The Impact of the COVID-19 Pandemic on the Demand for Density: Evidence from the U.S. Housing Market

Liu, Sitian and Su, Yichen

Queen's University, Federal Reserve Bank of Dallas

27 July 2020

The Impact of the COVID-19 Pandemic on the Demand for Density: Evidence from the U.S. Housing Market

Sitian Liu* Yichen Su†

July 27, 2020

Abstract

Cities are shaped by the strength of agglomeration and dispersion forces. We show that the COVID-19 pandemic has re-introduced disease transmission as a dispersion force in modern cities. We use detailed housing data to study the impact of the COVID-19 pandemic on the location demand for housing. We find that the pandemic has led to a greater decline in the demand for housing in neighborhoods with high population density. We further show that the reduced demand for density is partially driven by the diminished need of living close to jobs that are telework-compatible and the declining value of access to consumption amenities. Neighborhoods with high pre-COVID-19 home prices also see a greater drop in housing demand. While the national housing market partially recovered in June, we show that the negative effect of the pandemic on the demand for density persists, indicating that the change in the demand for density may last beyond an aggregate recovery of housing demand.

Keywords: COVID-19, Pandemic, Density, City, Neighborhood, Housing, Location, Telework, Amenity

JEL Codes: R2, R3, I1

*Department of Economics, Queen's University. Email: sitianliu@econ.queensu.ca.

†Federal Reserve Bank of Dallas. Email: yichensu@outlook.com. The views expressed in this article are those of the authors and do not necessarily represent the views of the Federal Reserve Bank of Dallas or the Federal Reserve System.

1 Introduction

The desirability and structure of cities are shaped by the strength of agglomeration and dispersion forces (Ahlfeldt et al., 2015). During much of human history, the dispersion forces such as communicable disease and overcrowding kept the sizes of settlements and cities from growing too large (Glaeser, 2011). Thanks to the infrastructure and medical advancements, these dispersion forces have been greatly reduced. With these dispersion forces in check, the agglomeration effect of human interaction for production and consumption is allowed to flourish, leading to the creation of large cities and dense neighborhoods (Glaeser et al., 2000; Glaeser and Gotllieb, 2000).

However, the COVID-19 pandemic has re-introduced the danger of disease transmission as a potentially serious dispersion force in modern cities (Autor and Reynolds, 2020). Due to the contagious nature of the disease, office workers no longer commute to crowded urban locations for work, and instead conduct businesses virtually at home (Bartik et al., 2020; Mas and Pallais, 2020; Bick et al., 2020). Consumption amenities such as restaurants have also seen much fewer visits, owing to policy mandates or consumers' concerns over the potential exposure to the virus in indoor public spaces (Allcott et al., 2020; Chen et al., 2020; Chetty et al., 2020; Cox et al., 2020). While some are hopeful that foot traffic may recover in a short time, many would argue that the aversion to crowded venues could continue well into the future.[1] Some companies even started to allow their employees to permanently work remotely.[2] If the aversion to crowded locations is expected to persist, it could reduce the desirability of large cities and dense neighborhoods, and lower the demand for housing in these locations.

In this paper, we use detailed geocoded housing data to study the spatial impact of the COVID-19 pandemic on location demand for housing across neighborhoods within cities and across cities in the U.S. We find that the COVID-19 pandemic has led to a greater decline in the demand for housing closer to central cities and in neighborhoods with higher population density. Home sales in central cities and dense neighborhoods dropped considerably more relative to other comparable locations since the outbreak of COVID-19. In contrast, although new listings have dropped as well in aggregate since the outbreak, we do not find differences in the decline by location characteristics such

[1] https://www.cnn.com/2020/04/21/opinions/bergen-osterholm-interview-two-opinion/index.html; https://www.hsph.harvard.edu/news/hsph-in-the-news/report-covid-19-will-likely-spread-for-up-to-two-more-years.

[2] https://www.npr.org/2020/06/22/870029658/get-a-comfortable-chair-permanent-work-from-home-is-coming.

as distance to downtown or density.[3] These results suggest that the pandemic may have changed people's demand for density.

We make several conjectures of potential reasons behind the decline, and test these conjectures in the data.

1. Dense neighborhoods tend to be close to large job centers (e.g., central business districts), which tend to have a greater share of jobs that are teleworking-compatible.[4] As the pandemic kicks in and people start to work from home, the need for living in these neighborhoods becomes obsolete. Therefore, the demand for the central locations could diminish.

2. Dense neighborhoods tend to have good access to a larger selection of consumption amenities.[5] Because of the need for social distancing, people are substituting away from visiting restaurants with home production. As a result, the value of living closer to consumption amenities may decline, leading to a lower demand for neighborhoods with premium access to amenities.

3. Dense neighborhoods tend to have higher costs of housing (Brueckner, 2011).[6] The demand for density had been rising in the U.S. before the pandemic (Autor and Reynolds, 2020), as large cities and dense neighborhoods were increasingly sought after.[7] As the driving forces behind such demand vaporize because of the need for social distancing, the value of bearing high housing costs to be in these locations is also decreased. Therefore, there could be an exodus of housing demand from locations of high house prices prior to the pandemic.

4. People may become more concerned about living in locations with high density per se due to the pandemic. First, people may perceive that dense locations are more prone to disease transmission due to overcrowding. For example, people are much more likely to bump into each other or share an elevator in a high-rise urban apartment than in suburban neighborhoods with

[3] One alternative explanation of higher-density areas seeing a greater decline in sales is that these places are more likely to be locked down because of the pandemic. Our results are robust to the concern because we control for MSA-specific effects of the pandemic, which could account for differential likelihoods of lockdowns across MSAs. Moreover, if there are differential likelihoods of lockdowns across neighborhoods, new listings should also exhibit heterogeneity, which we do not find.

[4] Table A1 in the Appendix shows that locations with higher population density tend to have a higher share of telecommute jobs in the surrounding areas.

[5] Table A1 shows that locations with higher population density tend to be close to a larger number of restaurant amenities per capita.

[6] Table A1 shows that locations with higher population density tend to have higher pre-COVID-19 home prices.

[7] Besides access to (short commutes) to jobs and access to consumption amenities, dense locations are able to support public amenities such as parks, museums and easier access to events and nightlife.

single detached houses. Second, dense neighborhoods are likely to rely heavily on public transit (e.g., buses, rails, and subways). Given the confined space of public transport vehicles, people may want to avoid settling in locations where residents rely mostly on public transit.[8]

The empirical evidence supports all of the conjectures. We find that the pandemic lowers home sales more in neighborhoods with a greater share of telework-compatible jobs nearby, more consumption amenities, higher pre-pandemic home prices, and lower income levels. After holding these observables constant, we still find that housing demand declines more in locations with higher residual density, which suggests that home buyers may be concerned about density per se owing to the fear of viral transmission in crowded places. Interestingly, we do not find evidence that the impact of COVID-19 on housing demand varies by neighborhood racial compositions, controlling for local density.

