Economic Studies at BROOKINGS

July 2015

# Are New Homes Special?

**N. EDWARD COULSON**
University of Nevada, Las Vegas

**ADELE C. MORRIS**
The Brookings Institution

**HELEN R. NEILL**
University of Nevada, Las Vegas

## Abstract

We attempt to disentangle the meanings of "new" in hedonic models of housing markets. "New" typically refers to the condition that the year of sale is equal to the year of construction, while the appropriate meaning ought to entail the unit's entry into the market, which is not the same thing. Some houses that are sold before, or some extended period of time after, completion are in effect, new homes, while not all homes of age zero are unused. We disentangle these definitions in a hedonic study of Las Vegas, where the spatial, and especially temporal, value of "newness" turns out to have substantial variation over the housing cycle.

The authors thank the Clark County Assessor's Office for providing the real estate data. Partial funding for this research was provided by Brookings Mountain West. The authors wish to thank Francois Des Rosier and participants at both the 2015 AREUEA National Meeting and the 2015 ARES Annual Meeting for helpful suggestions. The paper reflects only the views of the authors and not their respective employers, the Brookings Institution and the University of Nevada, Las Vegas.

## I. Introduction

Houses are idiosyncratic goods; no two houses are completely identical because the set of characteristics that are embodied in any particular unit are unique to that unit. The hedonic model of individual house prices maps the characteristics of houses to their transaction prices through a regression of sales price (or more often, its logarithm) on the vector of observable housing characteristics. In this way the implicit prices of those embodied characteristics can be estimated as the derivative of price on that characteristic (Rosen, 1974).

One of those characteristics is the age of the dwelling. In their survey of the empirical literature on hedonic pricing, Sirmans, MacPherson and Zietz (2005) note that, among the 125 studies they survey, dwelling age was included 78 times, more often than any other housing characteristic with the exception of interior square footage. In some ways this is puzzling, since age in and of itself is not an indicator of the utility flow from a dwelling, or of the cost of habitation (Rubin, 1993). While age is certainly correlated with increased maintenance costs or outmoded style characteristics, direct measurement of those features would more closely measure the utility or cost characteristics that really matter. Nevertheless, dwelling age has demonstrated itself to be a useful proxy for both physical and stylistic deterioration.

In a regression of log(sales price) on age, entering age in linear form maintains the assumption of a constant depreciation rate over the life of the property (and no vintage effects). But as an empirical matter the relationship between age and property value is seldom so simple; non(log)linear functional forms are usually observed. One reason is that the utility of ownership of a very old property may increase as the house gets very old. The cachet from owning (say) a 100 year old house may have so much value that the hedonic price of such a unit may turn positive. Certain style attributes, characteristic of buildings of a certain vintage, may have increased value in a market, and cause the price-age profile in the estimated hedonic function to be non-monotonic. Moreover, age may serve as a proxy for unobserved characteristics of the neighborhood, such as abundant mature trees or historic status. For this reason, flexible functional forms in building age are often thought to be beneficial in the estimation of hedonic functions. Studies that include age in higher order polynomials include Goodman and Thibodeau(1997) and Coulson and Lahr(2005). Coulson and McMillen (2008) use other nonparametric forms.

But more importantly, the physical depreciation rate may not be constant over time. Like other durable goods there may be steep depreciation at the beginning of a property's life, and then a flattening out (or other nonlinearities, as discussed above). One age of special interest is age zero, for example as illustrated in the separate reporting of new home prices and the number of newly completed homes by the U.S. Census Bureau. In a recent paper by Kahn and Kok (2014) where they estimate the impact of green labels on homes in California, they also control for alternative measures of age using vintage indicators relative to homes 50 years or older. They report evidence that age of zero increases sales price of homes by 20 to 22% and age of 1 to 2 years increases sales price of homes by 24 to 27%.

Developers and builders, the primary sellers of this class of homes, have a particular interest in the price of new versus existing homes. News stories also illustrate the interest in new homes as distinct from existing homes. In one recent article, Robison (2014) noted a truism that in the Las Vegas housing market "[i]f you want a new home, prepare to drop $100,000 more than you'd shell out for a resale." Some developers perceive the promulgation of this truism as something to be squelched, for fear of potential buyers would not search in the new home market, and on that account sponsored projects to educate buyers on the difference between unconditional and conditional price differences between new and old homes.

In this paper we closely examine, for what appears to be the first time in the hedonic pricing literature, the *new home premium*, the putative extra value placed on a home that is on the market for the first time, conditional on the embodied characteristics of a home. In coming to grips with the estimation of the new home premium we face three empirical challenges:

First, what would motivate a systematic difference in the price of new homes relative to existing homes while controlling for observable characteristics? We take the meaning of "new" to be "unused"-- that is, not previously occupied, and on that account, of recent construction vintage. This typically entails both positive and negative features of the home, which generally (following Rubin, 1993) are not otherwise captured in the characteristics vector of a hedonic regression. These features often include the ability to choose interior details of the house, such as paint color and floor coverings; newer technologies, such as smart thermostats or more sophisticated alarm systems; immature landscaping, which may be less attractive than established plantings, and so on. Thus the new home premium need not even be positive if the negative aspects of newness dominate the positive aspects.

Second, what is the definition of *new*? In a standard empirical implementation of a hedonic model there is no emphasis on "new" as such. The age of a property is usually the difference between the year of sale and the year of construction, and new homes are those with age equaling zero (and in fact not usually identified as new). This is not appropriate for our purposes. (1) above suggests that the new home premium should be identified with "unused" or "not previously occupied" rather than by the mere calculation of age. As we discuss in more detail below, we identify "new" in a couple of different ways: by noting whether the seller is a property developer; by identifying whether the parcel is appearing for the first time in the property assessor's database; by age; by methods which combine these three pieces of data. We will present results from multiple identification paths. Even then we need to confront special cases: if new construction sits vacant for two years before being sold to its first occupant, is it still new? And so on.

