A Non-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
1168 Agricultural Sciences Building
PO Box 6108
Morgantown, WV-26506
E-mail: praveena.jayaraman@mail.wvu.edu

Dr. Tesfa Gebremedhin
Professor
Davis College of Agriculture, Natural Resources and Design
1168 Agricultural Sciences Building
PO Box 6108
Morgantown, WV-26506
Phone: 304-293-5535
E-mail: tesfa.gebremedhin@mail.wvu.edu

Copyright 2013 by Praveena Jayaraman 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.
Introduction

Residential real estate investment has been recognized as an agent of economic growth since the 1970s 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 different 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) (Census, 2000), residential real estate was the largest class of assets held by individuals amounting to 78.7 percent of total household asset value. 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 (Deloitte, 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 (Census, 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 the increase in urban population and 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 (FDIC, 2010). In this region, commercial and residential real estate markets grew quickly due to strong regional employment and economic growth during 1982-1988. 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 year. In spite of the problem of acquiring affordable quality housing and available credit in rural areas, ownership of real estate is one of the best methods of asset accumulation for low-income rural households.

Carlino and Mills’ simultaneous equations model estimated employment and population changes in U.S. counties and explained the migration patterns in the US (Carlino and Mills, 1987). Additionally, a simultaneous equations approach is 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 will further attempt to delve into the question of whether “people follow jobs” or “jobs follow people” (Carlino and Mills, 1987). To stimulate regional economic growth, policy makers may need to know if they should pursue policies to influence the location and utility decisions of firms or of people. In conjunction with information about population and income this research plans to identify a causal relationship between residential real estate as a measure of the value people place on living in a location and employment as a measure of how firms value a location.

Regional 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 were compared to counties in the Northeast, bias reflecting amenities like climate would affect the results. This study will not 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 will be utilized, as they have been shown to be statistically similar to each other.
Literature Review

An often discussed idea in the housing literature is the assertion that a long-run equilibrium relationship between house prices and fundamentals, such as income, population and costs associated with home ownership. The validity of this assumption has important implications for how residential real estate dynamics are modeled. If the assumption is accurate, so that residential real estate prices are cointegrated with and fundamentals, then the error-correction specifications of Abraham and Hendershott (1996), Malpezzi (1999) and Capozza et al. (2002) are appropriate. Gallin (2006) used standard tests to show that there is little evidence for cointegration of house prices and various fundamentals at the national level. Additionally, Gallin (2006) showed that bootstrapped versions of more powerful panel-data tests, applied to a panel of 95 U.S. metropolitan areas over 23 years, also do not find evidence for cointegration. This shows that the level of residential real estate investment does not appear to have a stable long run equilibrium relationship with the level of fundamentals such as income or population.

Whether house prices are cointegrated with economic fundamentals such as incomes or employment was investigated by Zhou (2010). The existence of cointegration would imply that a reduction in the price of housing was caused by market forces. More specifically, cointegration means that there is a long run equilibrium relationship between residential real estate investment and economic fundamentals (Zhou, 2010). Conversely, failure to find cointegration suggests that a reduction in the price of housing would be random. Zhou (2010) selected ten cities (Boston, Chicago, Cleveland, Dallas, Los Angeles, New York, Philadelphia, Richmond, Seattle and St. Louis), and tested first for linear cointegration, then for nonlinear cointegration. Cleveland showed evidence of linear cointegration. Six of the nine remaining cities (Chicago, Dallas, Philadelphia, Richmond, Seattle and St. Louis) and the country as a whole showed evidence of...
nonlinear cointegration leaving three cities (Boston, Los Angeles and New York) with no
evidence of linear or nonlinear cointegration (Zhou, 2010). These results show that the
fundamentals of residential real estate investment are diverse at the city level, reflecting the local
nature of the housing market. Local conditions including employment and income affect
residential real estate investment decisions.

The interdependency of housing prices, rental prices, building land prices and income via
one simultaneous equilibrium analysis is used by Bischoff (2012) with unique cross sectional
data on the majority of German counties and cities for 2005, to estimate the equations in their
structural and reduced form. The results show significantly positive interaction effects of income
and real estate prices. Bischoff (2012) also uses 2SLS to model the impact of exogenous
changes on prices, income and population, showing the positive effect of population on
residential real estate investment and income.

