Undergraduate Economic Review

Volume 1 | Issue 1 Article 9

2005

Real Estate Prices: City Premiums and Neighborhood Effects

Derek Hagen
Minnesota State University Moorhead

Real Estate Prices: City Premiums and Neighborhood Effects

Abstract

I study the determinants of housing prices in the Fargo, ND-Moorhead, MN metropolitan area. I find no difference in the price of houses across state line even though Minnesota has much higher income taxes. In addition, I find support for the hypothesis that the prices of neighboring houses have a significant impact on the price of a house.

Undergraduate Economic Review

A publication of Illinois Wesleyan University

Vol. I – 2004-2005

Title "Real Estate Prices: City Premiums and Neighborhood Effects"

Author Derek F. Hagen

Affiliation Minnesota State University Moorhead

Abstract I study the determinants of housing prices in the Fargo, ND-Moorhead, MN metropolitan area. I find no difference in the price of houses across state line even though Minnesota has much higher income taxes. In addition, I find support for the hypothesis that the prices of neighboring houses have a significant impact on the price of a house.

INTRODUCTION

Is it possible for two comparable houses to have different prices in two different cities? Sure, but what if the two cities are next to each other, and there are no differences in the cities other than the state to which they belong? Realtors when pricing a house often use the price of comparable houses in the neighborhood as a yardstick. Is there a neighborhood price effect? In this paper, I attempt to answer these questions.

Fargo, North Dakota and Moorhead, Minnesota are next to each other separated by a river which is the state line. There is nothing else that differentiates the two cities geographically or topographically. While the law-of-one-price suggests that two comparable houses would have the same price in Fargo and in Moorhead, there is significant anecdotal evidence suggesting that there is premium for houses in Fargo. First, Fargo's population is much larger than Moorhead's and second people often comment how one could expect to pay more for houses in Fargo because of lower income taxes. Therefore, the first hypothesis to be tested in this paper is that there is a premium for houses in Fargo.

It is commonly known throughout the real estate community that prices of other houses in the neighborhood influence house values. Many people believe their neighbor's houses matter when they price their own homes. Therefore, the second hypothesis is that the price of houses is directly related to the prices of neighboring houses.

LITERATURE REVIEW

Each of the following papers studies the prices of houses in different areas across the United States. Do and Grudnitski (1998) analyze the effect golf courses have on house values in San Diego. Benson, Hansen, Schwartz, and Smersh (1998) study the effect of views on house prices in Washington State. Anderson (2000) attempts to find the premium open space has on houses in a suburb of Minneapolis, Minnesota. Finally, Dubin studies the effect of neighborhoods in Baltimore. I will further explain each of these studies below.

A. Quang Do and Gary Grudnitski (1995) analyze the effect of having a house placed directly on a golf course. Their study was done using data from a suburb of San Diego, CA. They choose certain houses located directly on golf courses, and near golf courses (properties that are across the street perhaps, but do not border the golf course). Then, to control for the effects of any other location specific factors, they choose certain houses from areas not near golf courses. Using these houses they arranged a grid to collect their data. The features they include in their regressions were age (years), number of bathrooms, number of bedrooms, total square footage of the house, number of fireplaces, lot size (square feet), time on the market (days), a dummy variable indicating if the property has a tin roof, a dummy variable indicating if the property has central air conditioning, and finally a dummy variable indicating if the house is on a golf course. They found that property on golf courses are about 7.6% more valuable.

Earl Benson, Julia Hansen, Arthur Schwartz, Jr., and Greg Smersh (1998) study the effects of a view on residential property values in Bellingham, WA. Typically, if a view variable was used, there were two options, view or no view. In this study they include several different types of views:

1) Full ocean view,

2) Superior partial ocean view (some obstruction),

3) Good partial ocean view (significant obstruction),

4) Poor partial ocean view (some water could be seen),

5) Mountain view (unobstructed view of snow-covered mountains),

6) Lake view (property has a view of a lake, but is not located on a lake), and

7) Lakefront property (a property located on a lake, where the benefits come, not only from the view of the lake, but also from the recreational uses of being on a lake).

They find that a full ocean view adds 59% when compared to similar houses with no view. The premiums for superior, good, and poor views were 31%, 29%, and 8%, respectively. A lakefront home adds 127% to the value, meaning a $200,000 home would sell for $317,600 if placed on a lake.

Soren Anderson (2000) studies the effect any open space has on the price of houses in a Minneapolis, MN suburb. The effect of a park, golf course, lake, etc., is not specifically analyzed, since all of them were counted as open space. He found that open space adds $43,000 to the price.