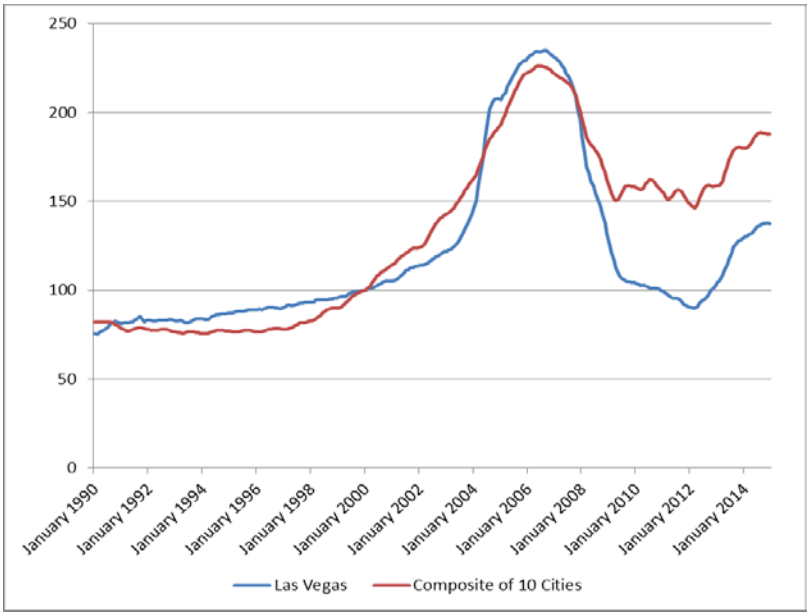
Finally, is the premium constant across time, place and property type? There is no reason to believe so. The supply and demand for new homes depends on the particular circumstances. When the number of new properties is small, the premium will be larger, other things equal. There can be certain neighborhoods which because of development restrictions have a small number of new properties, but are highly desirable. The premium might be very large in such circumstances. But it might also vary over time, as credit conditions restrict developers' ability to borrow against new home sales, and the premium might rise.

We use the Clark County, i.e. Las Vegas, housing market over the time period 1995 to 2014 as a testing ground for estimating the premium (and its variation). This market is ideal for such a study for a couple of reasons. First, and most obviously, the Las Vegas market exhibited wide price swings and large movements in the liquidity of the housing stock.

To illustrate the broad market context of our study, Figure 1 shows the S&P/Case-Shiller Home Sales Price Indices for Las Vegas and the broader Case-Shiller composite index for ten large U.S. cities.[1] The boom was steeper and the freefall in 2008 was more dramatic in Las Vegas than in most other cities.

[1]Data for this figure are the S&P/Case-Shiller Seasonally Adjusted Home Price Index Levels, downloaded April 7, 2015, from http://www.spindices.com/index-family/real-estate/sp-case-shiller.

**Figure 1. S&P/Case-Shiller Home Price Indices, Jan 1990 - Dec 2014**




The boom in home prices was accompanied by dramatic growth in population during the sample period. The population of Clark County grew dramatically over the duration of the data, with an average annual growth rate of 5 percent from 1990 to 2009. The county population more than doubled from about 986,152 in 1994 to more than 2,062,253 in 2013.[2] Second, there was a substantial variation in the amount of new construction over this time period, and across the various Las Vegas submarkets. In many Sunbelt markets, both natural and legal restrictions on development are binding and the amount of new construction minimal. This is not the case in the Las Vegas housing market. Third, we have a very large sample. As noted, we have a long span of data, and almost 800,000 property transactions. The large sample permits us to estimate the variation in the premium in multiple ways. Moreover, we can estimate the conditional relationship between age, newness, and price flexibly.

## II. Methodology

As noted in the introduction, we model houses prices as function of the embodied locational and structural characteristics. The literature notes that theory does not provide a guide to functional form

We model the inflation-adjusted selling price of a house as a function of its physical characteristics and other factors.

The typical log-log hedonic estimation equation for $y_{ist}$, the log price of house $i$ in census tract $s$ in the period of sale $t$, is:

1) $y_{ist} = c + \alpha_s + \gamma_t + \beta X_i + n_{ts}\alpha_s + \epsilon_{ist}$.

The variable c is a constant. The terms $\alpha_s$ are indicators for the census tracts in which the properties lie. Alone, they control for time-invariant characteristics of neighborhoods. Most census tracts are small, so in general the tract indicators are able to control for location; the median 2000 census tract in our data is only 1.5 square miles in area.

The terms $\gamma_t$ are indicators for the year of sale, and they control for broad trends in the Clark County and broader housing market.[3] The vector $X_i$ includes property characteristics that are standard in hedonic models, with coefficients $\beta$. The terms $n_{ts}\alpha_s$ control for interaction effects with time and location attributes. The variable $\epsilon_{ist}$ is an error term that reflects random variation in house prices.[4]

## III. Data

[2] See http://www.clarkcountynv.gov/depts/comprehensive_planning/demographics/pages/default.aspx last accessed on April 14, 2015.

[3] Including the interaction of census tract and year indicators is infeasible given the large number of years (35) and tracts (334) in our data.

[4] We estimate robust standard errors, and we cluster the errors by tract-year. See Nichols and Schaffer (2007) and Morris and Neill (2014) for similar approaches.

Data for this study comes from multiple sources. We obtain real estate property data through the Clark County Assessor's Office and consumer price index (CPI) from the U.S. Bureau of Labor Statistics. Census tract information for 2000 censuses were obtained using ArcMap program and shape files provided by the Clark County Office Geographical Information system Management Office. We combined data using parcel number, sale date, and geocoded information.

We use data from about 785,022 arms' length home sales in Clark County, Nevada, from January 1994 to December 2014, inclusive, obtained from the Clark County Assessor's Office.

Each unit of observation in the data is the sale of a residential property. The data include the actual selling price of the property, the sale date, and detailed characteristics of the home. As shown in Table 1, the property characteristics include lot size, square footage of living area, number of full baths, the age of the home at the time of sale, and indicators for amenities such as a pool.

Consistent with Clauretie and Daneshvary (2009) and Campbell et al (2011), we include an indicator for transactions the Clark County Assessor designates as linked to a foreclosure. This accounts for the well-established discounts associated with such forced sales.

