Exploring the impact of the natural gas leak on home values in Porter Ranch, California

by

Raymond Calnan

A Thesis Presented to the
Faculty of the USC Graduate School
University of Southern California
In Partial Fulfillment of the
Requirements for the Degree
Master of Science
(Geographic Information Science and Technology)

May 2019

To Adina, I could not do any of what I do without your continued support and encouragement. I love you and am grateful to have you in my life.

Table of Contents

List of Figures

List of Tables

Acknowledgements

I am grateful to my committee chair, Professor Wu, for the direction I needed and to Professor Kemp for her ongoing encouragement throughout the process. Both were integral to my success in this undertaking and to my mental health along the way.

List of Abbreviations

ACS	American Community Survey
DiD	Difference in Difference Analysis
GIS	Geographic Information System
GISci	Geographic Information Science
MLS	Multiple Listing Service
PSM	Propensity Score Matching
PUMA	Public Use Microdata Area
SRAR	Southland Regional Association of Realtors
SSI	Spatial Sciences Institute
USC	University of Southern California

Abstract

Families create and build wealth through their homes. What happens when the traditional trajectory of home values is interrupted by external events, and how can homeowners determine appropriate compensation? During late 2015 and early 2016, a natural gas leak from an underground storage tank went unchecked in the community of Porter Ranch, California. Following the capping of the leak, the area was cleaned so that no visible sign of the leak remained. With over 11,000 people and 2,800 households affected, the question lingered of whether the homes suffered in value from the past gas leak. This study considers the limitation of the traditional methods for determining home value, and applied an alternative method borrowed from the social sciences with spatial context in the design. Propensity score matching was used to pair the census tracts in the Porter Ranch area with similarly structured census tracts within Los Angeles County. A regression analysis was used to identify key characteristics that would be used to identify similar census tracts using census demographic data. Difference-in-difference analysis and graphical trend analysis were then used to determine that a loss in value did exist in homes of the Porter Ranch community. The results of the study suggest that the home values in the Porter Ranch area increased at a slower pace than the comparable areas. This suggests that an environmental stigma exists in the Porter Ranch community, and the home values may have been negatively impacted by the past gas leak.

Chapter 1 Introduction

Legacies of family wealth and security are often closely tied to the value of homes. Many people rely on the forced savings of their monthly mortgage payment to their homes to build security in retirement. Others use the value in their homes to access equity and send their children to school or pay for unexpected large expenses. However, many uncertainties in the home values are not controlled by the homeowners. The obvious one is change in home values as a result of changes in the overall economy (Leung 2003). Home values rise and fall due to overall market influences, such as mortgage interest rates, employment, and the strength of the economy. Home values can also be affected by natural or manmade changes in the local environment.

When an environmental disaster occurs, does the incident have impact on the housing market nearby and would the home value decrease even though this is no fault of the owner? In this research, my goal is to investigate the effect of a past natural gas leak on home values. Specifically, I used the Porter Ranch area in California as a case study and determine if there are ongoing or sustained losses in home values as a result of perceived environmental stigma associated with the past gas leak in this area. Based on Uniform Standards of Professional Appraisal Practice (USPAP) 2018-2019, the term "environmental stigma" is formally defined as "[an] adverse effect on property value produced by the market's perception of increased environmental risk due to contamination (Appraisal Standards Board 2018, p 81).

The understanding of whether an ongoing loss in home value exists or not as a result of the perceived environmental stigma will help in case law and in determining future losses to property owners. I hypothesize that there is a detriment to home value due to perceived stigma beyond the physical damage caused by the gas leak event in the affected area. The approach in this study may also shed light on using location-based methods in determining the loss in home values in other types of environmental stigma, such as ground water contamination or soil contamination. Having an alternative measure could provide better assessment of future disasters since there have not been any similar studies in the past.

1.1. Porter Ranch Gas Leak

On October 23, 2015, a leak of natural gas from an underground storage well was reported. The storage facility is a natural underground oil deposit that was pumped out in the past century. After the oil was removed, the underground formation was used as a natural gas storage facility. Most of the people living in the community were unaware of the gas storage facility or did not give it much thought until the gas leak occurred.

The leak continued until it was finally capped on February 11, 2016. This gas leak is one of the worst manmade disasters recorded due to the approximate 97,100 tons of methane gas that leaked over more than three months (Conley 2016). Due to the nature of the leak and wind patterns, a broad area of a community known as Porter Ranch was affected, along with surrounding communities, which had lessor exposure to the airborne plume of natural gas. The map in Figure 1 shows the gas leak location at the SoCal Gas' Aliso Canyon facility and the complaints reported by residents to the Gas Company during the gas leak period (South Coast Air Quality Management District 2018).

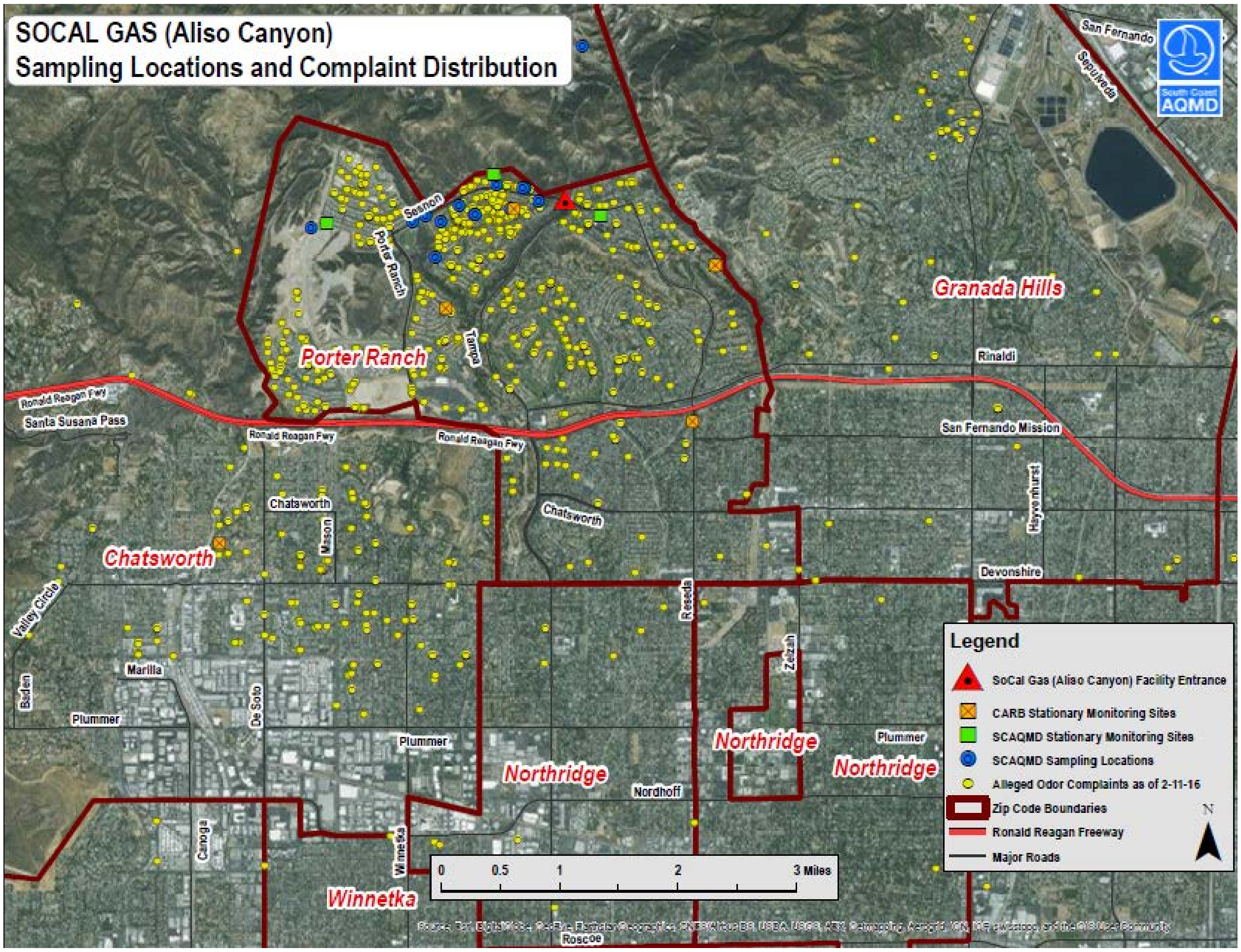

Figure 1 The locations of the SoCal Gas Facility in Aliso Canyon and the resident complaints in the surrounding neighborhood areas.

Over 11,000 people and approximately 2,800 households in this Porter Ranch area were evacuated during the gas leak (Blake 2016). During this period, residents stayed in local hotels or rental units. The demand of the temporary housing was so high that the tax revenue of hotel rooms in the neighboring areas increased by more than 50% over the previous year during the same period (Adkins 2015). The costs on temporary housing, relocation, clean-up, and monitoring were estimated at over $1 billion as of August 2018 (Reuters 2018). However, the total cost of the gas leak damage is difficult to measure when considering possible long-term impacts, such as healthcare costs and potential damage on biodiversity.

Following the capping of the gas leak, private vendors were hired by the SoCal Gas Company to undertake a clean-up effort to wipe away any physical signs of the gas leak (e.g. the effluent mist released during the gas leak). During this process, homes in the Porter Ranch area were cleaned inside and out, playgrounds were pressure washed, and air inside buildings was scrubbed for possible particulate matter. A stationary air quality monitoring report for the leak concluded that few physical signs remained as of January 2018 (South Coast Air Quality Management District 2018). Mobile monitoring by the South Coast Air Quality Management District along with the stationary monitoring by the Gas Company in the area surrounding the storage facility continue. With the physical clean-up and continuous monitoring, however, residents still have concerns about future leaks and worry that the potential harm to human health could have already occurred; lawsuits related to health problems due to the gas leak are now in process.

1.2. Real Estate Environmental Stigma

Environmental stigma can be defined as market resistance to a specific property as a result of some characteristic within the property or outside of it. Market resistance is usually determined by identifying the property's value as being lower than it would be without the deficit or negative characteristics. The difference between stigma and other forms of value loss is that a loss from stigma remains on the property after the problem has been remediated. A complete discussion relating to the concept of stigma and how it is determined is provided in Chapter 2.

1.3. Motivation

Losses in home values as a result of groundwater and soil contaminations by oil spills and hydraulic fracking have been documented. However, the research focuses on the loss in individual home values from the actual loss of use (Simons and Saginor 2006). This physical loss of use is fundamentally different from human perception and does not provide insight into the potential losses in home value that is unable to be remediated or compensated through substitution, such as finding a replacement home.

This research is motivated by the need to establish a method for analyzing the effects of natural and manmade disasters that may have long-lasting detrimental effects on home values and the homeowners. There is limited case law related to compensation for such events due to the desire of large organizations to settle out of court in order to avoid public scrutiny of case outcomes (Golberg and Calnan 2016). By establishing the existence, or non-existence, of a diminution in home values, claims and lawsuits against companies or government agencies can be handled more expeditiously and with greater fairness to all parties involved.

1.4. Research Goals

The goal of this research is to identify whether the home value changes in Porter Ranch area are due to the natural gas leak. The research objectives are (1) to identify the control areas that have comparable social and economic compositions of the gas leak-affected region, (2) to analyze the home value trends in the Porter Ranch region and (3) to determine if home value changes follow the occurrence of natural gas leak by comparing the changes in values between the Porter Ranch region and the identified control areas. I hypothesize that there is a diminution in value that remains unexplained by market, demographic, or property factors. The research will help explain the potential loss in home values as a result of the lasting perceived environmental stigma in a region.

1.5. Study Organization

The remainder of the thesis provides a literature review about the background of stigma and methods for analyzing the effect of stigma on housing values in Chapter 2. Chapter 3 describes the data and methodology of this study and Chapter 4 describes the results of the study. Chapter 5 provides the conclusion of the study and identifies limitations in this study and opportunities for future research.

Chapter 2 Background and Literature Review

This chapter provides a brief definition and discussion of the impact of stigma in real estate (Section 2.1), and an overview of the traditional approaches to determining the loss of value as a result of this real estate stigma (Section 2.2). The final section of this chapter (Section 2.3) identifies the alternative methods that are not traditionally used in determining property values effects that are the result of stigmatization of property.

2.1. Stigma

Many studies have highlighted the impacts of natural and manmade events on the desirability of residential and commercial properties (Boyle and Kiel 2001; Dale et al. 1999; Hastings 2004; Jackson 2001). The past studies tend to focus on events that are tangible in nature and leave lasting effects, such as lead pollution in soil. The literature also shows the relationship between the real estate values and desirability (Alfert, Collison and Tate 2005; Kinnard J and Worzala 1999; Lin 2014; Patchin 1988). This section summarizes the effects of real estate stigma on the desirability and therefore the value of real estate. Various types of stigma that have affected real estate values are also discussed. Stigma can often be long lasting and impact property values long after the original event/disaster has been mitigated/remediated.

