The Exodus from New York City during COVID-19: Evidence from Out-of-Town Home Purchases

Running Title:

The Exodus from New York City during COVID-19

Abstract

Based on the universe of individual-level property transactions, we provide direct evidence on the population redistribution from densely populated metropolitan areas to nearby locations after the COVID-19 outbreak. We focus on out-of-town property purchases—a novel proxy for corresponding migration flows—by New York City (NYC) residents in New York state and five adjacent States. At the highest point after the pandemic, out-of-town property purchases in urban and non-urban areas by NYC homebuyers increase by 79% and 187% relative to pre-COVID19 levels, respectively. We also find natural amenities, high-speed internet access, and the percent of the population with post-graduate education become more important factors in NYC homebuyers' location decisions after COVID-19, while urban amenities' attraction decreases. In addition, we show that higher COVID death and infection rates have a significant deterrence effect on NYC homebuyers' reallocation decisions.

Keywords: Migration, COVID-19, Work-From-Home, Housing Market, Property Transactions

JEL Classification: R21; R23; R28; I15

1. Introduction

The wide adoption of work-from-home (WFH) practice after the COVID-19 outbreak is reshaping the spatial distribution of population and economic activities (Liu & Su, 2021; Rosenthal et al., 2021). As WFH unshackles educated workers from the requirements of working on-site, workers can now relocate to less populated regions, including rural areas and small towns, for lower housing prices, better housing conditions, and more desirable location characteristics (Althoff et al., 2021; Low et al., 2022). These movements occur within metro areas and between metro and nonmetro areas, which creates a new spatial equilibrium across locations (Brueckner et al., 2021; Delventhal et al., 2021). Initially intended as a short-term policy, WFH is now likely to stay even after the pandemic (Barrero et al., 2021). Researchers predict that 20% of the work hours will be supplied from home after the pandemic, a drastic increase from 5% before (Bick et al., 2021). Survey data shows that a substantial proportion of respondents, especially those living in major cities, have already moved or are planning to move due to WFH arrangements (Ozimek, 2020). Therefore, the migration caused by COVID-19 has long-term implications for labor markets, housing markets, and local government finance, as well as the political landscape (Davis et al., 2021).

This paper investigates two pressing questions in this historical migration process. First, what is the scale of the spatial redistribution? Second, where are people moving to? Traditional migration data lack sufficient spatiotemporal granularities to evaluate the impacts of COVID-19 in detail, and most survey-based migration data sources have not been updated to cover the COVID-19 period yet. In this study, however, we show that out-of-town property purchases can serve as an efficient and convenient proxy for migration flows. Using a dataset of more than 16.8 million individual-level housing transactions in the entire United States between 2014 and 2021, we provide evidence for these two questions by analyzing changes in property purchases and shifts in location preferences of residents from New York City (NYC), one of the most populous urban agglomerations in the world.

We first document a significant surge of single-family home (SFH) purchases by NYC homebuyers in non-NYC urban and non-urban areas within the State of New York

(NY) and five adjacent states, including Pennsylvania (PA), New Jersey (NJ), Massachusetts (MA), Connecticut (CT), and Vermont (VT), after the first COVID-19 case was identified in March 2020[1]. The six study states account for more than 90% of all transactions by NYC homebuyers in each month during the entire study period. Event-study results show that, at the peak times, quarterly out-of-town SFH purchases by NYC homebuyers within the six study states surge by 79% and 187% in urban and non-urban areas, respectively. With this significant and abrupt surge, by the third quarter of 2020, the share of purchases made by NYC homebuyers rise from 3.2% pre-COVID-19 to 4.1% of all transactions in non-NYC urban areas and from 5.7% to 11.0% in non-urban areas.

We then investigate the destination characteristics that attract homebuyers from NYC. Previous research has identified pull factors such as distance to urban areas, natural amenities, and population density that affect migration from urban areas (Rupasingha et al., 2015). The demographic composition of migrants may be different after the COVID-19 outbreak. Even for migrants from the same demographic groups, location preference may change due to the pandemic (Guglielminetti & Rondinelli, 2021). Our results show that higher natural amenities and high-speed internet access are important pull factors while urban amenities, measured by the density of retail and food service establishments, lose attraction after COVID-19. Furthermore, we find that higher COVID-19 death rates and infected case rates in the destination locations have economically and statistically significant deterrence effects for homebuyers from NYC. Placebo tests with artificial COVID-19 starting days show that the case and death rates have no effect before COVID-19, suggesting that the significant results are not due to pre-existing trends.

