University of Nevada, Reno

Price gradient analysis of the Washoe County housing market with emphasis on the impact of recent employment gains in Sparks, NV Industrial Centers.

A thesis submitted in partial fulfillment of the requirements for the degree of Master of Arts in Economics

by
Kevin Jay

Dr. Shunfeng Song/Thesis Advisor

May 2018

THE GRADUATE SCHOOL

We recommend that the thesis
prepared under our supervision by

KEVIN JAY

Entitled

Price Gradient Analysis Of The Washoe County Housing Market With Emphasis On The Impact Of Recent Employment Gains In Sparks, NV Industrial Centers

be accepted in partial fulfillment of the
requirements for the degree of

MASTER OF ARTS

Shungfeng Song, Ph.D., Advisor

Mehmet Tosun, Ph.D., Committee Member

Deena Schmidt, Ph.D., Graduate School Representative

David W. Zeh, Ph.D., Dean, Graduate School

May, 2018

Abstract: This paper seeks to understand the role of the Tahoe Reno Industrial Center and freeway accessibility in determining housing prices in the Washoe County housing market in recent years as well as comparing different models. This is accomplished through hedonic, repeat sales, and difference in difference regression analysis. Findings show there are optimal distances from a freeway to increase home price as well as evidence to conclude the Tahoe Reno Industrial Center has increased housing prices by a small amount in recent years on the east side of Washoe County closest to the Industrial Center for a "medium" priced homes. This is in contrast to the popular opinion and news reports of the Industrial Center being the cause of large housing price increases in Sparks. While housing prices in Washoe County have recently seen large increases, the cause could instead be a recovery from the Great Recession or other reasons and not directly related to the increased employment and opening of the Industrial Center and its occupants in Sparks.

Table of Contents

1. Introduction:

The goal of this paper is to explain the changes in price gradients of the Washoe County housing market as a result of the increase in employment east of the City of Reno, specifically on USA Parkway where most famously Tesla has located it's new Gigafactory. Most renters and future homebuyers are acutely aware of the increase in prices taking place in the Reno/Sparks area and how prices have been affected. This paper seeks to give guidance to policy makers, future home buyers, and econometricians on the Washoe County Housing market and the different tools available to better understand its growth in recent years.

This paper seeks to make three main contributions to the body of research on price gradients and housing markets. First, when analyzing price gradients allows future Reno/Sparks residents to understand the housing market, it helps construction and home builders better plan new developments, and most importantly it helps policymakers understand the needs of the local housing market and where outsized gains are located. When policymakers discuss property taxes and community improvements, these gains should be accounted for in their analysis. A boom in employment such as the one Reno/Sparks has seen lately is unlike many other employment booms. If a large metropolitan area has an equally sized employment boom, the area can absorb that increase without a large shock to the housing market. For example, if the San Francisco Bay Area sees an additional 12,000 workers in their area, it will not shock their housing market as much as Washoe County. So, this analysis allows the reader to understand how a large shock affects an entire region, and not just a small sample of that region.

Secondly, it studies the impact of freeway accessibility and its relationship to housing prices in Washoe County. Policy makers should have an understanding of how increased freeway accessibility affects home prices so they know where outsized gains to home prices are located, and who those gains are going to. If policy makers want to increase freeway access to lower income communities, but these low income communities are mainly renters, then the gains to housing prices will go directly to the landlords. This can give guidance on how to use tax policy to remedy inequality issues such as this.

Thirdly, hedonic Price Regressions will be used in addition to a repeat sales analysis to analyze the market. The coefficients of both methodologies will be scrutinized to understand differences, strengths, and weaknesses of each model. Using these two analyses on the same data set will give a very clear picture of which is more effective when used to capture a shock to a housing market, and drawbacks of using either. These three contributions will guide this paper's methodology and reasoning throughout the design and execution of the analysis.

2. Literature Review:

The literature review will consist of the following parts: An understanding of what a price index is and the role of a CBD in determining land price. A section on the pros and cons of various house price index measures.

CBD & Price Gradients

CBD is a term used to describe monocentric population centers. This is when the bulk of economic activity occurs around the city center where land prices are highest and then outside this region there are housing units, then farm land and other land intensive uses. The term price gradient is used since the price of land and housing tends to fall exponentially away from CBD as the distance is increased from CBD (Osland, Thorsen, Gitlesen, 2007). A price gradient can be thought of a dimension of prices that help us to value land based on distance to a CBD (Antoniucci & Marella, 2017). Theory states that land closest to the CBD is the most valuable and therefore the most expensive. As you move further away the price will tend to decrease as the land is less valuable. This can applied to the context of Washoe County due to the downtown/midtown area being defined as the CBD where employment is located and where land is the most expensive. This decrease in price around the CBD can be explained by an exponential function with a negative exponent. The price decreases in this fashion around the CBD. Eqn 1 hows how this logarithmic transformation takes place.

Equation 1:

$$\ln(\text{Price}) = \beta Distance + \epsilon$$

where Distance is an X variable that we use to describe price. This log transformation can force the coefficient beta to be in front of the X variable distance and therefore can be easily used in an OLS regression. Later in the analysis this can be partitioned to east and west or other parameters to better understand the impact of Reno Tahoe Industrial Center. One thing to consider about price gradients over time is that when the price moves it usually does not do so in a uniform fashion. For instance, if the prices very far from the CBD rise and the prices at CBD rise at the same rate, then the prices rose uniformly. However if the prices at one of the ends stay constant while the other increases or decreases then it is difficult to determine the whole picture of the price movement. In the case of an ordinary price gradient where prices are decreasing with distance to the CBD, if the gradient from one year to the next gets steeper, it could be due to either prices at the CBD increasing in value, or prices away from the CDB getting lower, or a mixture of both, or prices at the CBD increasing in value at a faster rate than those away, or prices away decreasing faster than those at CDB. This is all in vice versa for situations in which the price gradient getting shallower. In these situations it is necessary to control for other variables that help to understand which situation is occurring.

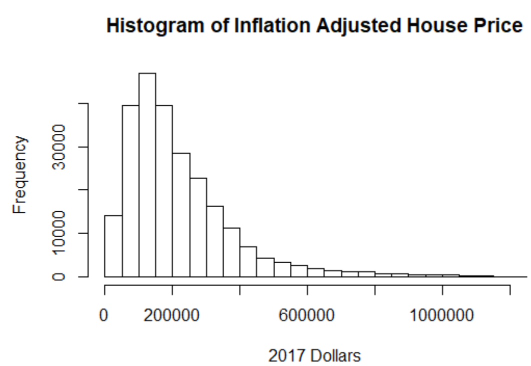
There does need to be a discussion of the difference in monocentric vs polycentric metropolitan areas. A quick glance at a map of Washoe County shows a very monocentric city model where there is a clear city center of downtown Reno with the rest of Reno and Sparks pushed outwards in all directions. This is heavy contrasted by monocentric city centers such as LA where there are different CBDs and the end of one CBD and the start of another are not so clear. Some people commute past another CBD on the way to their job in a polycentric CBD metro. For the analysis of Washoe County it is safe to assume that prior to the development of the Tahoe Reno Industrial Center, there is little evidence of a polycentric city model. Instead the use of a monocentric city model will be useful in the application of our analysis.

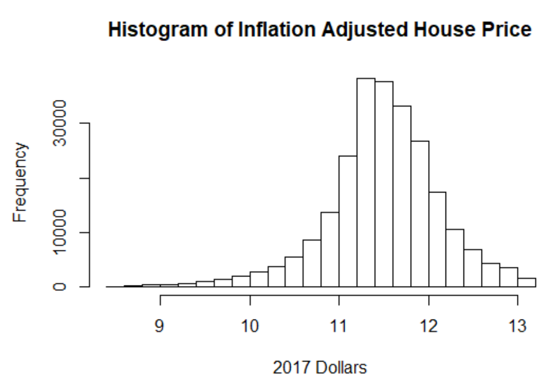
Hedonic Vs. Repeat Sales Models:

There are two predominant ways to estimate housing prices. One is to use hedonic price models and the other is to use a repeat sales model, which in many cases is an extension of the hedonic price model.

The hedonic price model is a model that studies the amount that consumers are willing to pay for certain attributes of a house. More clearly, a house has a final sale price without any line items for how much an additional bathroom costs, or 1 vs 2 car garage. hedonic models allow us to observe attributes as explanatory variables with the dependent variable of house price. This gives us an approximation of all β coefficients of the attributes to find out how much consumers pay for an extra square foot or in the case of this paper for being closer or further from the freeway. One common issue with hedonic price models are unobserved variables (Abbott, 2011). Of course one unobserved variable this paper contributes to the research is the estimation of freeway accessibility by method of "as the crow flies distance" to the nearest on ramp as well as the distance from the home to the freeway to control for noise and pollution. Abbots paper focuses more on spatial unobserved characteristics that have traditionally been controlled for through spatial fixed effects.

