Housing Affordability and Income Inequality: The Impact of Demographic Characteristics on Housing Prices in San Francisco

Olga Baranoff*

Senior Honors Thesis in Economics, Johns Hopkins University

April 10, 2016

Abstract: Housing affordability is a growing crisis for urban areas with constrained housing markets. In this paper, I examine the effects of income inequality and other demographic characteristics on housing prices in San Francisco. I find that income is not a determinant of housing prices, but that other housing and demographic characteristics do impact housing prices. That these findings differ from the existing literature is likely due to limitations of my study, but the implications of my study are nonetheless useful in the context of housing supply and pricing. As income inequality in San Francisco continues to grow, its impact on rising housing prices is relevant for policymakers seeking to address the city's housing affordability crisis.

* The author would like to acknowledge the JHU Economics Department for its support throughout this process. Specifically, she would like to thank her advisor, Robert Moffitt, and JHU graduate student Mikhail Smirnov for their guidance and assistance. The author would also like to thank her friends and family for their support.

Section I. Introduction

Housing supply and affordability plays a large role in creating differences in housing markets and labor markets across urban areas. Housing affordability within an urban area has implications for local labor markets and the local economy. A firm in a given region, for instance, cannot expand without affordably priced homes to house new workers. A robust literature in economics examines the relationship between housing affordability and income inequality, which has become a pressing issue for many of the country's most popular metropolitan areas.[1] This issue is especially relevant in the case of housing in San Francisco, which has been making national headlines in recent years for its rising rent and housing prices.[2] San Francisco's growing population and demand for housing have not been met by rising supply, and, at the same time, the city has become increasingly unequal in terms of income, further exacerbating housing prices.[3]

In this paper, my research question aims to determine what effect, if any, income has on housing prices in San Francisco. I hypothesize that greater income inequality exacerbates housing affordability, and that higher incomes increase housing prices. Specifically, I use hedonic price regressions to estimate the effect of house and household demographic characteristics on housing prices. I find that income does not seem to have an economically or

---

[1] See, for example: Janna L. Matlack and Jacob L. Vigdor, "Do Rising Tides Lift All Prices? Income Inequality and Housing Affordability," *Journal of Housing Economics* 17, no. 3 (September 2008): 212–24; Patrick Bayer, Robert McMillan, and Kim Rueben, "An Equilibrium Model of Sorting in an Urban Housing Market," Working Paper (National Bureau of Economic Research, November 2004); Shaila Dewan, "In Many Cities, Rent Is Rising Out of Reach of Middle Class," *The New York Times*, April 14, 2014.
[2] See, for example: Matthew Yglesias, "San Francisco Is Great—They Should Make More of It," *Slate*, April 11, 2013; "The Spectre Haunting San Francisco," *The Economist*, April 16, 2014; "Growing Pains," *The Economist*, December 7, 2013; Naureen Khan, "After Failed Propositions, San Francisco Housing Crisis Still Festering," *Al-Jazeera America*, November 4, 2015, sec. U.S.; Enrico Moretti, "How to Make SF Housing More Affordable," *San Francisco Chronicle*, June 9, 2013, sec. Opinion.
[3] Alan Berube and Natalie Holmes, "Some Cities Are Still More Unequal than Others—an Update" (Washington, D.C.: Brookings, March 17, 2015); Bayer, McMillan, and Rueben, "An Equilibrium Model of Sorting in an Urban Housing Market."

statistically significant effect on housing prices, but other house and household characteristics in the models do affect housing prices. While these findings may be limited due to data and methodology constraints, there is room to expand this study. Furthermore, my findings build on previous literature by providing policy implications for addressing the relationship between housing affordability and income inequality.

The remainder of this paper is organized as follows. Section II provides Background information on housing affordability in San Francisco. Section III reviews the existing literature. In Section IV, I describe the data used, and in Section V, I discuss my methodology. Section VI presents the main results of my estimations. Section VII considers limitations and suggests expansions of this study. Section VIII discusses potential policy implications and concludes.

Section II. Background

Rents and housing prices in San Francisco have been steadily rising for the past decade, creating a housing affordability crisis in the city.[4] At the same time, San Francisco has become one of the most unequal cities in the United States, with top incomes exceeding those of other American cities. Concurrently, and in contrast to its growing income inequality, San Francisco has also experienced high growth at the low-end of its income distribution; the city ranks second highest out of US cities ranked by highest increases in 20th-percentile incomes.[5] These findings suggest that, despite increasing income inequality in San Francisco, the city also offers socioeconomic mobility for those towards the bottom end of the income distribution. However, the city's rising income inequality impacts housing affordability and may deter lower-income individuals from residing in the city, and thus from being able to take advantage of the social

[4] "America's Most Expensive City Just Got Even More Expensive," *The Huffington Post*, accessed December 10, 2015; Moretti, "How to Make SF Housing More Affordable."
[5] Berube and Holmes, "Some Cities Are Still More Unequal than Others—an Update."

mobility of the city. Research has shown that increases at the top end of the income distribution are associated with significantly higher rents and crowding out for those at the lower end of the income distribution.[6] This crowding out effect also has potential impacts for local labor markets and economies. Aside from growing income inequality and decreased housing affordability, the supply of housing in San Francisco has failed to keep pace with demand, putting upward pressure on the existing stock of housing. This has further exacerbated housing prices and has contributed to San Francisco's housing affordability crisis.[7]

As shown in Figures 1 and 2, San Francisco has experienced a sharp increase in its total population since the 1980s, and total households have also grown. The total population by decade shows the number of individuals residing in San Francisco from 1970 to 2010. A household, in comparison, consists of one or more people living together within a single housing unit, and includes both family and non-family households.

While San Francisco's population has been growing at a steady rate, the growth of housing units has not kept pace. Figure 3 shows the number of housing units in San Francisco per decade. The total number of housing units in each decade has increased, with the the rate of new units increasing by 2%, 4%, 5%, and 9% between each decade. While the number of housing units has been growing at an increasing rate between each decade, supply has not kept pace with population growth or demand for housing. As Figure 4 shows, nearly half of San Francisco's housing units were built before 1940, and a majority of units were constructed before 1980. The existing housing stock, however, is in high demand; as Figure 5 shows, in every decade since the 1970s, there has been an increase in the number of occupied units (although the number of vacant units has fluctuated). The distinction between renters and owners in San

[6] Matlack and Vigdor, "Do Rising Tides Lift All Prices?"
[7] Thornberg, "California's Affordable Housing Crisis Spreads to the Middle Class."

Francisco's housing market is also noteworthy. Historically and presently, renters occupy roughly two-thirds of San Francisco's housing units, and owners occupy the remaining one-third (Figures 6 and 7).

In addition to growing demand and short supply of housing in San Francisco, housing affordability is of concern to those who rent and own property in the city. Housing is considered to be "affordable" if it costs no more than 30% of a family's income.[8] In San Francisco, Census estimates from 2012 suggest that over one-third of renters and owners pay more than 35% of their gross income on housing costs.[9] Figure 8 shows the percentage of renters and owners in 2013 that paid above and below 30% of their household income on their housing costs. Owners are broken out into those with and without mortgages. Roughly 40% each of renters and owners with a mortgage spend more than 30% of their household income on housing costs. Figure 9 shows median monthly costs for owners and renters in 2013. Renter costs are calculated as gross rent, while owner costs may include homeowner related fees and costs. Owners with mortgages spend the most on monthly housing costs.

Despite San Francisco's booming economy, the San Francisco Bay Area remains extremely difficult to build, which causes housing prices and rents to soar.[10] Supply constraints cause prices to increase, but the price of housing is not fully captured in land and construction costs; the expense of creating housing also stems from economic rents and zoning restrictions.[11] Zoning and land use restrictions in San Francisco induce property owners to extract economic

---

[8] "Affordable Housing," *U.S. Department of Housing and Urban Development*, accessed April 5, 2016, http://portal.hud.gov/hudportal/HUD?src=/program_offices/comm_planning/affordablehousing/.
[9] "SF Housing Data Hub | Housing Overview," accessed April 5, 2016, http://housing.datasf.org/overview/.
[10] Kim-Mai Cutler, "How Burrowing Owls Lead To Vomiting Anarchists (Or SF's Housing Crisis Explained)," *TechCrunch*, April 14, 2014, http://social.techcrunch.com/2014/04/14/sf-housing/; "The Spectre Haunting San Francisco."
[11] "Economic rent is the cost of non-produced inputs or advantages; the result of natural or contrived exclusivity." "The Spectre Haunting San Francisco"; Steven M. Teles, "The Scourge of Upward Redistribution," *National Affairs*, Fall 2015.

rents by allowing the "owners of capital [such as housing units] to extract an outsized share of the surplus generated by job creation."[12] Research on zoning and land use restrictions, described in Section III below, shows that higher-income individuals often have an outsize effect on maintaining restrictive land use and housing policies. These restrictive policies likely increase with rising income inequality. While zoning and land use restrictions are critical to understanding housing supply and affordability and therefore are described in length in the following section, data limitations make it difficult to conduct analysis of that topic. I therefore focus on the impact of income inequality and other demographic characteristics on housing prices.

Section III. Literature Review

While my research question focuses on the impact of income inequality on housing affordability, I review a larger body of economic literature on housing prices, supply and demand, and zoning. I examine literature on: the political economy of land use and zoning, economic rents and welfare implications, housing prices and supply and demand, and housing affordability and residential sorting.

Political Economy of Land Use and Zoning

A large body of literature from political science, economics, and law examines the political economy of land use and the role of politics and homeowners in determining land use and zoning policies.[13] In his article "City Unplanning," David Schleicher analyzes the limits on development in America's growing cities, which have not seen enough growth in housing supply to prevent prices from increasing exponentially. Although Schleicher focuses primarily on the role of local government regimes that make land-use decisions, his analysis provides insight into

[12] "The Spectre Haunting San Francisco."
[13] David Schleicher, "City Unplanning," *The Yale Law Journal* 122, no. 7 (May 2013): 1670–2150.

the influence of local politics on setting land use restrictions and attempting to mitigate the perceived negative externalities associated with new development. In San Francisco, for instance, local political groups and homeowner groups often form coalitions to fight development and rezoning. Schleicher and others also emphasize that resistance to greater density, especially by local government regimes and homeowner coalitions, slow job creation in productive metropolitan areas, such as the San Francisco Bay Area.[14]

William Fischel examines the history of American zoning and explains how homeowners came to dominate the zoning process.[15] Zoning was originally meant to protect homeowners in residential areas from devaluation by industrial uses. Fischel focuses on how the "homevoter" approach to local government explains why zoning became more exclusionary during the 1970s, and his approach is valuable in thinking about how homeowners in San Francisco may follow a similar model when lobbying their local government for certain zoning provisions.

Related to the "homevoter" approach to zoning, research on the political economy of land use regulation shows that, when higher income incumbents control the political processes that make decisions regarding local planning and zoning, regions can become less affordable as prices increase.[16] If a city, like San Francisco, has a growing population of high-income residents, and these residents, in turn, gain greater control over political processes that control zoning, then housing becomes even more skewed to favor the wealthy.

---

[14] Ryan Avent, *The Gated City*, 2011; Ryan Avent, "One Path to Better Jobs: More Density in Cities," *The New York Times*, September 3, 2011, sec. Opinion / Sunday Review, http://www.nytimes.com/2011/09/04/opinion/sunday/one-path-to-better-jobs-more-density-in-cities.html.

[15] William A. Fischel, "An Economic History of Zoning and a Cure for Its Exclusionary Effects" (Department of Economics, Dartmouth College, December 18, 2001), https://www.dartmouth.edu/~wfischel/Papers/02-03.pdf.

[16] John M. Quigley and Larry A. Rosenthal, "The Effects of Land Use Regulation on the Price of Housing: What Do We Know? What Can We Learn?," *Cityscape: A Journal of Policy Develompent and Research* 8, no. 1 (2005): 69–137; Emily Badger, "Voters in One of America's Most Expensive Cities Just Came up with Another Way to Block New Housing," *The Washington Post*, June 4, 2014.