2.1.1. *Defining Market Stigma*

In real estate appraisal, stigma can be thought of as market resistance related to the real or perceived risks by the market (Appraisal Standards Board 2018). Environmental stigma, a negative stigma generated from an outside event that is not directly a part of a property, can diminish the value of the property (Lin 2014). In some circumstances, this environmental stigma can result from a past event on a property that has been mitigated, however the knowledge of the past event is enough to make it less desirable and therefore increase the market resistance.

Losses in property value as a result of such market resistance is not the same as the property value being affected by the general economic conditions. The overall economy moves through cycles. These changes in the economy can include increases or decreases in interest rates, income, and buying power, which indirectly result in the change of home values (Wheaton 1999). The key difference with market resistance is that the overall economy can be improving, with home prices generally increasing, while at the same time stigmatized properties can suffer from a loss in value. This loss in value is attributed to something outside the general market conditions (e.g. an undesirable land use in the neighborhood) or from within the property (e.g. a lingering history). Homebuyers will factor the amenities of a neighborhood, good or bad, into the purchase price (Boyle and Kiel 2001). Environmental stigma associated with a neighborhood can negatively affect the desirability of the neighborhood and therefore reduce the perceived value of the neighborhood to a potential homebuyer (Hastings 2004).

The duration of the undesirable external event may also determine whether the diminution in value is permanent or temporary (Dale et al. 1999). For example, an existing airport or a freeway is often factored in when a buyer makes an offer on a home, therefore the loss in value is capitalized into the home and runs with the location. However, a new freeway being built next to a home may not have been factored into the original purchase price and therefore would negatively impact the existing owner of the home because it would be factored in by a future buyer. In contrast, a contaminated property in the final stages of cleanup might have much shorter value loss tenure. The owner of the property at the time of completion may reap the reward of purchasing a less desirable property to then sell it once it became more desirable due to the improvement of the area (Jackson 2001).

2.1.2. *Example of Stigma*

Stigma can take many forms; each of them can have different impacts on the desirability and value of property. A few examples of events that can result in lasting stigma include air pollution, ground water contamination, and landfills. These types of disasters may not always result in stigmatized properties. If the damage can be remediated and there is no further loss in value other than remediation costs, then there is no stigma (Roddewig 2001). However, it is common for lasting value effects even when the disaster has been remediated.

An example of manmade contamination that impacts property values is air pollution. Ridker and Henning (1967) recognized that property value varies based on air pollution and air quality. The research changes economists' views about the likelihood of measuring the willingness to pay for environmental public goods. Many of the applications proposed in the earlier studies relied on judgmental evaluation of hedonic estimates (Ridker and Henning 1967). However, later research used meta-analysis of several research studies to describe the existing estimates of the marginal value on reducing air pollution. The meta-analysis results confirmed the relationships between measures of the incremental value of reduction in air pollution, the level of air pollution in each city, the average income of its residents, and the value of property (Smith and Huang 1993). When the air pollution is relatively low, the house prices are relatively higher. The negative effects of this type of negative externality (e.g. air pollution) diminish as the externality is mitigated.

However, even as the problem is mitigated, the memory of the problem can have lasting effects (Dale et al. 1999). From 1934 to 1984, the RSR lead smelter in Dallas, Texas was found to contaminate surrounding neighborhoods' soil with toxic lead dust and slag (stony waste separated metal from raw ore during smelting) (Dale et al. 1999). Because of the negative environmental impacts, this area was believed to be 'stigmatized'.

As such, one would expect that individuals would pay a premium to live away from the smelter, and therefore result in property values to rise as distance from the smelter increases. In 1984, after the RSR smelter closed and series of cleanup measures were taken to clean the area, the price of nearby houses rebounded. However, the area in the closest proximity to the smelter also showed somewhat slower responses to price rebound (Dale et al. 1999).

The RSR smelter case demonstrated the implications of several policies concerning clean-ups and the effect on potential price rebounds. First, cleaning up a natural or manmade disaster can generate considerable benefits to the property values of nearby residences. Second, the value of the homes cannot be calculated simply by examining pre-cleanup price differentials since an accurate estimate of damages must account for both the period the site was affected, the speed of clean-up, the trend of price rebound, and the proximity to the disaster site (Dale et al. 1999). It should be noted that not all perceptions on negative environmental issues would result in a negative impact on home values nearby. While landfills are often thought of as a nuisance or negative externality, well-designed landfills have showed no significant effects on the value of properties within close proximity (Bleich, Findlay III and Phillips 1991).

2.2. Traditional Approaches to Appraise Property Values

There are three widely accepted general approaches to appraise properties for federally regulated loans (The Appraisal Institute 2008). These include the sales comparison approach, the income capitalization approach, and the cost approach. Determining the appropriate approach to use in a valuation assignment is critical and must start with an understanding of the purpose of the appraisal assignment and the type of property. The examples provided below are simplified to ensure a basic understanding of sales comparison and cost approaches to value since these are the most commonly used approaches in single-family appraisal practice and in cases of stigmatized properties associated with contamination (Kinnard J and Worzala 1999).

2.2.1. *Sales Comparison Approach*

The preferred method for determining the value for a single-family home is the sales comparison approach. Generally, it is the easiest approach to understand and it represents the common, practical thinking of most homebuyers. Single-family home appraisers often prefer to use this approach because it properly reflects the logic used by most market participants purchasing a single-family home (Clayton 1998).

The premise of this approach is based upon the concepts of substitution and competition (Rattermann 2009). A buyer would purchase a property at a value that reflects the adjusted sale price of similar properties that have recently sold. Sale and list prices of a property are adjusted based on the differences between the subject property and the comparable properties. Comparable properties are properties that have similar characteristics to the subject property but may differ in minor ways, such as size or amenities provided. Table 1 illustrates the procedure used in the sales comparison approach. Differences between the subject property and the recent sales of the five pre-identified comparable properties are compared in order to determine the adjusted value of the subject property. For simplicity, the comparable sales have similar numbers of bedrooms and bathrooms with the only differences being the existence of a garage or view. An example can be the subject property with no garage, but a recently sold comparable property contains a garage (Sale 4); a decrease in the sale price of the comparable property is necessary in order to reflect the added value associated with having a garage.

Table 1 The sale comparison approach compares the sale prices of recently sold comparable properties.

	Subject property	Sale 1	Sale 2	Sale 3	Sale 4	Sale 5
Sale Price	TDB	$650,000	$650,000	$625,000	$630,000	$650,000
Beds	4	4	4	4	4	4
Bathrooms	2	2	2	2	2	2
View	No	Yes	Yes	Yes	No	Yes
Garage	No	Yes	Yes	No	Yes	Yes

The first step would be to determine the difference in value between Sale 3 and the other three sales that have a garage. In this example, it is evident that the market attributes a $25,000 value to the existence of a garage since Sale 3 sold for $25,000 less than other sales with views and garages. Similarly, the existence of a view is valued at $20,000 in this example, because the home (Sale 4) lacking a view sold for $20,000 less than otherwise comparable homes. In actual practice, multiple pairs or sets of comparable sales are used to determine the market preference and value of a particular characteristic of a home (Rattermann 2009).

Once the market value of the individual characteristics is determined, the appraiser can compare the recent sales to the subject property. In Table 1, the subject property does not have a garage or view, so it can be concluded that its value should be $45,000 less than the recent sales ($25,000 for the garage and $20,000 for the view). Therefore, the concluded value can be determined as $605,000 for the subject property. This is a contrived example to highlight the process of the sales comparison approach. In actual appraisal practice, the adjustments are more complex and the adjustments to comparable sales do not neatly result in a perfect solution for the subject property value.

In the case of a diminution of value resulting from a potentially stigmatized property, the analysis is similar to the example provided. However, instead of an adjustment for a view or garage, an adjustment for "stigma" can be substituted (Bell 1998). If sales of homes with the negative impact have already occurred, then the appraiser can use the percentage reduction from an unimpaired home to determine the loss attributable to the stigma. As illustrated in the example, the value of the view can be determined using previous sales of homes both with and without views. The appraiser faces difficulty in finding the value for the first property that has suffered the loss if no previous sales of impacted homes are available. Events or impairments that are broad in geographic area compound this difficulty in using this approach because the impairment may impact many or all homes that are similar in specific attributes as well as location. Determining the loss of value is equally challenging when no sales have occurred, and all similar properties become similarly impaired, a situation commonly presented in large-scale disasters (Mundy 1992).

In the case of large-scale manmade disasters, even after the physical evidence of the leak is remediated, there may still be an ongoing stigma based upon fear or notoriety (Patchin 1988). The sales in the immediate area would be expected to be similarly affected therefore the sales comparison approach will capture the loss in desirability, at least those perceived by buyers active in the market. However, this perception may not reflect the actual cost and it would be difficult to ascertain what that cost is without comparing the values to homes outside the impacted area.

2.2.2. *Cost Approach*

The basic premise of the cost approach is that a buyer should not pay more for a property than what it would cost to purchase vacant land and build the property on it (The Appraisal Institute 2008). The steps for the cost approach are straightforward. First, the appraiser determines the value of the land and the cost to construct the improvements, and then considers any deductions required for economic depreciation. Depreciation estimates losses in economic value comes from multiple sources, including building deterioration and detrimental impacts to the property from both internal and external forces.

Since depreciation is a key element of the cost approach, the approach is very useful in valuing single-family homes. The exception would be in an urban area, where a buyer is less likely to consider this approach because densely developed areas have little land available. This approach can be less reliable than the sales comparison approach and tends to overestimate the value in single-family homes in situations where detrimental conditions may exist (Dotzour 1990). The approach can still be useful in the case of environmental stigma when the stigma is considered a form of depreciation.

As to the detailed estimate in the cost approach, the improvement cost depends upon contractor surveys or industry standards (Rattermann 2009). The depreciation can be determined based upon market comparisons, similarly to those used in the sales comparison approach (The Appraisal Institute 2008). As in the example shown in Figure 2 below, the cost of the construction ($450,000) is added to the cost or value of the land ($350,000) and then depreciation ($82,000) is deducted to compensate for improvements that are not in new condition. This results in a concluded value of $718,000 for the subject property.

Cost of Land	$ 350,000
Cost to build (new)	$ 450,000
less: Depreciation (for improvements that are not new)	$ (82,000)
Value	$ 718,000

Figure 2 Example of process for using cost approach.

When there is an issue surrounding diminution of value, the cost approach is very applicable if the cost to remediate can be determined. In a case of a property that has physical damage, the loss in home value should be equal to the cost of repair, plus potential inconvenience/time costs (Bell 1998).

2.2.3. *Determining Loss in Value*

Both of the sales comparison approach and cost approach are straightforward when a home value can be determined or a loss in value can be remediated. They can also be useful in situations where the damaged property after remediation cannot be repaired to its "unimpaired" state. The property value in the "unimpaired" state would then be compared with the value of the property in an "as is" or "impaired" state. The value difference between the two states is the loss in value. This step can be the key to determining the loss from marketing stigma versus a repairable cost of an environmental disaster since the cost to repair is not often accounted for the total potential loss as a result of market perceptions (Appraisal Standards Board 2018). Table 2 provides a simple illustration of how market stigma can be quantified using the example of a mold problem due to a leaking roof.

Table 2 An example how marketing stigma could be quantified by calculating the home value difference between the sale price and the 'impaired' state.

Unimpaired home value	$ 750,000
Cost to repair	$ 125,000
Impaired Value	$ 625,000
Home sale price	$ 575,000
Market Resistance (difference)	$ 50,000

In this example (Table 2), the home is worth $750,000, assuming it did not have a mold problem (Unimpaired home value). The owner has an estimate of $125,000 to repair the roof and remove the mold in order to bring the home to normal condition. Including the owner's repair cost, the home's "impaired" value is $625,000 ($750,000 - $125,000). If through a market survey, the owner can sell the home for $625,000, then there is no stigma or market resistance. However, if the market survey indicates that the owner must sell the property for less than $625,000, then there is market resistance to the 'idea' of having a mold problem. In this case, the buyers want a discount of $50,000 ($625,000 - $575,000) beyond the cost to repair the property. This discount is the stigma or market resistance cost attached to the property as a result of the past mold existence. In some situations, the loss from market resistance can be several times the actual cost of remediation (Simons and Throupe 2005).

