A HEDONIC REGRESSION ANALYSIS OF HUMBOLDT COUNTY PROPERTY DATA INTEGRATING THE EFFECT OF THE TSUNAMI EVACUATION BOUNDARY ON REAL ESTATE PRICE

By

Scott W. S. Harris

A Thesis Presented to
The Faculty of Humboldt State University
In Partial Fulfillment of the Requirements for the Degree
Master of Business Administration

Committee Membership
Dr. Harinder Singh, Committee Chair
Dr. Michelle Lane, Committee Member
Dr. David Sleeth-Keppler, Graduate Coordinator

December 2015
ABSTRACT

A HEDONIC REGRESSION ANALYSIS OF HUMBOLDT COUNTY PROPERTY DATA INTEGRATING THE EFFECT OF THE TSUNAMI EVACUATION BOUNDARY ON REAL ESTATE PRICE

Scott W. S. Harris

Using Multiple Listing Service real estate transaction data (n = 4,059, n = 4,352 and n = 9,765) spanning nine years, from January 2005 through September 2014, multiple linear regression analyses were conducted to obtain housing characteristic coefficients statistically correlated with final selling price for residential real estate in Humboldt County, California. The real estate attribute coefficients were utilized as inputs to develop a hedonic pricing model for Humboldt County residential properties. By analyzing the effect that common housing attribute independent variables exert on the dependent variable of selling price, regionally-appropriate significant predictors of real estate value were calculated.

In addition, the correlation between selling price and the location of a property within the Humboldt County Tsunami Evacuation Zone boundary delineated by the Redwood National and State Parks was analyzed. The location of a property within the Tsunami Evacuation Zone was associated with a reduced market price, with the average property within the tsunami zone selling for approximately $38,800 less than properties located outside of the tsunami zone. While a negative correlation was observed between
the tsunami zone and the market price, the correlation was generally not determined to be statistically significant for the analyzed dataset.
ACKNOWLEDGEMENTS

I would like to acknowledge and thank the Humboldt Association of Realtors and the Redwood National and State Parks for their integral data contribution to this project. I would also like to thank the Humboldt State University College of Professional Studies School of Business faculty for their technical support in all aspects of project completion. I would like to thank Amber Shows who contributed vital Graphic Information System capability to this project which made primary data analysis possible for this study.

I would also like to recognize Dr. Lori Dengler, professor of geology at Humboldt State University, as well as Vicki Ozaki and Judy Wartella, with the Redwood National and State Parks, for their input to this project and for their tsunami preparedness work which benefits the communities of Humboldt County. In addition, I would like to thank Dr. Patricia Harris, professor of nursing at Dominican University of California, for her technical review and support of this project.

Lastly, I would like to recognize Amber Harris, my wife and friend, for her strength and all the hard work which has made my pursuit of higher education possible. Thank you, Amber, for your devoted and loving support of our family.
# TABLE OF CONTENTS

ABSTRACT ........................................................................................................................................ ii

ACKNOWLEDGEMENTS .............................................................................................................. iv

LIST OF TABLES ........................................................................................................................ vii

LIST OF FIGURES ...................................................................................................................... viii

INTRODUCTION ........................................................................................................................ 1
  Hypothesis .................................................................................................................................. 3
  Project Environmental Setting ............................................................................................... 3
  Real Estate Value Hedonic Regression Modeling ..................................................................... 5
  Humboldt County Real Estate Data Collection and Modelling ............................................... 6
  Qualitative and Quantitative Regression Model Externalities .............................................. 7
  Tsunami Risk and Inundation Zone Mapping ......................................................................... 9
  Tsunami Zone Boundaries Within Humboldt County .............................................................. 9
  Hazard Boundaries and Real Estate ....................................................................................... 11
  Temporal Limits ..................................................................................................................... 12
  Geospatial Constraints .......................................................................................................... 13
  Governmental Boundary Formalization .............................................................................. 13
  Incorporation of Tsunami Zone Variable into Regression Model ......................................... 14

METHODS .................................................................................................................................. 16
  Data Management ................................................................................................................... 16
  Dataset Systemization and Data Exclusion .......................................................................... 18
  Tsunami Zone Variable Calculation and Regression Analysis Integration ......................... 20
LIST OF TABLES

Table 1: HAR MLS Final Data Set Summary Statistics ................................................... 18

Table 2: Humboldt County Property Hedonic Model Characteristic Data .................... 19

Table 3: Coefficients for Regression Analysis Array Excluding Tsunami Zone (First Array) ................................................................................................................................ 25

Table 4: Coefficients for Regression Analysis Array Incorporating Tsunami Zone and Lot Size (Second Array) ................................................................................................................... 27

Table 5: Coefficients for Regression Analysis Array Incorporating Tsunami Zone and Excluding Lot Size (Third Array) .............................................................................................. 29

Table 6: Date of Sale Regression Analysis Coefficients ................................................ 32

Table 7: Sub-Area Regression Analysis Coefficients .................................................... 34

Table 8: Tsunami Zone Variable Quantification ............................................................ 35
LIST OF FIGURES

Figure 1: Average Real Estate Selling Prices Over Time ................................................. 30
Figure 2: Average Sales Price by Location (Sub-Area) ................................................ 33
Figure 3: Tsunami Evacuation Map – Humboldt Bay Region ......................................... 49
Figure 4: Tsunami Evacuation Map – Orick Subregion .................................................. 50
Figure 5: Tsunami Evacuation Map – Big Lagoon Subregion ....................................... 51
Figure 6: Tsunami Evacuation Map – Trinidad Subregion .......................................... 52
Figure 7: Tsunami Evacuation Map – Moonstone Subregion ....................................... 53
Figure 8: Tsunami Evacuation Map – McKinleyville Subregion .................................. 54
Figure 9: Tsunami Evacuation Map – Arcata Subregion ............................................ 55
Figure 10: Tsunami Evacuation Map – Arcata Bay Subregion ..................................... 56
Figure 11: Tsunami Evacuation Map – Manila Subregion ............................................ 57
Figure 12: Tsunami Evacuation Map – Samoa Subregion ......................................... 58
Figure 13: Tsunami Evacuation Map – Eureka Subregion ........................................... 59
Figure 14: Tsunami Evacuation Map – South Bay Subregion ...................................... 60
Figure 15: Tsunami Evacuation Map – King Salmon & Fields Landing Subregion .......... 61
Figure 16: Tsunami Evacuation Map – Eel River Subregion ......................................... 62
Figure 17: Tsunami Evacuation Map – Shelter Cove Subregion .................................. 63
INTRODUCTION

A tsunami can be catastrophic event for a coastal area, causing flooding, major infrastructure and economic damage, injury or death. Tsunamis are typically caused by offshore earthquakes, thus Humboldt County, as coastal area that is one of the most seismically active in the continental United States, is particularly exposed to tsunami risk (Dengler, L., Hemphill-Haley, M., Felton, V., Monro, A., and Warren, J., 2011).

In an effort to mitigate anthropocentric safety risk resultant from potential tsunami events, many coastal communities have designated tsunami hazard and evacuation zones which define areas at risk from tsunami impacts. However, little is known about the relationship between the establishment of a tsunami zone and the subsequent economic impacts of the classification of such hazard areas, specifically the effects on the market value of homes within such zones.

Only a paucity of research assesses the magnitude of the impact that tangible and intangible residential property characteristics exert on real estate market prices in Humboldt County, California. To address this scarcity of regionally-specific studies calculating the impact of such characteristics, including location within a tsunami zone, a hedonic regression model was developed. Physical property attributes and geospatial data to evaluate and rank the effect of locally-appropriate significance of residential real estate characteristic independent variables on the dependent variable of final selling price. This study was completed at the invitation of the Humboldt Association of Realtors (HAR) in
conjunction with the Humboldt State University (HSU) College of Professional Studies
School of Business.