Economic Rents and Welfare Implications

A further issue to consider is the relationship among housing, local governments, and economic rents. In “Housing Supply Elasticity and Rent Extraction by State and Local Governments,” Diamond uses a spatial equilibrium model to shows that less elastic housing supply increases governments’ abilities to extract rents. Inelastic housing supply raises local governments’ tax revenues, which public sector workers extract as rent in the form of increased compensation, as opposed to governments using the revenues to provide taxpayers with additional government services. The author draws a comparison between San Francisco and St. Louis that is useful to consider. Diamond estimates that San Francisco’s land unavailability is 2.9 standard deviations higher than St. Louis’ and that, if San Francisco had the land availability of St. Louis, San Francisco could save over $11,000 per city worker annually.[17] The author’s estimation is purely hypothetical, however, since, unlike St. Louis, San Francisco is a on a peninsula and therefore has a very limited supply of available land. Nonetheless, Diamond’s research shows how government policies that impact housing supply can have economically large effects on government spending. Given this implication, a potential policy solution could be to channel these rents to alternative recipients, such as taxpayers or low-income housing recipients, instead of government employees.[18]

Affordable housing and low-income housing construction policies also contain potential welfare effects. The Low Income Housing Tax Credit (LIHTC) is one such policy. Analysis by Diamond and McQuade (2015) shows that LIHTC has spillover effects onto surrounding

[17] Rebecca Diamond, “Housing Supply Elasticity and Rent Extraction by State and Local Governments,” July 24, 2015, http://web.stanford.edu/~diamondr/research.html.

[18] Diamond’s work follows others who have examined zoning and tax. See, for example: Bruce W. Hamilton, “Zoning and Property Taxation in a System of Local Governments,” *Urban Studies* 12 (1975): 205–11.

neighborhood residents.[19] LIHTC development in low-income neighborhoods helps revitalize these areas, driving up housing prices, lowering crime rates, and attracting a more racially and income diverse population. LIHTC development in areas with higher median incomes and low minority population shares, however, causes price depreciation. Developing affordable housing in low-income neighborhoods appears to be a viable strategy for revitalizing these areas. These developments also seem to have large welfare impacts. In a city like San Francisco, that is facing growing income inequality and constrained housing supply, LIHTC or similar policies could increase support for rezoning and new development projects.

Another welfare implication of zoning has to do with the wealth transfer that can occur with rezoning. Groves and Helland estimate the transfer of wealth between owners of existing homes when municipal zoning ordinances are created.[20] The authors find that properties suited to residential use gain in value, while properties suited to commercial use decline in value. Since zoning's expressed purpose is to protect residential property from negative externalities associated with neighboring commercial development, zoning impact studies tend to focus on whether zoning raises housing values, but zoning could also create an opportunity for a transfer of wealth. The authors define two components that make up housing prices – the price of land and any structures on the land, and an option price, which is the value of changing the use of the land in the future. Some properties may have a higher option value as commercial developments, whereas others will be more valuable if they remain residential. The authors' study focuses on two communities in Harris County, Texas – one that passed a zoning ordinance more recently, and one that had a pre-existing zoning ordinance – to examine the wealth redistribution between

---

[19] Rebecca Diamond and Tim McQuade, "Who Wants Affordable Housing in Their Backyard? An Equilibrium Analysis of Low Income Property Development," July 2015, http://web.stanford.edu/~diamondr/research.html.
[20] Jeremy R. Groves and Eric Helland, "Zoning and the Distribution of Location Rents: An Empirical Analysis of Harris County, Texas," *Land Economics* 78, no. 1 (2002): 28–44, doi:10.2307/3146921.

existing homeowners when certain homeowners have a higher option value to zone their land as commercial, versus those whose land was more valuable remaining residential. The authors' study is useful in thinking about the benefits of zoning and rezoning and the potential implications of zoning for wealth distribution. For example, if low-income individuals benefit when the value of commercial development rises enough to warrant converting a property from residential to commercial use, then that could provide an impetus for increasing commercial zoning or allowing rezoning in a neighborhood.

Housing Prices and Supply and Demand

In addition to the demand-side issues of housing, the economic literature has also focused on rising housing prices and changes in housing supply. Glaeser et al. examine the cause of rising housing prices in the US in the second half of the 20th century. The authors show that, between 1950 and 1970, increases in housing prices reflected rising housing quality and construction costs, but that, since 1970, price increases grew due to the increasing difficulty of obtaining regulatory approval for building new homes.[21] The authors point out that supply of housing depends on three elements: land, a physical structure, and government approval to build the structure on the land. Rising prices, therefore, reflect increasing land costs, physical construction costs, and/or regulatory barriers to new construction. Glaeser et al. also examine whether the limits on new construction are the result of dwindling supply of land or of other barriers to new construction. The authors find that rising density is not enough to explain large declines in construction, i.e., that permit-issuing and other regulatory barriers play a large role in construction declines. They emphasize the impact of "man-made" housing scarcity, as oppose to geographic limitations, on housing supply and prices, in the sense that housing supply is

[21] Edward L. Glaeser, Joseph Gyourko, and Raven E. Saks, "Why Have Housing Prices Gone Up?," *NBER*, NBER Working Paper Series, February 2005, http://www.nber.org/papers/w11129.pdf.

constrained by government regulations. The authors find that a significant increase in the ability of local residents to block new projects impacts regulatory approval, which is relevant in the context of San Francisco housing, where neighborhood groups frequently embed themselves in the permitting and zoning processes for new developments.

Housing supply is also affected by new construction. In "A Dynamic Model of Housing Supply," Murphy develops a model in which parcel owners choose the optimal timing and nature of construction while accounting for expectations about future prices and costs. Murphy estimates the model using a dataset on parcel owners in the San Francisco Bay Area.[22] While the author is primarily concerned with understanding when and where construction costs occur, the dataset he constructs and his model present a useful methodology, since the framework he applies analyzes how individual behavior can explain aggregate patterns of construction and prices observed in macro data.

Inelastic housing supply is a recurring focus of economists studying urban economics, housing, and land use. Gyourko et al. document differences in house price appreciation across metropolitan areas, which the authors explain is the result of an inelastic supply of land, coupled with an increasing number of high-income households.[23] This causes high house prices and price-to-rent ratios, which crowd out lower-income households, thus decreasing housing affordability. The authors find that aggregate demand drives local house price and income trends. While demand is relevant, the authors point out the necessity of also considering supply. The authors define a "superstar" city as one that has both some inelasticity in its housing supply and that has excess demand, i.e., the location must be preferred by a large share of the population. By

[22] Alvin Murphy, "A Dynamic Model of Housing Supply," *Arizona State University*, July 13, 2015, http://www.public.asu.edu/~amurph10/supply.pdf.
[23] Joseph Gyourko, Christopher Mayer, and Todd Sinai, "Superstar Cities," *American Economic Journal: Economic Policy* 5, no. 4 (November 2013): 167–99, doi:http://www.aeaweb.org/aej-policy/.

this definition, San Francisco would qualify as a "superstar" city. This definition implies that the gap in house prices between cities can keep increasing, even when the inherent value of a location is constant, that a change in house price induces a change in the local income distribution (since land prices act as a clearing mechanism in the authors' model, by which higher-income households crowd out lower-income households from a scarce location), and that dispersion in expected house price growth rates are, in turn, expected to yield differences in price-to-rent ratios for houses (i.e., if home buyers expect their houses to appreciate over the long run, they should be willing to pay more, relative to renting, today). The supply and demand framework of housing implied by a "superstar" city is a relevant framework for considering the case of San Francisco's market for housing and for potential policies that could address housing affordability.

Glaeser et al. examine urban growth and housing supply by analyzing the elasticity of housing supply and the impact it has on generating either more populous cities or cities with higher paid workers and more expensive homes.[24] The model they present incorporates the heterogeneity of housing supply into urban development. The authors show that housing supply has become inelastic largely because of restrictive zoning and other land use regulations. Since housing is a durable good, the supply of housing, in theory, will be perfectly inelastic in areas where rents are not high enough to justify new construction. If demand for housing is sufficiently high, then the price of housing will be determined by the cost of new construction, i.e., the physical costs of construction plus the costs related to regulatory barriers. In places with high density and high levels of regulation, "urban success" is likely to leave population levels unchanged while leading to higher levels of housing prices and income (this is generally true for

[24] Edward L. Glaeser, Joseph Gyourko, and Raven E. Saks, "Urban Growth and Housing Supply," *NBER*, NBER Working Paper Series, January 2005, http://www.nber.org/papers/w11097.pdf.

San Francisco, where local economic success has driven population growth, but not by especially large amounts, and house prices and incomes have risen). The authors find that differences in housing supply cause higher housing prices and also impact cities' responses to increases in productivity. Their findings have policy implications for urban economics and the role of that housing supply in supporting labor markets and population growth in urban areas.

A large body of literature examines the impact of zoning on housing prices. In one paper, Glaeser et al. explain rising housing prices in Manhattan, where housing prices are roughly more than twice their supply costs.[25] The authors find that land use restrictions explain the gap between price and supply costs and show that regulation also constrains housing in other cities across the country. Since home-building is a highly competitive industry, with virtually no natural barriers to entry, price markups over construction costs are strong indicators of artificial barriers to new construction. In another paper, the authors find that the majority of US homes are priced close to their construction costs, but that, in coastal parts of the country, property values are much higher, at least in part because of strong demand for these high-amenity areas.[26] The authors find that the gap between total housing costs and the price of a structure is a combination of land costs and what they call a "zoning tax," which has a large impact on housing prices. The authors run three sets of regressions to test for the impact of zoning on housing prices. The first set of regressions determine the intrinsic value of land, with housing prices as a function of lot size and various control factors, and the results suggest that, on average, land tends to be cheap. The second set of regressions test whether high-cost areas have high population densities; if this relationship holds, then demand for land might be driving high housing prices, but, if these high-

[25] Edward L. Glaeser, Joseph Gyourko, and Raven E. Saks, "Why Is Manhattan So Expensive? Regulation and the Rise in Housing Prices," *The Journal of Law and Economics* 48 (October 2005): 331–69.
[26] Edward L. Glaeser and Joseph Gyourko, "Zoning's Steep Price," *Regulation*, Fall 2002.

price areas don't have exceedingly high population densities, then regulation might instead be driving high housing prices. The results suggest that amenities from locating in a certain city seem to have more of an effect on implicit zoning tax than on marginal cost of land. The third set of regressions use data from a survey of land-use restrictions to determine the correlation between the amount of zoning and land prices; the results find that high levels of zoning are ubiquitous in high cost areas.

The authors relay similar findings in another paper, in which they analyze whether or not America has a housing affordability crisis.[27] The authors argue that, in most of America, the price of housing is close to the marginal, physical cost of new construction, and, therefore, most of the country does not face a housing affordability crisis. The authors do find, however, that, in certain areas of the country, housing prices are significantly higher than construction costs and that these higher prices are due to zoning and land use controls. The authors provide evidence that, in cities like San Francisco, zoning and land use restrictions directly increase the price of housing.[28]

Residential Sorting and Housing Affordability

The concept of "sorting" also has a strong role in urban economics. Diamond (2015) examines the implication of residential sorting along education, skill, and amenities. Although changes in local labor demand within cities impacts sorting along labor market skills, amenities in higher-skill cities also influence sorting.[29] Changes in city amenities further exacerbate both

---

[27] Edward L. Glaeser and Joseph Gyourko, "The Impact of Zoning on Housing Affordability," *NBER*, NBER Working Paper Series, n.d., http://www.nber.org/papers/w8835.
[28] See: Joseph Gyourko, Albert Saiz, and Anita A. Summers, "A New Measure of the Local Regulatory Environment for Housing Markets: The Wharton Residential Land Use Regulatory Index," Forthcoming in Urban Studies (The Wharton School, University of Pennsylvania, March 29, 2007), http://real.wharton.upenn.edu/~gyourko/WRLURI/The%20Wharton%20Zoning%20Regulation%20Index-July%202,%202007.pdf.
[29] Rebecca Diamond, "The Determinants and Welfare Implications of US Workers' Diverging Location Choices by Skill: 1980-2000," August 27, 2015, http://web.stanford.edu/~diamondr/research.html.

sorting by education and inequality between high-school and college graduates. Using an equilibrium sorting model, Kuminoff et al. examine how households “sort” across neighborhoods according to their wealth and preferences for public and social goods.[30] Both the Kuminoff et al. and Diamond papers are useful to consider in the case of San Francisco, which is attracting an increasingly educated population because of local labor demand and because of the amenities the city offers to the majority college-educated population that fills the local labor market demand.

