ABSTRACT

Matthew Palm for the degree of Master of Public Policy presented on
June 4th, 2013.

Title: Population Density and Households' Transportation and Housing Cost Trade-Offs

Abstract Approved:

B Starr McMullen

As metropolitan governments explore density-promoting "smart growth" policies, finer analysis is needed to quantify the impact of such changes on households' transportation and housing costs. Existing research suggests that households in urban areas face a trade-off between living in areas with higher housing costs and lower transportation costs or the reverse, but does not explore how density changes explicitly impact this balance. This paper uses the 2000 Census Public Use Micro Sample (PUMS) data from twenty-three of the nation's most densely populated states to identify the impact of increased population density on household rents, housing unit values and monthly mortgage payments. The project additionally explores the possibility of a negative relationship, or trade-off, between what households spend on housing and the transportation options they face. Results suggest increased population density is a strong driver of higher housing costs even after controlling for available transportation variables. Results also confirm previous research that suggests households utilizing fixed route transit systems pay a premium for that access.

Population Density and Households' Transportation and Housing Cost Trade-Offs

By: Matthew Palm

An MPP Essay

Submitted to

Oregon State University

In partial fulfillment of
the requirements for the
degree of

Master of Public Policy

Presented: June 4th, 2013
Commencement: June 17th, 2013

Master of Public Policy thesis of Matthew Palm presented on June 4th, 2013.

APPROVED:

B Starr McMullen, representing Agricultural and Resource Economics

Kate Hunter-Zaworski, representing Civil Engineering

Haizhong Wang, representing Civil Engineering

I understand that my thesis will become part of the permanent scholarly collection of Oregon State University Libraries. My signature below authorizes release of my thesis to any reader upon request.

Matthew Palm, Author

ACKNOWLEDGMENTS

I would like to extend my sincerest thanks to Starr McMullen, Kate Hunter-Zaworski and Haizhong Wang for their support throughout the course of this project and my time here at Oregon State University. I would like especially acknowledge Starr, for chairing this committee and keeping me focused on narrowing down the scope of this project. I would also like to extend sincerest thanks to Brent Steel, the program director for the MPP program, for supporting and assisting me in completing this program at Oregon State.

Finally, I am thankful for the support for this research provided by the Furman Fellowships for Research in Transportation Economics.

Table of Contents

List of Tables & Figures

Abstract

As metropolitan governments explore density-promoting “smart growth” policies, finer analysis is needed to quantify impact of such changes on households’ transportation and housing costs. Existing research suggests that households in urban areas face a trade-off between living in areas with higher housing costs and lower transportation costs or the reverse, but does not explore how density changes explicitly impact this balance. This paper uses the 2000 Census Public Use Micro Sample (PUMS) data from twenty-three of the nation’s most densely populated states to identify the impact of increased population density on household rents, housing unit values and monthly mortgage payments. The project additionally explores the possibility of a negative relationship, or trade-off, between what households spend on housing and the transportation options they face. Results suggest increased population density is a strong driver of higher housing costs even after controlling for available transportation variables. Results also confirm previous research that suggests households utilizing fixed route transit systems pay a premium for that access.

A. Introduction

The World Bank states that climate change will increase costs, hardships and strains on human societies in the next century (World Bank 2012). Over five hundred state, regional and local governments began responding to this challenge by signing on to a compact to reduce emissions (US Conference of Mayors 2012). Oregon House Bill 2001 and Oregon Senate Bill 1059 charge the Lane Council of Governments (LCOG), the Portland regional metropolitan government (METRO) and several state agencies with assessing strategies for reducing greenhouse gas emissions from transportation (Oregon Senate 2010). The Oregon Department of Transportation (ODOT) Planning and Analysis Unit's Greenhouse gas Statewide Transportation Emissions Planning Model (GreenSTEP), developed by Brian Gregor, tests the efficacy of different transportation and urban form policies on reducing transportation related emissions (Gregor 2010). GreenSTEP modeling points to increasing urban population densities in Oregon communities as a solution to reduce vehicle miles traveled and greenhouse gas emissions from transportation. But regional governments must also consider negative or unanticipated side effects of policies promoting urban densification before embarking on major policy overhauls, particularly impacts on housing affordability. The impacts of increasing density on housing affordability could determine the political life or death of policies promoting 'smart growth' and sustainable urban design.

To explore the other side of effects of increased density on our communities, this project asks if population density increases household housing costs. The paper addresses this task using the 2000 Census Public Use Micro-Sample (PUMS) data for all metropolitan areas in 23 of the densest states in the United States. The metropolitan areas covered in this sample contain a population of over 165 million people.

The following literature review discusses the possibility of a relationship between housing costs and density. The data selected for this project is then introduced and discussed along with its limitations. The estimation model and methodology are presented followed by the basic ordinary least squares (OLS) results. The author estimates and presents results for the base models, followed by a section in which regional-mean dependent variables from the data are modeled using the 2000 Census PUMS data. A conclusion considers the implications of this research for urban and transportation planning and discusses next steps for continuing this research.

B. Theories On Population Density And Housing Costs

The multiple theories on why population density might correlate positively with housing costs can be divided into two categories: those which suggest that housing density and costs are endogenously and jointly determined, and those that suggest density drives higher costs because the benefits of living in a denser area attract more demand for housing there. The theoretical benefits of living in a denser area include:

1. Reduced transportation costs as destinations are closer.
2. Better social amenities, as denser urban areas contain populations that can support greater specialization in sectors ranging from health care to restaurants and entertainment.

Synthesizing these points, a large body of literature will be reviewed that demonstrates that:

3. Households in denser area can access more and better amenities at lower transportation costs than households in less dense areas, suggesting desire to locate in dense areas to achieve this accessibility drives up prices in dense communities.

The first subsection of this review discusses the evidence that population density and housing costs are jointly determined by other variables in urban housing markets. The two following subsections document the mounting evidence that increased residential density offers benefits to households that may induce more people to attempt to move into these areas and bid up prices further. The review concludes by discussing one study that suggests a negative relationship between population density and housing costs.

1. *Housing Density and Costs in A Concentric City: Jointly Determined, Highly Correlated*

The 'monocentric' model of a traditional city, and its application in evaluating household residential choices, suggests that population density is positively correlated with housing costs. This model posits a concentric city with a central business district (CBD) at its core and assumes that households prefer to live near the CBD to minimize their commute costs, all other things being equal. In response, builders attempt to maximize their gains from rising demand for housing near the CBD. They do this by building smaller and denser dwellings there, ceteris paribus (Alonso 1964; Muth 1969; Mills 1967). When the housing markets are in an equilibrium state, population density and housing costs are hypothesized to decline as the distance of those neighborhoods from the CBD increases. Mills first confirms this theory with an analysis of multiple monocentric urban areas in the United States from 1910 to 1963 (1970). The model suggests that population density and housing costs increase together as increased demand to live in a given area jointly increases prices and housing density in that area. This is expressed graphically in the figures below which show a negative relationship between population density and distance from the central business district (CBD). The second graph further illustrates the impact of this hypothesized relationship between housing rents (or prices) as the distance from the CBD increases:

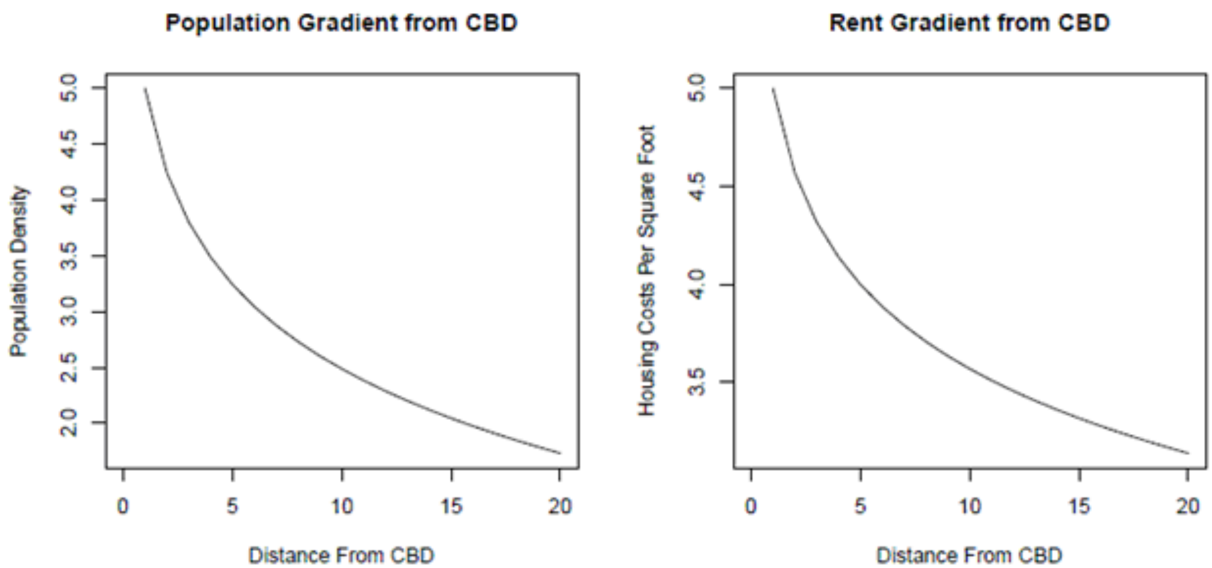

Figure One: Theoretical Population and Rent Gradients in a A Monocentric Urban Area

A broad, multi-city application of the monocentric model in previous decades comes from testing the model on urban areas that experienced suburbanization and sprawl in the decades after World War Two. Brueckner and Fansler (1983) tested model assumptions using census data on small cities for which they could derive estimates for agricultural rents and which they could show had a monocentric urban structure. Their model assumes assume sprawl takes place as a natural market response to changes in the urban equilibrium: incomes increase and populations increase. Higher incomes and populations increase the demand for housing, driving prices higher. The price increases pushes rents higher, inciting developers to produce denser dwellings to maximize returns (Fujita 1980). This concept is known in urban economics as Bid-Rent Theory (Fujita 1980). Lastly, the increasing demand for land in the city drives rents high enough that land owners at the margins are induced to switch from agricultural purposes to housing, increasing city size. Their test on this assumption found that incomes and city population did correlate positively with city size, while agricultural land rents around the cities tested correlate negatively with city size as suggested by the theories in the monocentric urban model.

Recent application of the model confirms the hypothesis of population density and housing costs jointly declining as distance from the CBD increases. Kulish, Richards and Gillitzer (2011) test the model's assumptions on postal code level median housing prices in major cities in Australia. While Australian cities are culturally and politically distinct from American cities, the study still offers insights into the validity of Bid-Rent Theory in a world of multi-centric megacities. The Australian postal code functions similarly to the United States zip code. Dividing Australia's population in 2001 by its number of postal codes that year produces an average of 7,800 residents per postal code, mirroring roughly 7,400 residents per zip code in the United States (Australian Bureau of Statistics 2013; United States Census Bureau 2013). They find that population density and median housing prices decline the further a postal code is from the CBD. They note that bordering a waterfront also contributes to higher postal code level population densities and median housing prices. Their results suggest that population density and housing prices correlate positively within an urban market. Their work does not suggest, however, that density drives prices itself.

Unfortunately, most cities in North America do not conform to a monocentric layout. Some researchers attempted to deal with this by testing the hypothesis that a neighborhoods' distance to multiple centers of employment positively impacts a neighborhoods' population density, rents and home values. In a working paper for the University of California Transportation Center, Song (1992) uses traffic analysis zone (TAZ) level data from the 1980 Census to test the assumptions of mono-centricity in the Los Angeles area. To test this, Song (1992) regresses TAZ resident worker populations on the zone distance to the CBD. Then Song regresses TAZ resident worker population on distances to five major employment centers in the Los Angeles area. The polycentric model performs far better with higher adjusted r-squareds and more significant coefficients. Song then states that in the polycentric city, one key assumption in the original Alonso-Muth-Mills model still holds: households locate with proximity to employment in mind, and this can be empirically observed even in cities with multiple employment centers.

Does proximity to multiple employment centers impact the price of housing directly? Ottensmann, Payton and Man (2008) find support for the significance of positive coefficients representing the distance to multiple employment centers on housing prices in a series of hedonic housing price models. They use housing sales data from the Multiple Listings Services (MLS) for Marion County, Indiana, and start out with a base regression using the following variables Palmquist (1988) deemed essential in a hedonic housing price model: square footage, number of bathrooms, number of rooms, lot size, unit age, neighborhood schools' SAT scores, tax rates, racial neighborhood make up and neighborhood median income. After running this base model regression, they run additional models including variables for distance to the CBD, a nearby township center, and a variety of employment centers. Results for the coefficients are positive and significant, but they contribute only marginally to explained variance in housing prices compared to a baseline regression without them. Instead, they conclude that structural attributes of houses and the characteristics of their immediate neighborhoods are the major determinants of housing prices.

2. Households in Denser Areas Face Reduced Travel Costs

According to bid-rent theory (Fujita 1980) , population density should correlate positively with housing costs and prices, as a denser area may guarantee not only closer proximity to employment, but closer proximity to shopping, entertainment and services like health care. This should make denser areas more attractive to residents, resulting in higher housing prices. Communities at certain levels of population density may even enable residents to access most of their amenities by walking and biking, providing a 'quality of life' factor that may also increase prices in such neighborhoods. Additionally, denser areas may provide a better transit network, which can mean lower travel time costs for individuals without vehicles. This section outlines evidence on the reduced travel costs of households in denser areas, and discusses a recent body of research stating that these travel benefits increase housing costs.

A study commissioned by the National Research Defense Council (2000) offers the most sweeping and dramatic conclusions that population density reduces vehicle ownership and vehicle miles traveled (VMT). Hotzclaw et al (2000) draw on travel analysis zone (TAZ) level data from the Chicago, Los Angeles and San Francisco regional governments' and use odometer readings and census data from those metropolitan areas to construct cross-sectional regression models for household auto ownership and vehicle miles traveled (VMT) per household at the TAZ level. After testing a series of population density, work center proximity and job accessibility measures on VMT, the authors find that each doubling of population density, specified in power form, correlates with a 33% to a 43% reduction in VMT. Across the three topographically and economically distinct metropolitan areas in the study, they find relatively consistent coefficients for relevant land use variables on VMT, particularly household residential density. The authors globalize conclusions and articulate their study's applicability to countries much different than our own. They do not apply the model to any other countries. But this study did not consider how residential self-selection may explain, in part, the reduced VMT of households in residential areas. They cannot disprove the possibility that the density effect on VMT is actually the result of dense communities attracting certain kinds of residents already predisposed to driving less, as their regressions did not account for the differences in attributes of the households living in the TAZs, such as the age and education level of the head of household or share of households raising children. As such, Hotzclaw et al (2000) cannot definitely conclude that denser communities induced people to drive less versus a counter theory that denser communities attract people who are pre-disposed to drive less for reasons like being young or low income.

Is the correlation between neighborhood population density and reduced VMT caused by higher population densities, or do denser communities attract certain kinds of individuals who drive less for other reasons? Brownstone and Golob (2009) use urban household data from the California sample of the 2001 National Household Transportation Survey (NHTS) to develop structural equation models for residential density, fuel usage and vehicle energy efficiency. Golob and Brownstone's work provides stronger conclusions than Hotzclaw et al 2000 because they use disaggregate household level data, enabling them to control for demographic variables that may better explain travel differences between households, such as age and gender of household head, number of children and number of workers. They still find that residential density correlates negatively with households' fuel consumption through reduced VMT. In addition, they suggest that residential density indirectly reduces households' fuel consumption because households in denser areas may buy more fuel efficient vehicles. They conclude that a household in a less dense neighborhood consumes more fuel than a household with equivalent demographic attributes in a higher density neighborhood. These demographic attributes include: household income, education, race, number of drivers and number of workers.

Building on Brownstone and Golob's work, Su (2011) uses the same 2001 NHTS data from all households in metropolitan areas nationwide to regress household fuel consumption on population density. Su goes a step further and attempts to capture transportation supply and congestion in his models by adding MSA level variables from the Texas Transportation Institute and the Urban Mobility Institute. These variables include: the spatial size of the households' MSA, the number of highway lane miles per square mile of MSA, the number of hours of annual delay per peak hour traveler, rail availability as reported by the household, bus availability as reported by the household, and MSA level transit revenue miles per peak-hour traveler. Su finds that even after controlling for the aforementioned transit supply and congestion variables, population density correlates negatively with household fuel consumption. Su's results also imply that road density, transit supply and peak hour traveler density correlate positively with population density, as Su reports the covariance between these variables and population density rises near or above .6.

Heres-Del-Valle and Niemeier (2009) also take Brownstone and Golob's work a step further by formulating a two-stage instrumental variable regression of VMT and automobile ownership with a broader sample of California household from the NHTS that includes rural areas. Heres-Del-Valle and Niemeier note, for example, that Golob and Brownstone (2009) consider how the number of children in a household impacts the household's residential location choice, but then do not consider how this variable also impacts VMT directly after controlling for residential location. Heres-Del-Valle and Niemeier apply instrumental variables for each land use characteristic included in their models to address residential self-selection. The authors choose three instrumental variables presumed to correlate with the endogenous regressors: the percentage of units in the neighborhood built before 1939, the percentage of the neighborhood population that is non-white, and the percentage of households containing families. They find an inelastic relationship between density and VMT: a ten percent increase in population density correlates with a 1.9% reduction in VMT, holding everything else constant. The application of instrumental variable regressions to address the relationship between VMT and population density still produces a significant negative coefficient for the effect of density on VMT. Combining these results with the discussion that follows provides a more complete picture of how high density communities impact what households spend on both housing and transportation.

Research on how transit development impacts housing prices suggests that potential commuter savings from taking subsidized public transit can come at the price of higher housing costs. Wadrip (2011) claims a consensus exists in the literature that proximity to public transit leads to higher home values and rents, particularly for multi-family, high density units, in his review of the literature for the Center on Housing Policy. Research produced by Center for Neighborhood Technology (CNT) finds that while residential unit sales values declined from 2006 to 2011, they rose for units within proximity to transit during that same period in Phoenix, Chicago, Boston, Minneapolis and St. Paul, and San Francisco (Becker et al, 2013). Writing for the Brookings Institute, Leinberger and Alfonzo (2009) study the relationship of housing and retail rents per square foot on neighborhood walkability scores and find walkability has a positive effect on rents in the Washington D.C. area. Cortright (2009) develops hedonic housing price models for over 90,000 properties in fifteen metropolitan areas and includes a "Walk Score" variable provided by ZipRealty.com to specify the impact of neighborhood walkability characteristics on housing prices. Walkability scores are point scores for housing units, with points awarded based on the number of services and amenities within a quarter mile of the property and partial point scores awarded for amenities within a full mile of the property. By definition, then, the Walkscore approximates a density of amenities and services within walking distance of the housing unit. Cortright develops separate models for each of the fifteen MSAs included in the study, and controls for a battery of variables traditionally included in hedonic housing price models: distance of property to CBD, number of bedrooms, baths, age of unit and number of jobs within a three mile radius. Cortright finds statistically significant, positive coefficients for walk scores in thirteen of the metropolitan areas covered. Cortright also finds mixed results for the theory that distance to CBD contributes to higher housing prices, with coefficients for distance to CBD varying in sign and significance across all thirteen regression models for all thirteen cities.