The issues with the hedonic Price Model begin first with a discussion of the use of transformed dependent variables with respect to the functional form of the model. (Chin, Chau 2002) Particularly when deciding to use a linear model, log-linear, or log-log model. Chin and Chau's paper notices many economists prefer the Box-Cox transformation to normality since it provides a better fit of the data. But, there is not much of a roadmap to determine when to use Box-Cox transformations, log transformations, to leave the data in its linear form, or to use another type of transformation. Some papers use both and find that there are better goodness of fits when using log transformed version of the housing price (Landajo, 2012). For the data set in this paper, the log transformation has been applied to the dependent variable of the home price. This has the effect similar to the Box-Cox transformation of normalizing the data, especially since the distribution of linear home prices is heavily skewed with many homes around a mean and then the right tail moving very far into the rich homes with only a few observations. The below graphs show the difference between the histograms of a levels and log transformed version of the home prices. It is clear that the graph on the left is the levels of the price and is highly skewed to the expensive homes. The right side graph of the log transformed homes are instead more normally distributed and in fact even a little skewed left. This is what the transformation looks like when performed on the same data set.





The final benefit to the log transformed model is the ability of the log transformation to reduce the impact of outliers on the regression. It is no secret that outliers can severely impact the effectiveness of the OLS regression, and data cleaning to remove these problem variables can at times be “arbitrary” (Bourassa 2016). Which in experience is true, if you remove any observation outside of 4 or 5 or more standard deviations your analysis will be impacted downward(in the case of expensive homes) or upward (in another case requiring removing small variables). Further, do you conduct the data cleaning on the linear home prices (or dependent variable) or do you conduct the cleaning on a transformed version of the variable? There is no gold standard for the rules of cleaning the data set to ensure important homes are not removed. One thing is clear is that the concept of the data cleaning is a much discussed issue in the field of empirical economics. Billings paper conducted an analysis of how severe different data cleaning methods were in biasing results. However, their main conclusion was that model specification and experiment design was more important than data cleaning in creating bias within a housing data model (Billings, 2015). Instead it must be argued in the paper why the observations are removed and a discussion of why they are removed must be included in the paper.

Repeat sales models focus on the changes of the house and house values over time and aggregate this data over all homes within a specified region or subset. repeat sales produces an index that is much different than indexes used for stock markets or other assets. Many assets are easily tracked in real time to find their prices, and these assets usually are all the same. A single stock for company X is the same as another individual stock for company X so the trades taking place can be calculated into the index. However, houses are different than these assets most notably that they are all different from one another (even between direct neighbors) and they are not sold at regular intervals. Some homes may be sold quite frequently and others may be bought new and never sold for many decades. repeat sales allows for these differences to take place and be accounted for properly.

The way the repeat sales analysis accounts for the time, sale price, and amenities offers serious advantages over other models. This model type can find the common appreciation rate of all homes over the time period specified (Melser, 2017). This allows a region to be studied and to find the rate of growth of housing prices for that region. The clear advantage of this method is being able to track the home sales and create this index. As opposed to the hedonic method which benefits from a larger dataset, however suffers in the inability to create an index of the same quality.

One issue with the data, however, is that there is no variation in the homes between sale times for the parameters of interest. Homes do not change location over time so their distance to an onramp does not change between sales. This is unfortunate because when the data is difference within the repeat sales model there is no way to control for variation in the distance to the freeway so repeat sales may not be the best estimator. repeat sales is also not without its own issues aside from data set specific issues. The most notable repeat sales issue is the huge reduction in observations that takes place to create the model.

In the hedonic approach, each sale is seen as an observation. However, in repeat sales one can think of the repeat sales pair as one observation. This in turn can heavily reduce the number of observations available in two ways. First, if there is a home that has only been sold once, potentially it is a brand new home or was bought a never sold, it will be eliminated from the dataset because it has no prior sale to reference that sale price against. The largest concern for this problem is that if a newly constructed home is purchased and more home sales are happening “around” this home in terms of time, then the index will disregard these homes. If these homes are statistically significantly different than the homes contained with the repeat sales model there will be bias in one direction.

This exact claim was tested by Meese and Wallace (1997) who tried to determine if there was bias repeat sales models due to this problem. Their hypothesis is that so called “starter homes” are uniquely different than other homes in terms of attributes, as well as in the frequency of sales. They figured that starter homes are smaller and in poorer quality, and are only lived in for a few years before the young family moves out of the home into a larger one that they stay in for longer periods of time. Therefore this means that repeat sales models have far more observations of starter homes than regular homes, and definitely more than the new homes that have only one sale attached to them. They confirmed this result by looking at the group of homes contained within the repeat sales model and the homes that were excluded and noted that the two groups showed statistically significantly different attributes.

Therefore, if the results of Meese and Wallace hold true for the analysis in this paper as well, then the price index will be skewed lower than is real because of the number of lowered priced “starter homes” in the data set. In fact this is the case and the results of this analysis are found in the data section of this paper. Without withholding this information any longer from the reader, the data does show a decline in both median and average price for a home as the number of sales for a particular home increase. This could be an issue with the most recent years of data since any newly constructed homes will be excluded, however all homes built and sold in the past will be counted and useful for the analysis Therefore the more recent years will have less observations than the years in the past when compared with the hedonic regression observations.

This is not necessarily a huge issue on its own, one method to understand this result is to only use homes that have repeated sales within the hedonic regression data set which would be a way of resolving the issue. However, this would not be comparing the same data sets and would require excluding certain homes from the analysis. If the analysis were performed in this manner it would be comparing the same homes and more specifically it would show the effect of TRIC on starter homes. This of course is not representative of the entire housing market but it would allow a concrete analysis of this subset of the market. The second way in which the dataset can be reduced is when there is a home that has been sold, for example, 3 times. Then the number of observations for this particular home is 2, because between sale 1 and sale 2 there is repeat sale, and between sale 2 and sale 3 another repeat sale. Therefore for any home in which there is data on the sales, the number of observations is effectively the number of sales minus one. So, if the data set is full of homes that have been sold only twice, then the dataset will be reduced by half, if the data set is full of homes that have been sold 3 times then it will be reduced by a third, and so on. So, depending on the distribution of the number of home repeat sales in the data set, the data set will be reduced by a large amount when in comparison to the hedonic sales observations. This is a large problem when the analysis begins with a small data set, however with all of the sales of Washoe County for the years from the mid 1900’s to the end of 2017 with 6 sales of each home, it can be confidently stated that there is no issue with the data reduction issue. The other predominant issue with the repeat sales model is the violation of the assumption that the asset sold between the two sales is the same asset. This is blatantly violated by the age of the home (Nagaraja, Brown, Wachter 2010). As time goes on the housing market typically goes up and appreciates even when controlled for inflation. However, as time increases so does age, which means the home at sale 1 is different than sale 2 due to age. Of course stocks and other assets cannot age, but homes definitely do and this can cause issues with the assumption that the asset is the same with the only difference being the price over time. Therefore controlling for age or depreciation is a common tactic to reduce the severity of this problem.

Price gradients are used to determine how distance to CBD affects the price of a home. Price gradients are in the shape of an exponential curve and provide insight towards how the price "falls" away from the CBD. As property values decrease away from the CBD it is typical for the shape of this decrease to be an exponential curve with a negative exponent and the challenge is to find the value of this decrease. In this paper we will discuss how the increase in employment centers east of the Reno/Sparks area has "elevated" the price gradients. We can think of price gradients as a three dimensional map of the Reno Sparks housing market. Visually, the Reno Arch can be thought of as a "mountain" where the prices are highest at that point and all prices around it begin to drop off like the slopes of the mountain. This gradient is in the form of an exponential variable with a negative exponent indicating the exponential decrease away from the CBD until prices are effectively flat.

Of course the models have their positives and negatives. The clear benefit from the hedonic approach is the ability to include all of the observations in the data set. The downside to the hedonic approach is the lack of tracking the sales between each house. For the repeat sales model there is a clear advantage of tracking sales for each property over the years for as many sales as the data set allows (in this case is 6 sales). But the downside is the removal of all homes that have only one sale, which is typically those homes that are very expensive and the homes that are very new.

Review of Each Model:

This section serves to differentiate the two different models that will used in the paper. First, a hedonic model has the following form.

$$\mathrm{Y} = \alpha_0 + \sum_{i=0}^{I} \beta_i X_i + \epsilon$$

where the price of the home or asset is a function of a constant, a summation of all characteristics with associated characteristics and an error term. Usually the price of the home is log transformed which gives the double attributes of reducing variability and allows the convenient method of interpreting the coefficients of the housing characteristics as percentages. So a coefficient of a bathroom would indicate the percent price increase of one additional bathroom to a home.

