Spring 5-10-2018

How the Proximity of Crime Impacts Housing Prices: A Hedonic Pricing Study of Inner-Loop Houston, TX.

Samuel D. Thompson

Samuel Thompson

Department of Geography and Environmental Studies

This thesis is approved, and it is acceptable in quality and form for publication:

Approved by the Thesis Committee:

Christopher Lippitt, Chairperson

John Carr

Robert Berrens

How the Proximity of Crime Impacts Housing Prices: A Hedonic Pricing Study of Inner-Loop Houston, TX.

By

Samuel Thompson

Bachelors of Science, Economics, Texas Christian University

Thesis

Submitted in Partial Fulfillment of the

Requirements for the Degree of

Master of Science

Geography

The University of New Mexico

Albuquerque, New Mexico

July 2018

How the Proximity of Crime Impacts Housing Prices: A Hedonic Pricing Study of Inner-Loop Houston, TX.

By

Samuel Thompson

B.S., Economics, Texas Christian University, 2011

M.S., Geography, University of New Mexico, 2018

ABSTRACT

This research used data on over 2200 house sales in inner-city Houston, TX to estimate the impact of crime on house prices. A GIS was used to tabulate home sale data from the MLS, neighborhood characteristics and crime data published by the Houston Police Department. Hedonic Pricing Models were built to assess how crime effects housing price and if proximity to the criminal event matters. Results show that crime does have a measurable impact on home sale price and that proximity of crime is important. An increase in number of violent criminal events leads to a discount in home price. An increase in all criminal events (violent and non-violent) leads to an increase in home price, suggesting a dichotomy in how/when crime is reported to police.

Table of Content

1. Introduction

The perception of place and the characteristics that affect that perception are undeniably important in the valuation of a community. As one of those characteristics, crime holds a unique place in this valuation in that crime creates fear; this is not a financial fear that rising and falling home prices may lead to, or a fear in the detraction in a families standard of living such that might come about with the removal of access to standard amenities. This fear is based off of direct, rather than indirect, personal harm.

How then are we to measure the level of fear that crime adds to a community? How can a concept as abstract as this be quantified for academic study? The answer to this brings us to a technique commonly used in economics to measure the monetary value of heterogeneous goods. Hedonic pricing (Rosen 1974) can be applied to measure the impact that crime has in the real estate market. By taking the sale price of a home and regressing that with a list of structural and neighborhood characteristics, we can pinpoint the price-impact that one of those characteristics has had on the price of the home. By including a measure of crime, we can show how crime is influencing the price of property. The ability to examine this relationship is important to better understand how home buying decisions are made. This can help city planners, police departments and real estate professionals improve a person's quality of life through better understanding of how people value the externalities of their community.

Specifically, this research is centered on determining how the effect of crime on housing prices varies by proximity. To do so, two sub-questions will be examined: 1)

Does crime rate have a measurable effect on housing prices in inner-loop Houston, and 2) does the proximity of crime affect this relationship? It is the goal of this project to model the extent (proximity by distance) to which reported crime impacts the purchase price of a residential property.

Value is added through this research by the application of hedonic pricing methods to the relationship between reported crime rates and real estate prices. Using a geographic perspective while keeping to a foundation of economic theory creates a more complete understanding of a problem through the inclusion of spatial context. This study borrows a commonly used practice in measuring the economic impact of urban greenspace by using proximity (Bark et al. 2011, De Vor et al. 2009, Price et al. 2010). Separating crime by distance is unique to this project.

2.0 Literature Review

2.1 Hedonic Pricing

Hedonic price models have been extensively utilized within urban, environmental, and real estate literatures (e.g., Colby & Wisehart 2002, DeVor & DeGroot 2011, Graves et al. 1988, Haab & McConnell 2002, Kim & Wells 2005, Mooney & Eisgruber 2001, Payton et al. 2008, Price et al. 2010). The estimation of housing price through multivariate regression analysis allows for decision makers in planning and government to utilize funds raised through the tax base to maximize value for residents and future homebuyers. By informing the future allocation of resources, these estimations have real value in their ability to predict changes over time and across space.

Hedonic pricing is an economic model for determining the price of a good without real price information. Hedonic pricing and hedonic models become concerned with the measure of price based on quality of characteristics of heterogeneous goods.

The usefulness of the hedonic model is maximized when the factors affecting value of a product being studied are heterogeneous in nature, because of the need to isolate differences between products in order to isolate difference in price. Rosen (1974) was the first to introduce this topic in economic literature. His paper was concerned with the development of a model to use product differentiation as a means to determine economic value within the market. His model of hedonic prices claimed that the price of a good is determined by a vector of other objectively measured characteristics. The known product price is a function of those measured characteristics. When multivariate regression is run between the observed price and the vector of characteristics we get the "hedonic" price, an estimated price of goods relative to the characteristics of said good.

Later, Dubin & Goodman (1982) described the principle behind the hedonic model in terms of a consumer's demand for food. They tell a tale of a grocery shopper wandering through the supermarket adding items to a shopping basket. An economist following the shopper through the market would be able to record to shopper's behavior and thus through several iterations learn about the sources of consumer demand. If however, the shopper was limited to purchasing only already filled shopping carts with varying "bundles" of goods, the economist would have a much more difficult time determining the shopper's valuation of specific goods, necessitating a hedonic modeling approach.

The real estate market acts in a similar way to the story of the grocery shopper. When purchasing a house, decision making is typically limited to an already existing stock of houses with their one bundle of characteristics; square footage, number of bedrooms, garage spaces, presence and type of heating and cooling functions, etc. Hedonic models isolate characteristics in the shopping cart by regressing the price of the product (the price paid for the already filled shopping cart or already built home) with the chosen characteristic. When applied, the hedonic price of a house is dependent upon the structural and neighborhood characteristics of that house.

The linear form of the model can be written as (Payton et al. 2008):

$$P = \beta_0 + \Sigma\beta_k S_k + \Sigma\beta_j L_j + u \qquad eq.1$$

where P represents a vector of housing prices, S_k represents a matrix of structural characteristics, L_j represents a matrix of neighborhood characteristics, b_0, b_k, and b_j, represent matching parameters and u represents a vector of random errors.

To put this into a more simple illustration (and ideal for more complete understanding outside of the economic discipline), Lynch & Rasmussen (2001) summarize the function as such:

$$P_i = f(S_i, N_i, C_k) \qquad eq.2$$

where P_i represents the selling price of a home, S_i represents structural characteristics, N_i represents neighborhood characteristics, and C_k represents the isolated variable of interest (in this case a statistic on crime).

More recent studies have concluded that simple linear models can be overly restrictive (Dubin & Goodman 2013) by implying that each observation of a characteristic has the same marginal valuation as the next. This would mean that the first and second bedroom of a house is valued equally as the seventh, eighth and beyond. Recent work has emphasized interactive forms (such as log-lin or log-log models) which provides more flexibility in the model.

2.2 Hedonic Pricing Studies and the Effects of Crime

Traditionally, hedonic models have been in the business of prediction – specifically in real estate and the forecasting of sales figures for potential homebuyers by real estate professionals. The study of the effects of neighborhood amenities including both positive and negative externalities has seen a rise in publications over the past couple of decades. The vast majority of these papers have been concerned with measuring the impact pollutants have on home purchasing behavior (De Vor et al. 2011, Graves et al. 1988). Additionally, environmental economists have become increasingly aware of the positive impact of green vegetation within a community such as percentage tree cover, vicinity to riparian corridors, presence of parks and recreational opportunities and different types of greenspace in general (Bark et al. 2011, Colby & Wishart 2002, Haab & McConnell 2002, Mooney et al. 2001, Payton et al. 2008).

Recently there has been an increasing emphasis on the effects of crime and crime patterns on the housing market. Lynch & Rasmussen (2001) attempt to model these effects through the weighting of the "cost of crime to victims" by seriousness of offense. Their results showed that crime has virtually no impact on house prices except in high crime areas. This dichotomy of effect between high and low crime areas is found in the majority of other studies. Tita et al. (2006) disaggregate crime to the census tract in order to localize the effect on the housing market. Using a time-series approach, another example can be found in a piece dealing with the willingness to pay to avoid violent crime by Bishop & Murphy (2011) – they find that crime is realized at varying rates for poor, middle class and wealthy neighborhoods. The results suggest that "high crime" and “low crime” is highly dependent on reporting tendencies and contrasts between high wealth and low wealth areas. It was found that violent crime tends to have the greatest impact on housing markets, which could be a simple function of the reporting rate of violent crime across lines of wealth. Buonanno et al. (2013) reiterated the limited effect of crime across housing markets except for the highly discounted high crime areas. Results showed that a one standard deviation increase in the perceived security of a district leads to a 0.57% increase in that districts valuation, while houses are discounted (1.27%) in districts perceived as being less safe than average. This supports the argument that crime has an impact that goes beyond its direct costs.

2.3 Accuracy of crime rates

Any study that utilizes some measure of crime within an area should be done with prudence. Ensuring that your data is the best representation of the true landscape is important when making claims that have societal implications. A review of those studies dealing with the issues of accuracy and reporting trends within crime data becomes necessary in order to accomplish this task. When focusing on police-reported crime statistics, the primary issues with statistical accuracy lies with two main areas: (1) the victim's decision to notify police (Boivin et. al. 2011, Bosick et. al. 2012, Greenberg et. al. 2004, Hart and Rennison 2003, Loftin and McDowall 2010, Melde and Rennison 2010, Tolsma et al. 2012), and (2) police discretion in recording of events (Gottfredson and Gottfredson 1988, Matrofski et. al. 1995, Boivin and Cordeau 2011, Nolan et. al. 2011). When making claims that involve the analysis of crime, it is important to remain transparent with the data and the inherent limitations. Unfortunately, there is no true way to test or correct for these problems. Law enforcement agencies gather crime information via guidelines set by the FBI called the Uniform Crime Reporting (UCR) Program. These guidelines are consistent across all agencies where it then becomes the responsibility of the agency itself to provide accurate data. In many cases, UCR is the only data available (Loftin and McDowall 2010). Nolan et al. (2011) offers an approach to estimating classification error in the UCR, but this only deals with police discretion and not the victim's decision in reporting of events.

Logic tells us that violent crimes (particularly homicide) are likely to be the most consistently reported criminal events. The problem with homicide is that it is such a rare occurrence that building a statistically significant sample size can be difficult. According the National Crime Victim Survey (NCVS), in 2015 only 47% of other (non-homicide) violent crime victimizations were reported to police. Of these, the most frequently reported were robbery and aggravated assault at 61.9%. As a comparison, property crime victimizations were reported at a measly 35% to police. This under-reporting of property crime could be explained by Greenberg & Beach (2004) conclusion that the reporting of property crime to police becomes more of a decision making process for victims. They explain that the decision to notify police after the victimization of property crime is influenced by the cognitively driven cost-benefit process of reporting (when the rewards outweigh the cost of calling the police), the affect driven process of reporting (an emotionally driven response) and the socially driven process of reporting (social influence and turning toward others for advise).

Out of fourteen hedonic pricing studies surveyed that utilized crime as a dependent variable, it was found that the representation of crime lacked consistency. A count or density of crime was used in four of the publications. De-min et al (2006) looked at counts of all crime, while Troy & Grove (2008) and Bishop & Murphy (2011) looked at counts of violent crime only. Lynch & Rasmussen (2001) looked at counts of criminal events but separated violent and property crime. There were five studies that used crime rates, usually per 1,000 inhabitants in a census tract or block group. Tita et al. (2006) and Vasquez Lavin (2011) looked at the crime rate for all crime reported. Ceccato & Wilhemson (2010) looked at crime rates for all crime then for each separate type of crime separately (rape, robbery, assault, theft, etc.). Boggess et al. (2013) looked at rates for all crime, violent crime and then each crime separately as Ceccato & Wilhemson (2010) did. Clark & Cosgrove (1990) was the only that utilized murder rate. The other five publications used data that was not tied to the reporting of crimes to police. Graves et al. (1988) used the FBI community crime rate. Buonanno et al. (2013) looked at crime and housing prices in Barcelona, Spain utilizing a crime victimization survey. Dubin & Goodman (1982) use principal components analysis in the treatment of crime characteristics, separating categories of crime.

