How Remote Work Will Reshape Our Cities: Exploring the Heterogeneity in Pandemic-Era Housing Price Changes

Ethan Nourbash
Advised by Professor Joel Mokyr
Northwestern University
May 1st, 2022

ABSTRACT: I analyze housing price changes throughout the pandemic at the zip-code level as a function of individual and regional characteristics of the zip-code. To this end, I use Zillow time-series housing data from January 2020 to January 2022, U.S. Census American Community Survey 2015-2019 5-year data release, and Google Maps API geographic data. I find that the share of work that can be done remotely adversely affects the change in housing demand over the two-year period for highly populated counties. This effect is less pronounced or non-existent in less populated counties. A model designed to analyze the changes in optimal worker housing location when offered remote work concludes this is a result of bid-rent functions differentiated by population size. Furthermore, there are statistically significant differences in remote work's effect on housing demand by worker characteristics, such as education, age, marital status, and income. The model identifies two potential causal channels for this effect: heterogeneous amenity preferences or relocation costs.

Keywords: Housing, Remote Work, Urban Economics
JEL No. R12, R23, R31
ACKNOWLEDGMENTS: I would like to thank Professor Mokyr for the effort he put into advising my thesis. His personal interest in the topic led to lively weekly meetings and inspired me to write the best paper I could. For exposing me to economic research practices through research assistantship, I thank Ameet Morjaria, Vivek Bhattacharya, and Deborah Weiss. For fostering a support group of undergraduate economists, I'm grateful to James Hornsten. Finally, I'm thankful for my mom's excellent research assistance, by which I mean her sending me Wall Street Journal articles about remote work.

1. INTRODUCTION

After the threat of infection subsides, there will surely be aspects of pandemic culture preserved in our post-pandemic world. We might have more stringent vaccination guidelines, contact tracing apps, or simply less crowded elevators. Aside from healthcare related changes, many companies have found success with work from home (WFH) after COVID-19 forced them to adopt the concept.

Not all companies experienced the same level of success with working from home. The nature of some businesses prevents remote work from even being an option. For example, a waiter cannot serve a customer if they are 10 miles apart, and neither can a janitor clean an office if he is still at home. As technology improves, the set of roles that require physical presence will likely diminish, further contributing to the adoption of remote work practices.

For those fortunate enough to work remotely during the pandemic, a home proximate to the office was no longer a necessity. Some took advantage of this freedom to temporarily live in locations with more open space. A local example of this effect is that Airbnb revenues in March 2020 tripled in Waukegan and shrank 11% in Chicago (Ren 2020). As people start to go back to the office a few days a week with a hybrid model, it will be interesting to see if location choice lies on a continuum. To answer this question, I hope to confirm the findings that workers who expect WFH post-pandemic are more likely to permanently reside further from the city center in the first part of this paper.

There is also the common sentiment that life-stage differences influence willingness to live close to the city center. As stated in Glaeser and Cutler's 2021 book *Survival of the City*:

> *"Middle-aged workers with life partners, children, and many community relationships are less likely to want to hang out at work until midnight. Consequently, the switch to remote work will have more appeal for the dull and middle-aged…than for the young and hip."*

If this hypothesis holds, then WFH will have a larger effect on middle-aged workers' housing choices than that of young workers. For example, a 2-day WFH schedule might not dissuade a young worker from living in SoHo but may persuade a middle-aged mother to relocate to the Hamptons (which experienced the largest positive rent change in New York during the pandemic according to Gupta et al. 2021). In the second part of my paper, I will attempt to discern which observable worker characteristics—age, education, marital status—influence willingness to relocate. This will be an original contribution to the growing literature on the pandemic's effect on the housing market.

Finally, the characteristics of a region likely contribute to its citizens' willingness to relocate. City population size has been found to be positively correlated with the likelihood of having a rent burden[1] (Oh 1995). If residents of large cities would experience a larger decrease in rent by moving out to a suburb than residents of small cities, they should be more willing to relocate with a rise in telecommuting. My final econometric model specification will differentiate

[1] Rent burden is defined as spending more than 30% of after-tax incomes on housing.

WFH's effect on housing demand first by population size and then by region as defined by the U.S. Census.

2. Literature Review

2.1. Brief History of the Urban Labor Force

Before the rise of computers, non-college-graduate workers were vital to white-collar work. They would operate in administrative support, clerical work, and sales. These positions are commonly referred to as middle-skill occupations. In the 1970s, the increase in computer technology allowed firms to start automating these jobs. In 1980, 43% of non-college-graduate workers worked in middle-skill jobs. This figure has since fallen to 29%. The rise of low-skill work employment explains 88% of this fall. The change in roles that less-educated workers filled had catastrophic effects on their livelihoods. It has contributed significantly to the decline of non-college workers' wages, urban wage premiums, and specialization (Autor 2019).

Since 2000, there has there been a migration to large cities driven by the young college-educated population. This migration was influenced by a change in consumption preferences amongst young college graduates. They now prefer a high density of leisure and hospitality amenities. This change in preferences can be explained by the income growth they experienced and the delay in family formation (Couture and Handbury 2020).

As a reaction to this change in the amenity preferences of wealthy city dwellers, leisure and hospitality has grown as a source of employment for uneducated workers in urban areas. From 1945 to 2020 employment in leisure and hospitality has grown from 4% to 10% of the total labor force[2]. At the beginning of 2020, the total employment in factories was equivalent to that of restaurants (Glaeser and Cutler 2021). The increase in leisure and hospitality employment, however, has not compensated for the loss of middle-skill employment given that the mean education level of city dwellers has been rising over the past few decades (Brinkman 2015).

The combined change in consumption preferences and rise in high-skill wages has led to a new industry coined "wealth work" by Autor and Salomons. Wealth work positions are those where the worker performs an in-person service for affluent customers. Examples include pet groomer, sommelier, personal trainer, and dance therapist. In 2019, wealth work accounted for 52% of new work. It is also a field not dominated by college graduates: 54% of wealth workers have not graduated college (Autor and Salomons 2019).

Before the onset of COVID-19, the urban labor force had reached a new equilibrium. Highly educated workers were compensated with large incomes by firms greatly in need of employees with skills that complement—rather than poorly substitute—advanced technology.

[2] This statistic was stated in Glaeser and Cutler 2021 using total employment. The total number of employees in leisure and hospitality went from 2.2 million to 16.9 million. To avoid misinterpretation, I accounted for the size of the labor force in the United States in both years: 1945 labor force data was taken from NBER 2022 and 2020 labor force data was taken from World Bank 2022.

These workers would then spend this income on a variety of in-person services operated by less-educated workers.

2.2. Immediate Impact of the Pandemic

The pre-pandemic urban labor market equilibrium was disturbed when a deadly virus started spreading through human contact. Educated workers worked from home instead of with colleagues in an urban office. Additionally, they could no longer spend their disposable income on in-person services such as bars and training gyms. This caused the desirability of cities to drop for educated workers and the availability of jobs to drop for uneducated workers.

The ability of a firm to continue operating early in the pandemic was contingent on how feasible remote work was. Manufacturing requires the labor to interact with large immovable machinery while financial services can be completed with a laptop and internet access. Consequently, the implementation of remote work had high variation across industries. As stated previously, the nature of an employee's job is highly contingent on her education level. In May 2020, only 5% of high school dropouts were telecommuting versus 66% of those with advanced degrees (Glaeser and Cutler 2021).

In the fourth quartile of industries grouped by mean worker education level, 64% of firms had some workers performing remotely. This value is only 36% for the first quartile. Restricting to firms that used WFH, there was variation in perceived productivity by education. In one regression, a 10% increase in the share of workers with a college degree was associated with a .33% rise in perceived WFH productivity (Bartik, Cullen, Glaeser, Luca, and Stanton 2020). Another survey confirmed the large variance in WFH productivity by education level and further claimed that the mean WFH productivity was 60%-70% of pre-pandemic levels (Morikawa 2021).

Unfortunately, when an industry is incapable of working remotely during COVID-19, it is less capable of staying profitable. Over 40% of leisure and hospitality workers lost their jobs compared to 7% of finance and insurance sector employees throughout the pandemic (Glaeser and Cutler 2021).

The pandemic caused workers to relocate, not necessarily because of a new set of preferences, but the availability of the amenities they had hitherto enjoyed. What is the value of a downtown Manhattan apartment when both the office and the bars are closed? The result was a dramatic change in rent for areas that met the following criteria: high presence of jobs that can be performed remotely, strict pandemic lock-down measures, and low housing supply elasticity. Rent, rather than home value, was analyzed for its ability to measure short-term housing demand. In aggregate, researchers found that remote workers left urban areas for less expensive suburbs still within their metropolitan area now that commuting is less frequent (Gupta et al. 2021). Ramani and Bloom 2021 define this rise in suburb prices paired with a decline in city center prices as "the donut effect".

2.3. Post-Pandemic Location Choices and Labor Market

After experiencing such dramatic change in housing and consumption choices because of work from home, one might ask: “How much of this change is permanent?” Many highly educated companies are having success with WFH and some workers appreciate the 0-minute commute. Surely not every urban employee appreciates access to fine dining over a larger apartment. After all we have learned from this large-scale work from home experiment, there are surely some aspects of work that have changed for good.

Since homes are investments with future revenue streams, changes in housing prices can be thought of as a measure of long-term effects on housing demand. While WFH influenced rents dramatically, there has been a less substantial change in prices (Gupta et al. 2021). The donut effect is not found outside of the largest American cities when using price rather than rent (Ramani and Bloom 2021). This suggests a long-term effect on housing demand exists but is limited.

Barrero, Bloom, and Davis (2021) show that work from home will have a large impact on the post-pandemic workforce commutes. Many workers have responded positively to the idea of having a hybrid work schedule in the future. According to their survey, nearly 40% of all respondents answered they want to work half of the time remotely. While the desire to have a hybrid work schedule transcends education and income level, this desire is only fulfilled if the employee’s desires agree with their employer’s post-pandemic plans. Employee desires and employer plans are the most similar for those earning $225,000 or more. For those earning $50,000 or less, the average remote days per week desired is 2.2 while the average employer’s plan is only .8. If employees are spending more time at home, this will influence consumption levels near their office. Barrero et al. go on to predict that work from home will adversely affect expenditures on meals, entertainment, personal services, and shopping in cities. This is because the inward commuters of a city have a significant impact on consumption spending in the area. Manhattan, for example, is predicted by Barrero et al. to have a 13% drop in consumer spending.

Given that amenity preferences—such as the number of restaurants nearby or the proximity to a park —vary across skill groups (Diamond 2016), the implications of a significant change in urban demographics include a shift in which businesses and industries can stay profitable in the city. Additionally, a new set of community preferences requires revision of government expenditures to avoid spending inefficiency. For example, if cities become more educated, city government expenditures on crime prevention might be less important than investing in library renovations.