The structure of this hedonic regression model is built upon secondary
independent variables consisting of quantifiable housing characteristics, including total
residence square footage and number of specific rooms, as well as qualitative data, such
as standardized descriptors of aesthetic appearance. When combined, the qualitative and
quantitative variables are cumulatively expressed as the dependent variable of real estate
market price.

Multiple regression analyses were performed using of residential properties sold
in Humboldt County between 2005 and 2014 (n = 4,059, n = 4,352 and n = 9,765) to
generate a Humboldt County hedonic pricing model. This model generates a weighted
value for each of the measured housing characteristics to describe the relative influence
of that independent variable on the dependent variable for all analyzed real estate.

This study expands on the above-described hedonic pricing model by addressing
an important aspect of coastal proximity: the economic effect of tsunami inundation
hazard boundaries on real estate value. To quantify the monetary impact on property
selling price associated with the delineation of the Humboldt County Tsunami
Evacuation Zone boundary (tsunami zone), developed by the Redwood National and
State Parks, primary geospatial data inputs were generated using the hedonic regression
model to incorporate the impact of the tsunami zone on the selling price of Humboldt
County residential real estate.
Hypothesis

The hypotheses tested for this study are as follows: the final selling price of residential real estate within Humboldt County is affected (increased or decreased) in a statistically significant manner as a result of the location of such real estate within the area defined by the tsunami zone boundary; versus the null hypothesis that the selling price of a property is unaffected, or not significantly affected, by the property’s location within the area enclosed by the tsunami zone.

Project Environmental Setting

Humboldt County is located on the northern Pacific coast of California, approximately 270 miles (440 kilometers) north of San Francisco and approximately 100 miles (160 kilometers) south of the Oregon border. Humboldt County is a rural forested region with the largest population centers generally situated in close proximity to Humboldt Bay, one of California’s largest coastal estuaries (Gutierrez, Redfern, and Richgels, 2012; Barnhart, Boyd and Pequegnat, 1992). Humboldt County’s western border is delineated by the Pacific Ocean and is bounded to the east by the Klamath Mountains and Trinity Alps.

The rugged coastline and dynamic topography of Humboldt County are products of the offshore and onshore geologic forces exerted by the conjunction of nearby tectonic plates (Dengler, Hemphill-Haley, Felton, Monro, Warren, 2011). The area of Northern California encompassing Humboldt County is one of the most seismically active in the
United States as a result of its close proximity to the convergence point of three tectonic plates known as the Mendocino Triple Junction and the tectonic plate collision boundary known as the Cascadia Subduction Zone (CSZ) (Dengler, et al., 2011). The historic geologic processes of this region produced a natural topographical obstacle to extensive anthropogenic development of Humboldt County. Due to Humboldt County’s limited population growth and urbanization since its founding in 1853, the County is designated as a nonmetropolitan (micropolitan) area (USDA, 2015).

Due, in part, to its geographic isolation and measured growth, Humboldt County remains generally verdant and bucolic, with forested land comprising approximately 80 percent of the county’s 2.3 million acres (9,300 square kilometers). Humboldt County’s population is primarily clustered into a small number of urbanite subareas, including seven incorporated municipalities, generally distributed along the United States Highway 101 transportation corridor and/or in close proximity to Humboldt Bay, an economic focal point for the region (County of Humboldt, 2013). As defined by the United States Census Bureau (USCB), the City of Eureka represents the largest urban cluster area within Humboldt County (USCB, 2015). The micropolitan urban clusters represent a significant portion of the housing and labor-market for Humboldt County’s population of approximately 135,000 residents. Humboldt County’s populace resides in approximately 62,000 housing units and maintains a homeownership rate of 56 percent (USCB, 2014). Nearly 70 percent of available housing units evaluated as part of this study were located within the few coastal urban clusters with the remainder distributed throughout the rural portions of the county.
Housing values in Humboldt County generally reflect the region’s weak economic forte. The median value of owner-occupied housing units in Humboldt County is estimated at $288,300, a value approximately 20 percent below the median housing price for California (United States Census Bureau, 2014). Due to the rural setting and small scale of the housing market, there has been few published real estate price analyses specific to Humboldt County (Busey, 2014). To address this lack of housing data, our team developed a residential real estate hedonic pricing model for Humboldt County. This analysis generally employed the hedonic regression modeling principals and methods, as described in the following sections.

Real Estate Value Hedonic Regression Modeling

Real estate market values are dependent on the characteristics of the properties being sold. In order to estimate the selling price for property with a given set of characteristics, or develop a price index for real estate within a given area, it must be known how each of the property’s’ qualities affect the overall market value (Sirmans, Macpherson and Zietz, 2009). The differences in various properties can potentially make estimation of market value difficult; however standard regression techniques can be used to estimate price indexes (Bailey, et al., 1963). The following sections discuss specifics of regression modeling relevant to this project.
Humboldt County Real Estate Data Collection and Modelling

The HAR is a membership organization for real estate professionals operating in Humboldt County, which maintains multiple listing service (MLS) data on all county property sales transactions. By performing an empirical study of secondary MLS residential real estate data collected over nine years for properties located within Humboldt County, California (county), our team quantified the magnitude of the influence various housing characteristics exerted on real estate selling price using standard hedonic regression analysis. These data were interpreted and the relative influence each housing attribute applied to the selling price was analyzed and ordered, allowing our team to develop a standardized index incorporating the numerous weighted property attribute variables.

Significant study has concentrated on the use of hedonic regression to assess real estate market value; however little study has focused specifically on pricing data for Humboldt County. As reported by the HAR, there have been no published regression analyses performed on real estate data applicable explicitly to Humboldt County, thus a normalized average for property attributes is needed for use in developing a localized price index. To address this data gap, the HSU project team analyzed residential property sales and attribute data compiled by the HAR.
Qualitative and Quantitative Regression Model Externalities

Neighborhood and general location characteristics, beyond quantifiable housing features, which affect the market price of residential real estate, including the qualitative valuation of streets, surrounding houses, green areas, neighbors and proximity to workplaces and services, are called housing externalities (Rossi-Hansberg, Sarte, Owens, 2010). Property-specific attributes that influence market price include a property’s physical characteristics, such as the number of rooms, number of bathrooms and lots size, but additional qualitative externalities cumulatively have as significant an effect on the price of residential real estate as the standard quantitative characteristics of such properties (Kain and Quigley, 1970). Quantitative property characteristics and housing externalities are collective determinants of real estate market price.

The challenge of isolating and measuring the cumulative qualitative and external characteristics of coastal real estate which affect real estate value, such as viewscape, is compounded by the highly correlated coastal risk characteristics, such as water inundation hazards, for such areas (Bin, Crawford, Kruse and Landry, 2008).

A further consideration when evaluating the consequences of tsunami zone establishment on coastal property is the effect of governmental regulations on real estate value. The Natural Hazard Disclosure Act (AB 1195), enacted in California in 1998 and codified in Section 1102 of the California Civil Code, requires that property sellers provide written disclosure to buyers when the property being sold lies within a mapped hazard area. Research evaluating the effect of AB 1195 on real estate value has shown
that the average sales price of a home within a flood hazard area was decreased by 4.2 percent compared to properties outside the flood hazard area (Troy and Romm, 2004). Additional research evaluating the effect of seller disclosures on property value found a four percent decline after the commencement of such disclosures for real estate located within flood zones (Pope, 2008).