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 depreciates and can be removed from the market. They also 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.
Using US post-war data for empirical analysis, Wen (2001) investigates the relationship between economic growth and fixed capital formation. The results show that differentiating residential investment from business investment is important in analyzing the relationship between capital formation and economic growth. Most household savings are in the form of real estate, and that economic booms often follow real estate booms and economic recessions follow real estate declines. He also shows that capital formation in the household sector unambiguously and unilaterally affects the growth of GDP, which in turn affects capital formation in the business sector.

Housing demand and population growth are also interconnected (Frame, 2008). Economic theory suggests that an increase in population in an area should increase the demand for housing, which can be measured as increases in the value of residential real estate. Although strong evidence of endogeneity problems hampers his research, Frame shows that changes in per capita aggregate income have a negative relationship with returns to housing and local population growth. This result is important for this research because it illustrates the interaction between population and residential real estate investment, and that the relationship is more complicated than traditional microeconomic theory suggests.

Using an error correction model, Johnes and Hyclak (1999) investigate some potentially significant interactions between housing and labor markets to explain the average wage, unemployment rate, labor force, and average house price in an urban area, using quarterly data from four cities, (Fort Lauderdale, FL, Hartford, CT, Houston, TX, and Milwaukee, WI) for the 1980s. Unemployment and labor force changes affect house prices and house prices have a significant effect on the size of the labor force. Labor markets have a value to potential jobseekers and employers, where a deeper labor pool increases the probability that employers
find qualified workers, and increases the probability that job seekers will become employed. The value of the labor market around a potential housing purchase should be capitalized into the price of housing.

Saks (2008) argues that the large variation in the house prices across the United States is due to the differences in the supply of housing. The elasticity of housing supply impacts local labor markets due to the influence of house prices on migration. Because of this influence on migration, local employment depends critically on the construction industry to accommodate increases in housing demand. The elasticity of housing supply affects the increase in labor demand on employment, wages and house prices. These predictions are assessed empirically by tracing out the effect of an increase in labor demand on metropolitan area housing and labor markets (Saks, 2008). A three variable Vector Auto-Regression (VAR) is used to estimate the shocks on the dynamics of housing prices, employment and wages. The results show that locations with higher housing regulations lead to lower residential construction and larger increase in housing prices in response to labor demand (Saks, 2008). Also, in the long run an increase in labor demand results in lower employment in metropolitan areas, resulting in lower elasticity of housing supply. These results show that interactions between local labor markets and the housing supply play an important role in the employment growth of a particular region.

Another approach to explain the changes in migration, housing market and labor market in response to employment shocks was taken by Zabel (2012). He explains that the households’ decision to move not only depends on the job prospects but also on the relative cost of housing. Household mobility in turn is significantly affected by the housing market as the ability to move depends on the relative cost of housing across metropolitan statistical areas. The housing market affects households through two channels. First, homeowners are less likely to move than renters
due to the higher cost of moving, the mortgage lock-in effect and loss aversion, thus affecting the employment prospects of homeowners. Second, the relative cost of housing across the markets, the likelihood that a household will move from one MSA to another depends on the relative cost of residing in each city.

Quigley (2002) evaluates the relationship between U.S. economic cycles and housing prices. He also indicated that few studies have explored the causal role between outcomes in the real estate market and the overall economy. Quigley blames this lack of research on the lack of an empirical model. Without an empirical model to test the relationship between real estate outcomes and economic growth, any theoretical relationship remains unverified. This research will attempt to fill in that particular gap, utilizing an empirical model to test the relationship that Quigley identified.

**Empirical 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 per capita income. The empirical model for the study is derived from the two-equation non-spatial simultaneous equation model of Carlino and Mills (1987). They formulated their model by modifying Steinnes’ model (1982). Deller et al. (2001) extended Carlino and Mills’ frame work to three simultaneous equations, which allowed them to incorporate interdependencies among income, population and employment change. Deller et al. (2001) has been extended to estimate simultaneous relationships of economic development with entrepreneurship (Mojica, 2009; Bashir, 2011), amenities (Kahsai, 2009), environmental regulation (Nondo, 2009), and modeling small business growth, migration behavior, local public services and median household income (Gebremeriam, 2006). This study will further extend the model of Deller et al. (2001) by
estimating the simultaneous relationship between residential real estate investment and population, employment, and median income.