3. THEORY

3.1. Set-Up

My model is inspired by Diamond 2016, with several key changes to provide insight into the impact of remote work on worker location choice: (1) The place of residence is determined separately from place of work; (2) a second Beckerian budget constraint is added (Becker 1965);

(3) the local good price is endogenous; (4) the amenity utility function has only two factors, commute time and leisure, that are both endogenous; and (5) the proportion of work done from home is introduced as a worker choice with an exogenous upper bound (a mechanic may desire 100% WFH, but his upper bound is 0%).

Assume that worker i has a job in the city center where he is paid a wage W_i. He chooses to live in region p that offers him the most desirable bundle of local good prices and amenities. He also is subject to a time budget constraint which forces the sum of time spent working, commuting, and leisure to be at most the amount of time he has in the day. Let this be represented by a Cobb-Douglas utility function of local goods M with price R_p, national goods O with a price P, and local amenities $A_{i,p}$. Let s_i be the worker i's linear amenity utility function. Let T_W be worker i's time spent working, T_c be the time spent commuting, T_l be the time spent on leisure, and T be the total time available. To emphasize remote work's effect on home life, I will assume T_W is fixed and observe the tradeoff between commuting and leisure.

$$U(M, O, p) := M^{\alpha} O^{1-\alpha} e^{s_i(A_{i,p})}$$

$$\text{s.t. } R_p M + PO \leq W_i$$

$$\& \; T_W + T_c + T_l \leq T$$

Let rent, R_p, be a function of distance from city center c_p such that $R_p{}'(c_p) < 0$ and $R_p{}''(c_p) > 0$. In particular, we will assume that the price of local goods takes this functional form:

$$R_p := \beta_{R_0} + \beta_{R_1} \log(c_p + 1)$$

where $\beta_{R_0} > 0$ and $\beta_{R_1} < 0$. This equation says that rent is highest in the city center ($c_p = 0, R_p = \beta_{R_0}$) and all other rent in the region is determined by a decreasing convex function of how long it takes to get to the city. Rent decreasing as a function of commute time, $R_p' < 0$, is inspired by bid rent theory which states that demand for real estate declines as distance from the city center increases. The convexity of the function, $R_p'' > 0$, is inspired by Gupta et al. (2021) who use $\log(X + 1)$ as the functional form of their bid-rent curve.

Let the amenity vector, $A_{i,p}$, contain two variables: worker i's commute time, T_c, and time spent doing leisure activities, T_l. Let the amenity utility function, s_i, have the following properties: $\frac{\partial}{\partial T_c} s_i < 0$ and $\frac{\partial}{\partial T_l} s_i > 0$. To achieve this goal, let s_i have the following form

$$s_i(A_{i,p}) := \beta_{s_0} T_c + \beta_{s_1} T_l$$

where β_{s_0} and β_{s_1} are preference coefficients such that $\beta_{s_0} < 0$ and $\beta_{s_1} > 0$. The interpretation of these signs is that the longer worker i must commute, the less happy he is, and the more time spent relaxing, the happier he is.

For further simplification, assume that leisure can only be spent on household affairs (childcare, reading, cooking) or entertainment venues (theatre, nightclub, restaurant). Let these be represented by T_h and T_e respectively. Moreover, assume worker i has a fixed ratio, $q_i \in \{0,1\}$, of how he chooses to allocate time between these two leisure groups.

$$T_l := T_h + T_e$$

$$T_h := q_i T_l$$

$$T_e = (1 - q_i) T_l$$

Since the effective commute time is partially determined by the amount of work done at home, T_c will respond endogenously to the amount of WFH individual i experiences. I assume that commuting to the city center is also required to take advantage of any entertainment venues. The number of days he commutes to the city for leisure is equal to the proportion of leisure time spent on entertainment venues. In this way, T_l influences T_c in a manner that varies by the leisure preferences of the worker. With these assumptions at play, worker i's effective commute time, T_c, has the following functional form:

$$T_c := (1 - r_i)c_p + (1 - q_i)c_p$$

where $c_p \geq 0$ is the commute time to the city center from place of residence and $r_i \in \{0,1\}$ is the share of work done remotely. The assumption behind this specified commute time function is that worker i cares about the mean commute time he experiences rather than the actual commute time of traveling to the office or the entertainment venue, c_p. For example, let worker i live 40 minutes from the city, he works from home 40% of the time and spends 15% of his evenings at a restaurant in the city. His mean commute time would be 30 minutes. When deciding whether to relocate, he considers his current commute to be 30 minutes rather than 40. I should note that an oversight of this specification of commute time is commute related time expenditures. If a worker is going to show up to the office, he or she will wear dress pants rather than pajamas. Since this aspect of commuting would be difficult to model and does not provide significant further insight, it is excluded. Furthermore, the optimal amount of WFH is always the upper bound in this simple model as no downsides to remote work, such as lack of socializing with colleagues, are factored into this model. This issue is further discussed in the conclusion.

3.2. Evaluation

3.2.1. Marginal Utility Equivalence

The first-order conditions of the household problem yield the following:

$$\frac{U_M}{R_p} = \frac{U_O}{P}$$

$$\frac{\alpha M^{\alpha-1}O^{1-\alpha}e^{s_g(A_{i,p})}}{R_p} = \frac{(1-\alpha)M^{\alpha}O^{-\alpha}e^{s_g(A_{i,p})}}{P}$$

$$\frac{\alpha O}{R_p} = \frac{(1-\alpha)M}{P}$$

$$O = \frac{(1-\alpha)}{\alpha}\frac{R_p}{P}M$$

3.2.2. Wage Budget Constraint

Substituting first-order condition results for O in the wage budget constraint yields

$$R_pM + P\frac{(1-\alpha)}{\alpha}\frac{R_p}{P}M \leq W_i$$

$$\left(1+\frac{(1-\alpha)}{\alpha}\right)R_pM \leq W_i$$

Assuming worker is using his entire wage to solve for M, we have

$$M = \frac{W_i}{\left(1+\frac{(1-\alpha)}{\alpha}\right)R_p}$$

3.2.3. Time Budget Constraint

Substitute T_c's functional form:

$$T_W + (1-r_i)c_p + (1-q_i)c_p + T_l \leq T$$

Assume equivalence in time budget constraint to solve for T_l.

$$T_l = T - T_W - (1-r_i)c_p - (1-q_i)c_p$$

3.2.4. Utility Maximization

After combining the equations, log-transforming, and simplifying (See Appendix 10.2), we get the following results:

$$\begin{aligned}\max_p log(U) \quad &= (1-\alpha)log\left(\frac{(1-\alpha)}{\alpha}\right) - log\left(1+\frac{(1-\alpha)}{\alpha}\right) + (-\alpha)log\left(\beta_{R_0} + \beta_{R_1}\log(c_p+1)\right) \\ &- (1-\alpha)log(P) + log(W_i) + \beta_{s_0}\left((1-r_i)c_p + (1-q_i)c_p\right) \\ &+ \beta_{s_1}\left(T - T_W - (1-r_i)c_p - (1-q_i)c_p\right)\end{aligned}$$

3.3. Moving Cost

Let there be an exogenous moving cost, H_i, that must be paid for worker i to change his housing location. Thus, if there is a change in one of the variables that influence optimal location[3], then the following must hold for worker i to be willing to move:

$$\Delta U_i \geq H_i$$

3.4. Example

Let my example worker, Jessica, be a real estate agent living in Houston. Assume the rent and entertainment functions have the following values for Harris County (See Appendix 10.3):

$$\beta_{R_0} = 2280.3, \beta_{R_1} = -635.1$$

Assume Jessica earns \$5500 monthly income and spends \$400 on national goods:

$$W_{i,t} = 5500, P_t = 400$$

Assume the following reasonable values for amenity utility function coefficients and utility elasticity of consumption:

$$\beta_{s_0} = -.085, \beta_{s_1} = .001, \alpha = .75$$

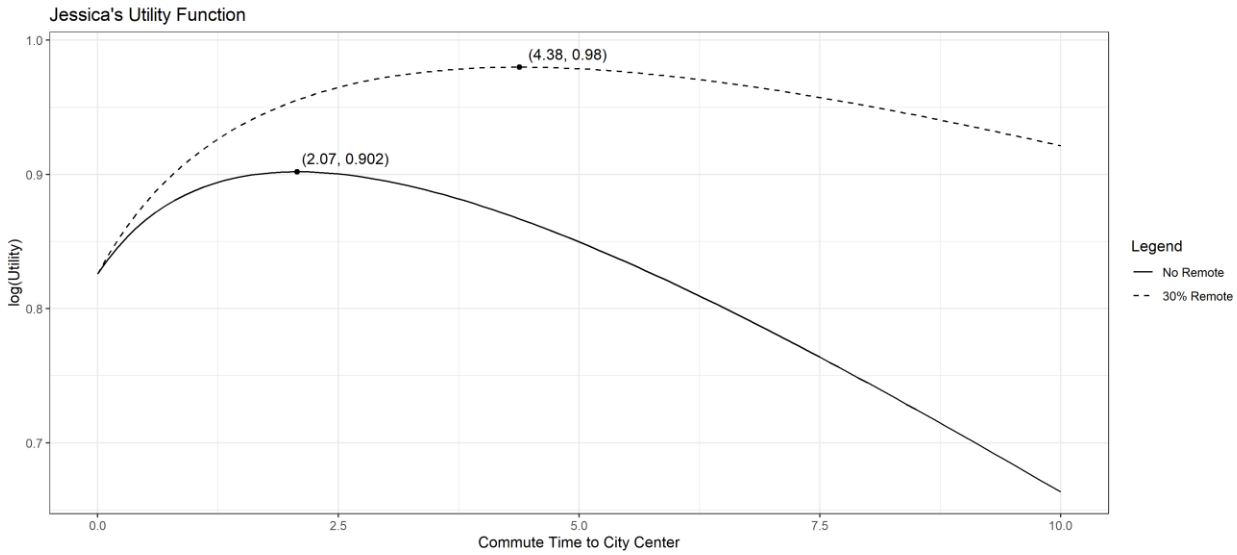
Assume she has 16 hours in the day, 8 of which are dedicated work; additionally, she likes to spend 15% of her nights off in the city.

$$T = 16, T_w = 8, q_i = .85$$

Assume she could complete 30% of her work from home. She experienced no WFH previously, but a company policy changed her level of WFH to its full potential. Assume that this policy change is not covid related. This allows her to keep her current q_i value below 1 and continue to travel to the city during her time off. This caveat is important to isolate the impact of WFH from the rest of covid's effects and get a sense of its long-term impact on location choice.