While some regulation decreases real estate value, as noted above, other studies conducted to measure the influence of governmental regulation on real estate value have confirmed that regulation raises housing prices (Malpezzi, 1996). As coastal communities in California are subject to regulatory constraints that inland areas are not, including the California Coastal Act, enacted in 1976 and codified in Section 30000 of the California Public Resources Code, the effect of coastal-specific regulation may represent an additional externality diluting the impact of tsunami zone establishment on selling price.

The objective attributes of real estate are readily quantified and their impact on selling price confidently ranked. In contrast, a property’s qualitative, interdependent and/or intangible features, such as viewscape, applicable regulatory jurisdiction and housing externalities, pose a challenge when attempting to incorporate such attributes into price indices, due to the difficulty of isolating such variables from other externalities and controlling for the multicollinearity inherent in overlapping property and location characteristics.
Tsunami Risk and Inundation Zone Mapping

Tectonic activity, such as fault movement, associated with the sea floor may result in the vertical displacement of overlying water, causing a series of water waves, or surges in a potentially catastrophic event known as a tsunami (Jain, 2005). Evidence from historical events and modeling of potential events indicate that tsunamis represent a hazard to populations and infrastructure distributed along the California coastline (Wood, Jones and Spielman, 2013). In response to the potential hazard of tsunami landfall, the California Governor’s Office of Emergency Services (CalOES), formerly known as the California Emergency Management Agency (CalEMA), and the State of California Geological Survey (CGS), in conjunction with academic and government agency partners, developed a tsunami-inundation zone map of populated areas of the California coastline to identify areas of potential water inundation during a tsunami event. The tsunami inundation zone incorporates low-lying areas in 20 coastal California counties that are at risk of flooding in the event of a tsunami in an effort to support tsunami preparedness and risk-mitigation in California (CalOES, 2009; Wood, et al., 2012). The tsunami zone boundary represents the area of land likely to experience flooding up to a maximum expected limit.

Tsunami Zone Boundaries Within Humboldt County

As a rural population center with over 100 miles (170 kilometers) of Pacific Ocean coastline located in proximity to the seismically active offshore CSZ, Humboldt
County is at risk of tsunami inundation and, thus was included in the Redwood National and State Parks tsunami analysis. The Humboldt County tsunami inundation analysis was completed by dividing the county coastal areas into 13 subregions, also referred to as mapped regions. The mapped regions are graphically depicted on Figures 3 through 17: Tsunami Evacuation Maps (Figures 3 – 17). Figures 3 – 17 are located in the report appendix.

As shown on the appended figures, the mapped regions do not include the entire length of the Humboldt County coastline, instead comprising only those areas of coastal land associated with the greatest population and infrastructure density. The mapped regions include the a range of littoral areas extending north from the Shelter Cove subregion, located at the southern border of Humboldt County, to the Orick subregion, located near the northern limit of Humboldt County (Redwood National and State Parks, 2014). The mapped regions exclude sparsely populated coastal regions, such as those around Cape Mendocino, California’s westernmost point located at the south central portion of Humboldt County.

The tsunami zone is physically demarcated in Humboldt County using standardized warning signage along roadways and coastal access points, alerting entrants to lowland coastal areas of the presence of the tsunami hazard or evacuation zone. Some areas within Humboldt County are equipped with emergency signals providing an auditory alarm which is activated in the event that a tsunami warning is issued by the United States National Oceanic and Atmospheric Administration (NOAA) National Weather Service (NWS) West Coast/Alaska Tsunami Warning Center (WCATWC). The
WCATWC issues tsunami warnings to potentially affected areas around the Pacific basin when seismic and oceanographic data indicate a tsunami risk to such regions (NOAA, 2014).

Hazard Boundaries and Real Estate

The effect of commonly reoccurring and/or seasonal natural hazards on property value, such as flooding, are well-studied, and the housing market has had a significant amount of time to acclimate to associated risk mitigation regulation. In contrast, risk mitigation regulations associated with tsunami hazards are a relatively novel occurrence. The establishment of United States Federal Emergency Management Agency (FEMA) Special Flood Hazard Areas (flood zones) is an example of a hazard boundary formation addressing a relatively common geospatially-confined natural hazard. We may look to flood zone boundaries to generally inform our discussion of the potential economic impacts of tsunami zones on the value of real estate within such areas.

Situation of a property within a flood zone has been shown to negatively impact the price of residential real estate when compared to characteristically comparable properties located outside of such zones (Harrison, Smersh and Schwartz, 2001). Correspondingly, the occurrence of flood events has shown to negatively impact real estate value. Prices for real estate within the 100-year floodplain have been observed to significantly decrease following a flood event; however the negative effect on property values from the manifested flood hazard fades between four to nine years following the inundation event (Atreya, Ferreira and Kriesel, 2013).
While previous study of flooding and flood zone impacts are useful to elucidate the general connection between natural disasters and real estate price, there are several compounding factors which obfuscate the utility of directly correlating the real estate value impacts from flood zones to tsunami zones. The following factors differentiate flood zones and tsunami zones, complicating practical application of flood zone impacts, specifically corresponding to real estate price, to tsunami zones.

Temporal Limits

The establishment of a tsunami zone boundary is a novel convention for Humboldt County when compared with the well-established flood zone delineations. The first iteration of the Humboldt County tsunami zone was mapped in 2004 with additional mapping generated by CalOES in 2009 (Patton and Dengler, 2004; CalOES, 2009). The recent advent of tsunami awareness and hazard delineation has allowed barely a decade for the impact of tsunami zone establishment to matriculate to the housing market.

Formalized flood zones came into existence with the passage of the National Flood Insurance Act of 1968, thus the housing market has had nearly fifty years to acclimate to the establishment of flood zones. Since the effects of tsunami zone institution have had only a fifth of that time to infiltrate public awareness and affect the housing market, the magnitude of the effect of nationally-standardized flood zones may outweigh that of the tsunami zones.
Geospatial Constraints

Impacts from tsunami events are constrained to low-lying areas within close proximity to an ocean coastline, whereas flood events are not so geographically limited. Given the expansive geographical distribution of flood impacts, the larger number of affected properties, and the increased frequency of occurrence, flooding impacts may more heavily weigh on the real estate market, than the impacts from tsunami events. Given the limited geospatial distribution and comparatively low frequency of events, tsunami hazard impact on real estate price may be viewed by the real estate purchaser as more arbitrary and less tangible than that resultant from flood hazards. Therefore, tsunami zone implementation may currently have a more diluted effect on property value than flood zone designation.

Governmental Boundary Formalization

While the federal government has designated some capacity to tsunami risk analysis, including the work of the National Tsunami Hazard Mitigation Program (NTHMP), there has been no national standardization of tsunami hazard evaluation or risk mapping. The establishment of the National Flood Insurance Hazard mapping system, administered by FEMA, has allowed for the normalization of flood zone efficacy across the United States; however, the delineation of tsunami zones is not federally formalized or calibrated. The tsunami zone delineation process has not been commonly standardized; relying on a diversity of inundation models from various academic and
governmental sources. Given the assortment of inundation models development and application, the tsunami zone impact on real estate price may be more inconsistent than is predicted by standardized flood zone determination.

Given the limited data associated with the above external stimuli, conclusions regarding perceived risk and the corresponding effect on property values from location within a designated tsunami zone may be consequently indeterminate and more difficult to accurately isolate and measure than flood zone impacts. Compounding the evaluation of tsunami zone impacts is a scarcity of relevant data calculating the impact of tsunami zone regulatory influence on real estate market price for rural areas within the United States. This data scarcity may be due to the fact that tsunamis which result in tangible economic and property impacts are rare in the United States, as well as the fact that general acknowledgements of tsunami hazards, and the subsequent adoption of tsunami regulatory boundaries, are a fairly recent occurrence.