A potential alternative explanation for the disproportionate drop in home sales near city centers and in dense neighborhoods is that these locations may have higher case rates, and thus be more prone to lock-down policies or have more cautious populations (Alexander and Karger, 2020). A greater decline in home sales may therefore reflect a freeze in owners' propensity to sell homes and buyers' propensity to buy homes due to COVID-19, which is not indicative of changes in the underlying location demand. If the likelihood of lock-downs is correlated with location characteristics such as distance to downtown and density, the correlation between the change in home sales and these location characteristics may simply reflect a "lock-down effect," not necessarily a change in the underlying demand for locations. In light of these potential concerns, we argue that our findings are likely to reflect changes in location demand for several reasons.

First, in our neighborhood-level analysis, we account for potentially different post-pandemic changes in home sales at the metropolitan statistical area (MSA) level, and therefore differences in the prevalence of COVID-19 across cities is not likely to affect our results. Second, it is still possible that different areas within the same MSA could experience different degrees of the lock-down effect due to different levels of the seriousness of the outbreak. We attempt to account for such variation by including county-level case rate as an additional location characteristic. Third, even though new

[8]Ideally, we would like to treat the point regarding the reliance on public transit as a separate conjecture and test it separately from density. As we attempt to analyze its role, we find that the public transit usage rate is extremely correlated with residential density at the neighborhood level. The residual variation in public transit reliance after controlling for density is too small to have enough power to identify its effect on housing demand.

listings have decreased in aggregate since the outbreak, changes in new listings do not vary by density or distance to downtown. If there is substantial within-city lock-down variation, it could also affect new listings, since sellers might be less likely to list their homes during a lock-down. The absence of variation in new listings by density and distance to downtown implies that the effect seen in sales is unlikely driven by the lock-down effect.

Finally, it is possible that the lower demand for dense and expensive locations is only a temporary (not a long-run) effect. As housing demand recovers in aggregate when the economy reopens, the demand for density could recover. So far in the data (by June), as the national aggregate home sales partially recover in June owing to the easing of the lock-downs, we do not see home sales in dense locations recover faster than other locations, and thus the negative effect of the pandemic on the demand for density remains even as the aggregate demand recovers, suggesting the dwindled demand for density may persist. This implies that the disease transmission may persist as a dispersion force in the urban spatial equilibrium.

2 Data

We combine several data sources for the analysis. The main outcome variables include sales, new listings, median home sale prices, rental prices, and online viewings of properties, which are sourced from Redfin Data Center, Zillow Research, and Realtor.com. We obtain ZIP code-, county-, and MSA-level characteristics from the American Community Survey (ACS), the National Historical Geographic Information System (NHGIS), and the Zip Code Business Patterns (ZCBP).

2.1 Sales, Listings, and Prices

We obtain information on home sales, new listings, and median home sale prices at monthly frequency at the ZIP code, county, MSA, and national levels from Redfin Data Center from year 2016 to June 2020. Redfin is an online real estate brokerage firm, which has direct access to local multiple listing services (MLS). Through the local MLS listings, Redfin provides information on all broker-listed homes.[9]

It is noteworthy that ZIP code-level variables are constructed with pooled three-month lagged

[9] The data include homes in foreclosure. Homes in pre-foreclosure, properties on sale outside of MLS (e.g., Craigslist), and commercial properties are excluded from the data.

data. For example, a ZIP code-level number of home sales in January, 2020 includes total home sales recorded from November 1, 2019 to January 31, 2020. Median sale price is the median price of homes sold in the three-month period. Variables constructed at the county level or above are one-month values only.

2.2 Online Views of Homes

Besides real sales and listings information, we use the online views of properties on Realtor.com as an alternative tracker of housing demand. Realtor.com is the largest real estate listing website in the U.S. They provide the ratio between the number of website views on the average property in each ZIP code and the number of website views on the average property nationwide at a monthly frequency between 2017 and June 2020. We use this monthly reported value as an alternative measurement for the underlying demand for housing in each ZIP code location.[10]

2.3 Rental Prices

We obtain monthly rental price data from the Zillow Observed Rent Index (ZORI) released by Zillow Research. The variable is a smoothed measure of the typical observed market rents in a given region. ZORI is a repeat-rent index that is weighted to the rental housing stock to ensure representativeness across the entire market, not just those homes currently listed for rent.[11] We use the monthly rent data from 2016 to June 2020.

2.4 Local Characteristics

We obtain local characteristics such as population density, income level, racial composition, public transit usage intensity from the summary tables of the 2013-2017 ACS through the NHGIS (Manson et al., 2020). The data come at the ZIP code, county, and MSA levels.

For each ZIP code, we calculate the Euclidean distance to the closest downtown. We geocode all the downtowns using the output of Holian and Kahn (2015).

[10]The advantage of using online views is that it is not subject the lock-down effect of the pandemic, since online views of properties are conducted virtually. The drawback of this measure is that people might simply browse properties without seriously considering purchasing a house.

[11]The index is dollar-denominated by computing the mean of listed rents that fall into the 40th to 60th percentile range for all homes and apartments in a given region, which is once again weighted to reflect the rental housing stock.

2.5 Telework-Compatibility

We compute the share of jobs that are telework-compatible for each ZIP code based on the spatial distribution of occupations, and an assignment of telework-compatibility for each occupation. Dingel and Neiman (2020) and Su (2020) use O*NET occupation characteristics to evaluate each occupation's suitability for telework, and assign a telework indicator to each occupation. We use the telework indicator developed by Dingel and Neiman for the main analysis.[12]

We use data from the 2016 ZCBP for local job distributions. The ZCBP comes at the NAICS level. We use the industry to occupation crosswalk to impute the local job distribution for each occupation. Based on the spatial job distribution at the ZIP code level and each occupation's telework-compatibility, we calculate the share of jobs within a 3-mile radius of each ZIP code that are telework-compatible. We use the ZIP Code Tabulation Area (ZCTA) Distance Database from the NBER website for distance measure between ZIP codes.[13]

We obtain MSA-level occupation compositions using micro-data from the 2013-2017 ACS obtained from the IPUMS (Ruggles et al., 2020). Specifically, for each MSA, we estimate the share of full-time workers aged from 25–69 in each occupation. Combined with the telework indicator for each occupation, we estimate the share of workers who are in telework-compatible occupations for each MSA.

2.6 Amenities (Restaurants)

We use the the 2016 ZCBP to estimate the number of restaurants by ZIP code in 2016. Using the similar method with which we construct the spatial profile of jobs, we calculate the number of restaurants (establishment with NAICS code 7225xx) within a 3-mile radius of each ZIP code. We then calculate the per capita number of restaurants (i.e., the number of restaurants within a 3-mile radius divided by the population of the ZIP code).

2.7 County-Level Case Rate of COVID-19

We download county-level case rates of COVID-19 from the Opportunity Insights Economic Tracker, which provides the source data for Chetty et al. (2020).[14]

[12] In Table A2, we show the results using Su (2020)'s calculation. The estimation results are almost identical.
[13] https://data.nber.org/data/zip-code-distance-database.html.
[14] https://opportunityinsights.org.

3 Empirical Analysis and Results

3.1 The Timing of the COVID-19 Pandemic in the U.S.

The COVID-19 cases started to spike in the latter half of March 2020 in the U.S.[15] Studies have also documented that physical movements and hours worked were dramatically cut back starting around the third week of March 2020, using data on people's real-time movements and business establishment activities (Atkinson et al., 2020; Kurmann et al., 2020).[16] Although the virus has started to spread in many other parts of the world before mid-March, the decline in people's physical movements and hours worked was small in the U.S.