Next, a repeat sales model has multiple functional forms. These different functional forms include simply just the sale date and price or can include adjustments to the property as well (Shimizu, Chihiro, 2010). A basic repeat sales model which only accounts for the sale price and time is as follows:

$$\Delta_{j,k} ln(Y_i) = \mathrm{D}_i d + \epsilon_{i,j,k}$$

where the left hand side has the difference of log price and the right contains dummy variables of time where the first sale is noted with a -1 at the time of first transaction and

1 at the time of the second transaction. Now, one concern with a model such as this is that it does not take any adjustments to the home into account when estimating with OLS. Therefore, we must find a new model that does allow us to account for changes. These changes could be anything from a new bathroom, the addition of a freeway onramp nearby, or even negative attributes such as increase crime rates or the nearby construction of a landfill. Clearly these changes will affect the housing price and should be accounted for. Including these changes results in the below model (Clapp 1998):

$$\Delta_{j,k} ln(Y_i) = \mathrm{D_i} d + \beta_i \mathrm{X_{j,k}} + \epsilon_{i,j,k}$$

where the coefficient beta on the variable that changes is shown. For many homes, this coefficient will not change. But for those homes in our sample who had a reduced access to the freeway this will be a determinant of how much that reduced distance has an effect on the housing price.

Another topic to consider is the change in valuation of amenities over time. For example, if many retirees move into Washoe County and desire one story homes rather than two story homes for ease of movement through their home in old age, then the value of one story homes will increase all else equal to two story homes. The interaction of this attribute with the time variables will give an indication if this is true and has a significant effect on the data and market (Cohen, Blinn, Boyle, Holmes, Moeltner, 2016).

Difference in Difference Analysis

A difference in difference analysis allows economists to turn an observational data situation into a quasi-experimental design that takes advantage of a policy change or other abrupt shift in circumstance and compares results between two similar groups. For example, Card and Krueger (2009) studied the impacts of the minimum wage increase in the state of New Jersey did not decrease employment in the state's fast food industry. They concluded this after comparing New Jersey and Pennsylvania employment figures for fast food and used Pennsylvania as the control in this "experiment" to show the difference in difference that the increase in minimum wage had on employment in New Jersey. The analysis must contain control and treatment groups and a time period that can be designated as pre and post treatment. Famous examples of the use of difference in difference include topics in labor economics such as how minimum wage changes in a particular state affected employment. Data was collected from two states, one state was a control and the other was the state that changed their minimum wage, then the data was also collected from before and after the new minimum wage went into effect. The DiD analysis was then used to determine if the minimum wage resulted in reduced employment. This type of "simulated" experimental design is useful in pretending that a controlled situation is happening and the effect is captured in these cases. The typical difference in difference analysis starts with a table of means that are compared to understand how at the means the analysis looks. The the below table, the variable Y is simply the variable of interest.

	Before Treatment	After Treatment	Difference
Treatment Group	12.55925	12.60315	0.0439
Control Group	12.60015	12.69319	0.09304
DiD			-0.04914

In the far right column, the difference between the treatment group before and after gives an idea of the effect of the policy change. The difference between the control group before and after gives a sense of the change that would have occurred if the policy change had not taken effect. Then the difference in difference shows the change in the treated minus the change in the control which is supposed to provide an idea of the effect of the policy change. The main issue with this is that both the treated and control group need to have similar trends in the pre treatment period (Pischke, 2008). This assumption, commonly referred to as the “common trend assumption” gives assurance that the two groups are similar in every way but in the policy change. Therefore with the policy change being the single variable that changes at that time, it is then valid to show the difference in difference is confidently the effect of the change.

What is left out of the above table are the independent variables that make up the rest of the dataset that is typically contained within the regression analysis. The final difference in difference regression will be the following model:

$$Y = \beta_0 + \beta_1(TreatmentGroup) + \beta_2(PostTreatment) + \beta_3(Interaction) + \beta_i X_i + \epsilon$$

where Treatment Group is a dummy variable to indicate which set of observations are the group treated, Post Treatment is dummy to indicate which observations take place after the observations occurred, and finally the interaction term indicates the multiplication of the Treatment Group and the Post Treatment variables to determine where and when the treatment is effect. The additional □□ variables are control variables that would typically be used to describe the data in a hedonic regression. These variables are still useful to control for variation in the observations during the difference and difference analysis. In fact they are essential to control for large variables that could potentially change the results of the study in the case they were not controlled for.

The main problem surrounding difference in difference analysis is the previously mentioned "common trend assumption". The common trend assumption assumes the control and treatment groups both follow a similar trend in the pre treatment period, and would have maintained that trend together given the treatment did not occur. This allows the analysis to use a situation referred to as the "counterfactual", or what the difference is for the treatment group from what would have happened if the treatment did not occur.

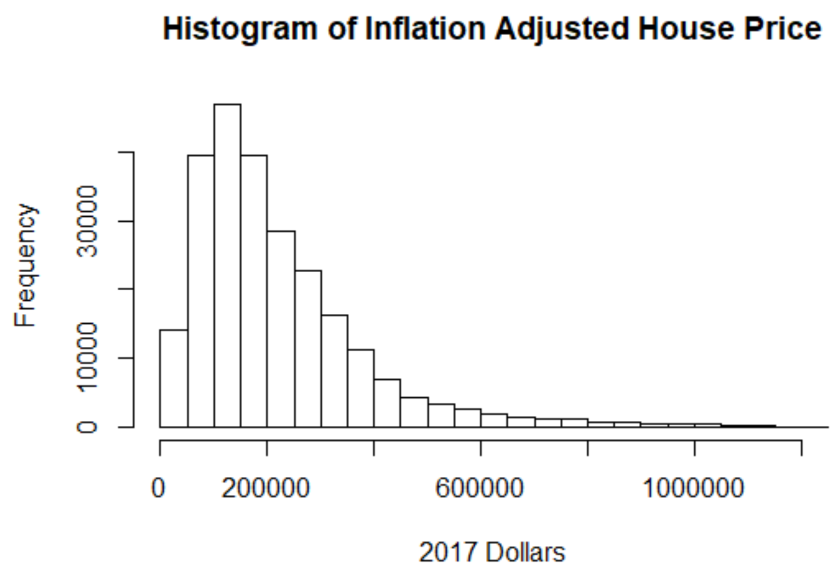
This is analysis is commonly shown in the above table using the appropriate average values for each box. But the amazing aspect of this analysis is the ability to control for descriptive variables in the regression. Whereas just using averages will ignore the descriptive data, the regression form of this model can give a holistic picture of the difference in difference.

3. Data and Study Area:

The study will be conducted on the housing market of Washoe County, NV. Washoe County sits on the Northwestern edge of Nevada with California directly on the West, and Oregon directly to the North. According to the US Census Bureau 2016 estimates Washoe County contains approximately 453,616 residents. Washoe County has 178,522 workers spread between 37,029 firms with a median household income of $54,955, while also having 12.5% of its residents living in poverty (US Census, 2016 estimates). Washoe County has seen a surge in economic activity after being devastated in the wake of the 2008 recession. Washoe County contains lots of rural land to the north, and most of its population is based around Reno (pop. 245,255) and Sparks (pop. 98,345), the rest of the residents live in so-called "Census designated places" and "other communities" which include unincorporated parts of the county. The amenities Washoe County provides include the University of Nevada, Reno, Truckee Meadows Community College, and the Desert Research Institute. Washoe County is relatively isolated from other metropolitan areas. Carson City (pop. 54,521) is located about 30 minutes south, Sacramento (pop. 495, 234) and the Bay Area (pop. 7 million) are 2 and 4 hours respectively from Washoe County. To the East, Fallon (pop. 8470) and Fernley (pop. 19,588) have small populations and many employees commute from those communities into Reno/Sparks for work. This allows the analysis to be confident that the housing market in Washoe County is not affected by commuters to Sacramento or that the market is too affected by Washoe County residents leaving to Fallon or Fernley as their relatively small populations have not seen any large increases in recent years.

The bulk of the data comes from the Washoe County Assessor's Office. There they have a repository of data that covers the last 6 sales of every piece of property in Washoe County. The data was downloaded in bulk form, however individual properties can be analyzed from this data set online and compared to other in terms of all attributes that this paper discusses. This is an extensively rich data set that allows for many variables to be included in the regressions. This dataset is also extensive in the length of time within the data. The earliest sale recorded in the dataset is from the 1940's. While sales from this early are not included in the hedonic regression, they are included in the repeat sales model. This is because the repeat sale occurs at some point later in time, potentially in a time of interest and provides information on the sale price for that future point. If these sales were disregarded it would reduce the data set by a certain degree. Specifically for our data set, the number of repeat sales homes in the data set is about

70% of the observations of the total data set. This combined with the information on multiple sales makes it easy for us to run the repeated sales model. Additionally, data was collected from the Federal Reserve Bank of St. Louis and the U.S Census Bureau, Nevada Department of Transportation, and google maps. Each of these data sources are explained below.