The inconsistency in which crime is measured in these studies further shows how little is known about the reporting of crime and how accurately it represents the true criminal landscape. As discussed previously, it is generally accepted that violent crime is more consistently reported than property crime. However, limiting variables to a single type of crime may have other unintended consequences. Because there is no standard set forth in the literature, this study takes a multipath approach in order to minimize the limitations that a single measure may provide; this study uses measures of all criminal events reported to police as well as singling out violent criminal events only.

Of interesting development is the trend to take advantage of Megan's law, which requires law enforcement agencies to make information regarding sex offenders public. Pope (2008) and Caudill et al. (2015) use the sex offender registry to show the nearest registered sex offender and a density of sex offenders within a given zone to measure a homebuyers devaluation with a convicted criminal in the area. While this certainly can't be used as a measure of real-time crime, it can be used as a way to measure the "fear" of future crime. The registry is a log of past criminal events, sometimes going back decades in time. Further, the registry does not show where the crime happened; rather it merely shows where the offender currently resides. It is for these reasons that the registry likely acts as a better measure of fear of crime rather than crime itself.

2.4 The Use of Scale and Aggregation of Data in Hedonic Studies

It is evident from the hedonic pricing and crime literature that there is a very real relationship to be found here, but evidence is conflicting. A contributing factor to this conflicting evidence is the various use of scale within the literature. How crime is aggregated or disaggregated is shown to have a very real impact on the results that are returned. Dubin & Goodman (1982) define the "neighborhood" as those places within and outside of the city, which showed a significant relationship. This is predated by Goodman (1977) where it was concluded that the neighborhood is best defined at the census block level for "goodness of fit" criteria. Goodman determined that the block level is this best predictor of homogeneity in a neighborhood. As you go further beyond the census block irregularities in the landscape are introduced. Tita et al. (2006) are one of the few to mention the aggregation impacts of crime data within their study, concluding that the impact of crime rates are misleading based on how crime is aggregated. A statistician can aggregate crime into makeshift neighborhoods or zones in order to influence conclusions that can be drawn from the data. This gerrymandering of data has real impacts politically and socially.

Clear throughout the literature, the Modifiable Aerial Unit Problem (MAUP), or the aggregation of data and its effects on statistical variation, is unpredictable and conclusions must hold transparency in their declarations. It is at least in need of careful consideration of the deeper meaning behind scale based statistical change. (Dungan et al. 2002, Fothering and Wong 1991, Hipp 2007, Kotavaara et al. 2012)

Outside of the Hedonic pricing and crime literature, scale effects are studied more frequently. Going back to hedonic models dealing with pollution, environmental quality and greenspace, we see a stronger grasp on the potential effects of not only the criteria variable to be studied, but also its effects across scales. These spatial effects are more commonly taken into account when looking at proximity measures - that is the vicinity of the house to the criteria variable (Bark et al. 2011, De Vor et al. 2009, Price et al. 2010). Payton et al. (2008) looked at average level of greenness around a property from 2 acre to 11 acre zones. This kind of analysis requires a set of continuous data, but can be very effective in the portrayal of scale effects by standardizing how scale is used throughout the study area. Each property can be assigned a value for each scale and later examined through statistical analysis.

In the case of this study, crime is not published as continuous data. However, crime count data can be manipulated to create a set of variables that can be approached in a similar manor to Payton et al. (2008) by assessing the count of reported crime within set distances from a property. Instead of using the 2 and 11 acre zones utilized by Payton et al. (2008), this study uses the same concept by creating zones within a set distance of a sold home (500, 1000 and 2000 yards). This way, each property can be assigned a value (number of criminal events) within the scale.

3.0 Methods

This research seeks to better understand how the effect of crime on housing prices varies by proximity. A hedonic price model was built to test this relationship of crime and housing price in Houston, TX, a city made unique by the absence of strict zoning laws which creates a heterogeneous landscape that is different than would be found in other cities with these zoning laws.

The following questions will guide the proposed research with the aim of contributing to the body of knowledge regarding the proximity at which crime has a measurable impact on housing prices. The primary question is answered through the examination of the following sub-questions:

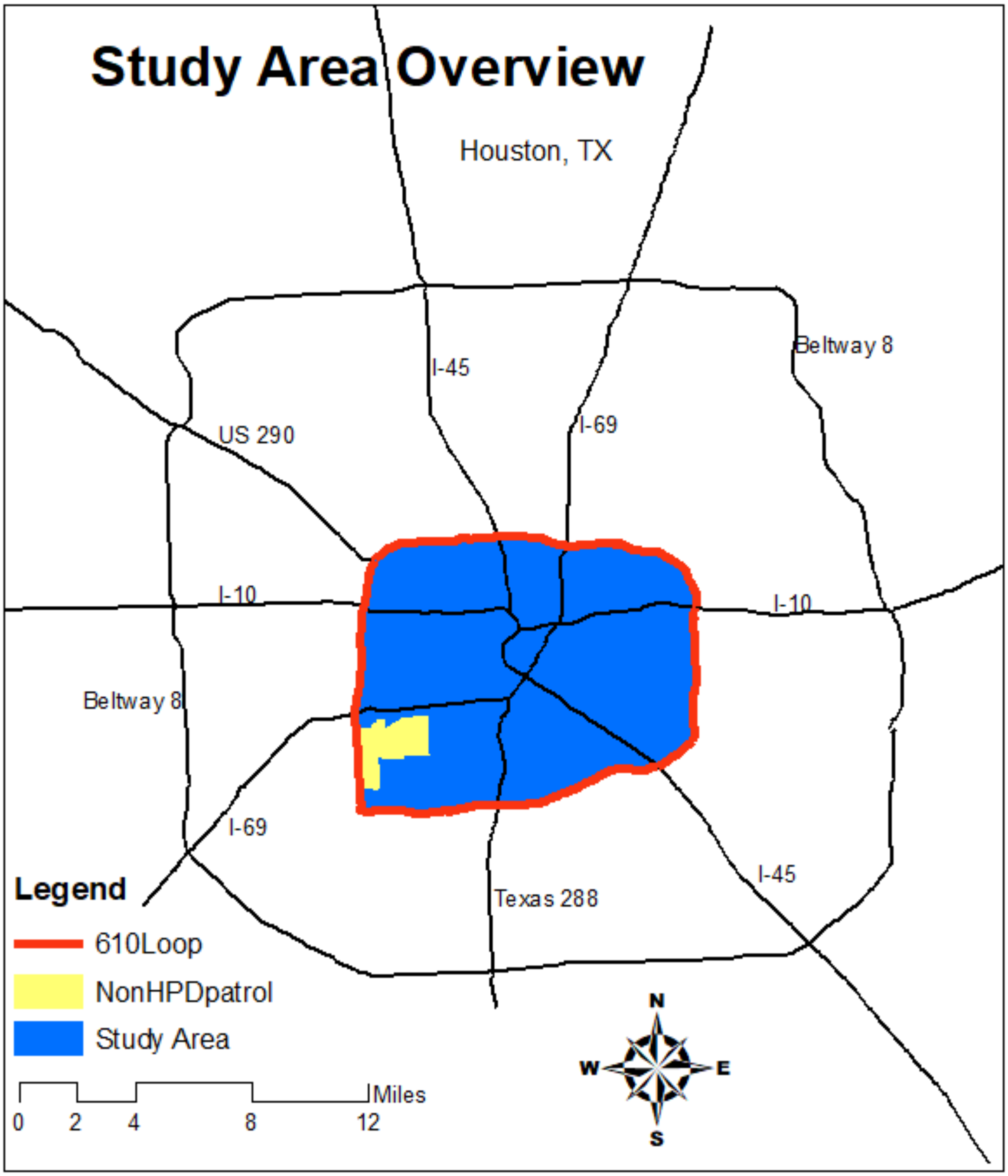
How does the effect of crime on housing prices vary by proximity?

- Is there a significant correlation between crime and housing price?
  - H_0: There is no significant correlation between crime and housing price.
- Does the proximity of crime affect this relationship?
  - H_0: Distance of crime to house has no effect on price.

3.1 Study Area

This project will be focused on the inner 610 loop of Houston, TX (Figure 1).

Figure 1: compiled with data from 2010 Census 'pcad.hdata.org'. Blue depicts study area.



The majority of the land area within the 610 loop is patrolled by the Houston Police Department. Only a small portion located in the southwest area of the loop is patrolled by the West University Place and Bellaire police departments.

The city of Houston is unique for the hedonic study in that the absence of zoning laws creates a more heterogeneous landscape than what would normally be found in cities with stricter zoning regulations. A heterogeneous landscape is important for a hedonic price model to function properly (Taylor 2003). This area of Houston is well established, highly urbanized, and diverse in its business and housing options. The population is spread evenly through the study area as depicted by Figure 2. Median household income as shown through 2010 Census Block Group is depicted in Figure 3.

Figure 2: compiled with data from 2010 Census by census block group. Dark blue depicts denser populations.