Since our housing data come at monthly frequency, we define April, May and June, 2020 as the period after the outbreak of COVID-19. The key assumption of our empirical analysis is that the observed housing market outcomes should have been the same as they were in these respective months in previous years, based on earlier data. The differences between the observed outcomes in April, May and June in 2020 and these months in previous years reveal the impact of COVID-19 on the housing market. We describe the empirical approach and its identification assumption with more details below.

3.2 The Aggregate Impact of COVID-19 on the U.S. Housing Market

We start our analysis with the aggregate impact of the pandemic on the housing market. Figure 1 presents the log number of homes sold, log number of new listings, log inventory, log number of days on the market, log median sale price, and log rental price at the national level by month in 2019 and 2020. The number of sales, new listings, and inventory all dropped dramatically in April 2020 relative to 2019, although sales and new listings started to pick up in May and June. Such wild swings are likely to reflect the sudden lock-down effects at the beginning of the pandemic and the effects of the subsequent re-openings on the housing markets. On average, sales declined around 26% and listings declined around 27% since the outbreak of COVID-19 in the U.S., compared to sales and listings from April to June in 2019. The aggregate housing and rental prices after the outbreak are not significantly different from the comparable months in 2019. The lack of effect on prices may

[15] https://coronavirus.jhu.edu/data/new-cases.

[16] Many technology firms post their data online, which confirm the sharp timing documented in the literature. One example is the Google Mobility Reports on the United States: https://www.google.com/covid19/mobility.

be because it takes longer for prices to adjust.

Next, we conduct our main analysis by drilling down into finer geography, and show the heterogeneous effects based on local characteristics like density.

3.3 Neighborhood-Level Analysis

We first discuss how we estimate the heterogeneity in the impact of the pandemic on the housing market across neighborhoods within MSAs. From here on, we interchangeably refer to MSAs as cities.

The empirical specification takes the following form:

$$\log(s_{ncmy}) = \beta_1 After_{my} \cdot x_{nc} + \beta_2 After_{my} \cdot \log(CaseRate_{nc}) + After_{my} \cdot \lambda_c + \pi_{my} + \delta_{nm} + \gamma_{ny} + \epsilon_{ncmy}, \tag{1}$$

where s_{ncmy} is a housing market outcome, such as homes sales in neighborhood n within MSA c in month m of year y.[17] $After_{my}$ is a dummy variable that indicates the outbreak of COVID-19, which takes 1 if $m \geq 4$ and $y = 2020$, and 0 otherwise. x_{nc} denotes some neighborhood characteristic of interest, such as density, distance to city center, etc., and we add additional characteristics in the regression equation interacted with $After_{my}$. We control for the county-level average case rate between April and June in 2020 of neighborhood n ($CaseRate_{nc}$), which could absorb some short-run effect of the pandemic. For instance, activities related to the housing market may be more likely to pause in areas with higher case rates because of lock-down policies or consumers' concerns. For that purpose, we also control for MSA-specific effects of the pandemic using $After_{my} \cdot \lambda_c$. If the effect of lock-downs varies by MSA, variation across neighborhoods over time within cities allows us to identify the effect of local characteristics on housing outcomes. π_{my} represents a time fixed effect, which absorbs common shocks in month m of year y faced by all housing markets in the country. δ_{nm} represents month-varying shocks to the housing market in neighborhood n, such as seasonal trends in housing activity.[18] γ_{ny} represents year-varying shocking to the neighborhood's housing market, such as local economic shocks. The coefficient of interest is β_1, which estimates differential effects of

[17] We add 1 to each outcome variable before taking log.

[18] Note that we allow the seasonal trends to differ potential by ZIP code.

the pandemic across neighborhoods of different characteristics.[19]

County-level analysis We first conduct the analysis at the county level. This is because Redfin Data Center provides monthly data at the county level—although ZIP code-level data are also monthly, the variables are constructed with pooled three-month lagged data. Table 1 presents the estimates of β_1. The dependent variable is log number of homes sold. The result in Column 1 suggests that the pandemic lowers home sales to a greater degree in counties with higher population density within a MSA. In Column 2, we further include interactions between (i) $After$ and log pre-COVID house price, (ii) $After$ and log average income, and (iii) $After$ and log share of whites. Including these controls attenuates the difference across counties by population density, but there are still significant differences. Moreover, holding population density constant, we find that home sales decline more in counties with higher pre-COVID house prices. Nevertheless, we do not find differences in the effect of the pandemic by income or racial compositions with the county-level analysis.

Decreased demand for density persists despite aggregate recovery in sales As mentioned in the introduction, one might speculate that such effects are short-term, and the reduced demand for density may rebound in the longer term as the housing market recovers. We address this by examining whether the reduced demand for density persists into June 2020. Figure 1 shows that national home sales have partially recovered to levels in 2019 by June. Specifically, we re-run the county-level regression by excluding observations in April and May, 2020. The results are presented in Columns 3–4 of Table 1. We find that the estimates of the coefficients on $After$ $\times$ log density and $After$ $\times$ log pre-pandemic house price even become slightly larger in magnitude and remain statistically significant. In other words, although the demand for housing seems to be recovering in aggregate since June, we do not find that home sales in counties with high population density recover faster. The results suggest that changes in the demand for density may persist.

ZIP code-level analysis: Unpacking the reasons for the change in location demand To explore the potential mechanisms that drive the differential sale reductions across neighborhoods, we move to ZIP code-level analysis. The upside of the ZIP code-level analysis is that we can examine to

[19]We weight observations by neighborhood population.

what extent the differential effects of the pandemic on home sales across neighborhoods are driven by spatial variation in the access to jobs, telework-compatibility, and consumption amenities. A potential downside of moving to the ZIP code level is that variables on home sales, new listings, inventory, and sale prices at the ZIP code level are constructed with pooled three-month lagged data. It may potentially reduce the magnitude of our estimates.

We estimate Equation 1 at the ZIP code level and the estimates of β_1 are presented in Table 2. The dependent variable in Columns 1–4 is log number of homes sold.[20] The result in Column 1 suggests that the negative effect of the pandemic on home sales is greater in neighborhoods closer to the central city with higher population density. It is possible that the differences are driven by local lockdown policies—higher density areas are more likely to have experienced lockdowns due to the pandemic. To mitigate the concern, as in county-level analysis, we control for county-level case rates and MSA-specific effects of the pandemic.

In Column 2 of Table 2, we add the interactions between $After_{my}$ and (i) log number of jobs within 3 miles of a ZIP code over the population of the ZIP code, (ii) log share of telecommute jobs, and (iii) log number of restaurants within 3 miles of a ZIP code over the population of the ZIP code. Including these terms slightly attenuates the difference in effect of the pandemic on sales by density, suggesting that some of the reduced demand for density could be attributed to the reduction in people's demand for living close to work and amenities. Indeed, the estimated coefficients suggest that the negative effect of the pandemic on home sales is greater in neighborhoods with more telecommute jobs and restaurants nearby.