There are however some negatives to the Assessor's office dataset. First, when the Assessor's office approves a new home sale, they update all values at the time of the assessment of the property. This means if an additional bathroom, garage space, sunroom, or other amenity was produced between two sale times, it will not exist in the data set. So, the values of the variables represent the characteristics of the home at the last sale. Meaning when a repeat sale regression is run, the home amenities could have potentially changed but there is no way to know from the data and we have to accept this possibility in our analysis. The only way to develop a method to find changes in the explanatory characteristics is to look at the changes over time that can be calculated after downloading the data from the assessor's office. This means if an onramp was built in a particular year it can be determined what the decrease in distance between the sales of the home. However, this paper uses the distance to the Reno Arch as a main determination of treatment versus control. And the distance between the arch and each home does not change throughout the dataset, therefore it must be accepted that any improvements, additions, or other changes to the homes are not accounted for over time with the exception of the age of the home which is simply the year built subtracted from the sale year.

The variables collected from the Assessor's office on about 160,000 single family homes are the following: Sale Amounts, Sale Dates, Building Square Footage, Lot Size, Year Built, latitude and longitude. Again, the sale amounts and dates are for the last 6 sales of the home. The average time between sales within the dataset is about 6.5 years. The log of sale price is used in all regressions. This is common in the literature and other research and is useful in reducing the large bias that expensive homes (i.e. outliers) create within the data. A handful of large home sales will skew the coefficients and using the log of price helps to reduce this problem. Building Square Footage give an indication of the interior building size. Lot Size gives the area of the land plot. Year built was not used as a stand alone variable. Instead it was subtracted from the sale date to find the age of the home at time of sale. Often times the square of age was also used in the regressions to better fit the data. This allows for greater variation in the data, if old homes increase/decrease in value at increasing/decreasing rates the square of age will help find it. The longitude and latitude were used for a handful of purposes. This geographic data actually came from a separate dataset in the Assessors office and was paired using a merge function using the PIN for each home, which is an identifying value for each property in Washoe County. Once the lats and longs were paired with the property, it was used to determine the distance to many different items. For the freeway access variable, the distance between the home and every single freeway on ramp in Washoe County was calculated, and the smallest value was then used as the distance to the nearest freeway. More specifically:

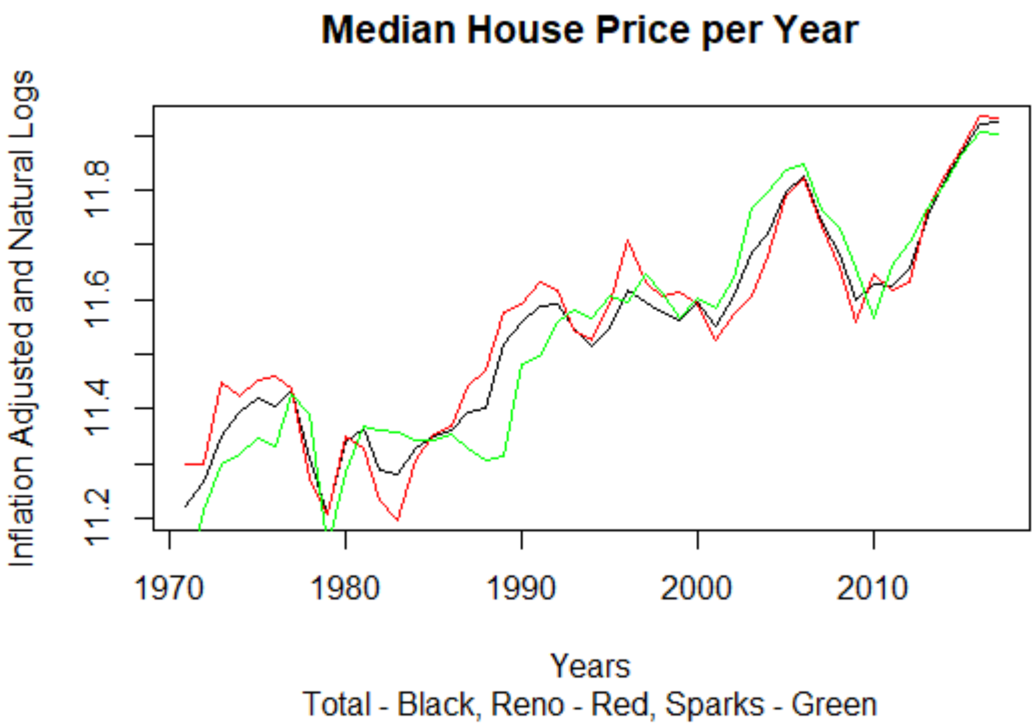
$$\text{FreewayDistance}_{i,j} = min(HavrsineDistance(Onramp_i, House_j))$$

Where the Haversine Distance (Mendoza, 1795) is given as

$$\text{HaversineDistance} = 2\text{rsin}^{-1}(\sqrt{sin^2((lat_2 - lat_1)/2) + cos(lat_1)cos(lat_2)sin^2((long_2 long_1)/2)}$$

The key benefit of using the Haversine Distance is that as distance between points increase, the curvature of the earth (□) is taken into account and one can imagine the Haversine Distance as walking across the earth's surface rather than simply ignoring the roundness of the earth and taking a straight line "through" the earth. However, this is still a "as the crow flies" distance between the house and the onramp. There are obviously issues with this approach. Using a "as the crow flies" approach does not account for twisty or windy roads that a commuter must use to actually reach the onramp. But this gives a close approximation of the distance from a home to a freeway onramp. The lats and longs were also used to compute distance in the same fashion from the Reno Arch. The Reno Arch is determined to be the CBD of Washoe County. When looking at previous years, the exponential coefficients on the log of the arch are negative in all directions so it was at one time the CBD, however this has changed due to the Reno Tahoe Industrial Center which has made Reno more of a polycentric metro area with potentially two CBDs. The Arch will be counted as CBD so that a comparison can be made between old and new within Washoe County.

The data from the St. Louis Federal Reserve Bank provided inflation, mortgage, and unemployment data. The CPI was used to convert all of the housing sale prices into $1983 dollars. The reason for the year of 1983 is because that is what FRED uses as their base year. The reason to inflation adjust however is more important. Doing this transformation allows the numbers to be a consistent dollar amount from year to year and accounts for changes in the price index so that 1 dollar in 1998 is roughly the same as a dollar in 2017. Inflation adjusting has huge impacts on the data set. When looking at home prices in 1998 (the start of the study period) versus home prices in 2017, the dollar amount change due to inflation is 50% between those two years. This means a dollar in 1998 is the same as one and half dollars in 2017. Without adjusting for this the results would be heavily biased and would cause many issues in the analysis. After inflation adjusting it is easier to understand how the impact of the explanatory variables impacts the house price. The graph below gives an visual representation of the Reno and Sparks housing market throughout the 1970's to the end of 2017. This graph was produced after inflation adjusting to 1983 dollars and then showing all of Washoe County, and then a Red line for Reno and Green line for Sparks.

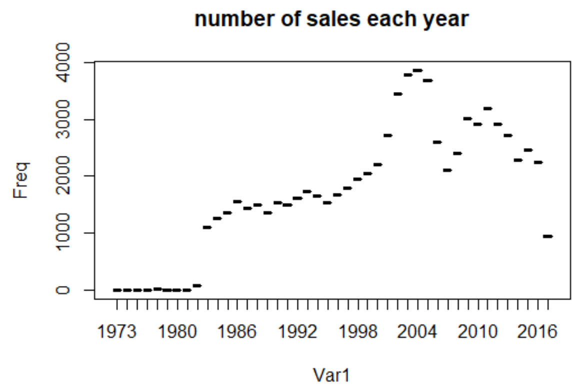


Unemployment gives us a macroeconomic health of the region. The data is only for Washoe County and therefore matches the housing price set. It is a monthly index and allows our regressions to capture the activity of the labor market in the county. The last data collected from FRED is the 30 year fixed rate mortgage rate given again in monthly intervals. The mortgage rate is a huge cost of buying a home and should therefore have an effect on the price of a home. The economic intuition is as the 30 year fixed interest rate rises, home prices should decrease as it is now more expensive to purchase the home over the life of the loan. Of course there are issues with this as well, namely that this the national 30 year rate, and not just for Washoe County, and finally that some property investors or even some homeowners pay in cash and do not use a mortgage. Therefore the rate will not affect those individual purchases. However, on the whole this gives a view of the macroeconomic conditions that exist within the US for each month of every year in the data set. All of the data from FRED is monthly data. And so the month and year of the housing sale is attached to the month and year of the FRED Data. The data is an average for the month, so even if a home was sold on the 1st of the month, it is still linked to the month and year of

4. Methods:

Hedonic Regressions by Year

For the regressions, each year has a different regression and then each year was split yet again for east or west of the arch. Then the coefficient of the log of distance from the arch can be analyzed against other years and east vs west. The study time for this analysis is the year the Reno Tahoe Industrial Complex was bought for its current purpose, 1998, to last full of year of data in the data set which is 2017. There is data in our data set for years prior, however the information is unnecessary when looking that far in the past. Instead, just limiting our study to 1998-2017 will give us the important timeline. The reasoning for running separate regressions each year is it allows us to compare the coefficients between each regression quickly and easily for each year as opposed to having the same coefficients for all of the dependent variables. The variables used in the regression are: Building Square Footage, age and its square, distance to the nearest freeway onramp and its square, Owner Occupied, Bathrooms, the 30 year fixed rate mortgage rate, Washoe County Unemployment rate, and distance from the Reno Arch. The mix of housing observations and macro economic conditions give the data a richer set of information to use in the regressions. Note that during the hedonic regressions in addition to inflation adjusting the home price, the natural log of the sale price was used rather than the level price. This is because the distribution of home prices are heavily inequal and using logs allows the data to be normalized better and reduce large outliers. The graphs below confirm this issue.



