Study Area Summary Report

EPC-14-037

March 2016

Prepared for
California Energy Commission

Prepared by
Center for Sustainable Energy

Center for Sustainable Energy®

Contents

I. Executive Summary

The California Energy Commission funded a series of projects designed to explore how sociocultural factors influenced customer adoption of home energy efficiency measures. This document is the Study Area Summary Report for one of those projects, EPC-14-037, led by the Center for Sustainable Energy (CSE). This report summarizes the evaluation, selection, characterization, and modeling of two regional energy efficiency markets in California. Data analyzed for this market characterization include residential HVAC change-outs recorded in the CalCERTS registry between January 2010 and June 2014, demographic and building characteristic data from the American Community Survey and census block group level voting records from the Statewide Database maintained by the University of California – Berkeley School of Law.

Hispanic and Latino households were selected as the focus of the research, due to their large and geographically dispersed population across the state and common cultural and linguistic characteristics. San Diego and Fresno counties were chosen as the study areas based on activity observed in the CalCERTS registry, climate, sociocultural composition, location within an IOU territory and the project team's relationships with stakeholders in these markets. Characteristics of the San Diego- and Fresno-area markets are summarized below.

- Both San Diego and Fresno County have large Hispanic and Latino communities and many block groups where over 25% of the population speaks Spanish at home. However, we discovered key differences in the geographic distribution of Spanish speakers between these two counties. Notably, the Spanish speakers in San Diego County are more evenly dispersed throughout the county (with the exception of the communities bordering Mexico). Understanding these patterns in the complex ethnic geography of the study areas will inform careful sampling strategies in the later stages of this research.
- The median home values in San Diego and Fresno counties have opposite patterns. In San Diego County, the most expensive homes are in the city center and near the coast, while in Fresno County the median home value in the city center is much lower than the surrounding areas. In Fresno County, there appears to be a slight positive relationship between home value and number of records in the CalCERTS registry.
- Political leanings in San Diego have a high degree of spatial heterogeneity, with concentrations of Democratic-leaning voters located in the City of San Diego's metropolitan core, the metro areas of inland valley cities, and along the Pacific coastal zone, while large swaths of the inland valleys and unincorporated areas lean more heavily Republican. Fresno shows a similar distribution of political leaning, with urban areas trending more Democratic and suburban/rural areas (particularly those to the northeast) trending more heavily Republican.

To gauge the relationship of sociocultural factors with the propensity of homeowners to perform an HVAC upgrade, a number of regression models were run. The results of this modeling exercise indicate a statistically significant but inconclusive correlation. Further investigation of this relationship will be conducted through the focus groups, structured interviews and surveys in the next phase of the project.

II. Introduction

The current policy framework used to assess the potential for and likely adoption of residential energy efficiency measures is driven primarily by estimates of cost effectiveness. While an important component of the equation, cost effectiveness calculations alone fail to accurately predict adoption and market potential, as they do not capture the multitude of factors influencing the decision-making process of individual market actors. In recognition of this limitation in the current policy framework, the California Energy Commission funded a series of projects designed to explore how sociocultural factors influenced customer adoption of home energy efficiency measures. This document is the Study Area Summary Report for one of those projects, EPC-14-037, led by the Center for Sustainable Energy (CSE).

This report summarizes the evaluation, selection, characterization, and modeling of regional energy efficiency markets in California with the intention of selecting two study areas in which to conduct focus groups, surveys, online experimentation and field work. The research team's goal was to identify markets that have large, dynamic residential energy efficiency markets with high demographic variability and sociocultural variation in customer participation. Preference was given to areas which together have a demographic and climatic composition that would be most useful for extrapolating results to the state at large, including a large coastal metropolitan area and a large inland metropolitan area within the state's investor owned utility (IOU) service territories. Additional preference was given to areas where the project team has stakeholder relationships, which will facilitate the field work called for in future tasks.

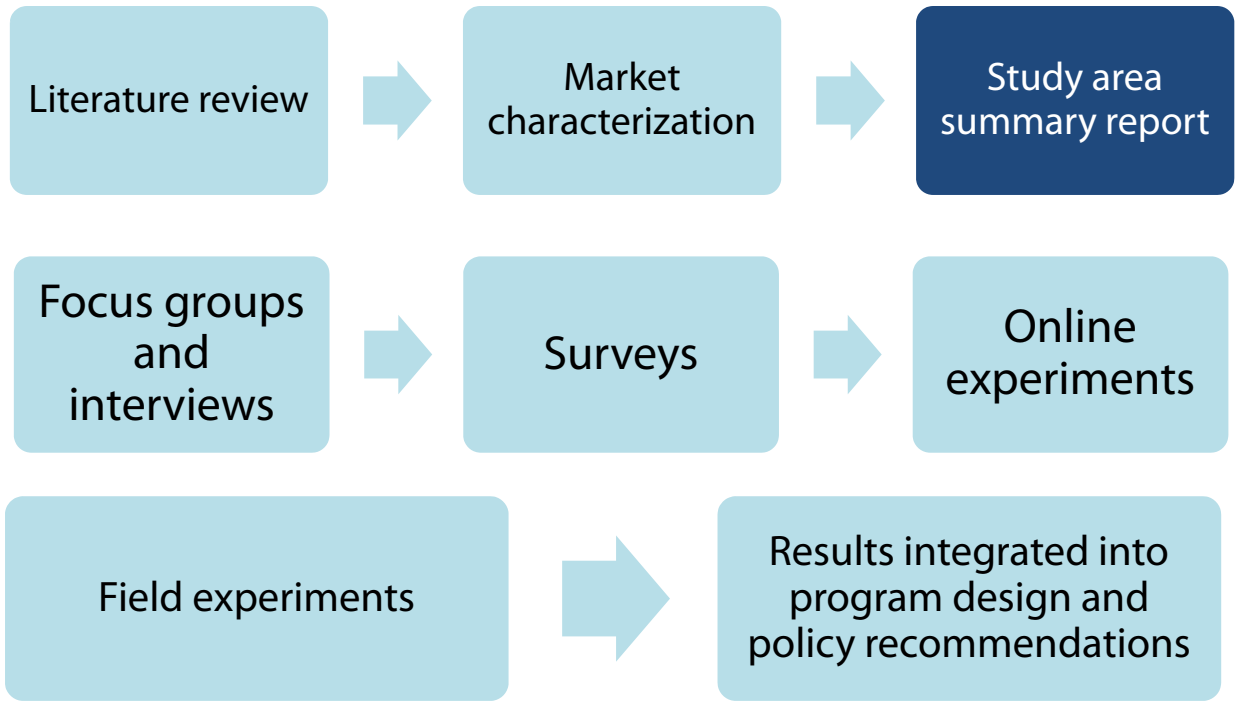


Figure 1: Research activities for project EPC-14-037

III. Data Overview

CalCERTS, Inc. Registry Data

This market characterization uses records of residential HVAC change-outs as an indicator of energy efficiency activity. In California, HVAC alterations must comply with the state's Building Energy Efficiency Standards, known as Title 24, Part 6. To demonstrate compliance with Title 24, Part 6, various forms must be uploaded to a Home Energy Rating System (HERS) registry. This study uses records from the CalCERTS HERS Provider registry, as CalCERTS is an approved HERS Provider by the Energy Commission. Registry records analyzed were from January 1, 2010 through June 30, 2014, when the 2008 code was in effect.

During the 2008 code cycle, Title 24, Part 6 required up to three compliance forms to be filed for each HVAC alteration. The CF1R-ALT and CF6R were submitted by the contractor (or consultant hired by the contractor) to confirm that energy features of the installation and performance specifications of the appliance conformed to the standards. In coastal climate zones, the submission of the CF1R-ALT and CF6R forms completed the Title 24, Part 6 compliance process. In inland climate zones (see zones 2 and 9-16 in Map1), the more extreme temperatures prescribed an additional level of assurance that the HVAC alteration met energy efficiency standards. A certified HERS Rater was required to test the system for proper refrigerant charge, fan watt draw and/or duct leakage, depending on what equipment was installed in the home. This field verification resulted in the submission of a third form, CF4R, to the HERS registry.

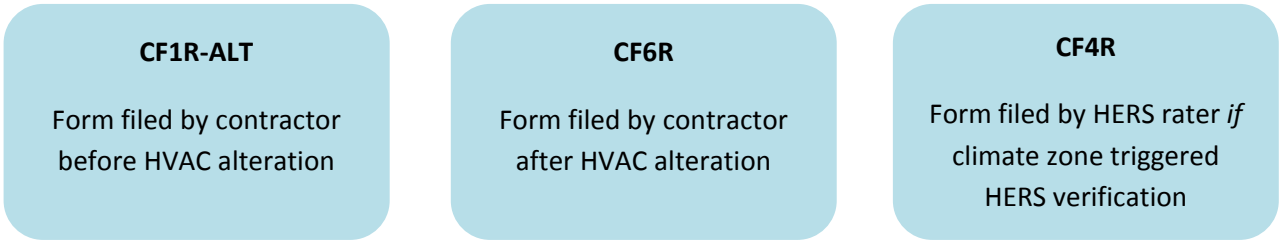
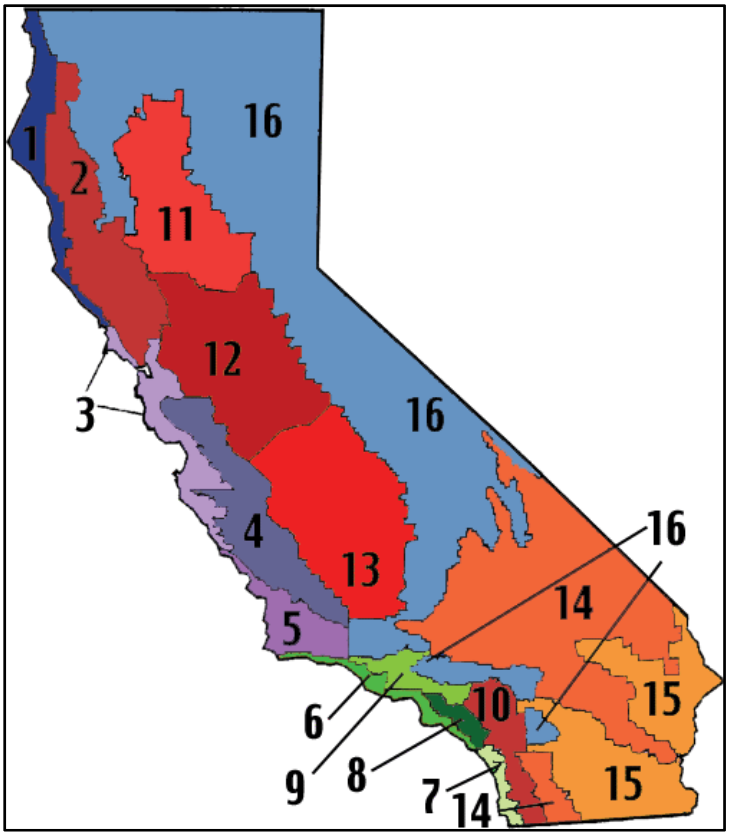


Figure 2: Title 24, Part 6 compliance forms for 2008 code cycle

The CalCERTS registry data offers value as a proxy for energy efficiency market activity because it represents upgrades performed across jurisdictions regardless of participation in rebate or financing programs. (Other analyses of energy efficiency market activity may rely on data from a particular incentive or financing program.) However, the data is limited in that it only reflects HVAC upgrades that went through the required permitting and Title 24, Part 6 compliance process. Estimates of permit

compliance rates for residential HVAC alterations in California range from 10 to 38 percent.[1] Furthermore, there are other types of energy efficiency improvements – such as insulation, air sealing, window upgrades, water heater replacements or lighting upgrades – that did not require HERS verification and thus are not reflected in the CalCERTS registry data. Some HVAC alterations may have been performed in conjunction with insulation, air sealing or other upgrades as part of the "whole house approach" to energy efficiency; other alterations were one-for-one replacements of broken or aging equipment and may not have been done with energy efficiency as an explicit goal.



Map 1: California Building Climate Zone map

American Community Survey Data

Demographic and building characteristic data at the block group resolution was obtained from the American Community Survey (ACS). This data is the result of yearly random sampling of the American public by the U.S. Census Bureau. It is publicly available in a weighted and aggregated form with personally identifiable information removed. Each year, the Census Bureau issues five-year estimates of

[1] DNV-GL, *HVAC Permitting: A Study to Inform IOU HVAC Programs*, 2014

each variable in the ACS. The ACS data presented in this report is drawn from the 2010-2014 five-year estimates.

The following ACS variables were used in the market characterization and modeling.

- Distribution of home value
- Distribution of household income
- Percentage of householders who speak Spanish at home
- Percentage of householders who identify as Hispanic or Latino
- Distribution of build year of home (in 10-year increments after 1930)

All variables reported in U.S. dollars (USD) are reported in 2014 inflation-adjusted dollars.

Political Party Affiliation Data

Block group level voting records were obtained from the publicly available Statewide Database maintained by the University of California – Berkeley School of Law. Voting records for the 2014 California gubernatorial election between Jerry Brown (D.) and Neal Kashkari (R.) were used as a proxy for political party affiliation in this analysis. Using an algorithm provided by the University of California – Berkeley Law School, the vote counts were aggregated or disaggregated from precincts to their corresponding block groups. The votes cast for each candidate were then normalized to the total votes cast in that block group to obtain a percentage of the vote. The Republican advantage in each block group was calculated as the percentage point difference between the percentages of votes cast for the Republican and Democratic candidates.

Study Area Selection

Prior to selecting the study areas, a determination was required on which sociocultural group(s) to focus on for the remainder of project. The project literature review examined three groups: Hispanics/Latinos, Asian Americans, and African Americans. Each of these groups presents their own characteristics in terms of definition, data availability, and conduciveness to research.

Asian Americans make up 13.5% of the California population; however, they are, as defined by block group-level census data, a broad group including distinct sociocultural groups ranging from Japanese to Indians to Saudis. This diversity of cultural and linguistic characteristics presents distinct challenges for the project.

According to the US Census Bureau's American Community Survey, African Americans represent 5.9% of California residents. In addition to representing a smaller percentage of the state's population than Hispanic and Asian Americans, their presence is highly concentrated, with large percentages in a few

metropolitan areas and relatively sparse populations in many other parts of the state. This creates challenges for conducting effective surveys and field research in later stages of the project.

Hispanic/Latino populations are both large (38.2% of the population) and geographically dispersed, making surveys and field research easier to conduct. In addition, many members of the state's Hispanic/Latino population are from a single country (83% from Mexico) and share a common language, Spanish. These common cultural and linguistic elements present an opportunity to explore the underlying influence that it plays within the wider cultural framework, and how that influences changes with a generational context. For these reasons, California's Hispanic/Latino population was selected as the focus of subsequent phases of this research and data associated with the prevalence of this population as used in study area selection.

Using the data outlined in the previous section, the project team performed a review of counties seeking to identify large, dynamic residential energy efficiency markets with high demographic variability and sociocultural variation in customer participation. In addition to the markets and sociocultural factors, the team sought areas of study with geographic and climatic variability to help extrapolate results to the state at large. To achieve this diversity, the goal was to identify a large coastal metropolitan area and a large inland metropolitan area within the state's IOU service territories. Table 1 provides a summary of total registry activity for the 10 California counties with the highest volume of CF1R-ALT records during the 2008 cycle of Title 24, Part 6. In addition to registry data, the table provides countywide statistics on statistics on owner occupancy among Hispanic households, percentage of households where Spanish is spoken at home, whether or not the county includes a coastal climate zone, and whether or not the counties' large metro areas are served by an IOU or a publicly-owned utility (POU).

County	CF1R-ALT count	Percentage of total data	% occupied homes with Hispanic householder	% households who speak Spanish at home	Primary climate	Metro utility service
Sacramento	44,133	17%	16%	13%	Inland	POU
Los Angeles	36,482	14%	36%	34%	Coastal	POU
Riverside	25,418	10%	34%	30%	Inland	POU
Orange	13,665	5%	23%	22%	Coastal	IOU
San Bernardino	13,500	5%	39%	33%	Inland	IOU
Placer	11,145	4%	9%	7%	Inland	IOU
Contra Costa	10,788	4%	17%	15%	Inland	IOU
Fresno	9,778	4%	41%	32%	Inland	IOU
San Diego	8,519	3%	23%	22%	Coastal	IOU
San Joaquin	8,162	3%	31%	25%	Inland	IOU

Table 1: Characteristics of 10 counties with the most CF1R-ALT records in CalCERTS registry

The top three counties, in terms of recorded CF1R-ALTs, are Sacramento, Los Angeles and Riverside, making them strong candidates for study; however, each of these counties has large portions of their metro area served by a POU. Of the seven remaining counties, registry activity ranges from 8,162 to 13,665, with two counties located largely in coastal zones and five in the inland portion of the state. The two coastal counties (Orange and San Diego), both have comparable percentages of Hispanic households and households where Spanish is spoken, although Orange has considerably more registry activity. Of the five remaining inland counties, three (San Bernardino, Fresno and San Joaquin) have large percentages of Hispanic households and households where Spanish is spoken.

Based on the climate, registry activity, and sociocultural composition of these counties, two counties were identified as candidate study areas in both the coastal and inland areas of the state. Ultimately, San Diego and Fresno were chosen as the coastal and inland study areas for this project based on the characteristics of these communities combined with the project team's relationships with stakeholders (e.g. contractors, local governments, CBOs) in these markets, which will be critical in the implementation of field experiments in later tasks.

IV. Study Area Characterization

The following maps and figures provide an overview of the study area, providing detail of the number of records in the CalCERTS registry and contextualizing these records with the demographic and housing variables introduced above. These demonstrations are designed to contextualize the following modeling section as well as provide inform the forthcoming survey work in these communities.

CalCERTS, Inc. Registry Records

Cleaning and Pre-processing

The analysis incorporates all records in San Diego and Fresno counties for which a CF1R-ALT form is registered in the CalCERTS database. The final dataset for analysis was filtered to remove multi-family dwellings. This was done by removing addresses with "unit", "apt", "#", or "number" in the address. Additionally, addresses with identical (to six decimal points) rooftop-level geocoding were removed from analysis. Block groups identified as having no occupied 1-unit homes were removed from the analysis. Twenty-five block groups were removed in this filtering step and, of these, two block groups had one home with a record in the CalCERTS registry. In this report, single family home (SFH) will refer to 1-unit occupied dwellings as defined by the U.S. Census Bureau. The cleaning and filtering process resulted in a combined 13,712 records to be analyzed for the study.

