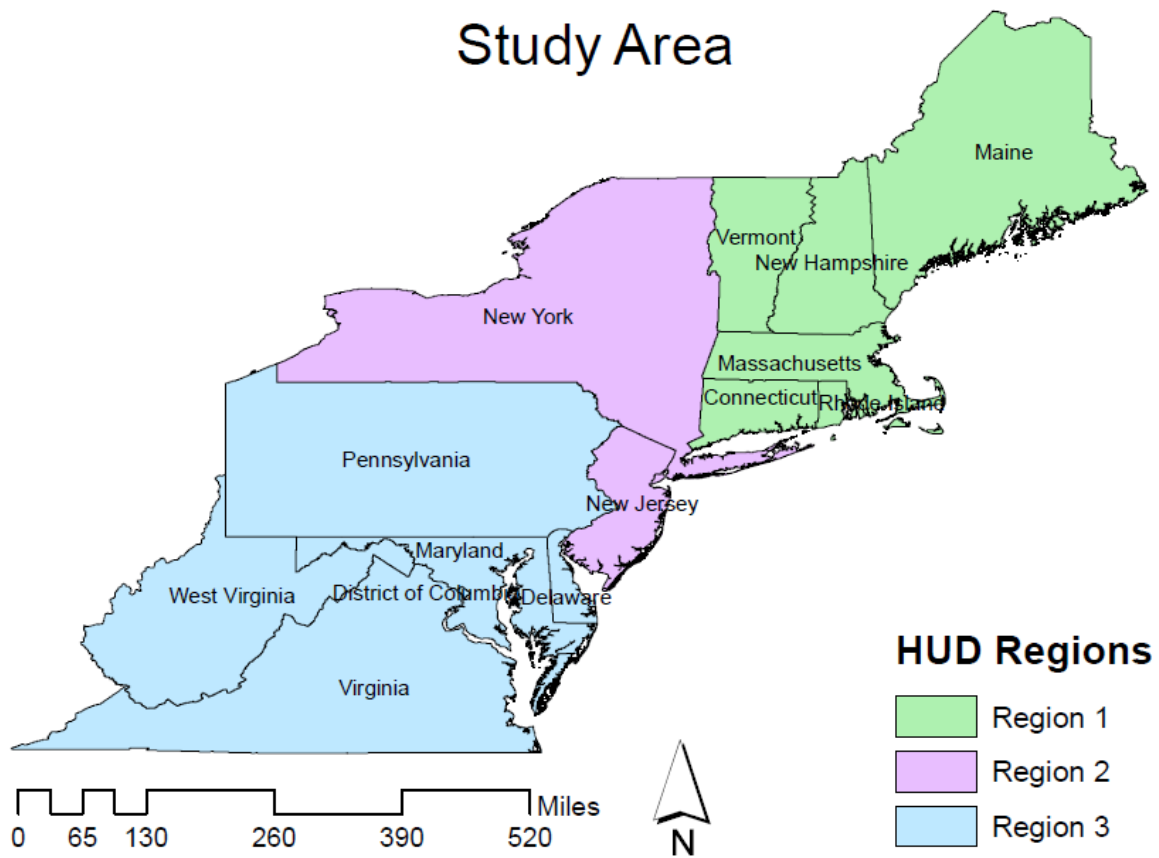
**Table 1: Definition of variables**

<b>Variable</b>	<b>Definition</b>
<b>Endogenous variables</b>	<b>Per county</b>
HVM	Median Value of residential real estate
<b>Exogenous variables</b>	
DEN	Population density per county
EMP	Employment
INC	Median income
MET	Dummy variable denoting metropolitan county
GOV	Federal government expenditure
BUS	Number of businesses
POV	Number of people below poverty line
BNK	Number of banks
PER	Number of new housing permits
OCC	Number of occupied housing units
VAC	Number of vacant houses

A dummy variable denoting metropolitan counties (MET) is also included in every equation. The metropolitan designation follows the USDA Economic Research Service definition of metropolitan counties. Some counties in the 1980 and 1990 time periods are not designated metropolitan counties, but become metropolitan in later time periods due to population growth.

This study focuses on the counties of the Northeast region and the District of Columbia of the United States for the census years between 1980 and 2010 as shown in Figure 1. County level data for endogenous and exogenous variables were collected from the US Census Bureau, American Community Survey, and the US Department of Agriculture, Economic Research Service. The study area coincides with the definition of the Department of Housing and Urban Development Regions 1, 2, and 3.