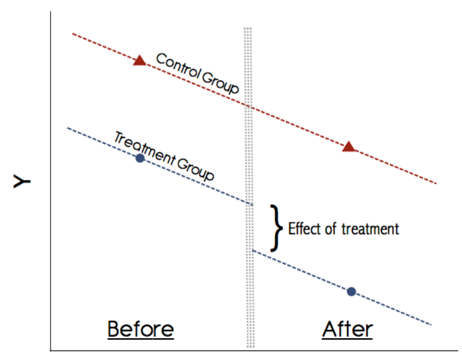
The square of age and distance to a freeway onramp was used to allow more flexibility in the estimation of their respective effects on the log of housing price. Based on the coefficients of these estimators it can be determined if there is a multiplicative effect to being near an onramp or if there is an optimum distance to the freeway a home should be for housing prices to be as high as possible for on ramp. Similarly with the age of the home, the square can indicate if homes lose value very quickly past a certain age or if there is an optimum age for a house, or if old homes retain value.

A CBD needed to be determined for Washoe County and the Reno Arch located on 4th and Virginia is used as the CBD. Why use the Reno Arch and not another area for CBD? It is because the downtown area has historically been the center of business activity in Reno. With the Casinos surrounding the Arch, City Hall, a Federal Courthouse, and much more surrounding the Arch it can safely be assumed to be the

CBD of Reno in historical terms until just very recently. This can be empirically proven by looking at the coefficients of hedonic regressions throughout the 90s. The coefficients of the log distance from the Arch are mostly insignificant or negative with significance. This will be expanded upon in the results section, but it clearly shows the Reno Arch can be used as CBD for the purposes of this paper. The data also had to exclude those values of home prices that were $0. The “sales” of these types are typically from the transfer of a home from an owner occupied situation to a trust in the owners name or to a family member. The Washoe County Assessor's office labels these transfers as “sales” regardless of whether or not it is what we would colloquially determine to be a “sale”. This is an important discussion to have in the methodology because when the natural log of the housing price is taken, a sale price of $0 results in a log value of negative infinity. Then the regression cannot be run and therefore these sales need to be excluded.

Difference in Difference

To better understand the conditions of the market, the exact same variables from the hedonic Regressions were used in a difference in difference analysis. The homes located west of the Arch are denoted as the control group, while the homes on the east are deemed the treatment group. For the time of the effect, the year of 2014 was used as the treatment time since that is when Tesla announced the Industrial Center as their home of the gigafactory. Using this methodology allows this paper to simulate an experiment design but using observational data. The experiment would be as if the western and eastern parts of Reno were two separate populations and then the treatment affected the eastern portion on the year of 2014 and after. Of course, it is necessary to discuss the issues inherent in difference in difference and how to control for them. In the literature review, a large portion of the paper was spent explaining how the parallel trend assumption must be maintained in order to gain reliable estimates. Practically what this means is that in time periods surrounding the treatment time, the trends are indistinguishable. In the pre treatment period, the years prior and post to 2014 were tested as proxy "treatment" periods in order to simulate testing the effect of TRIC in those previous years. If the difference in difference variable, recall this is the interaction term between treatment vs. control and pre vs. post treatment, is insignificant, then it means this proxy treatment time period has no effect on the outcome of the experiment. Therefore in the pre and post time periods the parallel trend assumption holds and the analysis is valid. This can be visualized in the following graph.



Notice both groups have the same trend in pre and post treatment and the only break is in the treatment time period. This gives the difference in difference coefficient. And, since the trends are parallel in the previous time period and there is no difference between the two, then the DiD coefficient will be zero or in other words, insignificant.

Repeat Sales

The overall method of the repeat sales model is explained in the literature review section. This section serves to explain the practical implementation of this model and understand its interpretation. The repeat sales model is again constructed using pairs of the sales and taking the log difference, or taking the log of the division of the sale pairs. This is explained below.

$$\ln(\text{Sale2}/\text{Sale2}) = \ln(\text{Sale2}) - \ln(\text{sale1})$$

Then, the dummy variables are created and these dummies are equal to -1 for the year of the first sale and 1 for the second sale with 0's for all other years in the dataset. Finally, the regression is run using all dummy time variables and all of the explanatory variables. The housing index can then be constructed from these coefficients for each year of the home sales to see how the market has moved up or down and be confident that the analysis takes into account the differences in housing prices within each housing sale. The explanatory variables can also be utilized to understand the housing price in the same manner that is used in the hedonic and difference in difference methods. A common choice to make with a repeat sales model is to determine what time period to use. If the data set is rich enough, it is potentially possible to compute an index for each week of the year. But if your dataset is smaller then the research must widen the time periods into months, fiscal quarters, or even into years. For this data set the year time interval was chosen for sufficiency of the data to have enough sales per time period to create the index, and for computing power restraints. At 16Gbs of RAM the computer used for this research was unable to handle any more time periods than at the year bands and further there are periods in the data set in which no repeat sales occured in a fiscal quarter, and therefore the decision was made to use a yearly index rather than another method. Note that for all analysis in this paper the log transformed versions of the housing price was used as the dependent variable. As stated previously there is no hard and fast rule regarding when to use levels, Box-Cox, or logs however logs are chosen for this paper since they are essential to the repeat sales model.

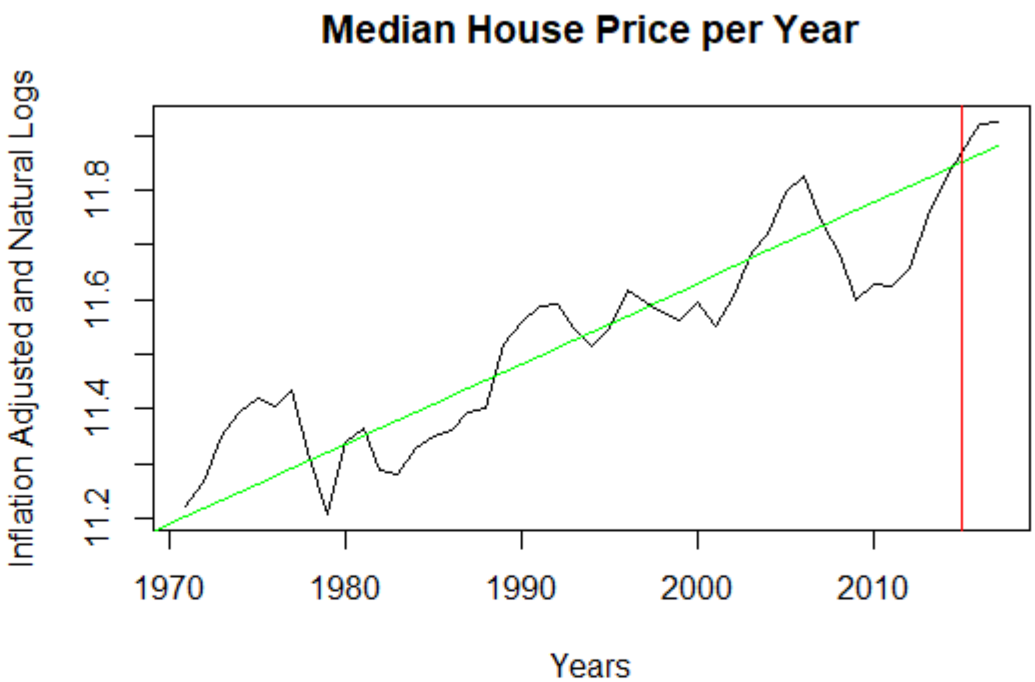
When the repeat sales model is run, there are 76151 observations in the repeat sales method as opposed to 107408 observations in the main model. Once results for the overall repeat sales model are obtained, the next goal is to compare the results of the repeated sales model to the homes in the hedonic regressions because these are the homes that allow this paper to compare these two methodologies equally. The mean and standard deviation of the inflation adjusted sale price are taken from the repeat sale regression and then a subset of the hedonic regression is run where all homes within two standard deviations from the mean are included. This methodology allows us to capture all homes that are similar in price to the ones that are included in the repeat sale regression but that may have not been sold more than once therefore comparing equally distributed data sets. There is some unexplained reason that the rest of the homes have not been sold more than once and this method has the added benefit of excluding those so they do not bias the results. Finally, the results of all of the regressions can be compared for the following purposes: the hedonic Regressions with all of the observations vs. the hedonic Regressions with the distribution from the repeat sales can give us an estimation of differences between those homes sold more than once and those only once. This has the benefit of showing if the repeat sales model is biased due to including many of the starter homes that were discussed in the previous literature review. The below table gives a reference of the price of homes based on the number of sales(note that for clarity these prices are in 2017 dollars to be easily referenced):

Number of Sales	Mean	Median	Observations
2	$283,192	$249,060	27,169
3	$257,678	$215,858	24,011
4	$254,792	$214,841	16,014
5	$243,721	$203,754	7,223
6	$231,674	$204,813	1,734

Clearly, it can be noted that the issue of the starter homes biasing the data in a repeat sales model is valid and it is indicated by the decreasing price for each band of home sales. From here, the values of each regression can be interpreted with this knowledge and better understood the meaning for each of the coefficients.