This evidence does not tackle population density and housing prices directly, but does so implicitly. Cortright's walkability measure is defined by the number of amenities within walking distance of the unit, and these amenity types are listed as: grocery stores, coffee shops, movie theaters, parks, bookstores drugstores, clothing and music stores, restaurants, bars, schools, libraries, fitness centers and hardware stores (Cortright 2009). As such, this walkability measure captures the density of the unit's neighborhood amenities itself. Cortright's findings thus indicate a positive relationship between amenity density and housing prices.

3. Increasing Population Density May Contribute To, Or Correlate With the Growth of Urban Amenities That Drive Housing Prices Higher

Multiple economists (2000) have noticed that housing costs in urban areas rose faster than wages in the last few decades before the Great Recession. They propose that a household's desire to access specific urban amenities drives this phenomenon, implying that travel costs and commute times are not the only spatial factors in residential location. Glaeser, Kolko and Saiz define the urban amenity premium as the social or consumptive returns to households locating in increasingly larger and denser cities (2000).

They arrange these returns into four broad categories: increasingly rich and diverse goods and services, aesthetics, good public services and speed. In explaining the last of those, speed, they articulate the same conclusions suggested by research in the previous section of this review. Glaeser, Kolko and Saiz (2000) note that rising incomes increase residents' value of time. Higher income earners may thus try to locate their residence in places that minimize their time spent in travel.

Population density itself may contribute to the existence of some of these returns to household for locating in urban areas. Schiff (2009) uses population, density, income and education levels across metropolitan areas for a regression modeling the number of types of cuisines listed in those cities on CitySearch.com. He finds that increased population and increased population density contribute to an increase in cuisine offerings in cities. Dense, large cities place large numbers of people within reasonable proximity to potential restaurant sites, increasing the likelihood that enough clientele for niche cuisines live close enough to those sites to support niche restaurants locating there. Rappaport (2008) identifies a powerful cross-sectional correlation nationwide between population density and a list of consumption amenities that includes availability of the arts, education, recreation opportunities, a good climate and low crime. Rappaport concludes that the population density of cities may be determined by households' desire to locate in areas with appealing amenities and a better "quality of life," in spite the loss of "quality of life" that occurs from congestion and crowding. Last, a state-level regression of gross state product per capita on state-level employment densities and income from the 1980s finds significant correlation between incomes and employment density at the statewide level (Ciccone and Hall 1997). This relationship may be an endogenous process, particularly in areas where agglomeration economies are present, like Silicon Valley in San Jose and the financial sector in New York, New York (Glaeser and Gottlieb 2009). Mounting evidence suggests that regardless of the transportation cost reductions associated with density, density may be endogenously related quality of life as desire to live in a "good" place leads to denser development there, which in turn enables more diversity of consumption amenities.

But increased density may not be purely beneficial for households, as research also identifies points at which density may contribute to nuisances like traffic and increased crime. In a recent study of crime rates in Baltimore, Maryland, Harries (2006) produces results that verify a large body of criminology research that suggests denser areas face higher crime. Using data from the Texas Transportation Institute, Hahn Chatterjee and Young (2002) note that population density correlates positively with the travel rate index, a measure of congestion, in a regression using the largest one-hundred and thirty eight metropolitan areas in the United States. Longer-term analysis of urban areas presented in the follow section helps reconcile the mixed impacts of density, with Cho (1997) noting that density correlates with higher prices up to a point until the negative impacts of density described above begin outweigh the benefits.

4. Research on Urban Growth Restrictions—Inconclusive on the Density-Cost Hypothesis

Unfortunately, research testing the impact of "smart growth" planning, from urban growth boundaries to "build up not out" policies offers no conclusive answers to the density, housing cost hypothesis. Compact development in the form of building taller multi-unit structures must increase if population density is to increase by definition, and a rental premium for building height has already been identified in the literature (Ali and Moon 2007). At a certain point, the cost savings of living in a smaller dwelling in a multi-unit structure are surpassed by construction and maintenance needs of larger, multi-story structures. This implies that prices per square foot should be lower for dwellings in buildings under four stories, but that above that increased building size will contribute to increased prices.

South Korea's greenbelt effectively "contained" growth in the capitol of Seoul starting in 1971, and its impacts on population density and housing prices suggest that encouraging density in this way may contribute to increased housing costs in the short run. Using data from 1970 to 1989, Cho (1997) confirms that such policies contributed to the development of high-density units and a greater wealth transfer from home owners to builders compared to a situation without growth controls. But unlike other research focusing on Seoul, Cho also models for the congestion effects of such intense development. Cho also looks at the impacts of a long run supply response to price increases in Seoul in the 1970s and the effects of congestion on returns to households' location decisions. Based on an assessment of these trends, Cho suggests that restrictions imposed by urban growth control policies do not lead to permanent increases in housing costs in the long run. Eventually, builders respond to demand for additional housing, producing more dense housing over time and causing prices to flatten. Cho argues that a restriction on land supply alone does not provide sufficient condition for housing price inflations if developers can respond by building smaller, denser units. Cho also notes that at the level of density when congestion costs surpass travel time savings benefits for residents, housing costs flat line or decline. The analysis suggests that a true evaluation of growth control or pro-density policies must also consider the diseconomies of density, namely: congestion pressures on transportation networks, water and sewer services, air quality and even schools.

Oregon also provides a useful but inconclusive case study on the impact of pro-density, "sprawl control" policies on the urban housing market. Using housing data at the census block level from 1990 and 2010, Jun (2006) runs hedonic housing price models for housing units in the Portland metropolitan area's three counties. Jun includes a dummy variable for if the census block is contained within the UGB. The dummy variable is insignificant, but Jun's variable on block level density is negative and significant, contradictory to theory. While this could give land-use planners some comfort in using density to control emissions without negatively impacting households' budgets, Jun's paper faces a few limitations. First, the analysis does not appear to differentiate between owner occupied housing for single-family detached units, attached units and units in a multi-unit structure. Second, use of block-level population density would not seem to capture the effect of a market or sub-market's density on housing prices. Individuals considering living in a city, or even in a given sub-section of a city, should be assumed to be choosing exclusively among houses within such a small segment of a housing market. Third, the model relies on block level averages of household attributes and not disaggregate, observed household attributes and costs.

In contrast to Jun(2006), a hedonic rent study focusing on the greater Boston area found density correlates positively with rents. Fisher, Polakowski and Zabel (2009) find a high significant log-log density coefficient for rents in the greater Boston area of .026 for single family units and .024 for condominiums (Fisher Polakowski and Zabel 2009). Their study uses disaggregate, housing-unit level data and finds that school quality measures and a neighborhood index for employment accessibility also both correlate very strongly and positively with rents.

This project adds to the literature discussing population density and housing costs in several unique ways:

1. It is first paper to use disaggregate household level data at the national level.
2. It both uses monthly housing costs reported by households and estimated property values.
3. It includes renters and home owners of single family and multi-family units, and produces estimates with positive signs for density across all these market segments.

C. Definitions, Model and Methods

1. Defining Density

This paper defines density as the residential population per square mile. Due to data limitations, the population density is defined at the PUMA-level. The PUMA, or Public Use Micro-Sample Area, is the lowest geographic unit associated with individual household responses in the 2000 Public Use Micro-Sample (PUMS) datasets used for this project. Each PUMA from the year 2000 is designed to encompass a minimum 100,000 residents counted in the 2000 Census, with a priority placed by designers on ensuring PUMAs line up consistently with county and city boundaries to the best extent possible (Missouri Census

Data Center 2011). So a county like San Francisco, with a population of around 770,000 residents at the time of the 2000 Census, contains only seven PUMAs, none of which include any areas outside of San Francisco County (Missouri Census Data Center 2011).

PUMA level density offers the best way to test the impact of density on housing prices given this limitation. An MSA/CMSA -level density variable would not capture how variation in population density across the same regional market contributes to price variation within that market. In the San Francisco example, an MSA/CMSA level density variable would treat rents in downtown San Francisco and distant suburban Walnut Creek as identically impacted by the density effects discussed previously in the literature. One location puts a resident within walking distance of a major employment center and hundreds of urban amenities ranging from art museums to a Chinatown, while the other may require residents to drive to reach the same amount of employment and just a dozen such amenities. Thus a density variable shared by these locations would not capture the effect of two of the three theoretical justifications for density effecting rents and prices that are bulleted in the introduction of the previous section.

In the papers discussed in the literature review, population density is calculated at a variety of levels smaller than the PUMA: traffic analysis zones for Hotzclaw et al (2000), census tracts for De Velle and Niemeier (2009) and census block groups for Jun (2006). While these may suffice for VMT research, they will not work for looking at density and housing prices. Census tracts, which are made up of block groups, are designed to contain close to an optimal number of 4,000 people according to Iceland and Steinmetz (2003). In very dense areas like San Francisco or Manhattan, such small scale precision becomes problematic for operationalizing the mechanisms by which density raises prices as discussed in the literature. Two such tracts in the downtown San Francisco example discussed earlier may have both very high but significantly different population densities because one contains higher condominium towers, yet both would still be within walking, biking and transit distance to the same employment and amenities as many are less than a kilometer wide. Furthermore, the urban amenity premium is defined in the early section as growing more specialized as areas become denser and more populous. Figure two below shows the difference between the census tracts and PUMAs in San Francisco, California, provided with the PUMS dataset from the Census Bureau (2003).

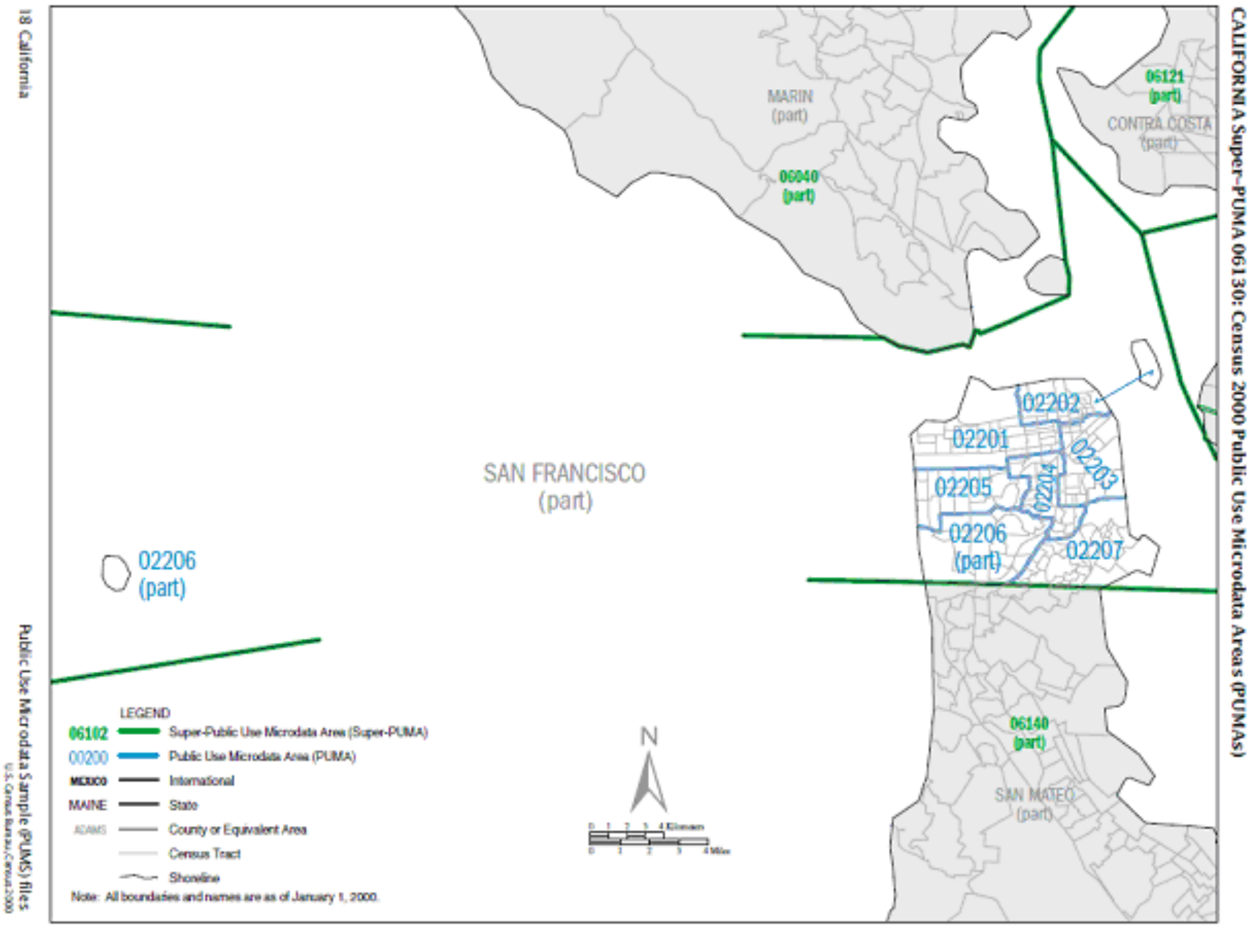

Figure 2 : PUMAs Overlaying Census Tracts in San Francisco (Census Bureau 2003)

2. *Definitions of Housing Prices, Rents and Costs and Their Strengths and Weaknesses*

The PUMS dataset includes three variables which form the primary basis of the analysis for this project: households' gross monthly rents for renter-occupied units, and estimated housing values for owner occupied units, and monthly mortgage payments for primary (first) mortgages. For some respondents in the dataset, mortgage payments included property taxes. Core models are developed to predict these variables, and then applied to monthly mortgage costs to explore the feasibility of specifying the impact of population density on monthly household owner costs through higher household spending. The following section details how rent and estimated value models are separated by building type: single and multi-family dwellings. To address concerns that respondents may not have had an accurate understanding of the value of their homes, an additional model run is regressed on housing unit value just among households who purchased their homes within the last year.

This paper adds to the literature evaluating the impacts of travel cost, urban density and other variables on housing markets by studying their impacts on home owners' monthly mortgage costs in addition to the estimated value of their homes. This approach originates in seminal research funded by the Brookings Institution and developed by the Center For Neighborhood Technology and Center for Transit Oriented Development (Haas et al 2006). Their analysis uses selected monthly owner costs, a variable in the PUMS dataset, to look at the relationship between housing costs and transportation costs on a month to month basis. They argue that this approach captures how households budget housing and transportation costs jointly, capturing the relationship that housing costs have to transportation costs when households on limited budgets must substitute one for the other. Their analysis demonstrates the impact increased density has on reduced travel costs.

To address concerns that a mortgage cost model is skewed by the years in which individuals first purchased their homes at a certain interest rate, an additional model run is regressed on primary mortgage amount just among households who purchased their homes within the last year to control for financial conditions influencing home buyers' decisions.

Lastly, the author notes that the mortgage amount variable available in the dataset, MRT1AMT, was altered to provide for consistency across households. Under forty percent of the households in the survey reported that their taxes were not automatically included in their mortgage payments. To correct for that, the TAXAMT variable was used to code the tax rate of those households, with those taxes added to the MRT1AMT variable where they weren't already to provide for consistency in the data. Additionally, for households that indicated insurance was automatically rolled into their mortgages, the insurance amount variable, INSAMT, was subtracted from the MRT1AMT to provide for consistency. The final variable created from this, labeled MRG, is thus a household's monthly mortgage payment on their primary mortgage with taxes included and insurance costs removed.

A final model is developed in response to Jun (2006), who found that median home prices at the zip code level correlated negatively with zip code level population density. As stated in the review of Jun's work in the previous section, the results may be a function of units in denser areas being cheaper because they are smaller, not because density correlates negatively with housing prices in properly specified models. This final model regresses on PUMA mean and median rents for the 1015 PUMAs included in the study using mean rents from the 2000 long form Census data (PUMS) and the 2006-2010 American Community Survey (ACS) wave PUMS data.

3. Data

The initial models are estimated using PUMS 5% sample data for the year 2000 for metropolitan counties in 23 states around the country, including the most populous and densely populated states. The combined metropolitan statistical areas included in this study rang from a size of 406,934 people to over twenty-one million people. In total, they include a population of over 101 million residents, over a third of the United States population in the year 2000. The 2000 PUMS data is used to avoid housing market distortions caused by the housing bubble and subsequent crash. Table 1 list the states included in the study.

Table 1: States Included in the Study

Alabama	Indiana	New York	Tennessee
California	Kentucky	North Carolina	Texas
Connecticut	Louisiana	Ohio	Virginia
Florida	Massachusetts	Oregon	Washington
Georgia	Michigan	Pennsylvania	Wisconsin

Illinois	New Jersey	South Carolina	

Figure three highlights all the states used in the study. This projects' range spans both coasts and includes states from all regions of the country save the rocky mountain area.

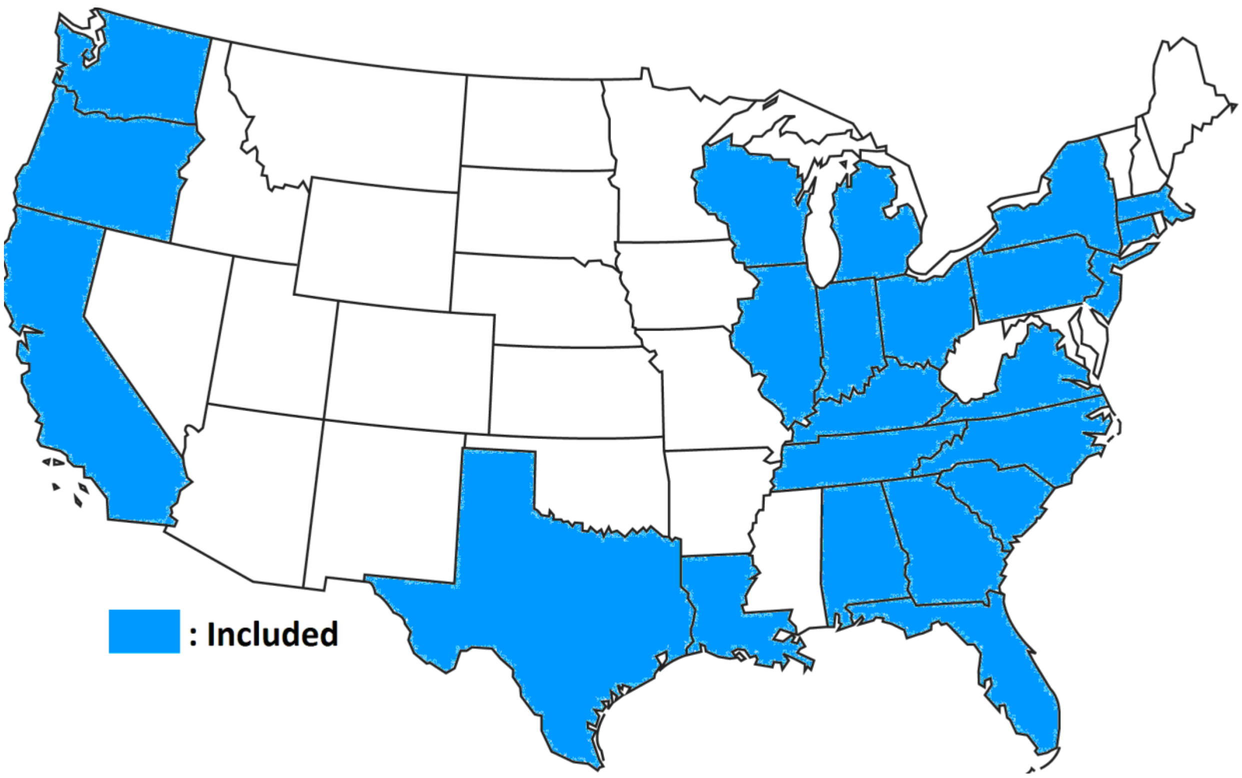

Figure Three: States Included In The Study

Several filters were applied to the PUMS data. The following were excluded:

• Households located in PUMAs coded by the Census Bureau as outside metropolitan areas (outside of MSA/CSAS)

• Vacant housing units

• Group quarters

• Housing units that were boats, RVs, vans, and similar

• Housing units without complete plumbing or kitchens

• Housing units that included meals in the rent

Several community variables were added to the PUMS data. The residential population density of each PUMA was calculated from year 2000 population and geographic area estimates provided by a web-based application hosted by the Missouri Census Data Center (http: //mcdc2.missouri.edu/websas/geocorr2k.html). The mean household income was calculated for each PUMA as a weighted average using the household income of the survey households in the PUMA (HINC variable) and the respective household weights (HWEIGHT variable). These variables will be discussed in more detail in the following section.