Our study fills important gaps in the nascent literature on COVID-19 induced migration based on individual-level property transaction data. While there is plenty of anecdotal evidence on COVID-19 induced migrations (Heather Kelly & Rachel Lerman, 2020; Hill, 2021), statistical evidence is scarce. Furthermore, existing evidence on the impact of the pandemic on residential location choices is almost exclusively based on the

[1] Because of the strong seasonality of the housing market, we use March 2020 as the cut-off point and select two commensurate time windows to define the pre- and post-COVID outbreak periods: the pre-COVID-19 period ranges from April 2018 to June 2019, while the post-COVID-19 period ranges from April 2020 to June 2021. Using two historical pandemic outbreaks, Francke and Korevaar (2021) find that pandemics cause large negative shocks on housing prices in the affected areas.

analysis of housing price, which is an indirect indicator of population movements[2]. To our knowledge, only two studies in progress directly investigate the flows of migration caused by COVID-19[3]. Using a dataset of interstate moves from three moving companies, Haslag and Weagley (2021) find that migration shifts toward smaller cities with lower living costs and fewer pandemic restrictions. However, their analysis excludes within-state migration, which far surpasses out-of-state moves historically (Molloy et al., 2011) and experiences much more change after COVID-19, as we demonstrate in this study. In addition, Ramani and Bloom (2021) use the US Postal Service National Change of Address dataset to analyze migration from large metropolitan areas in the United States. They focus on net migration from metropolitan areas and do not analyze the moving destinations.

This study contributes to the literature by providing direct evidence for COVID-19 induced migration based on the universe of individual-level property transactions. Furthermore, we highlight the importance of intermedium-range migration during COVID-19—migration within the state and adjacent states—which has been overlooked in previous studies. Finally, this study is the first to investigate location characteristics that attract COVID-19 induced migration.

Given the potentially large impact of migrants from population centers after the COVID-19 outbreak, local policymakers need to understand the factors that attract homebuyers from large urban centers. Our findings point to available policy levers: investments in transportation to decrease travel time to population centers, maintaining and improving natural amenities, improving internet infrastructure, and reducing the spread and impact of COVID-19. The fact that migrants emphasize urban amenities less after COVID-19 benefits small towns and rural areas that cannot support large retail and service sectors.

[2] For example, Liu and Su (2021) show that COVID-19 reduces housing demand in densely populated neighborhoods, driven by WFH and lower preference for urban amenities such as restaurants. Gupta et al. (2021) find that the pandemic decreases the housing price gradient from the city center to the suburbs. Brueckner, Kahn, and Lin (2021) demonstrate that high-productivity cities with higher WFH potential experience less housing price growth after COVID-19. D'Lima, Lopez, and Pradhan (2020) find that shutdown orders reduce property values. Francke and Korevaar (2021) leverage data from historical pandemics and show that housing prices experience transitory negative shocks during pandemics in affected areas.

[3] Couture et al. (2021) discuss quantifying movement and social contact by using cellphone data. Those data may work as indices for temporary movement and routine to grocery stores, restaurants, and recreational places, but more research is needed to utilize them to identify the actual migration flow due to WFH policy and COVID-19.

The remainder of the paper proceeds as follows. Section 2 introduces the data used in this study. Section 3 presents the econometric models. Section 4 discusses our empirical results. Finally, Section 5 concludes.

2. Data

2.1 Property Transaction

This study employs individual-level property transaction data from Zillow's Transaction and Assessment Database (ZTRAX)[4] which contains more than 400 million detailed transaction records across the United States. ZTRAX data has been increasingly used in academic studies (Leyk et al., 2020; Quesnel et al., 2020; Shen et al., 2021). Our analysis leverages the homebuyer information—the State and Zip Code where a buyer comes from—to identify and analyze potential migration flows from NYC to other urban or non-urban regions before and after COVID-19. Transaction data is first cleaned following the procedures recommended by Zillow[5] and Nolte et al. (2021) to identify arms-length sales, single-family houses, and other housing characteristics. The cleaned dataset includes 16,801,762 individual transactions of SFH in 50 states between January 2014 and June 2021. Each transaction contains information on various property characteristics (e.g., lot size, number of bedrooms, number of stories, fireplace, garage size, etc.), property coordinates, prior assessment valuations, and buyer and seller information. Most analysis focuses on the 102,106 SFH transactions during the study timeframe from 4,724 ZCTA areas in six study States, including New York, Pennsylvania, New Jersey, Massachusetts, Connecticut, and Vermont.