In Column 3, we investigate the role of pre-COVID house price and income level of a neighborhood. We find that the negative effect of the pandemic is greater in neighborhoods with higher house prices. Holding price constant, the effect is greater in neighborhoods with lower income levels. Never-

[20] Figure A1 presents binned scatter plots of the effect of the pandemic on the log number of homes sold against several neighborhood characteristics by ZIP code, including (i) distance to the central city, (ii) population density, (iii) share of nearby jobs (within 3 miles of a ZIP code) that are telework compatible, (iv) number of restaurants per capita (as a measurement of consumption amenity), (v) pre-COVID house price, and (vi) income. We find greater home sale declines in neighborhoods closer to the central city, and with higher population density, higher shares of telecommute jobs, more restaurants, higher pre-COVID house price, and lower income level. We obtain ZIP code-level effect of the pandemic on log sales by estimating β_{zc} in the following regression:

$$\log(s_{zcmy}) = \beta_{zc} After_{my} + \lambda_c + \pi_{my} + \delta_{zm} + \gamma_{zy} + \epsilon_{zcmy},$$

where all the variables are defined as in Equation 1. To construct the figures, we divide the x variable that measures a neighborhood characteristic into 20 bins, and plot the mean values of x and y variables within each bin, controlling for the average case rate of each neighborhood between April and June, 2020.

theless, controlling for house price and income level, we do not find differences across neighborhoods by racial composition.

In Column 4, we include all the observable location characteristics, and we find statistically significant estimates on all the interaction terms. Including all the control variables attenuates the difference across neighborhoods by distance to the central city. However, differences across neighborhoods by other characteristics remain large.[21]

Lastly, we use log number of online views of properties relative to the national average as the outcome variable in Column 5. We obtain this variable from Realtor.com, which is a real estate listing website. The number of website views of properties could be an alternative measurement of housing demand. On the one hand, the number of views can be a better measure of housing demand since website views are not likely to be affected by the lock-down policies. On the other hand, the variable may introduce significant measurement error since more views do not necessarily lead to more purchases. Most of the estimates are not statistically significant, which could be because views are a rough measure of housing demand. Nevertheless, we find relative large and accurate estimates of the coefficients on interactions between $After$ and log share of telecommute jobs, $After$ and log pre-COVID house price, and $After$ and log income.

Table 3 presents differential effects of the pandemic on other housing market outcomes, including new listings, inventory, days on market, median sale prices, and rents. First, although the number of new listings has declined since the outbreak of COVID-19 (as is shown in Figure 1), we do not find differences across locations by distance to downtown or population density.[22] The result suggests that the negative aggregate effect on new listings could be driven by an aggregate lock-down effect. The lack of correlation between changes in new listings and local characteristics suggests that the lock-down effect is not likely to be correlated with local characteristics, such as density and distance to downtown. Therefore, the heterogeneous effects of the pandemic on home sales shown in Table 2 are likely to reflect changes in people's demand for residence locations, not a lock-down effect. The result in Column 2 shows that the effect on inventory is lower in suburbs and less dense neighborhoods, which is consistent with the findings in Table 2 and Column 1 of Table 3. The result in Column 3 shows that the days on market are relatively longer for houses in dense neighborhood,

[21]The reliance of public transit usage can be another relevant neighborhood characteristic to include. However, public transit usage is extremely correlated with population density, the residual variation does not carry enough statistical power.

[22]We do not find differences across locations by other characteristics considered in Table 2 either.

which also suggests that the demand for dense neighborhoods might have declined more because of the pandemic. Lastly, we do not find discernible differences in the impact of the pandemic on house or rental prices across neighborhoods (Columns 4 and 5). This could be because it takes longer for prices to adjust to the underlying demand.[23]

3.4 MSA-Level Analysis

The results in the previous section suggest that the pandemic could have changed people's demand for residential locations within cities. Next, we examine whether the pandemic has affected the demand for certain characteristics of cities.

Specifically, we use the following empirical specification:

$$\begin{aligned} \log(s_{cmy}) = {} & \alpha_1 After_{my} \cdot x_c + \alpha_2 After_{my} \cdot \log(CaseRate_c) \\ & + \pi_{my} + \delta_{cm} + \gamma_{cy} + \epsilon_{cmy}, \end{aligned} \tag{2}$$

where s_{cmy} is a housing market outcome of MSA c in month m of year y; x_c is a city characteristic; $CaseRate_c$ is the average case rate of city c between April and June, 2020; δ_{cm} is a city $\times$ month fixed effect; γ_{cy} is a city $\times$ year fixed effect. Other variables are defined as in Equation 1.[24] The coefficient of interest is α_1, which estimates differential effects of the pandemic across cities of different characteristics. Identification is based on variation across cities over time.

Table 4 presents the estimates of α_1. The results in Columns 1-2 suggest that, similar to differences across neighborhoods, home sales decline relatively more in cities with higher pre-COVID house prices and lower income levels. Nevertheless, we do not find differences across cities by the share of workers in occupations that are likely to be teleworking-compatible and by the number of restaurants per capita. Interestingly, we find that new listings decrease less in cities with a larger share of workers in occupations that are teleworking-compatible. The effect is even larger after controlling for pre-COVID house price and income level of the city. The result suggests that it is possible that workers with jobs that can be done remotely may choose to sell their houses and move

[23]Figure A2 in the Appendix shows the effects of the pandemic on home sales and new listings for neighborhoods in different quartiles of density, pre-pandemic house price, share of telecommute jobs, number of restaurants per capita. The results confirm the previous findings that there is heterogeneity in the effect on sales across neighborhoods, but not in the effect on new listings.

[24]We weight observations by MSA population.

away from the city. We also find that house prices also drop more in cities where more workers are in telecommute jobs. Nevertheless, the effect on rental prices is smaller in such cities.

4 Conclusion

We study the impact of the COVID-19 pandemic on the demand for cities and neighborhoods, using highly localized data on the U.S. housing market. We show that the pandemic disproportionately lowers the demand for housing in cities and neighborhoods with higher population density. We further show that the decreased demand for density is partially driven by the diminished need for living near jobs that are telework-compatible owing to the increasing prevalence of telework due to the pandemic. The decreased demand for density is also partially driven by the dwindling attraction of consumption amenities (e.g., restaurants), thanks to the need for social distancing. We also show that cities and neighborhoods with higher pre-COVID-19 house prices see a greater drop in housing demand. Holding access to jobs, amenities, income, and initial house price constant, we show that neighborhoods with higher density still see a greater drop in housing demand, which suggests that population density per se may become a concern for home buyers.

The findings of the study suggest that in the wake of the sudden outbreak and continuing spread of the COVID-19 pandemic, contagious diseases emerge as an important dispersion force in the urban spatial equilibrium, at least in the short run. As the aggregate housing market partially recovers in June, housing demand in dense locations does not recover faster than other places, suggesting that the negative effect of the pandemic on the demand for dense locations persists even after the aggregate housing market recovers. This suggests that disease transmission as a dispersion mechanism may continue to be important in the long run.

References

Ahlfeldt, G., Redding, S., Sturm, D., and Wolf, N. (2015). The economics of density: Evidence from the Berlin wall. *Econometrica*, 83(6):2127–2189. DOI: `https://doi.org/10.3982/ECTA10876`.