Bayer et al. develop an equilibrium framework to analyzing residential sorting, using data on households in the San Francisco Bay Area.[31] The authors’ estimates provide characterizations for preferences of housing and neighborhood attributes, showing how demand for certain attributes varies with a household’s income, race, education, and family structure. The authors also test for the effects of an increase in income inequality. In the equilibrium model of a self-contained urban housing market, households sort themselves among a set of available housing types and locations. The authors find that increased spending by the top quartile of households is reflected in significantly higher housing prices rather than more spending in the local economy.

Matlack and Vigdor also examine the relationship between income inequality and housing affordability.[32] The authors find that, in housing markets with low-vacancy rates, increases in income at the upper end of the distribution are associated with significantly higher rents and prices of housing. Inequality is not universally associated with higher housing costs or reduced consumption by the poor. Instead, impacts of inequality are concentrated in metropolitan areas where housing supply is relatively inelastic. The authors show that raising the income of

[30] Nicolai V. Kuminoff, V. Kerry Smith, and Christopher Timmins, “The New Economics of Equilibrium Sorting and Policy Evaluation Using Housing Markets,” *Journal of Economic Literature* 51, no. 4 (2013): 1007–62.
[31] Bayer, McMillan, and Rueben, “An Equilibrium Model of Sorting in an Urban Housing Market.”
[32] Matlack and Vigdor, “Do Rising Tides Lift All Prices?”

the wealthy increases the price that the poor must pay for housing. The policy implications of the authors' findings suggest that "trickle-down" government policies do not yield favorable outcomes for all.

Section IV. Data

My goal in this paper is to present estimates of the impact of income and other demographic characteristics on housing prices in San Francisco and examine the policy implications of my findings. In my empirical approach, which I discuss in detail in Section V, I look at only one point in time, using data from the most recent year available, 2014. The data I use was compiled from the US Census, using the Census's American Fact Finder (AFF) tool. While I initially planned to use a dataset from IPUMS, I decided that I preferred to use data from a smaller geographic level. Since the smallest geographic metric that IPUMS makes available is the Public Use Microdata Areas (PUMA)-level, I gathered data from the Census instead, since AFF makes data available on the census-tract level. I chose census tracts over PUMAs because census tracts provided a better approximation for neighborhoods than PUMAs, and house prices likely vary by neighborhoods. The City and County of San Francisco contains 197 census tracts but only 7 PUMAs, so data on the census tract-level also provides more observations than that from PUMAs. Unlike IPUMS, AFF is not a very user-friendly tool; there is no easy way to combine various datasets. Since variables of interest are only available in separate datasets, each dataset must be individually downloaded, cleaned, and then combined to yield the final dataset used. To obtain all the variables I needed, I downloaded and sorted eleven individual datasets. I then combined these datasets into one file and further cleaned it up so that the file would be compatible with the statistical software, Stata, that I used. While this was a time-consuming and tedious process, it ultimately yielded the dataset I wanted, on the census tract-level. When selecting data, I identified both housing and household-demographic characteristics that were of interest and that seemed relevant. I drew on the characteristics used by Bayer et al. as an example for relevant housing and household variables, and I initially pulled far more variables than I ended up using in my estimations.[33]

Variables

Since my research question aims to address the impact of housing and household demographic characteristics on housing prices, I pulled data from the AFF for these characteristics in San Francisco, from the most recent year available. All the data is from 2014 and is from the Census's American Community Survey (ACS).

The variables I use in my regression can be organized into two sets – one set of variables describes housing characteristics, while the other describes household/demographic characteristics. Each observation of every variable occurs on the census tract-level; dropping blank values, this yielded 184 observations per each variable (i.e., 184 observations for each of the 184 census tracts for which full data was available). As I explain in detail in Section V, I report all variables in percentage terms, to control for differences in the number of housing units and households per census tracts.

The dependent variable of interest is housing price, which is given in the dataset by a variable that is an estimate of the value of owner-occupied units in median dollars. A limitation of this variable in the data is that housing unit values above $1,000,000 are not reported; they are simply listed as $1,000,000+. It is therefore difficult to know the exact value of owner-occupied units that are valued at over $1,000,000.

Independent variables that describe housing characteristics include the percent of units in a tract:

- with one to two bedrooms (note: units with zero bedrooms dropped for collinearity)

[33] Bayer, McMillan, and Rueben, "An Equilibrium Model of Sorting in an Urban Housing Market."

- with three or more bedrooms
- built 1969 or earlier
- built 1980 or later (note: built 1970-1979 dropped for collinearity)
- with a one-unit attached through up to four units in a structure (note: one-unit detached dropped for collinearity)
- with five or more units in a structure
- with two or more vehicles
- with no vehicles (note: one vehicle dropped for collinearity)
- owner-occupied, without a mortgage

Independent variables that describe household/demographic characteristics include

- Median household income in the past 12 months (in 2014 inflation-adjusted dollars)

and the percent of households per census tract:

- who moved recently (2000 or later)
- who moved 1990-1999 (note: moved 1980-1989 dropped for collinearity)
- who moved in 1979 or earlier
- that are Black
- that are Asian
- that are White alone
- that are Hispanic
- with a Bachelor's degree (BA) or more
- on food stamps
- with children under 18
- that have not worked in the past year
- that are age 44 and under
- that are age 45 to 64
- that are age 65 and older (note: ages 75-84 dropped for collinearity)

Descriptive Statistics

Tables 1 – 3 show summary statistics for the variables described above. Table 1, Table 2, and Table 3 provide summary statistics for the variables describing each of the three occupancy types: all occupied-housing units, only owner-occupied housing units, and only renter-occupied housing units, respectively. Tables 4 – 6 show detailed summary statistics on median household income for each of the three occupancy types.

Median household income of owner-occupied households, at $116,734, is almost double that of renter-occupied households, whose median household income is $65,547. The median household income for all occupied households is $81,427. The median housing unit price in the data is $796,050. However, since data on housing prices over $1,000,000 is not specified, i.e., if a housing unit is valued at over $1,000,000, the data only reports $1,000,000+, the reported median housing unit price of $796,050 may be lower than the actual median housing unit price in San Francisco. The number of observations for each variable, N, is 184, after dropping blanks and calculating percentages of each. The means of the independent variables across the occupancy types (all occupied, owner-occupied, and renter-occupied), were identical for housing characteristics (price, bedrooms, year built, units per structure, vehicles) and for food stamps, children, and work status. For household and demographic characteristics, the means of the independent variables across the occupancy types differed but were generally close in value. Overall, the summary statistics suggest that there is not wide variation across the variables for the three occupancy types.

Section V. Methodology: Econometric Models

The methodology I use in my analysis is quite simple. I use hedonic price regressions to determine the effect of various characteristics on housing price (I present a theoretical approach using a more structural model in Appendix C). Hedonic price regressions are ordinary least squares (OLS) regressions and are therefore fairly straightforward.

Hedonic regressions are a way of statistically estimating the relationship between a property's characteristics and its market value, thus allowing one to determine the value of the property itself.[34] A hedonic price regression is essentially a conventional regression that shows

---

[34] Edward Coulson, "Chapter 2: A Brief Survey and Interpretation of Hedonic Parameters," in *Hedonic Methods and Housing Markets* (Department of Economics, Penn State University, 2008), http://grizzly.la.psu.edu/~ecoulson/hedonicmonograph/monog.htm.

how the price of a housing unit is affected by its characteristics. A hedonic function links these characteristics, collectively defined as *X*, to the price of the real estate product, *P*, given by:

$$P = H(X)$$

The above function, *H(X)*, can be depicted in a linear function, in which:

$$P = \alpha_0 + \alpha_1 X_1 + \alpha_2 X_2 + \cdots + \alpha_k X_k$$

where X_1 through X_k are the attribute levels for *k* selected characteristics, and α_1 through α_k are the weights assigned to particular attributes. The linear function implies that a one unit increase in X_1 leads to an increase in the price of the property by α_1 dollars.

To determine the housing price, the parameters of the function, α_1 through α_k must be estimated. One can estimate the parameters using the model below, in which the number of equations (N) is greater than the number of unknowns (k+i), leaving us with an error term e_i,

$$P = \alpha_0 + \alpha_1 X_{1i} + \alpha_2 X_{2i} + \cdots + \alpha_k X_{ki} + e_i$$

The goal of the estimation problem is to minimize the errors through appropriate choices of α_i's (i.e., choose α_i's that minimize the sum of the squared errors), which allows one to claim that the hedonic price is estimated in a way that best represents how the characteristics relate to the housing prices. The examples above are for linear hedonic functions; to more accurately estimate price, I use the log-linear functional form.

Nonlinear pricing more accurately represents housing prices than linear pricing. Linear pricing is standard when competitive pressures can force the untying of tied bundles.[35] Consider an example of laundry detergent, for which a 10 pound box costs $10, but a 20 pound box costs $15. This is an example of nonlinear pricing because the first 10 pounds of detergent has a different price than the second 10 pounds ($10 versus $5). This kind of nonlinear pricing creates

[35] Ibid.

opportunity for arbitrage, but the fact that arbitrageurs do not actually take advantage of this suggests that it is not "economical" to do so, i.e., that the transaction costs of engaging in arbitrage are not zero. The case can be extended to housing units; physical housing characteristics are, in general, tied together in an inseparable bundle. For this reason, housing prices are more accurately estimated by nonlinear prices, and I therefore use a log-linear functional form in my estimations.

A generic log-linear functional form is given by:

$$\log(P_i) = a_0 + \Sigma_j a_j X_{ij} + e_i$$

where the coefficients are semi-elasticities; the coefficients give the percentage increase in price due to a unit increase in *X*.

I estimate three models in three sets of regressions, i.e., three models each for a) all occupied housing units, b) only owner-occupied housing units, and c) only renter-occupied housing units. The first model is a regression of the independent variables (household and housing characteristics), on the dependent variable, the logarithm of housing price. The regression in Model 1 is given by:

$$\begin{aligned}
&\log(median\ housing\ unit\ price) \\
&\quad = \beta_1\left(X_{bedrooms\ per\ unit}\right) + \beta_2\left(X_{year\ unit\ built}\right) + \beta_3(X_{units\ in\ structure}) \\
&\quad + \beta_4(percent\ without\ mortgage) + \beta_5(\log(median\ household\ income)) \\
&\quad + \beta_6\left(X_{vehicles\ available\ per\ unit}\right) + \beta_7\left(X_{year\ household\ moved\ into\ unit}\right) \\
&\quad + \beta_8\left(X_{households\ by\ race,ethnicity}\right) \\
&\quad + \ \beta_9(education\ status\text{: } percent\ with\ BA\ or\ more) \\
&\quad + \beta_{10}(percent\ of\ households\ on\ food\ stamps) \\
&\quad + \beta_{11}(percent\ of\ households\ with\ children\ under\ 18) \\
&\quad + \beta_{12}\left(X_{age\ of\ household}\right) + e_i
\end{aligned}$$

where

$X_{bedrooms\ per\ unit} = \begin{bmatrix} percent\ of\ 1\ to\ 2\ bedrooms \\ percent\ of\ 3\ or\ more\ bedrooms \end{bmatrix}$ (with 0 bedrooms per unit dropped for collinearity)

$X_{year\ unit\ built} = \begin{bmatrix} percent\ built\ 1969\ or\ earlier \\ percent\ built\ 1980\ or\ later \end{bmatrix}$ (with units built 1970-1979 dropped for collinearity)

$X_{units\ in\ structure} = \begin{bmatrix} percent\ with\ 1\ to\ 4\ units, attached \\ percent\ with\ 5\ or\ more\ units \end{bmatrix}$ (with 1-unit detached structures dropped for collinearity)

$X_{vehicles\ available\ per\ unit} = \begin{bmatrix} percent\ with\ 0\ vehicles \\ percent\ 2\ or\ more\ vehicles \end{bmatrix}$ (with 1 vehicle available dropped for collinearity)

$$X_{households\ by\ race/ethnicity} = \begin{bmatrix} Percent\ of\ black\ households \\ Percent\ of\ Asian\ households \\ Percent\ of\ white\ alone\ households \\ Percent\ of\ Hispanic\ households \end{bmatrix}$$

$X_{number\ of\ households\ by\ age} = \begin{bmatrix} Percent\ 44\ and\ under \\ Percent\ 45\ to\ 64 \\ Percent\ 65\ and\ over \end{bmatrix}$ (with households age 75 to 84 dropped for collinearity)

$X_{year\ household\ moved\ into\ unit} = \begin{bmatrix} percent\ moved\ 2000\ or\ later \\ Percent\ moved\ 1990 - 1999 \\ Percent\ moved\ 1979\ or\ later \end{bmatrix}$ (with moved 1980-1989 dropped for collinearity)

and e_i is the error term.