$$r_{i,0} = 0, r_{i,1} = .3$$

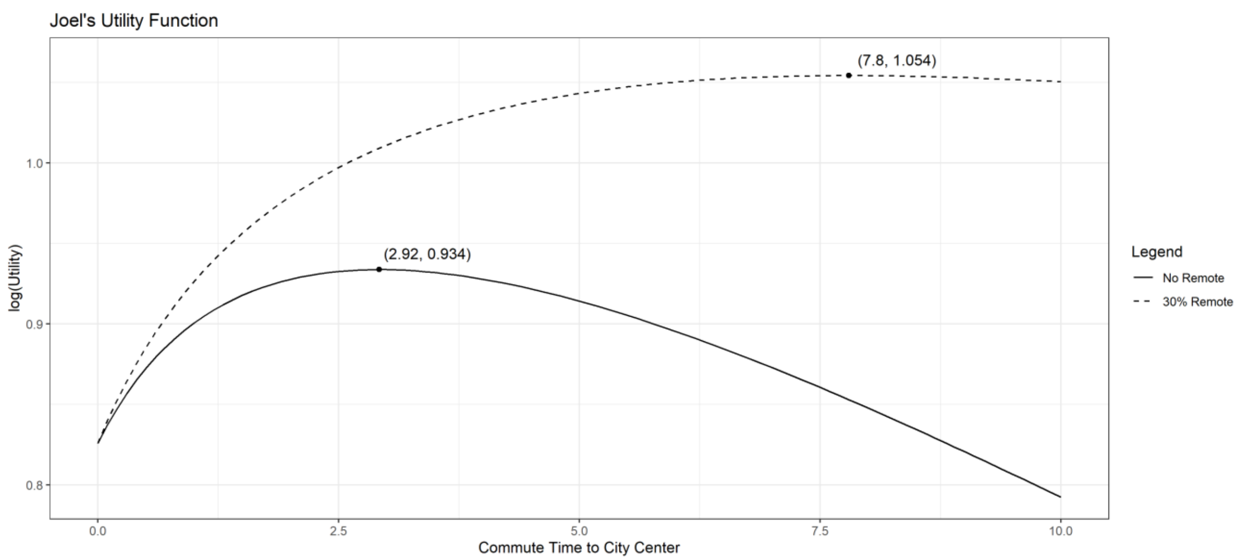
[3] All variables except P and W_i influence the optimal location of the worker.



When the amount of remote work increases, her optimal location choice gets further from the city. This reflects commute time being traded off for more local good consumption when mean work commute time decreases.

3.5. Comparison Example

Let Joel be a colleague of Jessica who enjoys spending all his leisure time at home reading, $q_i = 1$.



As can be seen in the above figure, the two workers have different initial and final optimal locations. Furthermore, Joel experiences a larger change in optimal Utility[4] and change in optimal

[4] ΔUtility = .324 for Joel and .2 for Jessica

distance from city center[5]. This is because Joel's preferred blend of leisure activities requires no trips to the city. He can therefore move further from the city and experience a smaller change in total commute time than if Jessica were to perform the same move. If there were to be a homogenous moving cost set at $H = .25$, we would only expect Joel to move further from the city. However, suppose they have different moving costs: $H_{Joel} = .3$, $H_{Jessica} = .15$. In this case, we would see Jessica move further from the city but not Joel. This model suggests that amenity preferences and moving costs can influence the willingness to move away from a city when the amount of time working remotely increases. If this theory holds along with the assumption that there is some homogeneity in amenity preferences within education, age, or marital-status groups, then the demographics of a region should influence the change in housing demand.

4. HYPOTHESIS TO BE TESTED

Given that amenity preferences cannot be directly observed, this paper will explore who is permanently leaving the cities by discerning which observable individual characteristics are inversely associated with expected long-term housing demand. Note that this method of analysis will not allow me to identify the causal channels at play in the residents' relocation decisions, only identify which types of residents are more likely to relocate.

Individual characteristics will be approximated by collecting data at the zip-code level. These will include zip-level averages for income, marital status, years of education, age, and amount of work that can be done remotely. Amount of work that can be done remotely will then be interacted with the other individual characteristics.

Characteristics of the region, zip-code in this case, will likely influence change in demand and will be included in the model. These include commute time to the city center and total population.

To estimate long-term housing demand changes, I will use the typical price of housing per zip-code. I will focus on the change in housing price from the onset of COVID-19 to the present.

I will explore interactions between worker characteristics and ability to work remotely to see what kinds of people are more willing to relocate given a WFH option. For example, is a 40-year-old mother more likely to leave the city than a 25-year-old single man if both are presented the opportunity to work remotely 3 days a week? These questions will be answered to the best of my ability by observing the sign and magnitude of the interaction variable.

Finally, to determine where we will see remote-capable workers leaving, I will run regressions exploring how regional characteristics and WFH interact. To determine the region sizes (large city, small city, suburb, rural) that are at risk of a WFH-induced exodus, I will estimate the effect of remote work on housing demand by population sizes. I do not expect a linear relationship between population and the effect size of being able to work remotely on housing demand. This is a result of Ramani and Bloom's (2021) finding that the pandemic-induced "donut effect" only existed in large cities when observing price changes. Thus, I will run a series of

[5] Change in optimal commute time is 4.88 minutes for Joel and 2.31 minutes for Jessica

regressions grouped by similar population sizes. To determine which geographic regions are more likely to see an exodus, I will run a series of regressions grouped by U.S. Census defined region.

5. DATA DESCRIPTION

5.1. Census

This paper uses 5-year data (2015-2019) at the zip-code level from the American Community Survey Data to get an estimate of urban community demographics (IPUMS 2021). Relevant variables include education level, median household income, population, commute time, household type (family or non-family), and occupation.

Education level was originally a set of numeric variables where the relevant sample (adults 25 or older) would indicate the highest grade level that they achieved. Each education variable indicated the total number of respondents with that grade being the highest that they achieved. To make this variable comparable across zip-codes, I converted each variable to the proportion that achieved that level by dividing by the total number of respondents. For ease of interpretation, I then took these proportion variables and used them as weights to calculate the average years of education variable *Education*. Household type is a categorical variable with two levels, family or non-family. I calculated the proportion of households that were families and named this variable *Family*. I generated a dummy variable, *Urban*, that is equal to 1 if the zip-code resides in a county with a population larger than 971352[6]. Median household income, *Income*; median age, *Age*; and total population, *Population*, were left unchanged.

Finally, I calculated the share of work that could be done remotely. Census data contains the total number of respondents per Standard Occupational Classification (SOC) major groups. Dingel and Neiman 2020 calculated the share of jobs that can be done at home by SOC major group. I summed the SOC major group totals weighted by the share of jobs that could be done remotely as calculated by Dingel and Neiman. I then divided by the total number of respondents to get the share of work that can be done remotely per zip-code, *Remote*. Other researchers have confirmed that this measure of remote work capabilities does a good job predicting industry-level remote work patterns during COVID-19 (Bartik, Cullen, Glaeser, Luca, and Stanton 2020).

Share of jobs that can be done at home, by occupation's major group (Dingel and Neiman 2020)

Occupation	O*NET-derived baseline
Computer and Mathematical Occupations	1.00
Education, Training, and Library Occupations	0.98
Legal Occupations	0.97
Business and Financial Operations Occupations	0.88
Management Occupations	0.87

[6] This value is 2 standard deviations higher than the mean county population

Arts, Design, Entertainment, Sports, and Media Occupations	0.76
Office and Administrative Support Occupations	0.65
Architecture and Engineering Occupations	0.61
Life, Physical, and Social Science Occupations	0.54
Community and Social Service Occupations	0.37
Sales and Related Occupations	0.28
Personal Care and Service Occupations	0.26
Protective Service Occupations	0.06
Healthcare Practitioners and Technical Occupations	0.05
Transportation and Material Moving Occupations	0.03
Healthcare Support Occupations	0.02
Farming, Fishing, and Forestry Occupations	0.01
Production Occupations	0.01
Installation, Maintenance, and Repair Occupations	0.01
Construction and Extraction Occupations	0.00
Food Preparation and Serving Related Occupations	0.00
Building and Grounds Cleaning and Maintenance Occupations	0.00

5.2. Google Maps

Though Census has commute time data, I decided to generate alternative commute time data using Google Maps Distance Matrix API. The motivation behind exploring alternative commute time variables is that a worker does not care about his neighbors' commute time when considering moving. He cares about his new commute time. Since the Census is reporting the commute times of people already living in each zip-code, this data is more representative of the worker's neighbors' commute times.

Hereafter, "Google commute variable" will refer to the commute variable created using the Google Maps Distance Matrix API. "Census commute variable" will refer to the commute variable created using the ACS 2015-2019.

For the Google commute variable, I calculated the distance between the observation's zip-code and the most populated zip-code in the observation's county. For better accuracy, I used Google's "Best Guess" traffic modeling with a 9 am arrival time in the observation's time-zone. Census asks the respondents within each zip-code to report their commute time to work. Originally, Census commute time data was a set of variables where commute time was grouped in 5-minute time intervals. I took the midpoint of each group's time window, converted each variable to proportion, and calculated the average commute time per zip code.

Similar to *Education*, the Census commute variable is a single numeric variable generated from a set of variables that reported the number of respondents per bin. Originally, Census binned commute time into 5-minute groups and reported the number of respondents that experienced a commute time within that range. I calculated the proportion of respondents in each commute time bin. I multiplied these proportions by the midpoint of each bin and took the sum of these values. This weighted mean became the Census commute variable.

In summary, the Google commute variable is my best approximation of a city emigrant's new commute time if they were to flee the "city" (most populated zip-code in their county) while the Census commute variable approximates the commute time experienced by the typical worker already residing in that zip-code. Surprisingly, there was little correlation between the two variables[7]. County is a reasonable estimate of the radius a city dweller might consider relocating to long-term given that research has found substantial relocation within a given metropolitan area but less so across metropolitan areas (Ramani and Bloom 2021).

As shown in the Appendix 10.4 regression tables, the Google commute variable (Table 2) always has more predictive power than the Census commute variable (Table 3) when estimating the price change variable. When state fixed effects are included, there is no difference in accuracy between using both commute variables (Table 1) and just the Google commute variable (Table 2).

As a result of the higher accuracy, increased clarity of interpretation, and comparability to my theory section, the Google commute variable is preferred over the Census commute variable and will henceforth be simply referred to as the *Commute* variable.

5.3. Zillow

To get an estimate of the change in demand for housing, I used Zillow's housing data at the zip-code level from January 2020 and January 2022 (Zillow 2022). My chosen measure of housing prices is the Zillow Home Value Index (ZHVI). ZHVI is a smoothed, seasonally adjusted measure of the typical home value in each zip-code[8]. I then calculated the percent change in housing price by taking the difference of ZHVI values for January 2022 and January 2020 and dividing by the January 2020 levels. This allows me to have one value that summarizes the change in housing prices over the time period of interest since the ACS data is not time-series. This variable is referred to as *PctChangePrice*.