Study Region	Number of records in CalCERTS registry after cleaning and filtering data
San Diego County	7,127
Fresno County	6,585

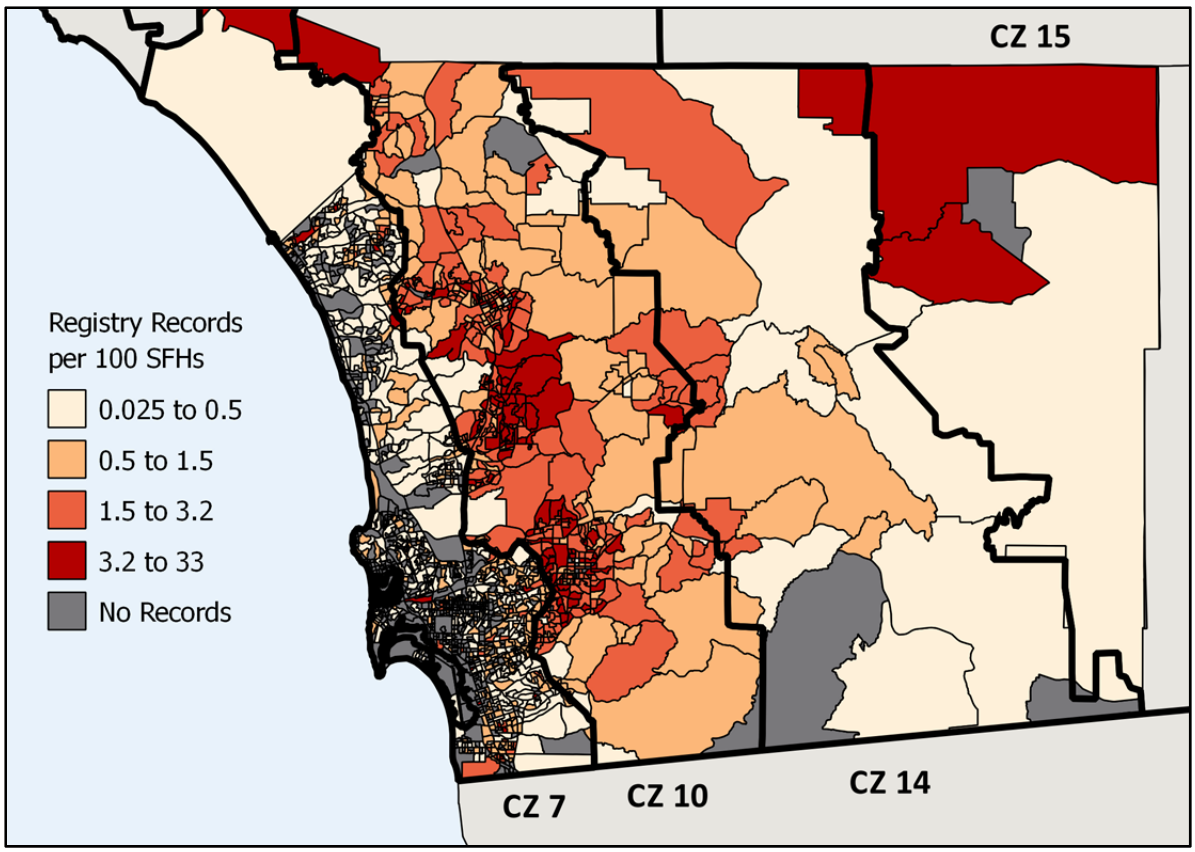
Table 2: Records in CalCERTS registry after cleaning and filtering data

Record Distribution

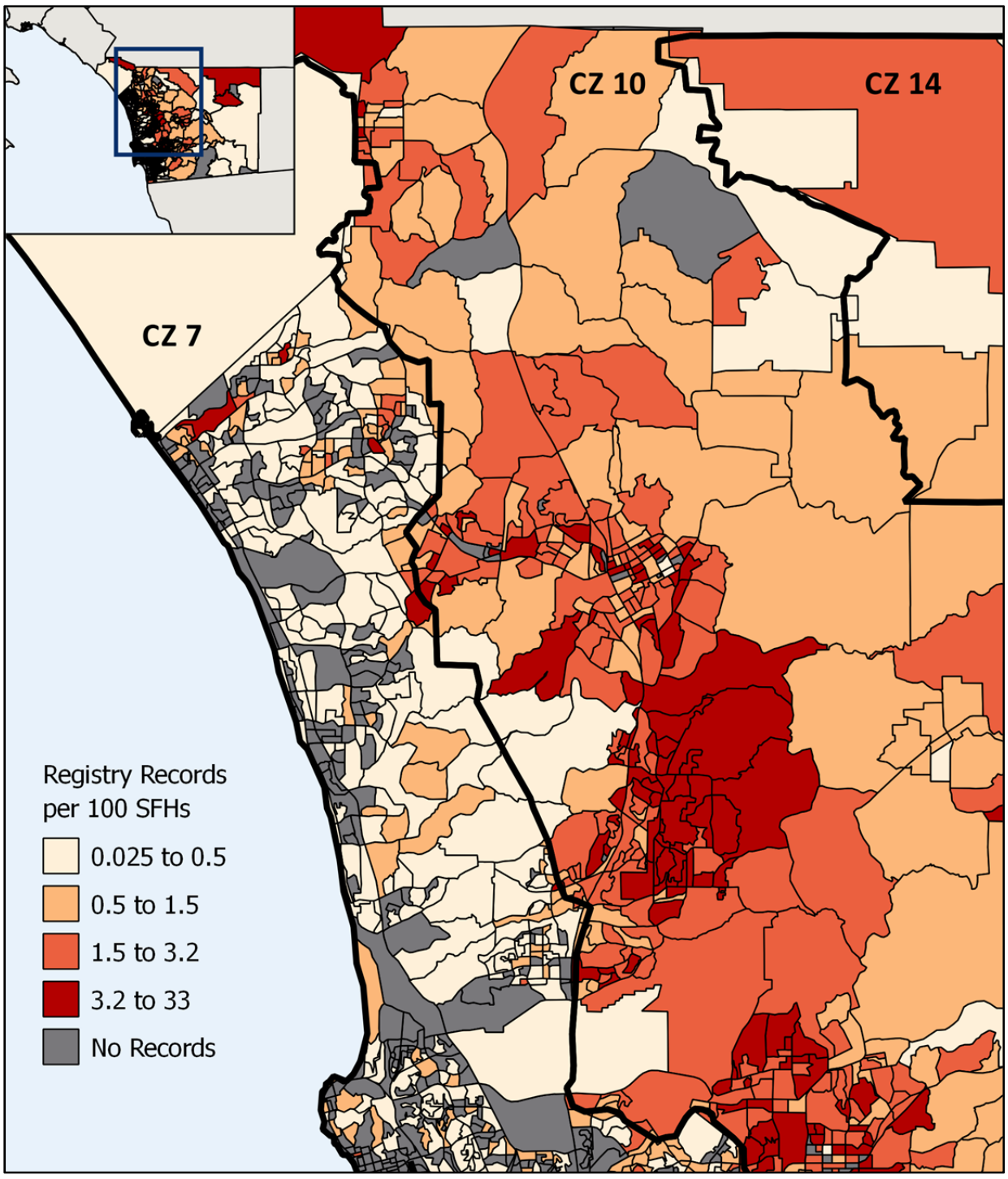
The following maps illustrate the saturation of HVAC alterations as measured by CalCERTS registry records throughout San Diego and Fresno counties.

In San Diego County, the city center has very low saturation of homes with records in the registry, even compared to other block groups in the coastal area. This may be because these are primarily old homes without central air conditioners and/or may be due to a higher density of multi-unit dwellings in the city center. The coastal region, climate zone 7, has lower levels of saturation of homes with registry records. This is likely due to a more temperate climate and the less rigorous Title 24, Part 6 requirements discussed in the previous section. An analysis of responses to the 2009 Residential Appliance Saturation Survey (RASS) administered by the Energy Commission indicated that 89.6% of single family homes in climate zone 10 have central air conditioning, compared to only 39.2% of single family homes in climate zone 7.

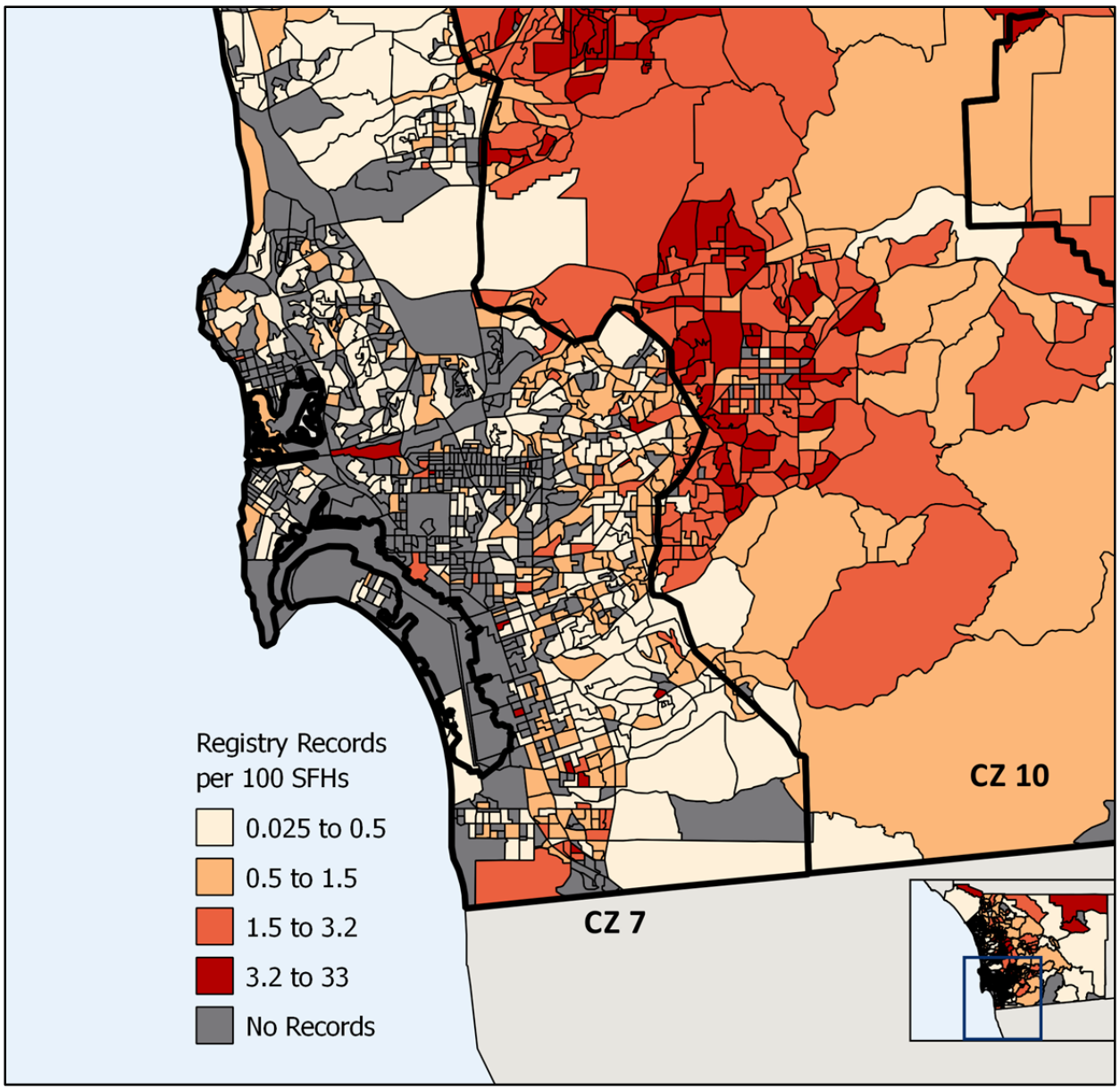
In Fresno County, there is a very different distribution of HVAC alteration registry records. The majority of the records are concentrated in the city center, with lower saturation in the rural areas. In addition, the block groups closest to, but not within, the city center of Fresno have lower saturation rates than those in either side.



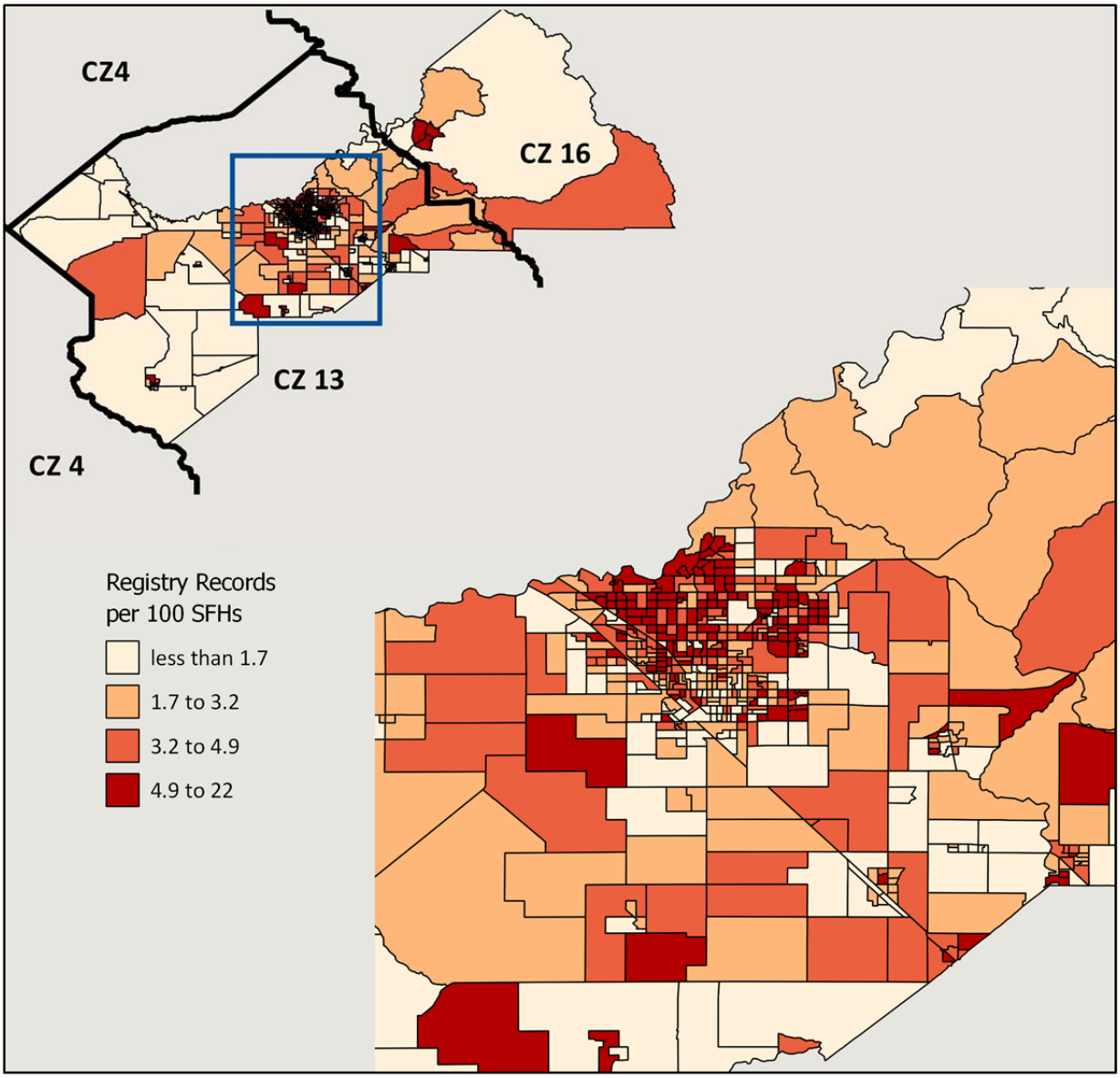
Map 2.1: Geographic distribution of CF1R-ALT normalized per 100 occupied single family homes by block group in San Diego County. Each color bin denotes one quartile. Block groups with no permits removed before calculating the quartiles. The block groups with no records are shown in gray.



Map 2.2: Geographic distribution of CF1R-ALT normalized per 100 occupied single family homes by block group in northwest San Diego County. Each color bin denotes one quartile. Block groups with no permits removed before calculating the quartiles. The block groups with no records are shown in gray.



Map 2.3: Geographic distribution of CF1R-ALT normalized per 100 occupied single family homes by block group in southwest San Diego County. Each color bin denotes one quartile. Block groups with no permits removed before calculating the quartiles. The block groups with no records are shown in gray.



Map 2.4: Geographic distribution of CF1R-ALT normalized per 100 occupied single family homes by block group in Fresno County. Each color bin denotes one quartile.

Spanish Speakers and Hispanic and Latino Householders

The plots and maps in this section illustrate the possible relationship between Hispanic and Latino communities and energy efficiency adoption as represented by HVAC alterations in the CalCERTS registry. Although there is a lot of heteroscedasticity in these demographic variables with respect to registry records, there is some useful information to be gained.

The geographic distribution of Spanish-speaking and Hispanic or Latino communities within San Diego and Fresno counties is quite different. Both counties have large Hispanic and Latino communities and many block groups where over 25% of the population speaks Spanish at home. Fresno County has 149 block groups with a Spanish-speaking population greater than 50% compared to 49 in San Diego County. San Diego County has many more block groups where householders who speak Spanish at home make up less than 25% of the population (as seen in Figure 3.1). San Diego County block groups where greater than 50% of the population speaks Spanish at home are almost all located in the southern-most part of the county near the Mexican border. There are few block groups in inland San Diego County in which Spanish-speaking or Hispanic or Latino populations make up greater than 70% of the total population.

The box and whisker plots in Figures 3.3 and 3.4 demonstrate a slight negative relationship between the percentage of householders who speak Spanish at home and the median number of records in the CalCERTS registry in Fresno County and inland San Diego County for block groups where the Spanish speaking population is less than 70%.

In Fresno County and inland San Diego County there is a slight leveling off or possible increase in the number of CF1R-ALT records per 100 homes in block groups as the Spanish-speaking population grows beyond 50%. Southwestern San Diego County has a large number of block groups with a high percentage of Hispanic and Latino householders as well as a range of CF1R-ALT record saturation (as seen in Map 4.2 and 2.3). The clustering of the Spanish-speaking population into one area, combined with the scarcity of CalCERTS records in coastal San Diego County makes interpretation of the relationship between the number of CF1R-ALT records and the Spanish-speaking population difficult.

There is some evidence of a parabolic relationship with the highest number of records in block groups with Spanish-speaking populations between 30 and 60% and fewer records with both less and greater density of Spanish speakers. However, the clustering of Spanish speakers makes it more likely than in other regions that there may be an unknown variable to explain the low number of records in block groups with speaking populations above 75%. Understanding these patterns in the complex climatic and ethnic geography of the study areas will inform careful sampling strategies in the later stages of this research.

Block groups with a high percentage of Hispanic and Latino householders and a high percentage of householders who speak Spanish at home are highly correlated geographically and exhibit the same patterns as described above. For this reason, these two variables were run separately in the regression modeling exercise.

As discussed previously, the large population of Hispanics and Latinos living in California makes them an ideal group for our research. However, the size of the Latino population also poses several challenges that must be carefully addressed in later stages of the research. There is heterogeneity in acculturation within all minority immigrant populations as they assimilate into a new culture. The large population of Latinos means there is a great deal of diversity of views on energy efficiency within this community. Putting aside the individual differences arising from each person's unique experiences, there are some community level effects that we aim to survey in this research project. One of these is the effect of the amplitude of saturation on individual views of energy efficiency. Does a Latino homeowner living in a majority Latino community have similar associations with energy efficiency messaging to a Latino homeowner living in a community where less than 20% of the population identifies as Latino or Hispanic? The answer to this question will help determine how to translate the findings of this research project to other areas of California where the Hispanic population is sparser.

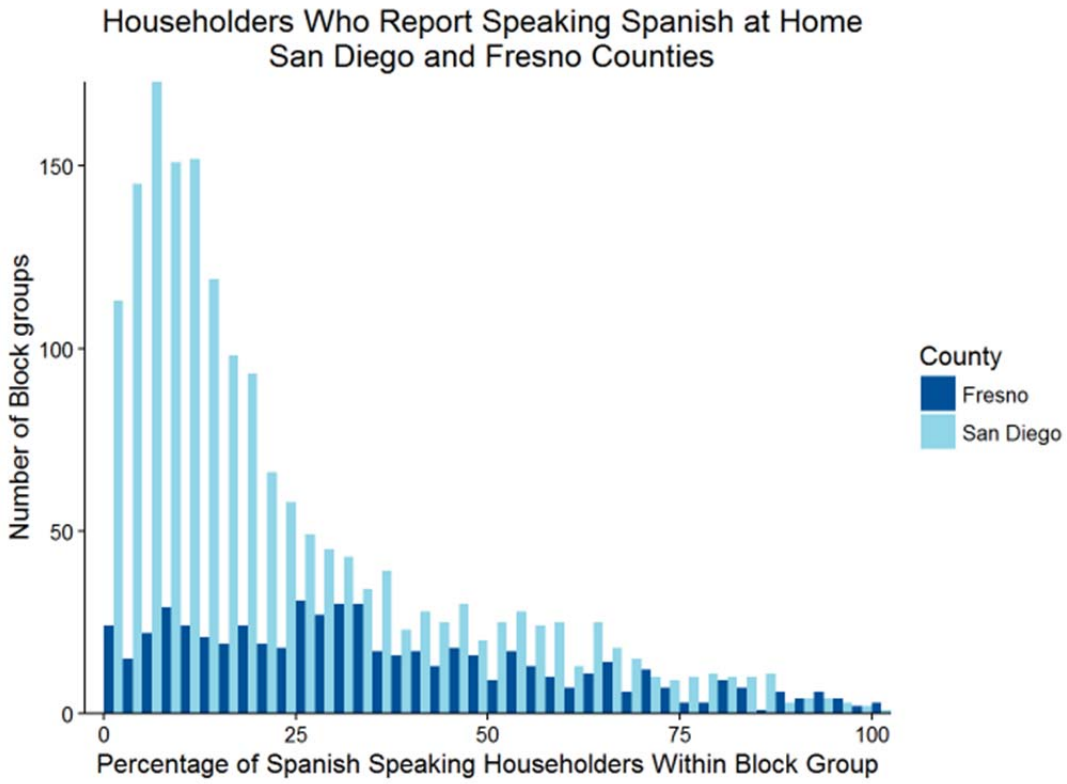


Figure 3.1: Distribution of Spanish speaking populations at block group level. The x-axis depicts the range of percentages of householders who speak Spanish at home within a given block group. The height of each bar is the count of block groups within San Diego (shown in light blue) and Fresno (shown in dark blue) counties that have a Spanish speaking population of the size depicted on the x-axis.