Alexander, D. and Karger, E. (2020). Do stay-at-home orders cause people to stay at home? Effects of stay-at-home orders on consumer behavior. *Federal Reserve Bank of Chicago Working Paper No. 2020-12*, 20(138):138. DOI: `https://doi.org/10.21033/wp-2020-12`.

Allcott, H., Boxell, L., Conway, J., Ferguson, B., Gentzkow, M., and Goldman, B. (2020). Economic and health impacts of social distancing policies during the Coronavirus pandemic. *Working Paper*. DOI: `https://dx.doi.org/10.2139/ssrn.3610422`.

Atkinson, T., Dolmas, J., Christoffer, K., Koenig, E., Mertens, K., Murphy, A., and Yi, K.-M. (2020). Mobility and engagement following the SARS-Cov-2 outbreak. *Working Paper*. DOI: `https://doi.org/10.24149/wp2014`.

Autor, D. and Reynolds, E. (2020). The nature of work after the COVID crisis: Too few low-wage jobs.

Bartik, A., Cullen, Z., Glaeser, E., Luca, M., and Stanton, C. (2020). What jobs are being done at home during the COVID-19 crisis? Evidence from firm-level surveys. *Harvard Business School Working Paper*, 20:138. DOI: `https://dx.doi.org/10.2139/ssrn.3574741`.

Bick, A., Blandin, A., and Mertens, K. (2020). Work from home after the COVID-19 outbreak. *Federal Reserve Bank of Dallas Working Paper No. 2017*. DOI: `https://doi.org/10.24149/wp2017`.

Brueckner, J. (2011). *Lecture on Urban Economics*. MIT Press.

Chen, H., Qian, W., and Wen, Q. (2020). The impact of the COVID-19 pandemic on consumption: Learning from high frequency transaction data. *Working paper*. DOI: `https://dx.doi.org/10.2139/ssrn.3568574`.

Chetty, R., Friedman, J., Hendren, N., Stepner, M., and the Opportunity Insights Team (2020). How did COVID-19 and stabilization policies affect spending and employment? A new real-time economic tracker based on private sector data. *Working Paper*. DOI: `https://doi.org/10.3386/w27431`.

Cox, N., Ganong, P., Noel, P., Vavra, J., Wong, A., Farrell, D., and Greig, F. (2020). Initial impacts of the pandemic on consumer behavior: Evidence from linked income, spending, and savings data. *Brookings Papers on Economic Activity, Conference Draft*.

Dingel, J. and Neiman, B. (2020). How many jobs can be done at home? *Journal of Public Economics*, 189. DOI: `https://doi.org/10.1016/j.jpubeco.2020.104235`.

Glaeser, E. (2011). *Triumph of the city: How our greatest invention makes us richer, smarter, greener, healthier, and happier*. New York: Penguin Press.

Glaeser, E. and Gotllieb, J. (2000). The wealth of cities: Agglomeration economies and spatial equilibrium in the united states. *Journal of Economic Literature*, 47(4):983–1028.

Glaeser, E., Kolko, J., and Saiz, A. (2000). Consumer city. *Journal of Economic Geography*, 1(1):27–50. DOI: `https://doi.org/10.1093/jeg/1.1.27`.

Holian, M. and Kahn, M. (2015). Household carbon emissions from driving and center city quality of life. *Ecological Economics*, 116:362–368. DOI: https://doi.org/10.1016/j.ecolecon.2015.05.012.

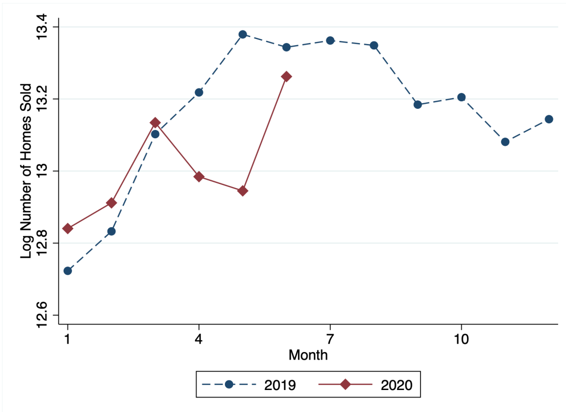
Kurmann, A., Lale, E., and Ta, L. a. (2020). The impact of COVID-19 on small business employment and hours: Real-time estimates with homebase data. *Working Paper*.

Manson, S., Schroeder, J., Van Riper, D., and Ruggles, S. (2020). IPUMS National Historical Geographic Information System: Version 14.0 [Database]. Minneapolis, MN: IPUMS. 2019. DOI: http://doi.org/10.18128/D050.V14.0.

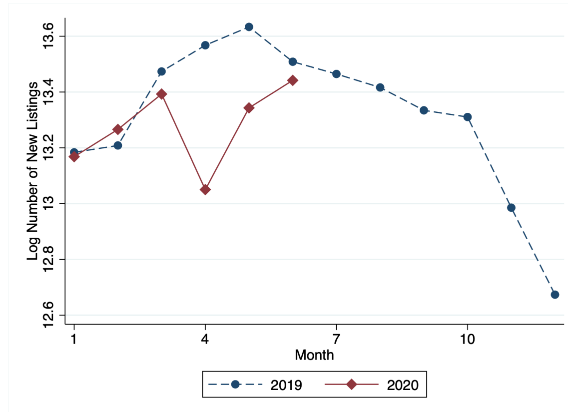
Mas, A. and Pallais, A. (2020). Alternative work arrangements. *Annual Review of Economics. Forthcoming.* DOI: http://doi.org/10.1257/aer.20161500.

Ruggles, S., Flood, S., Goeken, R., Grover, J., Meyer, E., Pacas, J., and Sobek, M. (2020). IPUMS USA: Version 10.0 [dataset]. Minneapolis, MN: IPUMS, 2020. DOI: https://doi.org/10.18128/D010.V10.0.

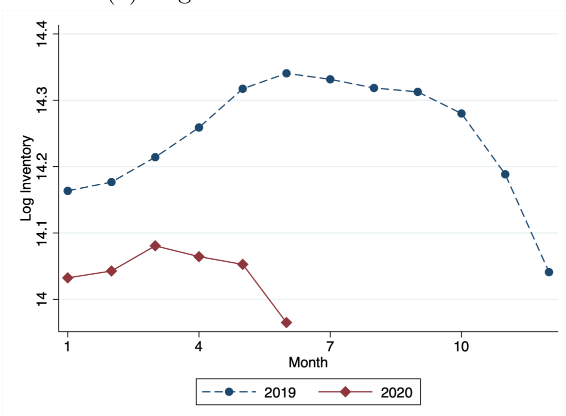
Su, Y. (2020). Working from home during a pandemic: It's not for everyone. *https://www.dallasfed.org/research/economics/2020/0407*.