Incorporation of Tsunami Zone Variable into Regression Model

In order to evaluate the effect of the tsunami zone regulatory boundary on real estate value, the tsunami zone applicability to each property was approximated with a binomial value in order to assess the resultant effect on the selling price using regression analysis. A tsunami zone buffer of 50 feet, extending inland beyond the mapped delineation of the tsunami zone, was added to the analysis to capture false negatives, such a properties located partially within the tsunami zone and/or properties located directly on
the tsunami zone boundary line. The tsunami zone buffer shall herein be incorporated into the definition of the tsunami zone for the purpose of this report.

Properties identified to be within the tsunami zone were assigned a numerical value of one, indicating the presence of a property boundary intersection with the tsunami evacuation zone. A value of zero, indicating the absence of property boundary overlap with the tsunami zone, was applied to properties located outside of the tsunami zone or tsunami zone buffer. The tsunami zone presence/absence binomial variables were added to the HAR dataset for calculation via regression analysis.

Having added the tsunami zone variables to the regression model, the subsets of properties included within and excluded from the tsunami zone were analyzed using hedonic regression analysis. Using this method, the relative impact of the tsunami zone boundary on Humboldt County real estate selling price was derived, as described in the following section.
METHODS

Three general arrays of regression analyses were performed for this project to develop the Humboldt County hedonic pricing model. Each regression array consisted of a collection of multiple regression analyses calculated using unique variable combinations. Multiple regression analyses were necessary to isolate the influence of specific independent variables and to mitigate the strong multicollinearity exhibited by many of the regression coefficients.

The first regression analysis array (n = 4,059) was performed using specific HAR-supplied MLS real estate secondary data and did not include tsunami zone primary data. The second regression analysis array (n = 4,352) was performed using the MLS data in conjunction with tsunami zone binomial primary data calculated for this study. The third regression analysis array (n = 9,765) consisted of MLS and tsunami data and excluded the Lot Acre variable, thereby allowing regression analysis of a data cohort more than twice the size of the first two arrays. The general data management and analysis methodology employed for this project are described in the following subsections.

Data Management

Tabulated Humboldt County MLS property attribute and sales data was used with permission and provided to HSU by HAR under the terms of a HAR confidentiality agreement. The dataset was maintained and analyzed electronically using Microsoft Excel software (version 2010) and IBM Statistical Package for the Social Sciences
(SPSS) software (SPSS Statistics package). The dataset matrix was arranged with the general real estate characteristic categories comprising the dataset vertical columns, while the individual real estate transactions, represented by unique MLS list numbers, were denoted horizontally in the dataset rows. Characteristic data applicable to each MLS list number was populated horizontally in cells corresponding to the relevant data columns.

The secondary MLS dataset provided for this project represented county residential real estate sales transactions from January 1, 2005 to September 18, 2014, including individual real estate transactions datum differentiated by a unique MLS numerical identifier (MLS list number) assigned to each property sale. The original HAR dataset provided to HSU (n=10,296) included invalid, incomplete or irrelevant data, thus such extraneous transaction data, applicable to 444 MLS list numbers, were excluded from the final regression analysis inputs. The final dataset was limited to MLS transactions consisting of operable data (n = 4,059, n = 4,352 and n = 9,765), containing relevant real estate attribute variable data necessary for regression analysis input. During regression analysis the dataset was further reduced by 5,674 MLS list numbers due to the limited number of data points for the Lot Acre variable.

The final MLS dataset consisted of quantitative and qualitative attribute information, as well as sales transaction specifics, including: the date of sale, final selling price and MLS list number. A summary of the final HAR MLS list number, transaction date and dependent variable data is provided in Table 1: HAR MLS Final Data Set Summary Statistics (Table 1).
Table 1: HAR MLS Final Data Set Summary Statistics

<table>
<thead>
<tr>
<th></th>
<th>Sold Date</th>
<th>Sold Price (USD)</th>
<th>House Area (ft.²)</th>
<th>Lot Size (acres)</th>
<th>Year Built</th>
</tr>
</thead>
<tbody>
<tr>
<td>Minimum</td>
<td>01/01/2005</td>
<td>11,000.00</td>
<td>144</td>
<td>0.02</td>
<td>1826</td>
</tr>
<tr>
<td>Maximum</td>
<td>09/18/2014</td>
<td>810,000.00</td>
<td>10,600</td>
<td>310.00</td>
<td>2014</td>
</tr>
<tr>
<td>Range</td>
<td>3,548 Days</td>
<td>799,000.00</td>
<td>10,456</td>
<td>309.8</td>
<td>188 Years</td>
</tr>
<tr>
<td>Mean</td>
<td>09/17/2009</td>
<td>293,022.27</td>
<td>1,578</td>
<td>3.23</td>
<td>1972</td>
</tr>
<tr>
<td>Median</td>
<td>09/03/2009</td>
<td>225,000.00</td>
<td>1,200</td>
<td>1.08</td>
<td>1981</td>
</tr>
<tr>
<td>Mode</td>
<td>07/29/2005</td>
<td>250,000.00</td>
<td>1,100</td>
<td>0.15</td>
<td>1950</td>
</tr>
<tr>
<td>Standard Deviation</td>
<td>-</td>
<td>124,295.68</td>
<td>604</td>
<td>13.24</td>
<td>-</td>
</tr>
<tr>
<td>Number of Data Points (n)</td>
<td>9,852</td>
<td>9,852</td>
<td>9,852</td>
<td>4,059</td>
<td>9,772</td>
</tr>
</tbody>
</table>

Note:
1. - = no meaningful value for the specified entry
2. ft.² = square feet
3. USD = United States dollar ($)

Dataset Systemization and Data Exclusion

Preceding initial regression analysis, the original data matrix was formatted and limited to create the final dataset, including common real estate attribute variables which comprise the hedonic pricing model. The original MLS data matrix included HAR user-defined information fields that were utilized for HAR operational statistics and other supplemental information not related to the description of the real estate being sold. The HSU project team systematized the data and identified the statistically relevant data fields, excluding immaterial and invalid information from the final dataset, arriving at 20
data categories (columns) containing information describing property location
information, sale specifics and building characteristic descriptions.

Of the 20 data categories, 11 were selected as variable inputs for the initial
regression analysis. The MLS variables utilized for the regression analysis are listed in
Table 2: Humboldt County Property Hedonic Model Characteristic Data (Table 2).

Table 2: Humboldt County Property Hedonic Model Characteristic Data

<table>
<thead>
<tr>
<th>Quantitative Characteristics</th>
<th>Qualitative Characteristics</th>
</tr>
</thead>
<tbody>
<tr>
<td>Final Selling Price (Units in US Dollars)</td>
<td>Standardized Appearance Valuation (Optional Descriptors Listed Below)</td>
</tr>
<tr>
<td>Number of Bedrooms</td>
<td>Poor</td>
</tr>
<tr>
<td>Number of Bathrooms</td>
<td>Fixer</td>
</tr>
<tr>
<td>Total Square Footage (units in square feet)</td>
<td>Fair</td>
</tr>
<tr>
<td>Number of Garage Stalls</td>
<td>Good</td>
</tr>
<tr>
<td>Lot Size (units in acres)</td>
<td>Very Good</td>
</tr>
<tr>
<td>Presence/Absence of Second Dwelling</td>
<td>Excellent</td>
</tr>
<tr>
<td>Sub-area (Location Name)</td>
<td>New Building</td>
</tr>
<tr>
<td>Date Sold</td>
<td></td>
</tr>
<tr>
<td>Presence/Absence of Tsunami Zone</td>
<td></td>
</tr>
</tbody>
</table>

For the selected attribute variables denoted in Table 2, some secondary data
contained in the HAR dataset were incomplete or inoperable, thus the dataset was
interpreted and regression calculated using a set of defined assumptions to inform the
data analysis. The assumptions used for this study are described in the following section.