Robin Dubin (1998) analyzes real estate prices in Baltimore, MD using the maximum likelihood method (ML). Dubin tries to show how ML is better than Ordinary Least Squares (OLS) when analyzing at the effect of neighborhoods. Knowing both that location matters in real estate pricing and that OLS has downfalls when it comes to accounting for neighborhoods, ML was used. The variables used are the number of rooms, number of bathrooms, number of stories, age (years), age squared, lot size (100 square feet), interior living area (100 square feet), month the house was sold (in 1978), x-coordinate on the Maryland grid, y-coordinate on the Maryland grid, dummies indicating detachment, patio, fireplace, central air conditioning, basement, if located in Baltimore county, one car garage, two car garage, and three car garage. He includes some interaction variables as well; the number of rooms times the number of bathrooms, and the detached dummy times the number of stories. Robin notices that correlations exist between the price of houses and the prices of neighboring houses. The paper shows how the correlations can be used when estimating the regression coefficients and predicting house prices. These data are then used to get results using OLS, OLS with Trend, and ML. Dubin (1998) finds that the ML method produced better results than OLS and OLS with trend.

While there are many twin cities separated by a state line in the United States, there are no studies analyzing their impact on housing prices. This is my main contribution to the literature. Second, I use a more direct measurement of the impact of neighboring house prices than Dubin (1998).

THEORY & MODEL

As stated above, the two working hypotheses are that the market price of a house with the same characteristics is higher in Fargo than Moorhead and that the price of a house is higher in neighborhoods where other houses sell for higher prices.

To test these hypotheses, a hedonic model is used. A hedonic model is the valuation of specific features of a product that is typically purchased as a single unit. Using OLS with market price as the dependent variable, a dummy variable is used to represent location. Since it is very difficult to quantifiably measure the neighborhood effect[1], a proxy is needed. As such, median income is used as the proxy for neighboring prices. The assumption is that houses will sell for higher prices in neighborhoods where more people with higher income live and vice versa. To control for other characteristics of the house, a number of variables were included. Table 1 lists the variables and their expected signs. These variables were chosen because these are features that buyers look for when buying a house, as well as what sellers advertise when selling a house. With the exception of the floor and age variables, I expect all the signs of the coefficients to be positive because they are desirable features for buyers. I expect the floor coefficient to be negative because adding an additional floor while keeping square footage constant reduces the area of each floor. I expect the coefficient for age to be negative because as the house gets older, it becomes worn out and starts to require repairs and such. The location variable is

1. Ideally an average price for the houses located in the same block would be used. But Multiple Listing Services do no provide such data.

Table 1: Variables and expected signs

Variable	Definition (measurement)	Expected Sign
P	Price (Dollars)	N/A
BED	Bedrooms (#)	(+)
BATH	Bathrooms (#)	(+)
LIV	Living Rooms (#)	(+)
DIN	Dummy: 1 if formal dining room 0 otherwise	(+)
AGE	Age (years)	(-)
GAR	Garages (#)	(+)
ATTGAR	Dummy: 1 if garage is attached 0 otherwise	(+)
FL	Floors (#)	(-)
FP	Fireplaces (#)	(+)
DECK	Dummy: 1 if deck is present 0 otherwise	(+)
SQFT	Square Feet (sq. ft)	(+)
FARGO	Dummy: 1 if in Fargo 0 otherwise	(+)
NEIGHBOR	Median Household Income (Dollars)	(+)

a dummy variable with a value of 1 if the house is in North Dakota (Fargo) and 0 otherwise. This is expected to be positive, reflecting my hypothesis that there is a premium to live in Fargo. The neighborhood variable is also expected to be positive reflecting my second hypothesis that neighboring house prices affect the house in question.

DATA

The data for this study were provided by a local real estate agent in Fargo. The information detailed all houses sold in the Fargo-Moorhead area between 1 February 2004 and 31 July 2004. Two different samples were used in this study. All houses that were sold in Fargo and Moorhead were used for the first

regression, with the exception of the few that had incomplete information. For the second regression, the sample had to be reduced due to the fact that not all the addresses were available and they were needed to determine the neighborhood effect. There were also observations from West Fargo, North Dakota and Dilworth, Minnesota in the second sample, so the location variable in the second regression refers to whether the house is in North Dakota (Fargo or West Fargo). Descriptive statistics pertaining to the first data set are given in Table 2.