Terms were dropped for the variables specified above to avoid collinearity. Additionally, all the independent variables are reported in percentage terms, to control for differences in the number of housing units and households across census tracts. For example, consider households by race per census tract. Let us assume there are 5 Black households in census tract A, out of a total of 20 households. This implies that 25% of households in tract A are Black. In census tract B, there are 5 Black households out of a total of 100 households, implying only 5% of households are Black in Tract B, even though there are 5 Black households in each tract A and B. Since we are interested in determining, for example, the impact of race on housing price, using percentages allows for all demographic characteristics to have the same denominator. Similarly, housing characteristics are also converted to percentages, so that, for instance, instead of a count of vehicles per housing unit in each census tract, we have a percentage of vehicles per unit in each census tract.

In Model 2, I add a quadratic term, Income^2, to test the quadratic relationship between income and housing price. The quadratic is useful to consider, because, presumably, housing prices increase with income (i.e., people with a higher income will likely purchase a more expensive house). However, at a certain threshold of income, housing prices will begin to increase at a decreasing rate, or even plateau (i.e., additional increases in income will not necessarily yield increases in housing price of comparable magnitude to the increase in income). Model 2 is the same as Model 1 above, but with an added Income^2 term:

$$\log(median\ housing\ unit\ price)$$
$$= \beta_1(X_{bedrooms\ per\ unit}) + \beta_2(X_{year\ unit\ built}) + \beta_3(X_{units\ in\ structure})$$
$$+ \beta_4(percent\ without\ mortgage) + \beta_5(\log(median\ household\ income))$$
$$+ \beta_6(X_{vehicles\ available\ per\ unit}) + \beta_7(X_{year\ household\ moved\ into\ unit})$$
$$+ \beta_8(X_{households\ by\ race,ethnicity})$$
$$+ \beta_9(education\ status{:}\ percent\ with\ BA\ or\ more)$$
$$+ \beta_{10}(percent\ of\ households\ on\ food\ stamps)$$
$$+ \beta_{11}(percent\ of\ households\ with\ children\ under\ 18)$$
$$+ \beta_{12}(X_{age\ of\ household}) + \beta_{13}(\log median\ household\ income)^2 + e_i$$

where log(*median household income*)^2 is the square of the logarithm of median household income.

In Model 3, I build on Model 1 by adding interaction terms. I add interaction terms between race/ethnicity and income. The interaction terms test the hypothesis that the impact of race on housing price differs by income. The interaction terms of *RaceXIncome* are meant to show that the effect of race/ethnicity on housing price is different for different values of income. Model 3 is the same as Model 1, but with added interaction terms for each race/ethnicity:

$$\log(median\ housing\ unit\ price)$$
$$= \beta_1(X_{bedrooms\ per\ unit}) + \beta_2(X_{year\ unit\ built}) + \beta_3(X_{units\ in\ structure})$$
$$+ \beta_4(percent\ without\ mortgage) + \beta_5(\log(median\ household\ income))$$
$$+ \beta_6(X_{vehicles\ available\ per\ unit}) + \beta_7(X_{year\ household\ moved\ into\ unit})$$
$$+ \beta_8(X_{households\ by\ race,ethnicity})$$
$$+ \beta_9(education\ status: percent\ with\ BA\ or\ more)$$
$$+ \beta_{10}(percent\ of\ households\ on\ food\ stamps)$$
$$+ \beta_{11}(percent\ of\ households\ with\ children\ under\ 18)$$
$$+ \beta_{12}(X_{age\ of\ household}) + \beta_{13}(X_{BlackXincome}) + \beta_{14}(X_{WhiteXincome})$$
$$+ \beta_{15}(X_{AsianXincome}) + \beta_{16}(X_{HispanicXincome}) + e_i$$

I run all three models for each of the three housing unit types – all occupied housing units, owner-occupied housing units, and renter-occupied housing units.

Section VI. Results

The regressions described above aim to to answer my research question by estimating the effect of income and other demographic characteristics on the median housing unit price in a census tract. The results suggest that income does not impact price as much as other factors, such as the number of bedrooms in a housing unit and the educational level of the household. While certain limitations, discussed in detail in Section VII below, may contribute to the lack of significance of income's effect on housing price, the results are nonetheless useful to consider in the context of demographic characteristics' impact on housing prices.

Table 7 shows the results of all three models, per each type of housing occupancy, i.e., for all nine variations of the model. The models are numbered 1, 2, 3, and letters a, b, and c denote the independent variables by type of housing unit. Letter (a) represents regressions that include data on all occupied housing types, (b) includes data only on owner-occupied housing units, and (c) includes data only on renter-occupied housing units. Robust standard errors of each coefficient, which measure the accuracy of an estimation's prediction, are listed below each variable in parentheses, and the statistical significance of a coefficient is starred. All variations of the model have a relatively high R^2 value, which indicates that the model fits the data well; in this case, the models explain that roughly 70-80% of the variation in median housing unit price is due to the variation in the independent variables. R^2 values, however, cannot determine whether coefficients are biased.

The effect of *log(median household income)* on *log(median housing unit price)* is only statistically significant in two iterations of the model, 1.a and 2.b, and the results are not economically significant. These results are somewhat unexpected and may be due to some of the limitations in this study that are discussed in Section VII. The number of bedrooms in a housing unit, both *1 to 2 bedrooms* and *3 or more bedrooms*, however, have a positive impact on median housing unit price and are statistically and economically significant across all nine variations of the model. For example, in model 1.a, increasing the percent of units per tract with 0 bedrooms to 1 bedroom or from 1 to 2 bedrooms by 1% leads to, on average, an estimated increase in the median housing unit price by 2%. Intuitively, it makes sense that more bedrooms in a housing unit would increase the value, or price, of that housing unit, and the estimated impact is indeed economically significant. Drawing on the descriptive statistics presented in Section IV, consider that the median housing unit price is roughly $800,000. A 2% increase in median housing unit price would therefore increase the price of the housing unit by roughly $16,000, an increase that, arguably, is economically significant.

Education also has a statistically significant, but not necessarily economically significant, impact on median housing unit price across all but one of the nine variations of the models. On average, a 1% increase in householders within a census tract with an education at the BA level or more raises median housing unit price in a census tract by 0.5%. For a median housing unit price of roughly $800,000, this would increase the price by about $4,000, an amount that may not be economically significant. Households with higher levels of education likely sort into housing that is more expensive, which can help explain why the median value of a housing unit in a census tract increases when more householders in that census tract have higher educational attainment. This phenomenon may also be explained by a self-selecting sorting model. For example, households with higher education levels may self-select into census tracts that have pre-existing high levels of educational attainment (for a related model of residential sorting, see Appendix C).

Other variables of note include vehicle ownership by housing units. The percentage of housing units with two or more cars per housing unit has a statistically significant, albeit negative, impact on housing price across all nine model variations. While this result may not seem intuitive, in the sense that households who can afford to have more vehicles per housing unit can likely also afford more expensive housing, perhaps householders spend more money on their vehicles and thus have less money to spend on housing. Also of note is that a one percent increase in the percentage of Black and Hispanic households per census tract is estimated to have a negative impact on median housing unit price, and that a one percent increase in the percentage of White alone households per census tract is estimated to have a positive impact on price, in certain statistically significant variations of the model. However, White alone households also have a negative impact on price in one of the models, which suggests that this variable may not be a good estimator for median housing price. Furthermore, none of the race/ethnicity estimations have large economic significance. Overall, the results of my estimations did not necessarily fit expectations from the literature, but this is likely due to the limitations of my study.

Section VII. Limitations and Expanding on this Study

Given the documented role of income inequality in driving up housing prices, my results – that income, on average, did not have an economically or statistically significant impact on housing prices – are somewhat unexpected. However, other house and household characteristics were statistically and economically significant in impacting housing price, although the extent of the estimations' economic significance varied. To reconcile my original hypothesis and the existing literature with my results, I discuss some limitations and areas for expansion.

One of the limitations of my regression estimates is the low number of observations. After changing variable counts to percentages and ensuring that all variables were available for all census tracts, this dropped the number of observations from 197 to 184. While neither are large to begin with, 184 is likely not enough observations to yield statistically or economically significant results, and my results reflect that limitation. An expansion of my study to improve this would be to increase the sample size by conducting the same study for all nine counties that make up the Bay Area; especially if this was done on the census tract-level, this would generate a large enough sample size to yield better results.

Furthermore, another limitation from the dataset I use is that the income variable is median household income for the entire census tract, i.e., there is only one measure of income for an entire census tract, which does not necessarily account for the variation in income that can occur within a tract. Ideally, this study would be conducted with data on the income of every household in a census tract, which would yield a much larger dataset and more statistically significant results. A related limitation of the current dataset is that the housing price variable only lists housing unit values up to $1,000,000; any unit valued at over $1,000,000 is simply listed as $1,000,000+ in the original data. In order to be compatible with the statistical software I used, all the $1,000,000+ values were changed to $1,000,000. This data likely does not accurately represent the true housing unit prices in San Francisco, and the median housing unit price in my dataset is likely much lower than the actual price. A better dataset would therefore also include more accurate housing prices. Furthermore, because I am using median housing unit price per census tract as my data point, this measure is only one value and does not capture the full range of housing prices that may be present in a given census tract. For example, there may be a handful of housing units in the Mission District census tracts that are priced at over $1 million, but others that are priced much lower, so the million-plus outliers might be pulling up the median. To mitigate this data issue, it would be ideal to have data on housing price for each housing unit within a census tract, but data on such a micro level is difficult to obtain.

Another limitation of my study is that there may not be enough or the correct housing and neighborhood characteristics included in the regressions. Since housing characteristics are what hedonic price regressions control for, including more of these, or different ones, could yield better results. For example, data on housing characteristics – such as square footage, number of bathrooms, pool, etc. – and on neighborhood variables – such as crime rates, schools, local businesses, proximity to amenities and public transit, etc. – would modify my estimations of housing unit price. Data on these characteristics, however, is difficult to obtain.

Related to data limitations is the potential problem of omitted variable bias (OVB). If OVB is present in the regressions above, that indicates that there is an unaccounted for component in the error term that is correlated with both the dependent variable and one or more of the independent variables, and is therefore biasing the coefficient on those variables either positively or negatively. For example, I do not include a house's square footage as an independent variable in my regressions, but one could imagine that square footage is correlated with both house price and the number of bedrooms per house unit (one of the independent variables I do include). To correct for OVB, more data would be needed on house characteristics or other potentially important variables. If the data is not easily available, then there are econometric methods that can solve for OVB.

One approach for correcting OVB is to use instrumental variables. Instruments must be uncorrelated with the errors and correlated with an independent variable. Determining a valid instrument is challenging, as is potentially obtaining data on that instrument, but pursuing this strategy would be a good expansion on this study. Another potential solution for OVB is using a fixed effects model, in which we assume there is some fixed effects estimator that fully justifies the correlation between an independent variable and the error term and that is constant over time. An example of a potential fixed effect is census tract (geography); however, census tracts do not stay constant over time, and therefore finding a valid fixed effects estimator is a further challenge to solving potential OVB in this study.