In addition to the traditional hedonic regression model categories described
above, two supplemental data categories (“Tsunami Zone” and “Tsunami Zone Buffer”)
were added to the final dataset to account for the presence or absence of the tsunami zone boundary with respect to the geographic location of the property associated with each MLS list number. The binomial inputs to the tsunami zone data categories captured a property’s inclusion within the tsunami zone, or exclusion from the tsunami zone. The tsunami categories were developed as described in the following section.

Tsunami Zone Variable Calculation and Regression Analysis Integration

Following the calculation of the initial regression analysis array, a second array of multiple regression analyses were performed to generate primary data and isolate the effect of a property’s location, specifically with reference to tsunami zone, on the selling price of that property.

To generate the primary data for the second regression array, the geospatial location of the individual properties contained in the HAR dataset were obtained using Esri ArcGIS (ArcGIS) Online address locator engine. Using ArcGIS software, unique identification points where established at the midpoint of each property, representing that property’s location. The points were arranged spatially on an ArcGIS map layer Shapefile. Once each of the properties from the MLS dataset were plotted on the ArcGIS map layer, the tsunami zone layer, provided by Redwood National and State Parks, was overlaid on the HAR property location data points. The number of intersections of the property location points with the tsunami zone layer was calculated, resulting in an output for the number of individual properties positioned within the tsunami zone. In addition, a second intersection analysis was performed that extended the search for boundary
intersection to 50 feet from the inland edge of the tsunami zone borderline to capture and account for properties that were partially included within the tsunami zone, but were excluded due to the position of the properties’ ArcGIS location identification point outside of the tsunami zone limit and/or excluded due to the location of property directly on the tsunami zone boundary line (i.e. properties with ArcGIS location points located partially within or contiguous to the tsunami zone).

Once the number of properties located within the tsunami zone and tsunami zone buffer were calculated, a binomial output was generated using ArcGIS to differentiate properties that were intersected with the tsunami zone and associated buffer from those located outside the tsunami zone and associated buffer. Properties located within the tsunami zone were designated with the number one and properties located outside the tsunami zone were designated with a zero. The binomial data was incorporated into the HAR dataset analyzed in conjunction with the other independent housing variables using Excel and SPSS software.

The dataset provided to HSU by HAR included properties located outside of Humboldt County, thus properties denoted as "Out of County" were removed from dataset to facilitate comparison of real estate within the study area only.

MLS list numbers with blank entries in the following variable categories were excluded from the regression analyses:

1. Total Square Footage
2. Bedrooms
3. Bathrooms
4. Appearance

Statistical Outliers

Statistical outliers in the data were identified for the sales price category by using the Z-score method. Real estate data were considered outliers if the Z-score of the final selling price was outside the range of $-3.5 < Z < 3.5$ according to the formula:

$$Z = \frac{(Value - Mean)}{(Standard Deviation)}$$

Data points located outside the range of $Z$, therefore designated as outliers, were removed from the dataset prior to regression analysis in order to achieve a more precise evaluation of independent regression coefficient significance.

Data Parameters

The following maximum threshold parameters were set in order to exclude extreme data values that would confound or exaggerate regression analysis results for the specified independent variables:

1. Maximum number of Bedrooms = 10
2. Maximum number of Bathrooms = 10
3. Maximum lot size (acres) = 310 (five properties deleted with >310 acres, deleted property acreage rage from 4300 to 50100 acres)
4. Maximum number of garage stalls = 10

Data points located above the above-denoted maximum thresholds were excluded from the regression analysis.
Assumptions

Our data analysis was informed and governed by assumptions as generally described herein. The primary assumption underlying this study is the accuracy of the secondary data provided to HSU by HAR. Further assumptions generally address the limitations of regression analysis in describing socioeconomic systems, such as the real estate market, as described below.

The discrete variables extracted from the continuum of housing market stimuli imprecisely measure the explanatory variables’ influence on market price as all components of the market exhibit multicollinearity and are approximated when specific variables are segmented and isolated. As such, this study is limited to the evaluation of the real estate attributes and data provided in the MLS dataset and does not represent a comprehensive evaluation of market forces (independent variables) acting on the final selling price (dependent variable). The presence of additional residential real estate variables, environmental externalities and socioeconomic drivers not quantified by the MLS dataset, and consequently not calculated as part of this study, may be assumed to exist in association with the study area. Such externalities may represent important uncalculated real estate value impact(s) beyond the significant market price predictors identified herein.
RESULTS

By performing an empirical study of primary and secondary data, our team quantified the magnitude of the impact various housing characteristics (independent variables) exerted on the final selling price (dependent variable) using multiple linear regression analyses. The data were interpreted and the relative impact each housing characteristic coefficients exerted on the selling price was analyzed and ordered, allowing our team to develop a hedonic pricing model. The results of the regression analyses are described in the following sections.

Multiple Regression Analysis Coefficients

Employing multiple regression analyses to interpret the relative impact each housing attribute exerted on the selling price, a set of dependent variable predictor coefficients was calculated consisting of a numerical value for each of the independent qualitative and quantitative variables which collectively determine the sales price of real estate in Humboldt County. The ranges of specific property characteristic coefficients for each of the three regression cohorts are listed in Tables 3 through 5. The coefficients associated with date of sale are denoted in Table 6. Coefficients linked with the real estate sub-area (property location) are summarized in Table 7.

Coefficients calculated during the first regression analysis array (n = 4,059), dataset excluding the tsunami zone variable, are summarized in Table 3: Coefficients for Regression Analysis Array Excluding Tsunami Zone (Table 3).
Table 3: Coefficients for Regression Analysis Array Excluding Tsunami Zone (First Array)

<table>
<thead>
<tr>
<th>Attribute Variable (n = 4.059)</th>
<th>Coefficient Range (USD)</th>
<th>Coefficient Mean (USD)</th>
<th>Significance Mean</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Minimum</td>
<td>Maximum</td>
<td></td>
</tr>
<tr>
<td>Total Square Footage</td>
<td>101.48</td>
<td>102.05</td>
<td>101.76</td>
</tr>
<tr>
<td>(units in square feet)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Lot Size</td>
<td>2,285.05</td>
<td>2,580.29</td>
<td>2,386.41</td>
</tr>
<tr>
<td>(units in acres)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Number of Bedrooms$^1$</td>
<td>13,372.34</td>
<td>13,372.34</td>
<td>13,372.34</td>
</tr>
<tr>
<td>Number of Bathrooms</td>
<td>11,987.31</td>
<td>61,033.78</td>
<td>33,712.00</td>
</tr>
<tr>
<td>Number of Garage Stalls</td>
<td>15,515.35</td>
<td>22,125.57</td>
<td>19,003.62</td>
</tr>
<tr>
<td>Presence/Absence of Second Unit</td>
<td>22,125.57</td>
<td>56,568.75</td>
<td>40,353.93</td>
</tr>
<tr>
<td>Age of Structure</td>
<td>(8.29)</td>
<td>(14.70)</td>
<td>(11.31)</td>
</tr>
<tr>
<td>(units in years)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Appearance Valuation</td>
<td>23,627.64</td>
<td>23,807.99</td>
<td>23,717.82</td>
</tr>
</tbody>
</table>