5. Results:

Before discussing the results, it may be worth again reviewing the median home price from the 1970s to present. In the figure below, a trend line has been added to show the trend from the 1970 to 2000. This allows an inspection of the entire Washoe County housing market and the trends since housing boom, bust, and recovery cycle.



This preliminary graph shows a trend line in green and a line at 2015 in red. Essentially, if this trend line shows the true rate of growth for Washoe County, then you can see the bubble peeking over the trend line, then the recession taking a huge dip with finally increasing back over the trend line in 2015. This means that the WC housing market was massively undervalued in the great recession and renters and home buyers have been the outsized winners in this housing market until around 2015. This is when the housing price peeked its head over the trend and we are now in an overvalued market if this is true. It is important to note this is a simply one variable regression analysis using years and median home price. This does not take individual homes or characteristics into account. However it does provide an overview of the market which can provide some context as the results are reviewed.

Equation (1) below shows the regression I ran to find the price gradient from CBD. The regressions are pooled by time and geography. Separating by left and right of CBD gives us an estimation of the difference in gradient for homes that are closer to USA Parkway. Separating by time allows us to analyze the construction process and shifts as of late.

$$\text{Ln(Sale Amount)} = \alpha + \beta_1(BldgSF) + \beta_2(LandSize) + \beta_3(OnrampDist) + \beta_4(OnrampDist^2) + \beta_5(Owner) + \beta_6(Baths) + \beta_7(Mortgage30) + \beta_8(Unemployment) + \beta_9(Age) + \beta_{10}(Age^2) + \beta_{11}(ArchDist) + \beta_{12}(TwoStory) + \beta_{13}(East) + \epsilon$$

The table below shows the coefficient on the variable for the log distance from the Reno Arch. This variable allows an estimation of the housing prices as one gets closer to USA parkway, or away from the Reno Arch. Recall for all years between 1990 and 1998 the coefficient for those homes to the east of the reno arch are negative or insignificant. In the early 2000s however they begin to pick up in significance and become permanently positive or insignificant for all years afterwards. The year 2002 does however show results that are counter intuitive to the hypothesis. Looking through the data there was a huge increase in home sales in that year which could explain the negative value if there were large outliers that biased the data. An increase from 2712 sales to 3450, so potentially there were a large number of sales of a specific type that biased the results. In 2012 there is a large spike in the price gradient which lasts until 2014, and then decreases. Notice starting in the year 2014, when Tesla announced Sparks as the site of their gigafactory, the price gradient is noticeably lower than in the previous recovery years. This likely means that the entirety of the Washoe County housing market was increasing, and the CBD was increasing faster than eastern Washoe County than it was from 2011-2014. This is easily double checked by looking at the median price of the homes in Sparks located 4 miles from CBD and homes more than 10 miles away. The homes less than 4 miles increased from $176,810 in 2014 to $252,057 in 2017 or 41% increase in price. Whereas the homes more than 10 miles away increased 15%. This vast disparity is what causes the gradient coefficient to drop in the exact years we would expect an increase.

Table 1
Gradient Coefficient for regressions by each individual year for all years 1998-current.

Time	All	Observations	West	Observations	East	Observations
1998	-0.0482*	6498	-0.1128*	2650	0.01962	3848
1999	-0.0281	7352	-0.0982**	3099	0.1088*	4253
2000	-0.02238	7129	-0.13146*	3268	0.10380*	3861
2001	0.01433	9065	-0.1340**	3679	0.10468*	5386
2002	-0.6559***	10448	-0.08103**	4309	-0.9877***	6139
2003	-0.09440***	11344	-0.1442***	4512	0.1957***	6832
2004	0.01517	12466	-0.0557	5242	0.14063***	7224
2005	-0.01369	12318	-0.09118**	4896	0.085249**	7422
2006	0.01270	9293	-0.099***	3744	0.14641**	5549
2007	0.02559	7120	-0.08694**	2742	0.11504***	4378
2008	0.0316	8566	-0.08811*	3212	0.14015***	5354
2009	-0.1564***	10081	-0.3273***	3747	0.1317***	6334
2010	-0.0219	9514	-0.1773***	3656	-0.11311	5858
2011	-0.3244***	10436	-0.1788***	4298	0.392***	6138
2012	-0.0411*	9379	-0.1622***	3650	0.29489***	5729
2013	-0.03844*	9099	-0.1361***	3703	0.23133***	5396

2014	-0.0176	8964	-0.09488***	3427	0.20291***	5537
2015	0.00301	9615	-0.04165**	3757	0.1337***	5858
2016	-0.11021***	9420	-0.2497***	3565	0.07076***	5855
2017	-0.0375**	4110	-0.1233***	1559	0.13869***	2551

Table 2

Hedonic and Repeat Sales models to compare “starter home” issue.

Table 2	Hedonic A	Hedonic B	Hedonic C	Hedonic D	RS
Intercept	12.2063160***	11.12841464***	11.80282278***	11.315372901***	-0.247440373***
Bldg SF	0.00032699***	0.000315597***	0.000287139***	0.000292480***	0.000051148521***
Land Size	-0.00000164***	-0.00000040247*	0.000000693***	0.000001689***	0.000006765318***
Dist to Onramp	0.000033074***	0.0000333522***	0.00001481612***	0.00001645964***	-0.00001943927***
Dist to Onramp Squared	0.00000000083***	-0.0000000007***	-0.0000000005***	-0.0000000004***	0.0000000011***
Owner	0.0007200712***	0.00018979384811	-0.000532707***	0.0002721271***	-0.00491773***
Baths	0.0565169853***	0.06136587764425***	-0.080863742***	0.0830688580***	0.0621499***
Mortgage 30	-0.0927410969***	0.07709605967166***	-0.0808131216***	0.0031988922	-0.011889341**
Un- employment	-0.067570262***	-0.00102133533994	-0.071525441***	-0.045774848***	-0.0066039***
Age	-0.0257147943***	-0.02481551862323***	-0.003247496***	-0.002855523***	-0.02094244***
Age^2	0.00026730672***	0.00025295280955***	-0.0000064432**	-0.000012784***	0.000822411***
Arch Distance	-0.000008869***	-0.00001000857240***	-0.0000073931***	-0.000008848***	0.0000067946***
Two Story	-0.0850900259***	-0.08714956449700***	-0.044830782***	-0.050125163***	-0.10731116***
East	-0.0560449896***	-0.06255829810534***	-0.047398614***	-0.051929629***	0.000971640742
Observations	208400	208400	183021	183021	67270
Full Data Set	Yes	Yes	No	No	No

RS $\mu + 2$□	No	No	Yes	Yes	Yes
Year Controls	No	Yes	No	Yes	Yes

Table 1 shows the results of the hedonic approach using different datasets and models that are comparable to the repeat sales model. This allows a comparison of the hedonic results with all of the sales with a reduced dataset comprised of the sales in the repeat sales model. Hedonic regression A uses the full data set with no controls for individual years. Hedonic B also uses the entire data set, but with controls for each individual year that allow a greater specificity in the model to control for any outlying years as well as to compare with the repeat sales estimates. Hedonic C and D both use a reduced data set that is less than the average of the repeat sales model on the far right column. This allows for an analysis of the "starter home issue" wherein the data set is potentially biased downwards due to the higher number of starter homes in the repeat sales model, and a lack of expensive homes that are not bought and sold as often. Due to these classifications it is now possible to see the differences between starter home datasets and full data sets as well as how controlling for each year makes a difference between the repeat sales model and the hedonic model.

The following areas are interesting results and deserve to be discussed. The coefficient for the variable Land Size is negative for the regressions A and B, but positive for all else. This indicates that the difference between the two datasets results in the possibility that there are some observations in the larger dataset that result in this change of signs. While there is no reason to believe a home price will go down because of a larger lot size, the magnitude of this coefficient indicates it does not contribute a large amount to the home price. The median land size in the dataset is 7,362 square feet, or under □ of an acre of land. Therefore at the typical land size, the home decreases in value by only 1.21%. Since there are so many observations in the dataset, it could be that everything has significance because the standard errors are so small.   For regression A both of the coefficients are positive indicating that housing price increases as the distance from the onramp increases and notably this happens at an increasing rate. For regressions B, C and D, the coefficient of the squared term is negative and the original term is positive meaning there is a optimum housing price at a certain distance away from the onramp. According to regression B, the optimum distance is approximately 12 miles away, and for C, and D, the optimal distance is around 8 miles away. Both of these values seem large especially considering the median distance from an onramp is approximately 3 miles. So there are outliers that must be biasing this result.