5.4. Summary Statistics

Income, *Commute*, and *Population* are scaled for ease of interpretation in Section 6's reported regressions. For the constructed *Commute* variable, the median value of 25 minutes is reassuring but a maximum value of 4,381 minutes indicates that the variable is occasionally mis-specified. The high completion rate supports the assumption that there will be no sample bias in the following analyses.

---

[7] The correlation across all observations is .045. The correlation by state has a mean value of .14. The correlation by state and restricted to *Urban* counties (as defined in Section 5.1) has a mean value of .23.

[8] For a complete methodology of the Zillow Home Value Index calculation, see the following website: https://www.zillow.com/research/zhvi-methodology-2019-deep-26226/.

	Completion Rate	Mean	Std. Dev.	p0	p25	p50	p75	p100
PctChangePrice	0.988086	0.262368	0.123254	-0.43784	0.187095	0.249324	0.328935	1.096018
Remote	0.975351	0.350836	0.112027	0	0.282067	0.338019	0.40944	1
Income ($10,000s)	0.940922	6.219697	2.623735	0.2499	4.551275	5.6472	7.19695	25.0001
Commute (10 min)	0.975023	3.01543	7.114163	0	1.676667	2.446667	3.455	438.1067
Family	0.978141	0.672136	0.123215	0	0.614943	0.680851	0.742737	1
Education	0.978732	14.33669	1.126969	6.181818	13.68047	14.23213	14.91053	20.02778
Age	0.974826	42.79162	8.461038	2.4	37.3	42	47.3	91.5
Population (10,000s)	0.981522	1.07675	1.509778	0	0.0942	0.3484	1.5418	12.8294

6. ECONOMETRIC MODEL

I will estimate the determinants of change in housing demand at the zip-code level using the following two equations:

$$PctChangePrice_z = \beta_0 + \beta_1 Remote_z + \beta_2 I_z + \beta_3 R_z \; (+\beta_4 S_z) + \epsilon_z \tag{1}$$

$$\begin{aligned} PctChangePrice_z &= \beta_0 + \beta_1 Remote_z + \beta_2 I_z + \beta_3 R_z \\ &+ \beta_4 Remote_z * I_z \; (+\beta_5 S_z) + \epsilon_z \end{aligned} \tag{2}$$

where z index zip-code; *PctChangePrice* is the percent change in ZHVI (estimated typical housing price) over the two-year period; *Remote* is as defined in Section 5; *I* represents an array of the individual characteristic variables *Income*, *Family*, *Education*, and *Age*; *R* represents an array of the regional characteristic variables *Commute* and *Population*.

To control for unobserved variables, these equations will be estimated with and without state-level fixed effects, *S*. This will allow me to account for unobserved variables that differ across states but are constant over time.

Again, the motivation for using house prices rather than rent is the common assumption that prices are forward-looking and therefore more capable of measuring future housing demand. The reason for using percent change in price rather than raw change is the assumption that the change in demand corresponding to a change in housing price is conditional on the current house price. For example, a substantial decrease in the demand for low-income housing might lead to a $10,000 decline in house listing price whereas a $10,000 change in housing price reflects a marginal decrease in housing demand for high-income housing.

The first equation intends to get a basic understanding of how the relevant covariates influence the change in housing price. Since I hypothesize there will be heterogeneity in post-pandemic housing demand by population size, this equation will be estimated for all zip-codes and restricted to *Urban* zip-codes. Initial findings support this hypothesis (Appendix 10.1). This will be further explored in the third econometric model.

I am interested in what types of people are more likely to take advantage of working from home, if given the option, and move further from their place of work. By interacting remote work with individual characteristics in equation 2, I will get an answer to this aspect of my research question. Again, this equation will be estimated for the entire data set as well as restricted to urban zip-codes.

Finally, I further explore regional heterogeneity in *Remote*'s effect. First, I split the sample into deciles by county-level population. Equation 1 with state fixed effects will be applied to each decile subset. This model specification will not only attempt to reproduce Ramani and Bloom's results that the donut effect only applies to large population cities, but directly relate this exodus to the level of WFH experienced. Second, I run regressions by region as defined by the U.S. Census. This will allow me to analyze the effect of the pandemic by culturally and environmentally similar areas of the U.S.

I predict equation 1 coefficient signs will confirm results found in previous research. Particularly, I expect to see the donut effect in urban areas, which would be realized through a positive sign on *Commute*, as well as negative demand change when working from home is prevalent, attained by a negative sign on *Remote*.

Concerning the heterogeneity in relocation of remote-capable workers (equation 2), I expect the following results: workers with children will leave cities when they can work remotely, *Remote*Family* negative, and older people will leave the city when they can work remotely, *Remote*Age* negative. These expectations are derived from the assumption that, other than proximity to office, city amenities are valued higher by young, single workers.

For the analysis of WFH's effect on relocation by county population size, I expect the magnitude of *Remote*'s coefficient to be negative and increase in magnitude by population size as a result of the relationship between rent burden and city size (Oh 1995).

I expect the South to have the smallest *Remote* coefficient as its cost of living is much lower than other regions, even restricting to urban areas. As a result, it is less likely that a remote-capable worker would need to relocate as a result of a rent burden in the South versus the West.

7. EMPIRICAL RESULTS

7.1. Overview

The four columns for sections 7.2 and 7.3 are specified as follows: all zip-codes with no fixed effects, *Urban* zip-codes subset with no fixed effect, all zip-codes with fixed effects, *Urban* zip-codes subset with fixed effect. Given that there is omitted variable bias concerns with columns 1 and 2 (Appendix 10.5), columns 3 and 4 are my preferred specification and are the subject of subsequent analysis. Regressions without state fixed effects are kept simply for completeness.

In section 7.4, the corresponding appendix table displays the regressions run within each county population decile, where the 1st column is the 1st population decile, etc. In section 7.5, the four columns in the appendix table correspond to sample restricted to Midwest, Northeast, South, and West in that order.

7.2. Equation 1 Results

As can be seen in the table below, equation 1 best explains the price changes in urban areas. The model has less success explaining the changes in price for all U.S. zip-codes. *Remote* has a negative, significant coefficient for all columns. This supports the hypothesis that the number of remote capable workers is inversely associated with percent change in housing price. As expected, there is large heterogeneity in the effect size of *Remote* when controlling for population size. Across the United States, a 10% increase in the number of remote-capable workers predicts a .6% decrease in percent change in housing price (Column 3). Restricting to *Urban* observations, a 10% increase in the number of remote-capable workers predicts a 3% decrease (Column 4). These results are consistent with existing literature (Gupta, Mittal, Peeters, and Nieuwerburgh 2021). *Income* is associated with a decrease in the dependent variable; heterogeneity across population sizes does not appear. *Education* and *Family* consistently predict an increase in percent increase in housing price. Consistent with my theory and existing literature (Ramani and Bloom 2021), communities with higher levels of *Commute* experience a significant increase in expected long-term housing demand across all zip-codes. As with *Remote*, the magnitude of the effect varies with population size with the effect being four times larger in *Urban* regions. Finally, *Age* has no significance in either Column 3 or 4.

Zip-Level Demographics' Effect on Percent Change in Housing Price (1)

	(1)	(2)	(3)	(4)
Remote	-0.0381***	-0.1817***	-0.0615***	-0.3137***
	(0.0109)	(0.0481)	(0.0093)	(0.0401)
Income	-0.0035***	-0.0083***	-0.0025***	-0.0024**
	(0.0005)	(0.0014)	(0.0004)	(0.0012)
Family	0.0875***	0.1801***	0.0674***	0.1625***
	(0.0076)	(0.0241)	(0.0065)	(0.0200)
Education	0.0110***	0.0188***	0.0075***	0.0184***
	(0.0011)	(0.0040)	(0.0010)	(0.0035)
Age	0.0009***	-0.0011***	0.00001	-0.0002
	(0.0001)	(0.0004)	(0.0001)	(0.0003)
Commute	0.0011***	0.0090***	0.0005***	0.0037***
	(0.0001)	(0.0012)	(0.0001)	(0.0011)
Population	0.0191***	0.0051***	0.0090***	-0.0005
	(0.0005)	(0.0013)	(0.0005)	(0.0011)
Constant	0.0194	0.0320	-0.0810***	0.0063
	(0.0160)	(0.0518)	(0.0239)	(0.0520)
State Fixed effects	No	No	Yes	Yes
Urban Subset	No	Yes	No	Yes
Observations	28,190	2,610	28,190	2,610
R^2	0.0562	0.0918	0.3459	0.4406
Adjusted R^2	0.0559	0.0893	0.3445	0.4328
Note:				*p<0.1; **p<0.05; ***p<0.01

7.3. Equation 2 Results

Every *Remote* interaction is significant, when state fixed effects are included, for all zip-codes and the *Urban* subset. This supports my hypothesis that there exist heterogeneous levels of willingness to relocate for workers that can work remotely.

Both interactions with a hypothesized sign have the opposite sign of what was expected. Highly remote communities with many families experience a demand increase in Column 3; this effect is over 40% larger in magnitude when restricted to *Urban* regions (Column 4). This implies highly populated regions with many families that can work remotely experienced a large increase in housing demand. Communities that are both remote-capable and older experience a larger increase in housing demand than the additive alone predicts. Both findings contradict my hypothesis on which individuals were most likely to relocate. The interaction with *Income* is negative for all zip-codes and positive when restricted to Urban zip-codes. The interaction with *Education* is always negative, but three times larger in magnitude for *Urban* zip-codes.