Notes:
1. $^1$ = Since there is significant multicollinearity in the explanatory variables, only the positive coefficient is utilized to calculate the mean.
2. USD = United States dollar ($)
3. The above table denotes variables included in the first regression analysis array
4. Negative values denoted in parentheses

As shown in Table 3, the significance mean for each variable is determined to be statistically significant for the first regression array with two exceptions. The second unit and age coefficient means are not determined to be significant. These variables are determined to be significant in two of the three regression calculations performed for the first array, however the coefficient means are influenced away from statistical significance by the correlated variables of square footage (second unit and age variables) and appearance (age variable only).
Additionally, two bedroom coefficients were calculated during the first regression array to be negative values. The negative coefficients are due to significant correlation between the number of bedrooms and all other variables and do not represent the actual influence of this housing characteristic on the market price. Multicollinearity between these variables results in negative regression coefficients for this housing characteristic, thus the mean of the positive coefficients calculated for the first regression array are used to estimate the weight that the number of bedrooms exerts on housing price. Using this approximation, each bedroom adds approximately $13,372 to the price of a home.

The coefficients calculated during the second regression analysis array (n = 4,352), incorporating the tsunami zone binomial variable, are summarized in Table 4: Coefficients for Regression Analysis Array Incorporating Tsunami Zone and Lot Size (Table 4).
Table 4: Coefficients for Regression Analysis Array Incorporating Tsunami Zone and Lot Size (Second Array)

<table>
<thead>
<tr>
<th>Attribute Variable</th>
<th>Coefficient Range (USD)</th>
<th>Coefficient Mean (USD)</th>
<th>Significance Mean</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Minimum</td>
<td>Maximum</td>
<td></td>
</tr>
<tr>
<td>Total Square Footage (units in square feet)</td>
<td>101.47</td>
<td>105.04</td>
<td>103.58</td>
</tr>
<tr>
<td>Lot Size (units in acres)</td>
<td>2,247.53</td>
<td>2,581.54</td>
<td>2,364.37</td>
</tr>
<tr>
<td>Number of Bedrooms¹</td>
<td>13,003.25¹</td>
<td>14,268.97</td>
<td>13,518.24</td>
</tr>
<tr>
<td>Number of Bathrooms</td>
<td>7,320.90</td>
<td>65,376.92</td>
<td>33,251.66</td>
</tr>
<tr>
<td>Number of Garage Stalls</td>
<td>15,408.95</td>
<td>22,475.00</td>
<td>19,361.67</td>
</tr>
<tr>
<td>Presence/Absence of Second Dwelling</td>
<td>30,178.10</td>
<td>56,618.98</td>
<td>42,166.52</td>
</tr>
<tr>
<td>Age of Structure (units in years)</td>
<td>(5.88)</td>
<td>(14.94)</td>
<td>(10.14)</td>
</tr>
<tr>
<td>Appearance Valuation</td>
<td>19,668.39</td>
<td>24,147.11</td>
<td>22,491.47</td>
</tr>
<tr>
<td>Presence of Tsunami Zone + Buffer</td>
<td>(6,024.02)</td>
<td>(18,425.63)</td>
<td>(11,324.47)</td>
</tr>
</tbody>
</table>

Notes:
1. ¹ = Since there is significant multicollinearity in the explanatory variables, only the positive coefficients are utilized to calculate the mean.
2. USD = United States dollar ($)
3. The above table denotes variables included in the second regression analysis array
4. Negative values denoted in parentheses

As shown in Table 4, the location of a property within the tsunami zone is correlated with a negative impact on housing value, decreasing the selling price of property by an average cost of $11,324; however the presence of the tsunami zone is not determined to be statistically significant. While each individual regression calculation comprising the second regression array resulted in a negative impact on real estate value associated with the tsunami zone, the effect was determined to be statistically significant.
in only two out of nine analyses. The minimum significance value for the tsunami zone was 0.021 and the maximum was 0.146, resulting in a mean significance value of 0.198. In both cases of tsunami zone significance, the appearance variable was excluded from the second array regression analysis.

The regression values summarized in Table 4 are generally similar to those denoted in Table 3. The age independent variable is, once again, not determined to be a statistically significant influence on the dependent variable. Also, as observed in the first array, negative values were described for the bedroom coefficient. As the multicollinearity resulting in negative coefficients is manifest in the second array, negative values were again excluded from the mean positive coefficients calculation to arrive at an estimate representative of the true bedroom coefficient.

The coefficients calculated as part of the third regression analysis array (n = 9,765), incorporating the tsunami zone and excluding the property lot size variable, are summarized in Table 5: Coefficients for Regression Analysis Array Incorporating Tsunami Zone and Excluding Lot Size (Table 5).
Table 5: Coefficients for Regression Analysis Array Incorporating Tsunami Zone and Excluding Lot Size (Third Array)

<table>
<thead>
<tr>
<th>Attribute Variable</th>
<th>Coefficient Range (USD)</th>
<th>Coefficient Mean (USD)</th>
<th>Significance Mean</th>
</tr>
</thead>
<tbody>
<tr>
<td>Total Square Footage</td>
<td>108.81 - 109.36</td>
<td>109.08</td>
<td>0.000</td>
</tr>
<tr>
<td>Number of Bedrooms</td>
<td>(17,777.92) - 10,528.90</td>
<td>(7,968.24)</td>
<td>0.000</td>
</tr>
<tr>
<td>Number of Bathrooms</td>
<td>8,475.98 - 61,064.75</td>
<td>31,377.68</td>
<td>0.000</td>
</tr>
<tr>
<td>Number of Garage Stalls</td>
<td>14,391.25 - 20,831.92</td>
<td>17,881.88</td>
<td>0.000</td>
</tr>
<tr>
<td>Presence/Absence of Second Dwelling</td>
<td>47,213.71 - 71,583.26</td>
<td>57,790.25</td>
<td>0.000</td>
</tr>
<tr>
<td>Age of Structure</td>
<td>(8.59) - (15.28)</td>
<td>(11.76)</td>
<td>0.039</td>
</tr>
<tr>
<td>Appearance Valuation</td>
<td>23,518.49 - 23,704.56</td>
<td>23,611.53</td>
<td>0.000</td>
</tr>
<tr>
<td>Presence of Tsunami Zone + Buffer</td>
<td>(6,327.80) - (16,395.66)</td>
<td>(9,714.01)</td>
<td>0.169</td>
</tr>
</tbody>
</table>

Notes:
1. Since there is significant multicollinearity in the explanatory variables, only the positive coefficients are utilized to calculate the mean.
2. USD = United States dollar ($)
3. The above table denotes variables included in the third regression analysis array
4. Negative values denoted in parentheses

Table 5 represents the largest regression dataset, containing over 5,000 additional MLS list data points than those which comprised the regression model summarized in Table 4. Despite the larger dataset, the coefficients derived as part of the third regression array are generally comparable to those calculated during the previous regression array, with the exception of the age variable which is determined to be statistically significant. The mean significance for the tsunami zone is, once again, not determined to be statistically significant. The minimum significance value for the tsunami zone coefficient
was 0.004 and the maximum was 0.279, resulting in a mean significance value of 0.169. As observed in the second regression array, appearance is determined to be significantly correlated with the tsunami zone in the third array.