The individual state datasets were split into subsets by tenure and housing type. The state datasets were then combined into national datasets by tenure and housing type. The multi-state datasets were used to estimate the following models detailed in Table 2:

Table Two: Models

Dependent Variable	**Unit Type**	**Tenure Type**	**Other Specifications**
Gross Monthly Rent	Single Family	Rented for cash rent	None
Gross Monthly Rent	Multi Family	Rented for cash rent	None
Estimated housing value	Single Family	Owned by residents	None
Estimated housing value	Multi Family	Owned by residents	None
Estimated housing value	Single Family	Owned by residents	Owners who have resided in unit for less than one year
Estimated housing value	Multi Family	Owned by residents	Owners who have resided in unit for less than one year
Primary Mortgage Monthly Payment With Taxes, Insurance Excluded	Single Family	Owned by residents with mortgage	Owners who have not paid off mortgage

Primary Mortgage Monthly Payment With Taxes, Insurance Excluded	Multi Family	Owned by residents with mortgage	Owners who have not paid off mortgage
Primary Mortgage Monthly Payment With Taxes, Insurance Excluded	Single Family	Owned by residents with mortgage	Owners who have not paid off mortgage and who have resided in unit for less than one year
Primary Mortgage Monthly Payment With Taxes, Insurance Excluded	Multi Family	Owned by residents with mortgage	Owners who have not paid off mortgage and who have resided in unit for less than one year

The author used Chow Tests in which these models are pooled together and then regressed separately to identify if significant structural differences existed in these data groupings. The author found a statistically significant and positive effect from the Chow Tests for the pooling and splitting of these groups by unit type (single family versus multi-family), and between recent home buyers (years of residence <1) and those who were not recent home buyers. All Chow test statistics reported significance at the .000 level.

2a. *Overview of Theoretical Model and Variables Included*

This paper utilizes the ordinary least squares (OLS) regression tool for analysis with robust, clustered standard errors that are explained in a following section. The regressions for each model run using the entire sample. Housing price models of the following form are estimated to assess the potential effect of urban population density on households' monthly base housing costs:

$$V = \beta_1 + \beta_{1\times M}M + \beta_{1\times H}H + \beta_{1\times N}N + \beta_{1\times L}L + \varepsilon$$

$$R = \beta_1 + \beta_{1\times M}M + \beta_{1\times H}H + \beta_{1\times N}N + \beta_{1\times Q}Q + \varepsilon$$

$$P = \beta_1 + \beta_{1\times M}M + \beta_{1\times H}H + \beta_{1\times N}N + \beta_{1\times L}L + \varepsilon$$

$$C = \beta_1 + \beta_{1\times M}M + \beta_{1\times O}O + \beta_{1\times N}N + \beta_{1\times L}L + \varepsilon$$

$$D = \beta_1 + \beta_{1\times M}M + \beta_{1\times O}O + \beta_{1\times N}N + \beta_{1\times Q}Q + \varepsilon$$

$$E = \beta_1 + \beta_{1\times M}M + \beta_{1\times O}O + \beta_{1\times N}N + \beta_{1\times L}L + \varepsilon$$

in which V is a vector of housing values for single family units, R is a vector of monthly gross rents for single family renters, and P is a vector of primary mortgage monthly payments for single family units. Additionally, C is a vector of housing values for multi-family units, D is a vector of monthly gross rents for multi-family renters, and E is a vector of primary mortgage monthly payments for multi-family units. M is the MSA/CMSA level dummy, H is a matrix of unit or structure characteristics found to be relevant for single family units, and O is a matrix of structure or unit characteristics found to be relevant for multi-family units (Palmquist 1988, Ottensmann, Payton and Mann 2008), N is a matrix of household characteristics to either control for household self-selection into certain kinds of communities (Golob and Brownstone 2009; Su 2011), or to serve as proxies for neighborhood attributes not available in the data. L is a matrix of PUMA-wide characteristics relevant to home owners and Q is a matrix of PUMA-wide characteristics for renters. The β1 is the constant term vector and ε is a vector of error terms. Since the distribution of housing prices does not fit the normal distribution required for ordinary least squares, we assume housing values, rent and mortgage payments will also be and thus adopt from the literature a transformation of the dependent variables for both markets into natural log form (Palmquist 1988; Ottensmann et al 2008). This is the most commonly-used specification in hedonic housing price models (Ottensmann et al 2008). The log transformation can reduce the heteroskedasticity associated with the use of highly-skewed price and cost variables.

Log-log relationships in ordinary least squares regression produce elasticity results between each explanatory variable and the dependent variable as discussed by Verbeek (2012),

$$(1) \qquad \frac{dE\{y_i|\mathrm{x}_i\}}{d\mathrm{x}_{ik}} * \frac{\mathrm{x}_{ik}}{E\{y_i|\mathrm{x}_i\}} = \frac{\mathrm{x}_{ik}}{\mathrm{x}_i'\beta}\beta_k$$

which produces the elasticity as a *constant* in a log linear model, β. This offers policy analysts the ability to easily interpret how percent changes in explanatory variables may correlate with percent changes in the dependent variable, holding all other things constant.

2.b. The MSA Level Dummy

The N-1 MSA level dummies control for the extent to which broader regional differences drive differences in home values and rents in this multi-state dataset. One of the major determinants of urban growth or decline in the United States after 1970 is the weather (Glaeser 2008). MSA level dummy variables are intended to capture this effect of weather impacting the growth of certain metropolitan areas. MSA level dummies also intended to capture the effect of agglomeration economies booming or imploding in certain economic sectors driving housing bubbles in specific cities, as with technology in Silicon Valley and banking in New York (Glaeser, Gyourko and Saiz 2008). MSA dummies may also capture the effect of the regional costs of construction and the effect of regional zoning and urban growth policies. Inclusion of these dummies ensures that the effect of density is not estimated arbitrarily across space in the United States. The baseline dummy is the first MSA in the compiled dataset: Birmingham, Alabama.

2.b. Housing Unit Characteristics

The housing unit characteristics vector, H, includes structure age, number of bedrooms, the total number of rooms in the unit and a dummy variable for if the property exceeded one acre in size. These variables have all been found to fit significantly in hedonic housing price models (Ottensman, Payton and Man 2008; Palmquist 1988; Rosen 1974). For the multi-family renter models, the data did not contain a variable to create the acreage dummy. For the multi-family models, two additional variables are included to capture variation in the size of multi-family structure: the number of units on the property and a dummy variable for large buildings, defined as buildings with over 50 units in the structure as defined by Blackley et al (Blackley et al 1986). Number of units and building age come from numeric, categorical variables in the PUMS dataset that cluster these variables arbitrarily. To address this, the author substituted the median value for each range of the variables. For example, a unit that is coded categorically as being five to ten years old is recoded as 7.5 years old.

These are the only variables used in hedonic housing price and rent research that exist in the PUMS dataset. The literature presents a wide array of other very detailed attributes to housing units and very specific amenities offered in rental complexes that have been found to correlate positively with housing prices and housing rents. Table Three presents a list of these variables and the recent papers which have found them to have a significant and positive impact on prices and rents:

Table Three: Overview of the Literature

Variable	**Literature**	**Sign**
Central Heating	Blackley 1986; Marks 1984	Positive
On Site Parking, Garage	Kestens et al 2006; Kim 1992; Marks 1984	Positive
Number of Baths	Blackley et al 1986; Kim 1992; Fisher Polakowski and Zabel 2009; Guntermann and Norrbin 1987	Positive
Porch Or Deck	Ottensman, Payton and Man 2008; Guntermann and Norrbin 1987	Positive
In Ground Pool	Marks 1984; Guntermann and Norrbin 1987	Positive
Fireplace	Ottensman, Payton and Man 2008; Kestens et al 2006; Kim 1992; Guntermann and Norrbin 1987	Positive
basement	Blackley et al 1986; Ottensman, Payton and Man 2008	Positive
Unit Floor Space	Ottensman, Payton and Man 2008; Guntermann and Norrbin	Positive

	1987	
Trees on Property	Kestens et al 2006;	Positive
Unit Has Air Conditioning	Kim 1992; Ottensman, Payton and Man 2008; Marks 1984	Positive

For this study, the impact of these variables on home values and rents will be included by the error term or absorbed by other coefficients. Of concern is whether the exclusion of these variables biases the population density coefficient in the model upwards or downwards and if so, by how much? This matters because ordinary least squares, the methodology used in this paper, requires that the error term be systematically uncorrelated, or exogenous, with variables included in the model:

$$(2) \qquad \mathrm{E}[\varepsilon|X] = 0$$

If these omitted variables, implicitly in the error term **ε**, correlate with our independent variables, listed above as **x**, this assumption is violated. Exclusion of unit floor space, garages and driveways, and pools are likely to push the population density coefficient downward towards zero. Failure to include them in the model means the model is not controlling for the fact units in less dense areas may be more likely to have pools, garages and larger floor spaces. Thus, the model may **under-estimate** the impact of density on housing because it does not account for these price-raising amenities which houses in less dense areas are theoretically more likely to contain.

Other variables in the literature that are omitted from this dataset, such as the type of heating and air conditioning and if the unit has a fireplace, may be a function of the age of a structure and structure type. Thus those variables' coefficients, which are included in the model, may be read as the effect of age of structure and all of its attributed impacts on the unit's utility for residents: such as whether or not it has a gas or electric stove. Beyond this, there is no reason to believe that households in denser areas are systematically more or less likely to have a fireplace or have electric heating instead of gas heating.

If valid justification is found in future research that attributes such as type of stove, type of heater and type of air conditioning system are in fact systematically correlated with population density, then the results presented in this paper need to be interpreted with greater caution.

2.b. Household Characteristics

Hedonic housing price models in the literature generally do not include characteristics of residents, but this model does because certain attributes of householders serve as effective proxies for unit attributes and as useful descriptors of neighborhood characteristics. The household characteristics matrix, H, consists of the household's annual income, number of persons, number of vehicles, a dummy variable for if the household is white or not, number of years spent living in that unit and a dummy variable equaling one for if the household included a commuter who took a fixed route transit system to work. The author considered the following mode choices applicable as fixed route transit: streetcar, subway and railroad. Using PUMS data, the author only needed to recode the years living in a unit from categorical to numeric using the same process described above for building age. This section will discuss this vector of variables and their limitations in more detail.

Household income is included to capture the impact of other utility-bearing attributes on the housing value or rent not already captured by existing unit attribute variables as housing is a normal good. This assumes that, all things being held constant, households with higher incomes will pay for better housing—which includes those things not captured in this model like a pool, porch or more floor space. Additionally, the household income may serve as an effective proxy for many immediate neighborhood level characteristics that affect housing values and rents like crime, school quality and urban blight. Household income works as a proxy for neighborhood quality because neighborhoods in the United States are highly segregated by income, and had become more so in the decades leading up to 2000

(Abramson, Tobin and VenderGoot 1995; Jargowsky 1996; Reardon and Bischoff 2011). Household income of tenants can thus capture many local effects on housing values not captured at the much large PUMA-wide level.

Similarly, the number of persons in the household is presumed to capture some of the missing effect of floor space in the model based on the assumption that, ceteris paribus, larger families will select into housing with more floor space. This implies that the number of persons in the family should correlate positively with floor space and thus housing values and rents. The number of vehicles in the household is also included in the model to serve a similar role. A household with two vehicles would be more willing to pay for a unit with a garage and driveway, whereas a household with no vehicles may be more accommodating to such a unit. Thus, the number of vehicles owned by the household is assumed to help positively explain home values and rents as proxy for the availability of parking for that housing unit.

Because this model relies so heavily on resident attributes as proxies for variables that normally explain home values, auxiliary models for home values and primary monthly mortgage payments are run, restricted just to those households who purchased within a year of providing information to survey gatherers. Additionally, the fact that home values are reported by respondents and not by professional appraisers suggests the benefit of restricting the home value market to recent buyers: presumably, the buyers will remember the price they paid on their homes less than a year prior to taking the survey. This precaution is not necessary for the much more flexible and demand-responsive rental markets.

Length of residency is a necessary variable to include because many rental units have agreements with tenants about the rate at which rents can rise when renewing contracts. Many cities also have rent control policies in place that prevent landlords from altering rents too much from year to year if a tenant remains in the same unit. This practice is referred to in urban economics as ‘soft’ rent control, and cities which implemented these policies as early as the 1970s retained them through the 1990s (Arnott 1995).

The dummy variable for whether or not the household is white captures the effects of redlining and discrimination in home valuation and lending practices throughout the latter half of 20^{th} century in the United States. The race of the household as a proxy for the race of the immediate neighborhood functions similarly as the income proxy did for area attributes due to the extreme racially segregated nature of residential housing in the United States (Massey and Denton 1993; Quillian 2012). This segregation continues to impact the ability of families of color to access better employment and housing opportunities (Gabriel and Painter 2012). This suggests housing values should be systematically higher in white communities, and that white individuals will thus live in units that are systematically of higher value. Even in the 1990s, the race or skin color of individuals living in a unit often contributed to appraisal decisions made by realtors in ways that adversely affected non-white residents (Louis 1997). However, other researchers using property value data from the Houston, Texas area failed to find support for the significance of race in housing values after controlling for local incomes (Holmes and James 1995). These differences could just be a function of 'white flight' and the implicit containment of communities of color to inner cities, but research from the Metropolitan Institute comparing white and non-white suburbs also finds racial differences also explain systematic differences in housing values in suburbs also (Anacker 2010). Using data similar to the one used in this study, Sykes finds that racial differences in housing values continued to exist from 1970 up to 2000 in the Integrated Public Use Micro Sample dataset, or IPUMS, with differences in home values between racial groups not closing significantly enough to disprove the importance of race in housing valuation by the year 2000 (Sykes 2008). Lastly, population density and racial segregation themselves may be related. Using time-series national data from 1980 to 2000, Pendall and Carruthers specify a quadratic relationship between the two variables in MSAs nationwide: segregation rises with density, levels off briefly and then declines at extremely high densities (Pendall and Carruthers 2003).

Lastly, the dummy variable for if the household contains a fixed route transit commuter captures the effect of proximity to transit on home values and rents. This variable may underestimate this effect because it does not capture households in proximity to transit with workers who may walk, bike or drive to work but still benefit from proximity to transit for use on non-work trips. Also, the variable will not capture the effect of proximity to transit on households without an employee commuting to work. In spite of this, prior research suggests this variable will positively impact housing values and rents. The Center for Neighborhood Technology (CNT) finds that while residential sales values declined from 2006 to 2011, they rose for units within proximity to transit during that same period (Becker et al, 2013). The CNT study only looked at five metropolitan areas: Boston, Chicago, Minneapolis-St. Paul, Phoenix and San Francisco and found the positive affect of proximity to transit on residential sales held regardless of if the transit option was heavy rail, light rail or bus rapid transit (BRT). Because the 2000 PUMS data did not differentiate between BRT and other bus service, the author opted to leave that mode out of inclusion of the fixed route transit commuter dummy.

2.d. Puma Level Variables

The matrix of PUMA level variables, N, captures sub-market and market level effects on housing costs. PUMA mean income is the weighted mean of all household incomes in the PUMA. For renter datasets, the authors' substitute this market attribute with the weighted mean income of all renters by PUMA, as renters compete with other renters for proximity to CBDs. PUMA density is the population per square mile, as discussed above. Finally, the model includes the average commute time for commuters in single and high occupancy vehicles at the PUMA level.

Inclusion of PUMA mean incomes captures sub-MSA regional characteristics much as individual household incomes may capture immediate neighborhood characteristics. More importantly, it ensures the PUMA-level population density variable does not represent other attributes of the PUMA that are themselves correlates of mean income. School quality and funding, by nature of the local tax systems used across the country, correlates positively with higher incomes (Chetty and Friedman 2005). Crime rates correlate positively with poverty, but also with communities wherein there is high income inequality

(Pratt 2001). Because reliable PUMA level data on these community attributes does not exist, the author includes PUMA mean incomes as a proxy for these community quality attributes based on this literature. Additionally, the author includes a measure for income inequality within PUMAs for two reasons that suggest significance of this variable. First, the criminology literature specifically notes that income inequality itself correlates with higher crime, the author includes a variable on PUMA level income inequality. This would lead to the assumption that an income inequality variable should correlate negatively with the dependent variables. Second, the author found in data exploration that PUMAs with income distributions skewed to the higher end of incomes had higher mean and median rents. This exploration suggested that PUMAs with greater income inequality, and therefore non-normally distributed incomes within the PUMA, would actually correlate with increases in the dependent variables. To capture these effects, this variable is specified as the difference between the household income at the 95th percentile of incomes in the PUMA subtracted from the income at the 20th percentile of incomes in the PUMA. Household weights were applied in the development of this variable. The author chose not to use standard deviations of income at the PUMA level because the pairwise correlation between PUMA weighted mean income and the standard deviation of the weighted mean income was .65. Econometricians generally consider pairwise correlations above .6 to raise issues of multicolinearity (Studenmund 2011). The pairwise correlation between PUMA weighted mean incomes and the inequality variable selected was only .4. With both variables in logarithmic form, that pairwise correlation dropped to .39.