Zip-Level Demographics' Effect on Percent Increase in Housing Price (2)

	(1)	(2)	(3)	(4)
Remote	-1.112***	-1.406***	-0.578***	-0.445
	(0.110)	(0.357)	(0.092)	(0.285)
Income	0.012***	-0.012***	0.011***	-0.013***
	(0.001)	(0.004)	(0.001)	(0.004)
Family	-0.254***	-0.488***	-0.272***	-0.365***
	(0.020)	(0.080)	(0.017)	(0.064)
Education	0.009***	0.020**	0.016***	0.043***
	(0.003)	(0.009)	(0.002)	(0.007)
Age	-0.004***	-0.005***	-0.004***	-0.003***
	(0.0003)	(0.001)	(0.0002)	(0.001)
Commute	0.001***	0.010***	0.001***	0.004***
	(0.0001)	(0.001)	(0.0001)	(0.001)
Population	0.019***	0.005***	0.009***	-0.001
	(0.001)	(0.001)	(0.0005)	(0.001)
Remote*Income	-0.034***	0.003	-0.027***	0.020***
	(0.003)	(0.009)	(0.002)	(0.007)
Remote*Family	0.805***	1.450***	0.793***	1.127***
	(0.049)	(0.168)	(0.041)	(0.134)
Remote*Education	0.014*	-0.003	-0.019***	-0.061***
	(0.007)	(0.021)	(0.006)	(0.017)
Remote*Age	0.012***	0.009***	0.009***	0.005**
	(0.001)	(0.003)	(0.001)	(0.002)
Constant	0.397***	0.633***	0.098**	0.161
	(0.041)	(0.146)	(0.039)	(0.120)
State Fixed effects	No	No	Yes	Yes
Urban Subset	No	Yes	No	Yes

Observations	28,190	2,610	28,190	2,610
R^2	0.080	0.137	0.367	0.479
Adjusted R^2	0.079	0.134	0.365	0.471
Note:				*p<0.1; **p<0.05; ***p<0.01

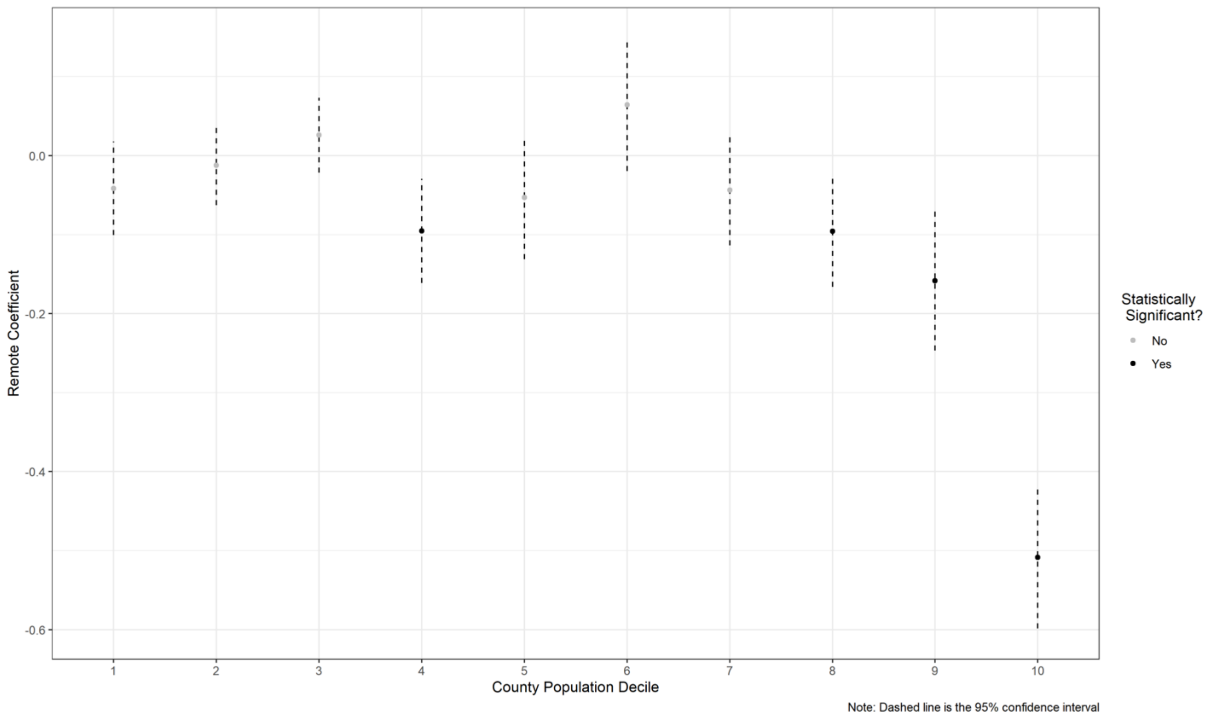
7.4. Equation 1 by County Population

The below graph displays a scatterplot of the county population deciles versus the coefficient on the *Remote* variable. The full regression table can be found in Appendix 10.6. I will provide representative examples[9] to better understand the types of communities in each decile. For the 1st decile, Cross Plains, TX is a town with about 1,000 people, one stoplight, and eight churches. In the 5th decile, Eau Claire, WI is home to a four-year college, a hospital, and a mall. In the 10th decile, San Diego, CA, a well-known large city, has its own airport, thousands of restaurants, and an MLB baseball team. Therefore, the 1st decile regression can be thought of as restricting the model to rural villages, the 5th is effectively restricting to medium-sized towns, and the 10th decile is a large city restriction on the model.

I ran the equation 1 regression model on each population decile separately. Significance only occurs once in the first 7 deciles. From the 8th decile onward, however, there is consistent significance. Only the 10th decile's effect size appears to defy an otherwise marginal effect size of the *Remote* variable. These results imply that communities with large amounts of work that can be done remotely experienced negative demand shocks in large urban areas and little to no change in demand in suburban and rural regions. This result is consistent with my hypothesis that larger cities will experience a larger exodus than other region types when commute time is less important for its residents. It is also consistent with Ramani and Bloom (2021) who claim that long-run housing demand effects only exist for large cities.

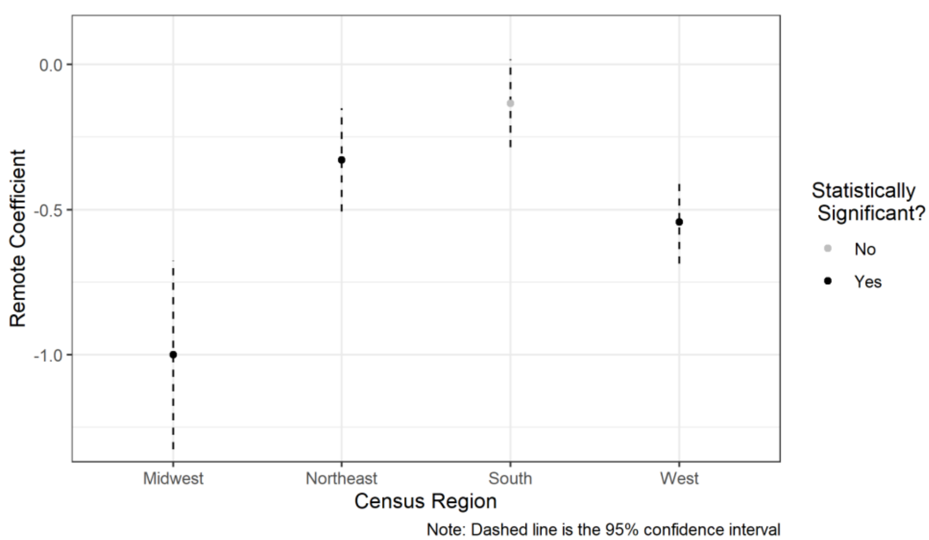
Additionally, the apparent homogeneity of effect size for the first 9 deciles acts as a robustness check for Remote's coefficient in Column 3 for sections 7.2 and 7.3. By splitting the regression by population size and observing similar effects for the first 9 deciles, we know that the all zip-code model specification was hiding heterogeneous effects in the top 10th decile for the *Remote* variable, but not for any other county population level.

[9] I selected counties with population sizes close to the mean population within each decile to serve as the representative community for the decile.



7.5. Equation 1 by Census Region

The below graph displays a scatterplot of census regions versus the coefficient on the *Remote* variable. The full regression table can be found in Appendix 9.7. Since Census regions have large variance in average population size[10] and the previous section revealed heterogeneity in *Remote*'s effect on housing demand by population, the region model is restricted to counties with populations in the 10th decile to ensure any heterogeneity discovered is not simply because the region control is acting as a proxy for population.



[10] Mean zip-code population 7,500 in the Midwest and 15,600 in the West.

Notably, the effect size of *Remote* in the South is less than half the size of the next smallest coefficient value and it is insignificant at the 95% level. *Remote* is statistically significant for the remaining three regions and peaks in magnitude in the Midwest.

8. CONCLUSION

Given that *Remote* has a significant, negative coefficient in every model specification, I broadly conclude that an increase in WFH leads many workers to move. Since *Commute* is significant and positive, their new housing location is likely to be further from the city center and their office.

I am led to believe there is heterogeneity in willingness to relocate when offered the option to work remotely given that the interactions between individual characteristics and *Remote* are all statistically significant. Through the lens of my Cobb-Douglas location choice model, this can be explained by either a set of amenity preferences or moving costs that are similar within income, education, marital-status, or age groups and dissimilar across demographic groups. While the expected signs were not realized, I retain the belief that young unmarried people benefit greater from a city's available amenities. Instead, I conclude that heterogeneous moving costs are a larger factor than initially supposed. A moving cost a family might face in leaving the city is forcing their children to make new friends in a new school district. A moving cost an older worker could face is difficulty integrating into a new community at a life stage when social outings are infrequent.

A simplification in the model admits issues. Remote work is treated as a strictly beneficial shock to the worker's amenity utility function, $s_i(A_{i,p})$. Additionally, it assumes that the amount of WFH possible is fully actualized. If there are costs to working remotely, this might not be the case. For example, if a mother of young children does not want to be seen on a Zoom call with a child crying in the background, she might continue to travel to the office to take advantage of an environment where she can be undisturbed. If she were to take full advantage of the remote work her employer offers, she might need to live in a home with an additional room to be her home office. This would be a cost associated with working remotely that might make relocation and fully realized WFH unprofitable. The problem with this hypothesis is that existing data is contradictory. As stated in the literature review, remote work was well received across demographics and largely desired post-pandemic (Barrero, Bloom, and Davis 2021). With respect to the young mother example, Barrero et al. find that women with at least some college and children under 12 desire ~2.5 days remote on average, ~.2 days more than the overall average. This leads me to believe that proximity to young children during the workday is a benefit from the worker's perspective. Additionally, they find that the average worker has invested 15 hours and $561 into fixed costs related to WFH. These include video-conference software education, computers, furniture, and internet connection. Given that many have already paid WFH related costs and workers most at risk of household related difficulties still desire WFH, I conclude that the incorporation of remote work cost into my theory and models would have a negligible impact on my findings. However, remote work cost may still be an essential factor when observing WFH's long-term impact from a firm's perspective. Productivity levels, for example, may be largely correlated with which employees receive WFH and to what extent.

While the model specification in section 7.2 determines there is a significant, negative relationship between amount of WFH and long-term housing demand, section 7.4 shows that the 10^{th} decile of population is an outlier driving this result. By referring again to my model and assuming that the individual characteristic variables controlled for in the model sufficiently capture amenity preference differences, this would imply that the bid-rent gradient is largest in highly urban areas.