Table 2: Descriptive Statistics for FM Housing Data 1

Sample 1	Fargo	Moorhead	All
Mean	$143,218.44	$132,929.27	$140,163.80
Median	$124,000.00	$119,900.00	$122,900.00
Range[2]	$409,900.00	$377,500.00	$426,900.00
Observations	495	209	704

There is much variability in the prices of homes, as shown by the range. Table 1 shows the volatility in the 704 observations used for the first regression.

Table 3 presents descriptive statistics for the 204 observations used for the second regression. Note that the median and the mean prices are higher in Fargo than in Moorhead providing some support to the working hypothesis.

Table 3: Descriptive Statistics for FM Housing Data 2

Sample 2	North Dakota	Minnesota	All
Mean	$146,132.59	$128,624.51	$141,240.62
Median	$121,000.00	$117,000.00	$121,000.00
Range	$437,000.00	$297,495.00	$450,000.00
Observations	147	57	204

2. The range in this study is the highest price minus the lowest price.

RESULTS

The first regression includes all the variables mentioned earlier except for for the neighborhood effect variable. The results of this regression are reported in Table 4. The number of bedrooms, existence of a formal dining room, age of the house, whether the garage(s) are attached or not, and whether the house has a deck or patio are some of the variables that are not significant. I experimented with the functional form on the age variable, thinking it may be quadratic rather than linear but the results were the same. The variables that were significant at the 95% level are in bold. The insignificance of the attached garage variable is very surprising because I expected a homebuyer would be willing to pay a premium to not have to go outside on cold winter mornings. The fact that it get so cold in the area might explain why the fireplace variable was so high and statistically significant. The regression shows that each additional fireplace adds about $14,000.00 (almost 10% of the mean) to the price of the house. Other surprising results are that the number of

bedrooms and the age of the house are statistically insignificant. The F-statistic is 178.1344, much higher than the critical value of F. The Adjusted R-Squared is 0.751468, so the equation explains approximately 75% of the variation in price.

The location variable turned out to be insignificant. The insignificance of this variable means that whether the house is in Fargo or Moorhead does not matter to homebuyers. These results did not support my hypothesis that there would be a premium to purchase a house in Fargo. The Law-of-One-Price holds and the differences in income and property taxes seem to offset each other.

Table 4: Regression 1 Results

Observations: 704			
Variable	Coefficient	t-Statistic	VIF
C	-$5,323.62	-0.6640	
BED	$2,336.21	1.3397	1.8
BATH	$17,956.75	**7.6680**	2.0
LIV	$11,100.55	**4.8205**	1.5
DIN	-$1,045.28	-0.3720	1.2
AGE	-$126.00	-1.8453	3.0
GAR	$17,586.25	**8.2469**	2.0
ATTGAR	-$160.79	-0.0426	2.5
FL	-$11,744.24	**-6.1966**	1.6
FP	$13,944.40	**6.0580**	1.4
DECK	-$1,366.10	-0.5473	1.2
SQFT	$42.00	**13.1613**	3.2
FARGO	$3,330.44	1.2678	1.1
R-Squared	0.755711		
Adjusted R-Squared	0.751468		
Durbin-Watson Stat	1.191131		
F-Statistic	178.1334		

People who choose to live in Fargo pay high property taxes, but low income taxes, with the opposite being true for people in Moorhead.

I suspect multicolinearity may be present in this first regression since some of the variables appear to measure the same thing, such as square footage and a number of other variables all measuring the size of the house. I tested for this using Variance Inflation Factor (VIF) tests. The results for these tests are in the right-most column of the Regression 1 Results table. The results do not support the suspicion that multicolinearity exists, since none of the VIFs are above 5.[3] The highest VIF result is SQFT as I thought it would be, at 3.2.

3. Five is the arbitrary number assigned to these tests to reveal multicolinearity. See Studenmund, p. 258.

The second regression includes all the variables from the first regression plus the neighborhood effect variable. The results of the second regression are reported in Table 5. Also, the smaller sample size was used in this regression because of the unavailable addresses. Again, the variables that are statistically significant are in bold. The neighborhood effect variable is significant and positive which supports my hypothesis. The coefficient of the neighborhood variable is 0.71, which is interpreted as an increase of one dollar in the average income a neighborhood increases the selling price of a home by $0.71. If there are two identical houses, one in a neighborhood with a median income of $50,000 and the other in a neighborhood with a median income of $100,000, we expect the house in the wealthier neighborhood to sell for $35,500 more.