We dropped sales before 1994 and after 2014 and removed outliers from the data. For example, given the specialized market for such properties, we exclude very large homes (greater than 8,000 sq. ft. of living area) and homes on very large lots (greater than five acres). We also drop sales of properties with prices under $40,000 or over $5 million ($2010).[5]

**Table 1. Summary Statistics**

Number of Observations = 785,022

| | Mean | Standard Deviation | Min | Max |
|---|---|---|---|---|
| Property price ($) | 250,193 | 156,469 | 40,001 | 2,300,612 |
| Living space (ft$^2$) | 2011 | 789 | 280 | 7,980 |
| Lot size (acres) | .16 | .14 | .01 | 5 |
| Developer | .22 | 0.42 | 0 | 1 |
| Pool indicator | .21 | .41 | 0 | 1 |
| Full bathrooms | 2.24 | .57 | 1 | 6 |
| Foreclosure indicator | .11 | .32 | 0 | 1 |

**Notes**: All dollar values appear in 2010 dollars, deflated using the Consumer Price Index for All Urban Consumers. The data set combines information from the Clark County Assessor's Office and the U.S. Census.

**Table 2. Age Indicator Variables**

Number of Observations = 785,022

| Variable | Mean | Std.Dev. | Min | Max |
|---|---|---|---|---|
| Home age (years) | 11.16 | 13.40 | 0 | 113 |
| New construction | 0.26 | 0.44 | 0 | 1 |
| Presold | 0.06 | 0.24 | 0 | 1 |
| Age0 | 0.20 | 0.40 | 0 | 1 |
| Inventory | 0.06 | 0.24 | 0 | 1 |
| New-construction developer | 0.21 | 0.41 | 0 | 1 |
| All new developer | 0.22 | 0.42 | 0 | 1 |
| Presold developer | 0.05 | 0.22 | 0 | 1 |
| Age0 developer | 0.16 | 0.36 | 0 | 1 |
| Inventory developer | 0.01 | 0.12 | 0 | 1 |

[5] This approach is similar to that in Morris and Neill (2014).

These dropped observations collectively represent 9 percent of the raw dataset.

Table 2 reports the variables we have constructed for age and newness. The variable age takes on the values of 0 to 113, calculated as the difference between the year the house was sold and when it was built, with the exception that it is equal to zero when the difference is negative

We turn to definitions of newness. As discussed above, newness has two dimensions. One is that of age. A new home, by this definition, was built not very long before its sale. The other dimension is freshness. By this we mean that the property has not been previously occupied, and may be customizable by the buyer. This "unused" quality may have great value to buyers. Previous hedonic studies have not made the distinction clear. Most studies (as noted) simply define new in terms whether the home is of age zero (or negative). To the extent that the regression allows for nonlinearity, it may pick up the premium that attaches to new homes. Other studies include an indicator variable for new construction, but do not distinguish between different ages of that construction.

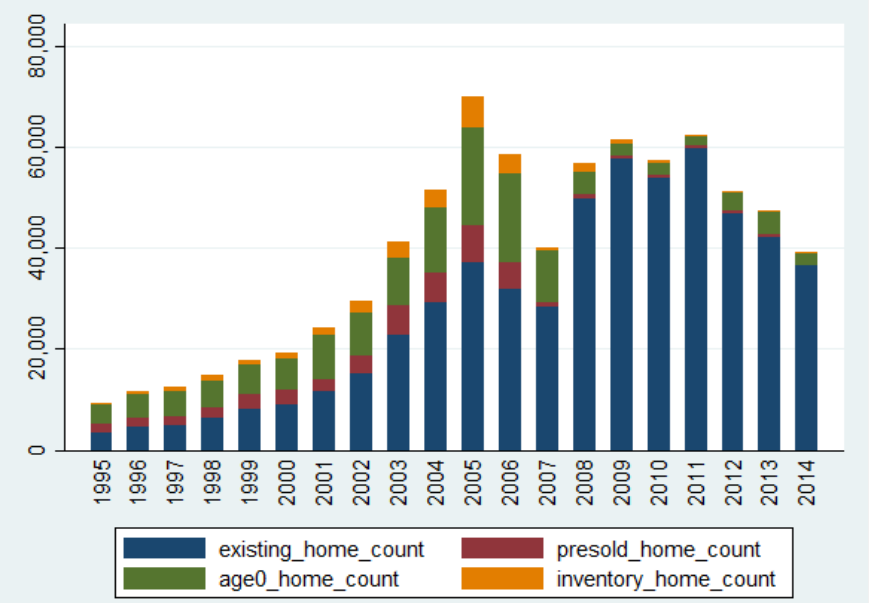
To that end, we propose the following variables. We define *presold* to account for homes that are built after sale. We define *age0* as homes that are built in the same year as sold and *inventory* as sales in the first two years after the house was built. The *new construction* indicator identifies homes that meet the criteria of either of the first two categories. In each case the seller might be the original developer or some other party.

The distinction is important, and we define three mutually-exclusive identifiers for sales by developers specifically to indicate categories of homes that have never been unoccupied. The *presold developer* indicator is 1 if the home meets all of three following conditions, and zero otherwise. First it was sold within the two years *prior* to the year it was built (for example as a pre-sale in a new development). Second, the sale must be the first time that property identifier appears in our dataset, and third, the property must have been sold by a developer, as indicated by whether the seller is not a bank and appears more than ten times in our overall dataset. The *presold developer* indicator captures the instance when a new home may be customized by its buyer, thus potentially adding an extra amenity to the new home sale. On the other hand, these pre-built sales may carry risks for buyers as they may lose money if the developer fails to complete the home, substantial delays occur, or the broader neighborhood development stalls. Thus we cannot predict whether this variable is more likely to be positive or negative.

The *age0 developer* indicator is one if the sale meets three similar conditions, and is zero otherwise. First the sale occurs *the same year* the home was built, as recorded in the county database; the sale is by a developer; and the sale is the first appearance of that property number in the database. This captures the case when a home may be built primarily "on spec," or speculatively, and is sold in short order. Many design decisions have already been made by the builder, and the homes could be sold in move-in condition.

Finally, the *inventory developer* indicator is one if the sale occurs one or two years *after* the year in which the home was built, as recorded in the county database; the sale is by a developer; and the sale is the first appearance of that property number in the database. This captures the case when a home is fully built on spec. All design decisions are made by the builder and the homes are likely to be in move-in condition.