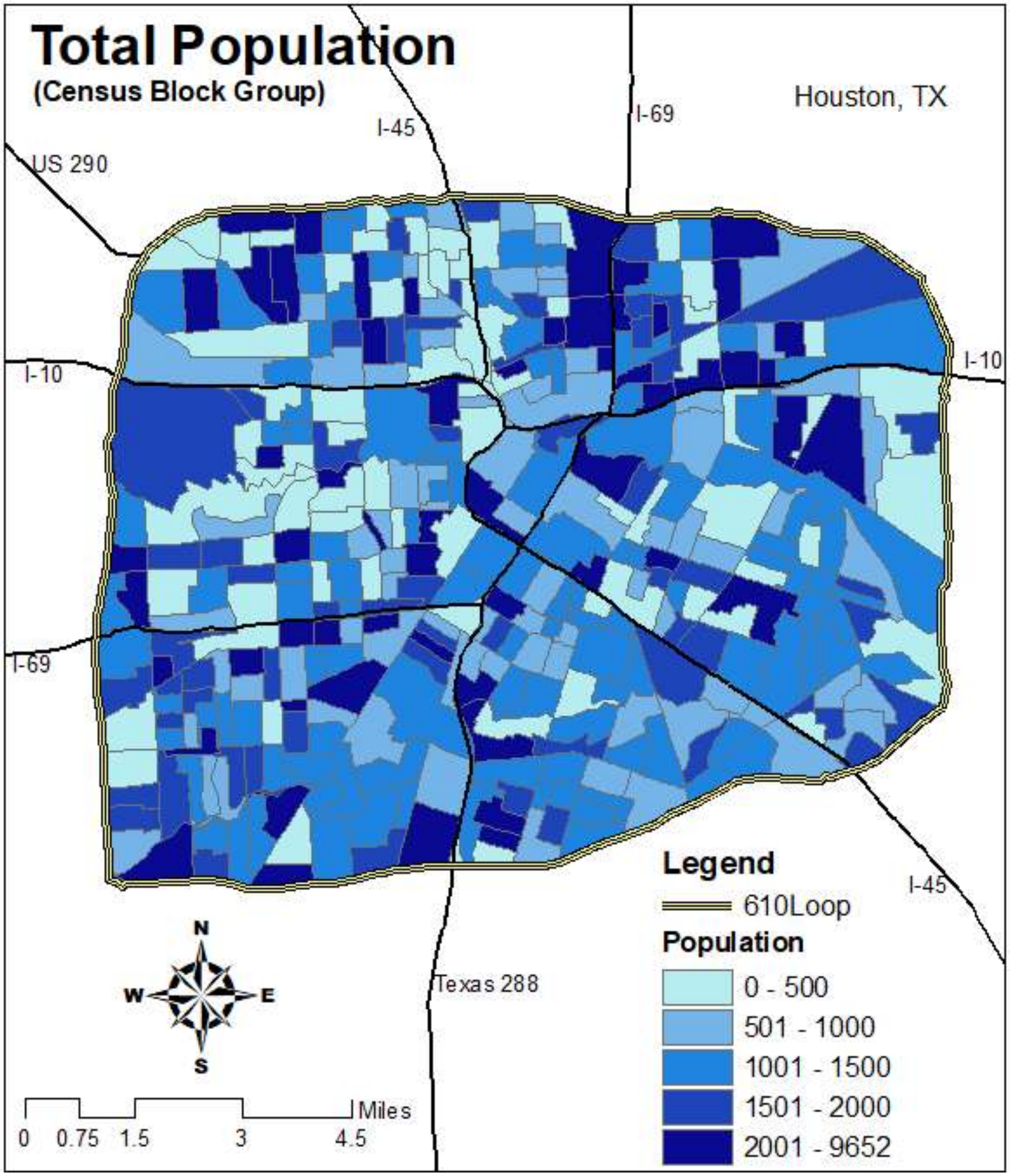
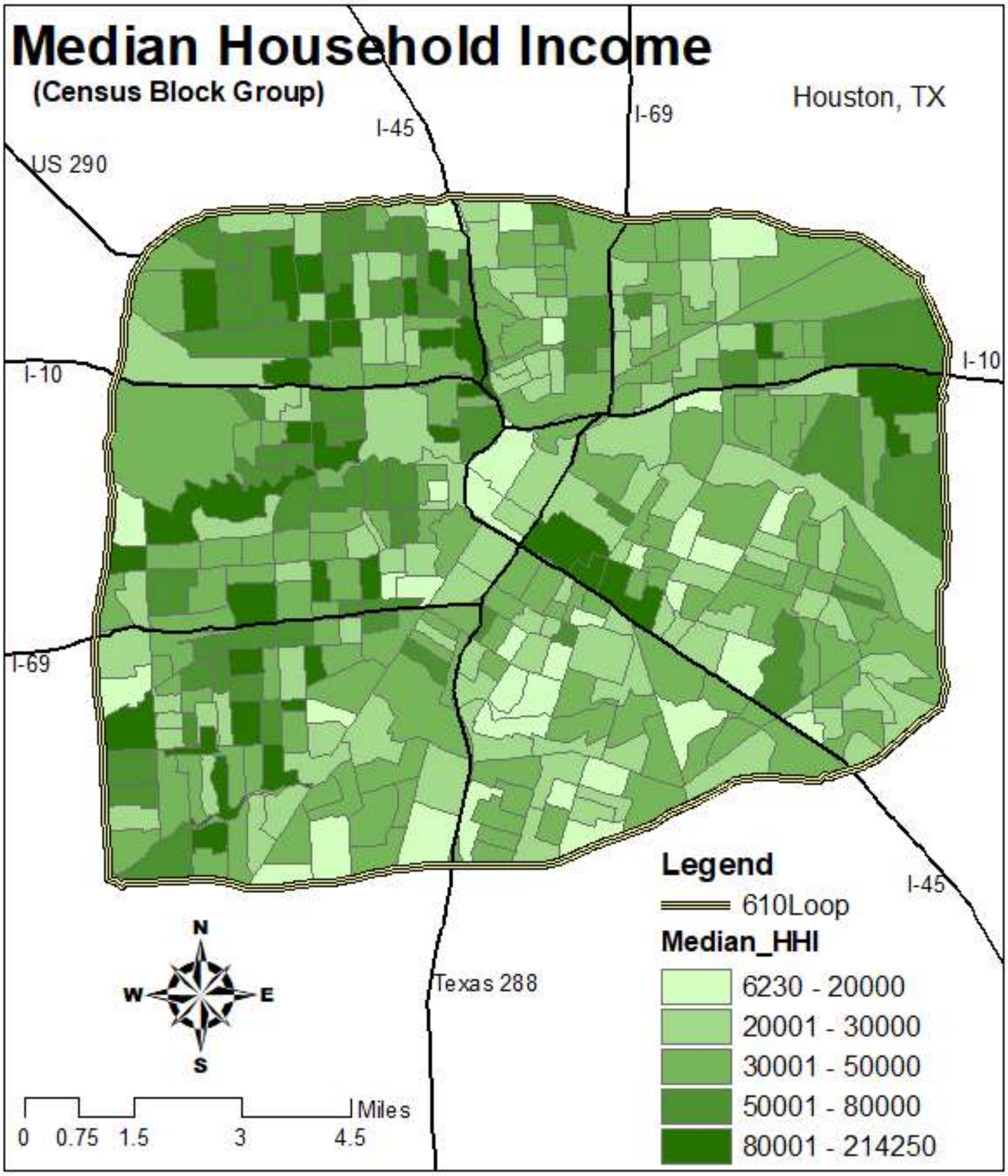


*Figure 3: compiled with data from 2010 Census by census block group. Dark green represents higher Median Household Income *(Median_HHI)*



3.2 Data

This study uses data from multiple sources. All variables are listed in Table 1. Variables were selected based on trends within the hedonic price modeling of real estate literature, including a measure of education background, demographics, and structural variables such as number of bedrooms, bathrooms and lot size (Kim & Wells 2005, Lynch & Rasmussen 2001, Payton et al. 2008). Selected variables were based on these trends as well as what is readily and freely available to the public within the study area. Some deviations include; 1) census data on education levels were used in lieu of school ratings – the study area is served by the Houston ISD which receives relatively poor ratings across the majority of schools with little variation, 2) the existence of A/C was omitted because of the prevalence in the Houston area due to climate, 3) floodplain was chosen over other commonly used measures of environmental quality such as air quality – this was substituted due to the frequent impact that flooding has in the area.

Housing data were obtained for the year 2010 to coincide with available crime data published by the Houston Police Department. Housing data comes from the Multiple Listing Service (MLS) database maintained by the Houston Association of Realtors.

Table 1: variable descriptions

Variable	**Units & Notes**	**Source & Year**
LogSale	*Log Sale price ($)*	MLS *(2010)*
SqFtBldg	Total sq. footage of home (hundreds)	MLS (2010)

LotSize	Total sq. footage of lot (hundreds)	MLS (2010)
GarageCap	Total number of garage spaces	MLS (2010)
Beds	Total number of bedrooms	MLS (2010)
BathsFull	Total number of full bathrooms	MLS (2010)
Median_HHI	Median annual household income by census block group	Census (2010)
Pwhite	% of white non-Hispanic population by census block group	Census (2010)
Phispanic	% of Hispanic population by census block group	Census (2010)
Pblack	% of black non-Hispanic population by census block	Census (2010)
PgradHS	% of population where highest level of education is that of a high school graduate	Census (2010)
p_bs	% of population where highest level of education is that of a college undergraduate degree	Census (2010)
P_Vacant	% of homes within census block group that are vacant	Census (2010)
CBDFeet	Distance from sold home to Central Business District (feet)	City of Houston GIS (cohgis.com)
ParkFeet	Distance from sold home to nearest park/greenspace (feet)	City of Houston GIS (cohgis.com)
F610Feet	Distance from sold home to the 610 loop (feet)	City of Houston GIS (cohgis.com)
HWYFeet	Distance from sold home major Houston highway (feet). Major highways include I-10, I-45, US59/I-69, Texas288 and US290.	City of Houston GIS (cohgis.com)
AirFeet	Distance from sold home to the nearest airport (feet)	City of Houston GIS (cohgis.com)
HospFeet	Distance from sold home to the nearest hospital (feet)	City of Houston GIS (cohgis.com)
FireFeet	Distance from sold home to the nearest City of Houston Fire Station	City of Houston GIS (cohgis.com)

	(feet)	
LibrFeet	Distance from sold home to the nearest public library (feet)	City of Houston GIS (cohgis.com)
PoliceFeet	Distance from sold home to the nearest City of Houston Police Station (feet)	City of Houston GIS (cohgis.com)
F100year	Dummy variable '1' if sold home is within 100 year floodplain as defined by FEMA. '0' if outside of 100 year floodplain	City of Houston GIS (cohgis.com)
F500 year	Dummy variable '1' if sold home is within 100 or 500 year floodplain as defined by FEMA. '0' if outside of floodplain	City of Houston GIS (cohgis.com)
Crime within 500 yards	All crime published at city block level within 500 yards of sold home.	HPD Crime Statistics (2010)
Crime within 1000 yards	All crime published at city block level within 1000 yards of sold home.	HPD Crime Statistics (2010)
Crime within 2000 yards	All crime published at city block level within 2000 yards of sold home.	HPD Crime Statistics (2010)
Crime between 500 and 1000 yards	All crime published at city block level between 500 and 1000 yards of sold home.	HPD Crime Statistics (2010)
Crime between 1000 and 2000 yards	All crime published at city block level between 1000 and 2000 yards of sold home.	HPD Crime Statistics (2010)

Crime data, which has been collected and published at the block level by the Houston Police Department (HPD), was categorized as a count of both all crime and violent crime (i.e., murder, rape, robbery and aggravated assault) within 500, 500 to 1000, 1000, 1000 to 2000 and 2000 yards of each sold home. The blocks are published as the locations of crime that fall on the same street with street numbers falling within a range of 100 as defined by HPD (100 to 199, 200 to 299, etc.). Assumptions were made that mid block would fall at the 50th increment within the range (150 for the 100 to 199 block). This does mean that criminal events are not geolocated to the precise location where the crime occurred. The ultimate size of the block is dependent on the distance from one intersecting street to the next. The HPD defined one block range as an increment of 100 houses (0 to 99, 100 to 199, etc) however this is not true throughout the entirety of the city. This does mean that in some cases the location of the criminal event will be close but not exactly where the event would occur (this would be the case regardless, because HPD does not publish the exact house number in order to protect victim privacy).

Crime Data for the year 2010 were obtained from the HPD website and were geolocated then aggregated into a table that includes MLS and census data. Census data for 2010 was obtained at the block level as recommended by Goodman (1977). Each single family detached home was geocoded and merged by location with location characteristics, neighborhood characteristics and reported crime data with the use of a geographic information system (GIS).

The resulting attribute table was loaded into a statistical package (SPSS) where the models were executed. A detailed summary of how the GIS database was built is below organized by data source:

Census:

Census data was downloaded from www.census.gov. The data set included an attribute table and a polygon shapefile, which were joined together to show boundaries in Harris County by census block group.

City of Houston GIS (COHGIS):

From COHGIS (www.cohgis.com) a road shapefile in line format for Harris County was downloaded and then the 610 loops and other major roadways were selected and a new shapefile created based on this selection. A polygon shapefile was then downloaded which had every property (parcel) for Harris County. The 610 loop shapefile was used to clip the parcels and a new shapefile was created for parcels only within the loop. Next, shapefiles for the remaining price model variables were downloaded and opened into the GIS. The near tool was used to calculate distance between the sold home and features from COHGIS (CBD, parks, highways, etc.).

Multiple Listing Service (MLS)

The MLS in Houston is managed by the Houston Association of REALTORS® (HAR). A Houston area Realtor allowed access to this database to download data on all houses sold in the year 2010 along with all of their attribute information. This information was downloaded by “market area” (zones 2, 4, 9, 16 and 17) which came in .txt format. These tables were appended together in Microsoft Excel and then a final MLS table was opened into the GIS. From here data for each sold property was joined to the existing parcel shapefile from COHGIS. Parcels that had MLS sales data for 2010 were selected and a new shapefile was created based on this selection.

Houston Police Department (HPD) Crime Statistics:

The HPD publishes statistics from the Uniform Crime Report (UCR) at www.houstontx.gov. This information was downloaded for each month of 2010 which came in a .txt format. These tables were appended together in Microsoft Excel and then opened into the GIS. The location attributes of this data was by city block (e.g. '100-199 Main St' or '400-499' Broadway). As previously discussed, assumption were made that a criminal event was located at the 50^{th} increment of a 100 house block range (0 to 99, 100 to 199, etc). The actual city block ranges vary between a 25 and 100 range, so the geolocator was adjusted to ensure the placed location of the event falls somewhere on the correct block if there are less than 100 houses. Actual criminal event location could vary throughout the block. From here each criminal event was geolocated based on address using the streets shapefile (creating a point shapefile). Spatial join was used between the located criminal events and the shapefile with MLS information to find counts of criminal events within the specified distance ranges (500, 1000 and 2000 yards).