2.2 Migration Flow

[4] Data provided by Zillow through the Zillow Transaction and Assessment Dataset (ZTRAX). More information on accessing the data can be found at http://www.zillow.com/ztrax. The results and opinions are those of the author(s) and do not reflect the position of Zillow Group.

[5] Zillow provides an example code to illustrate a standard data cleaning process of creating a hedonic dataset. See https://github.com/zillow-research/ztrax for more details.

We evaluate how well can out-of-town property purchases be a proxy for migration. County-to-county migration flows are obtained from the American Community Survey (ACS) 5-year (2015-2019) migration data, which can be compared to the annual number of NYC homebuyers in the 140 counties of the study region from 2015 to 2019. State-to-state migration data for the six study states is collected from the ACS 1-year (2019) dataset. In Section 3.1, we test the correlation between out-of-town property purchases and corresponding migration flows at both the county and state levels.

2.3 COVID-19

We adopt the county-level COVID-19 data maintained by the *New York Times*[6]. The dataset includes daily cases and deaths reported in each county and state across the US since the beginning of the pandemic. We aggregate the dataset at the county-month level. The death and infected case rates are measured by deaths/cases per 1,000 population.

2.4 Location Characteristics

Data on location characteristics of zip codes and counties are collected from several sources. The boundary of NYC is defined by a collection of 169 zip codes from the Bronx, Manhattan, Brooklyn, and Staten Island. We employ the 2010 Rural-Urban Commuting Area Codes (RUCA) developed by the USDA Economic Research Service (ERS) and define Rural Area and Small Town as *Non-Urban*, while zip codes in micropolitan and metropolitan areas as *Non-NYC Urban* in the study area. We employ the Natural Amenities Rank developed by USDA ERS[7] to measure natural amenities. The ranking system is based on climate, topography, and water environment and spans from one (the lowest natural amenity) to seven (the highest natural amenity). Among all counties in NY, the highest natural amenity rank is four, while the lowest is two. In the entire study area, the highest amenity rank is five, and the lowest is two. We merge 126 zip code areas with rank five with those with rank four for ease of comparison. For high-speed internet access, we use the 2018 county-level data maintained by the Federal Communications Commission. The

[6] More details about the dataset can be found at https://github.com/nytimes/covid-19-data

[7] The NAS is available for counties in the lower 48 States. More details about NAS are available at https://www.ers.usda.gov/data-products/natural-amenities-scale/

data contains high-speed (over 200 kbps) internet connections per 1,000 households for each county. Urban amenities are measured by the density of retail establishments and food services. We collect the total number of retail establishments (defined according to the North American Industry Classification System (NAICS) 44 and 45) and restaurants (NAICS 7224 and 7225) from the County Business Pattern database published by the US Census Bureau. The number of establishments is normalized by 1,000 population. Furthermore, we use the ACS 5-year data of 2019 as the primary source for social and demographic statistics of each zip code area. Specifically, we compute the ratio of residents from different racial groups (i.e., black, white, and others), the percentage of residents with different education backgrounds (i.e., Bachelor degree, Master/Ph.D. degree, and others), and total population. We use the Zillow Home Price Index (ZHPI)[8] in regression analysis to control for housing prices.

The linkage between zip code-level and county-level datasets is achieved through the USPS ZIP Code Crosswalk data (HUD-USPS Crosswalk File) published by the US Department of Housing and Urban Development. A zip code that spreads across counties is assigned to the county with most of its residential addresses. We calculate the Euclidian distance between New York City and each zip code area in *QGIS* (v3.18 Zurich). Finally, in mapping and regression analyses, all variables at the zip code level are linked to ZCTAs using a crosswalk developed by the Uniform Data System.