**Figure 1. Study Area**



### Empirical Results

The results from the LM tests discussed in equations 9, 12 and 13 are given in Table 2 below. As the LM statistic is larger than the corresponding  $\chi^2$  statistic with 1 degree for

freedom, the spatial models will not collapse to OLS, and OLS is not appropriate for this estimation. The results indicate that the most appropriate model is the spatial autoregressive model, as the LM statistic for the  $LM_{\delta}$  is larger than the LM statistic for the  $LM_{\rho}$  tests.

**Table 2: LM test results**

	$LM_{\delta}$	Robust $LM_{\delta}$	$LM_{\rho}$	Robust $LM_{\rho}$
LM statistic	1948.4	676.845	1293.3	21.7063
Probability	0	0	0	0.000003
Chi <sup>2</sup> 1 d.o.f.	6.64	6.64	6.64	6.64
# of observations	1736	1736	1736	1736
# of variables	11	11	11	11
Time periods	4	4	4	4

The results of the SAR model including both county and year fixed effects are given in Table 3. The spatial weight matrix used for this analysis is a 5 nearest neighbor weight matrix based on the physical centroids for the 434 counties in the study area.

Overall fit can be measured by the  $R^2$  of 0.9901 which is initially viewed as very high. Elhorst (2010) recommend that squared correlation coefficient be used as an alternative measure of goodness-of-fit because of the nature of fixed effects. The squared correlation coefficient for this model is 0.4067. The difference between the  $R^2$  and the squared correlation coefficient indicates how much of the variation in the dependent variable is explained by the fixed effects. The fixed effects portion of this model explains approximately 58% of the variation in the dependent variable.

The results from this SAR model indicate that there is a significant level of spatial autocorrelation in the dependent variable, with the  $\delta$  parameter equal to 0.84311 and is significant at the 1% level. This result confirms the usage of the SAR model as the appropriate model. This result is also used to calculate the proper marginal effects. LeSage and Pace (2009)

show that the marginal effect of a change in an explanatory variable is depends on both the direct and indirect effects which depend on the value of  $\delta$ .

The primary focus of this research is to determine the effects of changes in county level median income, population density, and employment on median housing values. Theory states that increases in these three explanatory variables should increase median housing values, as they should cause increases in the demand for housing.

The direct effect of a change in a county's income measures how a change in a particular county's income affects median housing value in that same county. From Table 3, the direct effect of a change in a county's median income on the median housing value is 0.011017 and is significant at the 5% level. This means that as a county increases its own income by 10%, median housing value in that same county increases by 0.11017%. A major advantage of the spatial econometric techniques is their ability to quantify spatial spillovers in the form of the indirect effects. The indirect effect estimate is 0.042484, and it is significant at the 5% level. As a county increases its median income, median housing value in neighboring counties (as defined by the 5 nearest neighbor weight matrix) increases. The indirect effect is larger than the direct effect, showing that an increase in median income in one county has a larger effect on median housing values in neighboring counties than in its own county.

The final effect estimate requiring discussion is the total effect, which is the sum of the direct effect and the indirect effect. Arguably, this is the most important quantity that needs interpretation in that the total effect measures how changes in the median income affects median housing value, considering own-county and neighboring county spillover effects. The point estimate for the total effect of a change in the median income is 0.053501 and is statistically significant at the 5% level. The total effect estimate shows that, as the median income increases

by 10%, median housing value increases by 0.53501%. An increase in income should lead to an increase in residential real estate investment, as consumers should spend more income on housing, increasing the demand for housing.

The direct effect of a change in a county's population density measures how a change in a particular county's income affects median housing value in that same county. From Table 3, the direct effect of a change in a county's population density on the median housing value is 0.088114 and is significant at the 5% level. This means that as a county increases its own population density by 10%, median housing value in that same county increases by 0.88114%. The indirect effect estimate is 0.339828, and it is significant at the 5% level. As a county increases its population density, median housing value in neighboring counties (as defined by the 5 nearest neighbor weight matrix) increases. The indirect effect is larger than the direct effect, showing that an increase in population density in one county has a larger effect on median housing values in neighboring counties than in its own county. The total effect of a change in the population density is 0.427942 and is statistically significant at the 5% level. The total effect estimate shows that, as the population density increases by 10%, median housing value increases by 4.27942%. An increase in population density should lead to an increase in residential real estate investment, as demand for housing should increase.