Concerning the small, insignificant coefficient on *Remote* in the South in section 7.5, two potential explanations can be found in Gupta et al. (2021). They find housing supply inelasticity and stringency of local lockdowns to flatten the housing price bid-rent gradient, but stringency is small and insignificant once inelasticity is controlled for. Given Texas is well known for its lenient development laws and has few geographic constraints (Saiz 2008), it's possible urban Texans capable of remote work do not need to relocate as a result of pressing financial constraints or there is not a large enough financial gain from relocating. Unfortunately, this explanation does not apply to the ordering of the other regions. Los Angeles, San Francisco, New York, Boston, and San Diego all have supply elasticities lower than Chicago, Detroit, or Minneapolis (Saiz 2008), yet the coefficient for Midwest is larger than that of the West or Northeast with 95% confidence.

In conclusion, this paper is able to address remote work's influence on the demand for housing both across the U.S. and restricted to urban areas. It identifies heterogeneity in workers' willingness to relocate controlling for the amount of work they can do from home. It also identifies highly populated counties as the most vulnerable to emigration. It does a poor job, however, in identifying the causal channels leading to the heterogeneity in remote work's effect. Further research could rectify this problem by attempting to recreate my econometric models at the individual level supplemented with survey information on resistance to relocation and amenity preferences. This would directly identify how the WFH-induced migration effect differs by the causal channels addressed in my theory, amenity preferences and relocation costs.

Given that the expected amount of work done remotely post-pandemic is positively correlated with income (Barrero et al. 2021), city policymakers should expect a smaller budget in the near future. On the other hand, less in-person work means a decrease in the amount work related transportation, which makes downsizing public transportation an evident choice to keep the city's budget under control.

9. REFERENCES

Autor, D. H. (2019). Work of the Past, Work of the Future. *AEA Papers and Proceedings*, *109*, 1–32. https://doi.org/10.1257/pandp.20191110

Autor, D., & Salomons, A. (2019). New Frontiers: The Evolving Content and Geography of New Work in the 20th Century [Extended Abstract].

Barrero, J. M., Bloom, N., & Davis, S. (2021). Why Working from Home Will Stick. *NBER Working Paper No. 28731*. https://doi.org/10.3386/w28731

Bartik, A. W., Bertrand, M., Cullen, Z., Glaeser, E. L., Luca, M., & Stanton, C. (2020). The Impact of Covid-19 on Small Business Outcomes and Expectations. *Proceedings of the National Academy of Sciences*, *117*(30), 17656–17666. https://doi.org/10.1073/pnas.2006991117

Becker, G. S. (1965). A Theory of the Allocation of Time. *The Economic Journal*, *75*(299), 493. https://doi.org/10.2307/2228949

Brinkman, J. (2015). Big Cities and the Highly Educated: What's the Connection? *Federal Reserve Bank of Philadelphia Business Review*, *98*(3).

Couture, V., & Handbury, J. (2020). Urban Revival in America. *Journal of Urban Economics*, *119*, 103267. https://doi.org/10.1016/j.jue.2020.103267

Diamond, R. (2016). The Determinants and Welfare Implications of US Workers' Diverging Location Choices by Skill: 1980-2000. *American Economic Review*, 106 (3): 479-524.

Dingel, J., & Neiman, B. (2020). How Many Jobs Can Be Done at Home? *NBER Working Paper No. 26948.* https://doi.org/10.3386/w26948

Glaeser, E. L., & Cutler, D. M. (2021). *Survival of the city living and thriving in an age of isolation*. Penguin Press.

Gupta, A., Mittal, V., Peeters, J., & Van Nieuwerburgh, S. (2021). Flattening the curve: Pandemic-induced revaluation of Urban Real Estate. *NBER Working Paper No. 28675.* https://doi.org/10.3386/w28675

Morikawa, M. (2021). Work-from-home productivity during the Covid-19 pandemic: Evidence from Japan. *Economic Inquiry*, *60*(2), 508–527. https://doi.org/10.1111/ecin.13056

National Bureau of Economic Research, Total Civilian Labor Force, 14 Years and Older for United States [M0825AUSM175NNBR], retrieved from FRED, Federal Reserve Bank of St. Louis; https://fred.stlouisfed.org/series/M0825AUSM175NNBR, April 7, 2022.

Oh, D. (1995). Households with Rent Burdens: Impact on Other Spending and Factors Related to the Probability of Having a Rent Burden. Financial Counseling and Planning, 6, 139-147.

Ramani, A. and Bloom, N. (2021). The Donut Effect: How COVID-19 Shapes Real Estate, SIEPR Policy Brief, January.

Ren, H. (2020). Suburban Airbnb Pivots to Long-Term Renting After Coronavirus Hit. *Medill Reports*. Retrieved April 18, 2022, from https://news.medill.northwestern.edu/chicago/suburban-airbnb-pivots-to-long-term-renting-after-coronavirus-hit/.

Saiz, A. (2008). On local housing supply elasticity. *SSRN Electronic Journal*. https://doi.org/10.2139/ssrn.1193422

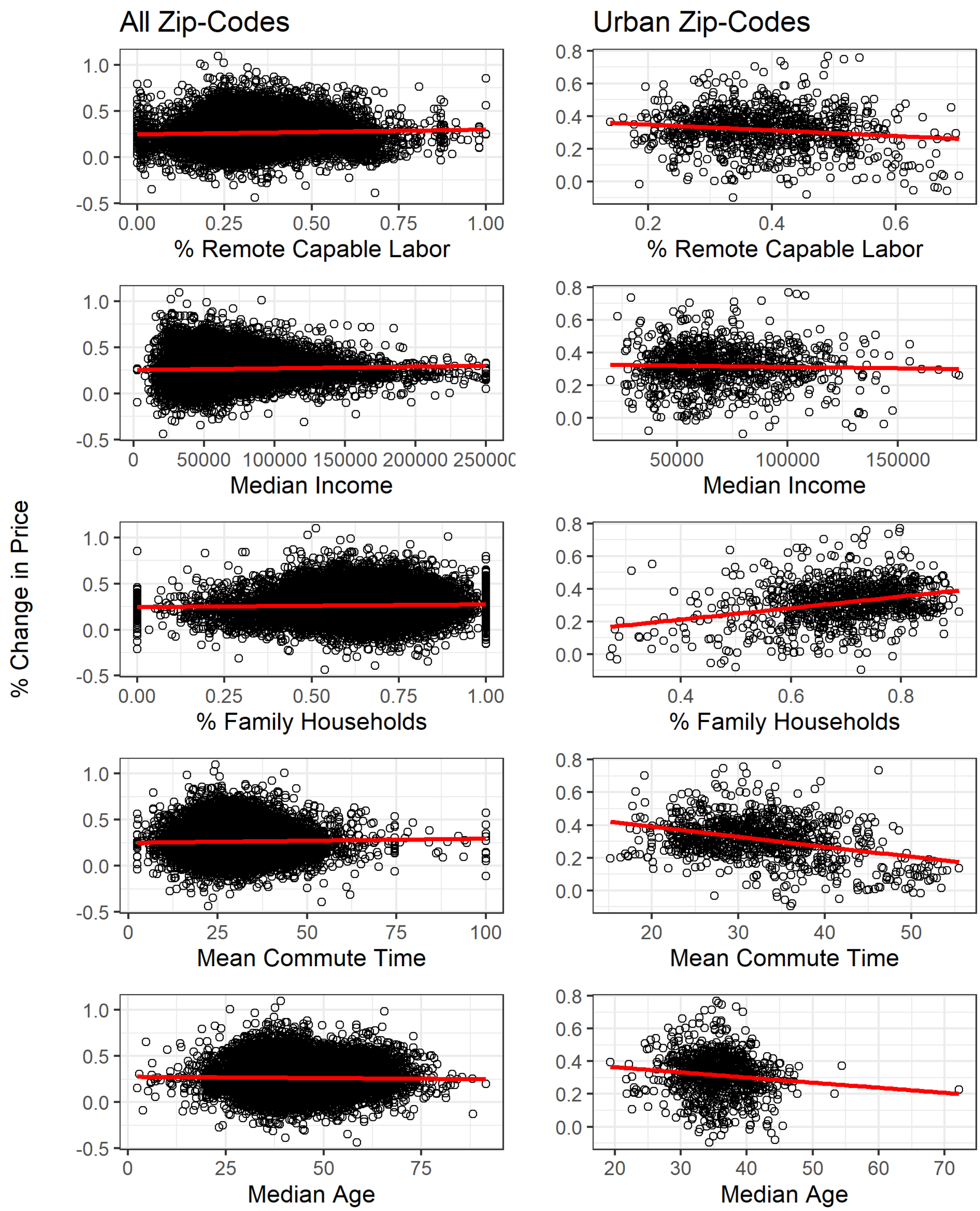
Speakman, M., Tucker, J., & Bachaud, N. (2022). Housing Data: Home Values. Retrieved March 2022, from https://www.zillow.com/research/data/.

Manson S., Schroeder J., Van Riper D., Kugler T., and Ruggles S. (2021). IPUMS National Historical Geographic Information System: Version 16.0 [dataset]. Minneapolis, MN: IPUMS. http://doi.org/10.18128/D050.V16.0

World Bank. Labor force, total – United States. World Development Indicators, The World Bank Group, 2022, https://data.worldbank.org/indicator/SL.TLF.TOTL.IN?locations=US. Accessed 7 April 2022.

10. APPENDIX

10.1. Heterogeneity in Covariates' Relationship with Price Change by Population



10.2. Utility Maximization Algebra

If the worker is maximizing his utility function, he is maximizing the log-transformed utility function.

$$\max_{M,O,p} ln(U) = \alpha ln(M) + (1-\alpha) ln(O) + s_i\big(A_{i,p}\big)$$

Substitute in first order condition.

$$\max_{M,p} log(U) = \alpha log(M) + (1-\alpha) log\left(\frac{(1-\alpha)}{\alpha}\frac{R_p}{P}M\right) + s_i\big(A_{i,p}\big)$$

$$\max_{M,p} log(U) = \alpha log(M) + (1-\alpha)\left[log\left(\frac{(1-\alpha)}{\alpha}\right) + log\left(\frac{R_p}{P}\right) + log(M)\right] + s_i\big(A_{i,p}\big)$$

$$\max_{M,p} log(U) = \alpha log(M) + (1-\alpha) log\left(\frac{(1-\alpha)}{\alpha}\right) + (1-\alpha) log\left(\frac{R_p}{P}\right) + (1-\alpha) log(M) + s_i\big(A_{i,p}\big)$$

$$\max_{M,p} log(U) = (1-\alpha) log\left(\frac{(1-\alpha)}{\alpha}\right) + (1-\alpha) log\left(\frac{R_p}{P}\right) + log(M) + s_i\big(A_{i,p}\big)$$

Substitute in wage budget constraint.

$$\max_{p} log(U) = (1-\alpha) log\left(\frac{(1-\alpha)}{\alpha}\right) + (1-\alpha) log\left(\frac{R_p}{P}\right) + log\left(\frac{W_i}{\left(1+\frac{(1-\alpha)}{\alpha}\right)R_p}\right) + s_i\big(A_{i,p}\big)$$

$$\max_{p} log(U) = (1-\alpha) log\left(\frac{(1-\alpha)}{\alpha}\right) + (1-\alpha) log\left(\frac{R_p}{P}\right) + log(W_i) - log\left(\left(1+\frac{(1-\alpha)}{\alpha}\right)R_p\right) + s_i\big(A_{i,p}\big)$$

$$\max_{p} log(U) \quad = (1-\alpha) log\left(\frac{(1-\alpha)}{\alpha}\right) + (1-\alpha) log\big(R_p\big) - (1-\alpha) log(P) + log(W_i) - log\left(1+\frac{(1-\alpha)}{\alpha}\right) - log\big(R_p\big) + s_i\big(A_{i,p}\big)$$

$$\max_{p} log(U) = (1-\alpha) log\left(\frac{(1-\alpha)}{\alpha}\right) - log\left(1+\frac{(1-\alpha)}{\alpha}\right) + (-\alpha) log\big(R_p\big) - (1-\alpha) log(P) + log(W_i) + s_i\big(A_{i,p}\big)$$

Substitute in amenity preference functional form.