3. Econometric Models

3.1 Correlation Between Out-of-Town SFH Transactions and Migration Flows

In this study, we use out-of-town SFH transactions as a proxy for migration flows. To examine the validity of this novel migration data, we test the correlation between SFH transactions and migration flows using historical data. Specifically, we summarize the annual average of SFH transactions made by NYC homebuyers between 2015-2019 in 140 non-NYC counties of the six study states. We then compare it to the migration flows from

[8] More information about the Zillow Home Price Index (ZHPI) is available on Zillow Research website (https://www.zillow.com/research/data/).

NYC to those 140 involved counties available from the ACS 5-year data for the corresponding years. The Pearson's product-moment correlation at the county level is $r_{county}(138) = 0.83, p < 0.001$. Considering that the ACS county-to-county migration data has substantial margins of error[9], SFH transactions seem to serve as a good proxy for migration flows from NYC to other areas in the study region.

We further examine the correlation between SFH transactions and migration flows across the six study states. We first summarize the total transactions by homebuyers from one state in all other five study states of the year 2019 and correlate the 6 x 5 flows with that from the ACS 2019 1-year state-to-state migration data. The Pearson's product-moment correlation between transactions and migration flows at the state level is $r_{state}(28) = 0.89, p < 0.001$. To account for factors, such as housing prices, property taxes, and income levels, in origin and destination states that can drive a wedge between migration and property purchase, we residualized SFH transactions and migration flows by origin and destination state fixed effects. The correlation with residualized variables is $r_{state,residual}(28) = 0.83, p < 0.001$.

3.2 Event Study

We apply an event study to separate the changes in transactions due to the COVID-19 outbreak from the typical seasonal fluctuations that existed before COVID-19. Fitting the seasonal dummies using the entire series would absorb some of the post-COVID-19 increase while overestimating the pre-COVID-19 seasonal fluctuations since the changes in transactions after COVID-19 would also follow a seasonal pattern. To address this challenge, we detrend the entire series using the seasonal effects estimated from the pre-COVID-19 data.

[9] The ACS county-to-county migration data are derived from responses to migration questions in the ACS questionnaire. The 5-year data are collected continuously over a 5-year period, which estimates where people lived when surveyed and where they lived 1 year prior during the survey periods. More details about the migration flow data can be found at ACS website: https://www.census.gov/data/developers/data-sets/acs-migration-flows.html

We first estimate the following Ordinary Least Squares (OLS) regression using data before the year 2020 to capture any pre-existing seasonal trends before the outbreak of COVID-19:

$$N_t = \alpha_0 + \sum_{n=Q2}^{Q4} \alpha_q QT_q \ + \epsilon_t \tag{1}$$

where N_t is the number of housing transactions in month t, QT_q represents three dummy variables for the second to the fourth quarter q of a year (i.e., Q2 ~ Q4), α_0 is the constant term, which can be interpreted as the average of first quarter transactions over the pre-COVID-19 years, and ϵ_t is the idiosyncratic error. We then apply the estimations of the above equation to detrend the entire time series of housing transactions (i.e., from 2018Q1 to 2021Q2) and obtain the detrended monthly total of housing transactions ($\widetilde{N}_t$). Lastly, we estimate the following event study equation:

$$\widetilde{N}_t = \beta_0 + \sum_{n=2018Q1\ (n\neq 2020Q1)}^{2020Q2} \beta_{qt}\delta_{qt} + u_t \tag{2}$$

where $\widetilde{N}_t$ is the detrended total of housing transactions in month t (normalized to the percentage of transactions in Q1 2020), δ_{qt} is a series of year-specific quarterly time-dummies ranging from the first quarter of 2018 to the second quarter of 2021 (the first quarter of 2020 is excluded as the base), β_0 is the constant term, and u_t is the error term. Estimations of β_{qt} together with the 95% confidence intervals are plotted in Figure 4 (Panel A-C and E-G) for different combinations of destination regions and homebuyers.

Based on the detrended monthly total of housing transactions ($\widetilde{N}_t$), we perform a Difference-in-Differences (DID) analysis to separate changes in the overall demand from the changes of demand from NYC homebuyers. The DID model is specified as the following:

$$\widetilde{N}_t = \beta_0 + \beta_1 NYC + \beta_2 \delta_{qt} + \gamma_{qt} NYC \cdot \delta_{qt} + u_t \tag{3}$$

where NYC is a dummy that equals one for homebuyers from New York City and equals zero for local home buyers and γ_{qt} is the DID estimator. Other terms are defined the same as in equation (3). Estimations for γ_{qt} are plotted together with the 95% confidence intervals in Figures 3D and 3H.