The direct effect of a change in a county's employment measures how a change in a particular county's income affects median housing value in that same county. From Table 3, the direct effect of a change in a county's employment on the median housing value is 0.319333 and is significant at the 1% level. This means that as a county increases its own employment by 10%, median housing value in that same county increases by 3.19333%. The indirect effect estimate is 1.229236, and it is significant at the 1% level. As a county increases its employment, median

housing value in neighboring counties (as defined by the 5 nearest neighbor weight matrix) increases. The indirect effect is larger than the direct effect, showing that an increase in employment in one county has a larger effect on median housing values in neighboring counties than in its own county. The total effect of a change in employment is 1.548569 and is statistically significant at the 1% level. The total effect estimate shows that, as employment increases by 10%, median housing value increases by 15.48569%. An increase in employment should increase residential real estate investment, as employees need places to live, increasing the demand for housing.

**Table 3: SAR spatial panel results**

Dependent Variable: Median Housing Value per county			
Explanatory Variables	Direct Effect	Indirect Effect	Total Effect
logINC	0.011017 ** (0.027248)	0.042484 ** (0.03244)	0.053501 ** (0.03091)
logDEN	0.088114 ** (0.039768)	0.339828 ** (0.04368)	0.427942 ** (0.04241)
logEMP	0.319333 * (0)	1.229236 * (0)	1.548569 * (0)
logBUS	0.087274 * (0.00024)	0.336659 * (0.0006)	0.423933 * (0.00047)
logPOV	-0.079319 * (0)	-0.305125 * (0.000017)	-0.384445 * (0.000012)
BNK	-0.0007 * (0.00005)	-0.002695 * (0.00013)	-0.003395 * (0.000095)
PER	0.000019 * (0)	0.000072 * (0)	0.00009 * (0)
logVAC	0.06845 ** (0.000022)	0.263633 * (0.000064)	0.332083 * (0.000045)
logOCC	-0.2647 * (0.00014)	-1.021003 * (0.00038)	-1.285702 * (0.00029)
MET	0.008151 (0.54714)	0.03129 (0.55114)	0.039441 (0.55009)
$\delta = 0.84311$ p-value=0			
Adjusted R <sup>2</sup> : 0.9901, Corr <sup>2</sup> : 0.4067			

The differential between the direct effects and the total effects for all of the coefficients shows that controlling for spatial dependence is important. Any research that neglects spatial dependence potentially drastically underestimates the effects of change in income, population density, and employment on housing values.

## **Conclusion**

The main objective of this study was to examine the impacts of residential real estate investments on economic development represented by population, employment, and median income by using the Spatial Autocorrelation (SAR) panel estimation methods. The study area was the Northeast US, and a panel data set was created from the Census years 1980-2010. Data was collected primarily from the US Census and also from the USDA Economic Research Service.

This study could be substantially improved with a more complete dataset. Data limitations are responsible for the choice of only census years. Census tract level data instead of county level would substantially increase the number of observations, and potentially the complexity of this problem, but should provide much better results due to the reduction in spatial aggregation bias.



## References

- Arku, G. and R. Harris "Housing as a Tool of Economic Development since 1929." *International Journal of Urban and Regional Research*, Vol. 29, (2005) pp. 895-915.
- Badcock, B.A. "Homeownership and the accumulation of real wealth." *Environment and Planning D: Society and Space*, Vol. 7, (1989) pp. 69-91.
- Baumont, C. "Neighborhood Effects, Urban Public Policies and Housing Values A Spatial Econometric Perspective", *A Spatial Econometric Perspective*: Université de Bourgogne Laboratoire d'Economie et de Gestion (CNRS), 2007).
- Carlino, G. and E. Mills "The Determinants of County Growth." *Journal of Regional Science*, Vol. 27, (1987) pp. 39-54.
- Carruthers, J.I. and G.F. Mulligan "Environmental Valuation: Connecting Theory, Evidence, and Public Policy", *Presented at the meetings of the Associated Collegiate Schools of Planning in Ft. Worth, Texas*, 2005).
- Carruthers, J.I. and G.F. Mulligan "A locational analysis of growth and change in American metropolitan areas\*." *Papers in Regional Science*, Vol. 87, (2008) pp. 155-171.
- Census "Survey of Income and Program Participation", in D.o. Commerce (ed.), (Washington DC, 2000).
- Census "Selected Historical Decennial Census, Urban and Rural Definitions and Data", in D.o. Commerce (ed.), (Washington DC, 2011).
- Deller, S.C., V. Lledo and D.W. Marcouiller "Modeling Regional Economic Growth with a Focus on Amenities." *Review of Urban & Regional Development Studies*, Vol. 20, (2008) pp. 1-21.