Figures 3.2, 3.3, 3.4 (see below): Relationship of HVAC alteration saturation to Spanish speaking communities (i.e., block groups) in coastal San Diego County (3.2), inland San Diego County (3.3) and Fresno County (3.4). The x-axis depicts the range of householders who speak Spanish at home within a given block group. Each observation in the box and whiskers plot is one block group. The blue area represents the 25%-75% range of the distribution (i.e., 2nd and 3rd quartile), the median is indicated by a solid black line in the center of the blue, and outliers are represented with a single black point.

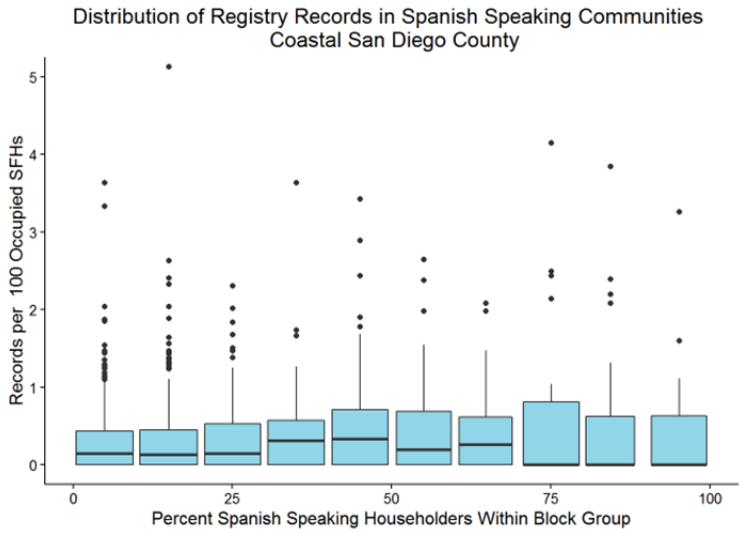


Figure 3.2

(The upper bound of the Y-axis has been reduced to 10 records per 100 SFHs to more clearly display the medians. This reduction prevents some outliers from being visible in the graphic.)

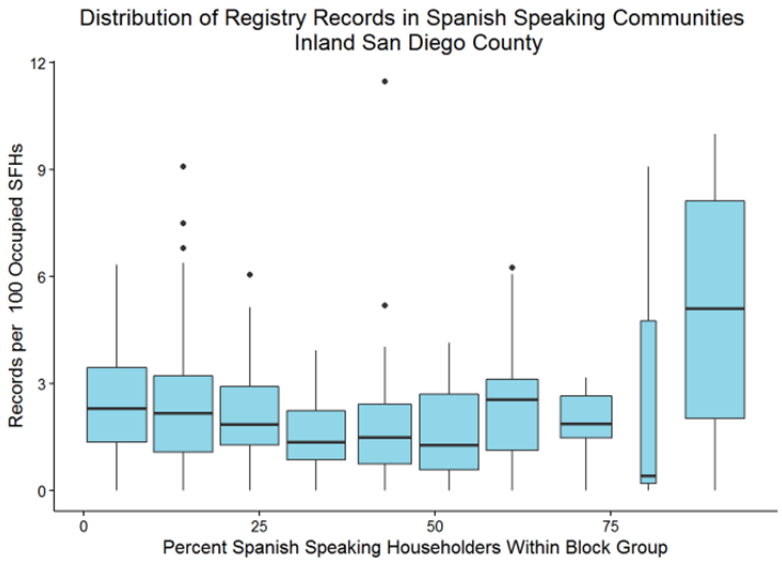


Figure 3.3

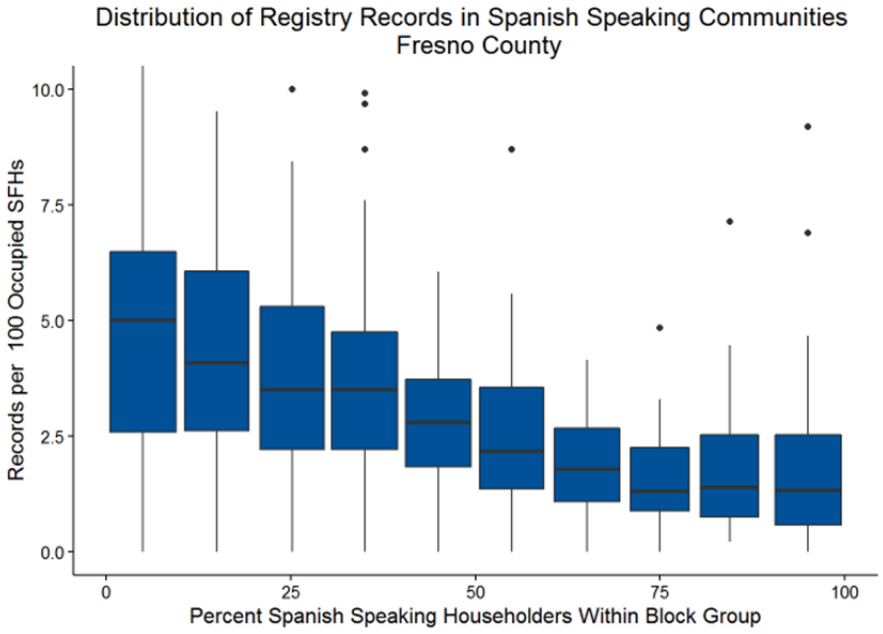
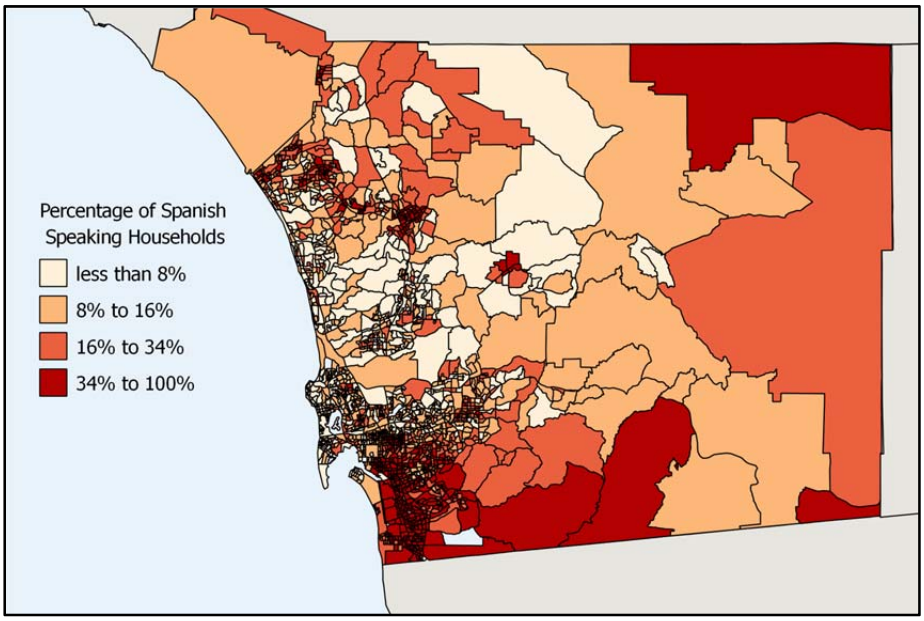
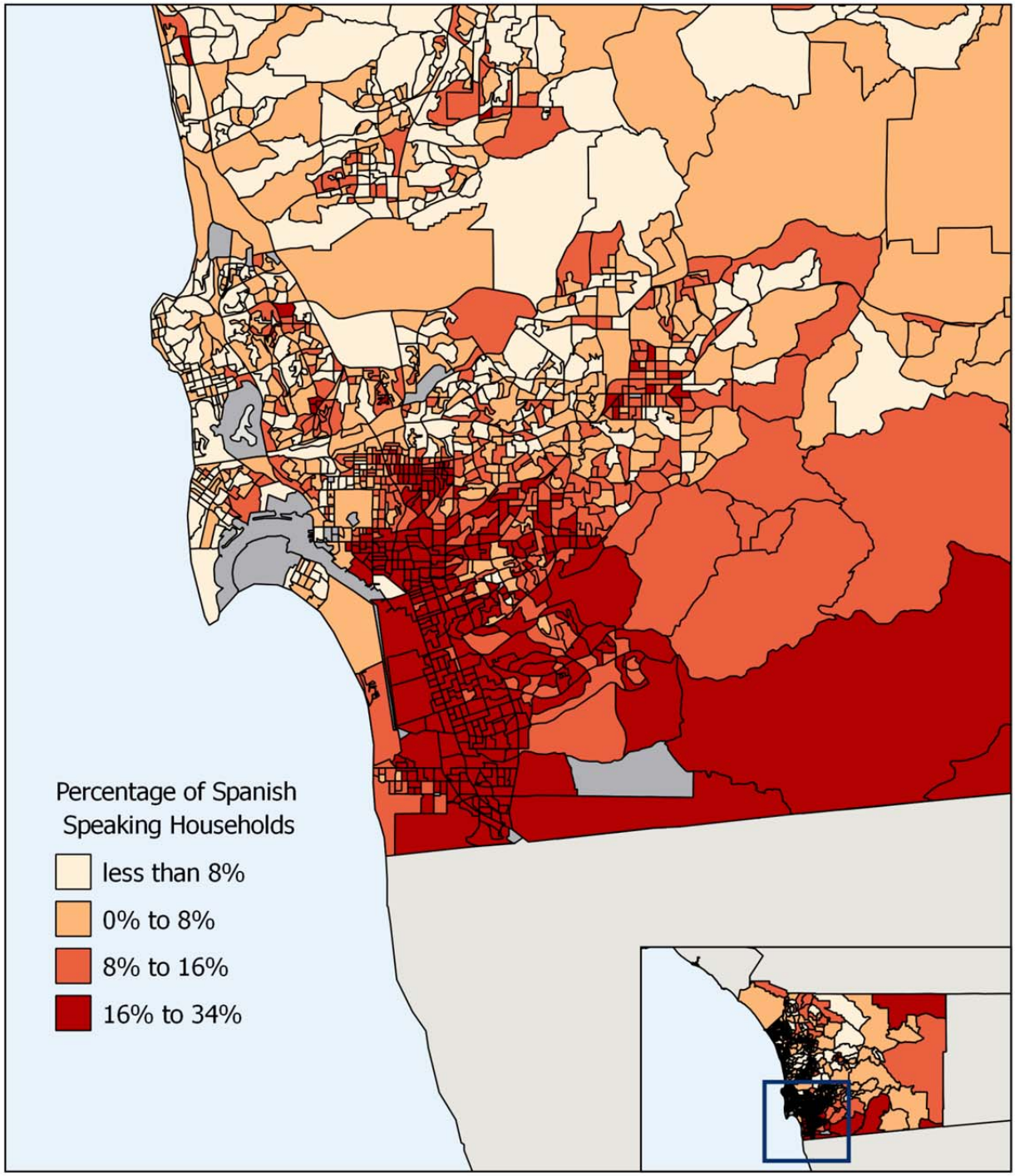


Figure 3.4

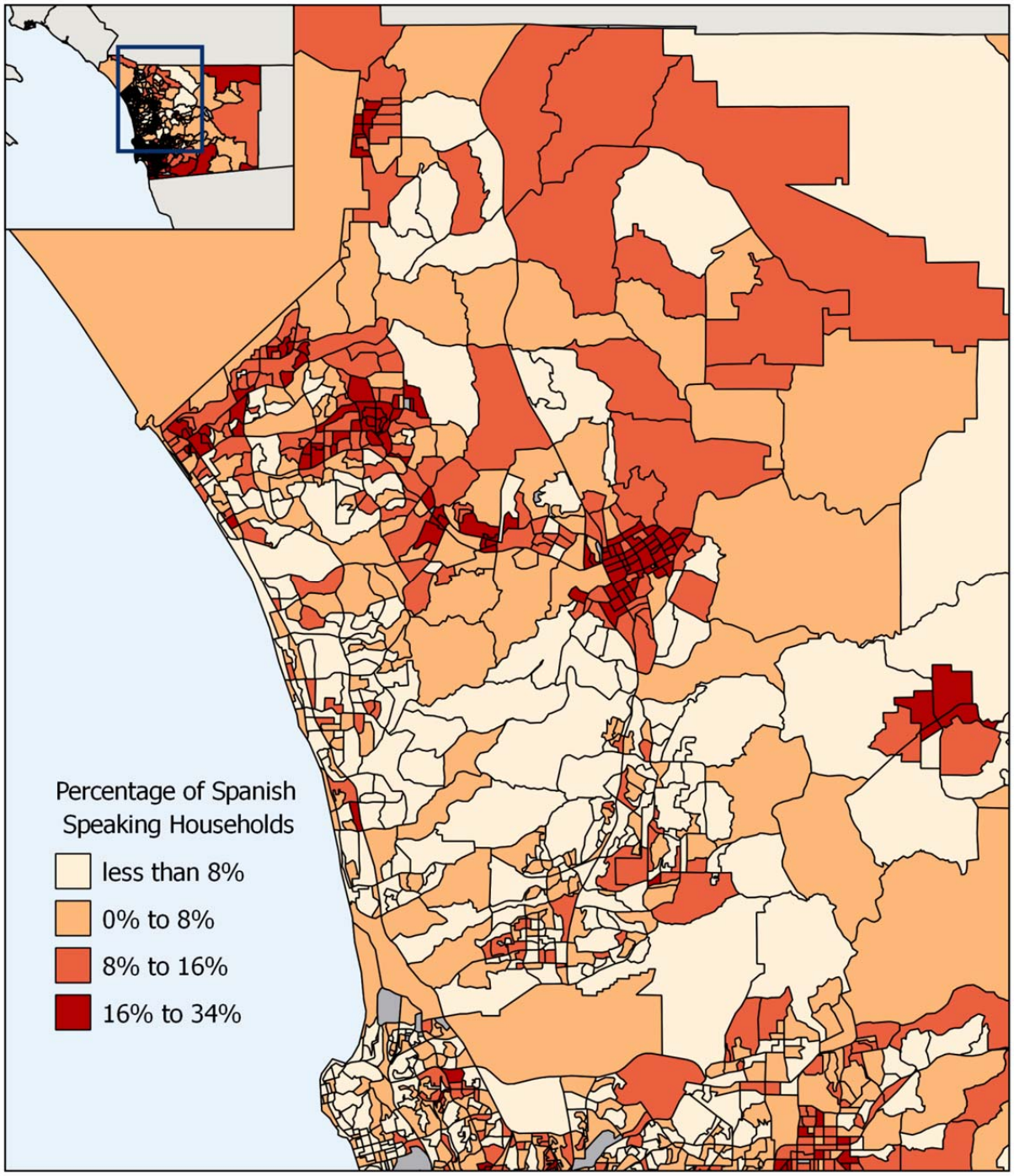
(The upper bound of the Y-axis has been reduced to 10 records per 100 SFHs to more clearly display the medians. This reduction prevents some outliers from being visible in the graphic.)



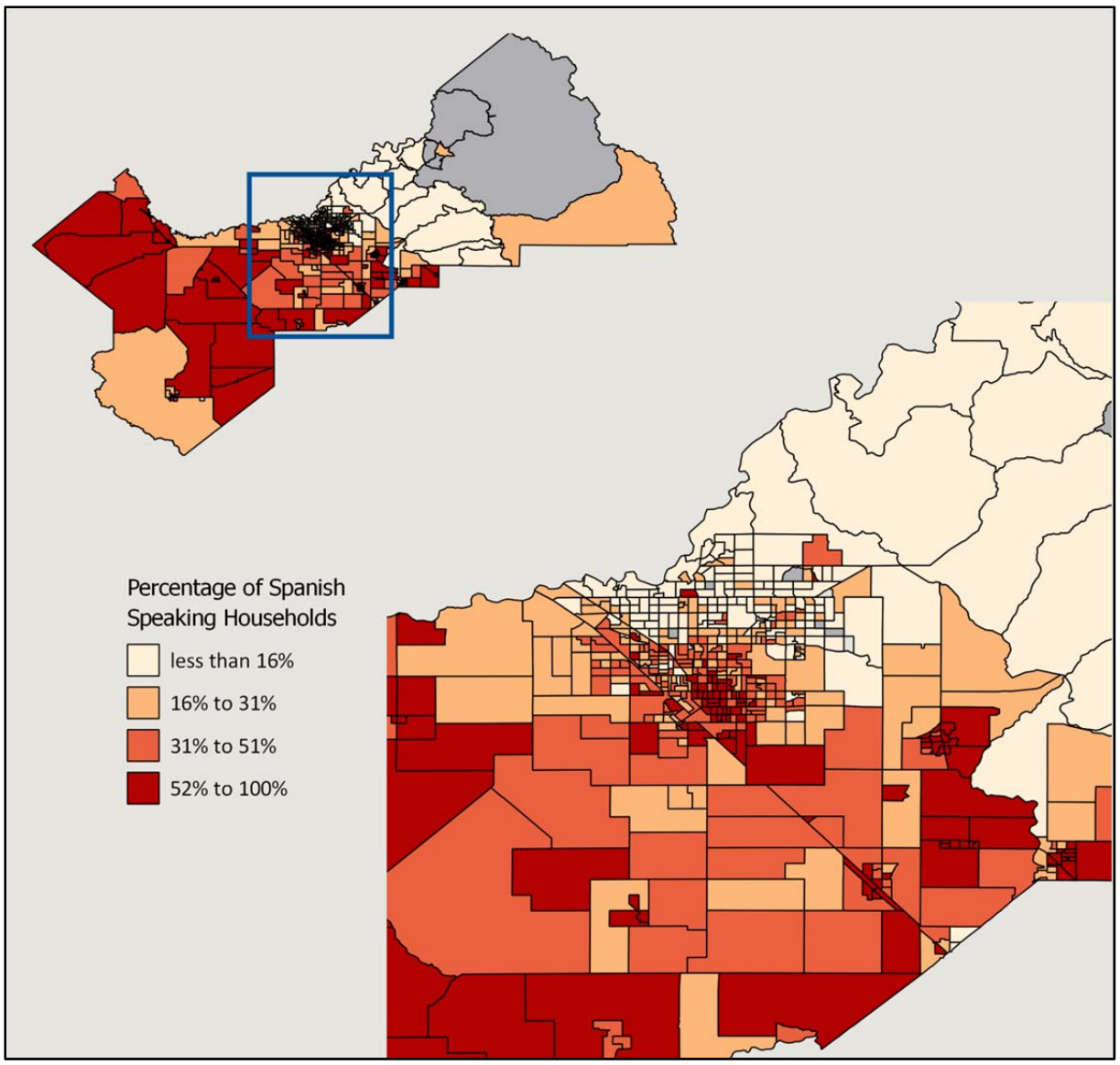
Map 3.1: Percentage of householders within each block group in San Diego County which has a householder who speaks Spanish at home. Each color bin represents one quartile in the distribution of Spanish speaking populations across all block groups.



Map 3.2: Percentage of householders within each block group in southwestern San Diego County which has a householder who speaks Spanish at home. Each color bin represents one quartile in the distribution of Spanish speaking populations across all block groups.



Map 3.3: Percentage of householders within each block group in northwestern San Diego County which has a householder who speaks Spanish at home. Each color bin represents one quartile in the distribution of Spanish speaking populations across all block groups.



Map 3.4: Percentage of householders within each block group in northwestern Fresno County which has a householder who speaks Spanish at home. Each color bin represents one quartile in the distribution of Spanish speaking populations across all block groups.