4. Clustered Robust Standard Errors

This model requires clustered robust standard errors because a vector of independent variables in the model, vector L, are defined at the PUMA level and not the level of the household, which is the level at which observations are defined. Wooldrige (2003) lays out the case for robust clustered standard errors in situations where observations are pooled into non-overlapping groups—such as schools, employers, or

PUMAs—that exert their own effects on observations at that level. The following variance estimators demonstrate the difference between the traditional OLS estimator and one that is cluster robust:

$$V_{OLS} = s^2 i = 0n(\mathrm{x}'\mathrm{x})^{-1} \tag{3}$$

$$V_{Cluster} = (\mathrm{x}'\mathrm{x})^{-1} \sum_{j=0}^{n_c} u_j' u_j (\mathrm{x}'\mathrm{x})^{-1} \tag{4}$$

$$u_j = \sum_j e_i \mathrm{x}_i \tag{5}$$

Equation (3) shows the derivation of the variance in ordinary least squares where *s* is the variance of the residuals of the regression and X is a vector of independent variables. In equation (4), *u* is the cluster level variance of the residuals calculated in equation (5). Cluster-robust standard errors therefore punish our beta coefficients by calculating their standard errors by K-1 degrees of freedom instead of N-1, where K is the number of clusters.

D. Data Exploration: Transportation Variables, Housing Stock And Population Density

The only transportation data available in the PUMS dataset are commute times and commute modes. This section provides a brief overview on the relationship between these variables and population density, the primary dependent variable of interest, to explore if the assumptions about density and transportation explored in the literature hold with this data. The figures below show PUMA population density on the x-axis and weighted commuter mode splits by PUMA on the y-axis. In figure four, driver mode share is in blue, fixed route transit users in green, all transit users in black, and walk and bike mode share combined in red:

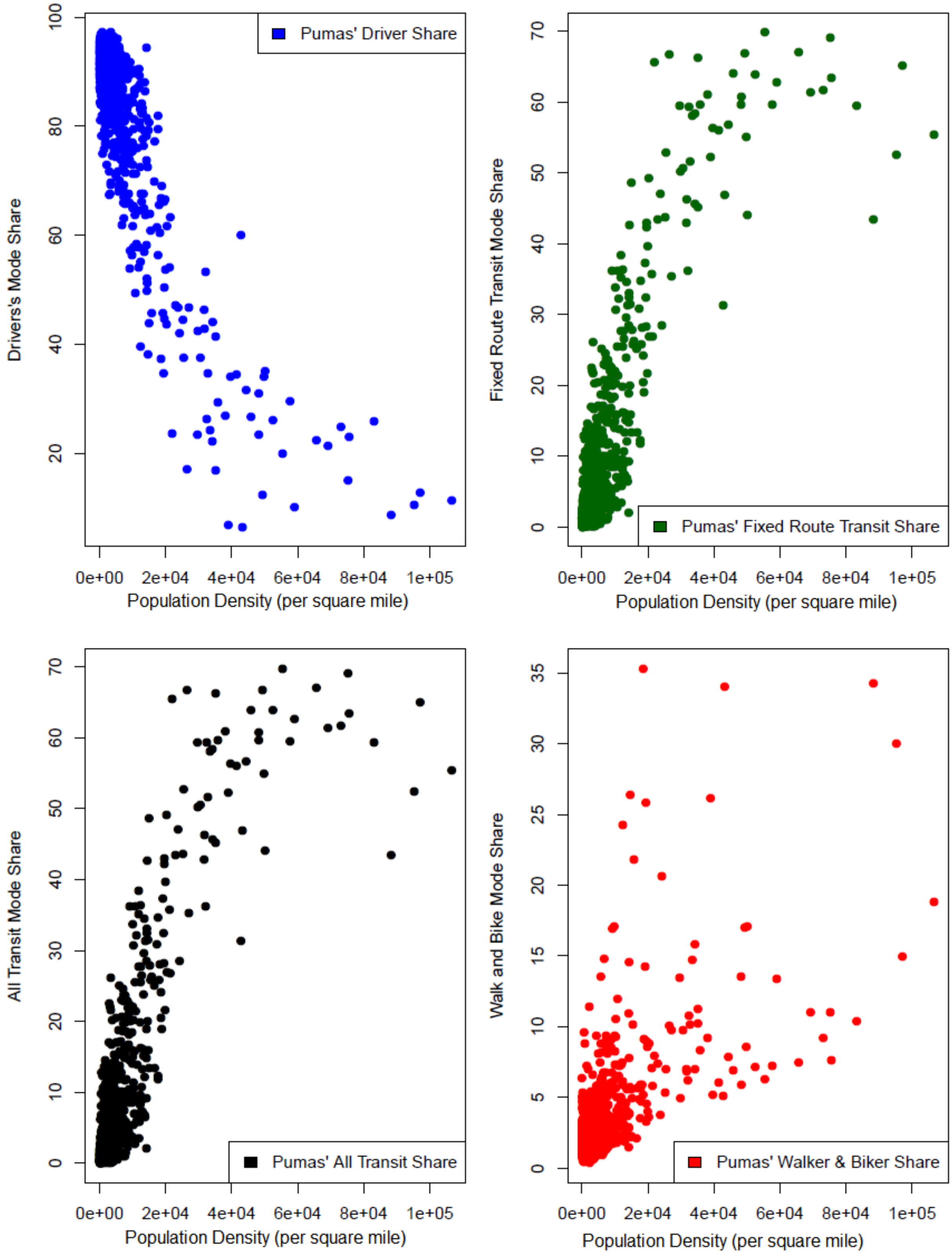

Figure 4: Puma Commuter Mode Splits By Density

The PUMS data confirms the literature on the negative relationship between population density and vehicle usage. As population density increases, vehicle mode share clearly declines. This does not necessarily mean that households in denser areas aren't driving for other purposes, but does show that this data is consistent with previous research. Transit mode share increases with density, and the relationship at the PUMA mode split level is clear. While a similar relationship appears for walking and biking and density, the relationship is highly heteroscedastic, particularly beyond a density of 10,000 persons per square mile.

The author also assessed the possible relationship between density and commute times. Figure five below shows PUMA mean commute times for drivers (in blue) and PUMA mean commute times for transit users (in green) plotted by PUMA density:

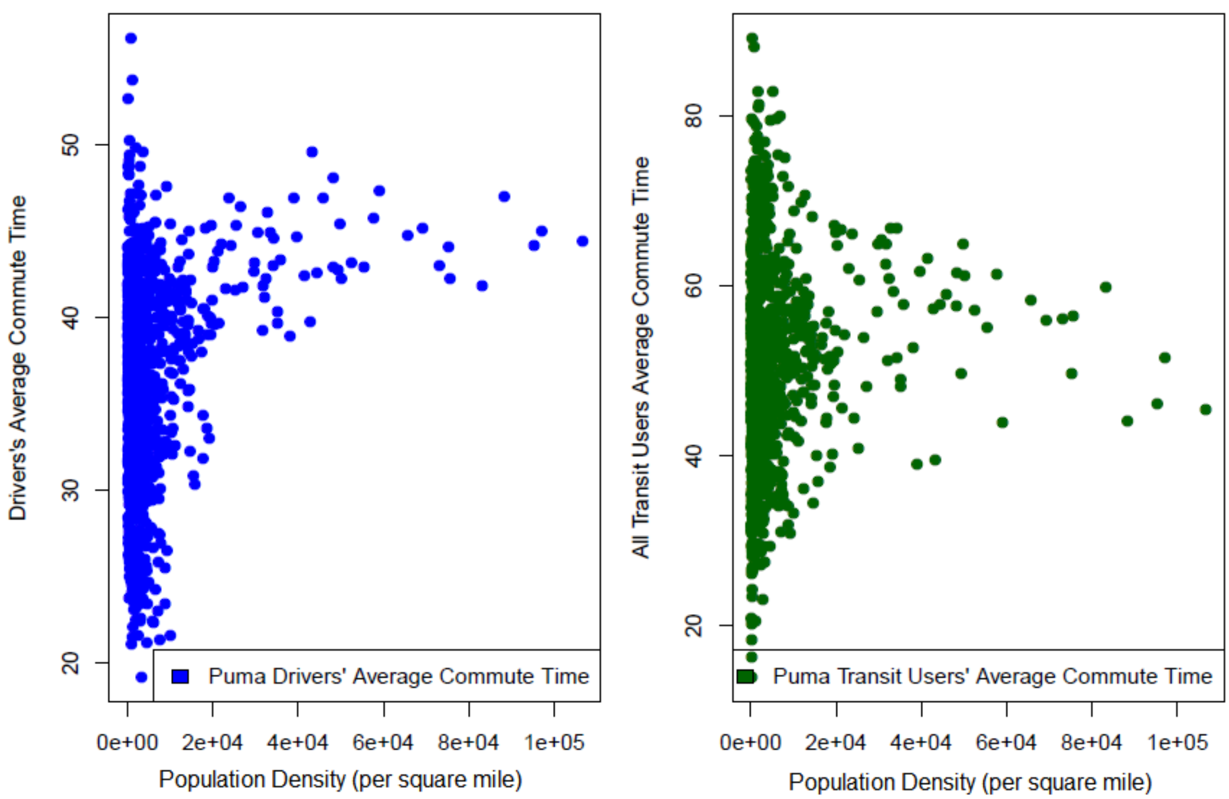

Figure 5: Puma Average Commute Times By Density and Modes

As population density increases, driver commute times appear uncorrelated up until a density of around 20,000 persons per square mile. Above that level of density, PUMA mean driver commuter times are fairly high, with most above forty minutes. Transit users' average commute times vary more dramatically across all levels of density. Although the author applied person weights from the 2000 PUMS data to improve the robustness of the sample estimates, the transit mean commute times may be more subject to error due to the small number of transit users surveyed in many PUMAs. But the data still confirms that commuters, as utility-maximizes, opt for the modes most convenient for their circumstances: commuters in high density areas with high driving commute times are more likely to take transit.

Because models are estimated separately (based on Chow Tests results) by structure type, tenure type and, in some cases, years of residence, care is needed in evaluating how this segmentation impacts the distribution of population density, PUMA average commute times and resident mode choices. Figure Six presents the cumulative density functions of population density for the four major segments used in this project: single-family renters, single-family owners, multi-family renters and multi-family owners:

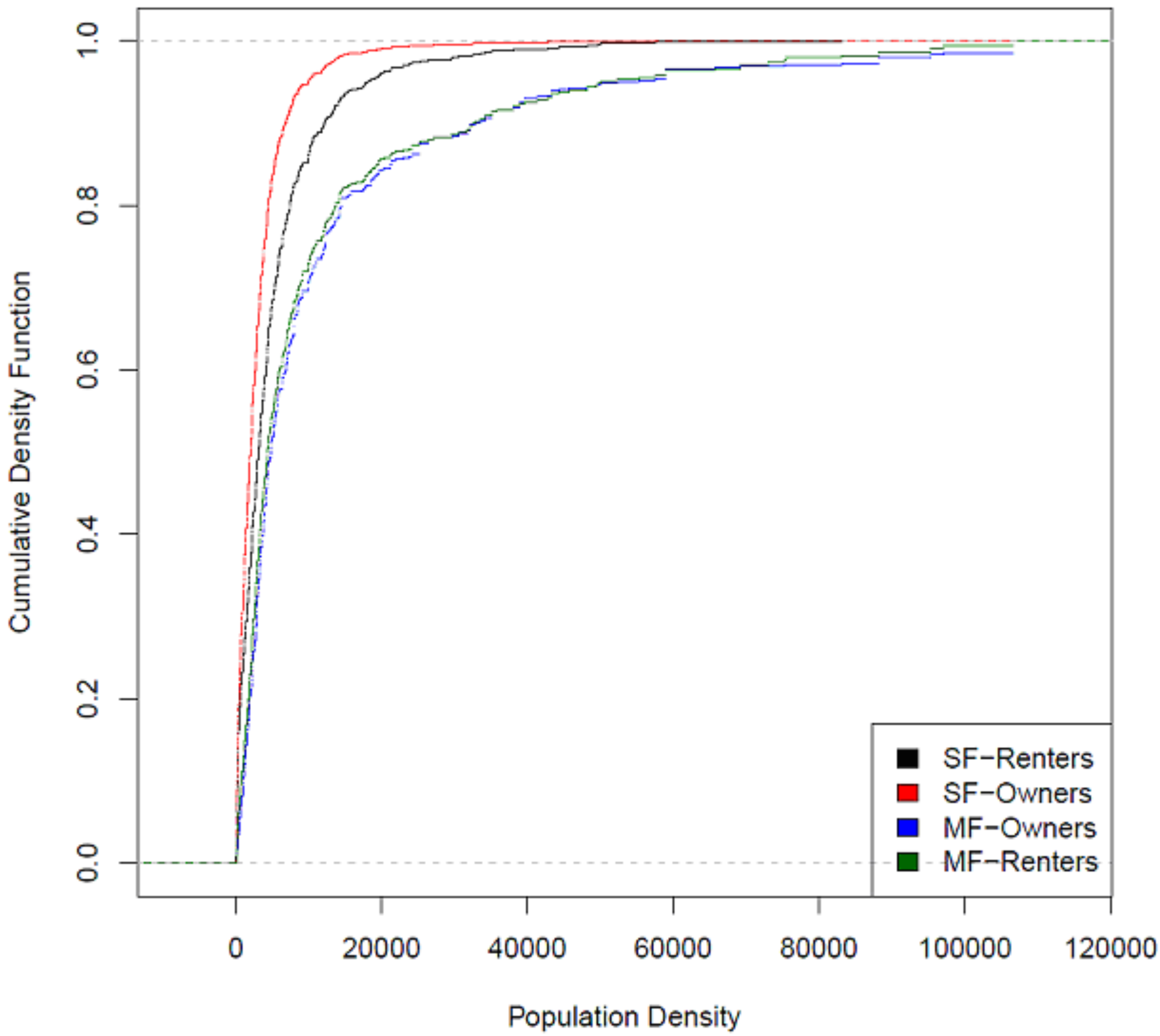

Figure 6: CDF Of Density by Segment

As would be expected by definition, multi-family units are more prominent in denser areas, so their cumulative density functions are not at as sharp from zero to twenty thousand residents per square mile. Within single-family unit dwellers, renters are skewed towards denser areas compared to home owners. The effect of population density on rents, unit values and mortgage payments may vary dramatically based on these different distributions within each segment.

Within these segments there is significant variation by years of residence in the unit. As stated earlier, home value and mortgage payment models are run first with the full samples and then just for recent home buyers (households that purchased their home within a year of taking the survey). This was done to attempt to capture the effect of density and commute variables' on home buyers whose purchases were made close to the times in which the density and commute data were collected, thus being more likely to reflect PUMA conditions at the time of the purchase. Figure seven shows the difference in the distributions of population density between recent home buyers and other home owners for single and multi-family segments:

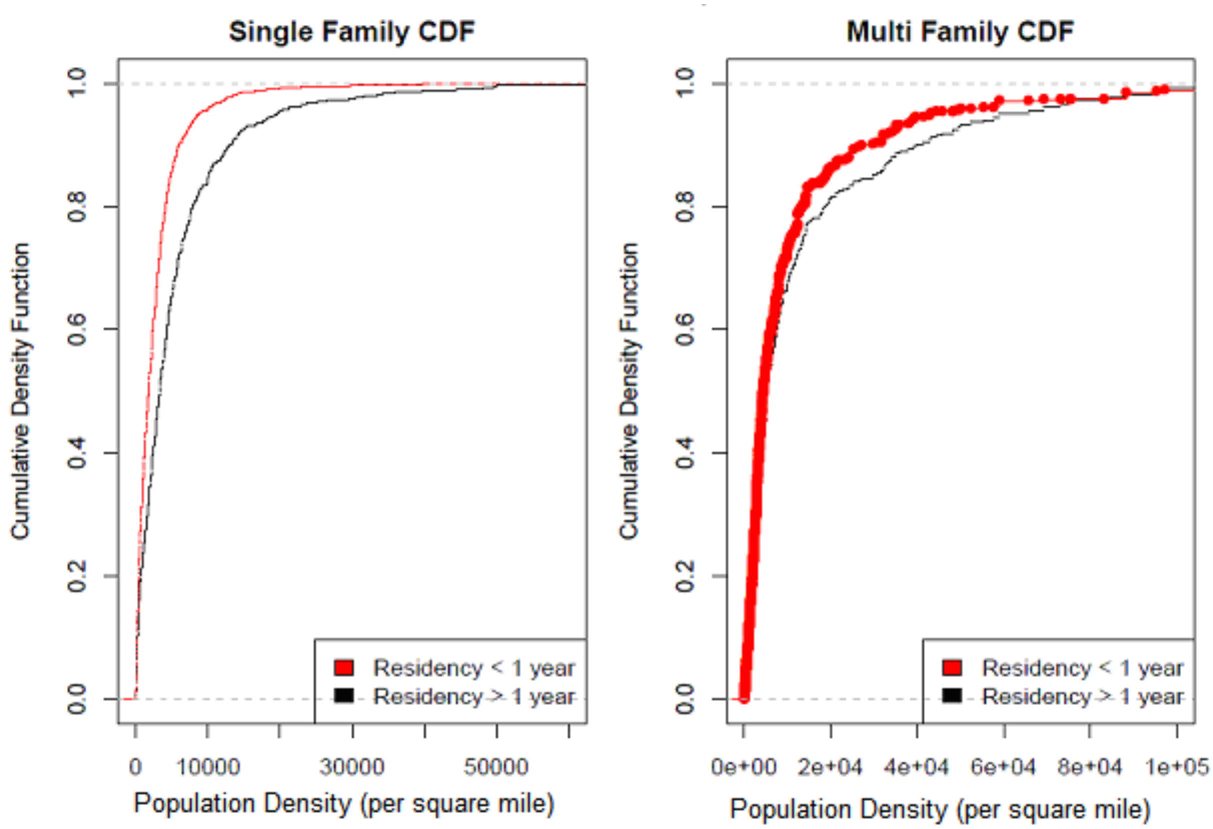

Figure 7: Differences in CDF By Years of Residence

In both cases, the cumulative density function rises faster for households who purchased their units within a year of taking the survey compared to those who purchased their units more than a year prior. If density does correlate or drive higher housing prices, then this result suggests that in the last year of the 1990s, home buyers were purchasing homes in areas that may have been systematically less dense than the country as a whole. This could lead to bias in the estimation of the effect of density on home values in models restricted just to these recent buyers.

E. Single Family Unit Models

1. Summary Statistics

This section reviews each segment of the single-family unit markets: renters, owners and owners still paying down their mortgages. The rental model has significantly lower incomes, with a mean income of around $45,000 compared to a mean of $80,836 for all home owners and $88,283 for home owners still paying down mortgages. In the case of the home owner models, those home owners who had purchased their home within the last year had higher incomes than the total group of home owners, with a mean of $86,375 in reported household incomes.

In terms of population density, the rental dataset had the highest mean population density at 5,289. This makes sense, as the rental markets tend to be larger in denser areas near urban cores. For both the Mortgage models and the Home Owner models, recent home buyers (with years of residence less than twelve months) were systematically living in less dense areas than their respective counterparts. T-tests confirmed a statistically significant difference in population density means between recent home owners and all home owners. T-tests also confirmed a statistically significant difference in population density means between those recent owners still paying down mortgages compared to other households paying down mortgages. This suggests that new home owners in the late 1990s were purchasing homes in less dense areas than the market as a whole.