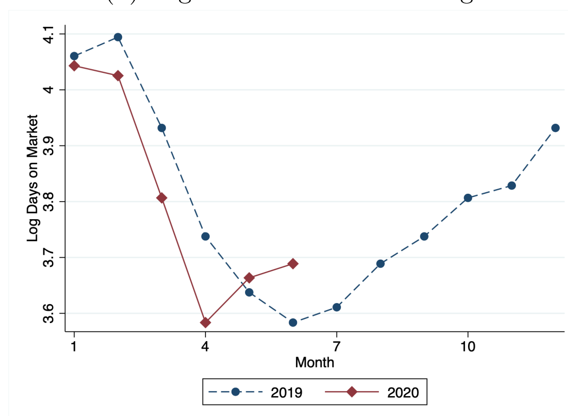
(a) Log Number of Homes Sold



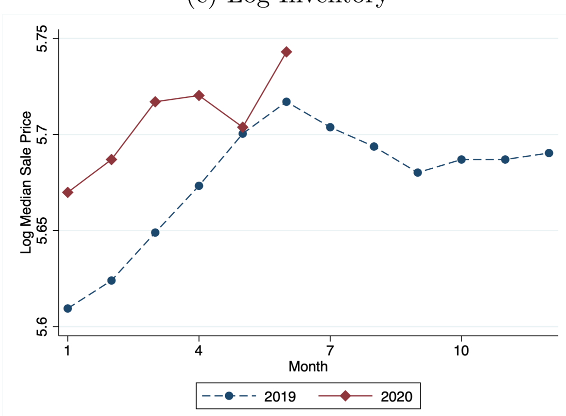
(b) Log Number of New Listings



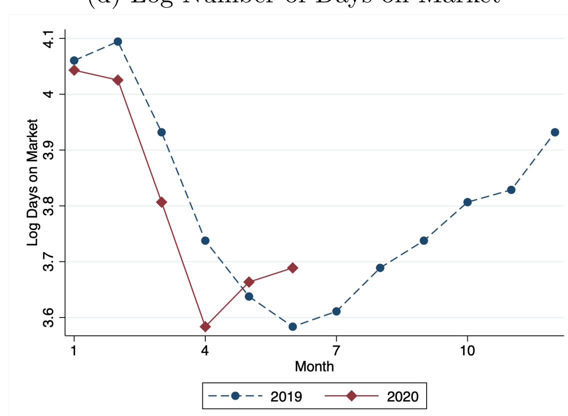
(c) Log Inventory



(d) Log Number of Days on Market



(e) Log Median Sale Price



(f) Log Average Rental Price

Figure 1: Changes in the Housing Market: 2019 vs. 2020

Note: The figures present the national housing market outcomes by month for 2019 and 2020. Panels a, b, c, d, and e use national monthly data on home sales, new listings, inventory, days on market, and median sale price from Redfin Data Center. Panel f uses monthly rental price data from Zillow Research at the ZIP code level. We use the national average rents, weighted by ZIP code population.

Table 1: Heterogeneous Effects of the COVID-19 Pandemic across Counties

	Log (Sales)			
	Full Sample		Excl. April & May, 2020	
	(1)	(2)	(3)	(4)
After $\times$ Log (Density)	-0.0387***	-0.0248***	-0.0372***	-0.0251***
	(0.00445)	(0.00529)	(0.00627)	(0.00822)
After $\times$ Log (Pre-COVID House Price)		-0.0906*		-0.116***
		(0.0512)		(0.0363)
After $\times$ Log (Income)		0.0188		-0.0504
		(0.0619)		(0.0768)
After $\times$ Log (Share of Whites)		0.0171		0.00844
		(0.0333)		(0.0186)
Observations	62,856	62,478	60,244	59,882

Note: The sample comprises all counties between January 2016 and June 2020 in Columns 1–2. Columns 3–4 exclude April and May, 2020. The dependent variable is log number of homes sold. $After$ is a dummy variable that is equal to 1 if the observation is between April and June, 2020, and 0 otherwise. All specifications include year $\times$ month, $After$ $\times$ MSA, county $\times$ year, and county $\times$ month fixed effects, and $After$ $\times$ log average case rate. Observations are weighted by the county's population. Standard errors are clustered at the MSA level: *** $p < 0.01$, ** $p < 0.05$, *$p < 0.1$.

Table 2: Heterogeneous Effects of the COVID-19 Pandemic across ZIP Codes

	Log (Sales)				Log (Views)
	(1)	(2)	(3)	(4)	(5)
After × Log (Distance to Downtown)	0.0195***	0.0112***	0.0149***	0.00867*	-0.00106
	(0.00492)	(0.00412)	(0.00521)	(0.00457)	(0.00364)
After × Log (Density)	-0.0168***	-0.0123**	-0.0158***	-0.0120***	-0.00109
	(0.00468)	(0.00548)	(0.00318)	(0.00410)	(0.00196)
After × Log (Jobs per capita)		0.0142**		0.0118**	-0.00408
		(0.00611)		(0.00559)	(0.00479)
After × Log (Share of Telecommute Jobs)		-0.0589***		-0.0504***	-0.00986
		(0.0140)		(0.0133)	(0.00782)
After × Log (Restaurants per capita)		-0.0226***		-0.0178***	0.000461
		(0.00561)		(0.00502)	(0.00489)
After × Log (Pre-COVID House Price)			-0.0424***	-0.0345***	-0.0129
			(0.00766)	(0.00713)	(0.00856)
After × Log (Income)			0.0347***	0.0318***	0.0139*
			(0.0110)	(0.0116)	(0.00775)
After × Log (Share of Whites)			-0.00512	-0.00289	-0.00449
			(0.00502)	(0.00495)	(0.00556)
Observations	529,902	499,392	509,814	484,650	447,113

Note: The sample comprises all ZIP codes between January 2016 and June 2020. The dependent variable in Columns 1–4 is log number of homes sold. The dependent variable in Column 5 is log number of views relative to the national average. $After$ is a dummy variable that is equal to 1 if the observation is between April and June, 2020, and 0 otherwise. All specifications include year × month, MSA × $After$, ZIP code × year, and ZIP code × month fixed effects, and $After$ × log average case rate. Observations are weighted by the ZIP code's population. Standard errors are clustered at the MSA level: *** $p < 0.01$, ** $p < 0.05$, *$p < 0.1$.

Table 3: Heterogeneous Effects of the COVID-19 Pandemic Across ZIP Codes (Additional Outcomes)

	Log				
	New Listings	Inventory	Days on Mkt	Prices	Rents
	(1)	(2)	(3)	(4)	(5)
After × Log (Distance to Downtown)	-0.000299	-0.0189**	0.0126	-0.00385	0.000364
	(0.00430)	(0.00763)	(0.0112)	(0.00577)	(0.000260)
After × Log (Density)	-0.0131	0.0106**	0.0131*	-0.000251	-4.62e-05
	(0.00841)	(0.00462)	(0.00781)	(0.00136)	(0.000219)
Observations	529,902	529,902	494,127	495,103	146,580

Note: The sample comprises all ZIP codes between January 2016 and June 2020. The dependent variables are log number of new listings, log inventory, log number of days on market, log median sale price, and log rental price. $After$ is a dummy variable that is equal to 1 if the observation is between April and June, 2020, and 0 otherwise. All specifications include year × month, MSA × $After$, ZIP code × year, and ZIP code × month fixed effects, and $After$ × log average case rate. Observations are weighted by the ZIP code's population. Standard errors are clustered at the MSA level: *** $p < 0.01$, ** $p < 0.05$, *$p < 0.1$.