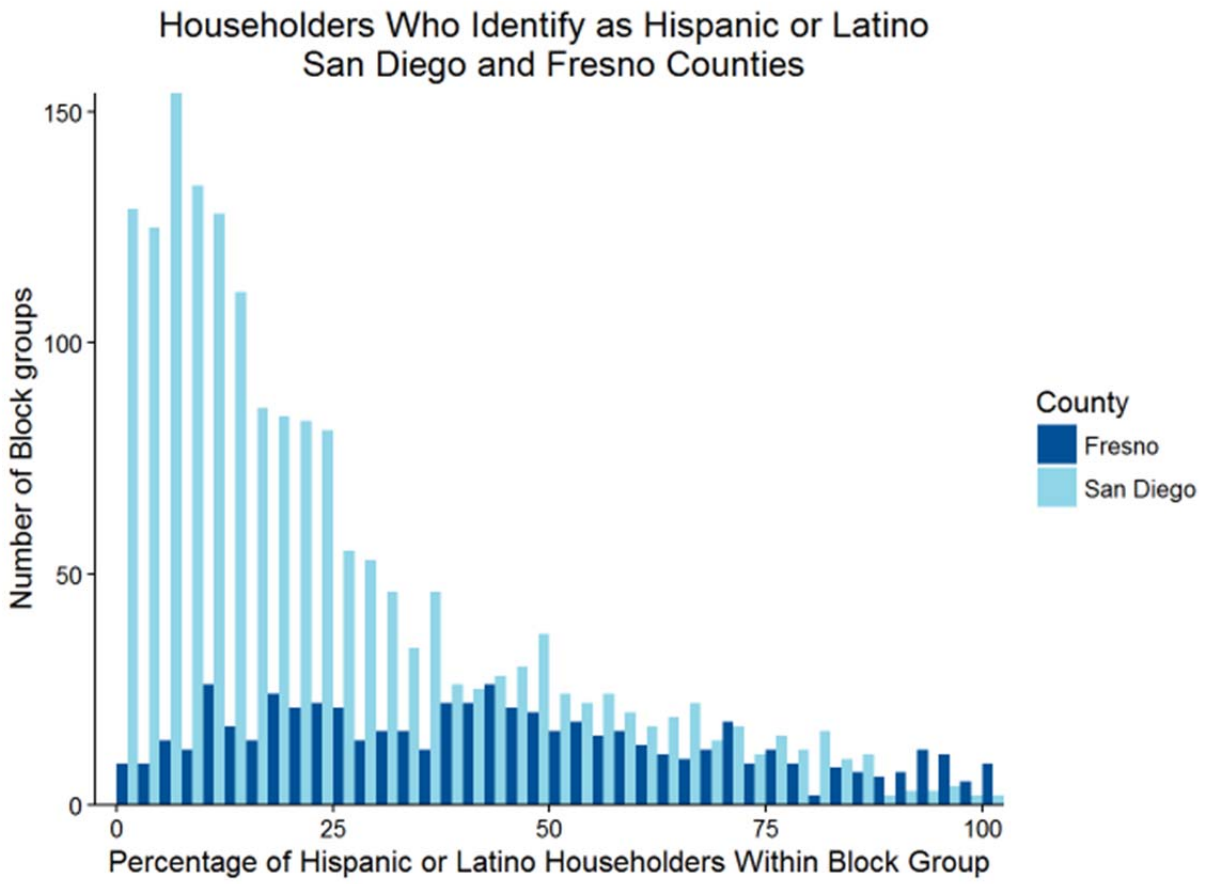


Figure 4.1: Histogram depicting the distribution of Hispanic or Latino populations at block group level across all block groups used in this study. The x-axis depicts the range of percentages of householders who identify as Hispanic or Latino within a given block group. The height of each bar is the count of block groups within San Diego and Fresno counties that have a Hispanic or Latino population of the size depicted on the x-axis.

Figures 4.2, 4.3, 4.4 (see following pages): Series of box and whiskers plots to display the variation in HVAC alteration saturation across a range of Hispanic or Latino communities (i.e., block groups) in coastal San Diego County (4.2), inland San Diego County (4.3) and Fresno County (4.3). The x-axis depicts the range of percentages of householders who identify as Hispanic or Latino within a given block group. Each observation in the box and whiskers plot is one block group. The blue area represents the 25%-75% range of the distribution (i.e., 2nd and 3rd quartile), the median is indicated by a solid black line in the center of the blue, and outliers are represented with a single black point.

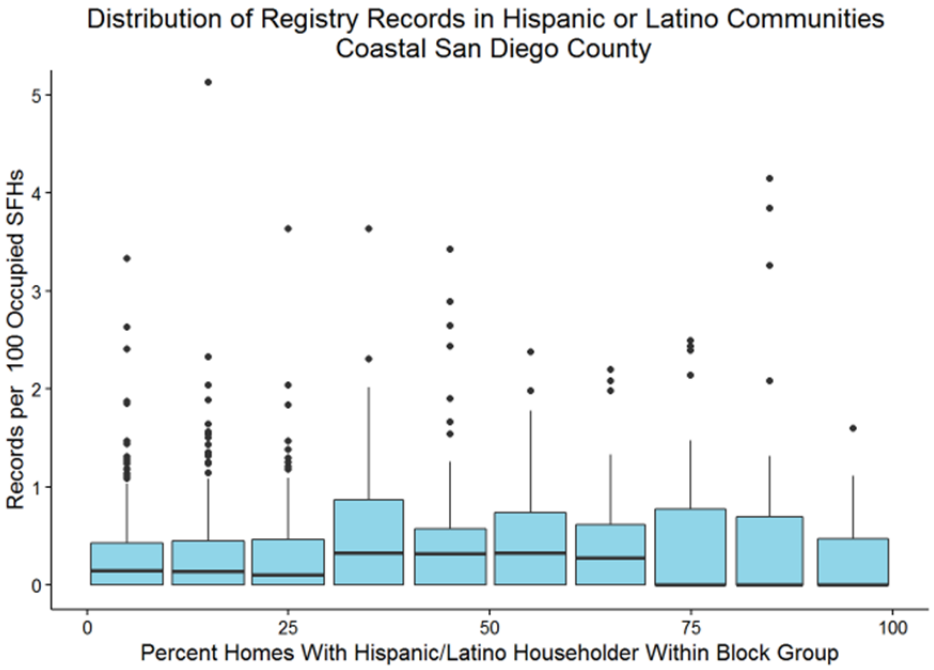


Figure 4.2

(The upper bound of the Y-axis has been reduced to 5 records per 100 SFHs to more clearly display the medians. This reduction prevents some outliers from being visible in the graphic.)

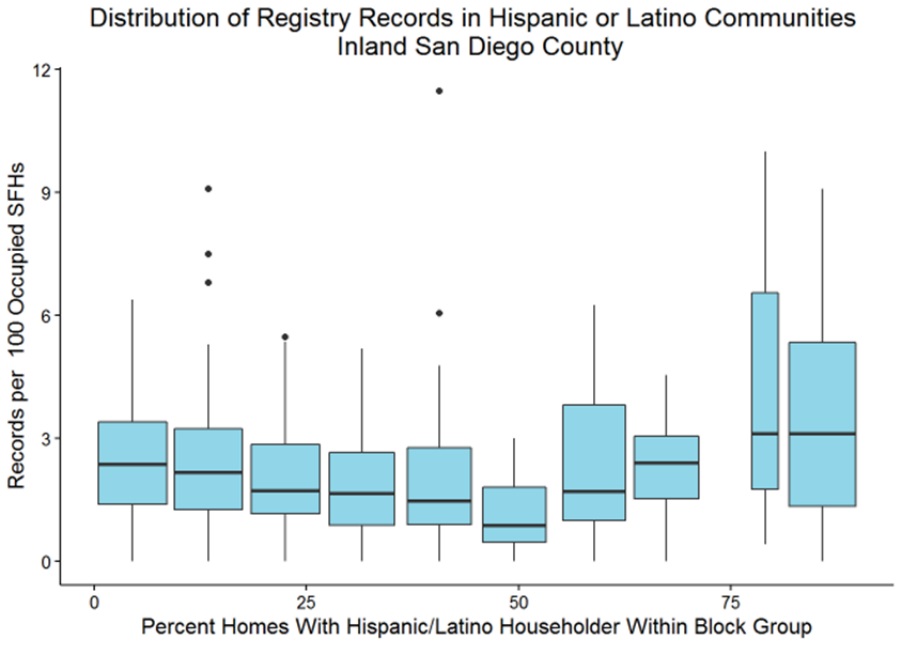


Figure 4.3

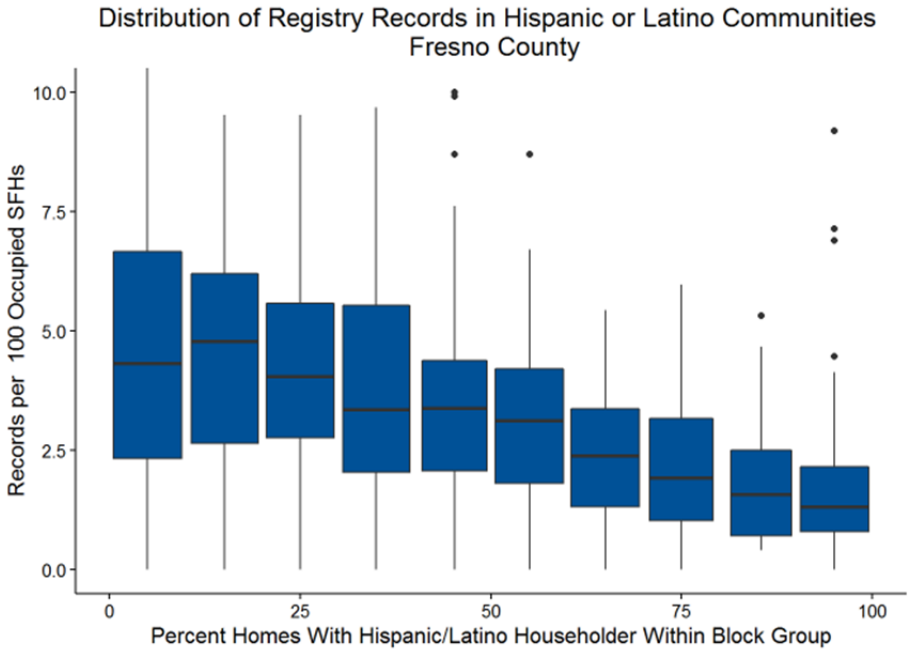
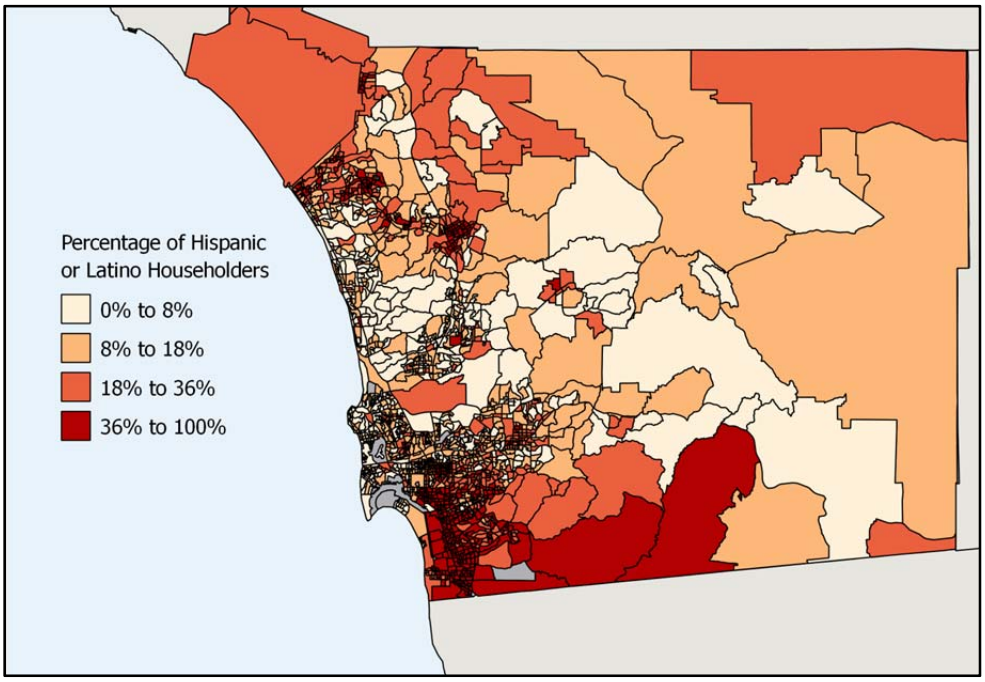
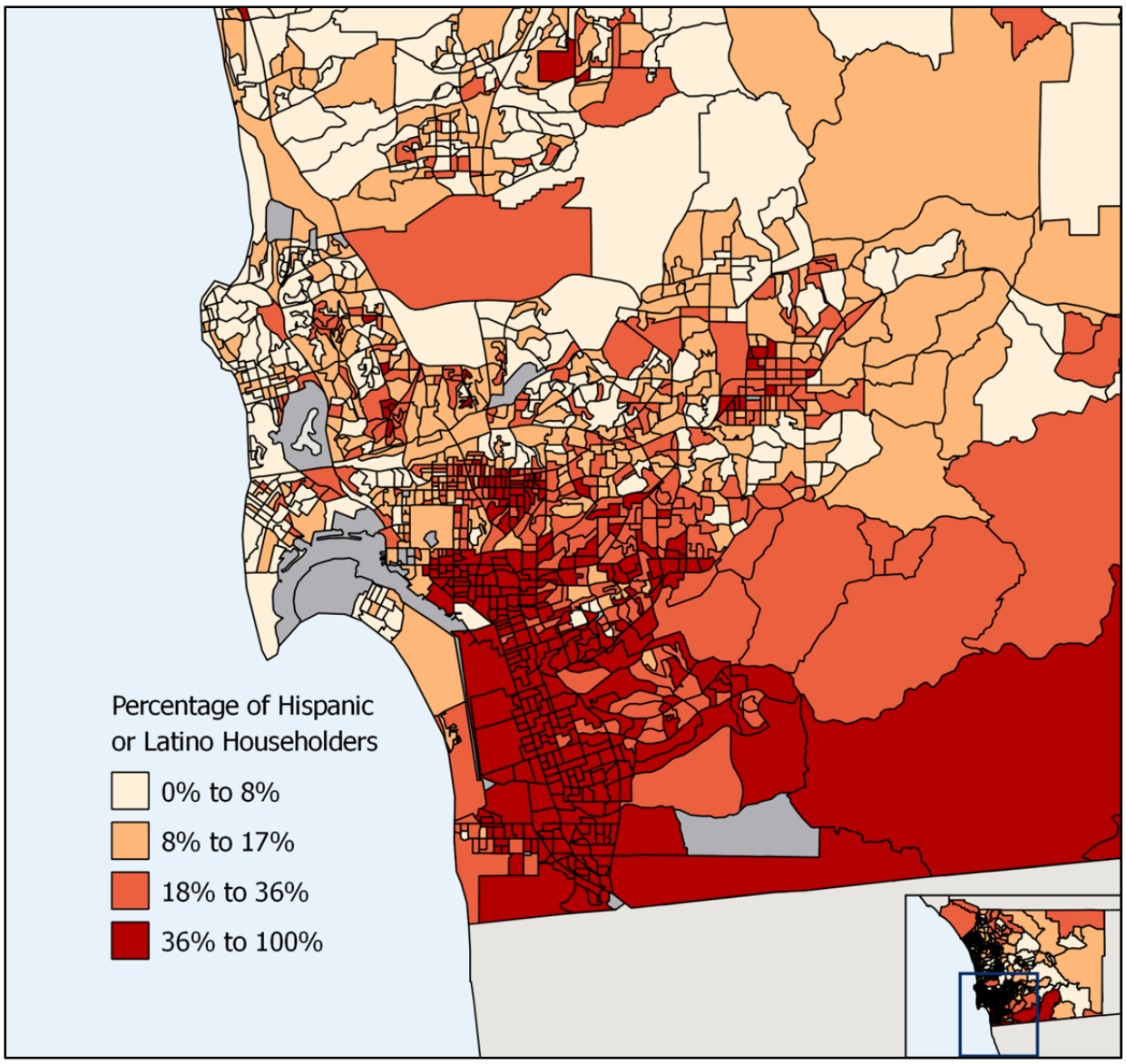


Figure 4.4

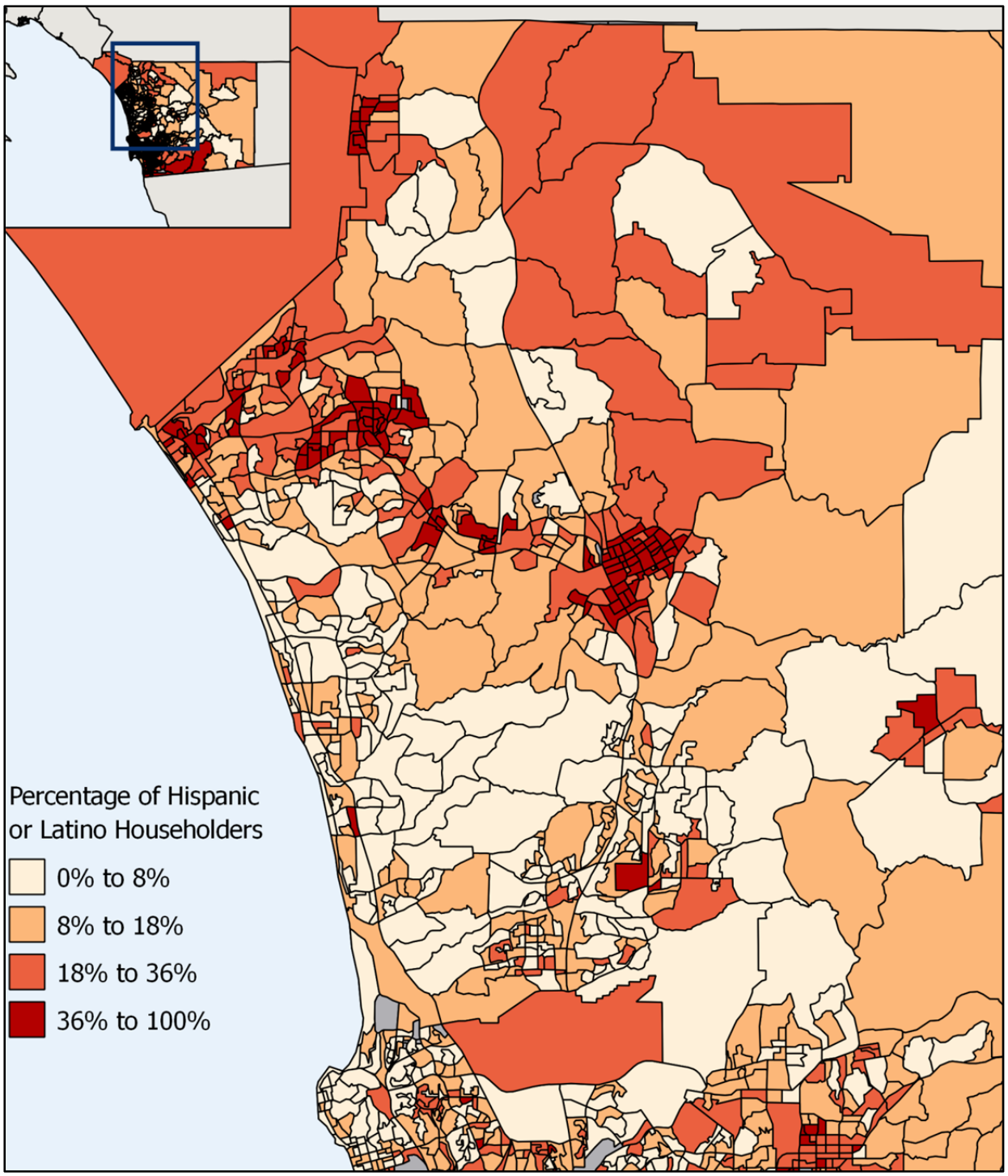
(The upper bound of the Y-axis has been reduced to 10 records per 100 SFHs to more clearly display the medians. This reduction prevents some outliers from being visible in the graphic.)