Temporal Effects on Real Estate Market Value

The influence of the sales transaction date on the selling price is depicted in Figure 1: Average Real Estate Selling Prices Over Time (Figure 1). The average final sales price for each year beginning on January 1, 2005 and ending on September 18, 2014 (study period), is shown in Figure 3 and the general market trend of increasing prices followed a marked decline and subsequent modest recovery is evident.

Figure 1: Average Real Estate Selling Prices Over Time

![Humboldt County Residential Real Estate Selling Prices Over Time](image)
As illustrated in Figure 1, the Humboldt County housing market experienced a maximum average final sales price in 2007 peaking at $351,679, after which time the real estate market was negatively impacted by the global economic recession which initiated a multi-year decline in real estate value beginning in 2008. Housing prices declined to a minimum average value of $241,274 in 2012, after which time the real estate market began to improve as shown by increasing average selling prices for each of the remaining two years within the study period.

Regression analysis data describing the impact of the year of sale on the dependent variable is summarized in Table 6: Date of Sale Regression Analysis Coefficients (Table 6).
Table 6: Date of Sale Regression Analysis Coefficients

<table>
<thead>
<tr>
<th>Date of Sale</th>
<th>Coefficient Range (USD)</th>
<th>Number of Listings (n)</th>
<th>Beta</th>
<th>Significance</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Minimum</td>
<td>Maximum</td>
<td></td>
<td></td>
</tr>
<tr>
<td>2005</td>
<td>56,891.95</td>
<td>65,051.97</td>
<td>1,357</td>
<td>0.158</td>
</tr>
<tr>
<td>2006</td>
<td>58,650.47</td>
<td>69,812.04</td>
<td>1,159</td>
<td>0.162</td>
</tr>
<tr>
<td>2007</td>
<td>56,453.93</td>
<td>60,610.71</td>
<td>1,043</td>
<td>0.137</td>
</tr>
<tr>
<td>2008</td>
<td>37,321.32</td>
<td>40,644.71</td>
<td>830</td>
<td>0.084</td>
</tr>
<tr>
<td>2009</td>
<td>5,613.49</td>
<td>4,959.84</td>
<td>853</td>
<td>0.011</td>
</tr>
<tr>
<td>2011</td>
<td>(24,875.43)</td>
<td>(28,070.74)</td>
<td>887</td>
<td>(0.057)</td>
</tr>
<tr>
<td>2012</td>
<td>(35,451.58)</td>
<td>(41,949.98)</td>
<td>1,032</td>
<td>(0.093)</td>
</tr>
<tr>
<td>2013</td>
<td>(19,984.90)</td>
<td>(23,389.44)</td>
<td>1,126</td>
<td>(0.057)</td>
</tr>
<tr>
<td>2014 (partial)</td>
<td>(6,375.52)</td>
<td>(10,421.89)</td>
<td>734</td>
<td>(0.019)</td>
</tr>
</tbody>
</table>

Notes:
1. 2010 not summarized in this table due to lack of regression data
2. 2014 includes listings from January 1, 2014 to September 18, 2014
3. Negative values denoted in parentheses
4. USD = United States dollar ($)

As denoted in Table 6, there is a statistically significant positive correlation between the date of sale and the sales price for the years up to and including 2008. The effect of the global recession is revealed in the significantly reduced coefficients for 2008 and 2009, as well as the subsequent statistically significant negative correlation between sales price and property sales occurring between 2011 and 2013.
Location Effects on Real Estate Market Value

The effect of the location (municipality, community sub region, or unincorporated area) on the selling price is depicted on Figure 2: Average Sales Price by Location (Sub-Area) (Figure 2). Figure 2 denotes the average market price for specific micropolitan locations throughout the approximately nine year study period.

Figure 2: Average Sales Price by Location (Sub-Area)
As shown in Figure 2, the Trinidad sub-area claims the highest average sales price ($595,978), while the largest urbanized area, the Eureka sub-area, retained the lowest average market value ($258,699) over the study period.

The location-specific regression analysis data is summarized in Table 7: Sub-Area Regression Analysis Coefficients (Table 7).

Table 7: Sub-Area Regression Analysis Coefficients

<table>
<thead>
<tr>
<th>Sub-Area Name</th>
<th>Coefficient Range (USD)</th>
<th>Number of Listings (n)</th>
<th>Beta</th>
<th>Significance</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Minimum</td>
<td>Maximum</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Arcata</td>
<td>56,267.88</td>
<td>58,889.93</td>
<td>931</td>
<td>0.13</td>
</tr>
<tr>
<td>Eureka</td>
<td>(23,019.84)</td>
<td>(20,267.91)</td>
<td>3,067</td>
<td>(0.09)</td>
</tr>
<tr>
<td>McKinleyville</td>
<td>11,762.66</td>
<td>16,917.79</td>
<td>1,381</td>
<td>0.03</td>
</tr>
<tr>
<td>Trinidad</td>
<td>147,861.90</td>
<td>148,726.17</td>
<td>89</td>
<td>0.11</td>
</tr>
<tr>
<td>Fortuna</td>
<td>(17,742.06)</td>
<td>(11,386.13)</td>
<td>1,185</td>
<td>(0.05)</td>
</tr>
<tr>
<td>Ferndale</td>
<td>23,037.10</td>
<td>24,932.78</td>
<td>188</td>
<td>0.03</td>
</tr>
<tr>
<td>Garberville</td>
<td>45,208.27</td>
<td>54,819.48</td>
<td>47</td>
<td>0.03</td>
</tr>
</tbody>
</table>

Note:
1. Negative values denoted in parentheses
2. USD = United States dollar ($)

As shown in Table 7, there is a negative correlation between the real estate value coefficient and the location of properties within the Fortuna and Eureka sub-areas, resulting in a respective selling price discount of $17,000 (Fortuna) and $23,000 (Eureka). In contrast, there is a positive correlation between the sub-area and selling price variables for properties in the following sub-areas (listed in descending order of
Increasing the value of the real estate by virtue of the location of property within these regions.

**Tsunami Zone Regression Analysis**

The number of properties that were identified within the tsunami zone and tsunami zone buffer is reported in Table 8: Tsunami Zone Variable Quantification (Table 8). As a result of the expanded area of the tsunami zone buffer, 41 additional properties are included within the cumulative area formed by the tsunami zone and tsunami zone buffer than are contained within the tsunami zone alone.

**Table 8: Tsunami Zone Variable Quantification**

<table>
<thead>
<tr>
<th></th>
<th>Tsunami Zone</th>
<th>Tsunami Zone Buffer (Tsunami Zone +50’)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Number of Properties</td>
<td>252</td>
<td>293</td>
</tr>
</tbody>
</table>

As shown in Table 6, there were a relatively small number of properties sold that were located within the tsunami zone (n = 252) or tsunami zone buffer (n = 293) throughout the date range limiting the HAR dataset. While the number data points within the tsunami zone and buffer zone was less than 300, a statistically viable analysis is possible. Consistent with the law of large numbers theorem, the generally-accepted minimum analysis threshold of 250 datum is exceeded by the observed occurrences of properties within the tsunami zone and buffer zone. Given the sufficiently large number of data points for real estate within the tsunami zone and buffer, the datasets may be
reasonably incorporated into the regression model calculation with confidence that the resulting output will result in statistically significant coefficients.
Applying the various locally-appropriate regression coefficients to a given property within Humboldt County allows for the general estimation of the selling price for that property. By applying the coefficient monetary value in direct proportion to a property’s observed physical characteristics an estimate for that property’s market price may be derived.