The general form of a four simultaneous equations model defining the interaction between population (P), employment (E), income (Y), and residential real estate (R) is specified as:

\begin{align*}
(1a) \quad P^* = f(P^*, Y^*, R^*/X^p) \\
(1b) \quad E^* = f(P^*, Y^*, R^*/X^E) \\
(1c) \quad Y^* = f(P^*, E^*, R^*/X^Y) \\
(1d) \quad R^* = f(P^*, E^*, Y^*/X^R)
\end{align*}

where \( P^*, E^*, Y^*, \) and \( R^* \) represent equilibrium levels of population, employment, median income, and residential real estate, respectively in the \( i \)th county; \( X^p, X^E, X^Y, \) and \( X^R \) are a set of exogenous variables that have either direct or indirect effects on population, employment, median income, and residential real estate.

Equations (1a) to (1d) represent actual population, employment, income, residential real estate, and exogenous variables in \( X \) that determine the equilibriums of population, employment, income, and residential real estate. Therefore, a non-spatial empirical estimation model is formed for a system of four simultaneous equations explaining population, employment, per capita income, and residential real estate, respectively. This system is defined as follows:

\begin{align*}
(2a) \quad P = \alpha_{0P} + \beta_{1P}E + \beta_{2P}Y + \beta_{3P}R + \sum \delta_{1P}X^p + u_1 \\
(2b) \quad E = \alpha_{0E} + \beta_{1E}P + \beta_{2E}Y + \beta_{3E}R + \sum \delta_{2E}X^E + u_2 \\
(2c) \quad Y = \alpha_{0Y} + \beta_{1Y}P + \beta_{2Y}E + \beta_{3Y}R + \sum \delta_{3Y}X^Y + u_3 \\
(2d) \quad R = \alpha_{0R} + \beta_{1R}P + \beta_{2R}E + \beta_{3R}Y + \sum \delta_{4R}X^R + u_4
\end{align*}
The endogenous variables $P$, $E$, $Y$, and $R$ indicate a county’s population, employment, median income, and residential real estate investment, respectively. Error terms are shown by $u_1$, $u_2$, $u_3$, and $u_4$ and the exogenous variable vector is represented by $X$. A panel data set is used, with 4 time periods, 1980, 1990, 2000, and 2010. These years are chosen because of the availability of data due to the census. Lag adjustment models assume that endogenous variables are adjusted over a period of time and not adjusted instantaneously to their equilibrium levels. Deller and Lledo (2007) and Deller et al. (2001) identified that the speed-of-adjustment coefficients are embedded in the coefficients $\alpha$, $\beta$, and $\delta$. This framework permits one to estimate structural relationships while simultaneously isolating the effects of real estate on regional economic growth. Thus, the estimation of equations (2a) to (2d) is from short-run adjustments of population, employment, per capita income, and residential real estate to long-run equilibriums ($P^*, E^*, Y^*$, and $R^*$).

**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 system of equations with endogenous variables as a function of residential real estate, human capital, economic, and demographic variables. All endogenous and exogenous variables are defined in Table 1.

The residential real estate dependent variable is the median housing value per county in dollars ($HVM$). 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 households per county (HHO). Banks and permits relate the availability of financing and new housing constructions, while
vacancy relates and households relate the supply of vacant housing and the potential demand for housing.

The population dependent variable is population per county (POP). This variable is directly related to HVM, as areas with an increase in population during the study period should have an increase in the demand for residential real estate. Government expenditures per county (GOV) and the number of households per county (HHO) are also explanatory variables to explain population.

The employment dependent variable is employment per county (EMP). This 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 dependent variable is INC. INC 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. Explanatory variables for income include people living in poverty per county (POV) and people 25 years or older with a college degree per county (EDU). Poverty is a reflection of actual income and the surrounding area. Education is an instrument for the ability of workers to earn a good wage.

Explanatory variables added to every equation include the population above 65 (P65), and population that is under 18 (P18). These variables relate segments of the population that are generally not in the labor force, and affect the amount of housing of housing that must be purchased by households. Additionally, the P65 group generally has a lower income, also
affecting residential real estate. A dummy variable denoting metropolitan counties is also included in every equation. The metropolitan designation is taken from and 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.