Table 4: Single-Family Models Summary Statistics

	Household Income			Population Density		
	Mean	Min	Max	Mean	Min	Max
Rent	45,107	-20,000	969,000	5,289.39	8.95	106,612.48
Home Owner Model	80,836.90	-30,000	1,771,000	3,167.84	8.95	106,612.48
Home Owner Recent Buyer	86,375.90	-30,000	1,237,000	2,824.22	8.95	88,217.10
Mortgage Model	88,283.40	-20,996	1,771,000	3,101.25	8.95	106,612.48
Mortgage Model Recent Buyer	88,046.30	-15,000	1,237,000	2,824.22	8.95	88,217.10

	White			Building Age		
	Mean	Min	Max	Mean	Min	Max
Rent	0.518	0	1	43.29	0.5	75
Home Owner Model	0.734	0	1	35.86	0.5	75
Home Owner Recent Buyer	0.67	0	1	26.33	0.5	75
Mortgage Model	0.704	0	1	33.24	0.5	75
Mortgage Model Recent Buyer	0.66	0	1	26.14	0.5	75

The Rental model was only 52.8% white, while the full Home Owner model was 73.4% white and recent home owners were 67% white. This suggests that people buying homes in 1999 were systematically less white than the existing home-owning population at that time, and this was confirmed by T-test. The Rent model had, on average, an older housing stock than the Home Owner and Mortgage models.

Table 5: Single Family Models' Dependent Variables

	Dependent Variables		
	Mean	Min	Max
Gross Rent (Renters)	806.68	4	3200
Home Value (All Owners)	200,036	5000	1,200,000
Home value (Recent Buyers)	221,155.60	5000	1,200,000
Mortgage Payments (All)	1,117.35	1.5	5100
Mortgage Payments (Recent Buyers)	1,327.32	4	5100

Renters paid, on average, \$806 a month for singly family units. Home owners paid, on average, \$1,117 a month for single family units. The mean home value for single family units was \$200,036, but was \$221,155 for recent home buyers.

The rental housing units were, on average, older, and their residents were less likely to be white and were of lower income than were residents who owned their own homes. Owners who had purchased their homes within twelve months of taking the long form Census were, on average, less white and of higher income than the broader population of home owners. They had also purchased homes in areas that were, on average, less dense than the areas inhabited by the larger home-owning segment.

2. Models Estimation and Discussion

The model coefficients cannot be compared across models beyond naïve analysis, but consistency in signs can be interpreted more strongly. Across all models, population density, the variable of interest, correlates positively and significantly with housing rents, monthly payments on primary mortgages and housing unit values. A one hundred percent increase in density increases rents by 4.7%, while it increases housing values by 10.8% with the same 100% increase in density. For mortgage payments, a 100% increase in density correlates with a 5.9% increase in mortgage payments for all households paying down mortgages, and 5.6% for those purchased their homes within twelve months of taking the Census, ceteris paribus.

As shown in Table 6, two separate models of single family rents were run to explore possible systematic biases in data availability. Nearly a quarter of renters did not report if their unit rested on a property greater than one acre or greater than ten acres. T-tests confirmed that the households which did not provide this information were systematically less white, of lower income, more likely to take fixed route transit and lived in denser areas, on average. It is possible that survey distributors in some areas may have excluded this question from their forms knowing no units with acreage existed, but the author could not confirm this in the records. Regardless, the density coefficients remain positive and strongly significant in both model runs.

Table 6: Single Family Rent Models

	Single Family Rent		Single Family Rent W/Land Dummies	
	b	**cluster robust t-stat**	**b**	**cluster robust t-stat**
Log Bedrooms	0.159	24.81	0.135	18.77
Log Rooms	0.182	34.08	0.19	29.88
Log Building Age	-0.053	-24.99	-0.059	-25.01
Log Years There	-0.055	-54.81	-0.057	-50.06
Land 1-9.9 Acres			-0.018	-8.88
Land 10 Plus Acres			-0.057	-12.3
Log Income	0.268	31.3	0.029	28.07
Log Residents	0.059	25.14	0.059	21.94
Log Vehicles	0.158	48.26	0.147	39.34
White Household	0.034	20.78	0.036	20.58
Fixed Route Commuter	**0.039**	**5.79**	**0.004**	**0.36**
Log Density	**0.047**	**20.92**	**0.047**	**18.07**
Log Puma Mean Inc				
Log Puma Mean Renter Inc	0.527	37.49	0.055	33.34
Log Income Difference	0.068	5.88	0.063	4.64
Log Puma Drivers Ave Commute	**-0.014**	**-0.457**	**-0.057**	**-1.51**
N	**220,676**		**144,795**	
K (Clusters)	**1015**		**1014**	
R Squared	**0.4122**		**0.4175**	
Adjusted R Squared	**0.412**		**0.4171**	

As expected from the literature, the number of bedrooms and rooms correlated positively with the log of gross rents, as did the number of residents, household income and the number of vehicles. The land dummies were negative and significant, along with the log of years there and log of the unit's age.

Mean renter income correlated positively with gross rents, as did the log difference between the PUMA 20th percentile income and 95th percentile income. This suggests that for every 100% increase in the difference between a PUMA's 20th percentile of income and 95th percentile of income, rents rise by 6.8%. This confirms the author's hypothesis that PUMAs with income distributions skewed towards higher incomes in fact face higher housing costs in the rental markets. If this inequality also drives increased crime, that effect does not influence rents in this data or is completely mitigated by the other impacts hypothesized income inequality would have on prices. The natural log of the PUMA drivers' average commute time correlated negatively with gross rent, but results were not statistically significant.

The fixed route commuter dummy was positive and significant for the model that excluded the parcel size dummies, producing an estimate that households with a fixed route commuter spent 3.9% more on rent compared to households without a fixed route commuter. This variable become insignificant in the land-dummy included model, which the author attributes to the fact that the quarter of observations lost by that inclusion were more likely to have a fixed route commuter, ceteris paribus. Running the same model with land dummies excluded on the groups with and without them, and then pooled as a Chow Test, found systematic differences in results produced by the two different groups. If the two pools of data running identical models are found to be systematically different, then coefficients estimated under different model specifications cannot be compared across pools or said to be equivalent even though both models appear similar.

Chow tests performed on the other single family models also found significant differences across models, with significant p-values for Chow test statistics on differences between all home owners and those who purchased their units within a year of taking the Census. This means the Chow test identified structural differences in the data pooled by years of residency when the divide was set at one year. As a result, both models were estimated separately to evaluate the impact of density on housing values and the impact of density on the housing market in 1999-2000.

Table 7: Single Family Home Owner Models

	Housing Value, All owners		Housing Value, Recent Buyer	
	b	cluster robust t-stat	b	cluster robust t-stat
Log Bedrooms	0.184	12.03	0.175	8.11
Log Rooms	0.743	50.28	0.641	37.02
Log Building Age	-0.15	-30.74	-0.099	-34.18
Log Years There	-0.042	-21.34		
Land 1-9.9 Acres	0.124	32.16	0.103	21.89
Land 10 Plus Acres	0.31	46.34	0.251	21.92
Log Income	0.102	53.31	0.117	29.52
Log Residents	-0.079	-18.98	-0.056	-8.4
Log Vehicles	0.221	29.62	0.2	16.64
White Household	0.124	24.53	0.092	21.08
Fixed Route Commuter	**0.068**	**7.32**	**0.069**	**5.97**
Log Density	**0.108**	**11.85**	**0.084**	**9.8**
Log Puma Mean Inc	1.18	25.8	1.02	24.44
Log Puma Mean Renter Inc				
Log Income Difference	-0.064	-1.59	-0.02	-0.043
Log Puma Drivers Ave Commute	**-0.154**	**-1.23**	**-0.448**	**-4.15**
N	**1,183,391**		**109,902**	
K (Clusters)	**1015**		**1010**	
R Squared	**0.4988**		**0.4856**	
Adjusted R Squared	**0.4987**		**0.4852**	

Log number of bedrooms, rooms, vehicles and acre dummies correlated positively with housing values, as anticipated in the literature. The number of years a household lived in the unit correlated negatively with home values, as did the unit's age. The number of residents actually correlated negatively and significantly with housing values, contrary to expectations. White households lived in units that were, on average, 12.4% higher in value than units owned by non-White households. In the recent buyer market, that difference was 9.2% higher.

Mean incomes correlated positively with housing values, but were elastic, with an 100% increase in mean incomes resulting in a 118% increase in home values. Population density was positively related to price, with a 100% increase in density leading to a 10.8% increase in home values for all units and an 8.4% increase for units of recent home buyers.

Households with at least one fixed route transit commuter lived in units that were, on average, 6.8% higher in value than the units of those without at least one fixed route transit commuter in the household. This was true both for the model which included all home owners and the model with just recent home buyers.

Both models found the relationship between PUMA drivers' average commute times and housing values to be negative, but this result was only significant for the recent home buyer model. The author does not interpret this as an invalidation of the housing and transportation cost trade off theory. Families that purchased their homes years before taking the survey cannot be interpreted as not making a housing and transportation cost trade off, because the transportation cost variable provided here, the Puma drivers' average commute time, reflects transportation costs in 2000 and not in the decades in which they purchased their homes. But this logic does support the idea that the commute time variables can represent transportation costs in the area relative to other areas that were presented to home buyers who opted to buy homes less than a year from taking the survey. One can think of many circumstances in which this could not be the case, such as residents locating on one side of a major bridge to cut down commute times and then the bridge then collapsing between the time of their housing purchase and the time that they took the survey. There is no reason to believe that such scenarios are systematic enough in the data—or took place frequently enough in the year 2000—to warrant concern.

Table 8: Single Family Mortgage Models

	Mortgage Payments, All Payers	Mortgage Payments, Recent Buyers

	b	cluster robust t-stat	b	cluster robust t-stat
Log Bedrooms	0.08	9.88	0.124	10.41
Log Rooms	0.41	59.39	0.391	39.33
Log Building Age	-0.001	-27.13	-0.001	-22.68
Log Years There	-0.111	-88.3		
Land 1-9.9 Acres	0.063	27.59	0.051	17.1
Land 10 Plus Acres	0.125	30.9	0.104	13.01
Log Income	0.092	55.38	0.1	29.97
Log Residents	0.03	14.15	0.006	1.56
Log Vehicles	0.051	17.55	0.053	8.83
White Household	0.02	11.06	0.024	11.04
Fixed Route Commuter	**0.031**	**6.3**	**0.056**	**6.26**
Log Density	**0.059**	**14.67**	**0.056**	**11.55**
Log Puma Mean Inc	0.578	33.31	0.593	27.94
Log Puma Mean Renter Inc				
Log Income Difference	0.043	2.73	0.044	2.4
Log Puma Drivers Ave Commute	**-0.115**	**-1.85**	**-0.288**	**-4.18**
N	**864,837**		**100,281**	
K (Clusters)	**1015**		**1010**	
R Squared	**0.4582**		**0.4843**	
Adjusted R Squared	**0.4582**		**0.4843**	

The mortgage payment models produced the same signs for coefficients compared to those of the housing value models, but these could not be tested for statistical equivalency and thus the comparison cannot extend further. PUMA mean incomes were significant and positively correlated with household monthly mortgage payments, but the relationship was inelastic in both models.

White households paid two percent more on their mortgages compared to non-white households. This difference was two and a half percent among recent home buyers. The much larger difference in home values for white and non-white households versus the small difference in mortgage payments between the two groups may reflect the lingering effect of discrimination in mortgage lending practices. However, such conclusions cannot be drawn because the coefficients in these models cannot be compared statistically given their different dependent variables. Future research should identify a study method that would enable such an analysis to be performed on this data.

Households with fixed route transit commuters paid 3.1% more than households without a fixed route transit commuter. Just among recent home buyers, the difference was 5.6%. A one hundred percent increase in population density correlated with a 5.6% increase in mortgage payments for the entire sample, and a 5.9% increase for recent home buyers. The log of PUMA drivers' average commute times correlated negatively with monthly mortgage payments. The coefficient was significant at the 10% level for the complete sample, and greater than 1% significance for recent home buyers. A one hundred percent increase in the PUMA drivers' average commute time correlated with an 11.5% drop in monthly mortgage payments for the full model and a 28.8% drop in mortgage payments for recent home buyers. These results, and the results on the fixed route transit dummy, can be interpreted as supporting the hypothesis that households tradeoff between housing and transportation costs.

3. Model Diagnostics and Discussion

The following plots included in this section were produced using R to identify the strength of the OLS models estimated in the previous section. This section begins with an explanation of the figures provided to determine whether the models conformed to the assumptions of ordinary least squares, followed by exploration of the results of diagnostics for each model.

Diagnostics were produced in R via the "plot" function for regression models, in which the saved models results are plotted. The diagnostics come in the form of four graphs which each explore different aspects of the models. The first graph, in the top left of each diagnostic figure, plots each observation with the model's fitted values for the observation on the x-axis and the observations' residual on the y-axis. The graph shows a trend line in red between these two points for all observations in the model. A flat trend line suggests homoscedasticity, a key assumption in OLS. The second graph in the top right plots the predicted standardized residuals of a perfectly normal distribution against the actual observed standardized residuals. The graph provides a 45% line on which these two distributions would lie if they were identical. Thus, any deviation of data points from this dashed line suggests non-normal distribution of the residuals. This raises concerns about possible violation of the normality assumption of the error term in OLS.

The bottom left graph shows the fitted values of each observation against the square root of the standardized residuals of each observation. Similar to the first graph in the top left, any systematic difference across the x-axis in standardized residuals suggests a violation of the assumption of homoscedasticity. Finally, the bottom right plot, titled "Residuals vs Leverage," plots the leverage each observation exhibits on the model on the x-axis against the standardized residuals of those observations on the y-axis. This allows for mapping of Cook's Distance, marked by a dashed red line, which is the point at which an observation is said to be influential in the regression.

The author notes that the use of cluster-robust standard errors automatically corrects for any inflated significance granted to coefficients as a result of heteroscedasticity. So while many of the results from diagnostics may show concern for violation of the assumption of homoscedasticity, this has been corrected in the production of standard errors and t-stats. This correction will not show up in the plots provided below.

Figure 8 shows the model diagnostics with the full single family rent sample in which the land dummies were excluded from the model. In Figure 8 the relationship between residuals and fitted values is very close flat, with the appearance of some heteroscedasticity at either end of the distribution of fitted values. This is not of concern, as the use of cluster-robust standard errors was an especially punitive control for how such heterscedasticity might artificially inflate significance (Verbeek 2012). . The results of the normal Q-Q plot suggest that at one end of the distribution the residuals are not normally distributed. Again, the lack of any relationship or discernible pattern in the relationship lends confidence to the model following the assumptions of OLS. Lastly, the bottom right plot of the residuals against each observation's leverage is a measure of potentially influential observations. The three household points labeled by R are labeled because they can be defined statistically as outliers. None of them appear to hold leverage in the model warranting concern.

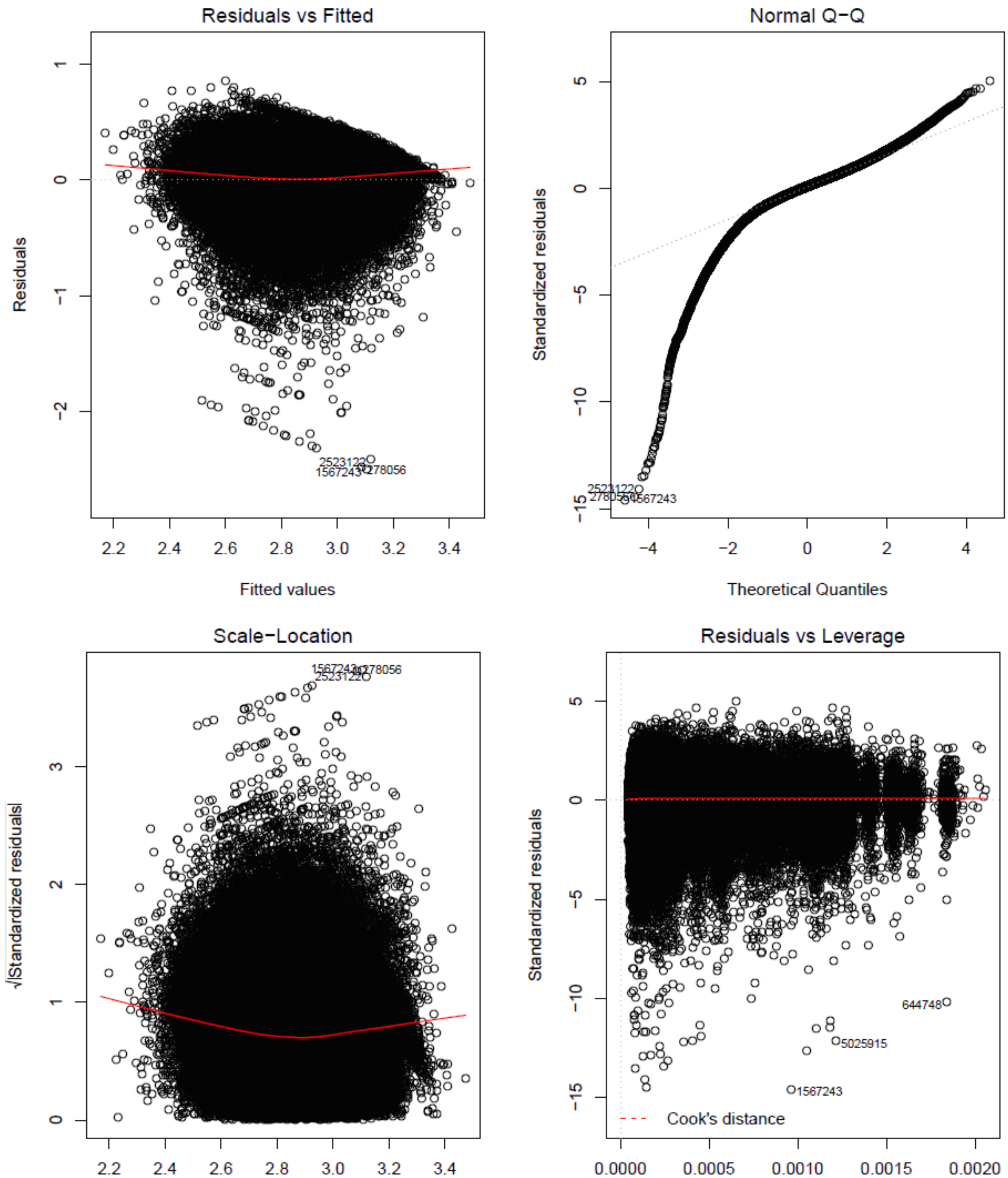

Figure 8: Single Family Rent Model Diagnostics

Figure nine shows the same diagnostics for the single family unit value dataset for recent home buyers.