Calculating the cost of stigma, as in this example, is straightforward when the sale information is available. However, when the property has not yet been sold or there is no intention of selling in the short-term, the cost associated to stigma cannot be determined through sales immediately (Bell 1998). Therefore, case studies of past market stigma are often used (Alfert, Collison and Tate 2005). The appraiser must rely on past home sales that had the stigma discounts as parts of the transaction and will have to estimate the subject property's level of market stigma. In many situations, this is accomplished through market surveys and other qualitative measures rather than quantitative analyses (Kinnard J and Worzala 1999).

The loss in value as a result of market stigma can be exacerbated by the market conditions at the time (Sanders 1996). In weak markets, the discount resulting from market stigma may be greater than in times when the overall market values are increasing. This is due to the expectations of risk that the potential buyers perceive.

2.3. Alternative Methods for Determining Loss in Value

Given the need for sale prices of comparable properties, the traditional methods used to determine value or loss in value of a property as a result of market stigma are limited. In any incidents where the environment of a geographic region is impacted, it is difficult to find sales of homes with similar locational attributes outside of the impacted area in order to compare them to the houses within the impacted area (Kinnard J and Worzala 1999). An additional limitation of the traditional approaches is to apply them when attempting to determine loss on a large spatial extent. One method to determine the impacts on many home values in an area is multiple regression analysis (Benjamin, Guttery and Sirmans 2004). A major issue with this approach is the selection of comparable sales to use in the analysis.

To minimize analyst bias and data limitations, alternative methods of analysis are considered in the following sub-sections. These are tools more commonly found in social science research and demographic analysis: Propensity score matching, Difference-in-difference analysis, and Trend analysis.

2.3.1. *Propensity Score Matching*

Propensity score matching (PSM) is a method for estimating the effects of a specific treatment in social sciences research (Caliendo and Kopeinig 2008). The PSM allows researchers to match observations from a treatment group to observations from a non-treatment (control) group based upon a set of criteria not affected by the treatment (Caliendo and Kopeinig 2008). For example, the researcher may use variables such as age, place of birth, wealth/income of parents, race, gender, etc. in an analysis if the treatment is job training and the outcome being tested in additional income at retirement.

A probit regression analysis is used to determine the average treatment effect of the observations with consideration for the two groups possible, treatment and control (Austin 2011). The treatment effect is the outcome after the treatment group has a specific treatment. For example, the treatment could be additional job training for human resources with the expected outcome of increased salary at retirement. The group that had the additional job training would be the treatment group. The individual observations are then scored based upon the probability of them being in the treatment group. The scores are used in order to determine how similar observations from each group are to each other, without considering the treatment itself.

Different models can be used for matching, and the most common model is a method based upon the closest match in the sum of the covariates using the average treatment effects (Bryson, Dorsett and Purdon 2002). In recent years, the use of PSM has expanded to many fields of research including migration decisions, labor market movements, and government policy (Bryson, Dorsett and Purdon 2002).

The main issue of PSM is that one cannot verify an expected outcome (from a treatment) that has not occurred. There is no way to test what the opposite outcome would be if the test subject was not given the treatment since the treatment has already been applied to the test subject (Caliendo and Kopeinig 2008). PSM might compare the treatment subjects to non-treatment subjects to determine the differences in the outcomes were based upon differences in treatment.

Variable selection in housing related regression analyses is often based on the type of research conducted and the purpose of the value being sought (Bloom, Noble and Nobe 2011). Critical to the analysis are structural variables, such as the number of bedrooms, bathrooms, house size, lot size, etc. (Kang and Reichert 1991). However, most analyses do not use all of the potential variables, but rather select those most critical to the analysis and relevant to the area under consideration (Bloom, Noble and Nobe 2011). The selection of variables is dependent on the variability within the geographic area under study and the aggregation level used, such as individual home level or the census tract (Boyle and Kiel 2001). Locational variables are often considered when determining the effect of particular variables on home values and when trying to predict values (Dubin, Pace and Thibodeau 1999).

2.3.2. *Difference-in-difference Analysis*

Difference-in-difference (DiD) analysis is a method of comparing outcomes to determine if there are differences between observations of two characteristics that are changing at different rates. This type of analysis may be one of the most important strategies in applied economics as it allows for both the treatment group and the control group to change over time and simply highlights the differences that cannot be explained (Puhani 2012; Bryson, Dorsett and Purdon 2002). The purpose of DiD is to identify the change in value not attributable to influences common to both the test group and the control group. The basic determination of DiD is calculating the change in values after an event. The change in value of the control group is compared to the change in the experimental (test) group as shown in the equation below.

Difference in Difference (DiD)

= (Change in values in test group) - (Change in values in control group)

Using the example of job training, DiD can determine how much the salary increased over time for each group (treatment and non-treatment) and identify how much of the salary increase could be attributed to the job training rather than other variables such as years of work, market change and etc. The analysis can be done based on the comparison of the salary for the observations in both the treatment group and the control group before and after job training in certain amount of timeframe to see if one group increased/decreased at a higher rate during this time period (Donald and Lang 2007).

2.3.3. *Graphical Trend Analysis*

Graphical analysis of trends is one of the most straightforward ways to communicate changes over time. This method of analysis can be as complex as determining quartile changes to the log of home values using statistical regression or as simple as plotting the individual values over time (Clark and Coggin 2009). It is very common to establish trends and changes using graphical means in policy and political science related topics as these methods tend to be more accessible to wider audiences and are easily interpretable (Kastellec and Leoni 2007).

By depicting changes in values over time to visually identify trends and differences between geographic areas the summarized data for a geographic area, such as a county or census tract, can be made more comprehensible with limited author/researcher influence (Weissgerber, Milic and Winham 2015). Using simple by effective graphs that represent true events help users of research and data understand changes and trends (Robbins 2012).

2.4. Summary

The literature on stigma shows that there is clear evidence that disasters can have an impact on home value and can have lingering effects even after the remediation of the physical detriment. The traditional approaches to value are often insufficient in the effort to appraise homes on a large geographic scale and have deficits in cases where a large community has been affected. This leads to the literature on the alternative methods that may be adopted from other disciplines of research. Successfully adapting these alternative methods of analysis may reduce some of the cost and time related to using the traditional approaches in cases of mass disaster.

Chapter 3 Data and Methods

The purpose of this analysis is to determine whether changes in home values exist as a result of lasting market resistance (property stigma) following the natural gas leak in Porter Ranch, California. The project was designed to first identify if there is an effect from the gas leak on the value of homes and, if the effect exists, to separate the change in market value that is related to the gas leak from overall market/economic change.

The overall process for establishing a difference in the market values for the gas leak affected area consisted of several steps and is shown, generally, in Figure 3 and described below. First, the affected area (the subject area) was defined and the required datasets for the analysis were acquired (Section 3.1 and Section 3.2, respectively). Next, the comparable areas to the subject area in terms of social-economic values were identified using propensity score matching, as detailed in Section 3.3. Changes to the market values of the subject area were compared to the identified comparable areas in Section 3.4. Finally, a trend analysis was conducted to ensure that the results were meaningful.

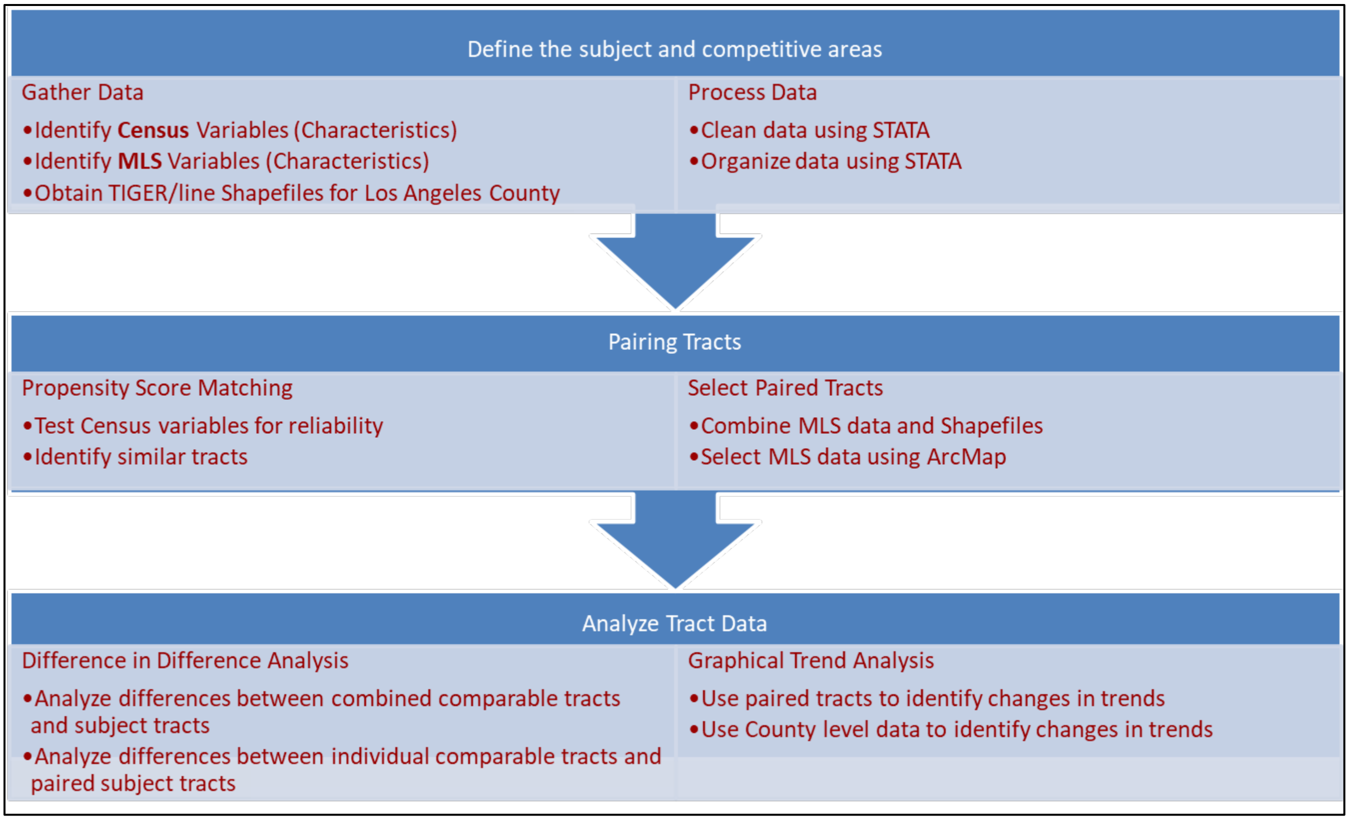

Figure 3 Overview of the workflow in this study.

3.1. Defining the Subject Area

The gas leak occurred in an open area north of the neighborhood called Porter Ranch in the City of Los Angeles. The City of Los Angeles defines the boundaries of neighborhoods, which are not recognized by the Census Bureau. Figure 4 shows the community map defined by the City of Los Angeles for the purpose identifying the Porter Ranch neighborhood. The neighborhood generally follows the boundaries of several combined census tracts.