Some real estate characteristic coefficients were determined to weigh more heavily on a property’s selling price than others. The real estate characteristics determined to significantly impact property selling price are summarized in Tables 3 through 5. The coefficients with the greatest impact on selling price include the following, listed in descending order of magnitude: presence of a second dwelling, appearance valuation, number of bathrooms, number of garage stalls and number of bedrooms. The presence of the tsunami zone was not found to affect the market price in a statistically significant manner when the appearance variable was included in the regression analysis. The appearance and tsunami zone variables are significantly correlated at the 0.01 level (two-tailed), thus the interaction between these variables should be further investigated.

The hypothesis that the price of Humboldt County property is significantly increased or decreased by the location of that property within the tsunami zone was generally not validated by the results of our study of the HAR dataset for the specified date range and number of transactions. As such, the hypothesis that Humboldt County
real estate selling price is significantly affected by property incorporation within the tsunami zone is rejected and the null hypothesis that tsunami zone does not significantly affect real estate selling price for the observed is accepted. This conclusion is qualified by the observation that the tsunami zone variable becomes significant when the appearance predictor variable is excluded from the regression analysis.

In contrast to inland property, real estate within the tsunami zone is subject to increased weathering, additional regulatory restrictions constraining renovation and/or development, and other socioeconomic factors that disproportionately impact the appearance rating of coastal property. The appearance and tsunami zone variables are significantly correlated at the 0.01 level (two-tailed), thus the interaction between these variables should be further investigated.

**Opportunities for Additional Research**

**Further Regression Models**

Further study may address the collinearity exhibited by the dataset and attempt to more precisely define the various independent variables which collectively describe a property’s selling price. It is recommended that additional regression models be generated using Humboldt County real estate transaction data and tsunami zone geospatial information. Such analyses should specifically explore the relationship between the appearance and tsunami zone variables. The appearance and tsunami zone coefficients should further investigated and refined coefficients individually derived to further study the effect (significance) of the tsunami zone as a predictor variable for real
estate market value. Additionally, the appearance variances between coastal and inland properties may be further scrutinized to more specifically measure the impact of the qualitative and relative appearance variable on selling price via further regression modeling.

**Tsunami Zone Public Awareness**

Precise measurement of the effect of the tsunami zone on selling price is subject to informed market participants acting with an understanding of the presence and consequence of the tsunami evacuation zone and tsunami hazard associated with a specific real estate transaction. It is necessary for sellers and buyers of real estate to be cognizant of the delineation of the tsunami zone and understand the consequences resulting from the inclusion of a subject property within such a hazard boundary, including the expectation for general awareness regarding tsunami warning systems, potential for evacuation and evacuation logistics.

Currently, it is not clear that a general understanding of the tsunami hazard and location of the tsunami zone exists amongst the general public in Humboldt County. While there is some public awareness regarding tsunami hazards and associated evacuation zones, it is not clear what the distribution of such an understanding is (how widespread and amongst which populations), or the depth of penetration of tsunami hazard acknowledgement (how significant is tsunami awareness and risk to individuals). Further, it is unclear that such an understanding, where it exists, informs the decisions of purchasers or sellers in general real estate transactions.
As such, a study of homeowner and homebuyer general knowledge regarding the location and implication of the tsunami zone should be conducted to evaluate the general awareness of the tsunami hazards in relation to applicable property transactions. Such further study should evaluate the presence, or absence of tsunami awareness, including distribution of general tsunami awareness within the population and significance of tsunami risk management in decision-making. Where tsunami hazard and tsunami zone awareness exists, further study should attempt to evaluate the cost-benefit analysis decision-making process employed by homebuyers and sellers in association with relevant real estate transactions. An educational intervention on tsunami preparedness may be appropriate as a next step, introducing an opportunity for a before-and-after study design.

**Longitudinal Study**

It is recommended that further study focus on the evaluation of housing prices for individual properties within the tsunami zone over time, comparing the price of such real estate in constant dollar units prior to and after tsunami zone establishment. A longitudinal study would account for the developing public awareness of tsunami hazards and tsunami zone location, allowing for the evaluation of evolving market reactions to such risk stimuli. It is recommended that longitudinal study control for the effect of inflation by evaluating housing prices in standard units. Standardization of monetary units for the selling price dependent variable over time may be accomplished for
regression analysis by selecting the reference point date with the least number of confounding variables as the standard evaluation metric.

**Homogeneous Property Comparison Study**

Further study is needed to more precisely assess the economic effect of the tsunami zone. Such additional study should attempt to control for externalities by identifying generally homogeneous properties located inside and outside of the tsunami zone and within a set distance of the tsunami zone boundary. By comparing similar real estate within close proximity, dissimilar in presence or absence of the tsunami zone, and analyzing the selling prices of such properties over time, patterns may be observed in market price disparities between the properties located within the tsunami zone and those located outside of the tsunami zone.

By limiting analysis to real estate having similar attributes within generally comparable areas, a regression analysis may be performed that attempts to control for unmeasured market stimuli by minimizing the number of dissimilar environmental and sociological externalities, thereby attempting to isolate the tsunami zone as the only divergent variable to more appropriately measure the significance of its impact on selling price.

**Conclusion**

The evaluation of housing price using hedonic models attempts to control for accumulative sociological and environmental factors that influence the spectrum of
human behavior by means of regression analysis coefficients which significantly correlate, or at least functionally approximate, our observations of the housing market. In doing so, regression analysis methodology refracts the spectrum of socioeconomic market influences into discrete components which represent the best estimate for the range of price indicators that determine real estate market value.

Cognizant of the limitations of regression analysis, this study provides a range of estimates for the economic impact that specific housing characteristics exert on property value in Humboldt County, concluding that a select few attributes as described herein collectively account for the largest proportion of the selling price. In addition, this study concludes that property location within the tsunami zone does not significantly correlate to the market price for applicable Humboldt County properties. Finally, this study provides the preliminary framework for further research into the locally-significant predictors of residential property market value, including the presence of regulatory boundaries such as the tsunami zone, for real estate in Humboldt County.


County of Humboldt (2013). Humboldt county general plan. Humboldt County Board of Supervisors Draft, Part 1, Chapter 1 (1-4-1-5).


United States Census Bureau (2014). State and County QuickFacts, Humboldt County, California.


APPENDIX

This appendix contains Humboldt County tsunami evacuation maps, Figures 3 – 17, produced by the Redwood National and State Parks. Figures 3 – 17 depict the projected maximum extent of tsunami inundation for various subregions throughout Humboldt County. Areas on the appended maps denoted with yellow highlight collectively represent the tsunami zone described by this study. The tsunami zones demarcated on the following figures were utilized for the primary data regression analyses described by this study.
Figure 3: Tsunami Evacuation Map – Humboldt Bay Region
Figure 4: Tsunami Evacuation Map – Orick Subregion
Figure 5: Tsunami Evacuation Map – Big Lagoon Subregion
Figure 6: Tsunami Evacuation Map – Trinidad Subregion
Figure 7: Tsunami Evacuation Map – Moonstone Subregion
Figure 8: Tsunami Evacuation Map – McKinleyville Subregion
Figure 9: Tsunami Evacuation Map – Arcata Subregion
Figure 10: Tsunami Evacuation Map – Arcata Bay Subregion
Figure 11: Tsunami Evacuation Map – Manila Subregion
Figure 12: Tsunami Evacuation Map – Samoa Subregion
Figure 13: Tsunami Evacuation Map – Eureka Subregion
Figure 14: Tsunami Evacuation Map – South Bay Subregion
Figure 15: Tsunami Evacuation Map – King Salmon & Fields Landing Subregion
Figure 16: Tsunami Evacuation Map – Eel River Subregion
Figure 17: Tsunami Evacuation Map – Shelter Cove Subregion
END OF APPENDIX