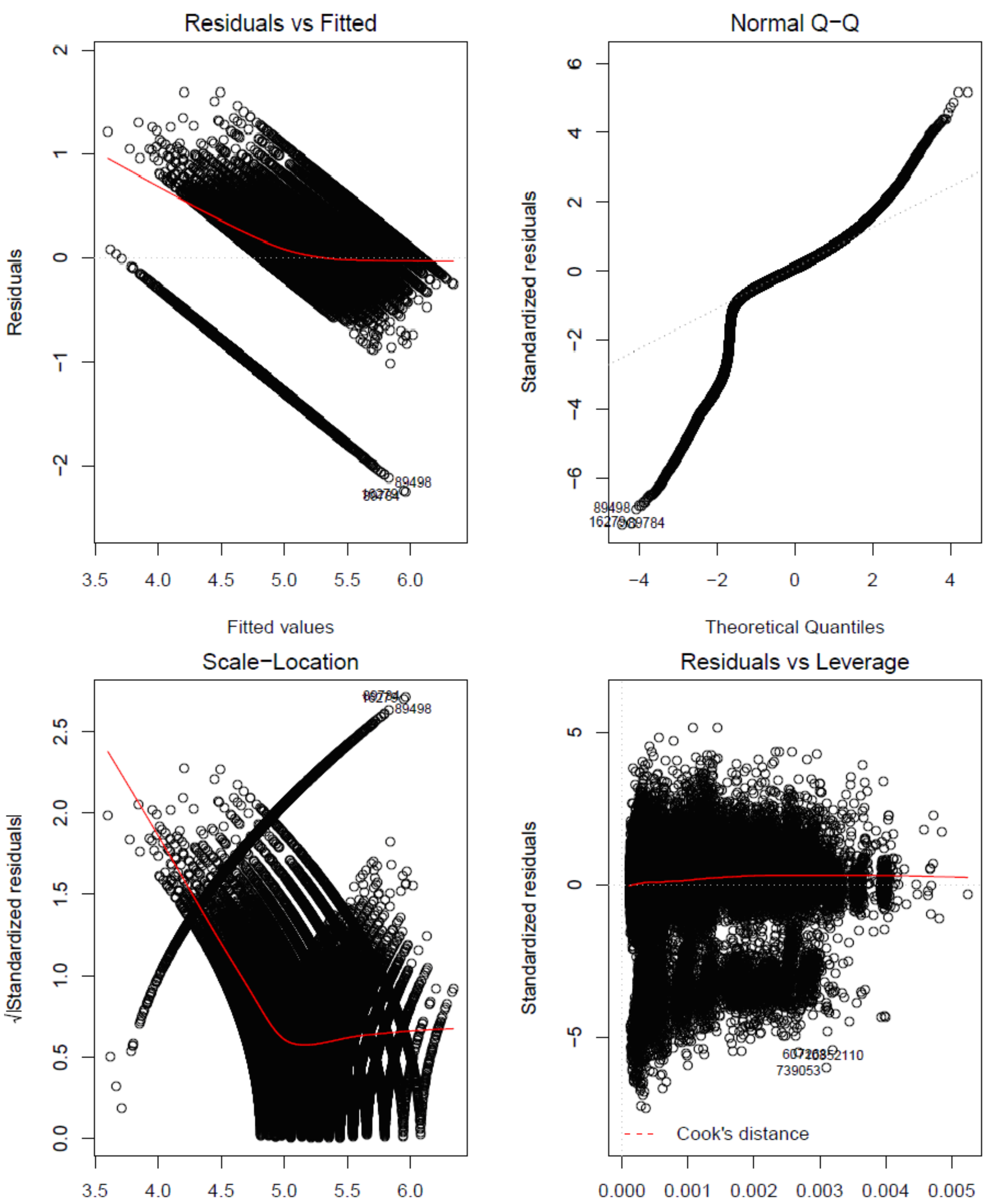

Figure 9: Single Family Unit Value Models, Recent Buyers

The scale-location and residuals versus fitted values plots demonstrate severe and confusing heterscedasticity. The author suspects this is due to the original categorical nature of the housing value data. The attempt to use the mean value of each category evidently failed to hold up to the assumptions necessary for Ordinary Least Squares to be BLUE, the Best Linear Unbiased Estimator. These results, coupled with the concerns about non-normality of the residuals presented in the Normal Q-Q plot in the top right, suggest the model does a particularly poor job in handling home values at the lower end of the distribution. The non-normality assumption being violated, this model is evidently invalid.

In cases in which the dependent variable is categorical, or in this case ordinal, Ordinal Logistic Regression might provide a better choice for analysis of the housing value data. Unfortunately, the categorical variables in the data

Figure Ten shows the same diagnostics for the single family primary monthly mortgage payment dataset for recent home buyers.

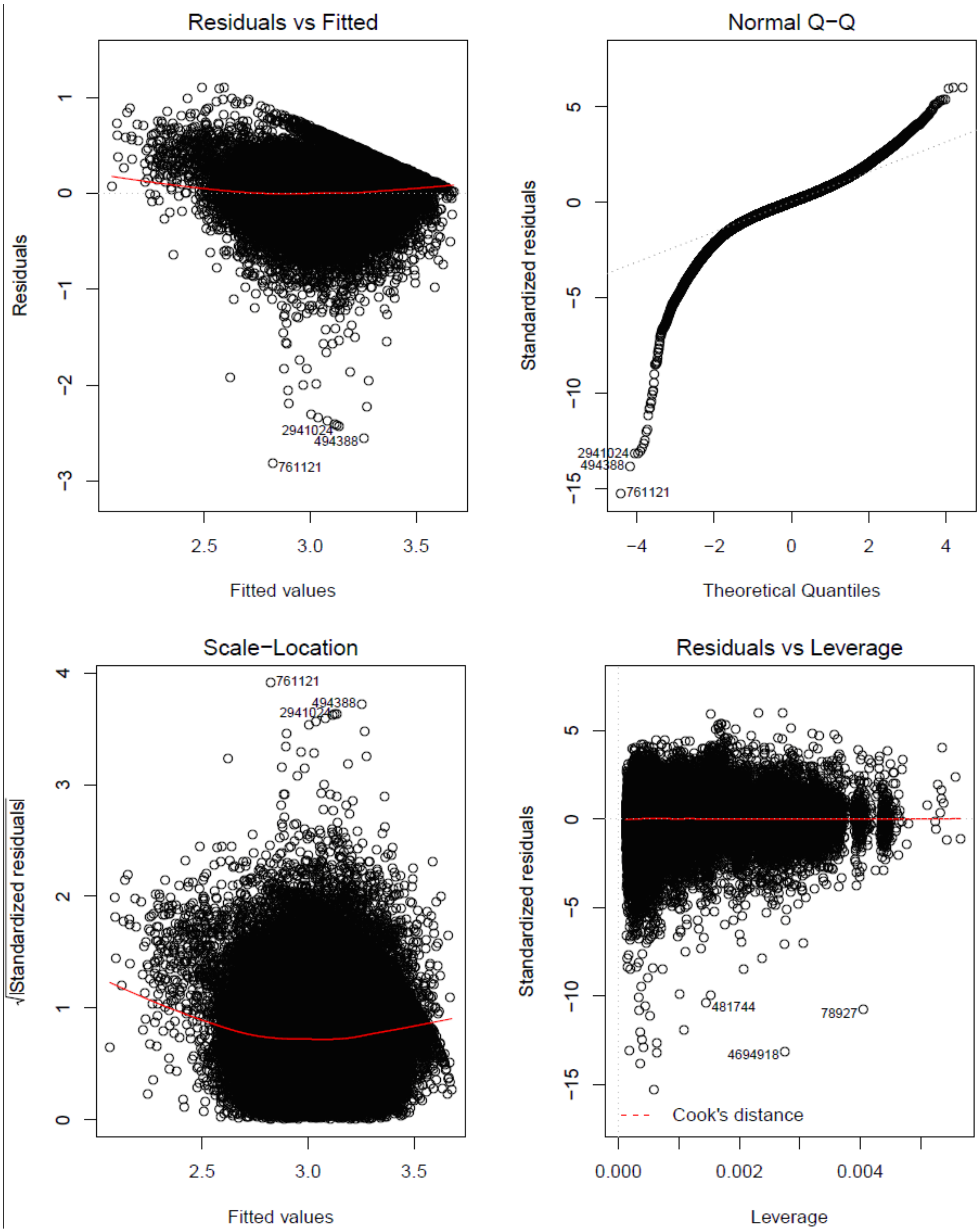

Figure 10: Single Family Mortgage Payment Diagnostics, Recent Buyers

In the fitted values versus residuals plot on the top right, there is the appearance of a relationship between the two points at the upper end of the fitted values. The author was unable to discern a reason for this pattern, but any heterscedasticity is controlled for via use of cluster-robust standard errors. The top right chart indicates that the model also suffers from a non-normally distributed lower end of the distribution of residuals, but that this is less severe of a problem in this model than in the renter model. The scale location plot suggests some heteroscedasticity at either end of the distribution of fitted values, but this has already been controlled for using cluster-robust standard errors. The residual versus leverage plot demonstrates that while no particular outlier is individually influential, there is a cluster of households with extreme standardized residuals towards the lower end of the chart. The models do not appear to do well in fitting data at the lower end of dependent variable distribution. This may be due to the unreliability of some respondents who could not accurately assess the values of their homes, and of the effects of special kinds of subsidized mortgages that individuals at the lower end of the distribution had access to but which there was no data available on.

The author ran variance-inflation-factor (VIF) tests for multicolinearity in each of the models and observed severe multicolinearity in the MSA/CMSA dummy variables. The VIF test regresses each independent variable on all the other independent variables in the model. Each of the R-squared values for those regressions is subtracted from one. The VIF score is one divided by those numbers for each observation. As such, the smaller the unexplained variance in the independent variable of interest not explained by other independent variables, the higher the VIF score. For all other variables, VIF scores did not rise above two. When VIF scores are only to be a concern when they rise above five, a score under two thus indicates multicolinearity is not a problem (Studenmund 2011). .

F. Multi-Family Unit Models

1. *Summary Statistics*

The multi-family data exhibited patterns similar to the single family one, with renters having lower incomes and a lower mean value for the white household dummy variable. Renters also had a higher mean value for building age, and this was found to be a statistically significant difference with all multi-family unit owners via T-test. As with the single family data, there were statistically significant differences between recent buyers and all home owners in the multi-family dataset. Table X presents the differences in density between recent buyers and all owners, and T-tests performed confirmed significance in the difference in means.

Table 11: Multi-Family Summary Statistics

	Household Income			Population Density		
	Mean	**Min**	**Max**	**Mean**	**Min**	**Max**
Rent	38,114	-20,000	1,620,560	11,419.59	8.95	1,066,612.00
Home Owner Model	68,306.64	-20,000	1,516,000	12,562.08	8.95	1,066,612.00
Home Owner Recent Buyer	69,188.10	-11,500	845,000	10,924.79	8.95	1,066,612.00
Mortgage Model	73,567.62	-20,000	1,516,000	12,174.65	8.95	1,066,612.00
Mortgage Model Recent Buyer	72,476.55	-11,500	764,000	11,028.10	8.95	1,066,612.00

*Density results significantly different between one years and Alls

	White			Building Age		
	Mean	**Min**	**Max**	**Mean**	**Min**	**Max**
Rent	0.47	0	1	34.52	0.5	75
Home Owner Model	0.73	0	1	32.63	0.5	75
Home Owner Recent Buyer	0.681	0	1	28.48	0.5	75
Mortgage Model	0.681	0	1	32.16	0.5	75
Mortgage Model Recent Buyer	0.658	0	1	28.72	0.5	75

Additionally, the difference in mean white households, or the share of households that were entirely white, was significantly different between recent home buyers and total home owners. Recent home buying households were less white. Also, as would be expected, the units of recent home buyers were, on average, younger than those of all home owners. That difference was also confirmed by T-test.

Table 12: Multi-Family Models' Dependent Variables

	Dependent Variables		
	Mean	**Min**	**Max**
Gross Rent (Renters)	689.27	4	3092
Home Value (All Owners)	174,130	5000	1,200,000
Home value (Recent Buyers)	173,784.60	5000	1,200,000
Mortgage Payments (All)	924.32	1.66	5000
Mortgage Payments (Recent Buyers)	1,029.95	4	5000

Multi-family unit renters reported an average rent of $689.27 a month, while individuals with mortgages were paying $924.32 on average with taxes included and insurance excluded. Recent home buyers paid more than the total group of home owners in monthly mortgages and the difference was found to be significant via T-test.

2. *Models Estimation and Discussion*

All the models in the multi-family tests produced weaker R-squared and adjusted R-squared scores than their complimentary single-family models. This means the models were less effective in explaining housing cost variables for multi-family units. Based on R-squared values, the multi-family renter model performed the best, explaining 36.07% of the variance in gross rents. The weakest performing model, for home values of households who had purchased their homes within the last year, explained 32.31% of the variance in the home values for recent home buyers.

Table 13: Multi-Family Rent Models

	Multi Family Rent	
	b	**cluster robust t-stat**
Log Bedrooms	0.13	16.93
Log Rooms	0.063	12.52
Log Building Age	-0.044	-16.21
Log Units	-0.022	-9.44

Log Years There	-0.075	-27.9
Log Income	0.036	28.78
Log Residents	0.035	10.32
Log Vehicles	0.284	50.16
White Household	0.047	16.85
Fixed Route Commuter	**0.083**	**16.13**
Log Density	**0.064**	**21.12**
Log Puma Mean Inc		
Log Puma Mean Renter Inc	0.636	33.71
Log Income Difference	-0.005	-0.036
Log Puma Drivers Ave Commute	**0.116**	**3.05**
N	**511,146**	
K (Clusters)	**1015**	
R Squared	**0.3607**	
Adjusted R Squared	**0.3606**	

The multi-family rental model upends several expectations. First, PUMA driver average commute times correlated positively and significantly with rents, not negatively as prior research would suggest. This may be due to omitted variable bias, as some things which correlate with the commute time variable may impact the decisions of those who choose to rent multi-family units in ways that are different from how they affect the decisions of those attracted to single family units. Also, the distribution of population density in this data pool is significantly different from the single family rents, and the correlation between congestion and driving commute times may also be skewing this result. The disparity in incomes at the PUMA level also correlated negatively with rents, but the results were not significant. Number of bedrooms, rooms, household income, number of residents and number of vehicles correlated positively with rents as anticipated. Building age, number of units in the building and the number of years the household lived in the unit correlated negatively with rents, as anticipated. A one-hundred percent increase in population density correlated with a 6.4% increase in rents for households in multi-family units. Multi-family unit renters with at least one fixed route transit commuter living in the household paid 8.3% more on their rents, ceteris paribus, than households without a fixed route transit commuter.

Table 14: Multi-Family Unit Value Models

	Housing Value, All owners		Housing Value, Recent Buyer	
	b	cluster robust t-stat	b	cluster robust t-stat
Log Bedrooms	0.632	14.69	0.696	10.27
Log Rooms	0.347	11.7	0.081	1.7
Log Building Age	-0.138	-8.81	-0.144	-11.12
Log Units	-0.036	-1.94	0	0.05
Log Years There	-0.03	-6.01		
Log Income	0.079	14.1	0.085	8.23
Log Residents	0.092	6.61	0.068	2.52
Log Vehicles	0.21	6.71	0.183	4.3
White Household	0.052	3.43	0.084	5.17
Fixed Route Commuter	**0.01**	**0.56**	**0.078**	**1.981**
Log Density	**0.289**	**12.18**	**0.285**	**11.28**
Log Puma Mean Inc	0.729	5.23	0.614	4.28
Log Puma Mean Renter Inc				
Log Income Difference	0.432	3.27	0.385	2.96
Log Puma Drivers Ave Commute	**0.175**	**0.52**	**-0.144**	**-0.42**
N	**86,153**		**12,938**	
K (Clusters)	**1013**		**951**	
R Squared	**0.3343**		**0.3231**	
Adjusted R Squared	**0.3337**		**0.3169**	

In the unit value models, the number of bedrooms, rooms, household income, years of residence, number of vehicles, PUMA mean income, the PUMA income inequality measure and number of residents correlated positively with housing values as anticipated. The age of the structure and the number of years the household resided in the unit correlated negatively with housing unit values. White households lived units that were 5.2% higher in value than non-white households. For recent home-buyers, this difference was similarly positive at 8.4%. PUMA average commute times did not correlate significantly with unit values.

A one-hundred percent increase in density led to a 28.9% increase in unit value among all owned multi-family units, and resulted in a 28.5% increase in unit value for those units purchased within a year of the Census. No statistically significant difference was found between the values of units in households with a fixed route transit commuter and those without for the full sample. Among units purchased within a year of respondents taking the survey, units with at least one fixed route transit commuter were 7.8% higher in value than those without a fixed route transit commuter. The author attributes this difference to the fact that the former group is more likely to contain households in close proximity to transit but without any workers at all. In those cases units that experience the same effect of what the fixed route transit dummy is attempting to capture are coded as not having that effect when they do, causing problems with the variable.

Table 15: Multi-Family Mortgage Models

	Mortgage Payments, All Payers		**Mortgage Payments, Recent Buyers**	
	b	**cluster robust t-stat**	**b**	**cluster robust t-stat**
Log Bedrooms	0.629	15.65	0.66	9.73
Log Rooms	0.257	9.58	0.077	1.52
Log Building Age	-0.11	-8.27	-0.118	-9.77
Log Units	-0.049	-3.29	-0.022	-1.4
Log Years There	-0.044	-9.07		
Log Income	0.095	14.15	0.101	8.24
Log Residents	0.088	6.03	0.077	2.79
Log Vehicles	0.156	5.58	0.128	2.93
White Household	0.067	5.97	0.088	5.98
Fixed Route Commuter	**0.018**	**0.99**	0.073	**1.88**
Log Density	**0.288**	**13.71**	**0.282**	**11.63**
Log Puma Mean Inc	0.613	5.26	**0.534**	4.02

Log Puma Mean Renter Inc				
Log Income Difference	0.38	3.77	0.364	3.01
Log Puma Drivers Ave Commute	**0.245**	**0.837**	**-0.233**	**-0.82**
N	**55,059**		**9,955**	
K (Clusters)	**1010**		**906**	
R Squared	**0.348**		**0.3421**	
Adjusted R Squared	**0.347**		**0.3367**	

In the mortgage payment models, the number of bedrooms, rooms, household income, years of residence, number of vehicles, PUMA mean income, the PUMA income inequality measure and number of residents correlated positively with housing values as anticipated. The age of the structure and the number of years the household resided in the unit correlated negatively with primary monthly mortgage payments. The PUMA driver average commute time, highly significant in the single family models, was insignificant in the mortgage payment models for multi-family units.