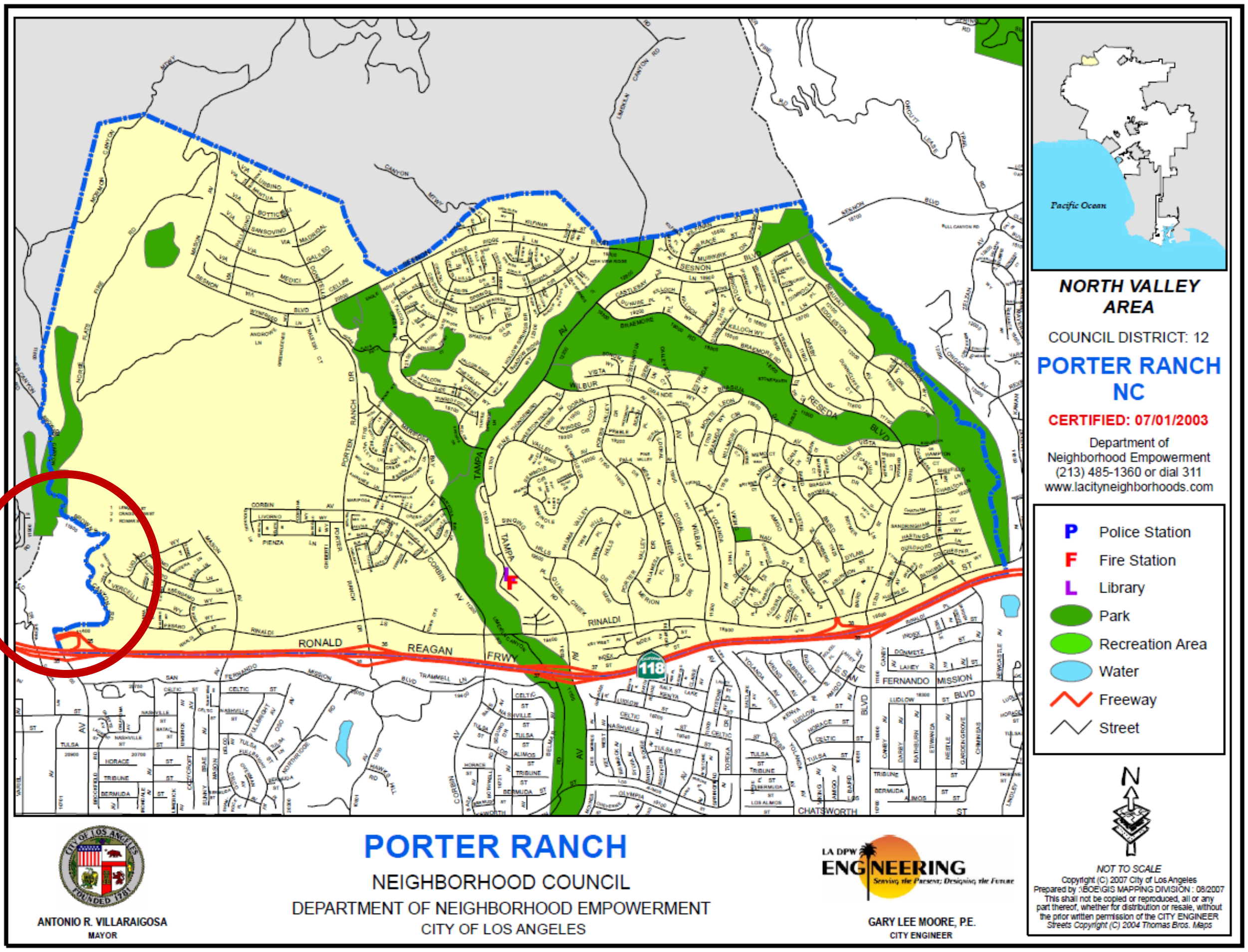

Figure 4 Porter Ranch Neighborhood Council boundary (Source: City of Los Angeles, 2007).

However, in order to conduct the analysis using the available data with defined administrative boundaries, an alternative subject area for this gas leak affected 'Porter Ranch neighborhood' must be redefined. For this purpose, the Census tracts that closely approximate the community of Porter Ranch were selected for this study. This includes six (6) census tracts. The Census tracts area and the City-defined neighborhood are very similar, but not exactly the same. The boundaries of the Porter Ranch subject area generally follows the outside boundaries of six contiguous census tracts as shown below in Figure 5 and City defined boundaries are very similar, except in the southwest corner of the subject area where the census tract includes slightly more area than the City-defined Porter Ranch neighborhood. Since there are virtually no homes located in this area (noted by a red circle in Figure 4), this difference should not alter the results of the analysis. Additionally, my study was intended to identify the immediate area affected by the gas leak and this portion of the census tract outside of the City-defined neighborhood area would be a part of the affected area.

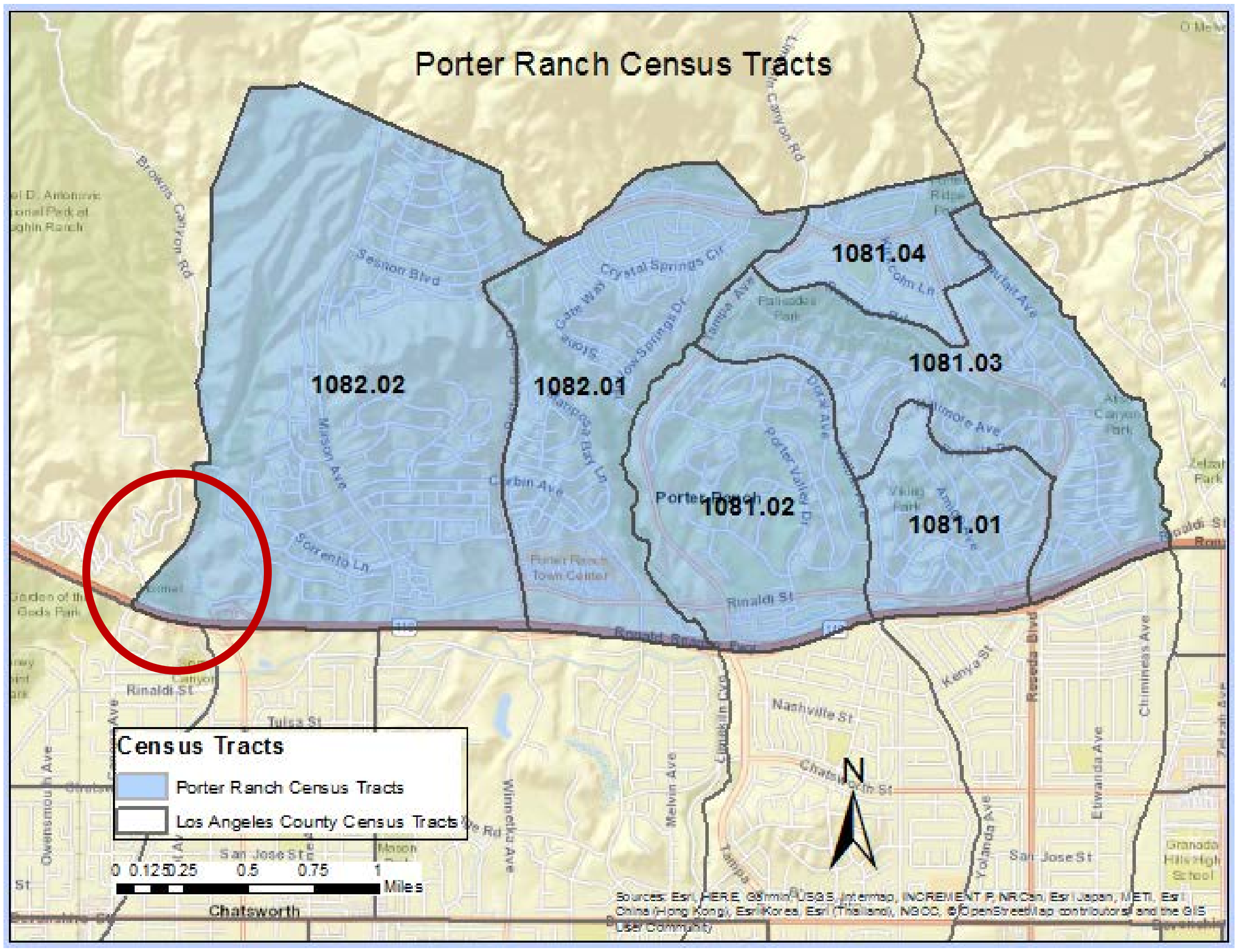

Figure 5 The subject area includes six census tracts covering the neighborhood of Porter Ranch.

3.2. Data Sources and Acquisition

Given the complexity of changes in real estate values, a number of sources were considered in order to illuminate the changes that result from the multiple market forces. Table 3 shows the list of the data sources used in this study.

Table 3 Table of data sources and examples.

Data source	Type of data	Example
American Community Survey	Demographic and housing	Year built, Household income
Southland Regional Association of Realtors	Home and sale transaction	Sale price, Sale date
TIGER/Line U.S. Census Bureau	Shapefiles	County and tract boundaries

3.2.1. *American Community Survey*

The primary data used in this analysis was from American Community Survey (ACS), provided by the U.S. Census Bureau. The ACS data is collected every year and has rich information and detail about the demographic composition of households. The ACS is based upon a sample of the population as compared to the decennial census that is a complete census of the entire population. Unlike the decennial census, the ACS survey is comprised of many questions that include specifics about individuals and households. Due to privacy concerns the ACS data is limited to the averages of multiple-year data for small geographic areas (e.g. Census tracts) only. In most cases, this is acceptable for longitudinal analysis that considers long-term trends.

In this study, I downloaded American Community Survey data (2011-2014) and included housing values and demographic information for analysis. The dataset would be used to consider overall market changes and find similar areas to the study area in terms of long-term trends in real estate value change and demographic shifts. The American Community Survey is considered a highly reliable data resources, although there is sampling error involved in any survey.

While there are numerous variables available in the ACS, the variables used in this analysis, shown Table 4 below, were selected in order to highlight the differences in Census tract composition in housing supply and demographics. These variables are further defined in Section 3.3.1.

Table 4 American Community Survey variables used in analysis.

Variable	Description
Total Units	Number of total units
Occupied	Percent of occupied units
Year Built	Year the home was built
Median Room Count	Median number of rooms per home
Owner Occupied	Percent of house that are owner occupied
Household Size	Median household size
Median Household Income	Median household income in US dollars

3.2.2. *Multiple Listing Service*

Real estate associations create and maintain private databases of the properties that are available for sale and are sold. These databases are known as multiple listing services (MLS). The MLS has a multitude of property characteristics, including specific address information, bedrooms, bathrooms, square footage, and sale price, etc. (Southland Regional Association of Realtors n.d.)

The MLS data used for this analysis was provided by the Southland Regional Association of Realtors® (SRAR). The majority of the basic characteristics in this MLS database are imported directly from the city and county of Los Angeles with additional information added by real estate agents/brokers that are members of the SRAR association.

While the data in the MLS is generally reliable, there are possible discrepancies between the data input and reality. This is a result of entry error or limitations of the public data. Incorrect aggregation of the agent input may result in total bathrooms of two (one full, two half), while the county data may show three bathrooms for the same property. Additional errors that may be present as a result of input issues including the existence of a pool, bonus room, or guest quarters. These characteristics can add or detract value to homes making the home value more or less accurate.

Although the MLS data is not perfectly accurate, it is regularly relied upon for analysis and is generally thought of as the most appropriate and useful when using the data in aggregate (Forgey, Rutherford and VanBuskirk 1994). Moreover, because the input errors are likely similar everywhere, it can be assumed that the overall data uncertainty of an area is the same as the overall data uncertainty for another area. The overestimate of one home is countered by the under estimate of another, therefore the tract as a whole is likely more accurate than an individual home.

One major advantage of MLS data over other sources of property data, such as public records, is that the MLS data has near real-time updates of the data and is more comprehensive than public records in regard to housing and transaction details. This type of data is also the data relied upon for federally regulated appraisal analysis used to determine the value of a property for lending and purchase purposes. This makes the MLS data the best qualified dataset in terms of home values for this analysis.

3.2.3. *Census TIGER/Line Shapefiles*

The TIGER/Line shapefiles for Los Angeles County were downloaded from Census website in order to align the MLS data with Census tracts of the subject area and the comparable Census tracts identified in the Propensity Score Matching. The TIGER/Line shapefiles are geospatial files that are constructed to align the data gathered from the Census bureau with the physical places they represent. The TIGER/Line shapefiles are the best available alignment files that can be used to align Census and other data with each other and with physical places. The shapefile used in this analysis is for the Census Tract and County as of 2016.

3.3. Methodology

In order to isolate and identify the potential change in value due to the gas leak, it is necessary to find comparable areas that show similar trends in value change over time prior to the gas leak. These similar areas are then used to compare to the subject areas for any deviation from the long-term trend in the subject area.

3.3.1. *Propensity Score Matching*

Propensity score matching was used in order to identify areas that were similar to the study area prior to the gas leak. As previously discussed in the literature review, propensity score matching selects comparable areas to the subject area based upon the results of a probit regression analysis. The matching controls for the differences between areas and the impact of the gas leak on the subject area. The analysis compares tracts within the subject area to those outside using several variables while controlling for the "treatment" effect, which in this case would be those that are within the subject area in which the gas leak occurred. This procedure is necessary to match comparable census tracts outside the Porter Ranch area with those within the area. For this study, a probit regression analysis was used to determine the likelihood of all observations (census tracts) being in the treatment group, based upon whether they were located in the Porter Ranch area, using the independent variables discussed above and the median home value as the dependent variable. The probit regression was used due to the binary nature of either being located within the gas leak area or not.

The next step of the PSM process is the scoring of the observations. The observations are scored based upon the probability of each observation being in the treatment group based upon the confounding (independent) variables (Pearl 2009). Using the STATA software, the matching was completed using a one-to-one pairing in order to find the most similar scores for comparable tracts to the scores of the subject tracts. Census tracts in Los Angeles County not impacted by the gas leak were then “matched” to the Census tracts within the gas leak area, based upon these characteristics. The “matched” tracts were then used in the difference-in-difference analysis discussed in the following section.

The 2015 ACS 5-year average data was used at the Census tract level for the propensity score matching in this study. This data was used to identify the Census tracts that are comparable to the subject Census tracts. However, multi-year data is less valuable when a comparison is necessary based upon a short period. Therefore, it is necessary to consider an alternative source, such as the MLS data, that have greater variability, such as sale price, and are available immediately prior to and after the gas leak. This more timely data was used in the next step of the analysis in order to highlight specific changes in value.