3.3 Models

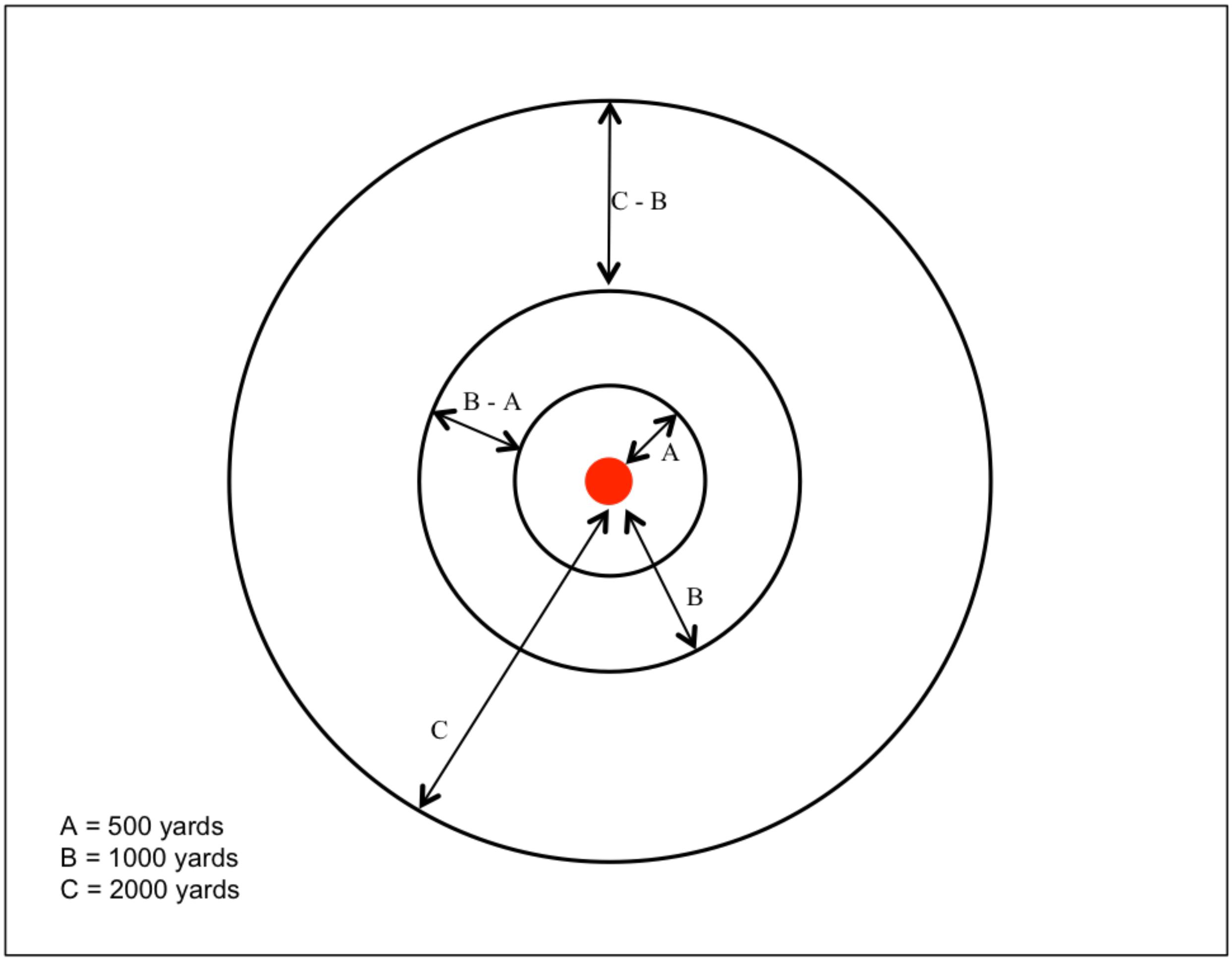
Ten models were built to measure the effect of crime on housing prices (Table 2). Each model measures the effect of crime on housing prices by extracting a count of all criminal events or violent criminal event within 500, 1000 and 2000 yards and from 500 to 1000 and 1000 to 2000 yards.

Table 2: Model descriptions

Model #	**Dependent Variables**	**Explanatory Variables**	**Crime**
1	Log Sale Price ($)	All housing & neighborhood characteristics	All crime within 500 yards
2	Log Sale Price ($)	All housing & neighborhood characteristics	All crime between 500 and 1000 yards
3	Log Sale Price ($)	All housing & neighborhood characteristics	All crime within 1000 yards
4	Log Sale Price ($)	All housing & neighborhood characteristics	All crime between 1000 and 2000 yards
5	Log Sale Price ($)	All housing & neighborhood characteristics	All crime within 2000 yards
6	Log Sale Price ($)	All housing & neighborhood characteristics	Violent crime within 500 yards
7	Log Sale Price ($)	All housing & neighborhood characteristics	Violent crime between 500 and 1000 yards
8	Log Sale Price ($)	All housing & neighborhood characteristics	Violent crime within 1000 yards
9	Log Sale Price ($)	All housing & neighborhood characteristics	Violent crime between 1000 and 2000 yards
10	Log Sale Price ($)	All housing & neighborhood characteristics	Violent crime within 2000 yards

Figure 4 below depicts a visual representation of the models.

Figure 4: The red dot represents a sold home. 'A' depicts models 1 & 6. 'B' depicts models 3 & 8. 'C' depicts models 5 & 10. 'B-A' depicts models 2 & 7. 'C-B' depicts models 4 & 9.



Neither Model 3 nor Model 5 separates the smaller scale in context of the larger (500 yards to 1000 yards and 1000 yards to 2000 yards). Model 3 shows us the effect of everything within 1000 yards without controlling for what Model 1 already showed us within 500 yards. Model 2 looks to separate the effect of the 500 and 1000 yard scales by including both measures. Combining the 500 yard count with the difference of the 1000 yard count around the 500 yard count allows for understanding the relationship of the surrounding criminal landscape on each property, controlling for the immediate criminal landscape around the property. The same can be said for combining the 1000 yards count with the difference of the 2000 yard count around the 1000 yard count (Model 4).

The hedonic model used is the log-linear expression Ordinary Least Squares method (OLS). The OLS model is the most commonly utilized in the literature (Payton et al. 2008). As discussed by Dubin & Goodman 2013, simple OLS can be overly restrictive while log-linear OLS can offer more flexibility to the model.

Results of the OLS were analyzed based on the expected versus realized signs and coefficients of dependent variables, as well as the coefficient of determination for the model. Dollar value summaries of the discount/premium associated with crime in the area is of primary interest, as measured by the coefficient (negative for a discounted price association) of the count of crime within the area. A negative coefficient on count of crime therefore indicates that for every 1 unit of increase in crime (reported incident) there is a set discount (the coefficient of the crime variable in dollars) in the price of the house.

3.3.1 Variable Selection

All variables were subjected to a Pearson Correlation test to check for collinearity and to show evidence that a relationship to house price exists. Table 3 below shows a Pearson Correlation matrix on a selection of variables. This selection was taken from the full list of starting variables for viewing ease. A manual forward selection process was used where all variables were sorted from highest to lowest correlation to 'LogSale' and then each variable was accepted or rejected based on correlations seen across other variables. A manual selection process was chosen in order to account for the logical associations between the data that algorithms aren't designed to see. Square footage of a house has a clear logical relationship to the number of bedrooms, but an algorithm can only see the correlation and not the reason behind this relationship (a higher number of bedrooms in most cases will mean a larger footprint for a home).

Final variable selections were based on an effort to decrease collinearity between the explanatory variables while maintaining a balance between structural and neighborhood characteristics. All variables evaluated based upon their logical importance to the equation along with the possible biases imposed between the variables. The final selection of explanatory variables are explained below:

LotSize:

'LotSize' had relatively low correlations across all variables. The highest coefficient to '*LotSize'* was '*SqFtBldg'* (0.302), which was not selected because of its consistently high correlations across most variables. The highest coefficient to a crime variable was '*Crime500'* at -0.146.

GarageCap:

'GarageCap' had high correlation to '*LogSale*' (0.505) with lower correlations across other variables that were ultimately selected. As with most structural variables, '*GarageCap*' was highly correlated with '*SqFtBldg*' (0.422), which was not selected. Logic tells us that houses with high garage capacity would lend itself to a house where the homeowner has possessions (whether it be vehicles or some other good) to fit in the extra space that a garage provides. The highest coefficient to a crime variable was '*VCrime200*' at -0.263.

Beds:

'Beds' exhibit a relatively high correlation to '*LogSale*' (0.452) while minimizing correlation with other variables. As discussed, '*SqFtBldg*' was not used because of consistently high correlations with variables. 'Beds' has the highest correlations with '*SqFtBldg*' at 0.696, but keeps correlation with other explanatory variables relatively low (the next highest is the correlation with 'GarageCap' at 0.289). Logic is consistent with what the data says – number of bedrooms should give a reasonable understanding of house size. Number of bedrooms should give a good representation of house size, without the problems introducing 'SqFtBldg' into the dataset would bring.

PgradHS

'PgradHS' represents the neighborhood characteristic that had the lowest levels of correlation across other structural characteristics. Most studies use some variable to

measure quality of education along with other variables to show socioeconomic characteristics. The data gathered for this study, however, shows that correlation is high throughout the socioeconomic variables ('PgradHS' to 'Median_HHI' is -0.560). Correlation is also high for socioeconomic variables with the crime data ('PgradHS' to 'Crime10_20' is -0.303. In an effort to reduce conflict between variables, '*PgradHS*' was the only socioeconomic variable chosen. 'PgradHS' was selected over 'Median_HHI' because correlations between the two are high (-0.560) while 'PgradHS' has higher correlation to 'LogSale' than does 'Median_HHI' (-0.651 vs 0.593). Including both variables would lead to a higher degree of collinearity.

HWYFeet, HospFeet, FireFeet

The majority of studies use some kind of proximity variable to add spatial context to the data. Distance to CBD and distance to greenspace or parks are common in other studies, but correlations with these were either insignificant to 'LogSale' or high between other explanatory variables in this dataset ('ParkFeet' had a not significant correlation to 'LogSale' at 0.008 while 'CBDFeet' was highly correlation to 'F500year' at 0.438). '*HWYFeet', 'HospFeet'* and *'FireFeet'* give the spatial context that other studies show by being a measure of distance, but keep correlations between explanatory variables low. The highest correlation was '*HWYFeet*' with '*PgradHS*' at -0.248.

F500year

In addition to the widespread use of proximity variables in the literature, a large portion of studies (Haab & McConnell 2002) also use some form of a negative environmental factor. Common is the use of urban air quality (Graves et al. 1988, Haab & McConnell 2002). Because of the issues that the city of Houston frequently has with flooding, the FEMA flood zone was included. A dummy variable was created to show whether or not a home fell within the 500 year floodplain as designated by FEMA.

The highest correlation between selected explanatory variables without the crime variables was between '*GarageCap*' and '*Beds*' with a coefficient of 0.289. The highest correlation between a crime variable and other explanatory variable was -0.285, which was between'*Crime2000*' and '*PgradHS*'.