**Figure 2. Number of observations by year of sale, existing, presold, age0, and inventory sales**




The above three variables (the *presold*, *age0*, and *inventory developer* indicators) are mutually exclusive values. We also define an indicator variable called *all new developer* that is 1 if the home meets the criteria of any of the three specific kinds of new homes. Finally, we also define an indicator variable called *new-construction developer* that includes sales that are characterized by *presold developer* and *age0 developer*.

We may mischaracterize some sales because we know the month of sale but not the month of completion (just the year the home was built). For example, a house can be sold in January that was completed the month before in December. We would characterize such homes categorized as inventory sales, while *age0* would perhaps be more apt. Similarly, a house built in January and sold in December of the same year might properly belong in the category of presold, but here we would label it *age0*. We conduct some limited robustness tests for January (and perhaps February sales). For our main results we adhere to the definitions above.

Figure 2 shows the distribution of the remaining observations by year of sale from 1995 through 2014. The figure shows sales of existing homes in blue. New home sales (both developer and non-developer) are identified as presold in red, age0 in green, and inventory in yellow. The observations are heavily weighted towards the past two decades due to the dramatic increase in homes built and sold from the mid-1990s through 2006, as evidenced by the heights of the bars.

**Table 3. Bivariate Models with Age Indicators; Dependent Variable: ln(property sale price)**

| | I | II | III | IV | V | VI | VII | VIII | IX |
|---|---|---|---|---|---|---|---|---|---|
| | b/se | b/se | b/se | b/se | b/se | b/se | b/se | b/se | b/se |
| newconstruct_dev | 0.379*** | | | | 0.387*** | | | | |
| | (0.017) | | | | (0.017) | | | | |
| allnew_dev | | 0.389*** | | | | | | | |
| | | (0.017) | | | | | | | |
| new_construction | | | 0.384*** | | | | | | |
| | | | (0.018) | | | | | | |
| presold_dev | | | | 0.383*** | | 0.312*** | | | |
| | | | | (0.021) | | (0.018) | | | |
| age0_dev | | | | 0.388*** | | | 0.358*** | | 0.392*** |
| | | | | (0.017) | | | (0.016) | | (0.018) |
| inventory_dev | | | | 0.421*** | | | | 0.340*** | 0.425*** |
| | | | | (0.017) | | | | (0.016) | (0.018) |
| inventory_dev | | | | | 0.421*** | | | | |
| | | | | | (0.017) | | | | |
| presold | | | | | | | | | 0.353*** |
| | | | | | | | | | (0.021) |
| _cons | 12.186*** | 12.178*** | 12.164*** | 12.178*** | 12.178*** | 12.249*** | 12.208*** | 12.259*** | 12.174*** |
| | (0.011) | (0.011) | (0.011) | (0.011) | (0.011) | (0.014) | (0.012) | (0.014) | (0.011) |
| N | 785022 | 785022 | 785022 | 785022 | 785022 | 785022 | 785022 | 785022 | 785022 |
| r2 | 0.07 | 0.077 | 0.084 | 0.077 | 0.077 | 0.014 | 0.05 | 0.005 | 0.079 |

**Notes**: All Asterisks ***, **, and * indicate statistical significance at the 1%, 5%, and 10% levels, respectively. Dollar values are in real $2010. Standard errors are clustered by tract-year, with 9758 distinct groups.

## IV. Results

### *Hedonic Model with Age and Newly Built Indicators*

We begin Table 3 with bivariate regressions of (log) price on the various indicators discussed above. Table 3 presents bivariate regressions of log (property sale price) on the various newness indicators. As might be expected, there is a very wide unconditional price gap of between 17 and 52 percent between new and used homes regardless of the definition of newness used.[6]

[6] Percentage differences throughout use exp(β)- 1 to calculate the effect of binary variables. Also note that age and its square are included in the regression. What the various newness indicators measure is the deviation of "new" homes from the "ordinary" path of prices as buildings age.

Interestingly, inventoried, homes, those homes that have time on the market long enough such that the year of sale is greater than the year of completion, have the highest percentage difference, more than either of the other two categories of developer sales, as exhibited in Column IV. Columns V through IX report estimates for alternative measures of new developer by themselves and as a group.

In Table 4, we add to these specifications the vector of hedonic characteristics as discussed above and