Table 4: Heterogeneous Effects of the COVID-19 Pandemic Across Metropolitan Areas

	Log (Sales)		Log (New Listings)		Log (Prices)		Log (Rents)	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
After × Log (Share of Teleworkers)	0.111	0.0448	0.112*	0.282***	-0.0254	-0.0497**	0.0183*	0.0321**
	(0.108)	(0.118)	(0.0657)	(0.102)	(0.0164)	(0.0208)	(0.0104)	(0.0124)
After × Log (Restaurants per capita)	-0.0393	0.0414	-0.103**	-0.0542	-0.00498	-0.00226	-0.0144**	-0.00762
	(0.120)	(0.0690)	(0.0476)	(0.0493)	(0.0124)	(0.0152)	(0.00570)	(0.00597)
After × Log (Pre-COVID House Price)		-0.184***		-0.00933		-0.0169		-0.00643*
		(0.0425)		(0.0356)		(0.0108)		(0.00355)
After × Log (Income)		0.271**		-0.147		0.0438		-0.00735
		(0.116)		(0.101)		(0.0279)		(0.0106)
After × Log (Share of Whites)		0.00819		-0.0708*		0.00396		-0.000665
		(0.0292)		(0.0411)		(0.00643)		(0.00430)
Observations	8,910	8,856	8,910	8,856	8,910	8,856	4,716	4,662

Note: The sample comprises all metropolitan statistical areas (MSAs) between January 2016 and June 2020. The dependent variables are log number of homes sold, log number of new listings, log median sale price, and log rental price. $After$ is a dummy variable that is equal to 1 if the observation is between April and June, 2020, and 0 otherwise. All specifications include year × month, MSA × year, and MSA × month fixed effects, and $After$ × log average case rate. Observations are weighted by the MSA's population. Standard errors are clustered at the MSA level: *** $p < 0.01$, ** $p < 0.05$, *$p < 0.1$.

Appendix

A1 Telework-Compatibility by Su (2020)

In the main analysis, we use the telework-compatibility indicator developed by Dingel and Neiman (2020). As a robustness check, we use an alternative telework indicator developed by Su (2020). Su uses a similar method to select telework-compatible occupations as Dingel and Neiman, with a slightly different and simpler set of criteria.

Specifically, Su assigns each occupation as either telework-compatible or not telework-compatible, based on five work context indices provided by O*NET. An occupation is remote-compatible if five criteria are all met:

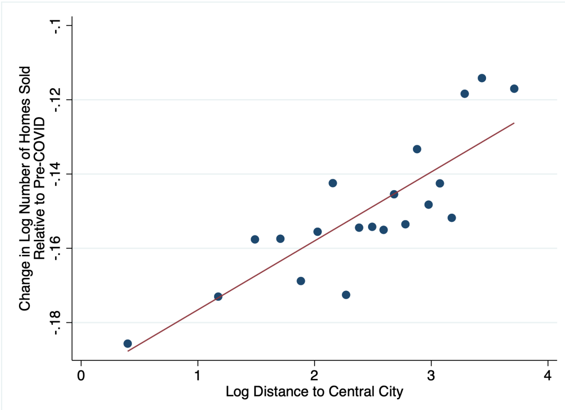
1. Work involves frequent use of email;
2. Work does not require physical proximity with other people closer than arm's length.
3. Work involves sitting at least half of the time.
4. Work does not involve significant kneeling, crouching, stooping or crawling.
5. Work does not involve significant bending, or twisting of the body.

The detailed selection criteria are listed as follows:

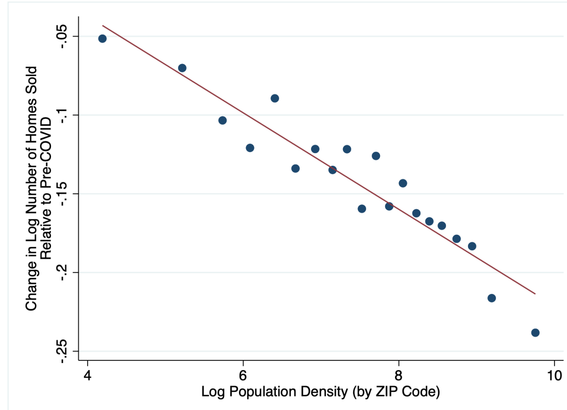
1. Work context variable "Electronic Mail" ≥ 87.5. According to the scale of the index, an index of 75 means using email at least once a week and not every day. An index of 100 means using email every day. "Frequent use of email" is likely close to every day. However, since O*NET is estimated statistically from national surveys, Su takes an average between 75 and 100 as the cutoff value to allow some room for statistical error.
2. Work context variable "Physical Proximity" ≥ 75. An index of 75 means physical proximity of an arm's length.
3. Work context variable "Spend Time Kneeling, Crouching, Stooping, or Crawling" < 50.
4. Work context variable "Spend Time Bending or Twisting the Body" < 50.

The occupation code used is *occ2010* defined in the IPUMS USA data. O*NET occupation codes are linked to *occ2010* with a SOC-occ2010 crosswalk.

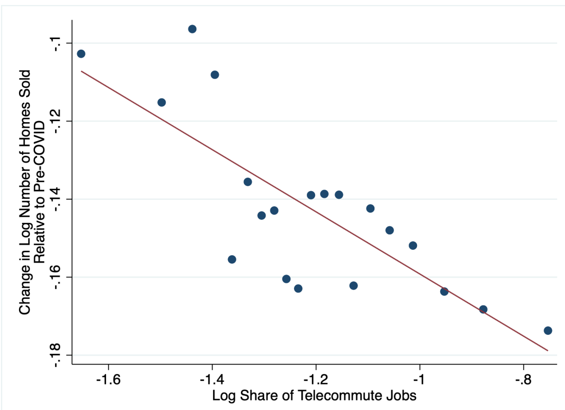
The regression results using Su is shown in Table A2. The magnitude of the coefficient on log share of telecommute jobs does not vary much by either definition of telework compatibility.