Finally, the repeat sales model flips the sign and shows that the home price is most hurt at near the freeway. Again, the result of being right near the freeway is suspect, but seeing the house is harmed at that distance makes more economic sense than the hedonic models. For the coefficient of the mortgage rate, the coefficients of hedonic regression B and D are positive which does not make intuitive sense. These regressions contain year controls so what effects the change of the mortgage rate could be captured by the year effects and then multicollinearity could be an issue with this variable. The coefficient of the two story home variable is consistently negative across all of the regressions. This means if the home is a two story building then the value of the home becomes lower, all else equal. This is intuitive because a two story house is larger than most one stories, so a one story home with the same square footage would have to be an expansive home. Additionally, a one story home is desirable for those who move to

Washoe County as retirees and do not want to handle stairs as they age. This is likely why the coefficient is negative. Finally, the hedonic regressions show the homes east of the Reno Arch are valued less than those on the west and the repeat sale model shows insignificance. Therefore the homes west of the arch are typically more expensive all else equal.

It is important to remind the reader that the dependent variable is different between the RS model and hedonic model, since the hedonic model has larger dependent variables (just the log of the house price) as opposed to the RS model which is the logged difference which has the ability to contain negative values. This is the case for many homes where there is a sale in the bubble (2005-2007) and one in the subsequent crash of the Great Recession (2008-2011). One would hope that regardless of this difference in dependent variable, the coefficients of the variables would maintain their sign and significance, while simply reducing their magnitude closer to zero. However, since this is not the case, the hedonic and RS models are then not to be trusted entirely on their own and the difference in difference models below will give more confidence to the results.

Table 3
Difference in Difference Results

Table 3	DiD - All	DiD - Less than $200,.000	DiD $200,001 - $400,000	DiD $400,001 - $600,000	DiD More than $600,000
Intercept	12.12609***	10.570597***	11.91353***	12.15444***	12.26232***
Bldg SF	0.000332***	-0.00002***	0.0001389***	0.0000469***	-0.00016***
Land Size	-0.0000006**	0.0000022***	-0.00000004	0.00000035***	-0.0000093***
Dist to Onramp	0.000032***	0.000011***	0.0000014**	-0.00000053	0.0000161***
Dist to Onramp Squared	-0.0000000006***	-0.0000000002***	0.00000000001	-0.000000000018	-0.0000000003
Owner	0.000349*	0.00062***	-0.000175**	-0.0001292*	0.00451***
Baths	0.0665594***	0.163557***	-0.005043**	0.01002***	-0.07878***
Mortgage 30	-0.072476***	0.01869***	-0.052529***	-0.0113531***	0.217062***
Un-employment	-0.0669932***	-0.020286***	-0.02668***	-0.00941***	0.06319***
Age	-0.0299908***	0.00193***	-0.00154***	0.000639***	-0.03072***
Age^2	0.0003264***	-0.000029***	0.000015***	-0.000006**	0.0003826***
Arch Distance	-0.0000083***	-0.0000019***	-0.00000145***	0.000001***	-0.0000018
Two Story	-0.091072***	0.01591**	-0.028462***	-0.02787***	0.071846***
Time	-0.04363***	0.094491***	-0.097645***	-0.03577***	0.628761***
East	-0.020814***	-0.031501***	-0.0003764***	-0.01216***	0.10500***
DiD	-0.116533***	-0.06306***	0.024375***	-0.02319***	-0.328816***
Observations	171,710	45,704	78,076	23,028	24,902

<u>Difference in Difference</u>

Utilizing the same equation from the hedonic regressions, there are interesting results for the difference in difference analysis. When the year of 2014 is used at the treatment year there is a decrease in price associated with the year and geography. This means if we view 2014 as the year that the tesla hired many workers and our "experiment" began, then many homes closest to TRIC reduced in value, but not the median priced homes. Which is fascinating, given that region should see a consistent increase in prices due to the Industrial Center due to gains in employment, however this increase is only in the median homes and not in the very inexpensive or very expensive homes. Using 2014 as the time variable also gives a good treatment time since the biggest tenant of TRIC, Tesla, made their announcement public at that time and no one could have supposed with certainty they would move there prior to their announcement. The results for the DID analysis are below:

Notice the negative sign and significance for the DID variable in the regression for all homes. This shows if this experiment design is valid, those homes east of the Arch suffered since the announcement of Tesla moving into the Area. Because of this surprising result, more regressions were computed in order to see if this result was consistent across all home prices. Potentially, the announcement of Tesla to TRIC caused only certain kinds of homes to increase and a full analysis of this is required. In the table above, results are shown for the entire dataset, and for $200,000 bands of house prices (recall these are logged prices with inflation accounted for and so for homes with a log price of 12.206 at $200,000 and so on for each of the bands of prices). The coefficients change depending on whether the homes are cheaper ($0-$200,000) or more expensive. The coefficients of the cheap homes (<$200,000) and very expensive homes

(>$400,000) are negative and significant. The homes in the middle show positive and significance for the difference in difference variable. This indicates that only median home prices rose in Sparks since the introduction of Tesla in TRIC. For cheap homes (less than $200,000) the coefficient is negative, and this could be that many of those homes have increased in value above $200,000 and therefore are excluded from that regression for recent years. Leaving only the worst and cheapest homes in the regression for the treatment years. Those homes more expensive than $400,000 are negative however. Most interestingly is the DiD coefficient on the homes more expensive than $600,000 has an even greater magnitude. This means that the more expensive the home is in Sparks past $400,000 the more it was harmed according to this experiment design. It could be that these rich homeowners care more about the potential pollution effects, or other reasons that are unknown. But expensive homes in the Sparks region are not increasing due to TRIC in recent years.

A highly interesting result is the way in which the effect of interest rates has on the price. For cheaper homes, the interest rate has a positive effect on home prices. This is counter intuitive but could potentially be explained by a strong economy helping investors buy rental property for those homes that are less than $200,000. But then the median homes become affected negatively by the interest rate. For homes between $200,000-$400,000 a 1 percent increase in the mortgage rate results in a 5% decrease in home price. For homes between $400,000-$600,000 the same increase in mortgage rate is only a 1% decrease in home price. Most astounding is that for homes greater than $600,000 a 1 percent increase in the mortgage rate results in a whopping 20% increase in home price! An explanation for this trend is that as the home price increases, it is more likely that the homeowners work compensation is related to the performance of the stock market or overall economy. If mortgage rates are high this indicates the economy is doing well, therefore these high income individuals can afford a more expensive home due to portfolio and/or business performance, rather than actually being enticed to purchase a more expensive home when it is more expensive to do so. This finding is incredibly interesting to witness in the analysis and gives insight into the differences in the people who have money and do not have money.

Related to this finding is the coefficient of unemployment. For all regressions the coefficient is negative, but for rich homes the coefficient is positive. This finding is odd and the logic applied to the mortgage rates cannot be applied in the same way to the unemployment rate.

Again, the only result that provides a positive coefficient are the homes within the $200,000-$400,000 range. Robustness checks were performed to see if this experiment design is valid and can be reliably used in this paper. To do this the year of 2014 is swapped out with other years to confirm that only 2014 holds a significant value as the DiD coefficient. When using 2013, the overall regression still shows significance and negativity and the $200k-$400k homes showed a reduced coefficient of 0.018 (down from 0.024) with significance at all levels. When using 2012 as the treatment year then finally the mid range home value DiD coefficient reduces in value. Therefore, there could have been increases in employment in 2013 that caused Reno to increase over the rate of Sparks that were not due to Tesla alone. This could have been the announcement of other employment in the area at this time. Then when using 2015 as the treatment yeat this coefficient increases to 0.028, indicating the difference between Reno and Sparks is expanding for these homes, but again the difference for all other homes are negative. So expensive homes are taking off in Reno, but mid range homes in Sparks are seeing a rise at an increasing rate.

6. Conclusion:

The goal of this paper was to study three things: the impact of Tesla and TRIC on the Washoe County housing market, to understand how freeway accessibility affects housing prices, and finally to compare repeat sales to traditional hedonic models and determine if the "starter home" issue exists in Washoe County. Each of these areas are important to study. From an economic point of view, they offer an interesting perspective for policy makers.

First, the freeway accessibility does not seem to be a large issue in Washoe County. All regressions show that an optimum distance from the freeway onramp is relatively far. The optimum distance is approximately 8-12 miles, which is further than the median home distance from an onramp. There is evidence to suggest freeway accessibility is becoming a more important issue over time however. So, as the population of Washoe County increases the transportation costs associated with driving to the freeway to get to your place of employment could be becoming more expensive (i.e. traffic), and therefore the optimum distance from the freeway could be getting smaller in the future. As policy makers and city planners decide where to place a freeway onramp, they should consider who cares the most about those accessibilities. Also, if they reduce transportation costs for a group of homes by construction of a freeway, they must consider who will receive the outsized gains of those homes. If those homes are rented out, then the residents will not receive these gains, but instead the landlords will receive them. If the goal is to increase welfare of a community this topic must be taken into account.