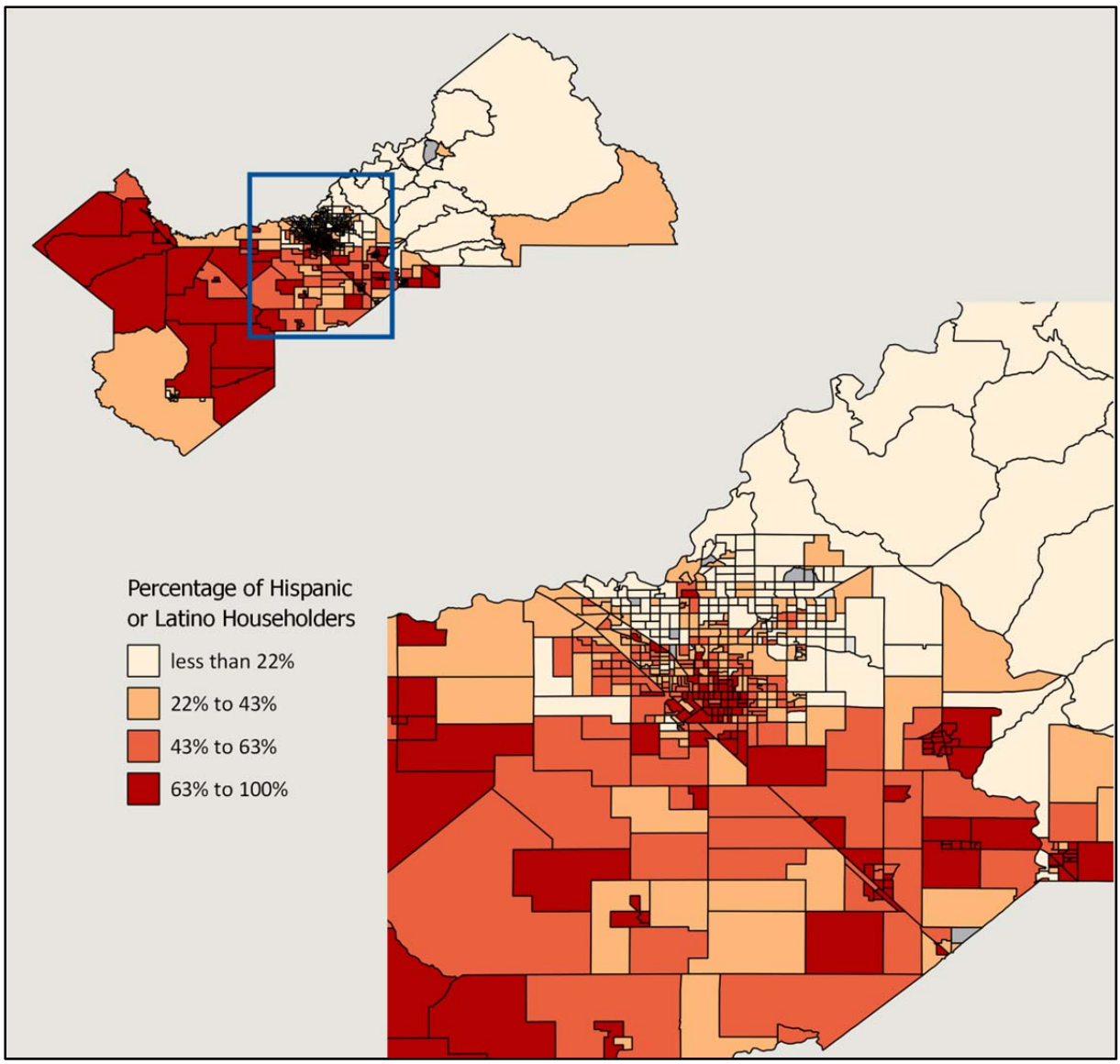
Map 4.1: Percentage of householders within each block group in San Diego County, which has a householder who identifies as Hispanic or Latino. Each color on the map represents one quartile in the distribution of Hispanic or Latino populations across all block groups.



Map 4.2: Percentage of householders within each block group in San Diego County, which has a householder who identifies as Hispanic or Latino. Each color on the map represents one quartile in the distribution of Hispanic or Latino populations across all block groups.



Map 4.3: Percentage of householders within each block group in San Diego County, which has a householder who identifies as Hispanic or Latino. Each color on the map represents one quartile in the distribution of Hispanic or Latino populations across all block groups.



Map 4.4: Percentage of householders within each block group in Fresno County, which has a householder who identifies as Hispanic or Latino. Each color on the map represents one quartile in the distribution of Hispanic or Latino populations across all block groups.

Home Value, Income and Build Year

Community level data on median home value, household income and build year of homes were collected to characterize the study areas. These data capture some of the key factors that may impact the decision-making process around energy use and efficiency measure adoption, and will be used as control variables in the model to isolate the impact that sociocultural factors may play.

The median home values in San Diego and Fresno counties have opposite patterns. In San Diego County, the most expensive homes are in the city center and near the coast, while in Fresno County the median home value in the city center is much lower than the surrounding areas. The home value data contains too many outliers to draw conclusions on the relationship between home value and registry record saturation. In Fresno County, there appears to be a slight positive relationship between home value and number of records in the CalCERTS registry. This aligns with our expectation that people with higher-value homes would be able to invest more in home upgrades.

The median household income and median home value are strongly correlated with each other. Median income was discarded in the regression modeling analysis described in Section V due to missing data in many block groups.

The build year of homes was reported in percentage of homes in 10-year increments and is not visually presented in this section due to the large number of plots and maps that would be necessary to display all data. However, this data is used in the following regression modeling analysis.

Figures 5.1, 5.2, 5.3 (see following pages): Series of box and whiskers plots to display the variation in HVAC alteration saturation across a range of median home values in coastal San Diego County (5.1), inland San Diego County (5.2) and Fresno County (5.3). The x-axis depicts the range of median home value within a given block group. Each observation in the box and whiskers plot is one block group. The blue area represents the 25%-75% range of the distribution (i.e., 2nd and 3rd quartile), the median is indicated by a solid black line in the center of the blue, and outliers are represented with a single black point.

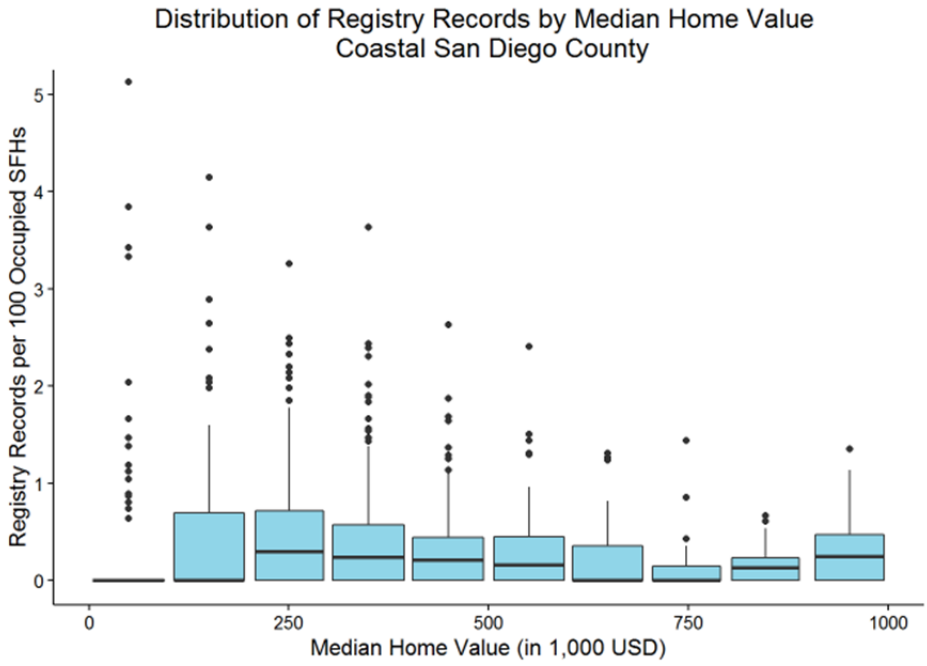


Figure 5.1

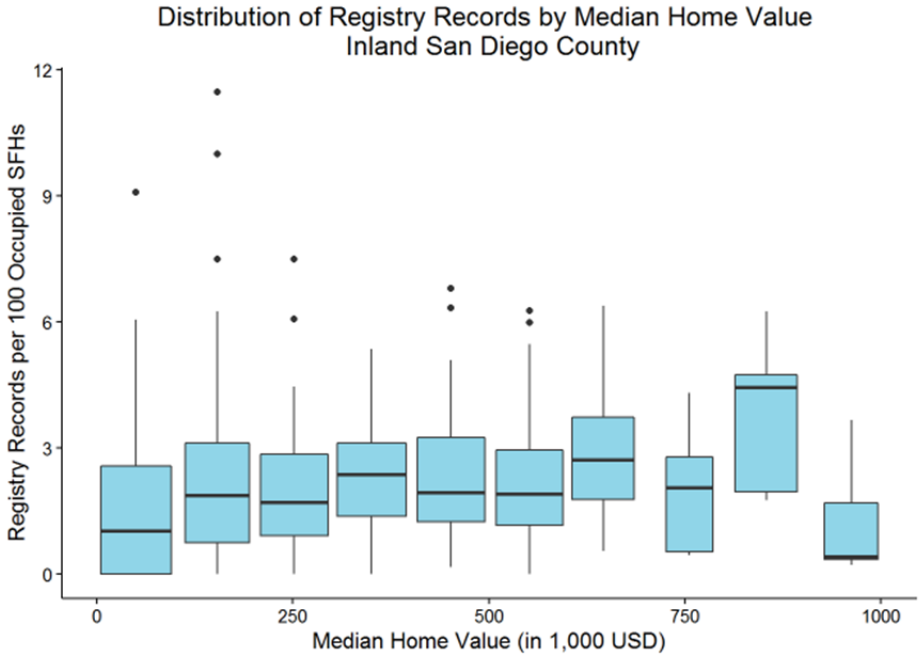


Figure 5.2

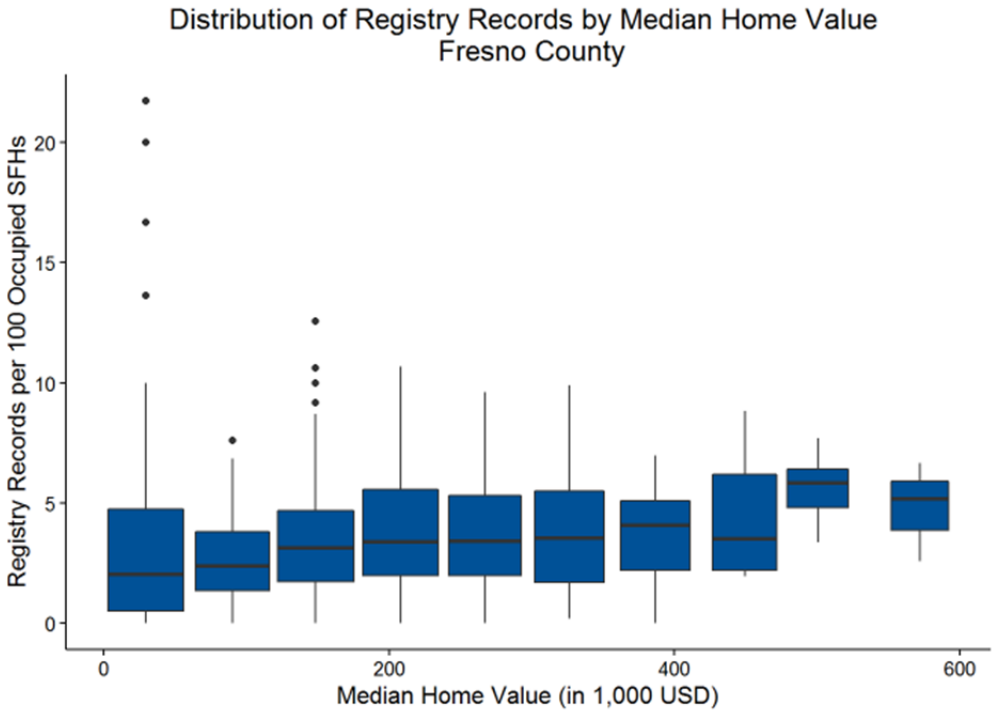
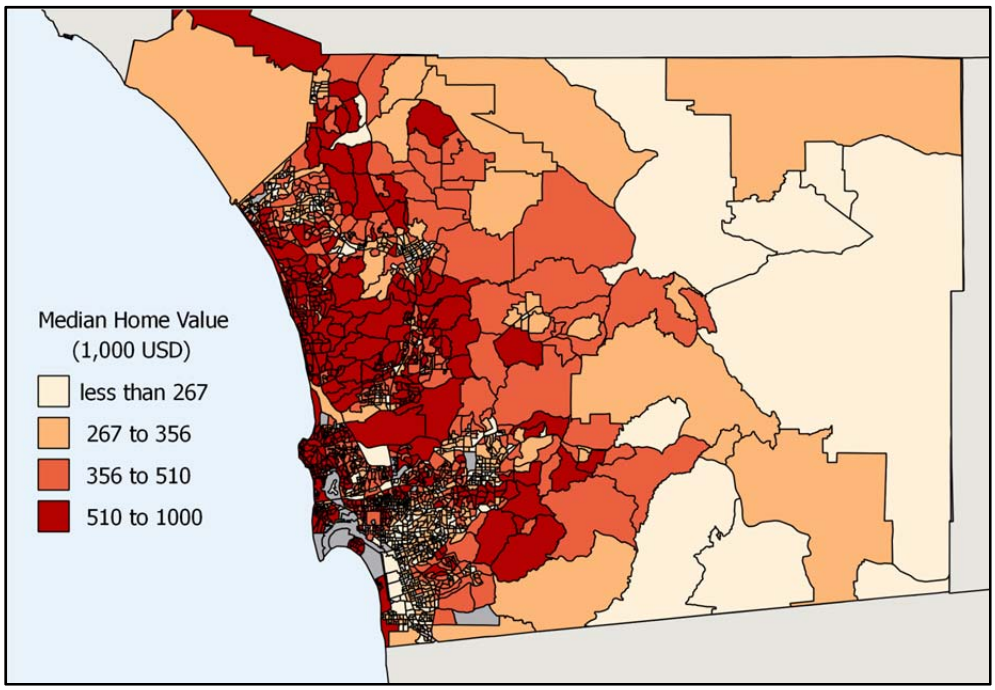
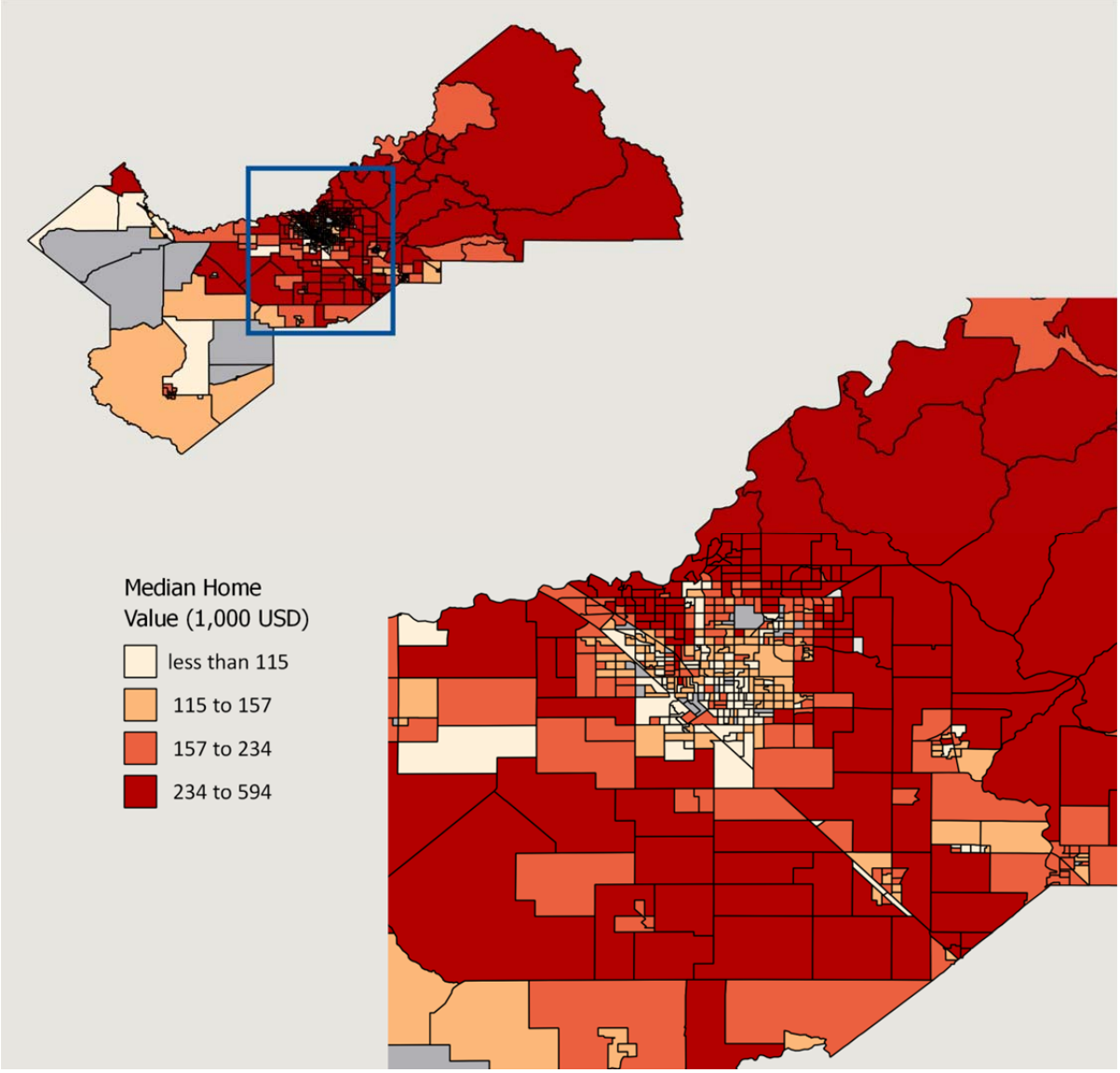


Figure 5.3



Map 5.1: Median home value (in 2014 inflation-adjusted 1,000 USD) within each block group in San Diego County. Each color on the map represents one quartile in the distribution of median home values across all block groups.



Map 5.2: Median home value (in 2014 inflation-adjusted 1,000 USD) within each block group in Fresno County. Each color on the map represents one quartile in the distribution of median home values across all block groups.

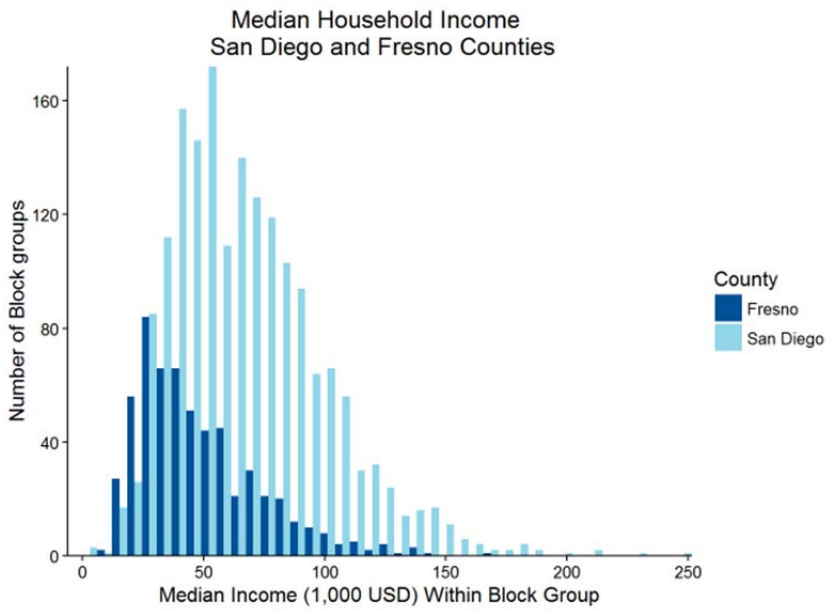


Figure 6.1: Histogram depicting the distribution of household incomes at block group level across all block groups used in this study. The x-axis depicts the range of household incomes (in 1,000, 2014 inflation-adjusted USD) across all block groups. The height of each bar is the count of block groups that have a median income corresponding to the value on the x-axis.

Figures 6.2, 6.3, 6.4: Series of box and whiskers plots to display the variation in HVAC alteration saturation across a range of median household incomes (in 2014 inflation-adjusted dollars) in coastal San Diego County (6.2), inland San Diego County (6.3) and Fresno County (6.4). The x-axis depicts the range of median incomes within a given block group. Each observation in the box and whiskers plot is one block group. The blue area represents the 25%-75% range of the distribution (i.e., 2nd and 3rd quartile), the median is indicated by a solid black line in the center of the blue, and outliers are represented with a single black point.