In addition to the data expansions and OVB corrections posed above, another potential expansion would be to isolate the effect of income inequality on housing prices in specific neighborhoods. With a larger dataset, e.g., one that included median income and housing unit price for every household within a census tract, I could group census tracts into neighborhoods and then estimate the effect of rising, or falling, income inequality in that specific neighborhood on housing prices. For example, in recent years there has been an influx of affluent households moving into the Mission District neighborhood of San Francisco, a previously predominantly working class and minority neighborhood. I could use hedonic price regressions to determine the impact of growing income inequality on house prices in that neighborhood, and potentially compare the changes over time and over different neighborhoods. For instance, I could examine the relationship between income and house prices in the Mission in previous decades, and run a comparison over time and across other neighborhoods that have had more constant median incomes, such as Pacific Heights, a traditionally affluent neighborhood, or Bayview-Hunters Point, a traditionally low-income neighborhood.

While this study has limitations and potential for expansions, the results offer a contribution to the role of income inequality on housing prices, and I discuss potential policy implications in the conclusion below.

Section VIII. Conclusion: Policy Implications

This paper contributes to a large literature in economics that shows how income inequality affects housing prices. While the main results of my estimations do not suggest a statistically or economically significant relationship between income and housing prices in San Francisco, this is likely due to the limitations described above, and the study could be expanded to overcome these limitations and yield better estimations.

While it is beyond the direct scope of this paper, in Appendix C I present a model of residential sorting. How households choose where to live impacts housing prices in that neighborhood. Residential choice is a function of information about the housing unit, neighborhood, affordability, and other factors. Households often sort into neighborhoods where there are other similar households. This further drives up housing prices, which, in turn, crowds out households with lower incomes and hurts local economies, since higher income householders are increasing spending on housing instead of on the local economy. Maintaining income diversity, therefore, helps keep housing prices affordable, but this is not possible if less affluent households are driven out of a city by high housing prices.

Another policy implication for housing affordability is that regarding the impact of housing supply on housing prices. Although this is outside of the models I estimate, it is useful to consider the role restrictive building regulations have on constraining housing supply. San Francisco has highly restrictive zoning and building codes, which likely impacts the city's inability to keep housing supply at pace with demand and which further raises housing prices. Furthermore, higher income households are more likely to be involved in resistance to rezoning, which has a negative impact on development and housing supply. In this way, the combination of San Francisco's already restrictive building environment, and a growth of higher income households who influence the land use regulation process, impacts housing affordability. Policies that loosen restrictive building regulations could increase housing supply and also potentially alleviate high housing prices.

Ultimately, housing affordability in cities such as San Francisco impacts the diversity of the city and the ability of the local labor force to live in the city where it works. While San Francisco has experienced growth in higher-wage sectors of its economy, a lower-wage economy that provides services for those with higher-incomes has grown alongside the higher-wage sectors of the economy. If lower-wage workers cannot afford to live in or near the city they work in, then this can decrease productivity in the local economy. For instance, firms in a region cannot expand without new, and affordable, housing to house their workers. While my analysis focuses on a small subset of the housing affordability issue in San Francisco, the implications are applicable to determinants of housing supply and pricing in any urban area with constrained housing markets and income inequality.

Figures and Tables

Figure 1: Total population by decade*

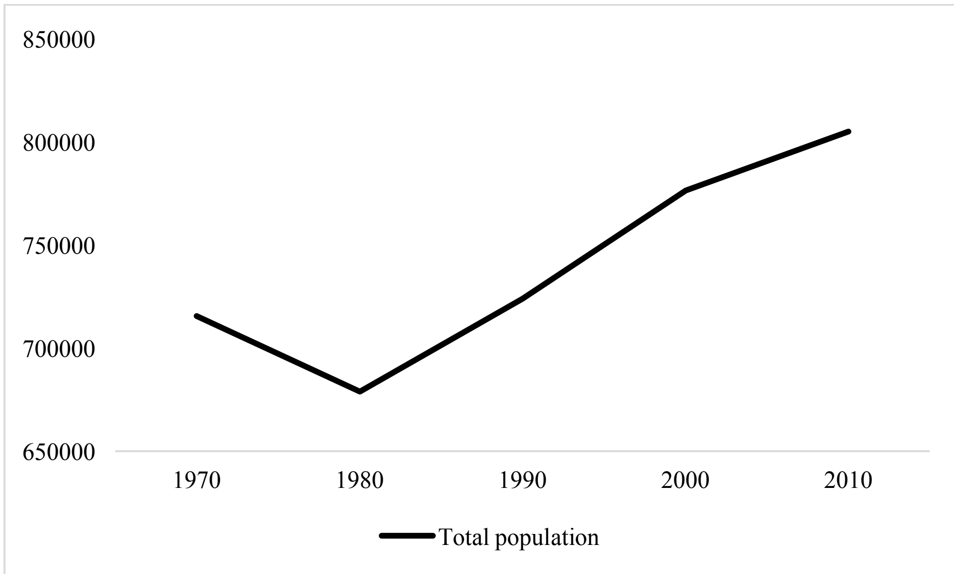
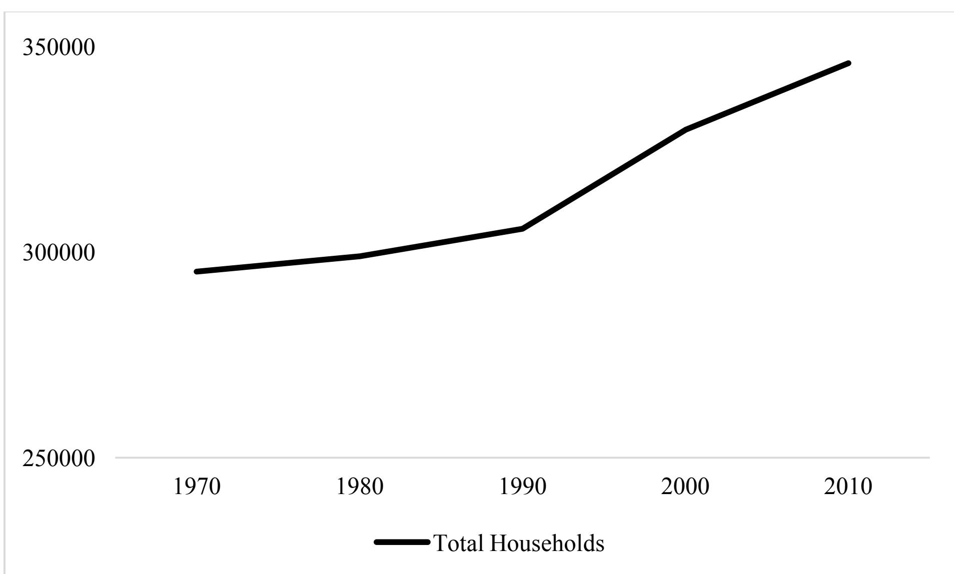


Figure 2: Total households by decade



* *Note: Sources for Figures 1-9 are SF Housing Data Hub; author's calculations.*