Secondly, the “starter home” issue holds true when analyzing the average and median home prices. As the number of sales of a home increase, the average inflation adjusted price of that home will go down. However, when analyzing table 2, some of the signs of Hedonic regression C change, which is the regression that shares a data set with the same attributes as the repeat sales regression. Because of these differences, one must consider the possibility that the starter homes are statistically significant than the full dataset of homes. Repeat sales are typically used to construct indexes of housing prices and costs at different geographical regions and this paper has shown the repeat sales index is biased downward in different ways that must be accounted for. Because of this, when constructing an index for Washoe County, care must be taken to understand how these starter homes affect the data. And, an index would be lower than the true housing market index.

Finally, the price gradient analysis of Washoe County did not show any obvious gains to Sparks that were not enjoyed by the rest of Washoe County in the years since Tesla announced their move to the Tahoe Reno Industrial Center for home other than median priced homes. In fact, in the years since the Tesla announcement the price gradient has shown that Sparks rose at a lower rate in the past relative to all else. This could mean that either the Central Business District gained value faster than Sparks did, or it could mean that Sparks substantially slowed its housing price increase in recent years. The far more likely reason is that all of Washoe County and specifically downtown/midtown areas increased the most out of any part of Washoe County with additional improvements to the median priced homes in Sparks. While it is impossible with this data set to know exactly what affect the TRIC had on the county, it is that a contributing role of the increase in price was due to a continual recovery from the recession or from retirees moving in from outside states. In the great recession houses became so massively undervalued that some of the recent gains could still be attributed to mean reversion rather than TRIC's impact. In the media reports lately describing the increase in housing prices in Sparks, the reporters are entirely attributing the increase to be due to TRIC, but they are not taking a market correction to the equilibrium price into account. This information is is special importance to lawmakers and policy makers who give tax incentives to companies that come and provide employment in the area. If policy makers should be careful to not give tax incentives to these companies based on any metric involving housing and instead should give them based on employment. It is difficult to impossible to tell the causal relationship between TRIC and housing prices, yet it is profoundly difficult to see the correlation between the opening of TRIC and housing prices and not make the assumption that there is a causal link between the two.

This paper could be benefited heavily by certain improvements. The data from the county assessor's office could have supplied tracked changes to the homes where each time the home was bought/sold the home values were updated to reflect any changes. For instance, if someone built an addition to the home, the value of square footage was increased and not accounted for in earlier sales. This could cause issues if many of the homes in the data set have had renovations made in prior years. The study could have also been aided by mortgage data for each home. If it was made clear the details of the mortgage and a price history for each year the analysis could have been incredibly detailed throughout the recession. While this is a substantial amount of personal data on each home, at least if the county provided information on the type of sale i.e. short sale, foreclosure, normal, etc… it would aid the study. Additionally, this paper could have been improved by waiting for more time to lapse since the announcement of Tesla to conducting this study. Potentially, there are effects of the expansion of TRIC that have not yet hit the housing market and we must wait longer to accurately study those effects.

Bibliography

"QuickFacts." *U.S. Census Bureau QuickFacts: Washoe County, Nevada*, www.census.gov/quickfacts/fact/table/washoecountynevada/PST045216.

Antoniucci, V., & Marella, G. (2017). Immigrants and the city: The relevance of immigration on housing price gradient. *Buildings, 7*(4), 91. http://dx.doi.org.unr.idm.oclc.org/10.3390/buildings7040091

"Impact of Highways on Property Values - ASU Digital Repository." https://repository.asu.edu/attachments/74139/content/HighwaysPropertyValues.pdf. Accessed 16 Oct. 2017.

Brownstone. 2003, "Drivers' willingness-to-pay to reduce travel time ... - ScienceDirect.com." http://www.sciencedirect.com/science/article/pii/S0965856402000216. Accessed 17 Oct. 2017.

Brugge, Durant, Rioux. "Near-highway pollutants in motor vehicle exhaust: A review of ...." 9 Aug. 2007, https://ehjournal.biomedcentral.com/articles/10.1186/1476-069X-6-23. Accessed 16 Oct. 2017.

Abbott, Joshua K., and H. Allen Klaiber. "An Embarrassment of Riches: Confronting Omitted Variable Bias and Multi-Scale Capitalization in hedonic Price Models." *Review of Economics and Statistics*, vol. 93, no. 4, 2011, pp. 1331–1342., doi:10.1162/rest_a_00134. https://www.researchgate.net/profile/Joshua_Abbott2/publication/261967645_An_Embarrassment_Of_Riches_Confronting_Omitted_Variable_Bias_And_Multiscale_Capitalization_In_hedonic_Price_Models/links/53e4fe950cf2fb74871104f1/An-Embarrassment-Of-

Riches-Confronting-Omitted-Variable-Bias-And-Multiscale-Capitalization-In-hedonic-Price-Models.pdf

Shimizu, Chihiro, et al. "Housing Prices in Tokyo: A Comparison of hedonic and repeat sales Measures." *JahrbÃ¼cher fÃ¼r NationalÃ¶Konomie Und Statistik*, vol. 230, no. 6, Jan. 2010, doi:10.1515/jbnst-2010-0612.

Clapp, J.M. & Giaccotto, C. The Journal of Real Estate Finance and Economics (1998) 16: 5. https://doi.org/10.1023/A:1007729214108

Bender, Bruce, and Hae-Shin Hwang. "hedonic housing price indices and secondary employment centers." *Journal of Urban Economics*, vol. 17, no. 1, 2 Mar. 1985, pp. 90–107., doi:10.1016/0094-1190(85)90039-7.

Housing Price Gradients in a Region with One Dominating Center Author(s): Liv Osland, Inge Thorsen and Jens Petter Gitlesen Source: The Journal of Real Estate Research, Vol. 29, No. 3 (2007), pp. 321-346 Published by: American Real Estate Society Stable URL: http://www.jstor.org/stable/24888210 Accessed: 13-02-2018 18:17 UTC

Meese, Richard, and Nancy Wallace. "The Construction of Residential Housing Price Indices: A Comparison of Repeat-Sales, Hedonic-Regression, and Hybrid Approaches." *Journal of Real Estate Finance and Economics*, vol. 14, 1997, pp. 51–73., www.census.gov/quickfacts/fact/table/washoecountynevada/PST045216.

Nagaraja, C. H., Brown, L. D., & Wachter, S. M. (2010). House Price Index Methodology. Retrieved from https://repository.upenn.edu/statistics_papers/145

Chau, K.W. and Chin, T. L., A Critical Review of Literature on the Hedonic Price Model (June 12, 2002). International Journal for Housing Science and Its Applications 27 (2), 145-165, 2003. Available at SSRN: https://ssrn.com/abstract=2073594

Bourassa, S. C., Cantoni, E., & Hoesli, M. (2016). Robust hedonic price indexes.*International Journal of Housing Markets and Analysis, 9*(1), 47-65. Retrieved from http://unr.idm.oclc.org/login?url=https://search-proquest-com.unr.idm.oclc.org/docview/1768596581?accountid=452

Landajo, M., Bilbao, C. & Bilbao, A. Empir Econ (2012) 42: 987. https://doi-org.unr.idm.oclc.org/10.1007/s00181-011-0485-9

Hedonic Amenity Valuation and Housing Renovations. Billings, Stephen B. Real estate economics, 09/01/2015. Volume: 43, Issue: 3, Page: 652-682, DOI: 10.1111/1540-6229.12093

Memoria sobre algunos metodos nuevos de calcular la longitud por las distancias lunares y explicaciones prácticas de una teoría para la solución de otros problemas de navegación, José de Mendoza y Ríos, Imp. Real, 1795 - 13 pàgines

Disaggregated Property Price Appreciation: The Mixed Repeat Sales Model Author: Melser, Daniel, Journal: Regional science and urban economics, ISSN: 0166-0462, Date: 09/01/2017, Volume: 66 Page: 108-118

Cohen, J., Blinn, C. E., Boyle, K. J., Holmes, T. P., & Moeltner, K. (2016). Hedonic valuation with translating amenities: Mountain pine beetles and host trees in the colorado front range. *Environmental and Resource Economics, 63*(3), 613-642. http://dx.doi.org.unr.idm.oclc.org/10.1007/s10640-014-9856-y Retrieved from http://unr.idm.oclc.org/login?url=https://search-proquest-com.unr.idm.oclc.org/docview/1771232891?accountid=452

Angrist, J. D.; Pischke, J. S. (2008). *Mostly Harmless Econometrics: An Empiricist's Companion*. Princeton University Press. pp. 227–243. ISBN 978-0-691-12034-8.