Table 3: Pearson Correlation matrix

Variable	LogSale	SqFtBldg	LotSize	GarageCap	Beds	PgradHS	CBDFeet	ParkFeet	HWYFeet	HospFeet	FireFeet	F500year	VCrime1000
VCrime1000	-.336**	-.196**	-.107**	-.262**	-.147**	.258**	-.481**	-.001	-.207**	-.173**	-.150**	-.250**	1
F500year	.068**	.038	.123**	.103**	.146**	-.135**	.438**	.002	.203**	-.065**	.067**	1	-.250**
FireFeet	.177**	.109**	.029	.077**	.050*	-.083**	.023	-.029	.140**	-.163**	1	.067**	-.150**
HospFeet	-.098**	-.053**	.036	-.041	-.050*	.034	-.035	-.141**	-.055**	1	-.163**	-.065**	-.173**
HWYFeet	.251**	.208**	.125**	.131**	.143**	-.248**	.107**	-.161**	1	-.055**	.140**	.203**	-.207**
ParkFeet	.008	.058**	-.005	.009	.077**	-.019	.208**	1	-.161**	-.141**	-.029	.002	-.001
CBDFeet	.116**	.146**	.106**	.184**	.191**	-.065**	1	.208**	.107**	-.035	.023	.438**	-.481**
PgradHS	-.651**	-.414**	-.086**	-.286**	-.242**	1	-.065**	-.019	-.248**	.034	-.083**	-.135**	.258**
Beds	.452**	.696**	.191**	.289**	1	-.242**	.191**	.077**	.143**	-.050*	.050*	.146**	-.147**
GarageCap	.505**	.422**	.047**	1	.289**	-.286**	.184**	.009	.131**	-.041	.077**	.103**	-.262**
LotSize	.168**	.302**	1	.047**	.191**	-.086**	.106**	-.005	.125**	.036	.029	.123**	-.107**
SqFtBldg	.714**	1	.302**	.422**	.696**	-.414**	.146*	.058**	.208**	-.053**	.109**	.038	-.196**
LogSale	1	.714**	.168**	.505**	.452**	-.651**	.116**	.008	.251**	-.098**	.177**	.068**	-.336**

Note: **significant at 1%, *significant at 5%

Table 4 below shows descriptive statistics for the final selected variables

Table 4: descriptive statistics for all variables with additions from Table 5b in results section

Variable	Minimum	Maximum	Mean	Std. Deviation
'LogSale'	9.306	15.684	12.483	1.008
'LotSize'	1328	150935	6448.95	6219.802
'GarageCap'	0	8	1.46	.891
'Beds'	1	10	2.97	.787
'PgradHS'	.000	.633	.155	.108
'HWYFeet'	22.558	11015.041	3389.285	2120.625
'HospFeet'	423.623	16457.457	6678.242	3493.219
'FireFeet'	108.105	8269.035	4310.411	1706.946
'F500year'	0	1	.200	.399
'CBDFeet'	.000	35579.059	15204.42	7071.736
'SqFtBldg'	360	9742	2230.64	1221.856
'Crime500'	0	200	31.01	24.644
'Crime1000'	4	667	128.03	88.367
'Crime2000'	29	1603	479.20	282.354
'VCrime500'	0	36	4.08	4.717
'VCrime1000'	0	85	16.46	13.754
'VCrime2000'	5	242	62.07	42.178

4. Results & Discussion

Table 5a shows the OLS results for each model that was introduced in Table 2 and equation 3. Each model explains more than 61% of the variance in housing prices (the models for Tita et al. 2006 explained roughly 43% to 44% of variance in housing prices while the models Dubin & Goodman 1992 explained 64% to 68% of variance in housing prices). The estimated coefficients for violent crime variables have the expected sign and are significant at $p < 0.01$. The estimated coefficients for all crime, however, have an unexpected positive coefficient and are not significant at all scales. In model 1 (all crime within 500 yards) the estimated coefficient for crime is not significant. Model 3 (all crime within 1000 yards) returns an estimated coefficient (.036) that is significant at $p < 0.05$. All explanatory variables other than crime have the expected sign and are significant at $p < 0.01$.

Because the models are expressed as log-lin (the dependent variable *LogSale* is in log form while the independent variables are not), the coefficients can be multiplied by 100 to show the percentage change in price for every one unit increase in the independent variable. Models 1 through 5 use a count of all crime as explanatory variables. Models 6 through 10 use a count of violent crime as explanatory variables.

Models 1, 3 and 5 use crime aggregated within 500, 1000 and 2000 yards respectively. As discussed above, crime was not significant at within 500 yards. Model 3 shows that for every criminal event reported and logged within 1000 yards leads to a 3.6% increase in housing price. Model 5 shows that for every criminal event reported and logged within 2000 yards leads to a 5.5% increase in housing price.

Both Model 2 and Model 4 indicate that that the relationship between their respective two scales is important. Model 2 shows that for every criminal event reported and logged between 500 and 1000 yards leads to a 4.2% increase in housing price. Model 4 shows that for every criminal event reported and logged between 1000 and 2000 yards leads to a 5.8% increase in housing price.

Models 6 through 10 follow the same course as 1 through 5, but use counts of violent crime only rather than all criminal events. Results are reversed and amplified. In model 6 (all violent crime within 500 yards) the estimated coefficient for crime (-0.088) is significant at $p < 0.01$ where in model 1 the estimate coefficient for all criminal events (0.011) was not significant. All other coefficients at scales other than all crime at 500 yards were significant.

Models 6, 8 and 10 use all violent crime aggregated within 500, 1000 and 2000 yards respectively. Model 6 shows that for every violent criminal event reported and logged within 500 yards leads to an 8.8% decrease in housing price. Correspondingly, Model 8 returns a 12.5% decrease and Model 10 returns a 12.8% decrease.

As with Model 2 and Model 4, Model 7 and Model 9 separate the effect of the 500 and 1000 yard scales. Model 7 shows that for every violent criminal event reported and logged between 500 and 1000 yards leads to a 11.9% decrease in housing price. Model 9 shows that for every violent criminal event reported and logged between 1000 and 2000 yards leads to an 11.5% decrease in housing price.

Table 5a: Model results

Variable	Model 1 (Crime500)	Model 2 (Crime5_10)	Model 3 (Crime1000)	Model 4 (Crime10_20)	Model 5 (Crime2000)	Model 6 (VCrime500)	Model 7 (VCrime5_10)	Model 8 (VCrime1000)	Model 9 (VCrime10_20)	Model 10 (VCrime2000)
LotSize	.072**	.074**	.074**	.073**	.069**	.066**	.066**	.066**	.068**	.067**
	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000
GarageCap	.286**	.286**	.286**	.291**	.290**	.271**	.265**	.263**	.264**	.260**
	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000
Beds	.236**	.236**	.236**	.233**	.234**	.236**	.232**	.233**	.233**	.233**
	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000
PgradHS	-.494**	-.485**	-.487**	-.475**	-.476**	-.487**	-.479**	-.478**	-.490**	-.485**
	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000
HWYFeet	.053**	.056**	.056**	.055**	.056**	.041**	.042**	.039**	.043**	.040**
	.000	.000	.000	.000	.000	.003	.002	.004	.002	.003
HospFeet	-.048**	-.042**	-.044**	-.038**	-.039**	-.062**	-.076**	-.076**	-.067**	-.073**
	.000	.002	.001	.005	.004	.000	.000	.000	.000	.000
FireFeet	.092**	.092**	.092**	.093**	.093**	.081**	.077**	.075**	.085**	.081**
	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000
F500year	-.089**	-.081**	-.081**	-.075**	-.075**	-.106**	-.113**	-.115**	-.117**	-.120**
	.000	.000	.001	.000	.000	.000	.000	.000	.000	.000
Crime Variable	.011	.042**	.036*	.058**	.055**	-.088**	-.119**	-.125**	-.115**	-.128**
	.453	.003	.012	.000	.000	.000	.000	.000	.000	.000
R^2	.613	.614	.614	.615	.615	.619	.624	.625	.624	.626

Note: **significant at 1%, *significant at 5%

We can test the stability of the model by introducing additional explanatory variables and observing the change in results. Table 5b shows the OLS results for each model that was introduced in Table 2 and equation 3 with the inclusion of variables 'CBDFeet' and 'SqFtBldg'. 'CBDFeet' was chosen because its central location in the study area gives additional spatial context to home location. 'SqFtBldg' was chose due to its high correlation with 'LogSale' at 0.714. Both variables add to the overall collinearity of the model.

With these added variables each model explains more that 70% of the variance in housing price. The estimated coefficients for violent crime variables retain the expected sign and are significant at $p < 0.01$. The estimated coefficient for all crimes maintains an unexpected positive coefficient and are not significant at all scales (Models 1 through 3 coefficients are not statistically significant while Models 4 and 5 are significant at $P < 0.05$).

From Table 5b we continue to see that proximity of a violent criminal event to home matters. Counts of violent crime from 0 to 500 yards shows that each additional criminal event yields a 9.7% decrease in home price. At 500 to 1000 yards this becomes a 13.9% decrease in home price. From 0 to 1000 yards each additional violent criminal event yields a 12.5% decrease in home price while this increases to 17.4% at 1000 to 2000 yards.

The addition of these variables yields a better overall fit of the model (violent criminal events within 2000 yards goes from .626 to .722). These variables may increase the overall collinearity of the model, but results show a similar story; violent crime is capitalized in the housing market through lower home prices. A similar result between the models, even with the increased collinearity, affirms the reliability of the model.

Table 5b: model results with inclusion of 'CBDFeet' and 'SqFtBldg'

Variable	Model 1 (*Crime500*)	Model 2 (*Crime5_10*)	Model 3 (*Crime1000*)	Model 4 (*Crime10_20*)	Model 5 (*Crime2000*)	Model 6 (*VCrime500*)	Model 7 (*VCrime5_10*)	Model 8 (*VCrime1000*)	Model 9 (*VCrime10_20*)	Model 10 (*VCrime2000*)
LotSize	-.013	-.012	-.012	-.013	-.012	-.015	*.016	-.017	-.012	-.014
	.292	.333	.332	.306	.333	.217	.171	.166	.310	.249
GarageCap	.190**	.190**	.190**	.192**	.192**	.179**	.174**	.171**	.174**	.169**
	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000
Beds	-.039*	-.038*	-.038*	-.039*	.039*	-.036*	-.036*	-.035*	-.028	-.028
	.017	.019	.018	.016	.018	.026	.023	.028	.083	.077
PgradHS	-.396**	-.393**	-.393**	-.386**	-.387**	-.385**	-.373**	-.372**	-.382**	-.374**
	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000
HWYFeet	0.032**	.033**	.033**	.034**	.034**	0.021	.021	0.017	.020	.016
	.009	.007	.007	.006	.005	.091	.080	.147	.100	.183
HospFeet	-.042**	-.039**	-.040**	-.036**	-.036**	-.057**	-.074**	-.075**	-.070**	-.079**
	.000	.001	.001	.003	.002	.000	.000	.000	.000	.000
FireFeet	0.071**	.072**	.072**	.073**	.072**	.059**	.053**	.051**	.060**	.054**
	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000
F500year	-.052**	-.051**	-.051**	-.052**	-.051**	-.055**	-.053**	-.054**	-.043**	-.045**
	.000	.000	.000	.000	.000	.000	.000	.000	.001	.000
Crime Variable	0.005	.018	.016	.036*	.032*	-.097**	-.139**	-.125**	-.174**	-.193**
	.691	.181	.240	.026	.041	.000	.000	.000	.000	.000
CBDFeet	.011	.016	.016	.030	.027	-.023	-.051**	-.056**	-.109**	-.116**
	.419	.251	.268	.061	.084	.086	.000	.000	.000	.000
SqFtBldg	.486**	.484**	.485**	.480**	.481**	.488**	.486**	.487**	.481**	.482**
	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000
R^2	.704	.704	.704	.704	.704	.711	.716	.718	.718	.722

Note: **significant at 1%, *significant at 5%

Figures 5, 6 and 7 below can be used to visualize crime within the study area as well as value of sold homes:

Figure 5: all criminal events (2010)