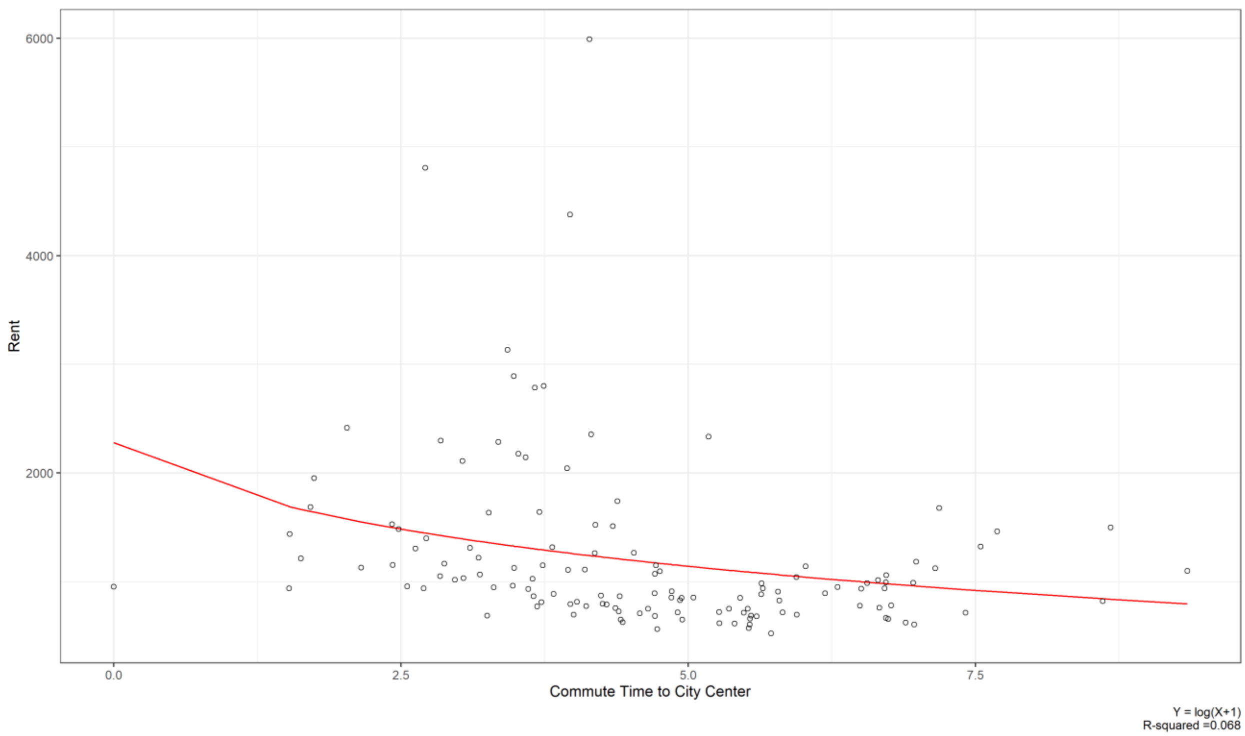
$$
\begin{aligned}
\max_p log(U) \quad &= (1-\alpha)log\left(\frac{(1-\alpha)}{\alpha}\right) - log\left(1+\frac{(1-\alpha)}{\alpha}\right) + (-\alpha)log\left(R_p\right) \\
&- (1-\alpha)log(P) + log(W_i) + \beta_{s_0}T_c + \beta_{s_1}T_l
\end{aligned}
$$

Substitute in R, T_c, and T_l's functional form.

$$
\begin{aligned}
\max_p log(U) \quad &= (1-\alpha)log\left(\frac{(1-\alpha)}{\alpha}\right) - log\left(1+\frac{(1-\alpha)}{\alpha}\right) + (-\alpha)log\left(\beta_{R_0} + \beta_{R_1}\log(c_p+1)\right) \\
&- (1-\alpha)log(P) + log(W_i) + \beta_{s_0}\left((1-r_i)c_p + (1-q_i)c_p\right) \\
&+ \beta_{s_1}\left(T - T_W - (1-r_i)c_p - (1-q_i)c_p\right)
\end{aligned}
$$

10.3. Estimated Rent and Entertainment Functional Form

To estimate Jessica and Joel's bid-rent function, R_p, I take advantage of the dataset I constructed as described in section 5. I wanted to have an urban county that was geographically large to have larger changes in optimal commute time. For this reason, I chose Harris County, the most populated county in Texas that is almost 2 million square miles large. To calculate rent, I divided the Zillow January 2020 housing price by 200. I then ran an OLS regression of R_p.



This regression results in the following coefficient values:

$$\beta_{R_0} = 2280.3, \beta_{R_1} = -635.1$$

10.4. Commute Variable Comparison

10.4.1. Table 1: Both Commute Variables

	Both Commute Variables			
	(1)	(2)	(3)	(4)
Remote	-0.038***	-0.179***	-0.061***	-0.314***
	(0.011)	(0.048)	(0.009)	(0.040)
Income	-0.004***	-0.006***	-0.003***	-0.002**
	(0.0005)	(0.001)	(0.0004)	(0.001)
Family	0.085***	0.194***	0.065***	0.163***
	(0.008)	(0.024)	(0.006)	(0.020)
Education	0.012***	0.014***	0.008***	0.018***
	(0.001)	(0.004)	(0.001)	(0.004)
Age	0.001***	-0.001**	-0.00004	-0.0002
	(0.0001)	(0.0004)	(0.0001)	(0.0003)
Commute (Google)	0.001***	0.010***	0.001***	0.004***
	(0.0001)	(0.001)	(0.0001)	(0.001)
Commute (Census)	0.0005***	-0.003***	0.0003***	-0.0001
	(0.0001)	(0.0003)	(0.0001)	(0.0003)
Population	0.019***	0.007***	0.009***	-0.001
	(0.001)	(0.001)	(0.0005)	(0.001)
Constant	0.0004	0.144***	-0.090***	0.009
	(0.017)	(0.054)	(0.024)	(0.053)
State Fixed Effects	No	No	Yes	Yes
Urban Subset	No	Yes	No	Yes
Observations	28,184	2,610	28,184	2,610
R^2	0.057	0.110	0.346	0.441
Adjusted R^2	0.057	0.107	0.345	0.433
Note:		*p<0.1; **p<0.05; ***p<0.01		

10.4.2. Table 2: Google Commute Variable

	Google Commute Variable			
	(1)	(2)	(3)	(4)
Remote	-0.038***	-0.182***	-0.062***	-0.314***
	(0.011)	(0.048)	(0.009)	(0.040)
Income	-0.004***	-0.008***	-0.003***	-0.002**
	(0.0005)	(0.001)	(0.0004)	(0.001)
Family	0.087***	0.180***	0.067***	0.163***
	(0.008)	(0.024)	(0.006)	(0.020)
Education	0.011***	0.019***	0.007***	0.018***

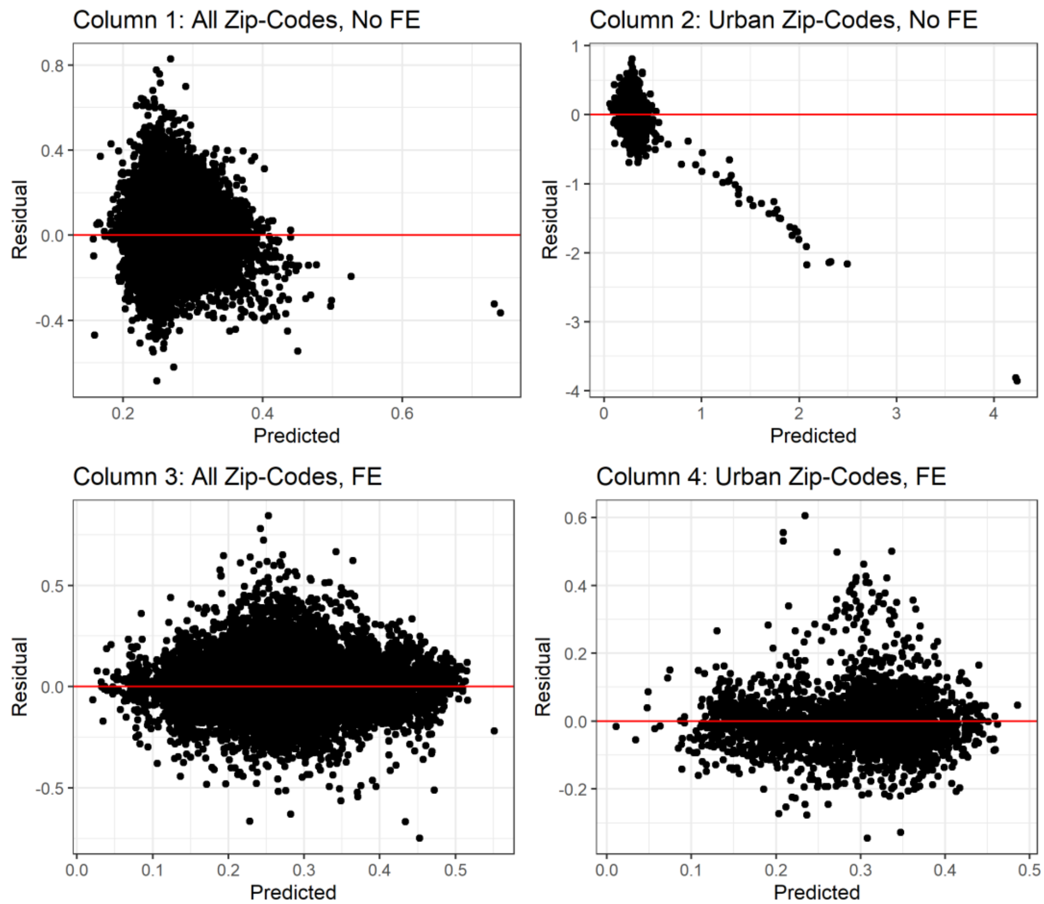
	(0.001)	(0.004)	(0.001)	(0.004)
Age	0.001***	-0.001***	0.00001	-0.0002
	(0.0001)	(0.0004)	(0.0001)	(0.0003)
Commute (Google)	0.001***	0.009***	0.001***	0.004***
	(0.0001)	(0.001)	(0.0001)	(0.001)
Population	0.019***	0.005***	0.009***	-0.001
	(0.001)	(0.001)	(0.0005)	(0.001)
Constant	0.019	0.032	-0.081***	0.006
	(0.016)	(0.052)	(0.024)	(0.052)
State Fixed Effects	No	No	Yes	Yes
Urban Subset	No	Yes	No	Yes
Observations	28,190	2,610	28,190	2,610
R^2	0.056	0.092	0.346	0.441
Adjusted R^2	0.056	0.089	0.345	0.433
Note:		*p<0.1; **p<0.05; ***p<0.01		