Identifying similar areas, using propensity score matching, is critical in order to conduct the next analysis, the difference in difference analysis. For the purpose of this analysis, similarity is defined as areas that portray similar property values and housing characters prior to the gas leak. By looking at the values of homes in multiple neighborhoods/communities in Los Angeles, "matching" communities were identified. The variables used in the final PSM analysis are listed in Table 5 and described in the following paragraphs. Initially, a broader set of variables was considered, including percentage of each race and educational levels of people in the tract, but these variables were found to not help in the differentiation of census tracts and where not found to have a determining effect on the home values in the census tracts. These variables were tested using STATA to determine if multicollinearity between variables exists and to see what impact the variables had on predicting the home value. In the end, the variables used match variables suggested in the literature regarding the determination of home values.

Table 5 Variable used for propensity score matching in this study.

Variable for PSM	Description	Source attribute in ACS
Total units (houses)	Number of total units	Total Units
Occupied units (houses)	Percent of occupied units	Occupied
yb_2014_later	Percent of housing built during or after 2014	Year Built
yb_10_13	Percent of housing built between 2010 and 2013	Year Built
yb_00_09	Percent of housing built between 2000 and 2009	Year Built
yb_90_99	Percent of housing built between 1990 and 1999	Year Built
yb_80_89	Percent of housing built between 1980 and 1989	Year Built
yb_70_79	Percent of housing built between 1970 and 1979	Year Built
yb_60_69	Percent of housing built between 1960 and 1969	Year Built
yb_50_59	Percent of housing built between 1950 and 1959	Year Built
yb_40_49	Percent of housing built between 1940 and 1949	Year Built
yb_39_earlier	Percent of housing built during or before 1939	Year Built
room_med2	Median number of rooms per home in tract	Median Room Count
own_occ	Percent of house that are owner occupied	Owner Occupied
hhsize_own	Median household size	Household Size
hh_income_med2	Median household income in US dollars	Median Household Income

- Total Units - In an effort to find comparable tracts, housing characteristics that define neighborhoods were used. The total number of units in a census tract was used to ensure that similar household structures exist in both the subject and comparable tracts.
- Occupied – The percent of occupied units in a census tract also provides an indication of the stage of the real estate and economic life cycle that a tract is in. For example, if a great number of units are vacant, then there could be a problem with the desirability of the neighborhood or there could be a lot of new development. Additionally, there could be a lot of run down units that are not habitable. Having similar occupancy rates helps to limit the differences in value that could result from this characteristic.
- Year built – The variable "Year Built" was categorized into 10 different variables that identify the percent of the housing supply that was built in each time grouping. This variable helps in the comparison of areas and identifies possible differences in areas that

may result from one area in a remodeling phase while another area is still in a declining phase. Another consideration in the comparison of Census tracts is that the median value of an area may be the result of the age of the housing supply. For example, differences in value changes may result from more new housing being built in one tract versus another. The intervals used are based upon the ACS data and provide for 10-year periods with the exception of the most recently built and oldest homes.

- Median Room Count – This variable represents the median number of rooms in the homes within the tract. The design and configuration of housing in the County of Los Angeles is so varied, that some areas may have homes that are very large with multiple rooms, while others may have fewer rooms in the same square footage. This variable identifies the total number of rooms, not just bedrooms. This provides a good indication of the utility of the home in a way that square footage does not.
- Owner Occupied – The percent of owner occupied homes in an area helps identify the make-up of the population and may indicate the stability of the values in the tract. A tract with a higher percentage of owner occupants is likely to have people with longer tenure and stability. The homes values in the area will also more closely represent the perceived value to the individual, rather than a value based upon the potential for future income to an investor.
- Household Size – This variable represents the median household size, the number of people in the household, for the Census tract. Given the diversity of households in Los Angeles, it is necessary to account for differing number of people living in similar sized homes. For example, the average two-bedroom house may have three people living in it in one neighborhood, but a similar home may have five people living in it in another

neighborhood. Including the household size in the analysis ensures that consideration of cultural and preferential differences are taken into account.

- Median Household Income – This variable represents the median household income for the Census tract. It is useful to include the household income because it helps to differentiate between Census tracts that have lower income households from those with higher incomes.

These variables are used in the propensity score matching to identify census tracts that have similar movements in property values and have similar characteristics in housing and household demographics out of the 2,340 total possible census tracts in the County of Los Angeles. Additional variables, such as the distance to the gas leak and distance to the city center were considered, but not included in this analysis. While variables of this type could have added richness to the analysis, some of the benefits that they represent (location) are captured by the home price. The demographic variables included in the analysis are not fixed, as people move and homes are destroyed and built, however, the location variables are mostly fixed in nature, with the exception of employment location changes over time.

A preliminary propensity score match was conducted using similar and additional variables at the PUMA level. The PUMA (Public Use Microdata Area) is a larger geography that provides a greater number of variables relating to individual household characteristics. Using the PUMA resulted in challenges since a single PUMA is larger than the subject area. This single geography did not provide enough variability for the propensity score matching to be successful. By reducing the number of variables and increasing the subject units (one PUMA to six Census tracts) the propensity score matching was successful in identifying comparable Census tracts. This process matched the subject census tracts to comparable census tracts outside the subject area. Once a comparable census tract was matched to a subject tract, it was not excluded from being matched to another subject tract. This was done to ensure that the best tracts were matched and not forced to match with the second best.

Spatial Join and Clipping were used to select and parse the MLS data by census tracts using the TIGER/Line shapefile in ESRI ArcMap ver. 10.6. The census tract features were used to select sale transactions in the subject tracts and comparable tracts for the difference-in-difference analysis and trend analysis. These sale transactions where then used in both analyses to help illuminate the differences between the area impacted by the gas leak and the peer Census tract.

3.3.2. *Difference in Difference*

The divergence in value changes was identified using a difference in difference (DiD) analysis. This considered the change in values over time in one area as compared to the changes in time in another area as shown in the equation below.

Difference in Difference (DiD)

= *(Change in values in subject area)* - *(Change in values in comparable area)*

The equation is comprised of four basic variables. The first step is to determine the change in the median price of homes in the Census tract affected by the gas leak. This was completed by subtracting the median home price in 2015 from the price in 2016. The difference was then divided by the price in 2015. The same process was used for the comparable Census tract. The next step was to subtract the percent change in the comparable Census tract from the percent change in the subject tract.

The difference in difference analysis acknowledges that the values of all areas can change over time but highlights the difference in the changes over time. If two areas are truly comparable and have historically moved in near unison, then there should be no difference in difference. However, if one area increases (decreases) in value at a faster rate, then there will be a larger difference-in-difference.

While DID analysis is often used to simply measure the differences in change, often without consideration for the direction of the difference (positive or negative). This analysis is concerned with the direction of the change in addition to the mere existence of a change. A negative difference in the change of the subject area would give greater merit to the notion that the gas leak contributed to a loss in value that has continued beyond the leak and clean-up.

For example, if the values increased by 10% in the subject area while the values increased by 15% in the comparable area and both had the same general economic conditions. Then it can be determined that the subject area increased at a slower rate (5% less). This could lead to the conclusion that the slower increase in values was due to the specific conditions of the subject area, namely the stigma related to the gas leak.

Using the areas that are comparable to the Porter Ranch area, as determined by the propensity score matching, a comparison of the changes in values in the subject area and the changes in values in the comparable area highlighted a divergence that could be attributable to the lingering effects of the gas leak (stigma).

A brief review of the data shows that since the gas leak was capped in February of 2016, there have been approximately 300 sale transactions of various types of homes in the Porter Ranch area each year since (Southland Regional Association of Realtors 2018). The sale transaction and housing data from the periods immediately before and after the gas leak was obtained from the MLS with the number of sales each year shown in Table 6.

Table 6 MLS sales in Porter Ranch from 2011 to 2016.

Year	Sales in Porter Ranch
2011	258
2012	285
2013	282
2014	273
2015	316
2016	276

3.3.3. *Trend Analysis*

As with any analysis where numerous characteristics (variables) are present, results can be misleading if the proper characteristics are not considered. Therefore, a trend analysis was performed to reinforce the results of the difference-in-difference analysis. The trend analysis highlights the changes in values for the subject Census Tracts along with the changes in values for the comparable tracts as identified in the propensity score matching. Additionally, a comparison of the combined subject area to the county was completed.

The trend analysis process relies upon the visual representation of the median home values graphed using Excel. The graphs are constructed using line graphs with specific points in time. The line graph is selected due to the visual representation of time and the trend that line graphs portray. The graphs were constructed for the periods both before the gas leak and after the gas leak with annual intervals on the horizontal axis and the median home value for the census tract on the vertical axis. The charts show the movement of median homes values obtained from the ACS data. The charts include values from 2011- 2016 in order to accurately portray the value movements and comparison both before the gas leak and after the leak was capped.

Chapter 4 Results

The results of the propensity score matching, difference-in-difference, and trend analyses are presented in this chapter to test the hypothesis of whether the gas leak had a lingering effect on the price of housing in the Porter Ranch community. The propensity score matching was used to identify appropriate similar Census Tracts to compare to the subject Census Tracts in the Porter Ranch area. The difference-in-difference results suggest that the home prices in the study area were negatively impacted by the gas leak. A trend comparison was conducted to ensure the reasonableness of the difference-in-difference analysis. The trend analysis results bolster the findings of the difference-in-difference analysis in that the conclusions suggest similar negative results for the Porter Ranch community.

4.1. Propensity score matching

As stated in Chapter 3, propensity score matching was used to determine the most similar census tracts to those within the Porter Ranch area. Initially, multiple variables were considering in the matching process. These included age grouping of the home, median income for the census tract, racial make-up of the tract, educational levels, etc. In the final analysis many variables were determined to not be relevant to highlighting the differences between the tracts. Therefore, only the variables identified in Chapter 3 were used in the final processing of the data.

The statistical software used in the cleaning of the data, analysis, and determination of the paired tracts was STATA. This software was selected due to the ease of use, having a GUI. The script used to clean and process the data is provided in the addenda.

After cleaning and organizing the data, the propensity score matching was conducted. The propensity score matching resulted in 5 comparable Census Tracts that mapped to the 6 subject Census Tracts. This is interpreted as 4 of the census tracts in the Porter Ranch area having a match that is solely comparable to them and two additional subject tracts pairing to the same comparable tract outside the Porter Ranch area. Figure 6 shows the subject tracts and the matching tract outside of the Porter Ranch area in relation to the County.

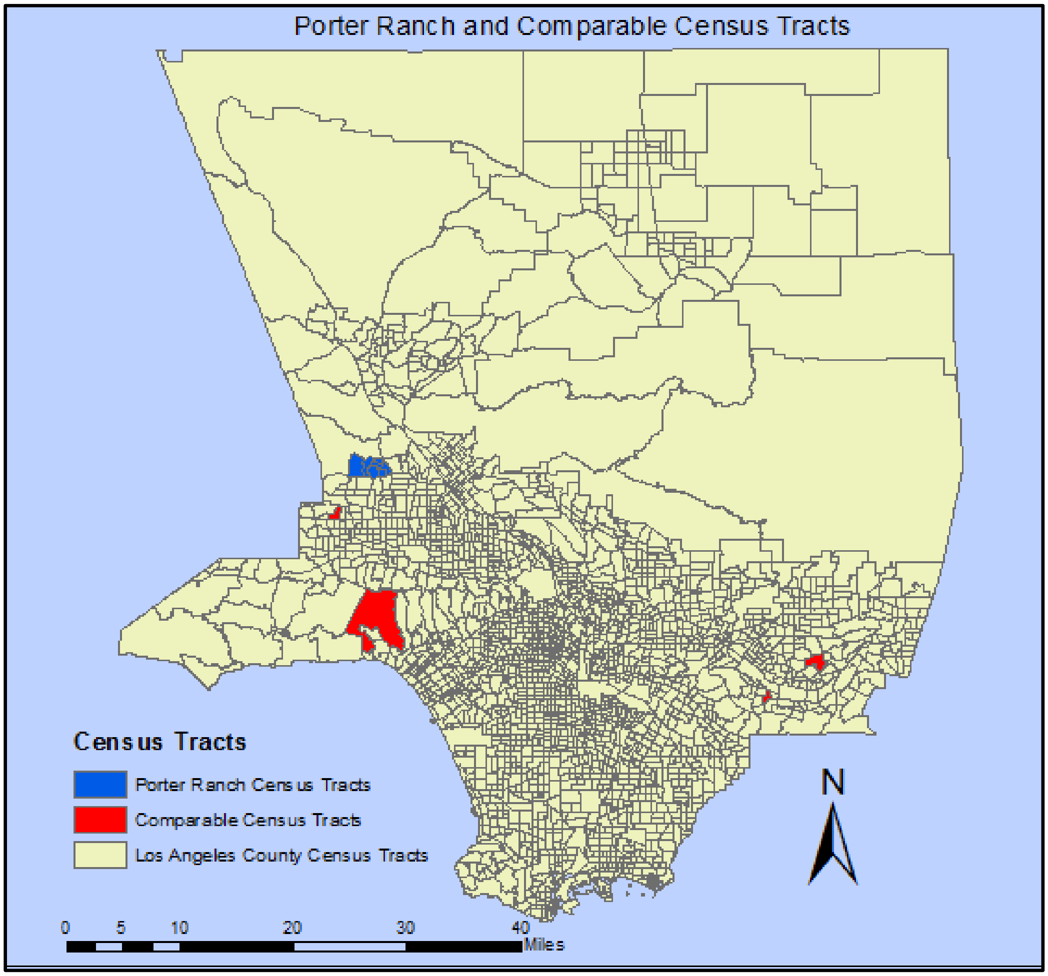

Figure 6 Porter Ranch and comparable census tracts in Los Angeles County.