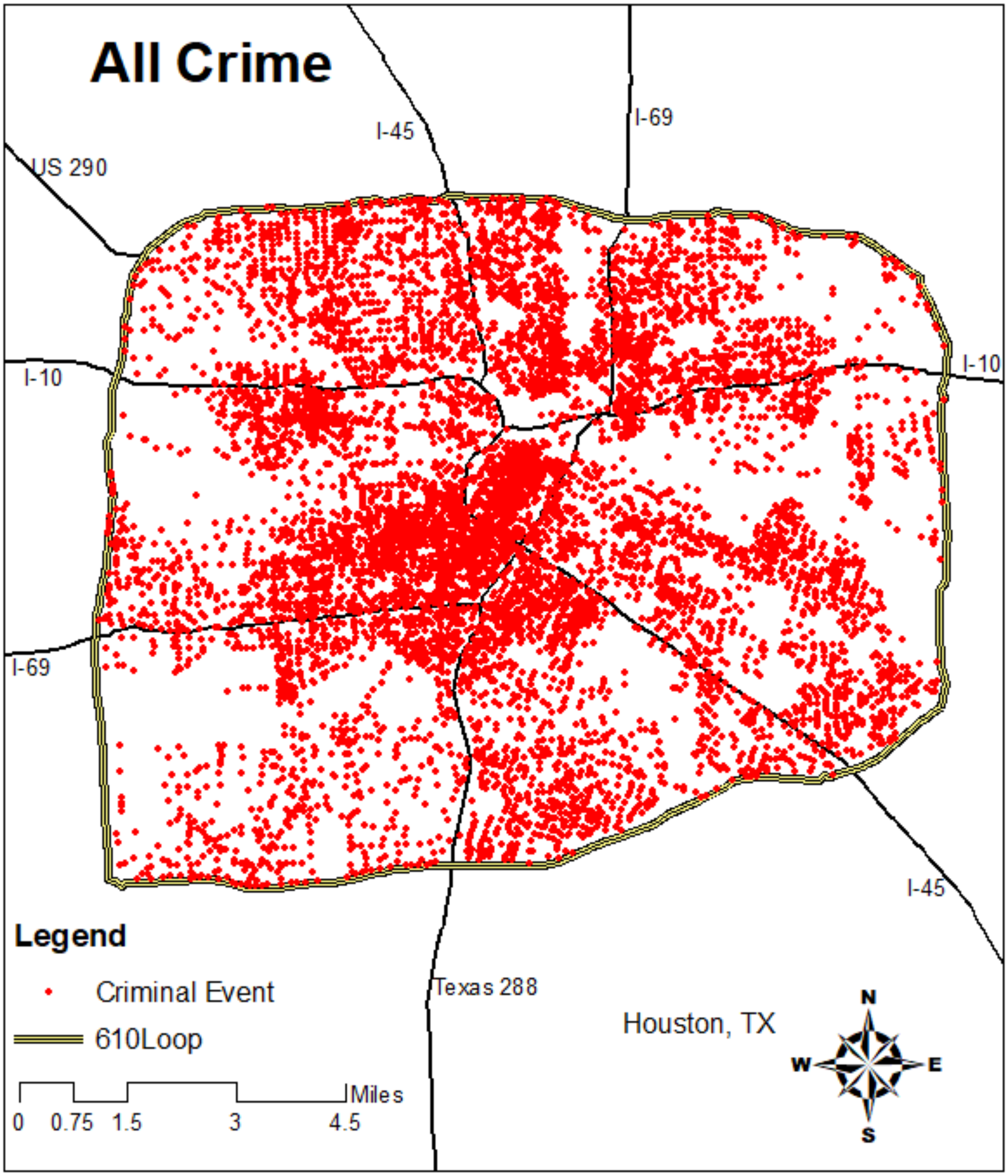


Figure 6: violent criminal events (2010)

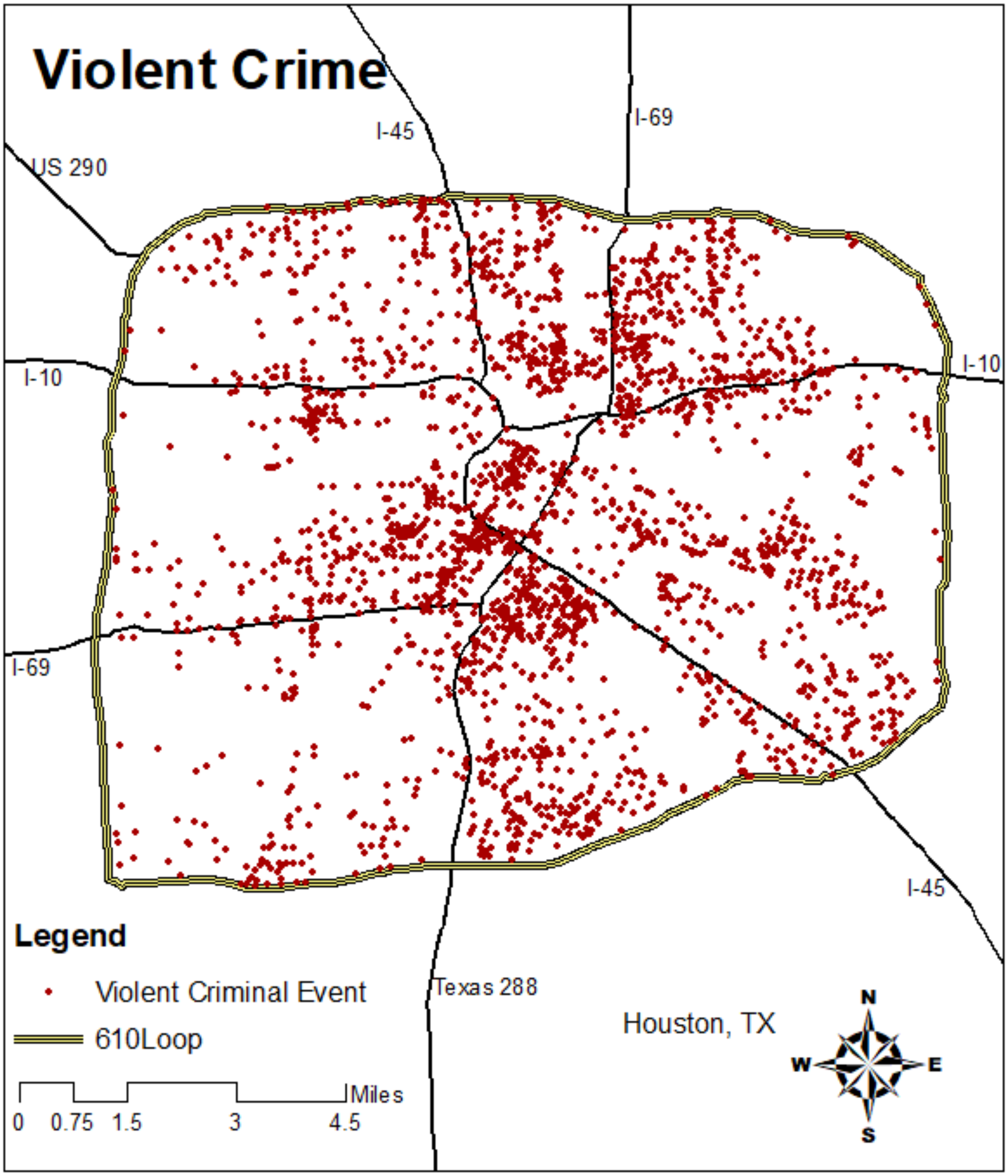
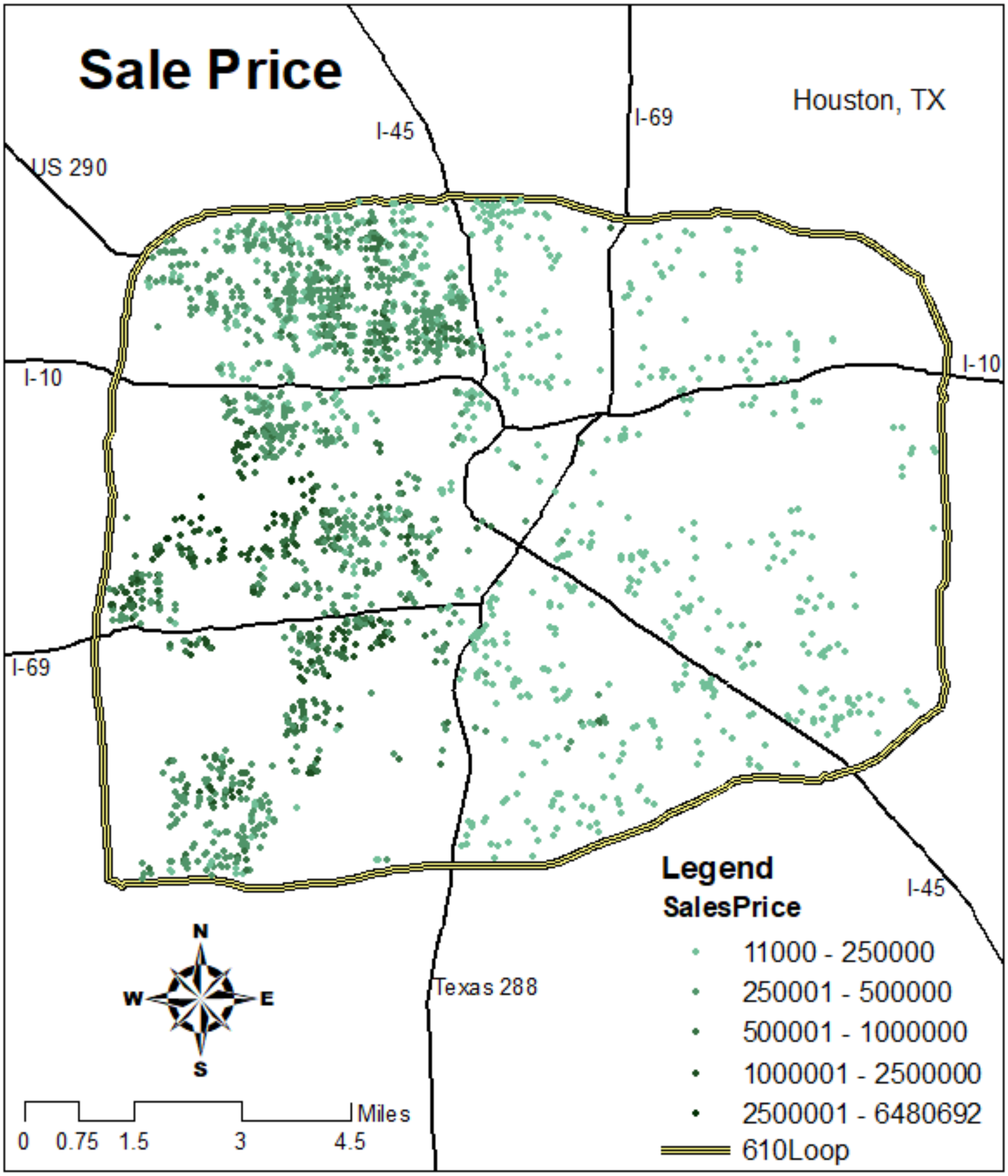


Figure 7: sold homes (2010)



Figures 8 and 9 below combine sales price and criminal events for better visualization:

Figure 8: All criminal events of 2010 (red). Sold homes (darker shades of green represent higher price).