**Table 4. Main Models with Age Indicators; Dependent Variable: ln(property sale price)**

| | I | II | III | IV | V | VI | VII | VIII | IX |
|---|---|---|---|---|---|---|---|---|---|
| | b/se | b/se | b/se | b/se | b/se | b/se | b/se | b/se | b/se |
| ln_living | 0.600*** | 0.599*** | 0.599*** | 0.599*** | 0.599*** | 0.599*** | 0.599*** | 0.599*** | 0.599*** |
| | (0.006) | (0.006) | (0.006) | (0.006) | (0.006) | (0.006) | (0.006) | (0.006) | (0.006) |
| ln_acres | 0.188*** | 0.188*** | 0.189*** | 0.188*** | 0.188*** | 0.188*** | 0.188*** | 0.188*** | 0.189*** |
| | (0.005) | (0.005) | (0.005) | (0.005) | (0.005) | (0.005) | (0.005) | (0.005) | (0.005) |
| pool | 0.072*** | 0.072*** | 0.072*** | 0.072*** | 0.072*** | 0.072*** | 0.072*** | 0.072*** | 0.072*** |
| | (0.001) | (0.001) | (0.001) | (0.001) | (0.001) | (0.001) | (0.001) | (0.001) | (0.001) |
| bath_full | 0.026*** | 0.027*** | 0.026*** | 0.026*** | 0.026*** | 0.027*** | 0.026*** | 0.027*** | 0.027*** |
| | (0.002) | (0.002) | (0.002) | (0.002) | (0.002) | (0.002) | (0.002) | (0.002) | (0.002) |
| foreclosure | -0.135*** | -0.135*** | -0.136*** | -0.134*** | -0.134*** | -0.135*** | -0.135*** | -0.134*** | -0.135*** |
| | (0.003) | (0.003) | (0.003) | (0.003) | (0.003) | (0.003) | (0.003) | (0.003) | (0.003) |
| age | -0.008*** | -0.007*** | -0.009*** | -0.007*** | -0.007*** | -0.007*** | -0.008*** | -0.007*** | -0.009*** |
| | (0.000) | (0.000) | (0.000) | (0.000) | (0.000) | (0.001) | (0.000) | (0.001) | (0.000) |
| age2 | -0.000** | -0.000*** | 0 | -0.000*** | -0.000*** | -0.000*** | -0.000** | -0.000*** | 0 |
| | (0.000) | (0.000) | (0.000) | (0.000) | (0.000) | (0.000) | (0.000) | (0.000) | (0.000) |
| newconstruct_dev | -0.016*** | | | | -0.011* | | | | |
| | (0.004) | | | | (0.004) | | | | |
| allnew_dev | | -0.003 | | | | | | | |
| | | (0.004) | | | | | | | |
| new_construction | | | -0.031*** | | | | | | |
| | | | (0.006) | | | | | | |
| presold_dev | | | | 0.002 | | 0.006 | | | |
| | | | | (0.006) | | (0.006) | | | |
| age0_dev | | | | -0.015** | | | -0.019*** | | -0.034*** |
| | | | | (0.005) | | | (0.005) | | (0.005) |
| inventory_dev | | | | 0.081*** | 0.081*** | | | 0.086*** | 0.065*** |
| | | | | (0.006) | (0.006) | | | (0.006) | (0.006) |
| presold | | | | | | | | | -0.068*** |
| | | | | | | | | | (0.008) |
| _cons | 7.886*** | 7.879*** | 7.908*** | 7.884*** | 7.884*** | 7.877*** | 7.886*** | 7.878*** | 7.911*** |
| | (0.057) | (0.057) | (0.056) | (0.056) | (0.057) | (0.057) | (0.057) | (0.057) | (0.057) |
| N | 784488 | 784488 | 784488 | 784488 | 784488 | 784488 | 784488 | 784488 | 784488 |
| r2 | 0.814 | 0.813 | 0.814 | 0.814 | 0.814 | 0.813 | 0.814 | 0.814 | 0.814 |

**Notes**: Asterisks ***, **, and * indicate statistical significance at the 1%, 5%, and 10% levels, respectively. Dollar values are in real $2010. Standard errors are clustered by tract-year, with 9758 distinct groups

displayed in Table 1. The coefficients are all significant at any usual level of Type I error, as would be expected with our large sample. The coefficient sizes are sensible. Across the various specifications we see that a one percent increase in house size yields a .6 percent increase in price, while one percent increase in lot size yields a 0.2% increase in price. A pool adds about 7.2% and an additional bath 3% to value. A house in foreclosure lowers the sale price by about 13%. The age variable indicates a depreciation rate of about 1 percent per year, at least for younger houses.

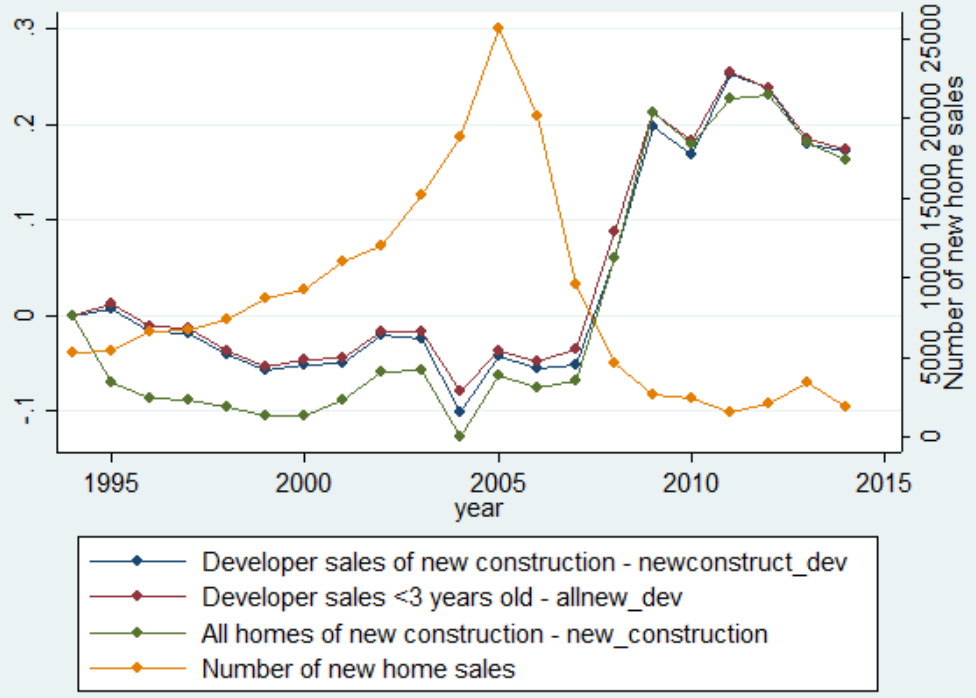
The magnitudes of the coefficients of the new house indicators are much reduced from Table 3, naturally enough. Indeed, they are negative except for that which indicates that the sale is from the inventory of the builder. This coefficient is around 8%; the others range from negative 1% to about 0 %, the latter for presold homes.

The negative signs on these coefficients are somewhat puzzling, since we might otherwise have expected that new homes would carry a premium. The small size and negative coefficients indicate that, in fact, new homes are not all that special.

Nevertheless an important methodological point is made by comparing columns I and III in the table. Column III presents a characterization of newness as measured by *new_construction* as it is usually done in hedonic studies, using those homes with an age of zero, whereas column I defines new as *new construction developer* sales. Note that the value of the latter coefficient is larger than that of the former. Merely defining new in the traditional manner includes properties that are not in fact new, and biases its coefficient downward.

Moreover, in a typical hedonic study, new homes that are one or two years old (and therefore not new) are here properly counted as new. The size of the coefficient of *inventory developer* in column IV is quite large. Therefore, to exclude "unused" home from the "new" category just because their age is not zero would seem to be erroneous. Columns V through IX provide estimates of the alternative measures of newness interacted with developer. Similar to the point made about the new construction coefficient, the presold coefficient is larger than presold developer and statistically significant at the 1 % level.