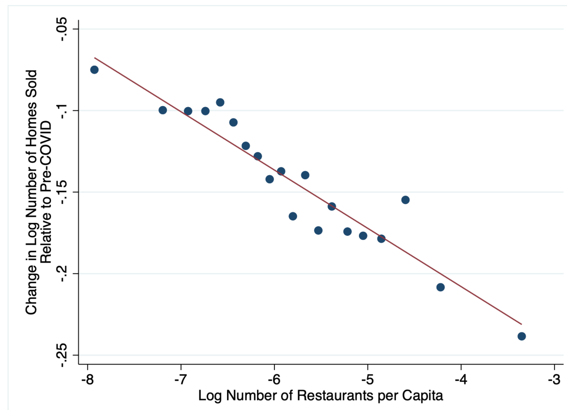
(a) Distance to the Central City



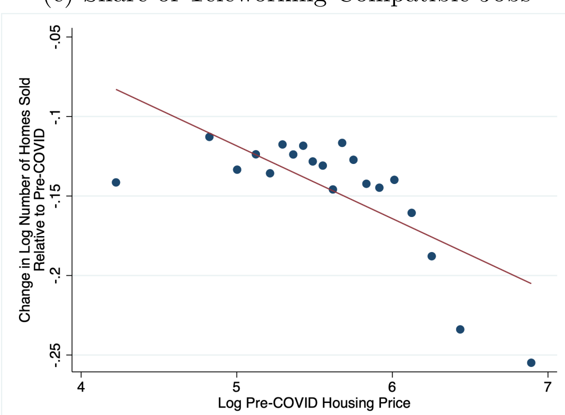
(b) Population Density



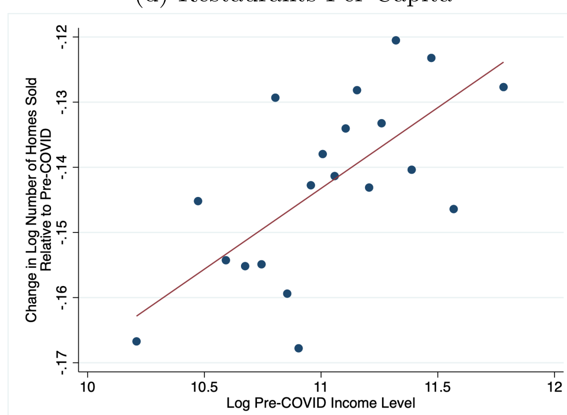
(c) Share of Teleworking-Compatible Jobs



(d) Restaurants Per Capita



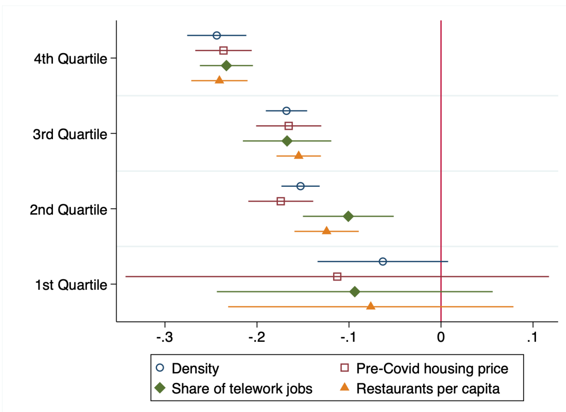
(e) Pre-COVID Housing Price



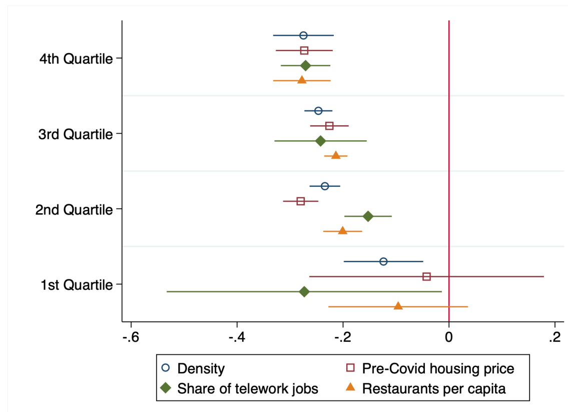
(f) Income

Figure A1: Changes in Log Home Sales (Relative to Pre-COVID) vs. Neighborhood Characteristics

Note: The figures present binned scatter plots of changes in log number of homes sold per month before and after the outbreak (y-axis) vs. neighborhood characteristics by ZIP code (x-axis). We obtain ZIP code-level changes in log sales, controlling for MSA, time, ZIP code $\times$ month, and ZIP code $\times$ year fixed effects. To construct the figures, we divide the x variable into 20 bins, and plot the mean values of x and y variables within each bin, controlling for county-level average case rate between April and June, 2020.



(a) Effect on Log Sales



(b) Effect on Log New Listings

Figure A2: Effects of the COVID-19 Pandemic on Home Sales and New Listings by Neighborhood Characteristics

Note: The figures present the estimates of β_1 of Equation 1. Specifically, we include four interaction terms $After_{my} \cdot x^n_{zc}$, $n = 1, ..., 4$, and x^1_{zc} is a dummy variable that indicates that the density of ZIP code z in city c is below the 1st quartile. x^2_{zc} is a dummy variable that indicates that the density of ZIP code z is between the 2nd and 3rd quartile. Similarly for x^3_{zc} and x^4_{zc}. We also separately consider the share of telecommute jobs, pre-COVID house price, and the number of restaurants per capita. The dependent variable is log number of homes sold in Panel a and log number of new listings in Panel b.

Table A1: Relationship Between Population Density and Other Neighborhood Characteristics

	Log (Density)			
	(1)	(2)	(3)	(4)
Log (Distance to Downtown)	-0.673***	-0.732***	-0.849***	-0.379***
	(0.0204)	(0.0198)	(0.0185)	(0.0204)
Log (Transits per capita)	0.340***			0.216***
	(0.0144)			(0.0132)
Log (Jobs per capita)		-0.183***		-0.285***
		(0.0285)		(0.0289)
Log (Share of Telecommute Jobs)		0.246***		0.358***
		(0.0565)		(0.0571)
Log (Restaurants per capita)		0.502***		0.561***
		(0.0283)		(0.0291)
Log (Pre-COVID House Price)			0.432***	0.221***
			(0.0345)	(0.0306)
Log (Income)			-0.573***	-0.327***
			(0.0478)	(0.0464)
Log (Share of Whites)			-0.231***	-0.217***
			(0.0182)	(0.0161)
Observations	7,744	9,204	9,095	7,446

Note: The sample comprises all ZIP codes. All columns include a MSA fixed effect. Observations are weighted by the ZIP code's population. Robust standard errors in parentheses *** $p < 0.01$, ** $p < 0.05$, *$p < 0.1$.

Table A2: Robustness Check: Using Telework-Compatibility Data from Su (2020)

	Log (Sales)		Log (Views)
	(1)	(2)	(3)
After * Log (Distance to Downtown)	0.0114***	0.00878*	-0.00107
	(0.00415)	(0.00463)	(0.00366)
After * Log (Density)	-0.0126**	-0.0121***	-0.00110
	(0.00555)	(0.00409)	(0.00197)
After * Log (Jobs per capita)	0.0141**	0.0118**	-0.00383
	(0.00629)	(0.00588)	(0.00491)
After * Log (Share of Telecommute Jobs)	-0.0558***	-0.0483***	-0.0103
	(0.0133)	(0.0127)	(0.00758)
After * Log (Restaurants per capita)	-0.0227***	-0.0180***	0.000219
	(0.00564)	(0.00508)	(0.00498)
After * Log (Pre-COVID House Price)		-0.0353***	-0.0130
		(0.00722)	(0.00855)
After * Log (Income)		0.0322***	0.0140*
		(0.0117)	(0.00772)
After * Log (Share of Whites)		-0.00250	-0.00440
		(0.00498)	(0.00555)
Observations	499,392	484,650	355,376

Note: The sample comprises all ZIP codes between January 2016 and June 2020. The dependent variable is log number of homes sold. $After$ is a dummy variable that is equal to 1 if the observation is between April and June, 2020, and 0 otherwise. All specifications include year $\times$ month, $After$ $\times$ MSA, county $\times$ year, and county $\times$ month fixed effects, and $After$ $\times$ log average case rate. The number of jobs per capita and the share of jobs within 3 miles of a ZIP code are estimated using data from Su (2020). Observations are weighted by the ZIP code's population. Standard errors are clustered at the MSA level: *** $p < 0.01$, ** $p < 0.05$, *$p < 0.1$.