Figure 3: Total housing units by decade

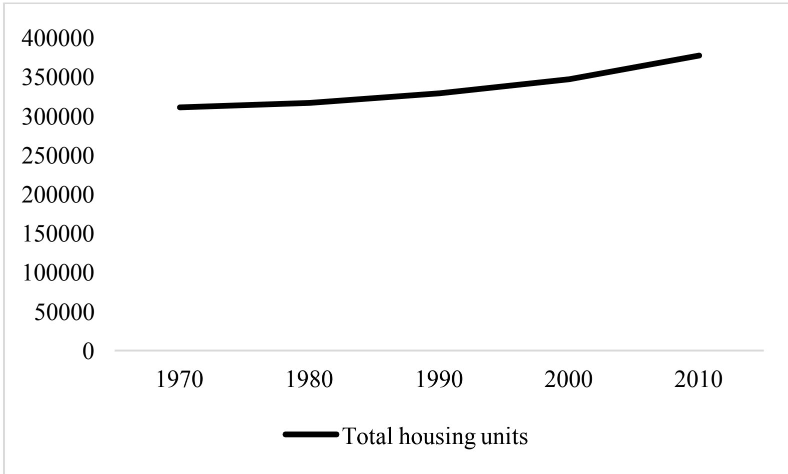


Figure 4: Number of housing units built by year

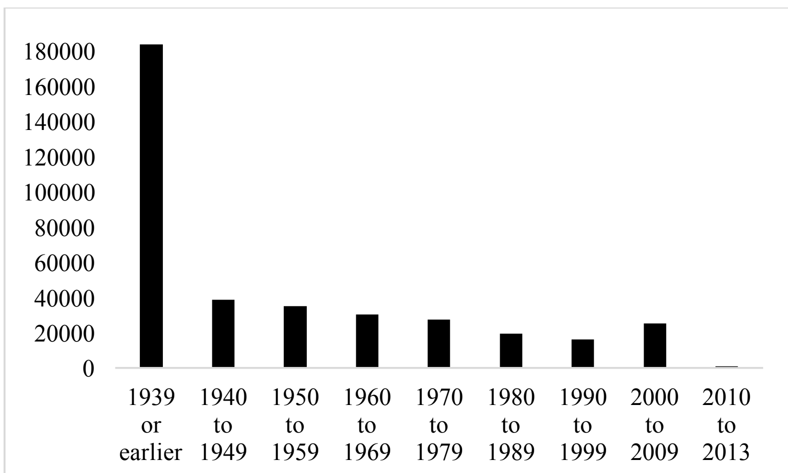


Figure 5: Occupied & vacant housing units by decade

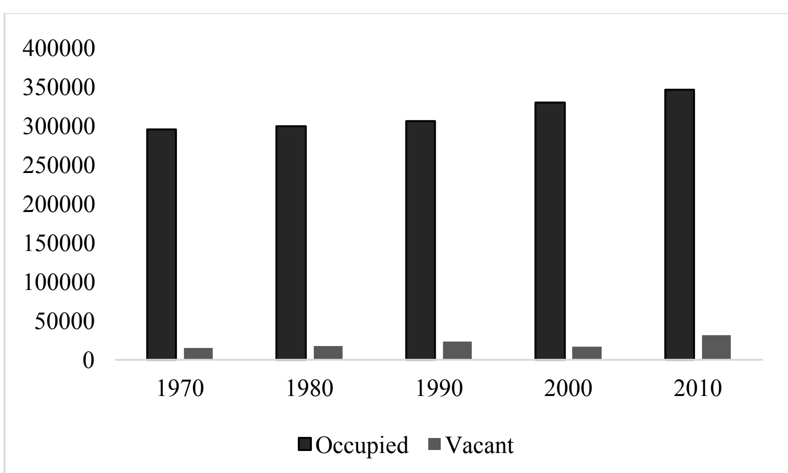


Figure 6: Share of renter and owner occupied units by decade

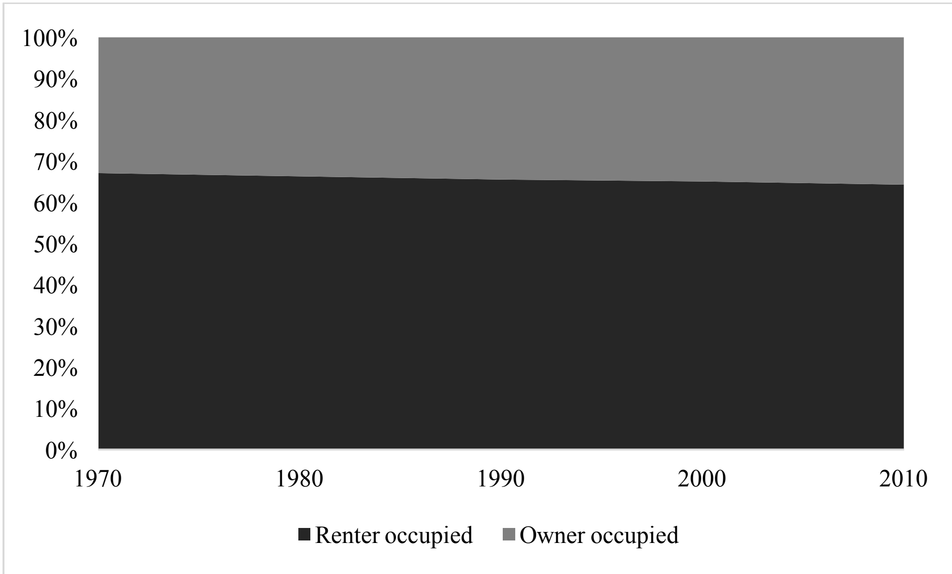


Figure 7: Count of renter and owner occupied units by decade

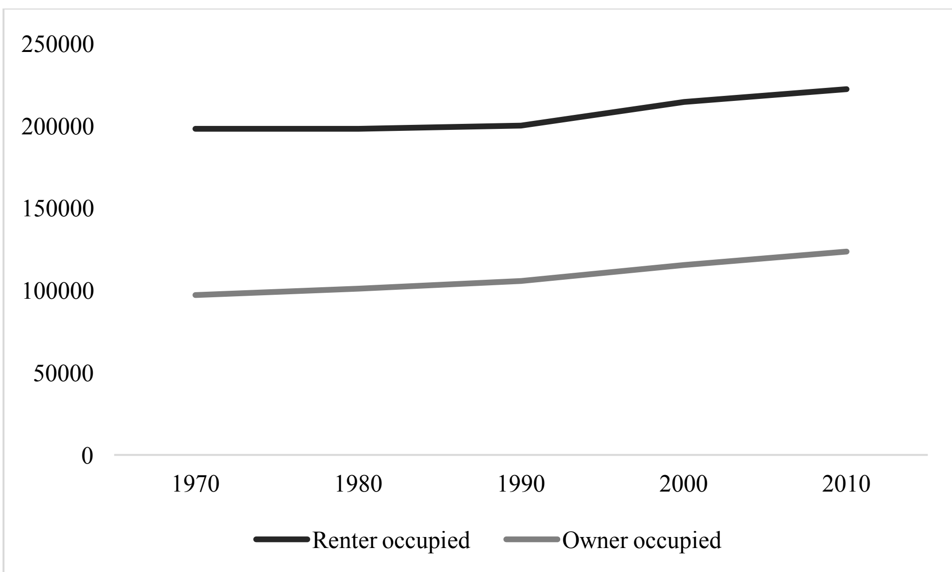


Figure 8: Median monthly housing costs, 2013

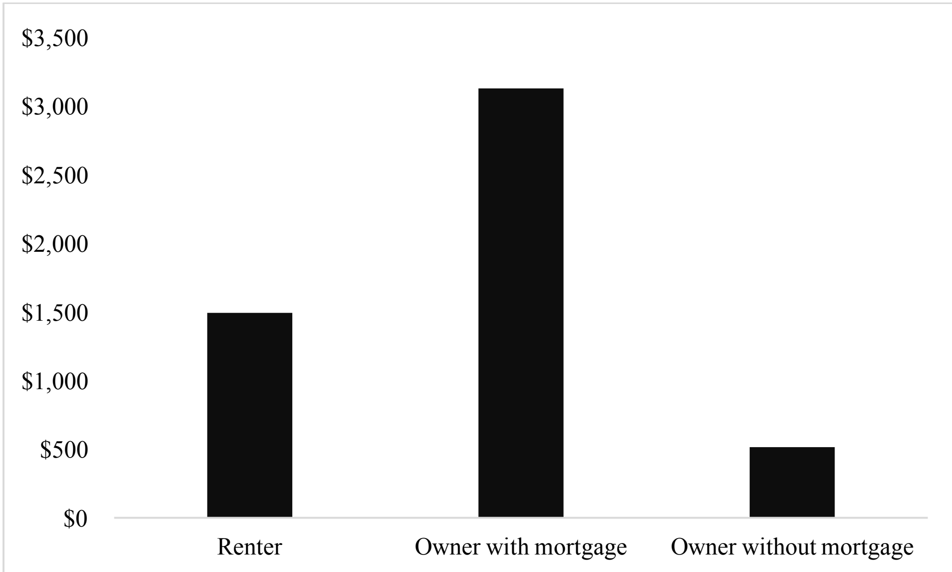


Figure 9: Housing costs as share of household income, 2013

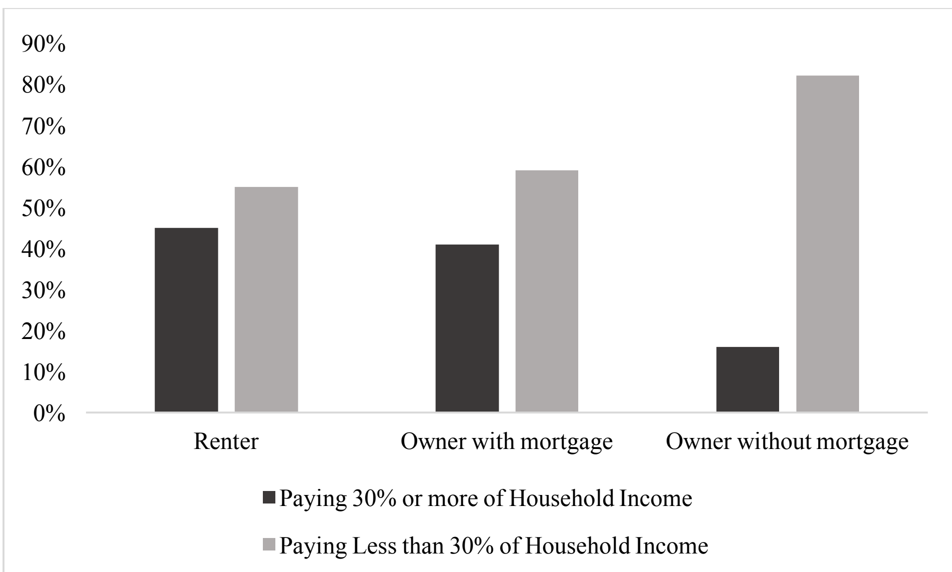


Table 1: All occupied housing summary statistics

	N	Mean	Std. Dev.	Min
Log of housing price	184	13.53	0.27	12.57
1 to 2 bedrooms	184	0.29	0.06	0.10
3 or more bedrooms	184	0.10	0.06	0.00
Built 1969 or earlier	184	0.19	0.05	0.00
Built 1980 or later	184	0.04	0.04	0.00
1 to 4 units, attached	184	0.12	0.06	0.00
5 or more units	184	0.13	0.11	0.00
Owner-occupied without mortgage	184	0.30	0.14	0.00
Log of median household income	184	11.24	0.49	9.49
2 or more cars	184	0.16	0.09	0.01
0 cars	184	0.27	0.19	0.02
Moved 2000 or later	184	0.33	0.06	0.19
Moved 1990-1999	184	0.16	0.05	0.01
Moved 1979 or earlier	184	0.05	0.03	0.00
Black households	184	0.07	0.11	0.00
Asian households	184	0.28	0.17	0.04
White alone households	184	0.49	0.22	0.04
Hispanic households	184	0.12	0.10	0.01
Education: BA or more	184	0.56	0.21	0.09
Food stamps	184	0.84	0.93	0.00
Households with children under 18	184	0.21	0.11	0.01
Work status in past year: no work	184	0.13	0.08	0.00
Age 44 and under	184	0.22	0.07	0.08
Age 45 to 64	184	0.18	0.05	0.06
Age 65 and older	184	0.08	0.03	0.00

Table 2: Owner-occupied housing summary statistics

	N	Mean	Std. Dev.	Min	Max
Log of housing price	184	13.53	0.27	12.57	13.82
1 to 2 bedrooms	184	0.29	0.06	0.10	0.45
3 or more bedrooms	184	0.10	0.06	0.00	0.27
Built 1969 or earlier	184	0.19	0.05	0.00	0.24
Built 1980 or later	184	0.04	0.04	0.00	0.24
1 to 4 units, attached	184	0.12	0.06	0.00	0.28
5 or more units	184	0.13	0.11	0.00	0.33
Owner-occupied without mortgage	184	0.30	0.14	0.00	0.88
Log of median household income	184	11.63	0.37	9.57	12.34
2 or more cars	184	0.16	0.09	0.01	0.35
0 cars	184	0.27	0.19	0.02	0.84
Moved 2000 or later	184	0.09	0.20	0.01	1.85
Moved 1990-1999	184	0.06	0.27	0.00	3.71
Moved 1979 or earlier	184	0.02	0.03	0.00	0.21
Black households	184	0.04	0.09	0.00	0.60
Asian households	184	0.34	0.22	0.00	1.00
White alone households	184	0.50	0.24	0.00	0.92
Hispanic households	184	0.09	0.09	0.00	0.49
Education: BA or more	184	0.63	0.21	0.10	1.00
Food stamps	184	0.84	0.93	0.00	5.33
Households with children under 18	184	0.21	0.11	0.01	0.56
Work status in past year: no work	184	0.13	0.08	0.00	0.45
Age 44 and under	184	0.15	0.07	0.03	0.37
Age 45 to 64	184	0.21	0.05	0.00	0.35
Age 65 and older	184	0.12	0.05	0.00	0.31

Table 3: Renter-occupied housing summary statistics

	N	Mean	Std. Dev.	Min	Max
Log of housing price	184	13.53	0.27	12.57	13.82
1 to 2 bedrooms	184	0.29	0.06	0.10	0.45
3 or more bedrooms	184	0.10	0.06	0.00	0.27
Built 1969 or earlier	184	0.19	0.05	0.00	0.24
Built 1980 or later	184	0.04	0.04	0.00	0.24
1 to 4 units, attached	184	0.12	0.06	0.00	0.28
5 or more units	184	0.13	0.11	0.00	0.33
Log of median household income	184	10.99	0.54	9.38	11.94
2 or more cars	184	0.16	0.09	0.01	0.35
0 cars	184	0.27	0.19	0.02	0.84
Moved 2000 or later	184	0.06	0.10	0.01	0.99
Moved 1990-1999	184	0.02	0.06	0.00	0.73
Moved 1979 or earlier	184	0.00	0.00	0.00	0.02
Black households	184	0.08	0.14	0.00	0.93
Asian households	184	0.23	0.15	0.00	0.89
White alone households	184	0.49	0.23	0.00	0.91
Hispanic households	184	0.15	0.13	0.00	0.54
Education: BA or more	184	0.53	0.23	0.05	0.90
Food stamps	184	0.84	0.93	0.00	5.33
Households with children under 18	184	0.21	0.11	0.01	0.56
Work status in past year: no work	184	0.13	0.08	0.00	0.45
Age 44 and under	184	0.28	0.06	0.08	0.41
Age 45 to 64	184	0.15	0.05	0.04	0.30
Age 65 and older	184	0.06	0.04	0.00	0.27

Table 4: Detailed summary statistics of median household income for all occupied housing units

	Percentiles	Smallest
1%	13983	13164
5%	25904	13983
10%	38977	15546
25%	61978	16677
50%	81427	
		Largest
75%	107528	147976
90%	126887	152143
95%	138021	156719
99%	156719	175313

Table 5: Detailed summary statistics of median household income for owner-occupied housing units

	Percentiles	Smallest
1%	42188	14333
5%	65909	42188
10%	72024	48438
25%	91240.5	50833
50%	116733.5	
		Largest
75%	149812.5	214643
90%	170234	218074
95%	187375	226250
99%	226250	229667

Table 6: Detailed summary statistics of median household income for renter-occupied housing units

	Percentiles	Smallest
1%	12367	11903
5%	19561	12367
10%	26850	12462
25%	47022	13079
50%	65547	
		Largest
75%	87145	138462
90%	109878	151484
95%	117543	152500
99%	152500	153750

(see Tables B1-B3 in Appendix B for more detailed statistics on Tables 4-6)