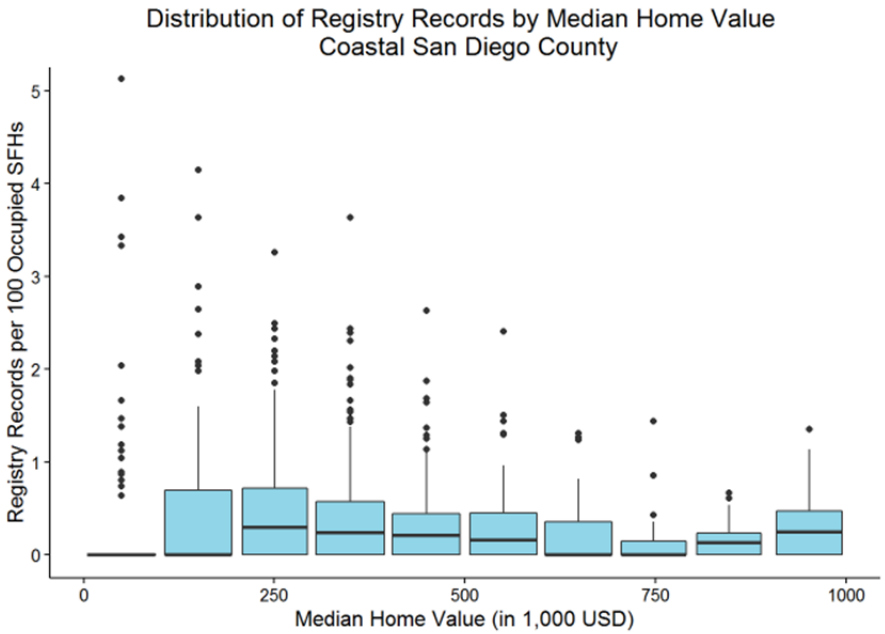


Figure 6.2

(The upper bound of the Y-axis has been reduced to 5 records per 100 SFHs to more clearly display the medians. This reduction prevents some outliers from being visible in the graphic.)

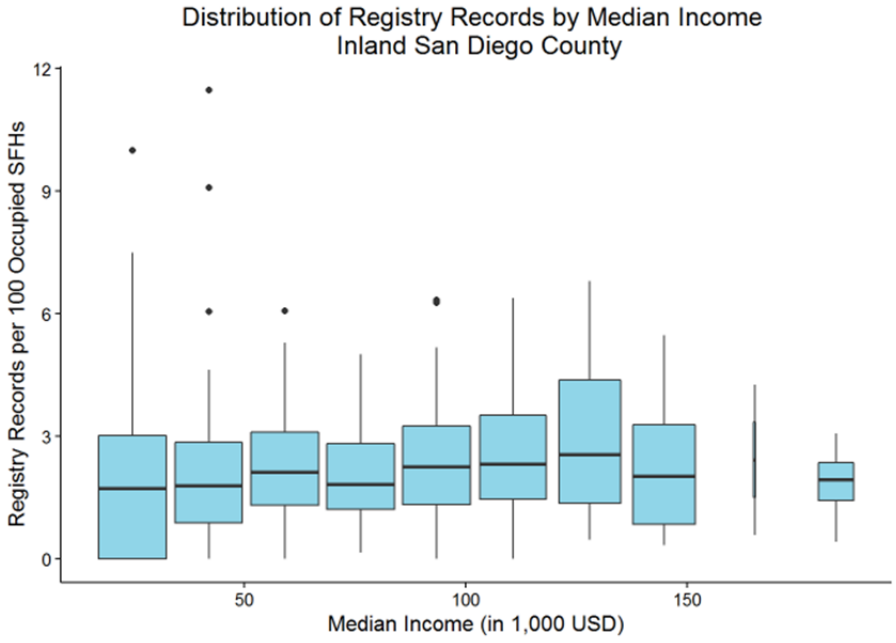


Figure 6.2

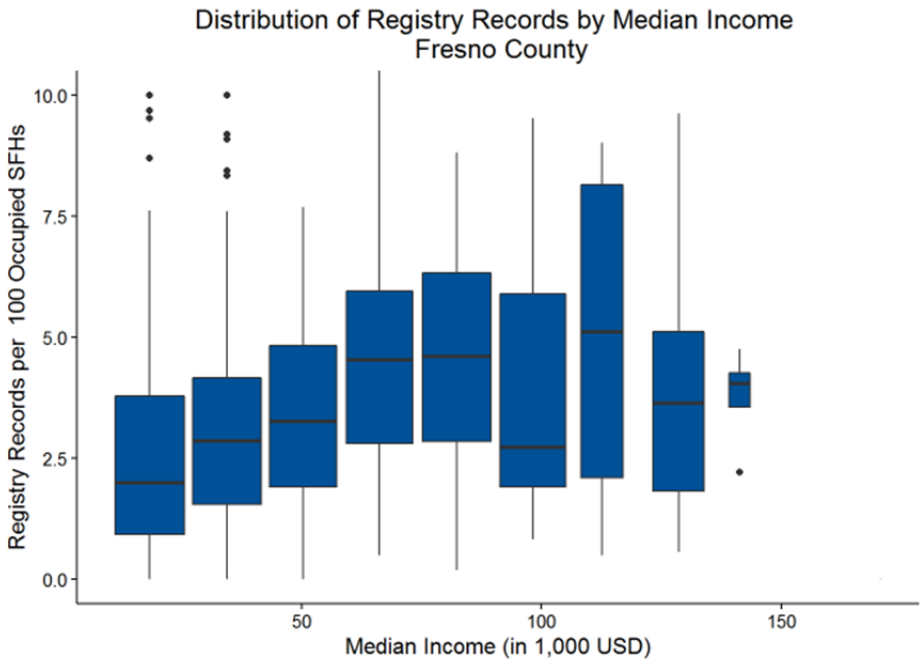
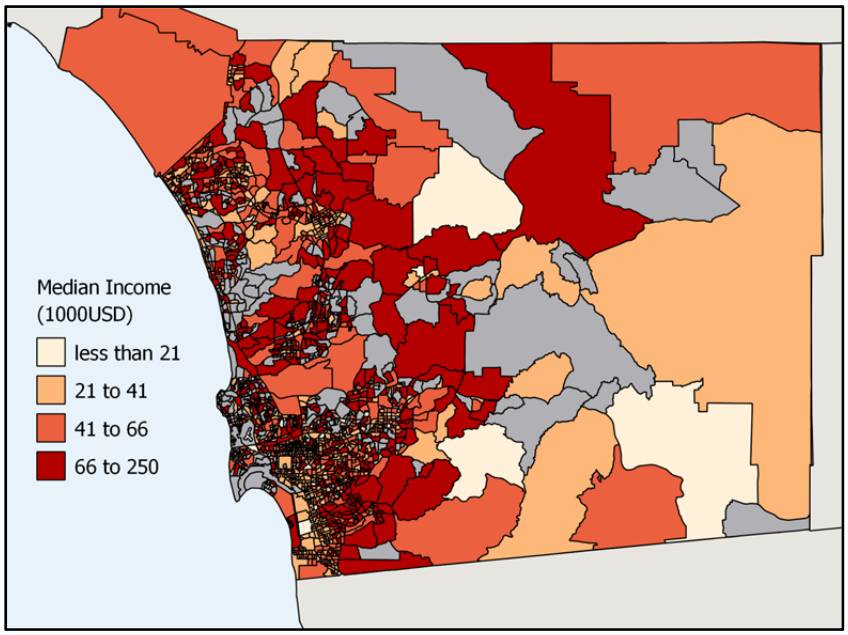
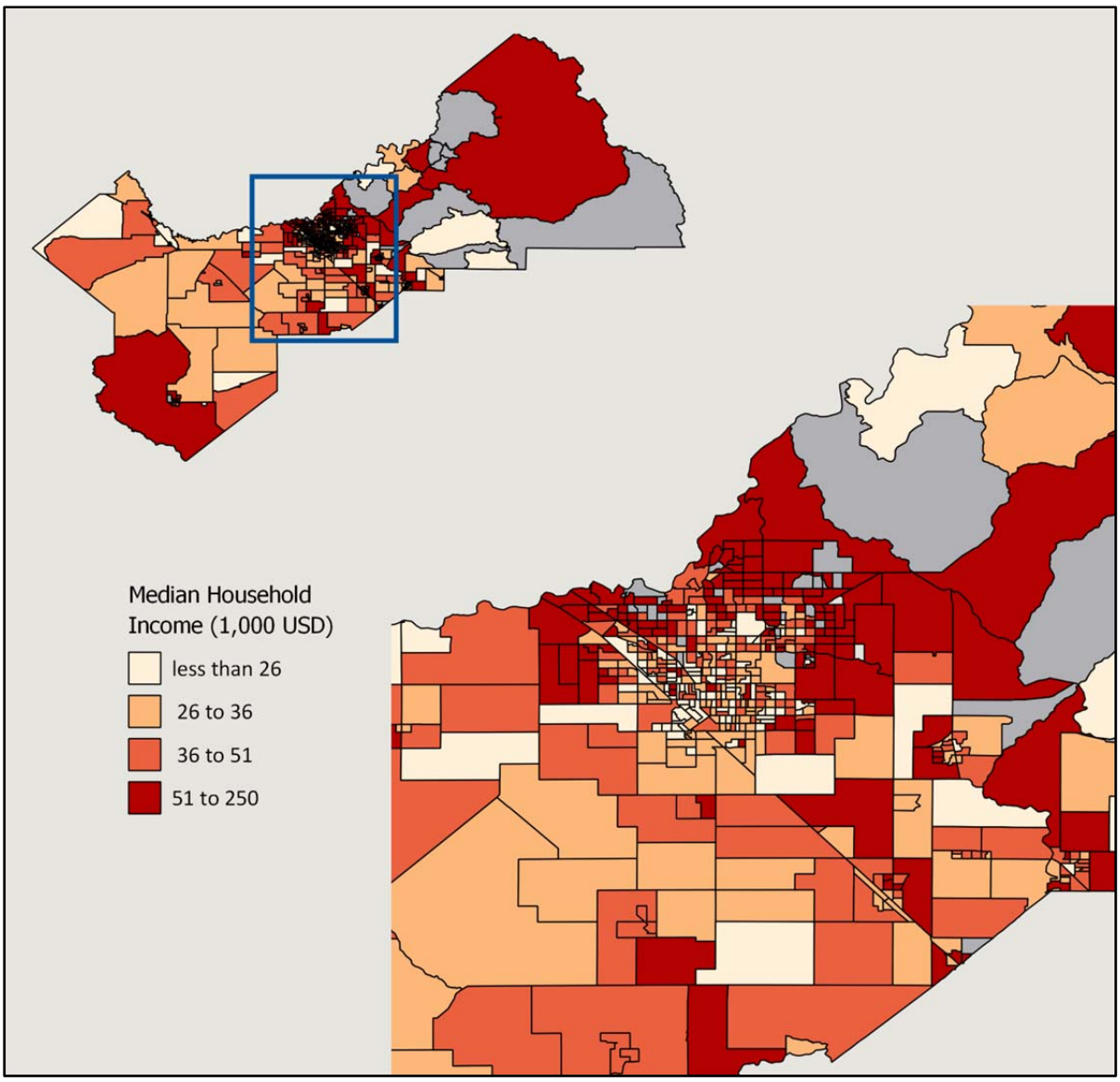


Figure 6.4

(The upper bound of the Y-axis has been reduced to 10 records per 100 SFHs to more clearly display the medians. This reduction prevents some outliers from being visible in the graphic.)



Map 6.1: Median income (in 2014 inflation-adjusted 1,000 USD) within each block group in San Diego County. Each color on the map represents one quartile in the distribution of median home values across all block groups



Map 6.2: Median income (in 2014 inflation-adjusted 1,000 USD) within each block group in Fresno County. Each color on the map represents one quartile in the distribution of median home values across all block groups

Political Party Affiliation

Previous research has investigated how political leanings influence energy-related behavior and attitudes (see literature review). To understand the scale, distribution, and direction of community level political leanings, block group level data on the Republican Party advantage from the 2014 gubernatorial election was included in this analysis.

Political leanings in San Diego (as measured by the party advantage variable) have a high degree of spatial heterogeneity, with concentrations of Democratic-leaning voters located in the City of San Diego's metropolitan core, the metro areas of inland valley cities, and along the Pacific coastal zone, while large swaths of the inland valleys and unincorporated areas lean more heavily Republican. Fresno shows a similar distribution of political leaning, with urban areas trending more Democratic and suburban/rural areas (particularly those to the northeast) trending more heavily Republican.

Figures 7.1, 7.2, 7.3 (see following pages): Series of box and whiskers plots to display the variation in HVAC alteration saturation across communities with a range of political orientation in coastal San Diego County (7.1), inland San Diego County (7.2) and Fresno County (7.3). The x-axis depicts the range republican advantage within each box group. Republican advantage is measured as the percentage point difference between votes cast for the republican candidate and votes cast for the democratic candidate. Negative values indicate a democratic advantage in that block group. Each observation in the box and whiskers plot is one block group. The blue area represents the 25%-75% range of the distribution (i.e., 2nd and 3rd quartile), the median is indicated by a solid black line in the center of the blue, and outliers are represented with a single black point.

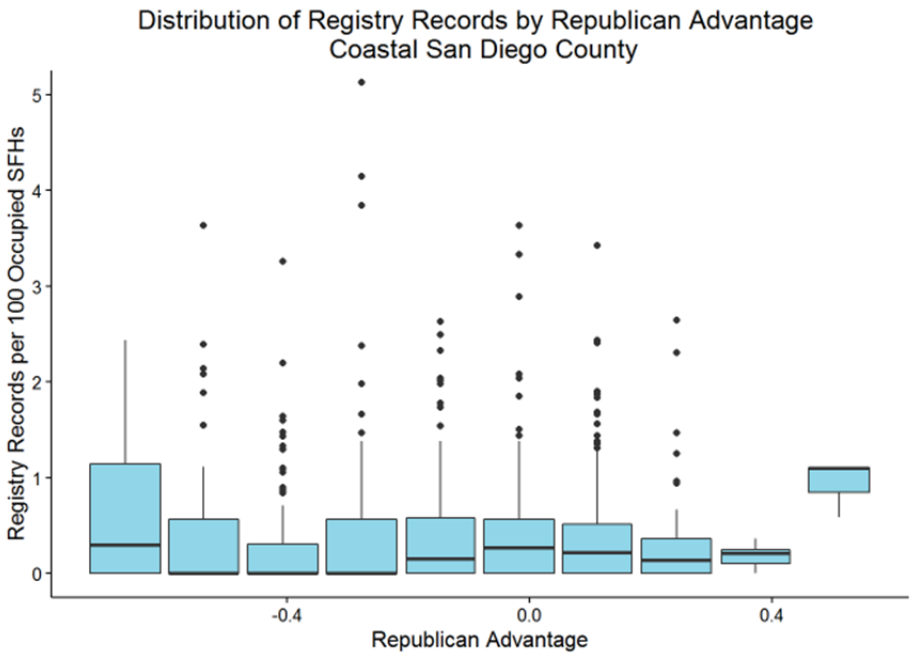


Figure 7.1

(The upper bound of the Y-axis has been reduced to 5 records per 100 SFHs to more clearly display the medians. This reduction prevents some outliers from being visible in the graphic.)

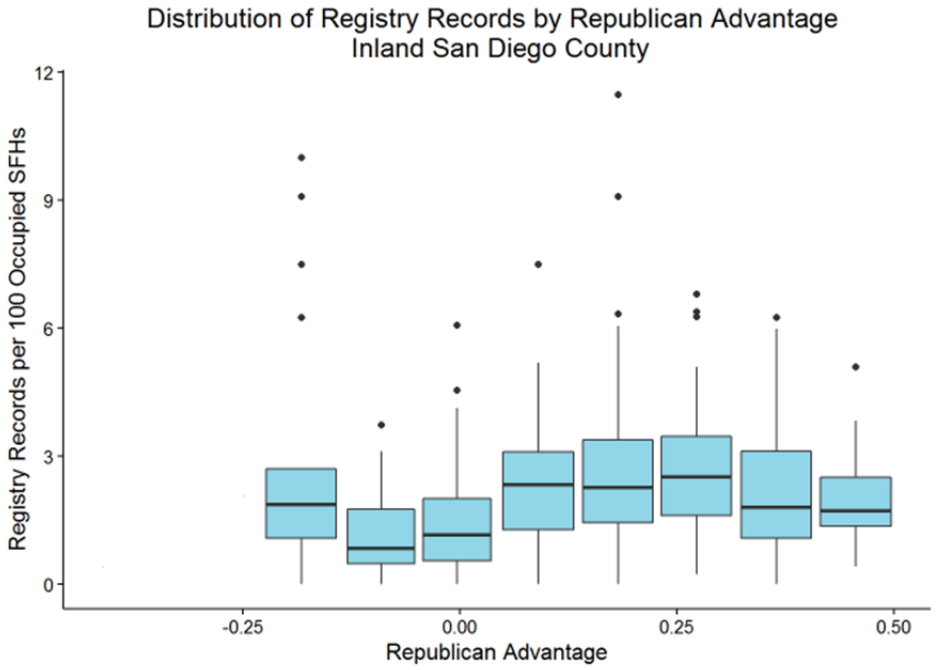


Figure 7.2

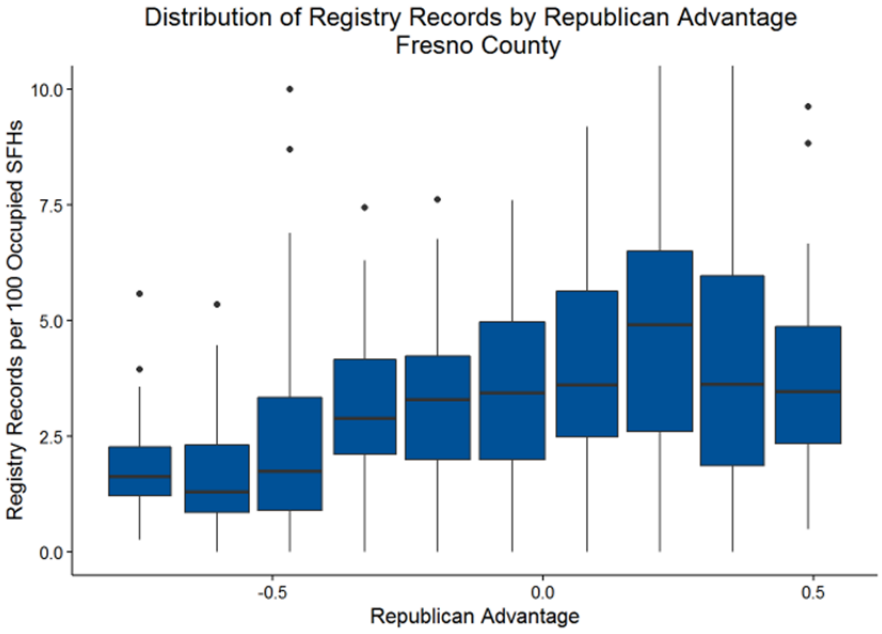
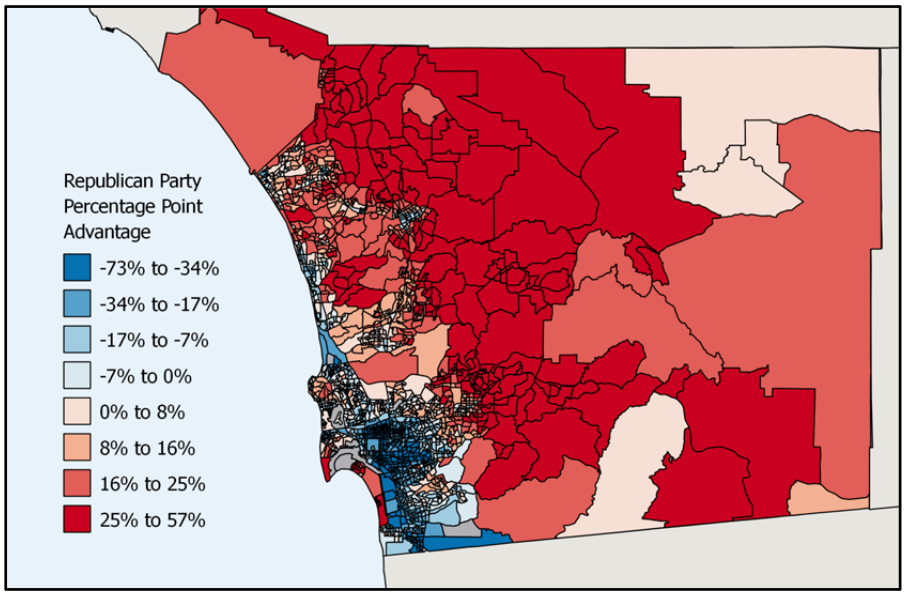
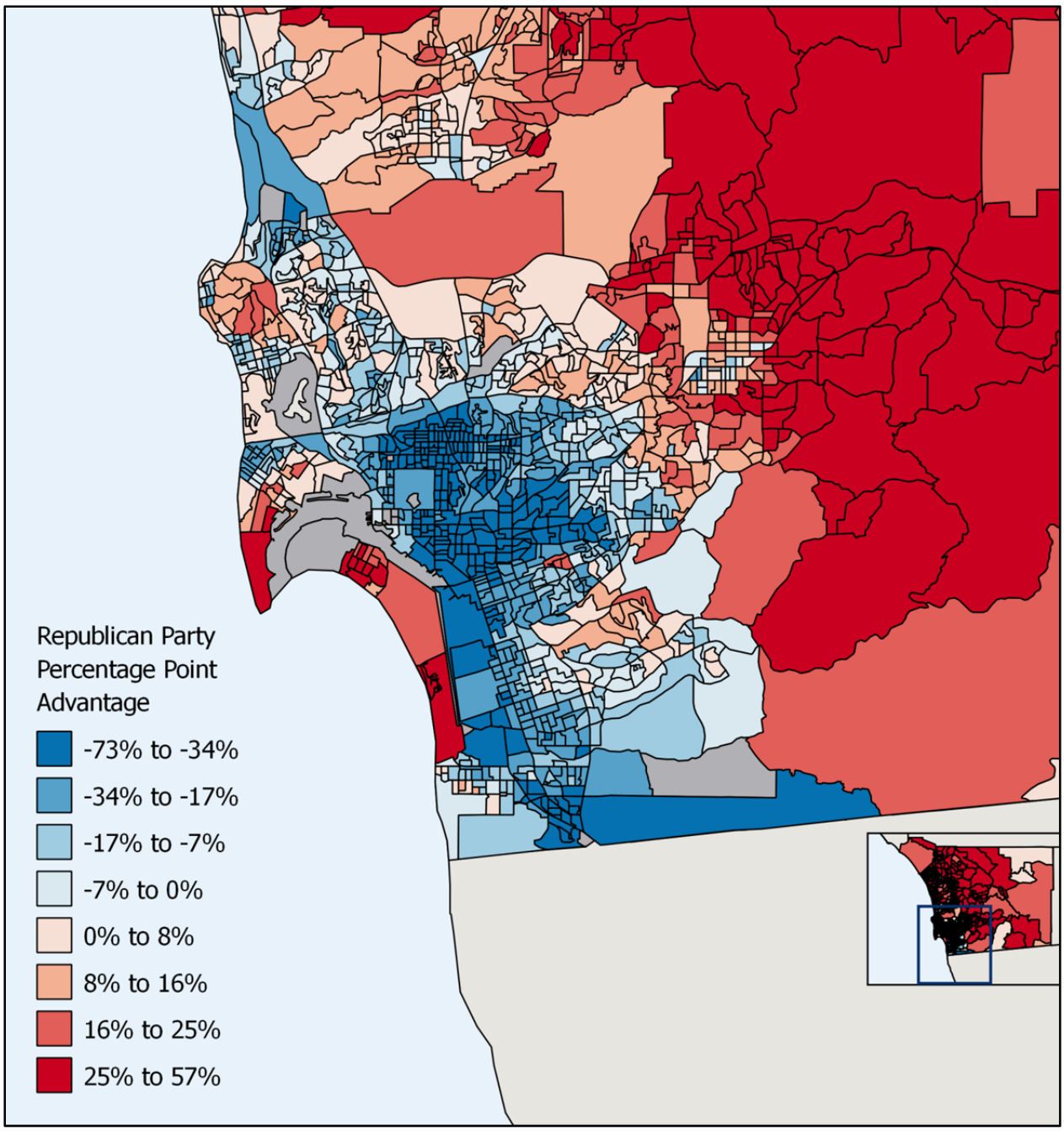