10.4.3. Table 3: Census Commute Variable

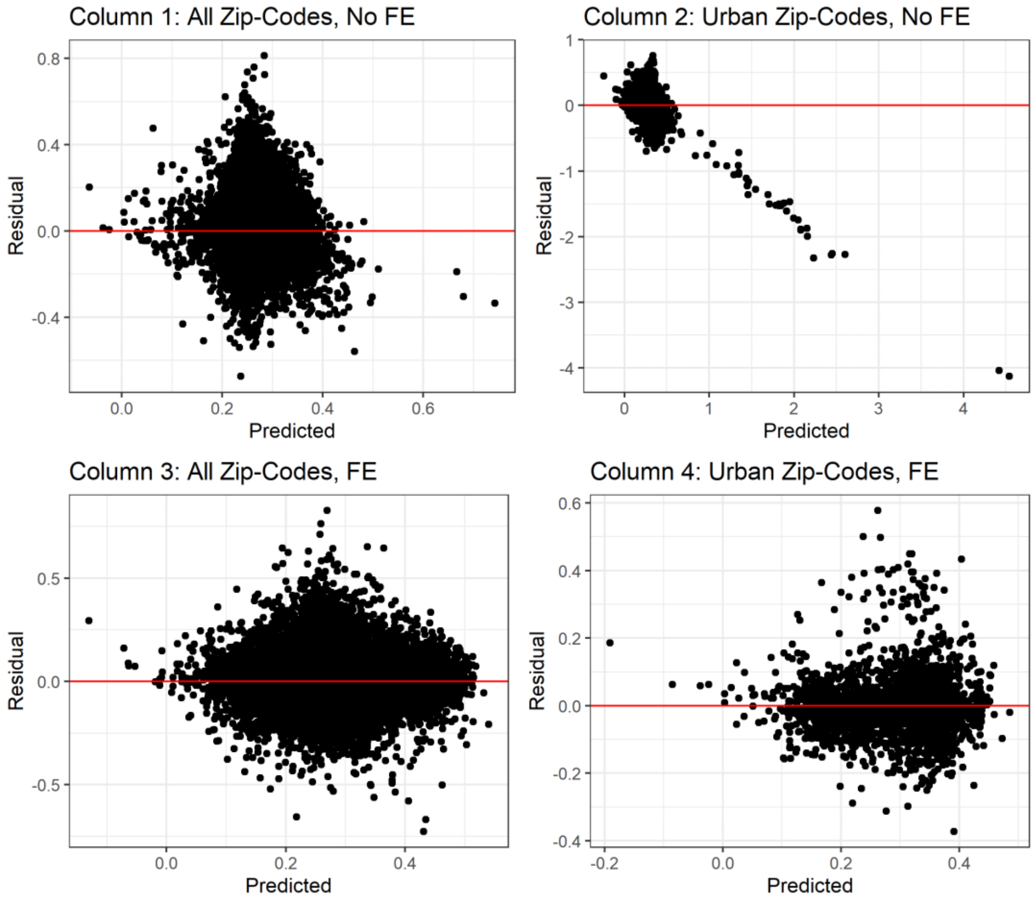
	Census Commute Variable			
	(1)	(2)	(3)	(4)
Remote	-0.041***	-0.185***	-0.063***	-0.320***
	(0.011)	(0.048)	(0.009)	(0.040)
Income	-0.004***	-0.006***	-0.003***	-0.003**
	(0.0005)	(0.001)	(0.0004)	(0.001)
Family	0.084***	0.198***	0.064***	0.169***
	(0.008)	(0.024)	(0.006)	(0.020)
Education	0.012***	0.015***	0.008***	0.020***
	(0.001)	(0.004)	(0.001)	(0.004)
Age	0.001***	-0.001	-0.00002	0.00002
	(0.0001)	(0.0004)	(0.0001)	(0.0003)
Commute (Census)	0.001***	-0.002***	0.0003***	0.00001
	(0.0001)	(0.0003)	(0.0001)	(0.0003)
Population	0.019***	0.006***	0.009***	-0.001
	(0.001)	(0.001)	(0.0005)	(0.001)
Constant	-0.005	0.144***	-0.039*	-0.008
	(0.017)	(0.054)	(0.020)	(0.053)
State Fixed Effects	No	No	Yes	Yes
Urban Subset	No	Yes	No	Yes
Observations	28,352	2,619	28,352	2,619
R^2	0.053	0.087	0.346	0.438
Adjusted R^2	0.053	0.084	0.345	0.430
Note:		*p<0.1; **p<0.05; ***p<0.01		

10.5. Regression Endogeneity Concerns

10.5.1. Equation 1 Residuals



10.5.2. Equation 2 Residuals



10.6. Equation 1 by County Population

Zip-Level Demographics' Effect on Percent Change in Housing Price by Region (1)

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
Remote	-0.0414	-0.0122	0.0259	-0.0953***	-0.0530	0.0641	-0.0437	-0.0958***	-0.1583***	-0.5086***
	(0.0303)	(0.0257)	(0.0242)	(0.0337)	(0.0399)	(0.0426)	(0.0356)	(0.0358)	(0.0452)	(0.0458)
Income	0.0003	0.0014	0.0033*	0.0052***	0.0058***	-0.0019	0.0007	0.0007	0.0004	-0.0058***
	(0.0021)	(0.0018)	(0.0017)	(0.0020)	(0.0020)	(0.0020)	(0.0017)	(0.0015)	(0.0014)	(0.0013)
Family	-0.0100	0.0380*	0.0110	0.0704**	0.0352	0.0440	0.0108	-0.0140	0.1056***	0.2378***
	(0.0267)	(0.0211)	(0.0213)	(0.0281)	(0.0321)	(0.0286)	(0.0249)	(0.0227)	(0.0241)	(0.0210)
Education	0.0086**	0.0091***	0.0010	0.0116***	0.0027	-0.0041	-0.0059	0.0037	0.0022	0.0323***
	(0.0042)	(0.0032)	(0.0035)	(0.0039)	(0.0043)	(0.0047)	(0.0036)	(0.0034)	(0.0043)	(0.0037)
Age	0.00001	0.0003	-0.0003	0.0010***	0.0008**	0.0012***	0.0007**	0.0003	-0.0008**	0.0004
	(0.0003)	(0.0002)	(0.0003)	(0.0003)	(0.0004)	(0.0004)	(0.0003)	(0.0003)	(0.0004)	(0.0003)

Commute	0.00003	0.00001	0.0001	-0.0041***	0.0002	-0.0003	0.0006	-0.0006	0.0072***	0.0033***
	(0.0003)	(0.0011)	(0.0002)	(0.0011)	(0.0003)	(0.0007)	(0.0004)	(0.0005)	(0.0016)	(0.0011)
Population	0.0153	-0.0030	-0.0021	0.0043	0.0103***	0.0082***	0.0088***	0.0048***	0.0042***	-0.0009
	(0.0102)	(0.0044)	(0.0029)	(0.0030)	(0.0025)	(0.0023)	(0.0019)	(0.0014)	(0.0014)	(0.0011)
Constant	0.1122	0.0343	0.0707	-0.0172	-0.0445	0.1537**	0.0711	0.2485***	0.2116***	-0.0762*
	(0.1096)	(0.0429)	(0.0714)	(0.0555)	(0.0721)	(0.0648)	(0.0554)	(0.0479)	(0.0595)	(0.0453)
State Fixed effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
County Population Decile	1st	2nd	3rd	4th	5th	6th	7th	8th	9th	10th
Observations	1,517	1,615	1,624	1,612	1,659	1,653	1,646	1,677	1,715	1,714
R^2	0.3092	0.4392	0.4611	0.4679	0.5621	0.4620	0.5124	0.5482	0.6097	0.4981
Adjusted R^2	0.2871	0.4216	0.4443	0.4519	0.5485	0.4462	0.4987	0.5349	0.6001	0.4907
Note:										*p<0.1; **p<0.05; ***p<0.01

10.7. Equation 1 by Census Region

Zip-Level Demographics' Effect on Percent Increase in Housing Price by Region (1)

	(1)	(2)	(3)	(4)
Remote	-1.0005***	-0.3288***	-0.1339*	-0.5425***
	(0.1660)	(0.0909)	(0.0773)	(0.0739)
Income	-0.0085**	0.0003	-0.0010	-0.0079***
	(0.0035)	(0.0028)	(0.0022)	(0.0023)
Family	0.1130	0.1106*	0.1243***	0.3551***
	(0.0705)	(0.0617)	(0.0310)	(0.0334)
Education	0.0526***	0.0032	-0.0024	0.0509***
	(0.0143)	(0.0084)	(0.0060)	(0.0058)
Age	0.0003	0.0002	-0.0005	0.0001
	(0.0012)	(0.0007)	(0.0005)	(0.0006)
Commute	-0.0074**	0.0034	-0.0027	0.0091***
	(0.0033)	(0.0030)	(0.0022)	(0.0015)
Population	-0.0051*	-0.0095***	-0.0009	0.0048***
	(0.0029)	(0.0030)	(0.0018)	(0.0017)
Constant	-0.2583	0.3028***	0.4109***	-0.4264***
	(0.1786)	(0.1167)	(0.0768)	(0.0700)
State Fixed effects	Yes	Yes	Yes	Yes
Census Region	Midwest	Northeast	South	West
Observations	243	313	511	647
R^2	0.5677	0.5535	0.3844	0.4394