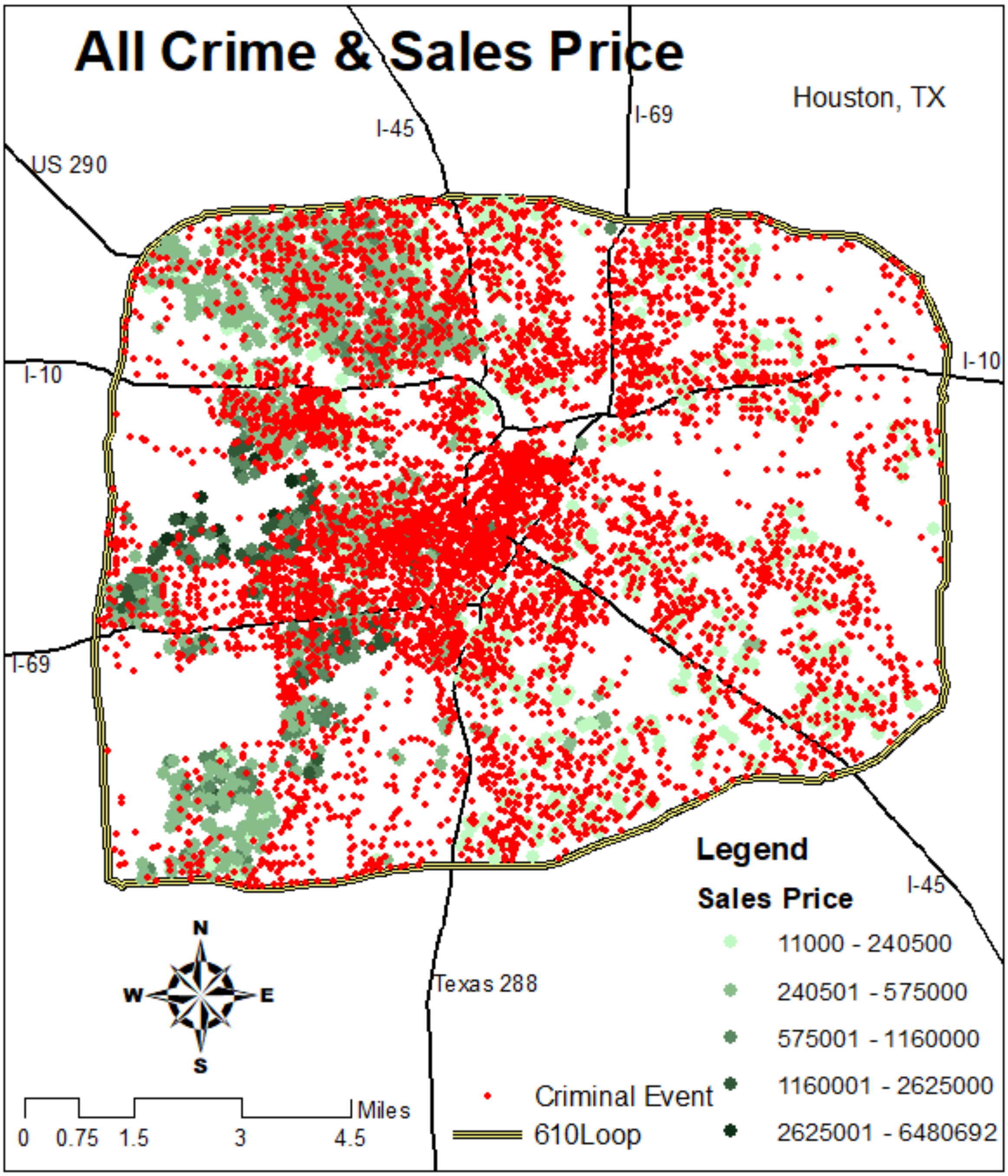
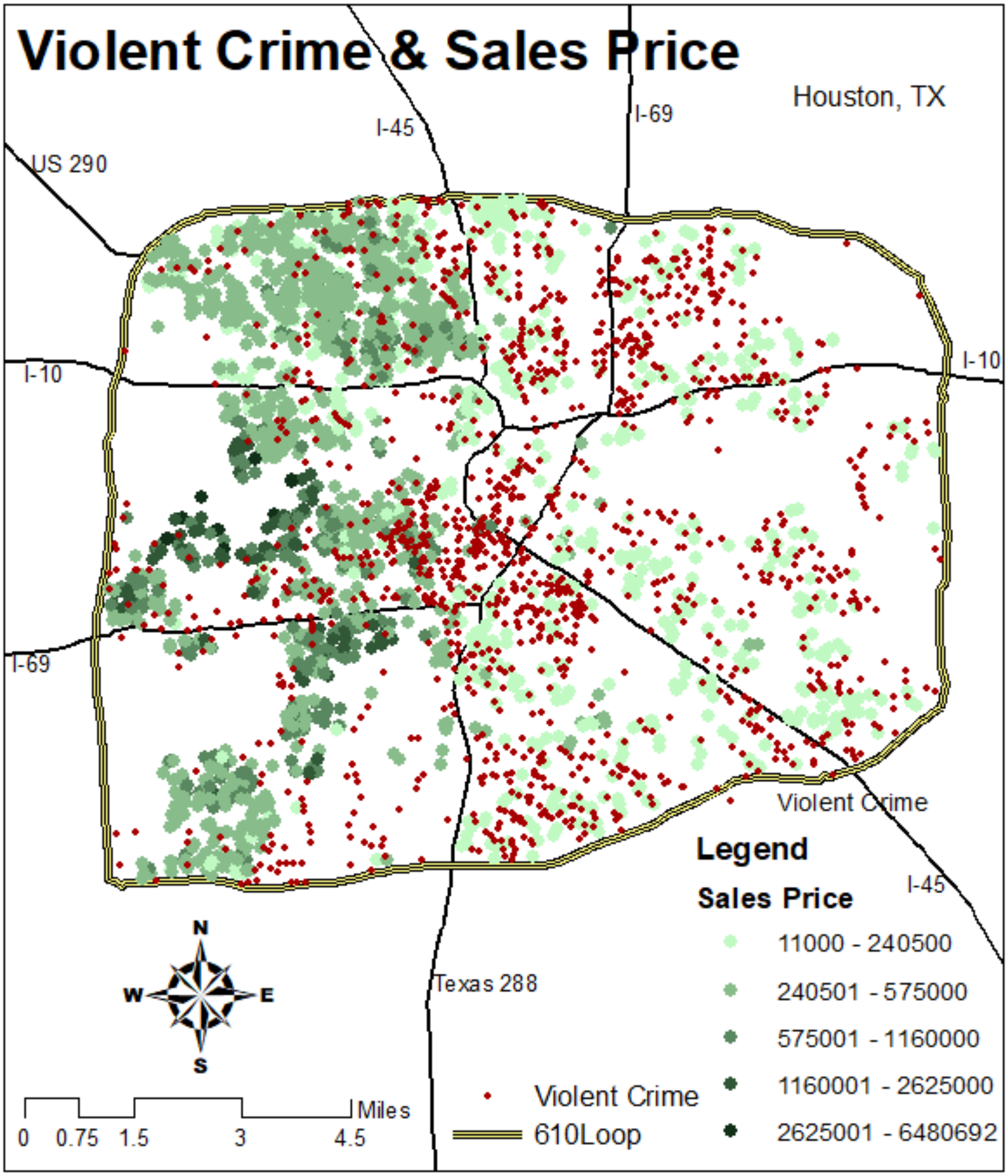


Figure 9: Violent criminal events of 2010 (red). Sold homes (darker shades of green represent higher price).



Results show that the estimation of the impact of crime on housing values across inner-loop Houston, TX is both intriguing and potentially misleading. There are several factors that must be considered before drawing conclusions.

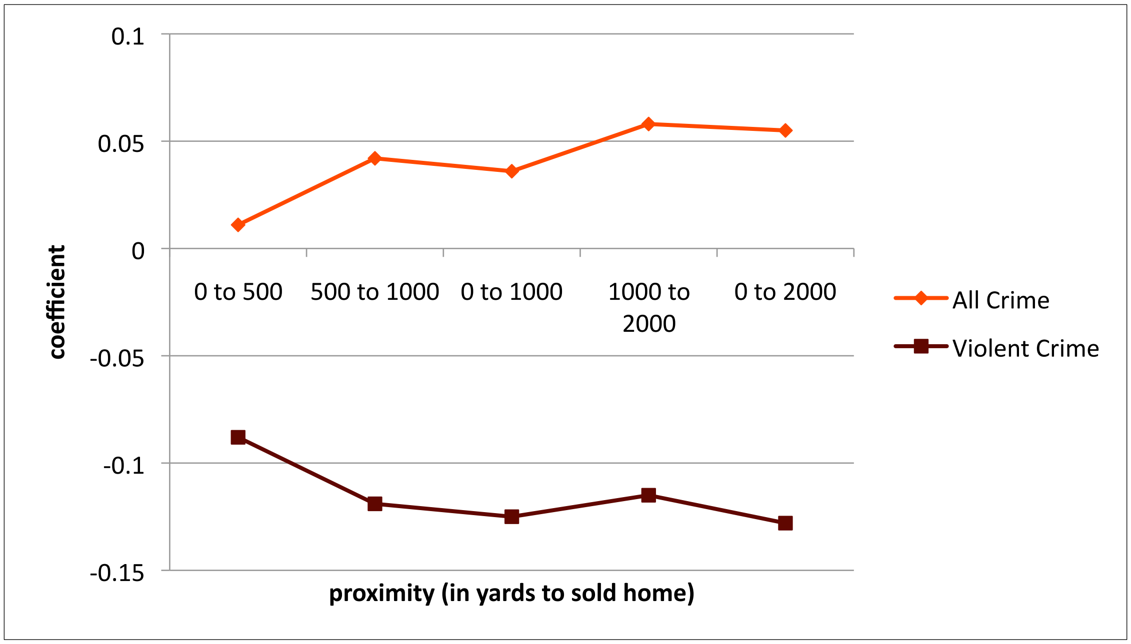
Based on the model results we can reject both of the null hypotheses. Data shows a statistically significant correlation between crime and housing price. In fact, returning to Table 5a, all but one of the models (all crime counts within 500 yards) were statistically significant at least at $p < 0.05$.

For all counts of crime within 500 yards, albeit an insignificant correlation, the coefficient is 0.011, which jumps to .042 (and becomes significant at 1%) from 500 to 1000 yards. For counts of all crime from 0 to 1000 yards we have an estimated coefficient at 0.036 which jumps to 0.058 at 1000 to 2000 yards.

For violent crime the sign becomes negative (which will be discussed later in this section), but we still see change as proximity shifts. For counts of violent crime, the coefficient from 0 to 500 yards is -0.088 and jumps to -0.119 from 500 to 1000 yards. At 0 to 1000 yards the coefficient is -0.125 which varies slightly from to -0.115 from 1000 to 2000 yards.

Figure 10 below shows how coefficients vary by space:

Figure 10: Crime coefficients at differing proximities to sold home. Data from Table 4a



To better understand the results, a few overarching themes need to be explored. The coefficients for all counts of crime do not hold the same sign (-) as those of the counts for violent crime only:

This could potentially be a result of how crime is reported. As discussed by Greenberg & Beach (2004), the reporting of property or non-violent crime is subject to a decision making process for the victim rather than an instinctual or emotional decision that you would see for violent victimizations. This tells us that non-violent crime is more influenced by the social norms that the person lives under, thereby resulting unreliable results. For example, Model 3 implies that for every additional reported criminal event within 1000 yards of the sold home you should see a 4.2% *increase* in sale price. Common sense tells us that the average person would not pay a premium to live in an area with increased crime; however, this could simply be explained as those living in communities with increased crime have become accustomed to crime and less likely to report victimization. This underreporting of crime would lead to distorted results, and when combined with other potential unseen biases in the data, could flip the sign from what you would expect.

Is it crime or the perception of crime that matters?

The notion that results can be altered based on a persons' familiarization with the criminal landscape lends us to the question of whether or not reported crime is the important issue. It may be the perception of crime, rather than crime itself, that has influence on home purchasing behavior (Buonanno et al. 2013). How then are we to measure how a person perceives crime in an area? This dilemma is perhaps even more difficult to overcome than the inaccuracies in published UCR crime statistics (Loftin and McDowall 2010, Nolan et al. 2011), but an important issue to tackle for future research.

Differentiating between correlations and causation:

This brings us back to the issue of whether or not other factors in a neighborhood are distorting results. Is crime affecting housing prices, or is it a large host of other variables (education, neighborhood amenities, etc.) and crime just happens to correlate? It is likely that crime does have some impact, as shown by this study and others (Dubin & Goodman 1982, Lynch & Rasmussen 2001, Tita et al. 2006, Buonanno et al. 2013), but there is no real consistency between studies on how much of an impact there is. Of course every market will measure differently, but this is still an issue that needs to be considered.

Other possible limitations:

The Hedonic framework has been applied for decades and is a generally accepted way of answering value-based questions of heterogeneous goods. However, there are limitations to the data and research design that must be considered – all of which represent opportunities for new research. For this project to work, one must assume that house value is impacted by crime rather than crime being a result of surrounding house values. This leads to a broader based question with real socioeconomic implications, does poverty cause crime, or does crime create poverty around it?