Table 7 shows each Census Tract in the subject area is matched to another Census Tract in Los Angeles County. The subject areas are identified in column 3 as “subject” with the matched tracts noted with the “comparable” designation. The tracts numbers are listed in the “geography” column with the matched pair identification number listed under the “match” column. Subject tracts 1081.01 and 1082.01 are both matched to the same comparable tract of

1395.03. As described in Chapter 3, the Census Tracts were matched based upon several variables in order to determine the comparable tracts where home price changed similarly to the subject tracts. These are the same tracts shown in the map in **Error! Reference source not found.**.

Table 7 Census Tract Matching.

Geography	Match	Area Type
Census Tract 1081.02	114	Subject
Census Tract 4034.07	114	Comparable
Census Tract 1081.04	117	Subject
Census Tract 1132.32	117	Comparable
Census Tract 1081.01	118	Subject
Census Tract 1082.01	118	Subject
Census Tract 1395.03	118	Comparable
Census Tract 1082.02	122	Subject
Census Tract 4086.24	122	Comparable
Census Tract 1081.03	124	Subject
Census Tract 9800.19	124	Comparable

Table 8 below shows the results for the regression analysis used in the propensity score matching. The pseudo R-squared in the modified model is .4362, meaning that approximately 44% of the variability in the median values in census tracts can be explained by the model created. While a higher R-squared is preferred in most statistical analyses, it is common for the R-squared to be lower in propensity score matching due to the smaller sample size, therefore, traditional significance testing is often discouraged in propensity score matching (Austin 2011).

Table 8 Propensity score matching regression results.

```
Probit regression                               Number of obs   =        131
                                                LR chi2(15)     =      21.25
                                                Prob > chi2     =     0.1289
Log likelihood = -13.733677                     Pseudo R2       =     0.4362
```

subjectarea	Coef.	Std. Err.	z	P>|z|	[95% Conf.	Interval]
total_units	.0012522	.0009939	1.26	0.208	-.0006958	.0032002
occupied	.3485645	.2011061	1.73	0.083	-.0455962	.7427252
yb_2014_later	8.173297	5.68768	1.44	0.151	-2.97435	19.32094
yb_10_13	11.06468	6.963458	1.59	0.112	-2.583442	24.71281
yb_00_09	10.35898	6.66074	1.56	0.120	-2.695829	23.41379
yb_90_99	10.39242	6.671689	1.56	0.119	-2.683849	23.46869
yb_80_89	10.41116	6.680666	1.56	0.119	-2.682705	23.50503
yb_70_79	10.43086	6.689706	1.56	0.119	-2.680721	23.54244
yb_60_69	10.43747	6.689291	1.56	0.119	-2.673295	23.54825
yb_50_59	10.36136	6.670916	1.55	0.120	-2.713398	23.43611
yb_40_49	0	(omitted)				
yb_39_earlier	9.295595	6.043529	1.54	0.124	-2.549506	21.14069
room_med2	1.166837	1.04834	1.11	0.266	-.8878725	3.221547
own_occ	-.1234473	.0914289	-1.35	0.177	-.3026447	.0557501
hhsize_own	.4986468	1.340569	0.37	0.710	-2.128821	3.126114
hh_income_med2	.0000847	.0000489	1.73	0.083	-.0000111	.0001805
_cons	-1086.594	684.4749	-1.59	0.112	-2428.14	254.9523

Note: 47 failures and 0 successes completely determined.

Table 9 shows the median home value range for the Porter Ranch tracts and the comparable tracts in 2011 and 2016. The range of values for the comparable tracts is broader than that of the subject tracts in both 2011 and 2016. The propensity score matching paired the tracts based upon the trends in values and the demographic characteristics of the tracts, therefore the range in values differs between the comparable tracts and the subject tracts. The year that the homes were built and size of dwelling in the different tracts has an effect on the median value of the home. The key to the matching was identifying tracts that have similar demographic make-up and similar home value trends prior to the natural gas leak, with the weight being on the housing and demographic characteristics. The initial median home value of the comparable tracts was not intended to match that of the subject tracts; rather the similarity is in the structural make-up and the trends.

Table 9 Median home value range.

	Median Home Value	
Porter Ranch Tracts	Low	High
2011	$616,000	$ 832,500
2016	$567,200	$ 783,400
Comparable Tracts	Low	High
2011	$332,900	$ 934,200
2016	$323,100	$1,133,700

4.2. Difference-in-difference analysis

Using the difference-in-difference analysis, the comparison of the home prices between Porter Ranch and the identified comparable census tracts suggests a difference on the median homes price change between 2015-2016. Overall, the median home prices for the census tracts in the Porter Ranch area had an increase of 3.8 % while those for the comparable census tracts had an average increase of 7.6 % after the gas leak (Table 10). Although changes in census tracts were expected to vary year to year due to the variety of homes sold and the number of sales, the overall trend of the home prices typically follows historic patterns and volatility between tracts should be negated when the aggregate change for the community is considered.

As shown in Table 10, the individual census tracts within the Porter Ranch area and their comparable tracts changed at varying rates. The average change in home prices for all the six tracts in the Porter Ranch area is 6.7% less than the comparable tracts during the period of 2015 -

2016. In contrast, prior to the natural gas leak during the period of 2011 - 2014, the tracts in the Porter Ranch area had an average change of 3.0% greater than the comparable tracts. This suggests that there was a change in the home value trend between the areas that is not related to the overall market changes.

Table 10 Difference in Difference for census tracts 2015-16 and 2011-14.

2015 - 2016 Difference-in-Difference Results		
Subject	Comparable	Difference
1081.01	1395.03	-9.00%
1081.02	4034.07	-4.15%
1081.03	9800.13	-7.15%
1081.04	1132.32	-5.94%
1082.01	1395.03	-10.91%
1082.02	4026.04	-3.28%

2011 - 2014 Difference-in-Difference Results		
Subject	Comparable	Difference
1081.01	1395.03	21.00%
1081.02	4034.07	-16.22%
1081.03	9800.13	-12.00%
1081.04	1132.32	18.76%
1082.01	1395.03	17.15%
1082.02	4026.04	-10.92%

All but one of the census tracts in the Porter Ranch area showed an increase in values between 2015 and 2016. However, the percentage increases of values in the Porter Ranch area tracts were less than those in comparable tracts. This resulted in a slowing in the increase of the overall house values in the subject area as compared to the comparable areas.

It is evident that the variation in price changes between tracts prior to the gas leak is much greater than following the gas leak. This could indicate that micro-locational factors, such as street frontage, cul-de-sac location, etc. had more influence on home prices prior to the gas leak. The importance of variability of location in the Porter Ranch area was diminished in all subject tracts following the gas leak. Therefore, it suggests that the proximity to the leak factored greater in the price determination than the location or specific property characteristics that had been more important prior to the leak.

4.3. Graphical trend analysis

In order to provide context and back-up for the difference-in-difference results, a graphical trend analysis was performed. The following figure shows the overall Porter Ranch trend in relation to the combined comparable tracts. The trends lines prior to the gas leak appear to rise and fall fairly similarly and the degree of difference stays relatively constant. However, following the gas leak the Comparable average rises at a faster pace than the Porter Ranch area.

This change in slope and degree of difference provides support for the results of the difference-in-difference. As indicated in Figure 7, historically the median values in Porter Ranch (blue) tend to be higher and follow a similar trend as the combined comparable tracts (red). This trend holds true from 2010 through to 2014, however it changes following the gas leak in 2015.

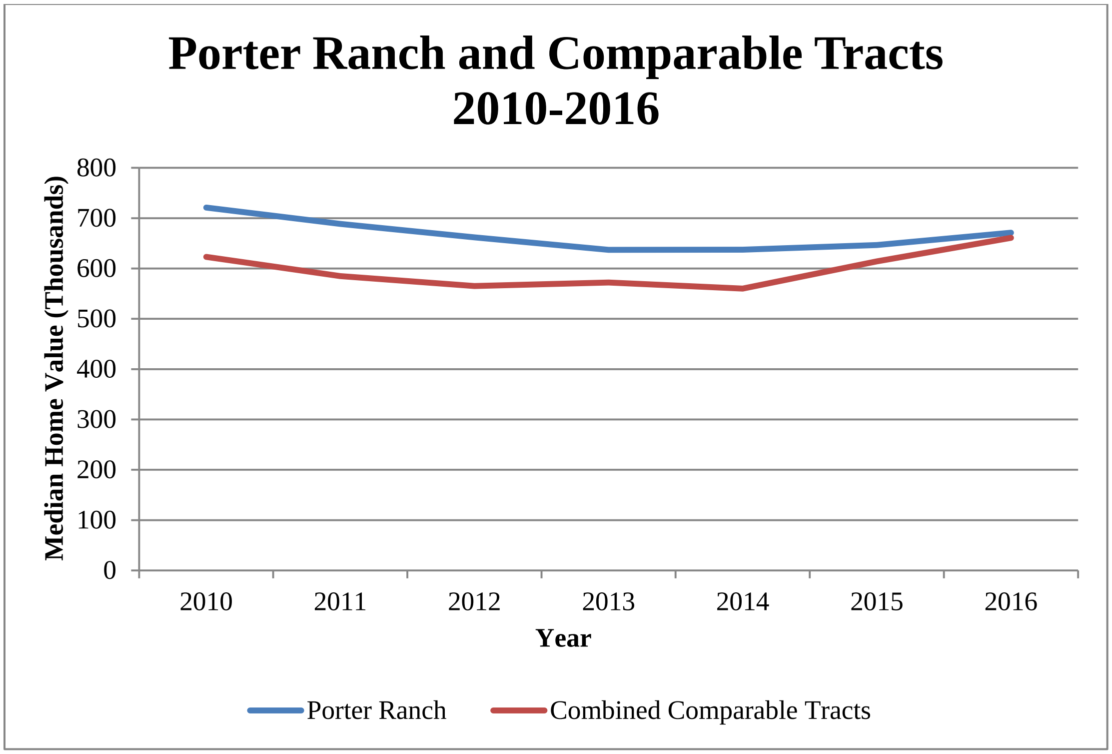

Figure 7 Trends of median home value for both Porter Ranch and comparable tracts between 2010 and 2016.

Both groups increase in value following the leak, however Porter Ranch does not increase as fast as the comparable tracts. Between 2010 and 2014, the median home value in Porter Ranch decreased by 11.6% while the reference group decreased by 10.1%. Between 2015 and 2016, the median home values increased in both groups with Porter Ranch increasing by 3.8% and the reference group increasing by 7.6%. Additionally, the historically higher home values in Porter Ranch converged in 2016 with the values in the reference group. The median home value was $670,950 for Porter Ranch and $660,760 for the reference tracts.

Comparing individual tracts within the Porter Ranch Community to the matched Census tracts shows more noise (volatility), but the overall trend is similar to that of the combined areas. The following figures (Figure 8 & Figure 9) show the individual Census tracts within Porter

Ranch (blue) with the matched tracts (red). These figures are further discussed in the paragraphs following the figures.