**Figure 3. Regression results from interactions between age indicators and year**



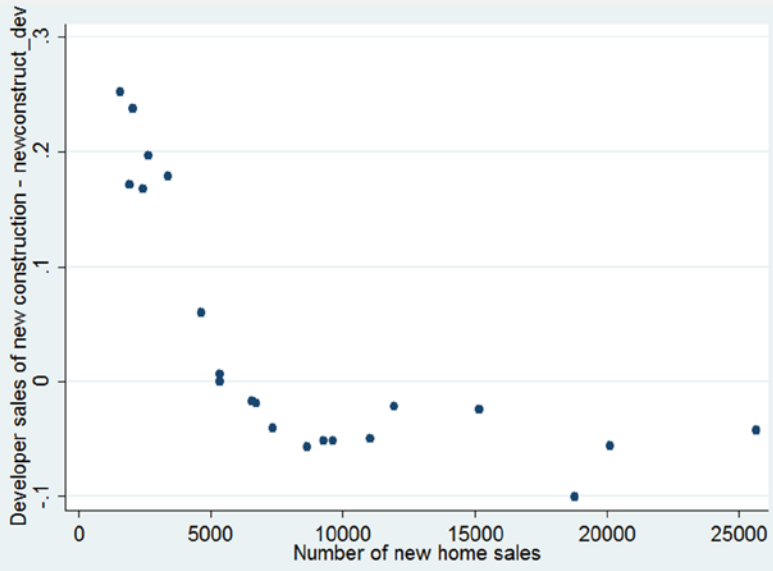
Nevertheless, the small magnitude and negative sign are unexpected and we are therefore moved to examine the data for variation in the new home premium across space and time. Focusing on time first we interact the new home measures in the previous analysis with sale year dummies. The coefficients of the background variables are much as before, and we forebear presenting them for space reasons. The key coefficients are presented in Figure 3 which tracks the coefficients of new construction, new construction developer, and all new homes developer across the sale years 1995 through 2014.

Several interesting results can be seen in Figure 3. First, in the boom years, new homes we call *new construction* from any type of seller are lower priced than homes that are sold by developers. Non-developer sellers of *age0* homes-- we will call these sales "flips", although we have not tried to characterize individual sellers in the manner of Bayer et al-- are evidently not getting the same price as developers, perhaps precisely because they are no longer thought of as new homes. Whether this is because buyers expect adverse selection, we cannot say, but it is of substantial interest that during the bust this gap narrows, and in fact during some years of the recovery these homes actually carry a premium. This gap narrows substantially during the bust and recovery.

Finally, and perhaps most robustly, the new home premium increases by a very substantial amount during the recovery period from 2009 on. All three measures show this. The reason seems clear; the yellow plot in Figure 3 tracks the number of new home sales (as measured by *age0* but any measure would have done) with the scale on the right hand axis. It can be seen that the number of new home sales plummeted between 2006 and 2008 and remained at low levels since. It was at this exact point that the new home differential began to rise from the negative range, or at best zero, to substantially positive numbers of between about 17 to 25%. Supply and demand provide the obvious explanation. The supply shift that lowered the quantity of new homes (for reasons having mostly to do with credit supplies and land use policies) drove up the price.

To illustrate demand for new construction by developers, we redraw the results illustrated in Figure 3. Figure 3A shows when there are relatively fewer new home sales, a premium exists for new homes. Figure 3A also shows that when there are more than 5,000 sales the premium disappears.

**Figure 3A. Regression results for developer sales of new construction vs. number of new homes sold in a given year**




As a final check on the distinction between different definitions of *new*, we estimate models which (a) remove the interaction terms between year of sale and newness indicator; (b) add one or more indicator variables which merely indicate whether the home was (i) 1 or 2 years old, 0 years old, or presold—whether or not it was sold by a developer; and (c) add interaction terms for both (sets of) indicator variables with an indicator for post-2007 sales. That is, we replace the simple indicator *new* (and its interactions) with:

$\beta_1$new + $\beta$2*new*post2007
+ $\beta$3*new*developer
+ $\beta_4$*new*post2007*developer

where *new* is a measure of the newness (as characterized by age) of the house regardless of the seller. The parameter $\beta_1$ represents the price impact of new, but non-developer, sales in the pre-2007 period, $\beta_1 + \beta_2$ the new home premium for non-developer sales post-2007, $\beta_1 + \beta_3$ is the impact of developer sales pre-2007, and $\beta_1 + \beta_2 + \beta_3 + \beta_4$ that of developer sales in the post-2007 period. In this specification, therefore, we can discriminate between newness as defined strictly defined by age versus that which is defined as a developer sale of new construction, recognizing that the circumstances under which a non-developer sale occurs is different for the different categories.

**Table 5. Coefficients from regressions treating both developer and non-developer sales as new**

| Ages | All (-2 to 2) | 0 | 1, 2 | -1&-2 |
|---|---|---|---|---|
| $\beta_1$ | -0.089*** | -0.024*** | -0.016*** | -0.104*** |
| $\beta_2$ | 0.207*** | 0.203*** | 0.011 | 0.343*** |
| $\beta_3$ | 0.027*** | -0.022*** | 0.029*** | |
| $\beta_4$ | 0.041*** | 0.018 | 0.120*** | |

**Notes**: Table entries are coefficients and significance levels for the indicated age group and measure as given in equation (2). The regressions include the same controls as specified in Table 4, including census tract and year indicators. *** indicate prob-values<.001.