Table 7: Regression Results

Dependent Variable	Log (median house price)								
	(1.a)	(2.a)	(3.a)	(1.b)	(2.b)	(3.b)	(1.c)	(2.c)	(3.c)
1 to 2 bedrooms	2.02***	2.09***	2.22***	1.78***	1.70***	2.01***	1.76***	1.82***	1.84***
	(0.56)	(0.56)	(0.57)	(0.61)	(0.60)	(0.63)	(0.61)	(0.61)	(0.60)
3 or more bedrooms	3.29***	3.45***	3.81***	2.09**	1.94*	2.90***	3.21***	3.40***	3.15***
	(0.98)	(0.98)	(1.01)	(1.04)	(1.02)	(1.05)	(1.10)	(1.41)	(1.10)
Built 1969 or earlier	0.85	0.75	0.87	1.40**	1.34**	0.82	0.37	0.35	0.73
	(0.62)	(0.57)	(0.58)	(0.58)	(0.57)	(0.65)	(0.68)	(0.66)	(0.56)
Built 1980 or later	-0.45	-0.61	-0.55	-0.30	-0.24	-0.51	-0.74	-0.78	-0.20
	(0.74)	(0.71)	(0.73)	(0.68)	(0.66)	(0.66)	(0.77)	(0.76)	(0.64)
1 to 4 units, attached	0.07	0.18	0.13	-0.73**	-0.57	-0.54	0.13	0.15	0.35
	(0.38)	(0.41)	(0.40)	(0.37)	(0.35)	(0.35)	(0.43)	(0.44)	(0.40)
5 or more units	0.11	0.23	0.43	-0.90**	-0.90**	-0.53	0.15	0.16	0.16
	(0.42)	(0.45)	(0.47)	(0.45)	(0.44)	(0.45)	(0.47)	(0.47)	(0.44)
Owner-occupied without mortgage	-0.17	-0.14	-0.21	0.09	0.05	-0.02	.	.	.
	(0.15)	(0.16)	(0.15)	(0.17)	(0.17)	(0.19)	.	.	.
Log of median household income	-0.04***	-1.47	0.28	0.09	-2.24***	0.02	0.01	-0.75	-0.01
	(0.08)	(1.22)	(0.52)	(0.07)	(0.84)	(0.10)	(0.06)	(0.92)	(0.12)
Income^2	.	0.07	.	.	0.10***	.	.	0.04	.
	.	(0.06)	.	.	(0.04)	.	.	(0.04)	.
2 or more vehicles	-0.90***	-0.92***	-0.8**	-0.81**	-0.82***	-0.82**	-0.90**	-0.91**	-0.82**
	(0.34)	(0.33)	(0.34)	(0.34)	(0.33)	(0.34)	(0.40)	(0.40)	(0.38)
0 vehicles	0.21	0.21	0.20	-0.08	-0.11	0.10	0.04	0.06	0.08
	(0.20)	(0.20)	(0.20)	(0.17)	(0.16)	(0.18)	(0.21)	(0.21)	(0.21)
Moved 2000 or later	0.20	0.09	0.09	-0.01	-0.02	0.04	0.77**	0.64*	0.71**
	(0.75)	(0.74)	(0.73)	(0.20)	(0.21)	(0.21)	(0.34)	(0.35)	(0.32)
Moved 1990-1999	-0.84*	-0.84*	-0.93**	0.04	0.01	0.02	-1.48***	-1.28***	-1.44***
	(0.46)	(0.46)	(0.46)	(0.06)	(0.06)	(0.06)	(0.44)	(0.47)	(0.43)
Moved 1979 or earlier	0.88	0.95	1.24	1.00	0.76	1.01	0.52	0.63	1.99
	(0.85)	(0.86)	(0.87)	(0.89)	(0.86)	(0.98)	(4.77)	(4.80)	(3.28)
Black households	0.16	0.19	3.03	-1.11**	-1.14**	-0.90*	-0.06	-0.07	0.22
	(0.60)	(0.59)	(7.55)	(0.43)	(0.44)	(0.49)	(0.39)	(0.39)	(0.64)
Asian households	0.78	0.81	7.06	-0.48	-0.52	-0.81*	0.42	0.43	0.69
	(0.62)	(0.61)	(6.75)	(0.42)	(0.43)	(0.48)	(0.36)	(0.36)	(0.50)
White alone households	1.01	1.00	3.75	-0.49	-0.56	-0.87*	0.70*	0.70*	0.26
	(0.66)	(0.65)	(6.17)	(0.42)	(0.43)	(0.45)	(0.40)	(0.39)	(0.49)

Hispanic households	0.59	0.70	1.28	-1.08***	-1.05***	-1.27***	0.36	0.38	0.82
	(0.58)	(0.57)	(7.04)	(0.39)	(0.40)	(0.42)	(0.37)	(0.37)	(0.52)
BlackXincome	.	.	-0.26	.	.	0.00	.	.	0.00
	.	.	(0.67)	.	.	(0.00)	.	.	(0.00)
AsianXincome	.	.	-0.57	.	.	0.00	.	.	0.00
	.	.	(0.60)	.	.	(0.00)	.	.	(0.00)
WhiteXincome	.	.	-0.26	.	.	0.00	.	.	0.00
	.	.	(0.54)	.	.	(0.00)	.	.	(0.00)
HispanicXincome	.	.	-0.07	.	.	0.00	.	.	0.00
	.	.	(0.62)	.	.	(0.00)	.	.	(0.00)
Education: BA or more	0.54***	0.50***	0.57***	0.56***	0.55***	0.47***	0.35**	0.31*	0.13
	(0.17)	(0.18)	(0.17)	(0.13)	(0.13)	(0.14)	(0.15)	(0.15)	(0.14)
Food stamps	-0.01	-0.01	0.00	0.00	0.00	0.01	0.00	-0.01	0.00
	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)	(0.03)	(0.03)	(0.3)
Households with children under 18	0.07	-0.04	0.02	-0.14	-0.17	0.06	-0.23	-0.28	-0.19
	(0.21)	(0.22)	(0.22)	(0.20)	(0.20)	(0.22)	(0.23)	(0.24)	(0.22)
Work status in past year: no work	-0.06	-0.09	-0.01	-0.39*	-0.30	-0.24	-0.06	-0.08	0.06
	(0.26)	(0.27)	(0.27)	(0.23)	(0.22)	(0.25)	(0.24)	(0.25)	(0.23)
Age 44 and under	-1.33	-1.32	-1.18	-0.67	-0.93*	-0.50	-0.56	-0.61	-0.55
	(1.17)	(1.10)	(1.05)	(0.53)	(0.55)	(0.51)	(1.24)	(1.22)	(1.17)
Age 45 to 64	-0.19	-0.21	0.04	-0.84	-1.04*	-0.34	-0.15	-0.21	-0.15
	(1.08)	(0.99)	(0.95)	(0.58)	(0.59)	(0.56)	(1.16)	(1.14)	(1.09)
Age 65 and older	-1.03	-1.22	-1.35	-0.53	-0.54	-0.35	0.00	-0.17	-0.58
	(1.30)	(1.23)	(1.20)	(0.50)	(0.52)	(0.49)	(1.36)	(1.34)	(1.29)
Constant	12.34***	20.13***	8.67	12.45***	25.76***	12.61***	12.20***	16.34***	12.27***
	(0.94)	(6.68)	(6.12)	(0.85)	(4.88)	(0.86)	(0.80)	(5.04)	(0.77)
N	184	184	184	184	184	184	184	184	184
R^2	0.75	0.76	0.77	0.77	0.78	0.79	0.71	0.71	0.74

Note: ***, **, and * indicate statistical significance at the 1%, 5%, and 10% levels, respectively.
For the c set of regressions, which are estimated for renter-occupied only, I drop the variable "owner-occupied without mortgage).

Appendix A: Initial Motivation, Data and Research Question Evolution

This thesis was born out of research I did in the spring of my junior year for one of my professors, Dr. Steven Teles of the JHU Political Science Department. Steve was consulting for a San Francisco-based non-profit that wanted to know more about issues of land use in the San Francisco Bay Area. He asked me to conduct some of the initial background research on this topic, focusing primarily on national and Bay Area organizations and policies around land use, land restrictions, open space, and building and growth. Steve, as a political scientist, had me focus in particular on various political coalitions that formed around land use and open space protections. He also had me examine the rent-seeking nature of land use restrictions and the implications of economic rents for building and zoning. While this research, on its own, was too political science-oriented to be used for an economics thesis, researching the topic heightened my interest in land use and housing issues in the San Francisco Bay Area; as a native of San Francisco, I was already quite familiar with the pressing housing issues facing the city. I wanted to learn more about what was driving housing prices and (a lack of) housing supply, and explore policy implications based on my findings.

I initially planned to study the impact of zoning regulations in San Francisco on housing prices. I found, however, that there was not any good data easily available for that area of research. I did find two different datasets on zoning, both of which I attempted to use, until I spoke with advisors and realized constraints were too great to allow me to pursue this line of research further.

One of the datasets I accessed was from the Wharton School at University of Pennsylvania. This dataset contained information on the “restrictiveness” of land use regulation and zoning policies on a national level. Large-scale zoning data tends to be difficult to access, since zoning is highly irregular across municipalities. The Wharton Residential Land Use Regulation Index (WRLURI) is based on research by Gyourko et al. that tries to understand why housing supply appears to be inelastic, especially in larger, coastal metropolitan areas. The researchers conducted a nationwide survey of local land use control environments, and then used the data from the survey to create a measure of stringency, or restrictiveness, of the local regulatory environment in each community (the "index" or WRLURI). The survey asked fifteen specific questions, and was completed by 2,649 individuals. The data showed that metropolitan areas tended to be more highly regulated than rural areas. The data also showed that community wealth is strongly and positively correlated with the degree of local land use restriction – the higher the median family income, median house value, or share of adults with college degrees, the greater the community's WRLURI value. Based on the index and survey data, San Francisco had a high WRLURI (on average, one standard deviation above the national mean), meaning that San Francisco has fairly restrictive land use policies.[36]

I decided against using the WRLURI in my research for several reasons. One reason was that the data was relatively old; the survey was conducted in 2004, and the researchers used Census data from 2000. Data that is over 10 years old may not be an accurate indicator of current zoning policies, and demographics since 2000, which is the Census year the researchers drew demographic data from, have likely changed. Another consideration was that I was interested in examining zoning, housing, and demographic changes on the neighborhood level. WRLURI data was only available on a city-wide level, which would have forced me to expand my analysis to multiple cities, since there was no available data on the restrictiveness of certain neighborhoods or census tracts within just the city of San Francisco.

[36] Gyourko, Saiz, and Summers, "A New Measure of the Local Regulatory Environment for Housing Markets: The Wharton Residential Land Use Regulatory Index."

One of the other datasets on zoning that I was initially considering incorporating into my analysis was from the City and County of San Francisco (CCSF). I contacted a data administrator who works for the city, and he provided me with access to data on zoning in San Francisco. The dataset provided zoning information on the parcel-level. Parcels are essentially plots of land, and are even smaller geographic areas than census tracts; a census tract may contain up to hundreds of parcels. The dataset from CCSF listed the zoning code of each parcel within San Francisco; the city contains 215,388 parcels. However, data from the Census on housing and household and demographic characteristics is on the census-tract level, so there was no straightforward way to aggregate these two datasets; for example, one option would have been to assign the mean type of zoning code of all the parcels within one census tract to that census tract, but that would have lost the variation of zoning within a tract.

I ultimately decided not to pursue a research question relating to zoning at all because of the limitations I faced. Conducting the kind of econometric modeling that such a project requires would have taken me months, in terms of gathering data, developing econometric models, and coding the models, and was beyond my undergraduate knowledge. After speaking to my advisor and other members of the economics department, I determined that the constraints I faced prevented me from being able to pursue this. Instead, I chose to simplify my research question and focus primarily on the impact of income and demographic characteristics on housing prices. In doing so, I could still address the issues I was interested in – namely, housing in the Bay Area and the implications of its rising cost – by examining what demographic variables affect housing prices.

Appendix B: Additional Data

Table B1: Detailed summary statistics of median household income for all occupied housing units

Obs	Sum of Wgt.	Mean	Std. Dev.	Variance	Skewness	Kurtosis
184	184	83782.72	32721.63	1.07E+09	0.08	2.68

Table B2: Detailed summary statistics of median household income for owner-occupied housing units

Obs	Sum of Wgt.	Mean	Std. Dev.	Variance	Skewness	Kurtosis
184	184	119731.7	39757.9	1.58E+09	0.37	2.84

Table B3: Detailed summary statistics of median household income for renter-occupied housing units

Obs	Sum of Wgt.	Mean	Std. Dev.	Variance	Skewness	Kurtosis
184	184	67456.17	31135.55	9.69E+08	0.42	2.76

Table B4: Detailed summary statistics of median housing unit price per census tract

	Percentiles	Smallest
1%	314300	287000
5%	446600	314300
10%	510400	373400
25%	630500	384000
50%	796050	
		Largest
75%	946500	1000000
90%	1000000	1000000
95%	1000000	1000000

Obs	Sum of Wgt.	Mean	Std. Dev.	Variance	Skewness
184	184	773313.6	185297.2	3.43E+10	-0.42

Appendix C: Residential Sorting Model

Although I use hedonic price regressions in my study to determine the effect of income inequality on house prices, a more robust methodology would be to use an equilibrium model of residential sorting. I did not use this approach in my study because I was constrained by both time and resources (such as data availability, and computing and coding capabilities). In this appendix, I present an equilibrium model of residential sorting that provides a framework for analyzing how income inequality and residential segregation impact housing prices. This kind of framework is useful in the policy implications it provides; for instance, understanding how households make a housing choice can help policymakers develop policies that take these residential sorting choices into account.