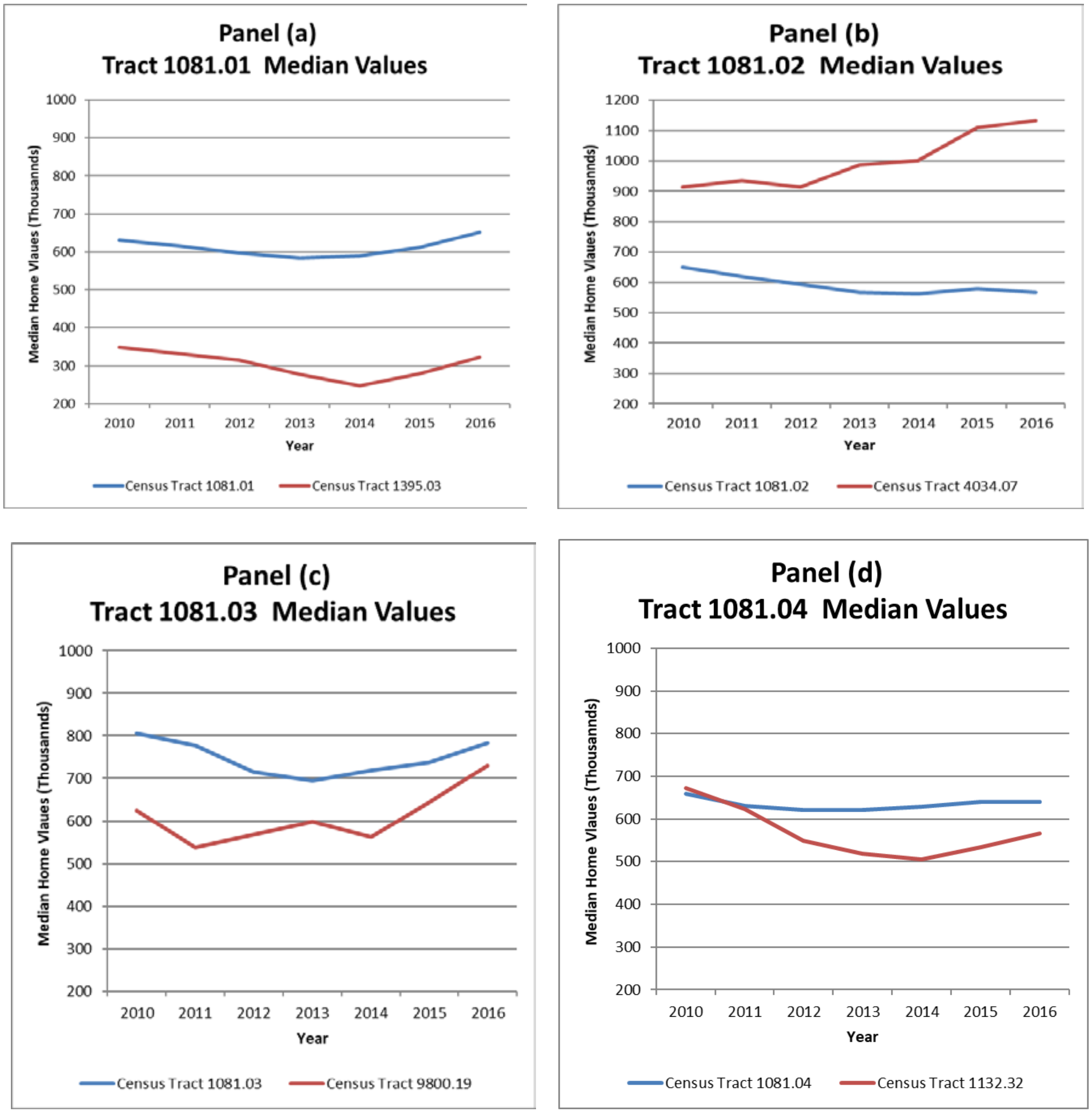

Figure 8 Trends of median home values for individual census tracts in Porter Ranch and comparable tracts (2010-2016), panels (a)-(c).

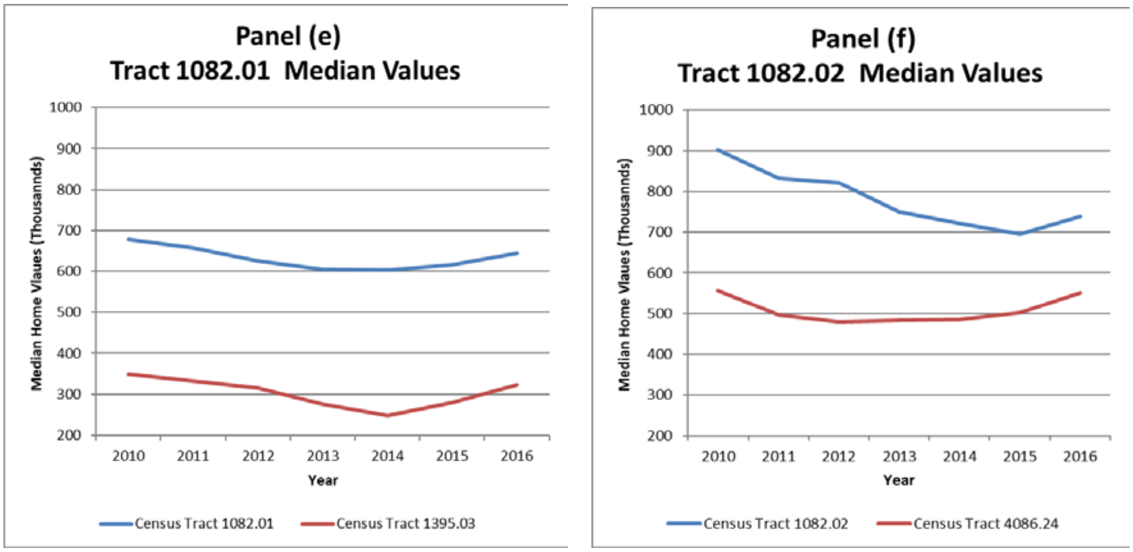

Figure 9 Trends of median home values for individual census tracts in Porter Ranch and comparable tracts (2010-2016), panels (e) and (f).

The values and value changes in census tracts, whether within the Porter Ranch area or for those comparable tracts, were highly variable (Figure 8 & Figure 9). Nonetheless, there was a consistency in the direction of the value movements. The trends for the median home values for all but one pair of the investigated tracts had a drop from 2010, and an increase in the later years (Figure (a), (c), (d), (e) and (f)). It is evident that this value increase trend in the Porter Ranch tracts (blue) does not show as fast an increase in the year following the gas leak. In Figure 8(b), Tract 1081.02 in the Porter Ranch subject area had a similar decrease in values since 2010 but continued the decrease trend following the gas leak. Its comparable tract (Tract 4043.47), on the other hand, had an overall increase trend from 2010 to 2016 (Figure 9(b)).

Comparing the median home values of the Porter Ranch subject area to that of Los Angeles County as a whole, a similar result to the individual tract comparisons is found (See Figure 10 below). The trend of recovery in home prices for the census tracts in the Porter Ranch area had a one-year delay compared to the entire Los Angeles County area. The increase in

Porter Ranch is slower than that of the county and the overall values are converging. This shows that the median value in Porter Ranch is not rising at the same pace as the County overall.

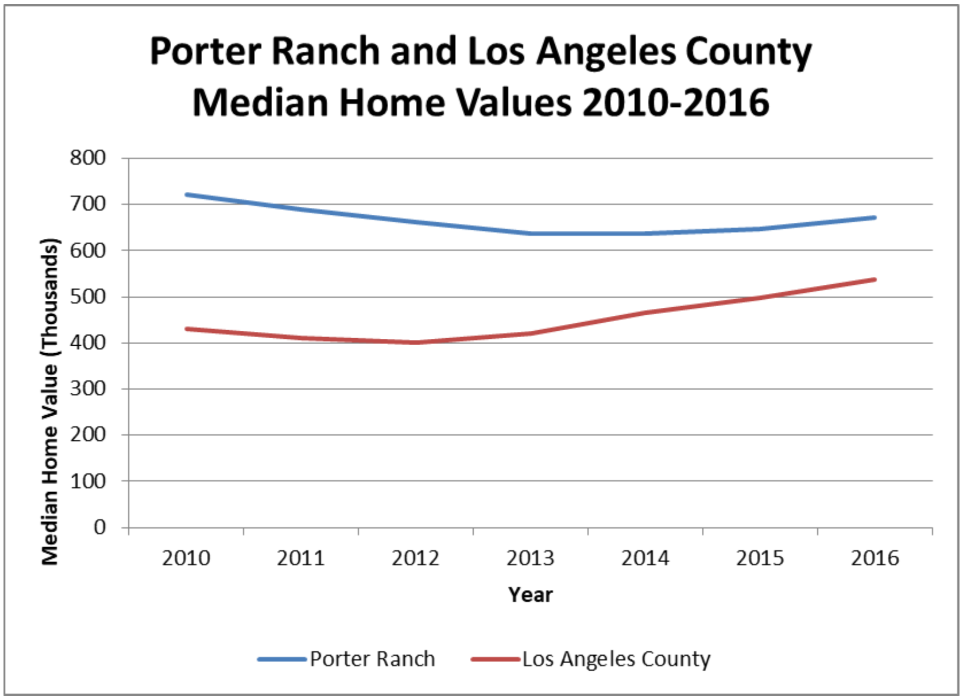

Figure 10 Comparison of median home values for Porter Ranch and Los Angeles County.

4.4. Discussion of results

The results of the Difference-in-Difference analysis suggest that there is a negative difference in the median home values between the Porter Ranch area and the reference areas, compared to the period before the gas leak. This is evidence that the gas leak could have had a detrimental impact on the values in the area.

The Trend Analysis shows differences between the comparable areas and Porter Ranch with most tracts having lower rates of increase as compared to the comparable tracts. This also provides support to the possibility that the median home values in the Porter Ranch area were inhibited by the gas leak, even after the gas leak was capped. The overall trend in the combined tracts and the County also increase faster than the Porter Ranch area, after the gas leak.

Both analyses support each other in the results, indicating that the home values in the comparable areas increased at a faster rate than the Porter Ranch area. If there are no overlooked attributes that could have contributed to the differences in the home value changes, the analysis suggests that the gas leak was the factor that slowed the increase in home values in the Porter Ranch area.

Chapter 5 Discussion and Conclusions

This study successfully suggests that differences of changes in values for homes in the Porter Ranch area exist after the gas leak occurring in 2015. Using propensity score matching, I was able to determine comparable Census tracts to the six Census tracts in the Porter Ranch subject area by using the housing and demographic variables. These comparable tracts were analyzed before and after the gas leak to determine if any significant changes in home values occurred during this period. Based upon the difference-in-difference analysis the study suggests that the Porter Ranch community was negatively impacted by the gas leak as far as home values are concerned. The trend analysis further reinforced the results by showing that the housing market performance of the comparable tracts, and the county as a whole, surpassed the tracts in the Porter Ranch community.

The following sections further discuss the known limitations of the study, potential future avenues of research, and the broader impact of this study.

5.1. Limitations of Study

With the attempt to capture the effects of an environmental incident on the value of individual homes, the results of this study were compelling. However, there are limitations to the results and to the study itself. These limitations are discussed in the following sub-sections.

5.1.1. *Extrapolation*

While the results of this research suggesting a difference in the home values in the Porter Ranch area as compared to the reference areas under the influence of the gas leak incident, care should be taken in applying the results to other use cases. The changes in home values for this study may not occur in the same manner in a different region or for other types of environmental incidents. Variations should be expected if the natural or man-made disaster is of a different type, has greater notoriety, or has lingering/visible effects (Cavallo and Noy 2009).

5.1.2. *Variable selection*

The variables selected for the propensity score matching were critical in determining the census tracts comparable to those in the Porter Ranch area. In an ideal scenario, additional variables, such as homeowner's age, wealth, ethnicity, and education may be used to improve the matching. However, the types of demographic characteristics available for the Census tract data used for this study was limited, given the accuracy and temporal data requirements. Additional variables, such as the distance to the gas leak and distance to the city center were considered, but not included in this analysis. These geospatial variables could have provided the insight of spatial relationships to the propensity score matching. The demographic variables included in the analysis are not fixed, as people move and homes are destroyed and built, however, the location variables are mostly fixed in nature, with the exception of employment location changes over time.

In the original plan of this study, the unit of analysis for the propensity score matching would have been the Public Use Microdata Area (PUMA) level data from the Census. This PUMA level of data is rich in the number of demographic variables and is available on an annual basis. However, the PUMA for the Porter Ranch community includes multiple other neighborhoods that were not impacted by the gas leak. A pilot study of propensity score matching using the data of this PUMA resulted in a singular comparable PUMA, which would not have provided the variation in housing age and housing type (small/large, condo/detached home) needed between the housing within the Porter Ranch community as the level of the census tracts data did. Therefore, the Census tract level data was used.

5.2. Future Work

There are several potential avenues of future work in this research subject in order to deepen the understanding of the gas leak impact on the home values in the Porter Ranch area. Future research can take different directions or include additional variables and considerations, as shown in the following sections.