Table 5: Regression 2 Results

Observations: 204			
Variable	Coefficient	t-Statistic	VIF
C	-$49,727.85	-3.0077	
BED	$1,198.29	0.3377	1.8
BATH	**$25,168.42**	**4.2213**	3.0
LIV	$3,083.99	0.6212	1.6
DIN	-$3,091.90	-0.5186	1.3
AGE	$88.97	0.6365	2.6
GAR	**$15,915.62**	**3.8928**	1.8
ATTGAR	-$4,608.09	-0.5842	2.5
FL	-$10,609.23	**-2.4865**	1.8
FP	**$20,213.50**	**4.5403**	1.6
DECK	$475.82	0.0875	1.2
SQFT	**$47.19**	**6.2671**	4.1
FARGO	$4,831.93	0.8437	1.1
NEIGHBOR	**$0.71**	**4.07516**	1.6
R-Squared	0.783304		
Adjusted R-Squared	0.768477		
Durbin-Watson Stat	1.723955		
F-Statistic	52.83098		

The t-Statistics of all the variables that are significant in regression 2 fell relative to regression 1. Adding a variable to control for the neighborhood effect took away from the explanatory ability of all the other variables. The coefficient for number of living rooms went from being very large and significant to being very small and insignificant. The fireplace coefficient went up to more than $20,000. The F-Statistic fell but remains relevant and the R-Squared and Adjusted R-Squared both rose, meaning the new regression explains more of the variability in the prices. I used VIF tests again (in the right column) to test for multicolinearity and the results show it is not serious enough for concern.

POTENTIAL WEAKNESSES

It is very difficult to pinpoint a single selling price when talking about an investment such as real estate. The decision to buy a house means giving up a substantial sum of money. As such, buyers and sellers are very careful about how much they are willing to pay, as well as how much they are willing to accept. Buyers want to get the lowest price possible, while sellers want to get the highest price possible. The actual selling price of the house depends on a number of things that we are not able to measure in a model, such as the negotiating skills of each party, the need for cash by the seller, the desire for the property by the buyer, current interest rates on mortgages, and the many other factors that arise in real estate deals (Dubin, 1998). Hence, the same house could sell for different prices as these factors change, although nothing about the house is changing. Such factors are nearly impossible to take into account in a model like this.

The proxy used to capture the neighborhood effect may be flawed. The problem with using pre-set neighborhood boundaries (like census data on median income by census tract) is that the neighborhoods are pre-determined. People may consider their neighborhood to be the land that is, for instance, within a one mile radius of their house. In which case, there are not set neighborhoods because neighborhoods change from different perspectives (Dubin, 1998).

I used linear form for all variables used. The potential problem with using the linear form in OLS is that the interpretations of the variables are not allowed to reflect other variables. For instance, the value of an additional bathroom does not take into consideration the number of bedrooms or the square footage of the house (Dubin, 1998). The use of interaction variables may help solve this problem. However, they remain the most common functional form because of the ease of interpretation.

Although I tested for multicolinearity in both regressions, it may still exist. Some of the variables may be related to each other, which would make one or the other insignificant. For instance, the age of the house may very well be related to the neighborhood the house is in or to the number of garages. Square feet measures size directly, but number of bedrooms, bathrooms, family rooms, and floors also measure size indirectly. It might be worthwhile to take a look at these seemingly related variables and see if and how the results change.

CONCLUSION

As hypothesized, there is a neighborhood effect when determining real estate prices in Fargo-Moorhead. It does matter what other houses in the neighborhood are selling for when pricing a house. If someone builds a million-dollar house in the neighborhood, expect the value of surrounding houses to increase. If people do not take care of their houses and the property does not look presentable, expect neighboring houses to decrease in value.

In the Fargo-Moorhead region, there does not seem to be a premium to live in one city over another, as some people seem to suggest. Location matters to people when buying a home but only on a micro-level. The neighborhood matters, while the city does not.

References

Anderson, Soren. "The Effect of Open Space on Single-Family, Residential Home Property Values." Macalester College, 2000.

Benson, Earl D., Hansen, Julia L., Schwartz, Jr., Arthur L., and Smersh, Greg T. "Pricing Residential Amenities: The Value of a View." Journal of Real Estate Finance and Economics 16, no. 1 (1998): 55-73.

Do, A. Quang and Gary Grudnitski. "Golf Courses and Residential House Prices: An Empirical Examination." Journal of Real Estate Finance and Economics 10, no. 3 (1995): 261-270.

Dubin, Robin A. "Predicting House Prices Using Multiple Listings Data." Journal of Real Estate Finance and Economics 17, no. 1 (1998): 35-59.

Studenmund, A. H. Using Econometrics: A Practical Guide. 4th ed. New York: Addison Wesley Longman, 2001.