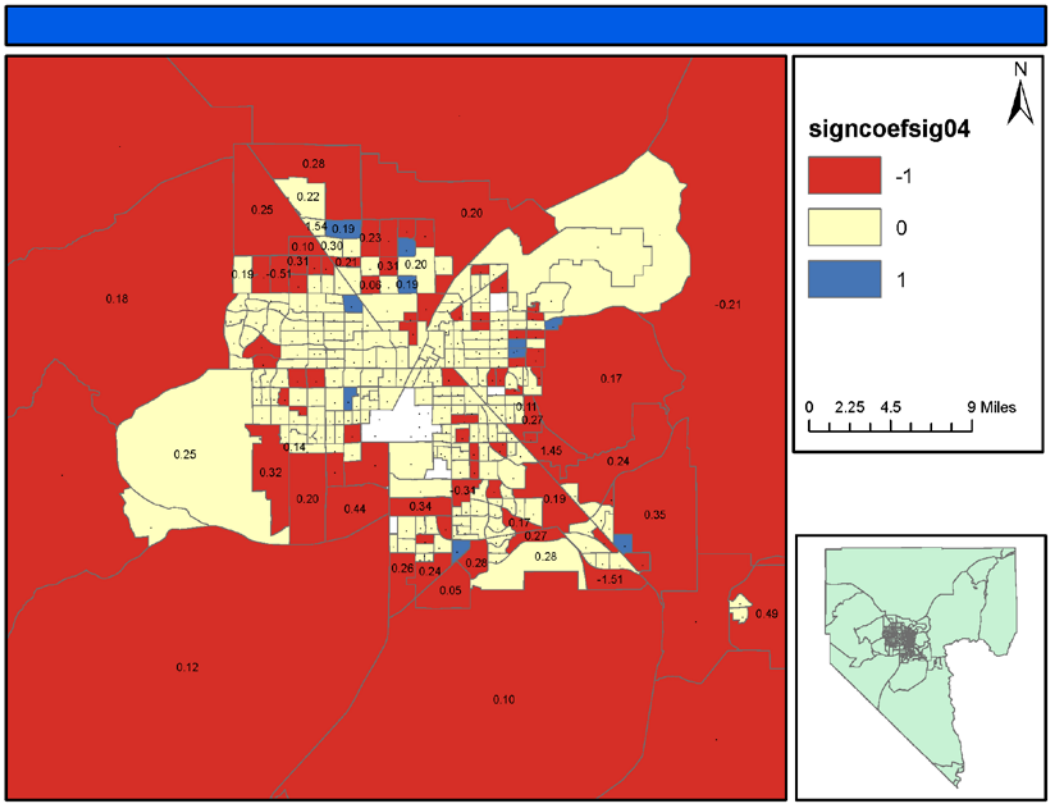
The first column in Table 5 provides results where each of the age categories has the same impact. As can be seen there, the effect of newness on home sales prices in the pre-2007 era was negative, -8.5 percentage points for non-developer sales of young houses to -6 percentage points for developer sales. The sales differences jumps to a positive 15 percentage points after 2007 for non-developer sales and 20 percentage points for developer sales. This puts the results of Figure 4 in a slightly different light; it indicates that any measurement of the new home premium might usefully be broken into two parts, that which arises because the home is new, and that which arises because the home has not been previously owned by another party. According to these results, the first factor is about three times as important as the second.

However, this may be because we cannot really observe which houses are in fact unused. It is s very likely that many non-developer sales of *age0* homes are "flips" (especially before 2007) and on that account never occupied. In order to gain further insight we break this down further by reporting on results of a second regression which replaces the newness indicators with our more finely detailed categorizations.

In the second column we have the four indicators/interactions for houses of age zero, and in the third column the same four, but as they apply to houses that are one or two years old. In the final column we report on the coefficients $\beta_1$ and $\beta_2$ for *presold* homes. We do not distinguish between developer and non-developer sales (i.e. we set $\beta_3$ and $\beta_4$ equal to zero) in this category since both types of sales will be of unused and customizable homes.

houses, the newness is enough; the premium for age zero homes that are non-developer sales is about 20 percentage points and that of developer sales is slightly higher (and the difference is statistically insignificant). This is congruent with the fact that even non-developer sales of age zero homes are likely to be previously unoccupied homes. In the post-2007, era these are rare.

**Figure 4. Map of Interaction Effects Between New Home Indicator for 2004 and Census Tracts**




The results are instructive. Before 2007, the premium that accrues to non-developer sales of age 0, 1, and 2 are negative and small in magnitude. After 2007, the premium for both developer and non-developer sales jumps, which is consistent with the prior results and with Figure 3.

The difference between the two categories lies with the source of the jump. In the case of age zero

However, for the category of one- and two- year old homes, the newness indicator is quite small, even in the post-2007 world. The source of the jump in Figure 4 for this category is from developer sales-- the coefficient on $\beta_4$ is ten times that of $\beta_2$. Premia for non-developer sales of "non-quite-new" homes are, roughly speaking, non-existent, even in the post-2007 world. This gets precisely at the distinction between the two concepts of newness. A one- or two- year old home sold by a non-developer is much more likely to

have been used in some way, and so not in the condition an age zero home would be in, regardless of the seller.

### *New Home Effects by Location*

In this section we examine the new home premium as it varies across locations in the Clark County housing market. We observed in the previous section Figure 3, shows that there were two regimes. New home premia for developers were negative before the bust and positive after the bust. We focus on the years 2004 and 2011 as exemplar years from each regime to estimate the geographic variation. Accordingly, we interact the *new construction developer* indicator[7] with census tract binaries for each of these years.

Figure 4 summarizes the interaction effects between *new construction developer* and census tracts for 2004 relative to all other years. Red illustrates negative values that are statistically significant ($p <= 0.1$) while blue illustrates positive values that are statistically significant ($p <= 0.1$). There are 92 negative coefficient estimates that range from -2.18 to -0.02 in Figure 4 which align with the negative coefficient values in Figure 3 for 2004.

Yellow illustrates coefficients that were not statistically significant while white illustrates tracts where there were not enough residential properties sold to estimate a value. The majority of new construction during the housing boom took place in the outer ring of the housing market, where buildable land was plentiful, and it was in these tracts that the new home premium was negative.