5.2.1. *Qualitative research on Porter Ranch*

The existing study could be enhanced through the process of interviewing participants in the real estate market of the Porter Ranch and surrounding areas. The interviews can include questions to both real estate professionals and homeowners on the motivations and criteria of home purchases and selling. Through these interviews, the motivations and criteria information can be used as calibers to improve the variable selection for analysis.

The motivation of the participants in home transactions immediately after the gas leak can also be used to improve the understanding on the root causes of the home price change at the time. It may become evident that sellers were selling at lower prices because they had expected the real estate stigma of the gas leak that might deter buyers. Alternatively, it may be found that buyers used the past leak as a bargaining tool. Another possible scenario would be the exclusion of the Porter Ranch area for home purchases from some homebuyers due to the real estate stigma or the perceived future threats in the area. Therefore, a truly useful qualitative study would need to include interviewing the real estate professionals who helped homebuyers select and complete purchases in areas similar to the Porter Ranch area.

5.2.2. *Additional variable selection*

As stated in Section 5.1, the selection of variables and geospatial unit of analysis were not ideal in this study. Rather, these were the best available data of the area and the time. Future research on this topic may include a broader, larger set of variables in both demography and property data. This could be achieved through the use of the restricted Census data at a bona fide Census Data Center as well as the creation of geospatial-related variables using a GIS software. With access to the restricted data, detail into household characteristics is obtainable. This may provide different and potentially more accurate results in the propensity score matching process and possibly different results in the difference-in-difference and trend analyses.

5.2.3. *Longitudinal data availability*

Four years of demographic and home value data before the gas leak and immediately following the capping of the leak and clean-up were taken into consideration in this study. The difference-in-difference analysis and trend analysis may have been more accurate if more longitudinal data was available. Having a longer window of demographic data and home values would help to validate the results, but given the recentness of the incident, a larger temporal scale of data is not yet available. The impacts on home values identified in this analysis can be short-term impacts that may be negated with time. However, this study is still crucial in terms of the timeliness to people affected by the gas leak.

Adding more years of data could also improve the accuracy of the results. After a few more years pass, more sales transactions in the Porter Ranch area and in comparable areas will be available for analyzing. Stable and long periods of the data can ensure that the findings of this study are not simply the results of market changes or structural changes to the housing supply. On the other hand, a larger temporal scale of the data may also reveal a turning point at which the gas leak no longer has a significant impact of the home values. This potential result could prove a real estate stigma to environmental disasters (e.g. a gas leak) diminishes as time passes.

5.2.4. *Disaster comparison*

Another potential avenue of research could include the comparison of the after effects of other natural/man-made disasters to the Porter Ranch incident. Selection of the comparison area needs to consider the type of real estate stigma or other "lingering" effects that may be actual or perceived, since some environmental disasters have physical lasting effects or a clear potential for a recurrence (Sanders 1996). For example, a dam failure has lasting effects, such as flooding, until repairs are completed, and poses a future threat if the dam is not decommissioned. The Porter Ranch gas leak has been capped, highly monitored, and will not likely be pressurized to the capacity it was in the past. Therefore, the real losses in home values are likely due to perception rather than future actual failure.

5.3. Broader Impact of the Study

This study provides the evidence of losses in home values within the Porter Ranch community because of the gas leak. The finding is important to the people living in the community and the surrounding areas because it supports their ongoing battles to maintain safeguards and the pending lawsuits surrounding the gas leak. Settlements related to the cases of negative environmental impacts are often confidential. Therefore, limited past case information is available to litigants. By conducting the analysis of the home values in the Porter Ranch area, it is clear that home values did not increase at the same rate as they had been prior to the gas leak and this supports the claims of individual homeowners who may have suffered a loss due to the gas leak.

One of the highlights of this study is to provide a new approach in determining losses in home values at a larger spatial extent. Prior to this study, most cases in home value losses rely upon individual appraisals that compare the subject home to three or more similar properties or based upon the cost approach (Chan 2001). However, this traditional approach becomes challenging to evaluate home values when a smaller spatial-scale (larger spatial extent) event, such as the gas leak, occurs. The challenge here is in the ability of the appraiser to find homes with similar characteristics in a similar area that have encountered similar changes in the overall market. By using propensity score matching, the need to identify a comparable area through trial and error is eliminated and the use of difference-in-difference analysis reduces the subjectivity of the analyst (appraiser). Standardized and quantitative methods, such as the approach used in this study, provides statistically sound results to rely upon when mass appraisal is necessary.

As our living earth experiences climate change and more manmade and natural disasters are expected, the need for determining mass losses inevitably increases. This study furthers the research by employing a quantitative approach with traditional and geospatial datasets, and the integration of traditional software tools with geospatial tools, resulting in a valuable addition to the field of property valuation in large-scale disasters.

References

Adkins, Jana. 2015. *SCV Hotel Room Demand Exceeded Supply During Gas Leak.* Last modified: May 25, 2015. http://archive.signalscv.com/archives/152568/.

Alfert, R., H. W. Collison, and G. W. Tate. 2005. "Recovering "Stigma" Damages in Mold-related Construction Defect Cases: Making the Property Owner Whole." *Florida Bar Journal* 79: 78-82.

Appraisal Standards Board. 2018. *Uniform Standards of Professional Practice (USPAP). Accessed April 16, 2018.* Washington DC: The Appraisal Foundation.

Austin, Peter C. 2011. "An introduction to propensity score methods for reducing the effects of confounding in observational studies." *Multivariate Behavioral Research 46*: 399-424.

Bell, Randall. 1998 "The impact of detrimental conditions on property values." *The Appraisal Journal* 66: 380-391.

Benjamin, John, Randall Guttery, and C.F. Sirmans. 2004. "Mass appraisal: an introduction to multiple regression analysis for real estate valuation." *Journal of Real Estate Practice and Education* 7: 65-77.

Blake, Paul. 2016. *How many cars and burping cows equal the California gas leak?* Last modified: January 11, 2016. http://www.bbc.com/news/world-us-canada-35258036.

Bleich, Donald H, M Chapman Findlay III, and G Michael Phillips. 1991. "An evaluation of the impact of a well-designed landfill on surrounding property values." *The Appraisal Journal* 59: 247-252.

Bloom, Bryan, MaryEllen Noble, and Michael Nobe. 2011. "Valuing green home designs: A study of ENERGY STAR® homes." *Journal of Sustainable Real Estate* 3: 109-126.

Boyle, Melissa, and Katherine Kiel. 2001. "A survey of house price hedonic studies of the impact of environmental externalities." *Journal of real estate literature* 9: 117-144.

Bryson, Alex, Richard Dorsett, and Susan Purdon. 2002. *The use of propensity score matching in the evaluation of active labour market policies.* Working Paper, London: Policy Studies Institute and National Centre for Social Research.

Caliendo, Marco, and Sabine Kopeinig. 2008. "Some practical guidance for the implementation of propensity score matching." *Journal of economic surveys* 22: 31-72.

Cavallo, Eduardo A, and Ilan Noy. 2009. *The economics of natural disasters: a survey.* Working Paper, Inter-American Development Bank.

Chan, Nelson. 2001. "Stigma and its assessment methods." *Pacific Rim Property Research Journal* 7: 126-140.

Clark, Steven P, and Daniel T Coggin. 2009. "Trends, cycles and convergence in US regional house prices." *The Journal of Real Estate Finance and Economics* 39: 264-283.

Clayton, Jim. 1998. "Further evidence on real estate market efficiency." *Journal of Real Estate Research* 15: 41-57.

Conley, S., G. Franco, I. Faloona, Donald R. Blake, J. Peischl, and T. B. Ryerson. 2016. "Methane emissions from the 2015 Aliso Canyon blowout in Los Angeles, CA." *Science*: aaf2348.

Donald, Stephen G, and Kevin Lang. 2007. "Inference with difference-in-differences and other panel data." *The review of Economics and Statistics* 89: 221-233.

Dotzour, Mark. 1990. "An empirical analysis of the reliability and precision of the cost approach in residential appraisal." *Journal of Real Estate Research* 5: 67-74.

Dubin, Robin, Kelley Pace, and Thomas Thibodeau. 1999. "Spatial autoregression techniques for real estate data." *Journal of Real Estate Literature* 7: 79-895.

Forgey, Fred, Ronald Rutherford, and Michael VanBuskirk. 1994. "Effect of foreclosure status on residential selling price." *Journal of Real Estate Research* 9: 313-318.

Golberg, Hilary M., and Ray Calnan. 2016. "On Parr: The Use and Propriety of Appraisal Methods in Computing Fracking Awards." *Journal of Law & Commerce* 35: 55-77.

Hastings, Annette. 2004. "Stigma and social housing estates: Beyond pathological explanations." *Journal of Housing and the Built Environment* 19: 233-254.

Jackson, Thomas. 2001. "The effects of environmental contamination on real estate: A literature review." *Journal of Real Estate Literature* 9: 91-116.

Kang, Han-Bin, and Alan K Reichert. 1991. "An empirical analysis of hedonic regression and grid-adjustment techniques in real estate appraisal." *Real Estate Economics* 19: 70-91.

Kastellec, Jonathan P, and Eduardo L Leoni. 2007. "Using graphs instead of tables in political science." *Perspectives on politics* 5: 755-771.

Kinnard J, William N, and Elaine M Worzala. 1999. "How North American appraisers value contaminated property and associated stigma." *The Appraisal Journal* 67: 269-279.

Leung, Charles Ka Yui. 2003. "Economic growth and increasing house prices." *Pacific Economic Review* 8: 183-190.

Lin, Steven AY, ed. 2014. *Theory and measurement of economic externalities.* New York: Academic Press.

Mundy, Bill. 1992. "Stigma and value." *The Appraisal Journal* 60: 7-13.

National Association of Realtors. n.d. *Multiple Listing Service (MLS): What Is It.* Accessed March 26, 2018. https://www.nar.realtor/nar-doj-settlement/multiple-listing-service-mls-what-is-it.

Patchin, Peter J. 1988. "Valuation of Contaminated Properties." *Appraisal Journal* 56: 7-16.

Pearl, Judea. 2009. *Causality*. Cambridge: Cambridge university press.

Puhani, Patrick A. 2012. *The treatment effect, the cross difference, and the interaction term in nonlinear "difference-in-differences" models.* Discussion Paper, Bonn: Institute for the Study of Labor: 1-7.

Rattermann, Mark. 2009. *The student handbook to the appraisal of real estate.* Washington DC: Appraisal Institute.

Reuters. 2018. *SoCalGas lifts cost of Aliso Canyon natgas leak to over $1 billion*.

Ridker, Ronald G, and John A Henning. 1967. "Ridker, Ronald G., and John A. Henning. "The determinants of residential property values with special reference to air pollution." *The Review of Economics and Statistics*: 246-257.

Robbins, Naomi B. 2012. *Creating more effective graphs.* Wiley.

Roddewig, Richard J. 2001. *Valuing Contaminated Properties: An Appraisal Institute Anthology*. Washington DC: Appraisal Institute.

Sanders, Michael V. 1996. "Post-repair diminution in value from geotechnical problems." *Appraisal Journal* 64: 59-66.

Simons, Robert A, and Ron Throupe. 2005. "An Exploratory Review of the Effects of Toxic Mold on Real Estate Values." *Appraisal Journal* 73: 156-166.

Simons, Robert, and Jesse Saginor. 2006. "A meta-analysis of the effect of environmental contamination and positive amenities on residential real estate values." *Journal of Real Estate Research* 28: 71-104.

Smith, Kerry V, and Ju Chin Huang. 1993. "Hedonic models and air pollution: twenty-five years and counting." *Environmental and Resource Economics* 3: 381-394.

South Coast Air Quality Management District. 2018. *Aliso Canyon Natural Gas Leak: Air Monitoring Results.* Los Angeles: South Coast Air Quality Management District.

—. "Aliso Canyon Update. 2018. "*South Coast Air Quality Management District.* Accessed May 12, 2018. http://www.aqmd.gov/home/news-events/community-investigations/aliso-canyon-update.

Southland Regional Association of Realtors. n.d. *MLS Matrix Information.* Accessed December 21, 2018: http://www.srar.com/mls/?p=matrix_information&sub=matrix_information.

Steven Ruggles, Katie Genadek, Ronald Goeken, Josiah Grover, and Matthew Sobek. 2017. "Integrated Public Use Microdata Series: Version 6.0 [Machine-readable database]." Minneapolis: University of Minnesota.

The Appraisal Institute. 2008. *The Appraisal of Real Estate.* 13th Edition. Chicago, IL: The Appraisal Institute.

Weissgerber, Tracey L, Natasa M Milic, and Stacy J Winham. 2015. "Beyond bar and line graphs: time for a new data presentation paradigm." *PLoS biology 13*: e1002128.

Wheaton, William C. 1999. "Real estate “cycles”: some fundamentals." *Real estate economics,* 27: 209-230.