Deloitte "Perspectives on Real Estate Uncovering Opportunity in a Distressed Market.", in D. LLC. (ed.), 2009).

DePasquale, D., R. Forslid and E. Glaeser "Incentives and Social Capital: Are Homeowners Better Citizens?" *Journal of International Economics*, Vol. 50, (2000) pp. 497-517.

Engelhardt, G. "House prices and home owner saving behavior." *Regional Science and Urban Economics*, Vol. 26, (1995).

FDIC "2007 Annual Report", in F.D.I. Corporation (ed.), 2010).

Forrest, R. and A. Murie "Differential accumulation: wealth, inheritance and housing policy considered." *Policy and Politics*, Vol. 17, (1989) pp. 25-39.

Gebremeriam, G.H., T.G. Gebremedhin, P.V. Schaeffer, T.T. Phipps and R.W. Jackson "A Spatial Panel Simultaneous-Equations Model of Business Growth, Migration Behavior, Local Public Services and Household Income in Appalachia ", *American Agricultural Economics Association* (Portland, OR, 2007).

Grebler, L., D. Blank and L. Winnick *Capital Formation in Residential Real Estate* (Princeton, National Bureau of Economic Research and Princeton University Press, 1956).

Gyourko, J., C. Mayer and T. Sinai "Superstar Cities", *NBER Working Paper No. 12355*: National Bureau of Economic Research, 2006).

Hamnett, C. "Home ownership, housing wealth, and wealth distribution in Britain", *Conference Housing Policy as a Strategy for Change* (Oslo, Norway, 1991).

Henderson, J.V. and Y. Ioannides "A model of housing tenure choice." *American Economic Review*, Vol. 73, (1979) pp. 98-113.

- Jeanty, P.W., M. Partridge and E. Irwin "Estimation of a spatial simultaneous equation model of population migration and housing price dynamics." *Regional Science and Urban Economics*, Vol. 40, (2010) pp. 343-352.
- Kalenkoski, C.M. and D.J. Lacombe "Minimum Wages and Teen Employment: A Spatial Panel Approach", *Working Paper Series Research Paper Number 2011-08* (Morgantown, WV: Regional Research Institute, 2011).
- Kelejian, H.H. and I.R. Prucha "Estimation of Simultaneous Systems of Spatially Interrelated Cross Sectional Equations." *Journal of Econometrics*, Vol. 118, (2004) pp. 27-50.
- Kendig, H.L. "Housing careers, life cycle and residential mobility: implications for the housing market." *Urban Studies*, Vol. 4, (1984) pp. 271-283.
- Leung, C. "Macroeconomics and housing: a review of the literature." *Journal of Housing Economics*, Vol. 13, (2004) pp. 249-267.
- Mayer, C.J. and C.T. Somerville "Residential Construction: Using the Urban Growth Model to Estimate Housing Supply." *Journal of Urban Economics*, Vol. 48, (2000) pp. 85-109.
- Ofori, G. and S.S. Han "Abraham J.H., P., Bubbles in Metropolitan Housing Markets, *Journal of Housing Research* 7(1996), 191-207." *Abraham J.H., P., Bubbles in Metropolitan Housing Markets, Journal of Housing Research* 7(1996), 191-207., Vol. 27, (2003) pp. 37-62.
- Plaut, S.E. "The timing of housnig tenure transition." *Journal of Urban Economics*, Vol. 21, (1987) pp. 312-322.
- Schiller, R. "Understanding Recent Trends in House Prices and Home Ownership", *NBER Working paper No. 13553*: National Bureau of Economic Research, 2007).

Turner, T.M. and H. Luea "Homeownership, Wealth Accumulation and Income Status." *Journal of Housing Economics*, Vol. 18, (2009) pp. 104-114.