Table 1: Definition of variables

<table>
<thead>
<tr>
<th>Variable</th>
<th>Definition</th>
</tr>
</thead>
<tbody>
<tr>
<td>POP</td>
<td>Population</td>
</tr>
<tr>
<td>EMP</td>
<td>Employment</td>
</tr>
<tr>
<td>INC</td>
<td>Median income</td>
</tr>
<tr>
<td>HVM</td>
<td>Median Value of residential real estate</td>
</tr>
<tr>
<td>P18</td>
<td>Number of people below the age of 18</td>
</tr>
<tr>
<td>P65</td>
<td>Number of people above the age of 65</td>
</tr>
<tr>
<td>MET</td>
<td>Dummy variable denoting metropolitan county</td>
</tr>
<tr>
<td>GOV</td>
<td>Federal government expenditure</td>
</tr>
<tr>
<td>HHO</td>
<td>Number of households</td>
</tr>
<tr>
<td>BUS</td>
<td>Number of businesses</td>
</tr>
<tr>
<td>POV</td>
<td>Number of people below poverty line</td>
</tr>
<tr>
<td>EDU</td>
<td>Number of people 25 years and older with college degree</td>
</tr>
<tr>
<td>BNK</td>
<td>Number of banks</td>
</tr>
<tr>
<td>PER</td>
<td>Number of new housing permits</td>
</tr>
<tr>
<td>VAC</td>
<td>Number of vacant houses</td>
</tr>
</tbody>
</table>

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

This section concentrates on estimation of the empirical models for determining the relationship between residential real estate investment and regional economic development. The simultaneous equation empirical model is estimated using the three stage least square (3SLS) method and the statistical package Stata. 3SLS is used to overcome the problem of correlation of the error terms of each equation, especially the population and employment equations. 3SLS also accounts for all restrictions on parameters in the system of simultaneous equations. The endogeneity of the four variables, population, employment, income, and residential real estate investment is tested with a Hausman test. A strongly balanced panel data set with 1736
observations for all the counties in the Northeast region is used with to estimate a fixed effects model using 3SLS via Stata.

**Population Equation**

The results of the population growth equation for the Northeast region using 3SLS are presented in Table 2. The population growth equation is estimated against endogenous variables of employment (EMP), median income (INC), and residential real estate investment as measured by median house values (HVM); and control variables are included to measure economic effects. The overall fit (R²) of the empirical population equation is 93 percent.

The empirical results show that population is positively and significantly related to employment which explains that an increase in the number of jobs also increases population. A significant and positive relationship between population and median income indicates that population increases as median income increases. The relationship between residential real estate and population is insignificant. Local government spending programs are amenities that improve living conditions within a county, and as expected government expenditures has a significant and positive relationship to population, but the effect is small. Households is barely significant, and metropolitan counties are not statistically significant. The variables households and metropolitan are both correlated with population, as they are both directly related to population, these variables were included because a Wald test showed they were jointly significant.
Table 2: Three Staged Least Square Results of Population Equation

<table>
<thead>
<tr>
<th>Variable</th>
<th>Coefficient</th>
<th>Std. Error</th>
<th>z statistic</th>
<th>p value</th>
<th>Confidence</th>
<th>Interval</th>
</tr>
</thead>
<tbody>
<tr>
<td>EMP</td>
<td>1.272415</td>
<td>0.0360514</td>
<td>38.50</td>
<td>0.000</td>
<td>1.207633</td>
<td>1.337197</td>
</tr>
<tr>
<td>INC</td>
<td>0.1509771</td>
<td>0.0088794</td>
<td>17.03</td>
<td>0.000</td>
<td>0.133597</td>
<td>0.1683572</td>
</tr>
<tr>
<td>HVM</td>
<td>-0.0055509</td>
<td>0.0118157</td>
<td>-0.47</td>
<td>0.639</td>
<td>-0.0287093</td>
<td>0.0176075</td>
</tr>
<tr>
<td>P18</td>
<td>1.239722</td>
<td>0.0319155</td>
<td>38.84</td>
<td>0.000</td>
<td>1.177168</td>
<td>1.302275</td>
</tr>
<tr>
<td>P65</td>
<td>0.2525964</td>
<td>0.0374633</td>
<td>6.74</td>
<td>0.000</td>
<td>0.1791696</td>
<td>0.3260232</td>
</tr>
<tr>
<td>GOV</td>
<td>0.0003922</td>
<td>0.0000596</td>
<td>6.58</td>
<td>0.000</td>
<td>0.0002753</td>
<td>0.000509</td>
</tr>
<tr>
<td>HHO</td>
<td>-0.0113248</td>
<td>0.0055973</td>
<td>-2.02</td>
<td>0.043</td>
<td>-0.0222952</td>
<td>-0.0003543</td>
</tr>
<tr>
<td>MET</td>
<td>126.6866</td>
<td>978.5033</td>
<td>0.13</td>
<td>0.897</td>
<td>-1791.145</td>
<td>2044.518</td>
</tr>
<tr>
<td>CONSTANT</td>
<td>4499.727</td>
<td>1448.697</td>
<td>3.11</td>
<td>0.002</td>
<td>1660.332</td>
<td>7339.122</td>
</tr>
</tbody>
</table>

Employment Equation

The results of the employment growth equation for the Northeast region using 3SLS are presented in Table 3. The employment growth equation is estimated against endogenous variables of population (POP), median income (INC), and residential real estate investment as measured by median house values (HVM); and control variables are included to measure economic effects. The overall fit ($R^2$) of the empirical population equation is 78 percent.