White households in multi-family units paid 6.7% more on their mortgages than non-white households, ceteris paribus, and this result was significant. Just among recent home buyers, that difference was also positive and significant at 8.8%. In the model with all mortgage-paying families, families with fixed route transit commuters paid 1.8% more on their mortgages than households without, but this result was not significant. For families who purchased their homes within a year of taking the survey, households with at least one fixed route transit commuter paid 7.3% more on their mortgages than those without. The author attributes the different results in the two models to the fixed route transit dummy better operationalizing the effect of the unit being near fixed route transit infrastructure in the recent home buyer model. This is because individuals who purchased a new home and then began commuting by those modes to work would be more conscious of this benefit when purchasing the unit. Also, the latter model could include families that purchased their homes and then, a decade later, began using fixed route transit because a new route went online during that decade near their home but which did not affect its value during the time in which they purchased the unit. For the reasons, the author lends more weight the recent home buyer models on the variable.

3. *Model Diagnostics and Discussion*

As with the single family models, the author computed diagnostics on the multi-family models in R. Figure eleven presents the model diagnostics for the multi-family rent model. The left hand residual versus fitted and scale-location plots suggest the violation of the assumption of homoscedasticity on the lower end of the distribution of predicted values. As in the single-family models, the multi-family rent model suffers from non-normal distribution of residuals at one end of the residual distribution according to the normal Q-Q plot on the top right.

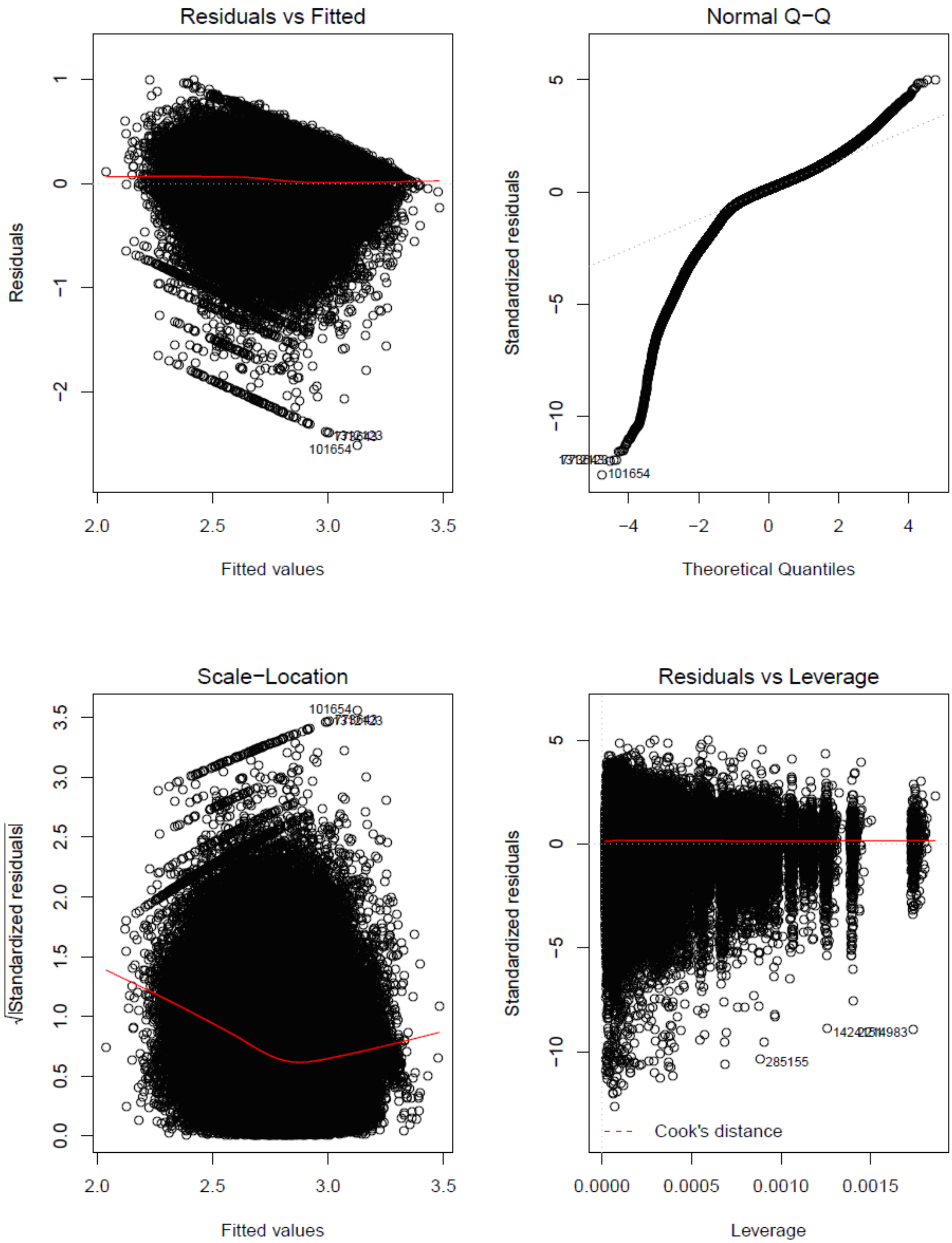

Figure 11: Multi-Family Rent Model Diagnostics

As with the single family unit model, the first year home owner model for multi-family unit values showed similar problems of severe heteroscedasticity and non-normal distribution of error terms:

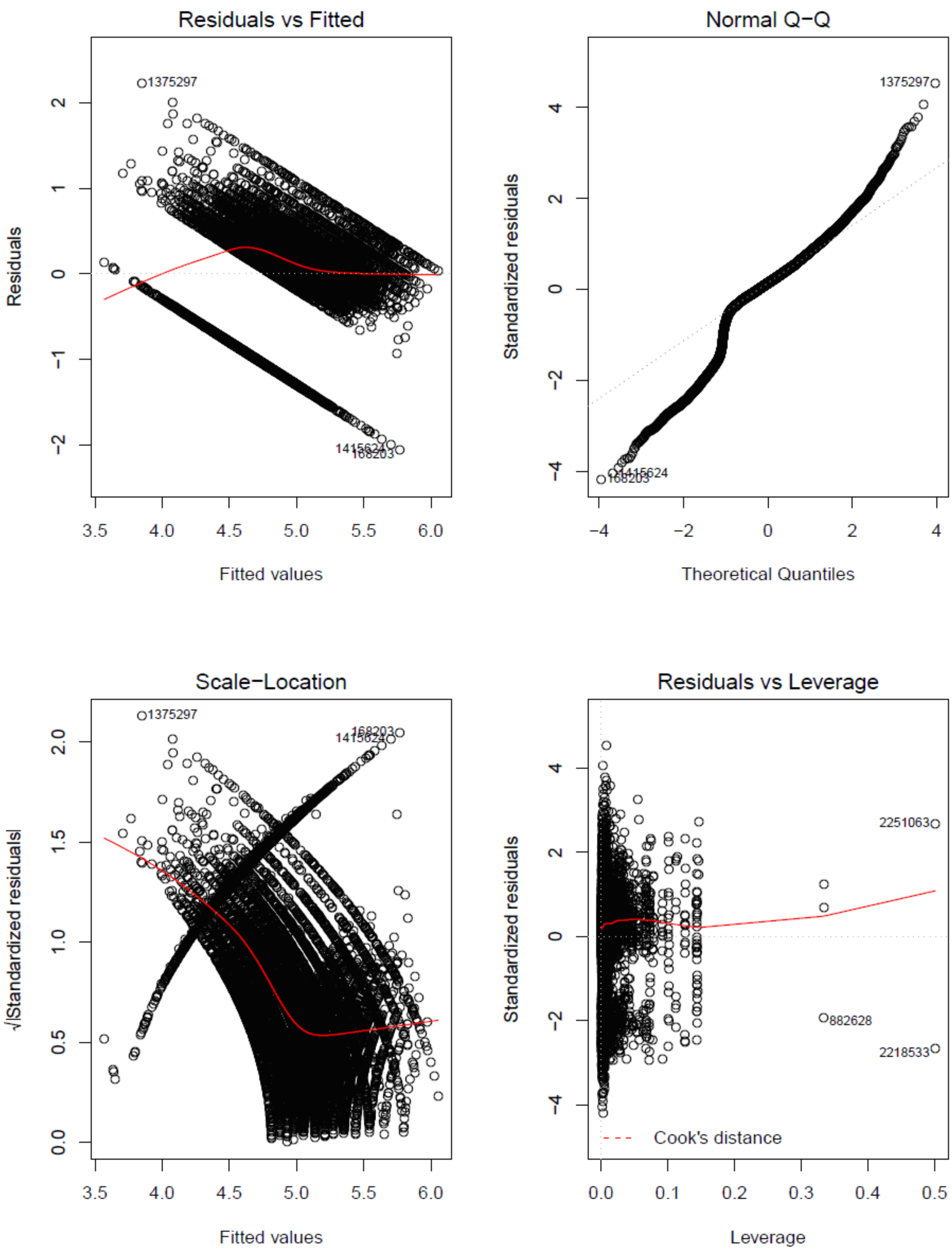

Figure 12: Multi-Family Unit Value Model Diagnostics, Recent Buyers

The author explored the three observations that were categorized by R as outliers and which exhibit some leverage in the data. The two furthest to the right in the residuals versus leverage plot in the bottom right corner of figure twelve were the only two observations in the one year dataset found in the MSACSMA area codded 4880, listed as McAllen-Edinburg-Mission Texas. The other observation recognized as an outlier by R, labeled as "882628" in Figure Twelve, was the only observation in PUMA 2201 in MSACMSA 1320, Canton-Massillon, Ohio. That PUMA listed a low population density of 538.82 persons per square mile and its only observation recorded a unit value of $5,000. That unit was coded as 25 years old, containing four rooms, two bedrooms with one resident who possessed three vehicles.

To understand the effects of these observations, the author ran auxiliary regressions for this model which limited the observations based on their standardized residuals in the original model diagnosed above. The author ran an auxiliary model for units where standardized residuals in the model above were less than the absolute value of three, and another where they were less than the absolute value of two. These models produced results with identical signs, but the fixed route transit dummy lost significance and the log of PUMA drivers average commute time gained significance. Those restricted models showed improved R-squared values (.44 and .45 respectively), but still showed serious issues similar to those presented in the scale-location and residual-fitted charts in Figure Twelve. The author believes this is due to the categorical nature of the variable.

Figure Thirteen presents model diagnostics for the multi-family primary mortgage model for recent home buyers. Like the previous model, which largely draws upon the same observations, this model suffered from heterskedasticity, non-normally distributed errors and a few units that exhibited leverage and may have influenced results.

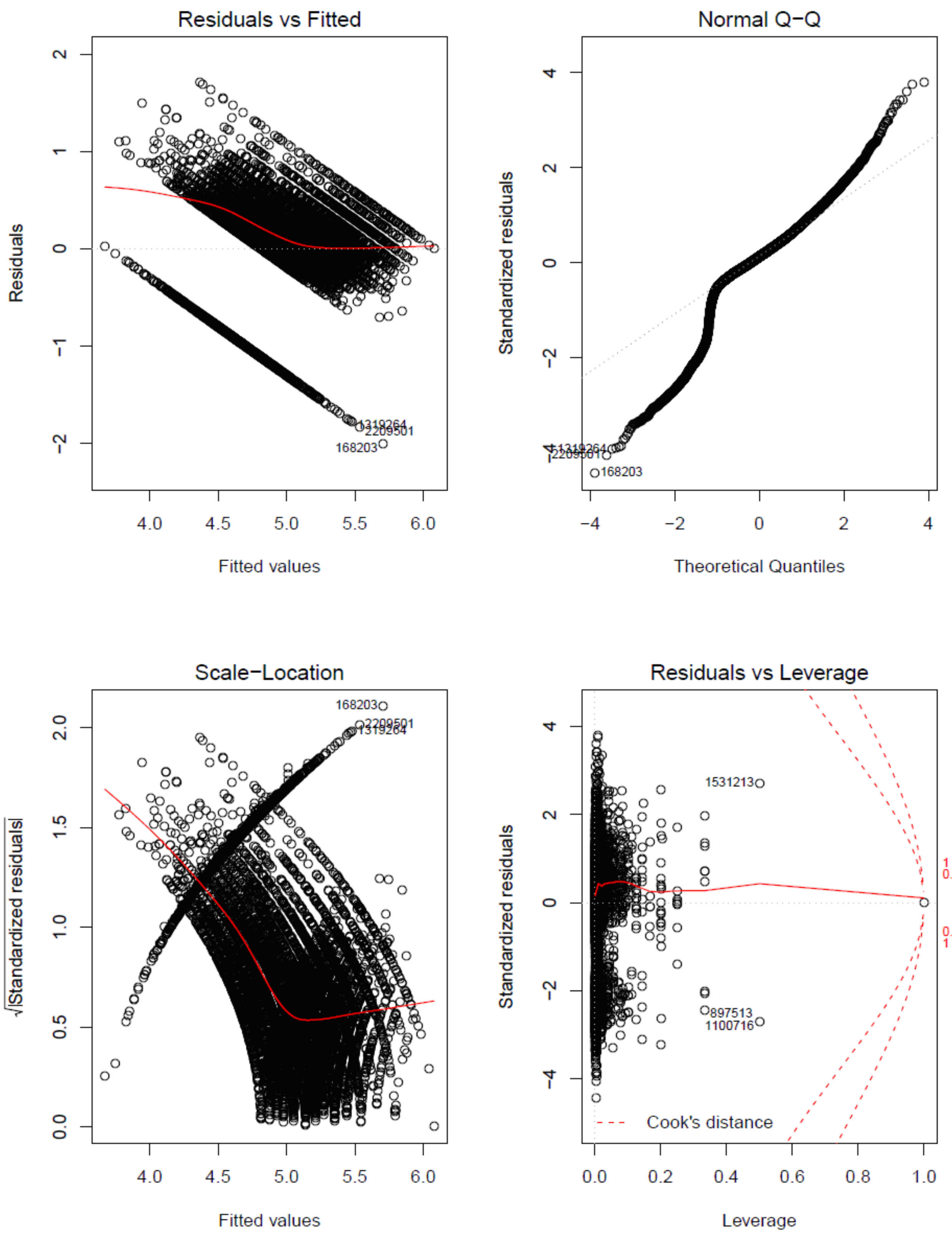

Figure Thirteen: Multi-Family Primary Mortgage Model, Recent Buyers

To address these problems, the author applied the same tactics used to address influential observations in the home value models. The author produced two jackknifed regressions, one which cut out observations with standardized residuals above the absolute value of 2, and another which cut out the value of observations with absolute values above 1.5. This process removes the observations which exhibited leverage in Figure 13 and their resultant biasing of the model coefficients. These models proved statistically more robust, with normally distributed error terms as documented by their diagnostic analysis provided in the appendix.

Table 16: Multi-Family Mortgage Payment Models Corrected For Outliers:

	Baseline From Table X		**Original Standardized Resid < |2|**		**Original Standardized Resid < |1.5|**	
	b	**cluster robust t-stat**	**b**	**cluster robust t-stat**	**b**	**cluster robust t-stat**
Log Bedrooms	0.66	9.73	0.429	12.86	0.403	13.95
Log Rooms	0.077	1.52	0.01	0.41	0.026	1.29
Log Building Age	-0.118	-9.77	0	0.136	0	-1.02
Log Units	-0.022	-1.4	-0.025	-3.3	-0.026	-3.87
Log Years There						
Log Income	0.101	8.24	0.083	12.77	0.09	15.69
Log Residents	0.077	2.79	0.067	5.65	0.055	5.54
Log Vehicles	0.128	2.93	0.084	4.38	0.062	4
White Household	0.088	5.98	0.011	1.94	0.005	1
Fixed Route Commuter	0.073	**1.88**	0.062	**3.95**	0.057	4.26
Log Density	**0.282**	**11.63**	**0.141**	**11.56**	0.139	13.65
Log Puma Mean Inc	**0.534**	4.02	**0.237**	3.85	0.221	4.18
Log Puma Mean Renter Inc						
Log Income Difference	0.364	3.01	0.242	4.04	0.238	4.56
Log Puma Drivers Ave Com	**-0.233**	**-0.82**	**-0.137**	**-1.2**	-0.189	-2.17
N	**9,955**		**9,434**		**8,571**	
K (Clusters)	**906**		**898**		**886**	
R Squared	**0.3421**		**0.4521**		**0.5238**	
Adjusted R Squared	**0.3367**		**0.4473**		**0.5194**	

These results suggest that the influential observations in the data may have severely biased model coefficients, the density coefficient in particular. After removing just over one hundred observations with large standardized residuals, the elasticity of population density on mortgage payments drops from .282 to .141. A stricter jackknife analysis on the far right of the table shows the density coefficient dropped further still to 13.65, but statistically significant differences between the two coefficients cannot be identified. With influential observations gone, the T-stats on both commute variables begin to rise significantly. They also produce signs that suggest corroborate the theory that households tradeoff between housing and transportation costs, with mortgage payments rising as PUMA driver average commute times fall.

G. Conclusions and Policy Implications

The results of this study have significant implications for long range planning priorities in major metropolitan areas. The regional government of Portland, Oregon's 2040 planning documents specifies ideal density requirements for distinct types of neighborhoods across the region (METRO 2004). A similar long range planning document produced by the San Francisco Bay area's regional government proposes increasing population densities in select neighborhoods across the region to levels denser than Manhattan, New York. This is proposed to minimize carbon emissions and prevent the expansion of the region's urban footprint even as it incorporates a projected 2.2 million additional residents (Association of Bay Area Governments 2013). The results of this paper suggest that these plans may have significant implications for the equity and welfare effects of these planning policies and goals.

Urban projects and policies must meet three criteria to be considered fair or just under a normative definition offered by Fainstein (2010): democratic governance, support for diversity and equitable outcomes . The models outlined in the results sections demonstrate that attaining higher densities in urban areas to reduce carbon emissions may not qualify as a just policy given the criteria of equitable outcomes. Most models support previous literature that suggests that fixed route transit commuters pay a premium to live in proximity to transit, mitigating potential benefits of transit orientated development for low income households. While increased density reduces transport costs for households via reduced vehicle miles traveled and increased ease of use of modes like walking and biking, all models run for this project show increased densities correlate with higher housing prices, ceteris paribus. To truly get at the question of whether or not density policies are equitable and thus potentially "just" urban policies, a finer analysis of household spending on transportation and housing is needed. A survey instrument which captures both transportation behavior and household expenditures on housing and transportation could better meet the task of modeling how increased population densities would affect household budgets in total. Such analysis could also get at how this joint impact of density on housing and transportation spending by households may vary among different income groups or ethnic groups based on their responses to increased transit availability or fluctuations in housing markets. Fainstein's model is one of many for defining just or equitable public policies, and the author does not include this discussion in this conclusion to take a stance on whether nor not density promoting policies are just or fair. Rather, juxtaposing Fainstein's criterion with what this project confirms about the relationships between density, accessibility and affordability suggests that future research must ascertain strategies to achieve increased population densities without increased housing costs or gentrification.

More importantly, the significance of density in defining agglomeration economies and the urban amenity premium, as discussed in the literature review, suggest that the ability to access and participate in these spaces may be key to an individual or communities' economic success and survival in future decades. Combing the work of Glaeser, Kolko and Saiz (2000) and Glaeser and Gottlieb (2009) with David Harvey's concept of the right to the city (2013) a picture emerges regarding the importance of cities maintaining their accessibility and affordability to enable all segments of society to participate in these spaces. With this framework applied to the data in this project, density promoting policies clearly run the risk of pricing city residents out of the areas of cities that host the employment sectors and urban amenities central to our society.