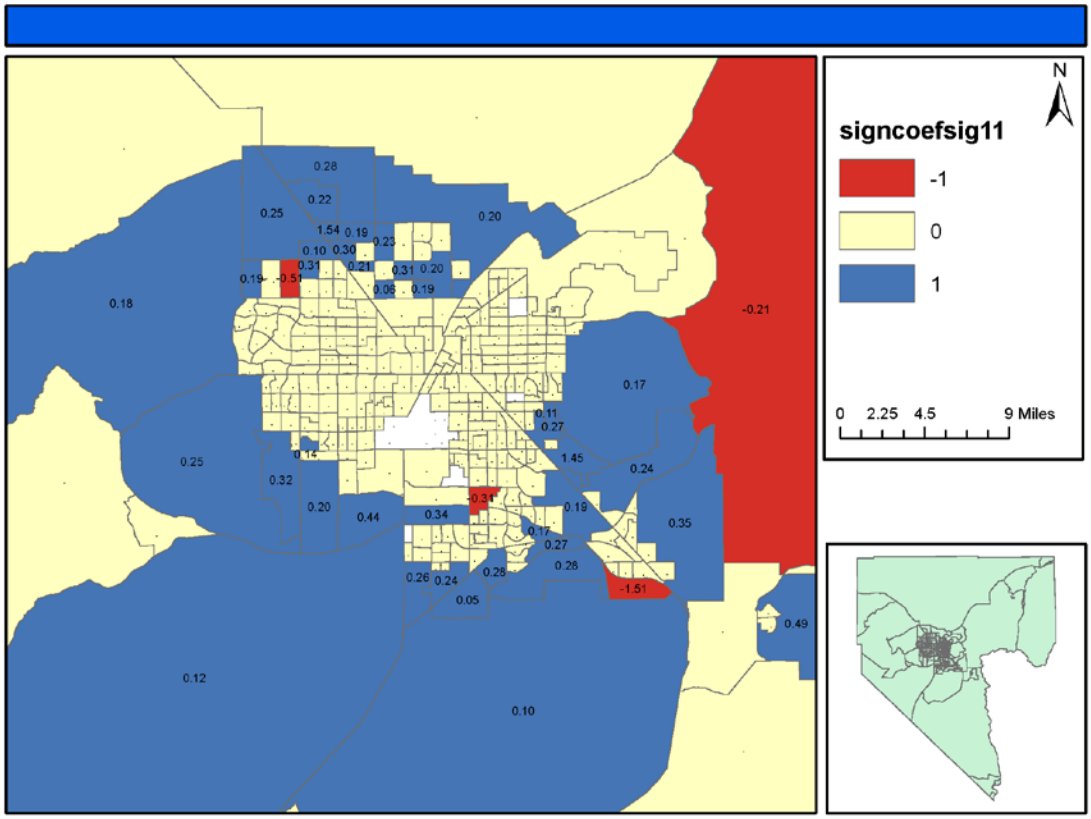
Nevertheless, note that the premium was positive in the ring just inside. Here, there was less new construction, and (congruent with the analysis of the previous section) with that restricted supply, a positive new home premium.

Figure 5 summarizes the interaction effects between *new construction developer* and census tracts in 2011 relative to all other years. Blue illustrates coefficient values that are positive and statistically significant ($p <= 0.1$) while red indicates negative coefficient values that are statistically significant. There are 44 coefficients that are positive and statistically significant which aligns with the positive estimates of new homes reported in Figure 3 for 2011.

Here, the new home premium is positive almost everywhere, precisely because even in the outer ring, new homes are fairly rare.

[7] That is, houses that are presold or age0. This could be done with the broader definitions of new without changing the nature of the results.

**Figure 5: Map of Interaction effects between new home indicator for 2011 and census tracts**




## V. Conclusion

We studied new home premia of developers using 785,022 home sales in Clark County Nevada during1994 to 2014. Our results suggest that the definition of a "new" home matters. Our results show that traditional methods of controlling for new construction might provide a downward bias as compared to measures that separate out new construction by builders and sold for the first time.

We also report evidence that premia associated with new construction varies before and after the housing bust in Las Vegas as well as across locations within the city. These variations are congruent with shifts in the supply of new houses.

When and where new homes are in short supply, prices rise, and when and where new houses are plentiful, the premium falls.

## VI. References


Bayer, P., Geissler, C., & Roberts, J. W. (2011). *Speculators and middlemen: The role of flippers in the housing market*. National Bureau of Economic Research.

Campbell, J, S. Giglio, and P. Pathak. (2011). "Forced Sales and House Prices." *American Economic Review*. 101. no. 5: 2108-2131.

Clauretie, T. M., & Daneshvary, N. (2009). Estimating the house foreclosure discount corrected for spatial price interdependence and endogeneity of marketing time. *Real Estate Economics*, 37(1), 43-67.

Coulson, N. E., & Lahr, M. L. (2005). Gracing the land of Elvis and Beale Street: historic designation and property values in Memphis. *Real Estate Economics*, *33*(3), 487-507.

Coulson, N. E., & McMillen, D. P. (2008). Estimating time, age and vintage effects in housing prices. *Journal of Housing Economics*, *17*(2), 138-151.Cropper, M., L. Deck, and K. McConnell. "On the Choice of Functional Form for Hedonic Price Functions." *Review of Economics and Statistics.* 70. no. 4 (1988): 668-675.

Goodman, A, and T Thibodeau. (1997). "Dwelling-Age-Related Heteroskedasticity in Hedonic House Price Equations: An Extension." *Journal of Housing Research*. 8. no. 2: 299-317.

Kahn, Matthew E., and Nils Kok. (2014). "The capitalization of green labels in the California housing market." Regional Science and Urban Economics no. 47: 25-34.

Morris, Adele C. and Helen R. Neill. (2014). "Do Gasoline Prices Affect Property Values" Climate and Energy Economics Discussion Paper, December 4, 2014.

Nichols and Schaffer (2007), "Clustered Errors in Stata," presentation, September 10, 2007. http://www.stata.com/meeting/13uk/nichols_crse.pdf

Robison, Jennifer (2014) "New Study Questions New-Home Resale Price Gap" *Las Vegas Review-Journal,* September 1, 2014

Rosen, Sherwin, (1974). Hedonic prices and implicit markets: Product differentiation in pure competition, *Journal of Political* Economy, 82:34-56.

Rubin, G. M. (1993). Is housing age a commodity? Hedonic price estimates of unit age. *Journal of Housing Research*, *4*(1), 165-184.

Sirmans, S. G., Macpherson, D. A., & Zietz, E. N. (2005). The composition of hedonic pricing models. *Journal of real estate literature*, *13*(1), 3-43.