The empirical results show that employment is positively and significantly related to population which explains that an increase in the number of people also increases employment. A significant and negative relationship between employment and median income indicates that employment increases as median income decreases. The relationship between residential real estate and employment is positive and significant. The number of businesses in a county is positively and significantly related to employment, which makes sense as all businesses should have at least one, and probably more than one, employee. Metropolitan counties are again not statistically significant.
Table 3: Three Staged Least Square Results of Employment Equation

<table>
<thead>
<tr>
<th>Variable</th>
<th>Coefficient</th>
<th>Std. Error</th>
<th>z statistic</th>
<th>p value</th>
<th>Confidence Interval</th>
<th>Lower</th>
<th>Upper</th>
</tr>
</thead>
<tbody>
<tr>
<td>POP</td>
<td>0.3455495</td>
<td>0.0093782</td>
<td>36.85</td>
<td>0.000</td>
<td>0.3271685</td>
<td>-0.107136</td>
<td>0.3639305</td>
</tr>
<tr>
<td>INC</td>
<td>-0.0964914</td>
<td>0.005431</td>
<td>-17.77</td>
<td>0.000</td>
<td>-0.107136</td>
<td>-0.0858469</td>
<td></td>
</tr>
<tr>
<td>HVM</td>
<td>0.0436752</td>
<td>0.0081121</td>
<td>5.38</td>
<td>0.000</td>
<td>0.0277758</td>
<td>0.0595746</td>
<td></td>
</tr>
<tr>
<td>BUS</td>
<td>3.184037</td>
<td>0.1788044</td>
<td>17.81</td>
<td>0.000</td>
<td>2.833587</td>
<td>3.534487</td>
<td></td>
</tr>
<tr>
<td>MET</td>
<td>-1561.609</td>
<td>822.8972</td>
<td>-1.90</td>
<td>0.058</td>
<td>-3174.458</td>
<td>51.23998</td>
<td></td>
</tr>
<tr>
<td>CONSTANT</td>
<td>7892.042</td>
<td>1151.693</td>
<td>6.85</td>
<td>0.000</td>
<td>5634.766</td>
<td>10149.32</td>
<td></td>
</tr>
<tr>
<td>N</td>
<td>1736</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>R²</td>
<td>0.7895</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Income Equation

The results of the median income equation for the Northeast region using 3SLS are presented in Table 4. The median income equation is estimated against the endogenous variables population, employment, and residential real estate investment as measured by median house values and control variables are included to measure economic effects. The overall fit (R²) of the empirical population equation is 33 percent.

Population is positively and significantly related to median income. This implies that a large population represents a large potential supply of labor, leading to increased output and median income. A significant and negative relationship between median income and employment indicates that median income increases as employment decreases. This may be because decreases in employment generally occur at the lower end of the income scale, thus increasing the median income when lower wage workers are dropped from the employment rolls.

The relationship between residential real estate and median income is positive and significant. As housing values increase, the median income of the residents of those houses should also increase in order to maintain equilibrium rents. Poverty is positively and significantly related to median income, which implies that rich counties have a lot of poor people living in them. Education is
also positive and significant, displaying the returns to education. Population under 18 and population over 65 are significant and negative, as these segments of the population are generally not in the workforce and so do not earn a large income. Metropolitan counties are negative and statistically significant, showing counties with more residents have a lower median income, as there are more low wage jobs than high wage jobs.