In conclusion, increased population density correlates with higher rents, higher mortgage payments and higher housing unit values in the 2000 PUMS 5% samples from twenty-three of the country's most populous states. In the late 1990's individuals were purchasing housing in newer and less dense areas. These home buyers were less likely to be white and had higher incomes than other home owners. Households that located such that a commuter in the household opted to commute via subways, street cars and rail lines paid a premium for that access, regardless of how long they had lived in their home. Households that purchased detached, single-family units within a year of taking the survey paid more on their mortgages as their PUMA's average driver commute time rose. This effect was initially not found in the multi-family models, but did appear significant in jackknifed models designed to reduce the effect of outliers and influential observations in the multi-family owner dataset.

The finding that increased density and proximity to transit, or at least transit use, correlate positively with housing costs across these three variables in most market segments should raise concern among policy makers about the best tactics to use to promote smart growth. Higher density development and expansion of transit systems my reduce vehicle miles traveled (VMT) for residents in a given area, but if these changes also raise rents and housing prices too high, they may merely 'shift' VMT emissions around as low income households re-locate to less accessible but more affordable areas. Steps should be taken to ensure affordable housing near transit stations and in density promoting land use policies.

Future research should build on this analysis by evaluating the impacts of density at the PUMA or tract level over time, assessing both transportation accessibility changes and housing affordability. Future research should also strive to identify communities that managed to achieve increased residential densities, reduced VMT and increased non-single occupancy vehicle mode shares without seeing a significant response in housing markets. Understanding how some communities have managed to make change while preserving existing communities could serve as models for urban and regional policy making nationwide to ensure equitable outcomes.

References

Aanacker, K. (2010). Still paying the race tax? analyzing property values in honogenous and mixed-race suburbs. *Journal of Urban Affairs*, *32*(1), 55-77.

Abramson, A. J., Tobin, M. S., & VanderGoot, M. R. (1995). The Changing Geography Of Metropolitan Opportunity: The Segregation Of The Poor In U.S. Metropolitan Areas, 1970 To 1990. *Housing Policy Debate*, *6*(1), 45-72.

Alonso, W. (1964). *Location and Land Use: Toward a General Theory of Land Rent*. Cambridge: Harvard University Press.

Arnott, R. (1995). Time for revisionism on rent control?. *Journal of Economic Perspectives*, *9*(1), 99-120.

Association of Bay Area Governments. (2013). *Plan bay area building on a legacy of leadership*. San Francisco, CA: Association of Bay Area Governments. Retrieved from: http://onebayarea.org/regional-initiatives/plan-bay-area.html

Blackley, D. M., Follain, J. R., & Lee, H. (1986). An evaluation of hedonic price indexes for thirty-four large SMSAs. *Association of Real Estate and Urban Economics Association Journal*, *14*(2), 179-205.

Brownstone, D., & Golob, T. F. (2009). The impact of residential density on vehicle usage and energy consumption. *Journal of Urban Economics*, *65*, 91-98.

Brueckner, J. K., & Colwell, P. F. (1983). A spatial model of housing attributes: theory and evidence. *Land Economics*, *59*(1), 58-69.

Brueckner, J. T., & Fansler, D. A. (1983). The economics of urban sprawl: theory and evidence on the spatial size of cities. *The Review of Economics and Statistics*, *65*(3), 479-482.

Census 2000. Public Use Microdata Sample, (PUMS), United States, prepared by the U.S. Census Bureau, 2003.

Chetty, R., & Friedman, J. (n.d.). Does local tax financing of public schools perpetuate inequality?. *Harvard Facult*. Retrieved April 6, 2013, from http://obs.rc.fas.harvard.edu/chetty/proptax_nta.pdf

Cho, M. (1997). Congestion effects of spatial growth restrictions: a model and empirical analysis. *Real Estate Economics*, *25*(3), 409-n/a.

Ciccone, A., & Hall, R. E. (n.d.). Productivity and the density of economic activity. *NBER Working Paper Series*. Retrieved January 16, 2013, from www.nber.org/papers/w4313.pdf?new_window=1

Cortright, J. (n.d.). Walking the Walk: How Walkability Raises Home Values In U.S. Cities. *Walking the Walk*. Retrieved February 27, 2013, from blog.walkscore.com/wp-content/uploads/2009/08/WalkingTheWalk_CEOsforCities.pdf

Fanstein, S. (2010). *The just city*. New York, New York: Cornell University Press.

Fisher, L. M., Pollakowski, H. O., & Zabel, J. (2009). Amenity-based housing affordability indexes . *Real Estate Economics*, *37*(4), 705-746.

Gabriel, S. A., & Painter, G. D. (2012). Household location and race: a 20-year retrospective. *Journal of Regional Science*, *52*(5), 809-818.

Glaeser, E. (n.d.). The life and death of cities. *Federal Reserve Bank of Cleveland*. Retrieved April 15, 2013, from www.clevelandfed.org/research/conferences/2008/4_3-4/glaeser.pdf

Glaeser, E.L. & Gottlieb, J.D. (2009). The wealth of cities: agglomeration economies and the spatial equilibrium. *NBER Working Paper No. 1480*. Retrieved November 30, 2012, from http://www.nber.org/papers/w14806

Glaeser, E. L., Kolko, J., & Saiz, A. (2000). Consumer City. *NBER Working Paper*, *No. w7790*.

Retrieved November 15, 2012, from http://www.nber.org/papers/w7790

Guntermann, K. L., & Norrbin, S. (1987). Explaining the variation of apartment rents. *The American Real Estate and Urban Economics Journal*, *15*(4), 321-340.

Hahn, E., Chatterjee, A., & Younger, M. S. (2002). Macro-level analysis of factors related to areawide highway traffic congestion. *Transportation Research Record*, *1817*, 11-16.

Harries, K. (2006). Property crimes and violence in united states: an analysis of the influence of population density. *International Journal of Criminal Justice*, *1*(2). Retrieved May 10, 2013, from http://www.sascv.org/ijcjs/harries.html

Harvey, David. (2013). *Rebel Cities: From the Right to The City to The Urban Revolution*. London: Verso Books.

Holmes, A., & James, J. F. (1995). Discrimination, lending practices and housing values: preliminary evidence from the houston market. *The Journal of Real Estate Research*, *11*(1), 1996.

Hotzclaw, J., Clear, R., Dittmar, H., Goldstein, D., & Haas, P. (2002). Location efficiency: neighborhood and socioeconomic characteristics determine auto ownership and use- studies in Chicago, Los Angeles and San Francisco. *Transportation Planning and Technology*, *25*, 1-17.

Jargowsky, P. (1996). Take the money and run: economic segregation in the u.s. metropolitan areas. *American Sociological Review*, *61*(6), 984-998.

Jud, D., & Winkler, D. (1991). Location and amenities in determining apartment rents: an integer programming approach. *The Appraisal Journal*, *NA*, 266-275.

Jun, M. (2006). The effects of portland's urban growth boundary on housing prices. *Journal of the American Planning Association*, *72*(2), 239-243.

Kestens, Y., ThÃ©riault, M., & Rosiers, F. D. (2006). Heterogeneity In Hedonic Modelling Of House Prices: Looking At Buyersâ€™ Household Profiles. *Journal of Geographical Systems*, *8*(1), 61-96.

Kim, S. (1991). Search, hedonic prices and housing demand. *The Review of Economics and Statistics*, *74*(3), 503-508.

Kulish, M., Richards, A., & Gillitzer, C. (n.d.). Urban structure housing prices: some evidence from Australian cities. *Reserve Bank of Australia*. Retrieved August 30, 2012, from www.rba.gov.au/publications/rdp/2011/pdf/rdp2011-03.pdf

Leingberger, C. B., & Alfonzo, M. (n.d.). Walk This Way: The Economic Promise of Walkable Places in Metropolitan Washington D.C.. *Walk This Way*. Retrieved March 5, 2013, from www.brookings.edu/~/media/Research/Files/Papers/2012/5/25%20walkable%20places%20leinberger/25%20walkable%20places%20leinberger.pdf

Louis, E. T. (1997, May. - Jun.). The price is wrong: appraisal 'redlining" understate millions in community assets. *National housing Institute Shelterforce Online*, *na*, 1:5.

Marks, D. (1984). The effect of rent control on the price of rental housing: an hedonic approach. *Land Economics*, *60*(1), 81-94.

Mills, E. (1967). An aggregate model of resource allocation in a metropolitan area. *The American Economic Review, Papers and Proceedings*, *57*(2), 197-210.

Mills, E. (1970). Urban Density Functions. *Urban Studies*, *7*(1), 5-20.

Muth, R. (1969). *Cities and housing*. Chicago: University of Chicago Press.

Oregon Metro. (2010). *Chapter 3.07. urban growth functional plan*. Portland, Oregon: Metro. Retrieved from: http://www.oregonmetro.gov/index.cfm/go/by.web/id=274

Ottensmann, J. R., Payton, S., & Man, J. (2008). Urban location and housing prices within a hedonic model. *Journal of Regional Policy Analysis*, *38*(1), 19-35.

Pendall, R., & Carruthers, J. I. (2003). Does Density Exacerbate Income Segregation? Evidence From U.S. Metropolitan Areas, 1980 To 2000. *Housing Policy Debate*, *14*(4), 541-589.

Pratt, T. (2001). *Assessing the relative effects of macro-level predictors of crime: a meta-analysis*. Cincinatti: University of Cincinnati.

Quillian, L. (2012). Segregation and poverty concentration: the role of three segregations. *American Sociological Review*, *77*(3), 354-379.

Rappaport, J. (2008). Consumption amenities and city population density. *Regional Science and Urban Economics*, *38*(6), 533-552.

Reardon, S. F., & Bischoff, K. (2011). Income inequality and income segregation. *American Journal of Sociology*, *116*(4), 1092-1153.

Schiff, N. (n.d.). Cities and product variety in the restaurant industry. *Sauder School of Business*. Retrieved January 28, 2013, from http://strategy.sauder.ubc.ca/schiff/density_diversity_011309.pdf

Song, S. (1992). Monocentric and polycentric density functions and their required commutes. *Working Paper*, *UCTC No. 198*, na.

Su, Q. (2011). The effects of population density, road network density, and congestion on household gasoline consumption in U.S. urban areas. *Energy Economics*, *33*(3), 445-452.

Sykes, L. L. (2008). Cashing In On The American Dream: Racial Differences In Housing Values 1970-2000. *Housing, Theory and Society*, *25*(4), 254-274.

Tse, R. (2002). Estimating Neighbourhood Effects In House Prices: Towards A New Hedonic

Model Approach. *Urban Studies*, *39*(7), 1165-1180.

Valente, J., Wu, S., Gelfand, A., & Sirmans, C. (2005). Apartment rent prediction using spatial modeling. *Journal of Real Estate Research*, *27*(1), 105-136.

Wardrip, K. (n.d.). Public Transit's Impact on Housing Costs: A Review of the Literature. *Insights from Housing Policy Research*. Retrieved October 29, 2012, from www.nhc.org/publications/index.html

Watson, T. (2009). Inequality and the measurment of residential segregation by income in american neigborhoods. *The Review of Income and Wealth*, *55*(3), 820-844.

Appendix

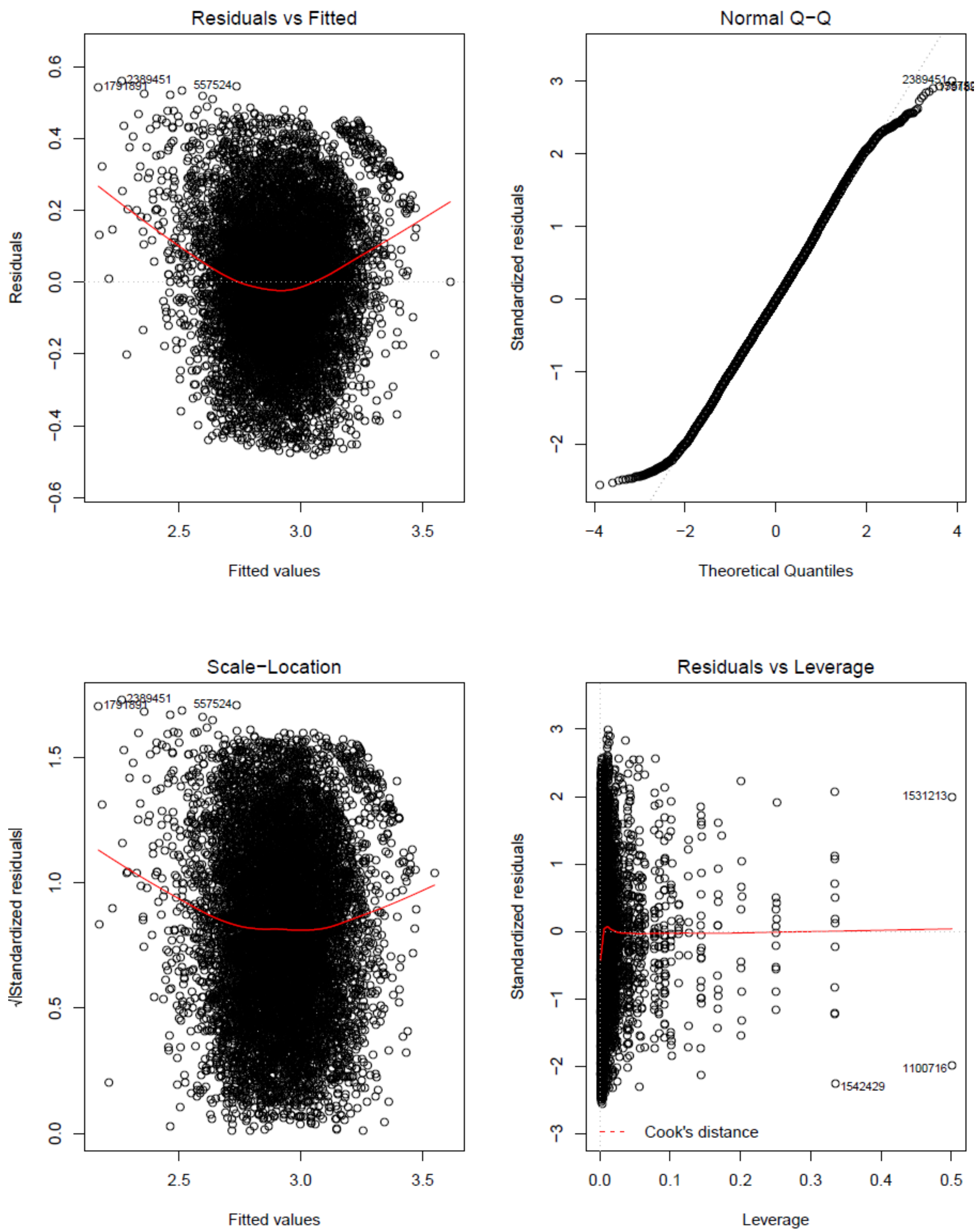

Figure 14: Auxilliary Regression Diagnostic for Multi-Family Mortgage Model restricted at |r| < 2

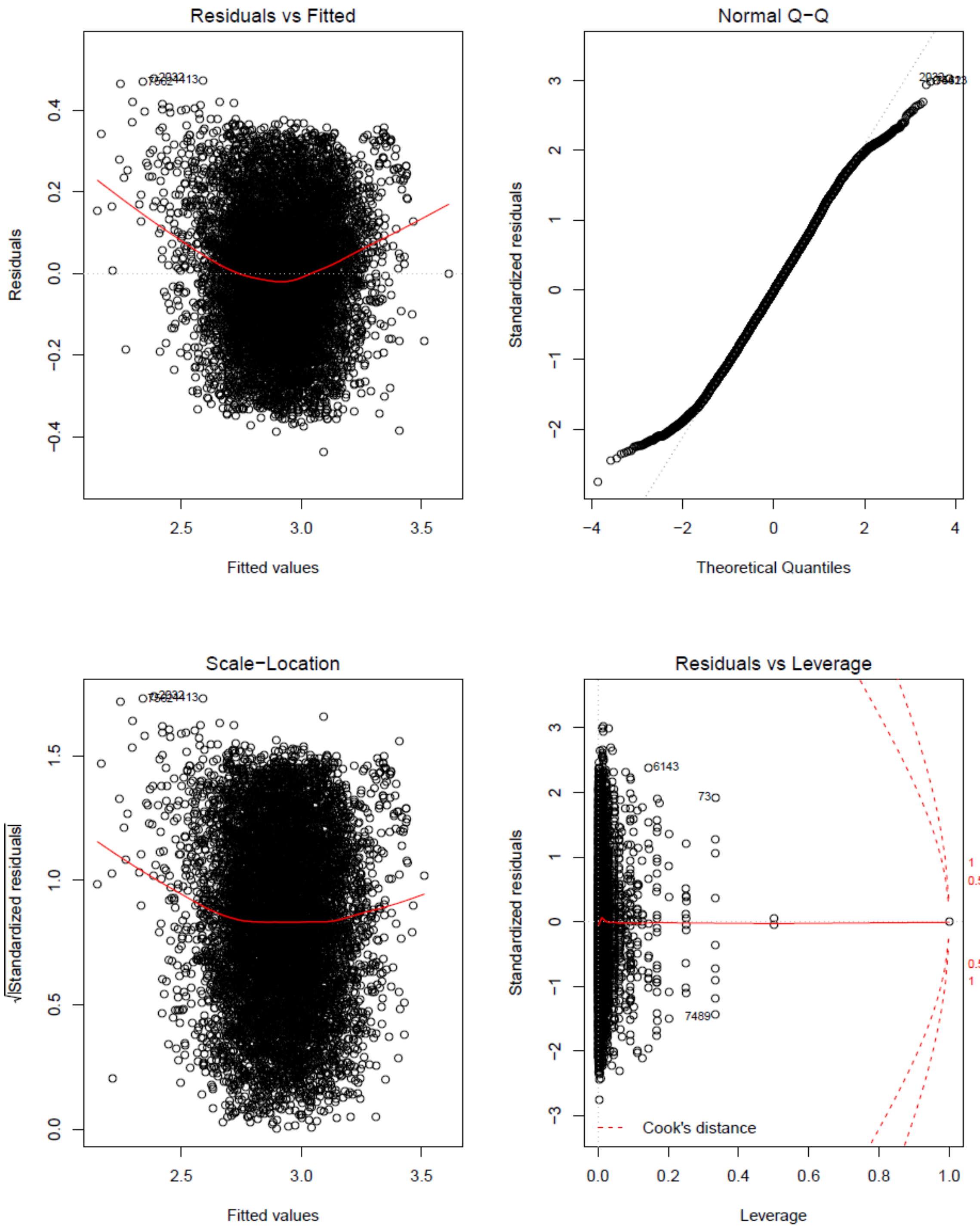

Figure 15: Auxilliary Regression Diagnostic for Multi-Family Mortgage Model restricted at |r| < 1.5