Bayer et al. estimate an equilibrium sorting model; their estimates provide a characterization of preferences for housing and neighborhood attributes and show how demand for these attributes varies with a household's income, race, education, and family structure.[37] The authors also expand on their preference estimates to estimate the impact of a simulated increase in income inequality on the housing market equilibrium. Their results suggest that increased spending in the top quartile households lead to significantly higher house prices, especially for the most desirable houses and neighborhoods in metropolitan areas. As a result, the consumption of housing and neighborhood attributes by households in the rest of the income distribution is adversely affected by the increased spending power of households in the top quartile.

Following Bayer et al., I present an equilibrium model of a self-contained urban housing market, in which households sort themselves among the set of available housing types and locations. There are two key elements to the model – a household residential location decision

[37] Bayer, McMillan, and Rueben, "An Equilibrium Model of Sorting in an Urban Housing Market."

problem, and a market-clearing condition. By examining how location decisions vary, on average, with household characteristics (income, education, race, etc.), one can understand how preferences for housing and neighborhood characteristics vary with these socio-demographic characteristics.

The residential location decision of each household can be modeled as a discrete choice of a single housing unit from a set of house types available in the market. A household chooses its residence to maximize its indirect utility function. Let X_h represent the observable characteristics of housing choice h, including house characteristics (e.g., size, age, type), tenure status (rented vs. owned), and neighborhood characteristics (e.g., school, crime, land use). Let $Z_h$represent the average socio-demographic characteristics of the corresponding neighborhood, p_h denote the price of housing choice h, and d_h^i denote the distance from residence h to the primary work location of the household i. This model is represented by the following equation:

$$\underset{(h)}{Max}\, V_h^i = \alpha_X^i X_h + \alpha_X^i Z_h - \alpha_P^i p_h - \alpha_d^i d_h^i + \xi_h + \varepsilon_h^i$$

with the error term of the indirect utility split into two components. ε_h^i is an individual-household specific error term, while ξ_h captures the unobserved quality of each housing choice. Each household's valuation of housing/neighborhood choice characteristics varies with that household's socio-demographic characteristics z^i (e.g., education, income, race, employment status, household composition, etc.). Each parameter associated with housing and neighborhood characteristics and price, α_j^i, varies with a household's own characteristics:

$$\alpha_j^i = \alpha_{0j} + \sum_{r=1}^{R} \alpha_{rj}\, z_r^i$$

The two equations above allow for a horizontal model of residential sorting, in which households have preferences over each choice characteristic, in contrast to a vertical model. In this model, households have different preferences over different locations. The model shows that a household maximizes its lifetime utility over its choice set by choosing where to live.

In order to smooth the individual residential location decision and thus establish the sorting equilibrium, I next characterize the housing market by a set of housing types that is a subset of the full set of available houses. The supply of housing of a type h is given by S_h. Household i chooses housing type h if the utility it receives from that choice is greater than the utility it received from all other possible housing choices, k, i.e., when

$$V_h^i > V_k^i$$

The probability that a household chooses any particular housing choice depends on the characteristics of the full set of possible housing types. The probability $P^i{}_h$ that household i chooses housing type h is a function of the housing and neighborhood characteristics and the household's own characteristics.

In order for the housing market to clear, the demand for houses of type h must equal the supply of houses, such that:

$$D_h = S_h \Rightarrow \sum_i P_h^i = S_h$$

and prices are assumed to adjust to clear the market. Excess demand for a housing type causes price to be bid up, and excess supply leads to a fall in price.

These models allow for an estimation of how households make their housing choice, and thus can be useful when considering policies that attempt to address residential sorting. These models can also estimate house price in a more precise way than hedonic price regressions, since the characteristics in the model are more detailed and exact. Furthermore, the model can be a tool to analyze how economic or policy changes can affect residential sorting and housing prices. For example, using a more advanced version of the model described above, Bayer et al. simulate a

10% increase in income of the top quartile of households. The authors find that the increased income of the top quartile impacts the overall housing market by raising the prices of the most desirable houses and neighborhoods, and, as a result, affects the consumption of neighborhood and housing attributes by households throughout the entire income distribution, with negative welfare implications for households who are not in the top quartile of the income distribution. This kind of model and simulation would be useful to consider in the case of San Francisco's present housing market, which has tightened and gotten increasingly expensive, while, at the same time, the city's income inequality has grown.

References

"Affordable Housing." *U.S. Department of Housing and Urban Development*. Accessed April 5, 2016. http://portal.hud.gov/hudportal/HUD?src=/program_offices/comm_planning/affordablehousing/.

"America's Most Expensive City Just Got Even More Expensive." *The Huffington Post*. Accessed December 10, 2015. http://www.huffingtonpost.com/2015/02/03/san-francisco-rent-2015-most-expensive-city_n_6609396.html.

Avent, Ryan. "One Path to Better Jobs: More Density in Cities." *The New York Times*, September 3, 2011, sec. Opinion / Sunday Review. http://www.nytimes.com/2011/09/04/opinion/sunday/one-path-to-better-jobs-more-density-in-cities.html.

———. *The Gated City*, 2011.

Badger, Emily. "Voters in One of America's Most Expensive Cities Just Came up with Another Way to Block New Housing." *The Washington Post*, June 4, 2014. https://www.washingtonpost.com/news/wonk/wp/2014/06/04/voters-in-one-of-americas-most-expensive-cities-just-came-up-with-another-way-to-block-new-housing/.

Bayer, Patrick, Robert McMillan, and Kim Rueben. "An Equilibrium Model of Sorting in an Urban Housing Market." Working Paper. National Bureau of Economic Research, November 2004. http://www.nber.org/papers/w10865.

Berube, Alan, and Natalie Holmes. "Some Cities Are Still More Unequal than Others—an Update." Washington, D.C.: Brookings, March 17, 2015. http://www.brookings.edu/research/reports2/2015/03/city-inequality-berube-holmes.

Coulson, Edward. "Chapter 2: A Brief Survey and Interpretation of Hedonic Parameters." In *Hedonic Methods and Housing Markets*. Department of Economics, Penn State University, 2008. http://grizzly.la.psu.edu/~ecoulson/hedonicmonograph/monog.htm.

Cutler, Kim-Mai. "How Burrowing Owls Lead To Vomiting Anarchists (Or SF's Housing Crisis Explained)." *TechCrunch*, April 14, 2014. http://social.techcrunch.com/2014/04/14/sf-housing/.

Dewan, Shaila. "In Many Cities, Rent Is Rising Out of Reach of Middle Class." *The New York Times*, April 14, 2014. http://www.nytimes.com/2014/04/15/business/more-renters-find-30-affordability-ratio-unattainable.html.

Diamond, Rebecca. "Housing Supply Elasticity and Rent Extraction by State and Local Governments," July 24, 2015. http://web.stanford.edu/~diamondr/research.html.

———. "The Determinants and Welfare Implications of US Workers' Diverging Location Choices by Skill: 1980-2000," August 27, 2015. http://web.stanford.edu/~diamondr/research.html.

Diamond, Rebecca, and Tim McQuade. "Who Wants Affordable Housing in Their Backyard? An Equilibrium Analysis of Low Income Property Development," July 2015. http://web.stanford.edu/~diamondr/research.html.

Fischel, William A. "An Economic History of Zoning and a Cure for Its Exclusionary Effects." Department of Economics, Dartmouth College, December 18, 2001. https://www.dartmouth.edu/~wfischel/Papers/02-03.pdf.

Glaeser, Edward L., and Joseph Gyourko. "The Impact of Zoning on Housing Affordability." *NBER*, NBER Working Paper Series, n.d. http://www.nber.org/papers/w8835.

———. "Zoning's Steep Price." *Regulation*, Fall 2002.

Glaeser, Edward L., Joseph Gyourko, and Raven E. Saks. "Urban Growth and Housing Supply." *NBER*, NBER Working Paper Series, January 2005. http://www.nber.org/papers/w11097.pdf.

———. "Why Have Housing Prices Gone Up?" *NBER*, NBER Working Paper Series, February 2005. http://www.nber.org/papers/w11129.pdf.

———. "Why Is Manhattan So Expensive? Regulation and the Rise in Housing Prices." *The Journal of Law and Economics* 48 (October 2005): 331–69.

Groves, Jeremy R., and Eric Helland. "Zoning and the Distribution of Location Rents: An Empirical Analysis of Harris County, Texas." *Land Economics* 78, no. 1 (2002): 28–44. doi:10.2307/3146921.

"Growing Pains." *The Economist*, December 7, 2013. http://www.economist.com/news/united-states/21591187-californias-new-technological-heartland-struggling-its-success-growing-pains.

Gyourko, Joseph, Christopher Mayer, and Todd Sinai. "Superstar Cities." *American Economic Journal: Economic Policy* 5, no. 4 (November 2013): 167–99. doi:http://www.aeaweb.org/aej-policy/.

Gyourko, Joseph, Albert Saiz, and Anita A. Summers. "A New Measure of the Local Regulatory Environment for Housing Markets: The Wharton Residential Land Use Regulatory Index." Forthcoming in Urban Studies. The Wharton School, University of Pennsylvania, March 29, 2007. http://real.wharton.upenn.edu/~gyourko/WRLURI/The%20Wharton%20Zoning%20Regulation%20Index-July%202,%202007.pdf.

Hamilton, Bruce W. "Zoning and Property Taxation in a System of Local Governments." *Urban Studies* 12 (1975): 205–11.

Khan, Naureen. "After Failed Propositions, San Francisco Housing Crisis Still Festering." *Al-Jazeera America*, November 4, 2015, sec. U.S. http://america.aljazeera.com/articles/2015/11/4/san-francisco-housing-crisis-vote.html.

Kuminoff, Nicolai V., V. Kerry Smith, and Christopher Timmins. "The New Economics of Equilibrium Sorting and Policy Evaluation Using Housing Markets." *Journal of Economic Literature* 51, no. 4 (2013): 1007–62.

Matlack, Janna L., and Jacob L. Vigdor. "Do Rising Tides Lift All Prices? Income Inequality and Housing Affordability." *Journal of Housing Economics* 17, no. 3 (September 2008): 212–24. doi:10.1016/j.jhe.2008.06.004.

Moretti, Enrico. "How to Make SF Housing More Affordable." *San Francisco Chronicle*, June 9, 2013, sec. Opinion. http://www.sfchronicle.com/opinion/openforum/article/How-to-make-SF-housing-more-affordable-4590271.php.

Murphy, Alvin. "A Dynamic Model of Housing Supply." *Arizona State University*, July 13, 2015. http://www.public.asu.edu/~amurph10/supply.pdf.

Quigley, John M., and Larry A. Rosenthal. "The Effects of Land Use Regulation on the Price of Housing: What Do We Know? What Can We Learn?" *Cityscape: A Journal of Policy Develompent and Research* 8, no. 1 (2005): 69–137.

Schleicher, David. "City Unplanning." *The Yale Law Journal* 122, no. 7 (May 2013): 1670–2150.

"SF Housing Data Hub | Housing Overview." Accessed April 5, 2016. http://housing.datasf.org/overview/.

Teles, Steven M. "The Scourge of Upward Redistribution." *National Affairs*, Fall 2015.

"The Spectre Haunting San Francisco." *The Economist*, April 16, 2014. http://www.economist.com/blogs/freeexchange/2014/04/housing-markets.

Yglesias, Matthew. "San Francisco Is Great—They Should Make More of It." *Slate*, April 11, 2013. http://www.slate.com/blogs/moneybox/2013/04/11/san_francisco_zoning_needs_more_density_and_tall_buildings.html.