**Table 4: Three Staged Least Square Results of Income Equation**

<table>
<thead>
<tr>
<th>Variable</th>
<th>Coefficient</th>
<th>Std. Error</th>
<th>z statistic</th>
<th>p value</th>
<th>Confidence</th>
<th>Interval</th>
</tr>
</thead>
<tbody>
<tr>
<td>POP</td>
<td>3.203113</td>
<td>.3401578</td>
<td>9.42</td>
<td>0.000</td>
<td>2.536416</td>
<td>3.86981</td>
</tr>
<tr>
<td>EMP</td>
<td>-5.495857</td>
<td>.3958025</td>
<td>-13.89</td>
<td>0.000</td>
<td>-6.271615</td>
<td>-4.720098</td>
</tr>
<tr>
<td>HVM</td>
<td>.3480396</td>
<td>.0622058</td>
<td>5.59</td>
<td>0.000</td>
<td>.2261184</td>
<td>.4699608</td>
</tr>
<tr>
<td>POV</td>
<td>1.033383</td>
<td>.2269625</td>
<td>4.55</td>
<td>0.000</td>
<td>.5885449</td>
<td>1.478222</td>
</tr>
<tr>
<td>EDU</td>
<td>.449182</td>
<td>.0541476</td>
<td>8.30</td>
<td>0.000</td>
<td>.3430547</td>
<td>.5553093</td>
</tr>
<tr>
<td>P18</td>
<td>-3.752701</td>
<td>.3705581</td>
<td>-10.13</td>
<td>0.000</td>
<td>-4.478981</td>
<td>-3.02642</td>
</tr>
<tr>
<td>P65</td>
<td>-.473765</td>
<td>.1796389</td>
<td>-2.64</td>
<td>0.008</td>
<td>-.8258509</td>
<td>-.1216792</td>
</tr>
<tr>
<td>MET</td>
<td>-15802.27</td>
<td>6157.337</td>
<td>-2.57</td>
<td>0.010</td>
<td>-27870.43</td>
<td>-3734.114</td>
</tr>
<tr>
<td>CONSTANT</td>
<td>35918.15</td>
<td>9241.413</td>
<td>3.89</td>
<td>0.000</td>
<td>17805.31</td>
<td>54030.98</td>
</tr>
</tbody>
</table>

**Residential Real Estate Equation**

The results of the residential real estate equation for the Northeast region using 3SLS are presented in Table 5. The residential real estate equation is estimated against the endogenous variables population, employment, and median income and control variables are included to measure economic effects. The overall fit ($R^2$) of the empirical population equation is 36 percent.

Population is negatively and significantly related to residential real estate investment. This implies that a large population negatively impacts residential real estate prices. A significant and positive relationship between residential real estate and both median income and employment indicates that residential real estate investment is positively affected by areas with high median incomes and employment. As employment and incomes increase, consumers invest
some of their income in housing. Residential real estate investment is negatively and significantly related to the number of banks in each county, and to the number of new building permits issued in each county. These instruments measure the availability of financing, and the construction of new residential housing. Potentially newly constructed residential real estate is not as valuable as previously constructed residential real estate, due in part to neighborhood effects. Older construction may be in established neighborhoods with amenities like schools, roads, and access that are showing up in the analysis as reduced prices on newly constructed homes. The number of vacancies and the number of households are both positively and significantly related to residential real estate investment. These instruments attempt to measure the demand for housing in each county, as each household should need a place to live, and each vacant house is a potential place for a household to live. Metropolitan counties are positive and statistically significant. Showing that there is increased residential real estate investment in counties with a greater population.

Table 5: Three Staged Least Square Results of Residential Real Estate Investment

<table>
<thead>
<tr>
<th>Variable</th>
<th>Coefficient</th>
<th>Std. Error</th>
<th>z statistic</th>
<th>p value</th>
<th>Confidence Interval</th>
<th>N</th>
<th>R²</th>
</tr>
</thead>
<tbody>
<tr>
<td>POP</td>
<td>-0.8278177</td>
<td>0.2015442</td>
<td>-4.11</td>
<td>0.000</td>
<td>-1.222837, -0.4327983</td>
<td>1736</td>
<td>0.3652</td>
</tr>
</tbody>
</table>
Conclusion

The results show that policies to increase residential real estate investment or housing values at the county level are ineffective at attracting more people to an area, but these policies may help increase residents' employment and incomes. The increase of employment and income may then attract more people to a county. Additionally, policies designed to increase an areas' population, employment opportunities, or median income also affect the value of residential real estate. This feedback effect should help reinforce the initial policies, encouraging economic growth in the county.

Further Research

This study could be substantially improved with a more complete dataset. Data limitations are responsible for the choice of only census years. In the future, as the cost of data collection falls, a larger dataset will be available for research of this type. Significant data is already becoming available for the late 2000's, but a longer time series will be helpful.

Also a spatial panel model would better explain the spillover effects of housing and employment location choices. Some cities reside in multiple counties, with suburbs and exurbs affecting the commuting behavior of residents. A spatial panel model may better explain these behaviors, and potentially lead to a differing policy analysis.
References