Figure 7.3

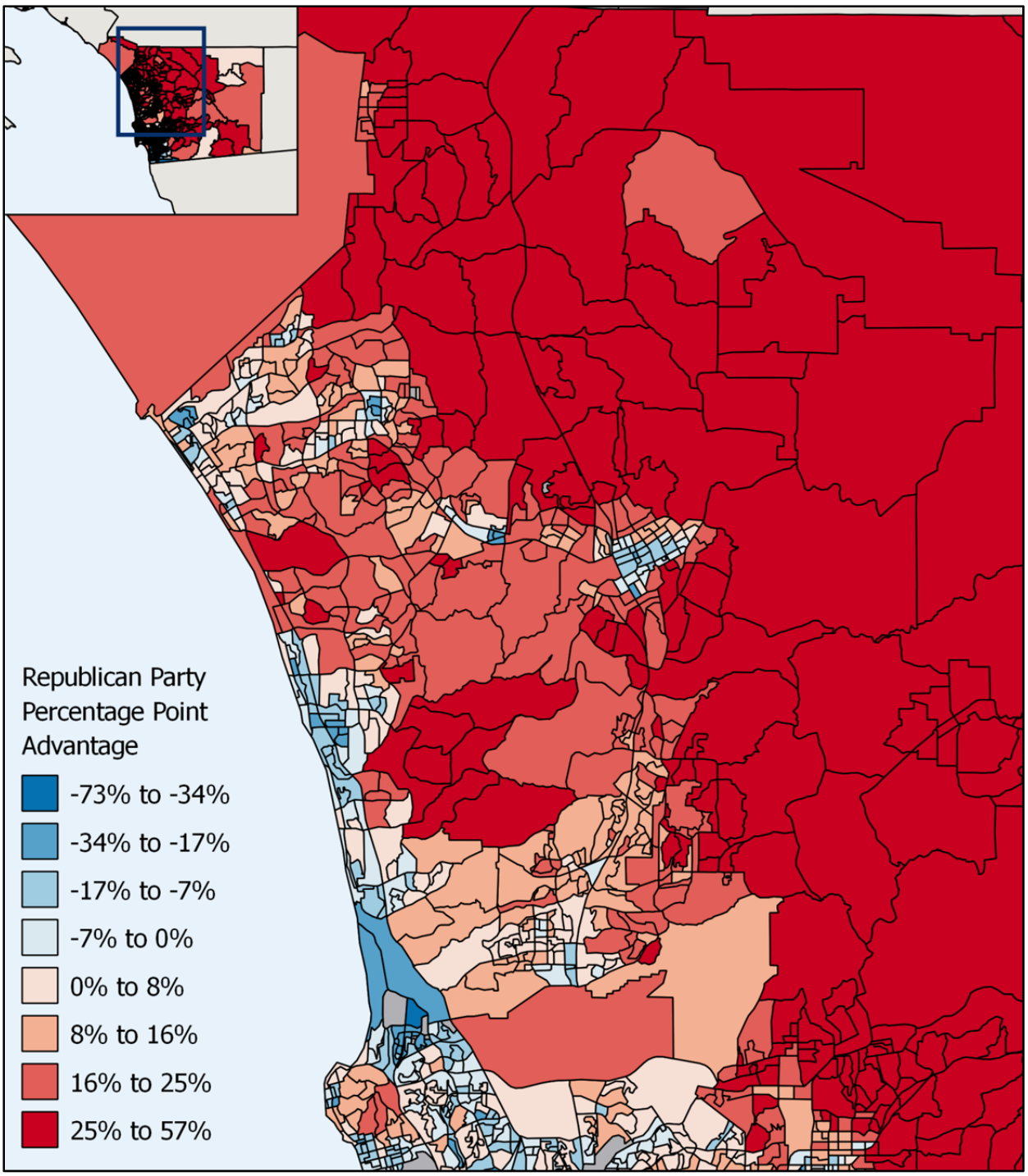
The upper bound of the Y-axis has been reduced to 5 records per 100 SFHs to more clearly display the medians. This reduction prevents some outliers from being visible in the graphic.



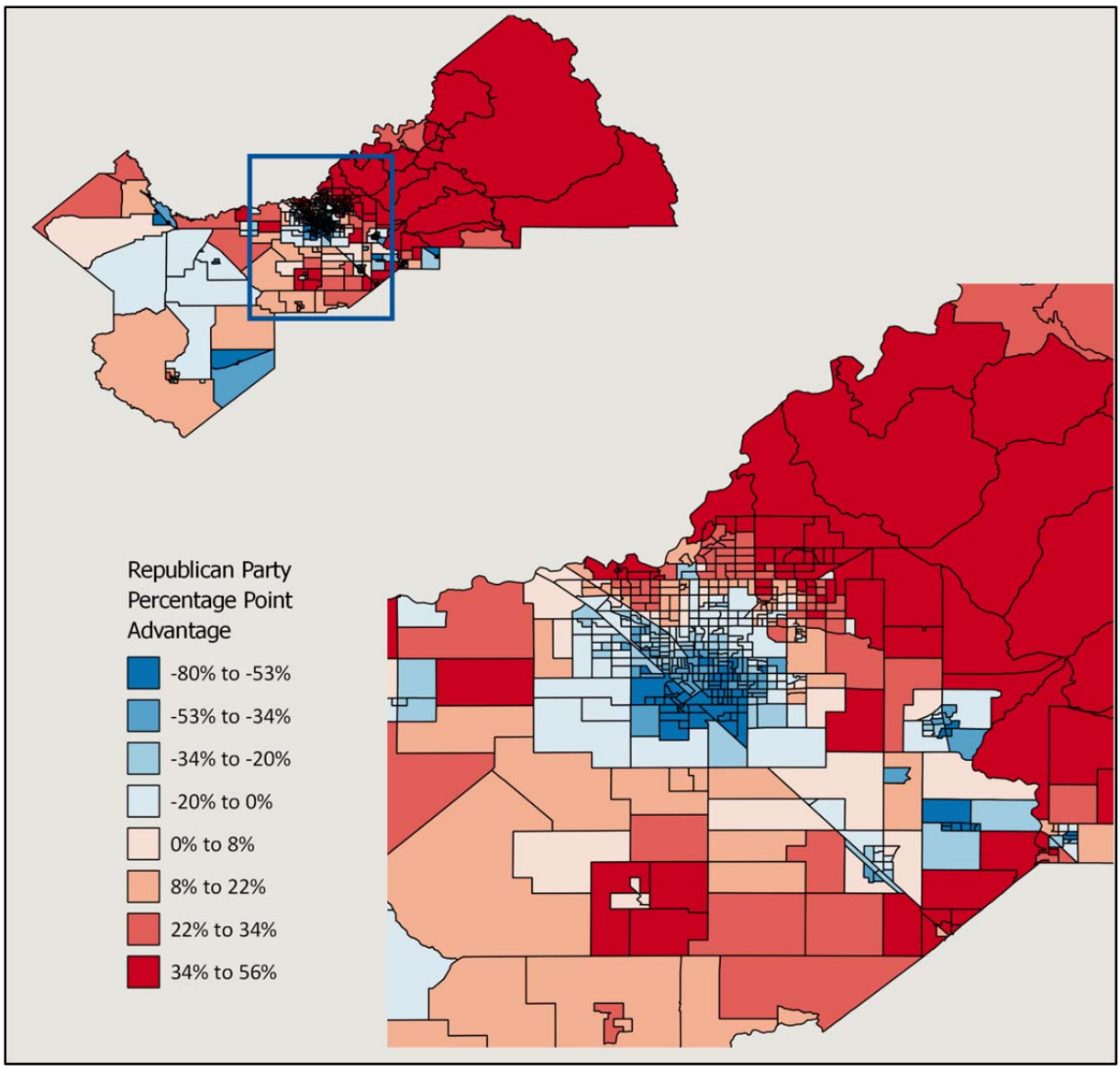
Map 7.1: Republican party percentage point advantage (in 2014 California gubernatorial election) within each block group in San Diego County. Republican advantage is measured as the percentage point difference between votes cast for the Republican candidate and votes cast for the Democratic candidate. Negative values indicate a Democratic advantage in that block group.



Map 7.2: Republican party percentage point advantage (in 2014 California gubernatorial election) within each block group in southwestern San Diego County. Republican advantage is measured as the percentage point difference between votes cast for the Republican candidate and votes cast for the Democratic candidate. Negative values indicate a Democratic advantage in that block group.



Map 7.3: Republican party percentage point advantage (in 2014 California gubernatorial election) within each block group in northwestern San Diego County. Republican advantage is measured as the percentage point difference between votes cast for the Republican candidate and votes cast for the Democratic candidate. Negative values indicate a Democratic advantage in that block group.



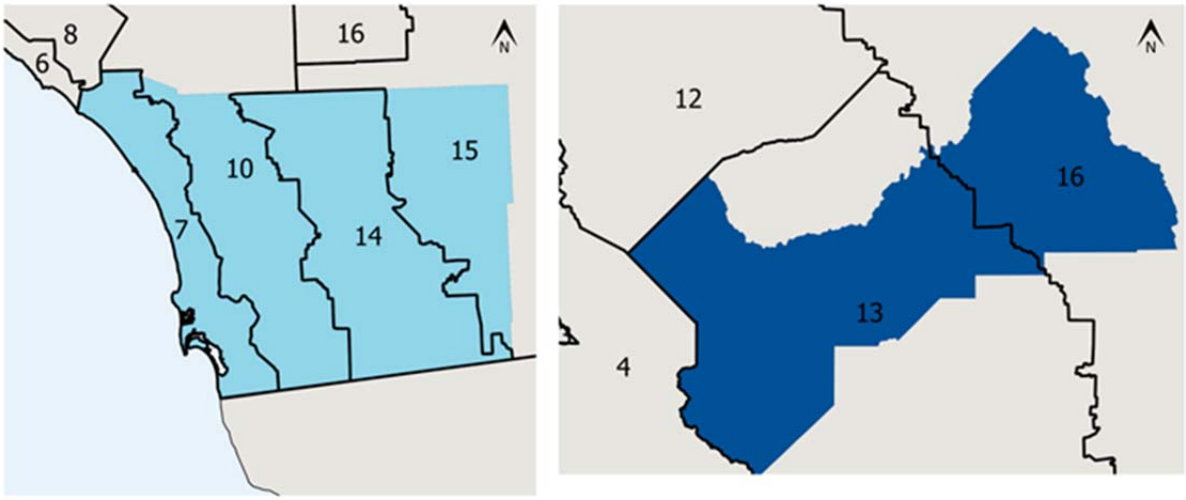
Map 7.4: Republican party percentage point advantage (in 2014 California gubernatorial election) within each block group in Fresno County. Republican advantage is measured as the percentage point difference between votes cast for the Republican candidate and votes cast for the Democratic candidate. Negative values indicate a Democratic advantage in that block group.

Climate Zones

Title 24, Part 6 requirements for energy efficiency levels and field verifications depend on the climate zone in which the measure is installed. One metric used to define the Energy Commission climate zones are the range of heating and cooling degree days in a region. Cooling and heating degree days (CDDs and HDDs) are a measure of energy consumption required to heat or cool a home to a comfortable temperature. For example, a single day with a temperature 10 degrees above a determined comfortable (e.g., 65°F) would have a CDD of 10. The annual CDD is the summation of the total CDDs in one year.

In San Diego County, the climate is more temperate near the ocean, in zone 7. Moving east through San Diego County, the dry climate and hilly topography contribute to more extreme variations in temperature. A home in climate zone 7, where the average annual CDDs range from 505 to 1,110, requires less energy to stay comfortable than a similarly constructed home in climate zone 10, where the average annual CDDs range from 1,714 to 1,937.

Climate zone 13 in Fresno County has a similar climate to zone 10, but with slightly hotter summers and cooler winters. Climate zone 16 is a mountainous region with a semi-arid climate and much colder winters and milder summers than the other climate zones analyzed in this study. Climate zone 16 has a small population, and the CalCERTS data records include few forms from this area.



Map 8: California building climate zones in San Diego (left) and Fresno (right) counties.

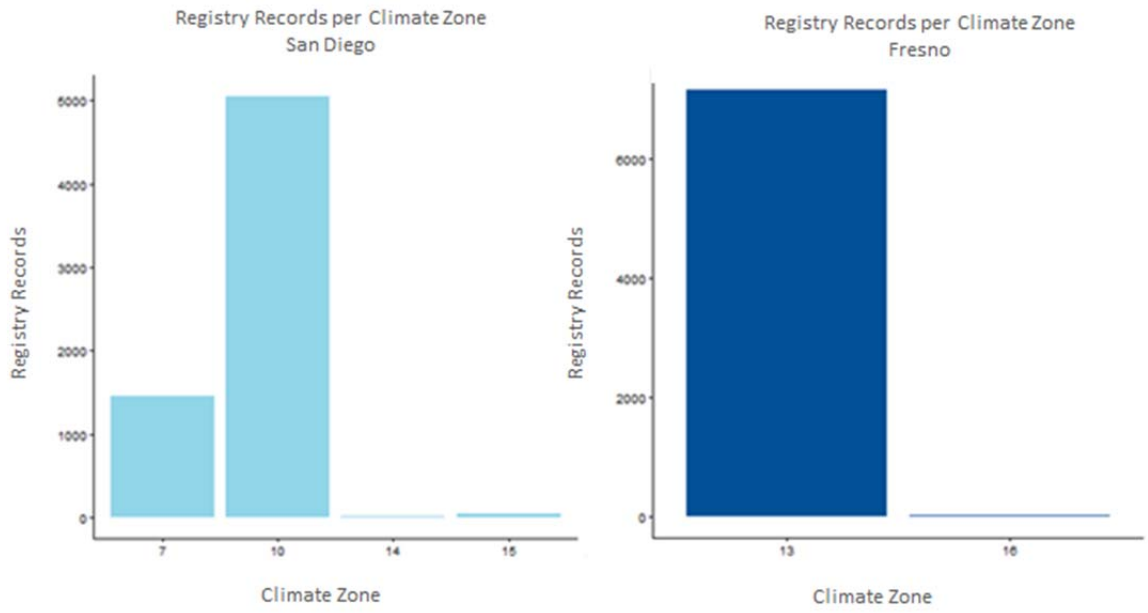


Figure 8: Distribution of homes with records in the CalCERTS registry across climate zones in San Diego and Fresno counties.

V. Regression Modeling

Model Specification

To gauge the relationship of sociocultural factors with the propensity of homeowners to perform a mechanical system upgrade, a number of regression models were run. The first model examined if the proportion of the Spanish-speaking community has a measurable impact on the upgrade activity. The second set of models tested the impact that the size of a Hispanic or Latino population has on upgrade activity. Both models were tested at the block group level.

The data was modeled using a binomial regression with a log link function (logistic regression) at the block group resolution. The response variable was formulated as a $\boldsymbol{n\ x\ 2}$ matrix, where $\boldsymbol{n}$ is the number of block groups in each study region (e.g., Fresno or San Diego county). The first column lists the number of homes with an associated entry in the CalCERTS registry and the second column lists the number of homes associated with an entry in the CalCERTS registry subtracted from the total number of occupied single-family homes in that block group. This is analogous to a binomial regression with the vector of successes and failures as the response. In this model the successes and failures have been aggregated to the block group level because this is the resolution of the predictor variables.

The predictor variables listed below were chosen because of their suspected relationship with householder's likelihood to replace an air conditioner or furnace.

Variable	Type	
Number of CalCERTS records/ Occupied single family homes	Dependent	Y
Percentage of Spanish speakers	Independent	X_1
Percentage of Hispanic or Latino homeowners	Independent	X_1
Median home value	Independent	$\mathbf{X_2} \ldots \mathbf{X_n}$
Republican advantage	Independent	X_3
Age of home	Independent	$X_4 \ldots X_n$

Table 3: Variables used in regression modeling.

Results

The results from the regression modeling are displayed in the following tables. All variables in all models were shown to have a statistically significant effect on the response variable. The coefficients given for each predictor can be interpreted as the log odds change in the response given a 1 unit change in the predictor variables. The exponentiated coefficient gives the multiplicative change in the odds ratio of the response that results from a 1 percentage point change in the predictor, all other effects held constant. For example, if the current number of registry records in a block group was 5 and that block group contained 10 occupied SFHs, the odds ratio would be 0.5. A 1 percentage point change in the number of Spanish speakers in that block group, using the coefficient from San Diego Coastal Model A, would result in and odds ratio of 0.5*1.01 = 0.505.

The effect of a 1% increase in the percentage of a block group that is Spanish speaking and the percentage of a block group that is Hispanic or Latino is nearly identical with each region that the model was run. In the San Diego coastal region, this change would result in a one percent increase in the log odds of the response. In the San Diego inland region, this increase would result in a less than one percent decrease in the response. In Fresno, this change would result in a one percent decrease in the response (all other effects held constant).