We also must consider how crime is reported. As previously discussed in the literature review, statistical inaccuracies can come from the victim's decision to notify police and police discretion in the recording of events. The UCR may be the only data available, so potential bias in the data must be acknowledged.

Finally, because of the inaccuracies in reported crime data, we must consider other avenues to measure the perception of crime. In this study we are measuring how crime is valued or discounted in housing prices. Perhaps a survey system would be a better representation of the true landscape of how somebody perceives the crime in a neighborhood.

The overall conclusions drawn from this study are twofold; 1) crime does affect housing prices, and 2) the proximity of that crime matters. Through the investigation of this in Houston, it was found that the reporting of crime and how crime is measured can leave unintended impressions on conclusions drawn. Logic says that increased crime should decrease home value, but the data shows a different story. One possible explanation goes back to the Greenberg & Beach (2004) conclusion that the reporting of non-violent crime is more of a decision making process than the instinctual reaction of reporting violent victimizations. If crime is a normal fact of life for you, then you may be less likely to report to police. If, however, crime is not familiar, then you may be more prone to report. This familiarization with the criminal landscape leads to biases in the data, which means inaccurate conclusions are made.

5. References

Bark, Rosalind H., Daniel E. Osgood, Bonnie G. Colby, and Eve B. Halper. "How Do Homebuyers Value Different Types of Green Space?" *Journal of Agricultural and Resource Economics* 36.2 (2011): 395-415.

Bishop, Kelly C., and Alvin D. Murphy. "Estimating the Willingness to Pay to Avoid Violent Crime: A Dynamic Approach." *American Economic Review*. 101.3 (2011): 625-629.

Boggess, Greenbaum and George Tita. "Does crime drive housing sales? Evidence from Los Angeles." *Journal of Crime and Justice*. 36.3 (2013): 299-318

Boivin, Remi, and Gilbert Cordeau. "Measuring the Impact of Police Discretion on Official Crime Statistics: A Research Note." *Police Quarterly*. 14.2 (2011): 186-203.

Bosick, Stacey J., Callie Marie Rennison, Angela R. Gover, and Mary Dodge. "Reporting violence to the police: Predictors through the life course." *Journal of Criminal Justice*. 40. (2012): 441-451.

Box, G. E. P. and D. R. Cox. "An analysis of transformations." *Journal of the Royal Statistical Society*, Series B, 26, (1964): 211–243

Buonanno, Paolo, Daniel Montolio, and Josep Maria Raya-Vilchez. "Housing prices and crime perception." *Empir Econ*. 45. (2013): 305-321.

Chappell, Allison T., John M. MacDonald, and Patrick W. Manz. "The Organizational Determinants of Police Arrest Decisions." Crime & Delinquency. 52.2 (2006): 287-306.

Caudill, Affuso and Ming Yang. "Registered sex offenders and house prices: An hedonic analysis." *Urban Studies*. 52.13 (2015): 2425-2440

Ceccato, Mats Wilhemson. "The impact of crime on apartment prices: evidence of Stockholm, Sweden." Sustainable Regional Growth and Development in the Creative Knowledge Economy. *50th Congress of the European Regional Science Association.* (2010): 19-23

Clark, James Cosgrove. "Hedonic Prices, Identification, and the Demand for Public Safety." *Journal of Regional Science*. 30.1 (1990): 105-121

Colby, Bonnie G., and Steven Wishart. "Riparian areas generate property value premium for landowners." Agricultural & Resource Economics, University of Arizona, College of Agriculture and Life Sciences, 2002.

Csillag, Ferenc, Marie-Josee Fortin, and Jennifer L. Dungan. "On the Limits and Extensions of the Definition of Scale." *Bulletin of the Ecological Society of America*. 81.3 (2000): 230-232.

De-min, Jian-chun, and Hui-li. "GIS Supported Hedonic Model for Assessing Property Value in West Oakland, California." *Chinese Geographical Science.* 16.1 (2006): 70-78.

De Vor, Friso, and Henri L.F. De Groot. "The Impact of Industrial Sites on Residential Property Values: A Hedonic Pricing Analysis from the Netherlands." *Regional Studies*. 45.5. (2011)

Dubin, Robin, and Allen C. Goodman. "Valuation of Education and Crime Neighborhood Characteristics Through Hedonic Housing Prices." *Population and Environment*. 5.3 (1982): 166-181.

Dungham, J. L., J. N. Perry, M. R. T. Dale, P. Legendre, S. Citron-Pousty, M.-J Fortin, A. Jakomulska, M. Miriti, and M. S. Rosenberg. "A balanced view of scale in spatial statistical analysis." *Ecography*. 25. (2002): 626-640.

Fotheringham, A S, and D W S Wong. "The modifiable areal unit problem in multivariate statistical analysis." Environment and Planning. 23. (1991): 1025-1044.

Gehlke, C. E., and Katherine Biehl. "Certain Effects of grouping Upon the Size of the Correlation Coefficient in Census Tract Material." *Journal of the American Statistical Association*. 29.185 (1935): 169-170.

Gottfredson, Michael R., and Don M. Gottfredson. *Decision making in criminal justice*. Vol. 3. Springer, 1988.

Goodman, A.C. A comparison of block group and census tract data in a hedonic housing price model. Land Economics. 5 (1977): 483-487.

Graves, Phil, James C. Murdoch, Mark A. Thayer, and Don Woldman. "The Robustness of Hedonic Price Estimation: Urban Air Quality." *Land Economics*. 64.3 (1988): 220-33.

Greenberg, Martin S., and Scott R. Beach. "Property Crime Victims' Decision to Notify the Police: Social, Cognitive, and Affective Determinants." *Law and Human Behavior*. 28.2 (2004): 177-186.

Haab, Timothy C., and Kenneth E. McConnell. *Valuing Environmental and Natural Resources: The Econometrics of Non-market Valuation*. Cheltenham, U.K. : E. Elgar Pub. , 2002.

Hart, Timothy C., and Callie Rennison. United States. Department of Justice. *Reporting Crime to the Police, 1992-2000*. 2003.

Hipp, John R. "Block, Tract, and Levels of Aggregation: Neighborhood Structure and Crime and Disorder as a Case in Point." *American Sociological Review*. 72.5 (2007): 659-680.

Kim, Yeon-Su, and Aaron Wells. "The impact of forest density on property values." Journal of Forestry 103.3 (2005): 146-151.

Kotavaara, Ossi, Harri Antikainen, Mathieu Marmion, and Jarmo Rusanen. "Scale in the effect of accessibility on population change: GIS and a statistical approach to road, air and rail accessibility in Finland, 1900-2008." *Geographical Journal*. 178.4 (2012): 366-382.

Learnihan, V., K.P. Van Niel, B. Giles-Corti, and M. Knuiman. "Effect of Scale on the Links between Walking and Urban Design." *Geographical Research*. 49.2 (2011): 183-191.

Loftin, Colin, and David McDowall. "The Use of Official Records to Measure Crime and Delinquency." *Journal of Quantitative Criminology*. 26. (2010): 527-532.

Lynch, Allen K., and David W. Rasmussen. "Measuring the impact of crime on house prices." *Applied Economics*. 33. (2001): 1981-1989.

Manley, David, Robin Flowerdew, and David Steel. "Scales, levels and processes: Studying spatial patterns of British census variables." *Computers, Environment and Urban Systems*. 30. (2006): 143-160.

Mastrofski, Stephen D., R. Richard Ritti, and Debra Hoffmaster. "Organizational determinants of police discretion: The case of drinking-driving." Journal of Criminal Justice. 15.5 (1987): 387-402.

Mastrofski, Stephen D., Robert E. Worden, and Jeffrey B. Snipes. "Law Enforcement in a Time of Community Policing." Criminology. 33.4 (1995): 539-563.

Melde, Chris, and Callie Marie Rennison. "Intimidation and Street Gangs: Understanding the Response of Victims and Bystanders to Perceived Gang Violence." *Justice Quarterly*. 27.5 (2010): 619-666.

Mooney, Sian, and Ludwig M. Eisgruber. "The influence of riparian protection measures on residential property values: the case of the Oregon plan for salmon and watersheds." The Journal of Real Estate Finance and Economics 22.2-3 (2001): 273-286.

Mu, Lan, and Fahui Wang. "A Scale-Space Clustering Method: Mitigating the Effect of Scale in Analysis of Zoe-Based Data." *Annals of the Association of American Geographers*. 98.1 (2008): 85-101.

Nickels, Ernest L. "A note on the status of discretion in police research." Journal of Criminal Justice. 35. (2007): 570-578.

Nolan, James J., Stephen M. Haas, and Jessica S. Napier. "Estimating the Impact of Classification Error on the "Statistical Accuracy" of Uniform Crime Reports." *Journal of Quantitative Criminology*. 27. (2011): 497-519.

Openshaw S. and Taylor P. 1979 A million or so correlation coefficients: three experime nts on the modifiable area unit problem.
In N Wrigley (ed). Statistical Applications in the Spatial
Sciences. pp. 127-144 , London. Pion.

Openshaw S. 1984. *The Modifiable Areal Unit Problem*. Geobooks, Norwich, England.

Payton, Seth, Greg Lindsey, Jeff Wilson, John R. Ottensmann, and Joyce Man. "Valuing the Benefits of the Urban Forest: A Spatial Hedonic Approach." *Journal of Environmental Planning and Management* 51.6 (2008): 717-36.

Pope, J. “Fear of crime and housing prices: Household reactions to sex offender registries.” *Journal of Urban Economics*. 64 (2008): 601-614

Price, James I., Daniel W. McCollum, and Robert P. Berrens. "Insect infestation and residential property values: A hedonic analysis of the mountain pine beetle epidemic." Forest Policy and Economics 12.6 (2010): 415-422.

Rennison, Callie Marie. "An investigation of reporting violence to the police: A focus on Hispanic victims." *Journal of Criminal Justice*. 38. (2010): 390-399.

Rosen, Sherwin. "Hedonic Prices and Implicit Markets: Product Differentiation in Pure Competition." *Journal of Political Economy*. 82.1 (1974): 34-22.

Shihadeh, Edward S., and Wesley Shrum. "Serious Crime in Urban Neighborhoods: Is There a Race Effect?." *Sociological Spectrum: Mid-South Sociological Association*. 24.4 (2004): 507-533.

Studenmund, A. H. *Using Econometrics: A Practical Guide*. 5th Edition. Boston: Pearson Addison Wesley, 2006. 414-418.

Tagashira, Naoto, and Atsuyuki Okabe. "The Modifiable Areal Unit Problem in a Regression Model Whose Independent Variable is a Distance from a Predetermined Point." *Geographical Analysis*. 34.1 (2002): 1-20.

Taylor, Lara. "The Hedonic Method." A Primer on Nonmarket Valuation. : Springer Netherlands, 2003.

Tita, George E., Tricia L. Petras, and Robert T. Greenbaum. "Crime and Residential Choice: A Neighborhood Level Analysis of the Impact of crime on Housing Prices." *Journal of Quantitative Criminology*. 22. (2006): 229-317.

Tolsman, Jochem, Joris Blaauw, and Manfred te Grotenhuis. "When do people report crime to the police? Results from a factorial survey design in the Netherlands, 2010." *J Exp Criminology*. 8. (2012): 117-134.

Troy, M. Grove. “Property values, parks, and crime: A hedonic analysis in Baltimore, MD.” *Landscape and Urban Planning*. 87.3 (2008): 233-245

Truman, R. Morgan. “Crime Victimization.” *U.S. Department of Justice: Bureau of Justic Statistics* (2015).

Vasquez Lavin, Dresdner and Aguilar. “The value of air quality and crime in Chile: a hedonic wage approach.” *Environment and Development Economics*. 16 (2011): 329-355

Wang, Fahui. "Job Access and Homicide Patterns in Chicago: An Analysis at Multiple Geographic Levels Based on Scale-Space Theory." *Journal of Quantitative Criminology*. 21.2 (2005): 195-217.