The results of this modeling exercise indicate that we can reject the null hypotheses that the percentage of householders in a block group that are Hispanic or Latino or Spanish speaking has no correlation with the number of homes with records associated with an HVAC replacement in the CalCERTS registry. However, the small size of the coefficients for these sociocultural variables, as well as the differing results from one region to another, indicate that the elucidation of this relationship will require further investigation.

Models with Hispanic/Latino Response	San Diego Coastal (A)		San Diego Inland (A)	
	$\widehat{\beta_X}$ (Std. Error)	$e^{\widehat{\beta_X}}$	$\widehat{\beta_X}$ (Std. Error)	$e^{\widehat{\beta_X}}$
Intercept	-5.778*** (0.0730)	0.00	-3.392*** (0.0518)	0.03
% Hispanic/Latino	0.008*** (0.0016)	1.01	-0.005*** (0.0013)	1.00
Republican Advantage	0.005** (0.0017)	1.01	0.003** (0.0013)	1.00
% Build Year > 2000	-0.011*** (0.0021)	0.99	-0.015*** (0.0009)	0.98
% Build Year 1990-2000	-0.011*** (0.0025)	0.99	-0.008*** (0.0010)	0.99
% Build Year 1960-1970	0.005*** (0.0017)	1.00	0.006*** (0.0013)	1.01
% Build Year 1930-1940	-0.018*** (0.0043)	0.98	-0.024*** (0.0048)	0.98
% Build Year < 1930	-0.010*** (0.0036)	0.99	-0.029*** (0.0056)	0.97
% Home Value 100-300k	0.006*** (0.0013)	1.01	-0.003*** (0.0009)	1.00
% Home Value > 100k	0.008*** (0.0028)	1.01	-0.005*** (0.0016)	0.99

Table 4.1: Results from regression models in which percentage Hispanic/Latino in each block group is an independent variable and data is drawn from San Diego County CalCERTS records.
Stars represent the significance of coefficients: * $P \leq 0.05$ ** $P \leq 0.01$ *** $P \leq 0.001$

Models with Spanish Language Response	San Diego Coastal (B)		San Diego Inland (B)	
	$\widehat{\beta_X}$ (Std. Error)	$e^{\widehat{\beta_X}}$	$\widehat{\beta_X}$ (Std. Error)	$e^{\widehat{\beta_X}}$
Intercept	-5.7806*** (0.073)	0.00	-3.4036*** (0.0518)	0.03
% Hispanic/Latino	0.0083*** (0.0016)	1.01	-0.0040*** (0.0014)	1.00
Republican Advantage	0.0055** (0.0018)	1.01	0.0029* (0.0013)	1.00
% Build Year > 2000	-0.0108*** (0.0021)	0.99	-0.0156*** (0.0009)	0.98
% Build Year 1990-2000	-0.0113*** (0.0025)	0.99	-0.0082*** (0.0010)	0.99
% Build Year 1960-1970	0.0048*** (0.0017)	1.00	0.0062*** (0.0013)	1.01
% Build Year 1930-1940	-0.0173*** (0.0042)	0.98	-0.0235*** (0.0048)	0.98
% Build Year < 1930	-0.0102*** (0.0036)	0.99	-0.0285*** (0.0056)	0.97
% Home Value 100-300k	0.0057*** (0.0013)	1.01	-0.0032*** (0.0009)	1.00
% Home Value > 100k	0.0077*** (0.0028)	1.01	-0.0054** (0.0016)	0.99

Table 4.2: Results from regression models in which the percentage Spanish speakers in each block group is an independent variable and data is drawn from San Diego County CalCERTS records.
Stars represent the significance of coefficients: * P ≤ 0.05 ** P ≤ 0.01 *** P ≤ 0.001

Fresno (A)	$\widehat{\beta_X}$ (Std. Error)	$e^{\widehat{\beta_X}}$
Intercept	2.5563*** (0.0428)	0.08
% Hispanic/Latino	0.0068*** (0.0008)	0.99
Republican Advantage	0.0031*** (0.0007)	1.00
% Build Year > 2000	0.0202*** (0.0007)	0.98
% Build Year 1990-2000	-0.0006 (0.0008)	1.00
% Build Year 1960-1970	-0.0027** (0.0011)	1.00
% Build Year 1930-1940	0.0064*** (0.0018)	0.99
% Build Year < 1930	0.0032** (0.0013)	1.00
% Home Value 300-750k	0.0022*** (0.0007)	1.00
% Home Value > 100k	0.0043*** (0.0009)	1.00

Fresno (B)	$\widehat{\beta_X}$ (Std. Error)	$e^{\widehat{\beta_X}}$
Intercept	2.5745*** (0.0403)	0.08
% Spanish Speaking	0.0084*** (0.0008)	0.99
Republican Advantage	0.0030*** (0.0007)	1.00
% Build Year > 2000	0.0200*** (0.0007)	0.98
% Build Year 1990-2000	-0.0010 (0.0008)	1.00
% Build Year 1960-1970	0.0032*** (0.0011)	1.00
% Build Year 1930-1940	0.0066*** (0.0018)	0.99
% Build Year < 1930	-0.0033** (0.0013)	1.00
% Home Value 300-750k	-0.0018** (0.0007)	1.00
% Home Value > 100k	0.0038*** (0.0009)	1.00

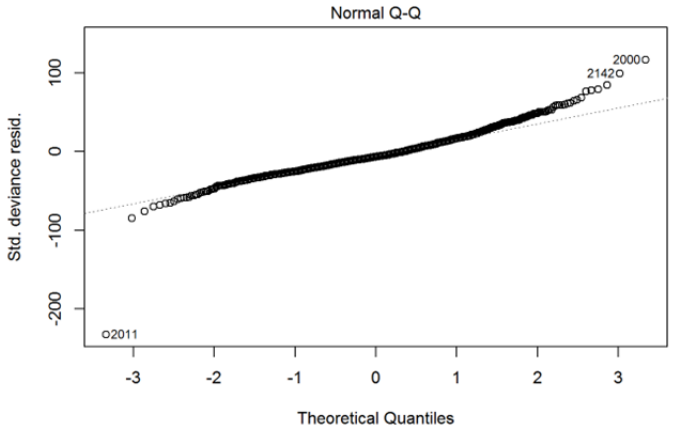
Tables 4.3 and 4.4: Results from regression models using data drawn from San Diego County CalCERTS records. Stars represent the significance of coefficients: * $P \leq 0.05$ ** $P \leq 0.01$ *** $P \leq 0.001$

Appendix: Model Validation

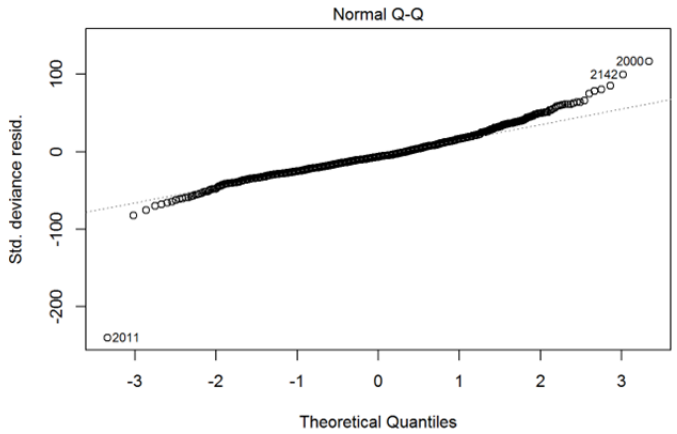
Several tests were run to test the validity of the model. These tests are employed to verify that the necessary assumptions of a logistic regression model can be applied to the chosen dataset and model specification. These tests and the results from their application to the models run in this analysis are detailed below.

Normality of Residuals

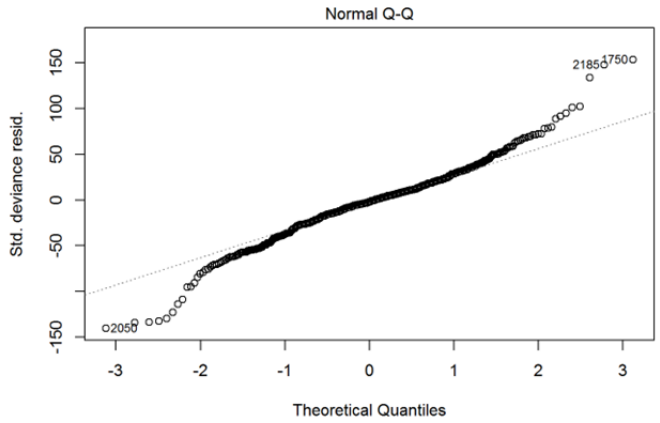
San Diego Coastal Model A



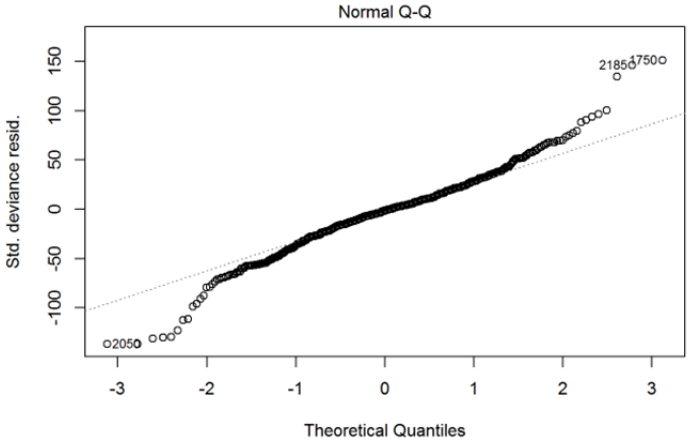
San Diego Coastal Model B



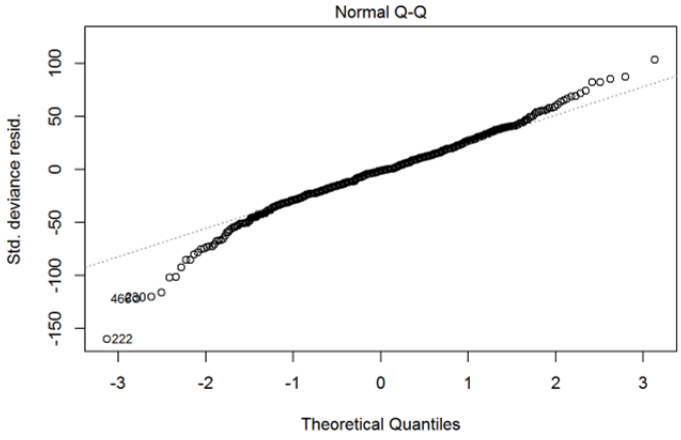
San Diego Inland Model A



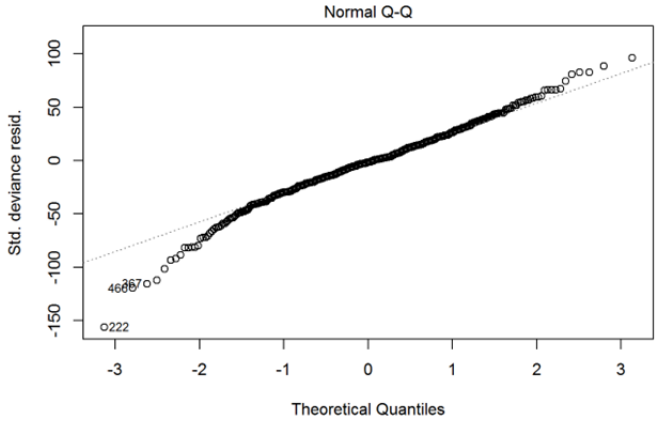
San Diego Inland Model B



Fresno Model A

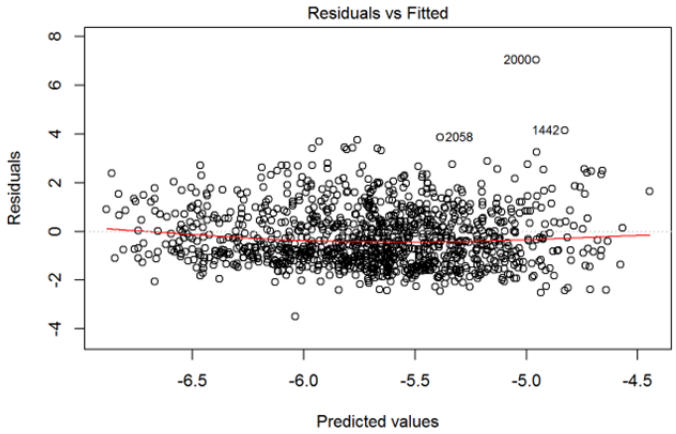


Fresno Model B

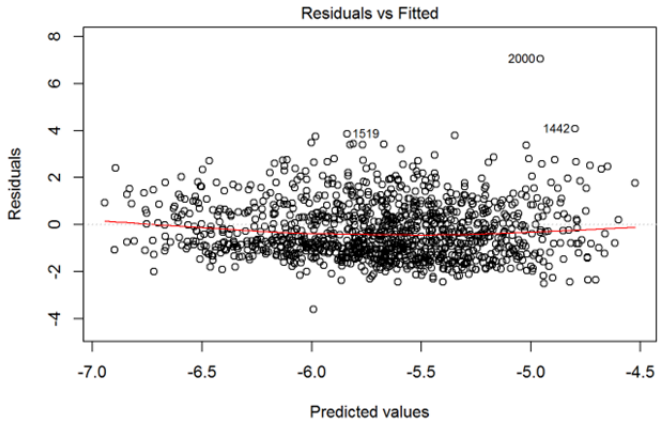


Constants Variance of the Residuals

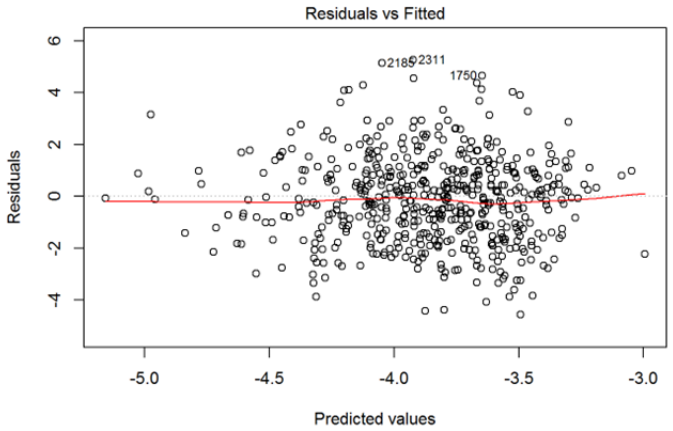
San Diego Coastal Model A



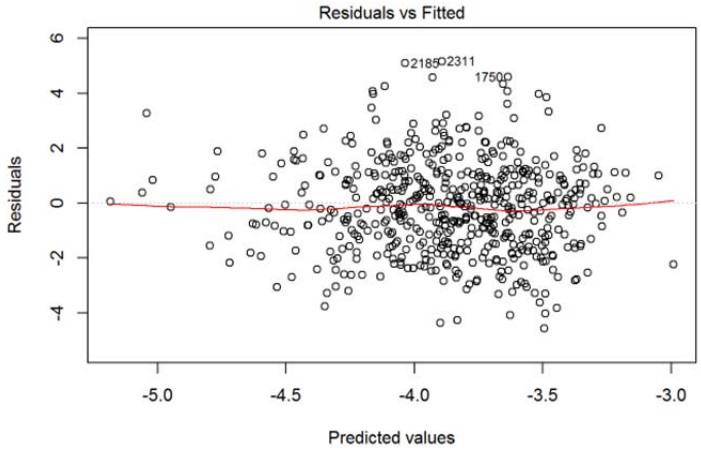
San Diego Coastal Model B



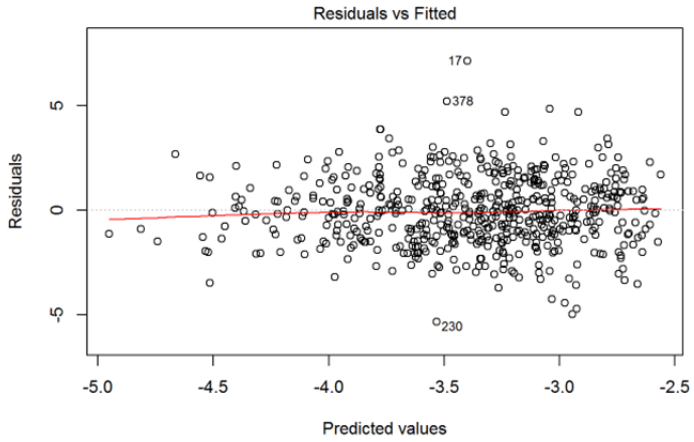
San Diego Inland Model A



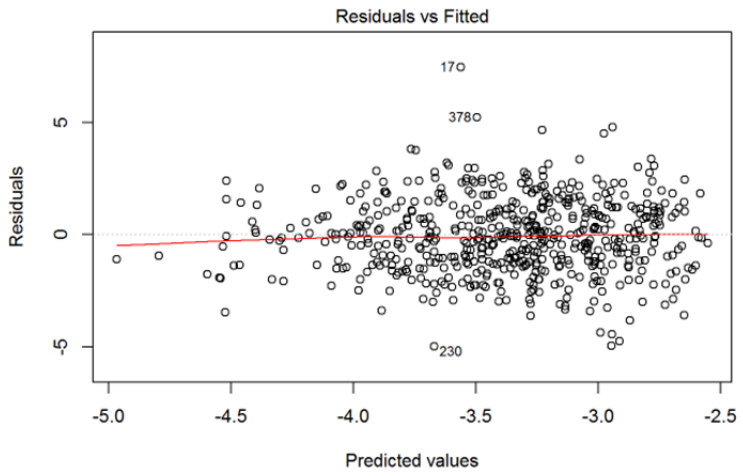
San Diego Inland Model B



Fresno Model A



Fresno Model B



Multicollinearity of Predictors (GVIF)

To test for multicollinearity of the independent variables, generalized variance inflation factors were calculated (as described in Fox and Monette, 1992) using the 'car' R library. The value of VIF can be interpreted as the inflation in size of the confidence ellipse for the coefficients in comparison with what would be obtained for completely uncorrelated data. A value of GVIF > 4 indicated multicollinearity. The values for the independent variables in all fitted models are below 4 and the assumption of no multicollinearity is supported by the following tables.

GVIF San Diego Models	San Diego Coastal (A)	San Diego Inland (A)	San Diego Coastal (B)	San Diego Inland (B)
% Spanish Speaking	2.07	1.67	2.04	1.70
Republican Advantage	2.03	1.36	2.00	1.36
% Build Year > 2000	1.55	1.14	1.16	1.14
% Build Year 1990-2000	1.31	1.19	1.31	1.20
% Build Year 1960-1970	1.26	1.35	1.27	1.35
% Build Year 1930-1940	1.25	1.14	1.25	1.14
% Build Year < 1930	1.34	1.15	1.35	1.14
% Home Value 100-300k	1.87	1.46	1.88	1.49
% Home Value > 100k	1.10	1.02	1.10	1.029

Fresno (A)	GVIF
% Spanish Speaking	1.84
Republican Advantage	2.97
% Build Year > 2000	1.27
% Build Year 1990-2000	1.47
% Build Year 1960-1970	1.40
% Build Year 1930-1940	1.60
% Build Year < 1930	1.39
% Home Value 300-750k	1.74
% Home Value > 100k	1.88

Fresno (B)	GVIF
% Hispanic/Latino	2.12
Republican Advantage	3.09
% Build Year > 2000	1.27
% Build Year 1990-2000	1.48
% Build Year 1960-1970	1.40
% Build Year 1930-1940	1.60
% Build Year < 1930	1.39
% Home Value 300-750k	1.76
% Home Value > 100k	1.85

Goodness-of-Fit

McFadden's Pseudo R^2 was calculated for each model as a measure of the goodness of fit. This statistic is somewhat analogous to the R^2 and provides an indication of the amount of variance explained by the model compared to a null model with no predictors. However, McFadden's R^2 is much lower than traditional R^2 values, with values greater than 0.2 indicating an excellent model fit. The relatively poor fit of both San Diego Coastal models may be due to the large number (nearly half) of block groups in this region with no CalCERTS records.

Model	McFadden's R2
Fresno - Spanish Speaking	0.49
Fresno – Hispanic/Latino	0.48
San Diego Coastal- Spanish Speaking	0.14
San Diego Coastal - Hispanic	0.14
San Diego Inland - Spanish Speaking	0.30
San Diego Inland – Hispanic/Latino	0.31