3.3 Cross-Sectional Regressions on Location Choices: Location Characteristics

For the analysis of the attractiveness of fixed location characteristics, we use cross-sectional regressions specified as:

$$\Delta N_i = \alpha_0 + \beta_i L_i + \varepsilon_i \tag{4}$$

where ΔN_i is the difference in the number of housing transactions in zip code area i at the two time windows, and L_i is a vector of location characteristics of zip code area i including natural amenity, internet access, the logarithmic form of distance to New York City, the density of retail business and food services establishments, and demographic statistics including the logarithmic form of the total population of zip code i , population ratios from different racial groups (i.e., white, black, and others), and population ratios with various education backgrounds (i.e., Bachelor degree, Master/Ph.D. degree, and others). Location characteristics also include the change in HPI (housing value index) between January 2018 and January 2020. In Figure 5, standardized coefficients are presented where the dependent variable and continuous independent variables (i.e. all except the natural amenity dummies) are normalized to a mean of zero and standard deviation of one. Regression results with unstandardized variables are included in Appendix Table B1.

3.4 Panel Regressions on Location Choices: Pandemic Impacts

The panel regression investigates the dynamic shocks of COVID-19 in each zip code area on the number of property transactions by NYC homebuyers. The dependent variable, ΔN_{it}, is the difference in the number of housing transactions in zip code i between a post-COVID month t and a corresponding month two years earlier before COVID-19 ($t-24$), measured in standard deviation. This differencing removes seasonal trends and time-invariant location characteristics. The independent variable of interest is COVID-19 death or infection rates per 1,000 population. We control for the housing price indices ($\Delta HPI_{i\tau}$) in the previous three months before transactions, also measured as the difference between a given post-COVID-19 month and two years earlier. Three lags of the COVID-19 variables are also included.

$$\Delta N_{it} = \sum_{\tau=t-3}^{t-1} \beta_\tau \Delta HPI_{i\tau} + \sum_{\tau=t-3}^{t} \varphi_\tau COVID_{i\tau} + \sigma_t + \rho_i + \varepsilon_{it} \quad (5)$$

We control for the month-year fixed effects σ_t and county fixed effects ρ_i. Falsification tests are carried out with achronological COVID-19 data at 24 months and 48 months earlier. We thus construct two different time windows between 2016 and 2019 (April 2016 to June 2017 is the artificial pre-lockdown period; April 2018 to June 2019 is the artificial post-lockdown period) and between 2014 and 2017 (April 2014 to June 2015 is the artificial pre-lockdown period; April 2016 to June 2017 is the artificial post-lockdown period).

4. Results

4.1 Descriptive Analysis of Out-of-Town Property Purchases by NYC Homebuyers

This section presents a descriptive analysis of property transactions in NY and five adjacent states of NYC. Specifically, we are interested in the geographical distribution of residential properties purchased by NYC homebuyers before and after the COVID-19 outbreak and how the composition of homebuyers (i.e., local homebuyers, NYC homebuyers, all other homebuyers) changes after COVID-19. Individual property purchases are aggregated at the zip code level and categorized by within versus outside NY or NYC and *urban* versus *non-urban* status.

[Insert Figure 1 Here]

Figure 1 presents the number of property transactions made by NYC homebuyers from January 2018 to June 2021. We find that the number of monthly property purchases drops to the lowest point in March 2020, then rebounds and surpasses previous years. The main drivers of this rebound are out-of-town property purchases in the study regions, which increase to 74.2% of all purchases by NYC homebuyers in August 2020 from a two-year average of 52.5% before COVID-19 (see Figure A2 in Appendix). While non-NYC urban areas experience a larger absolute increase in the number of NYC homebuyers, non-urban regions witness a higher relative increase. In addition, our data show that NYC homebuyers mostly purchase properties within NY throughout the study period (>75% of the study area all the time, see Figure A3 in Appendix).

[Insert Figure 2 Here]

The drastic and abrupt increase of NYC homebuyers in other NY areas makes NYC homebuyers a more prominent source of housing demand in those areas. Figure 2-a shows that while the total number of transactions after COVID-19 is similar to previous years, the percentage of purchases made by NYC homebuyers has increased in non-NYC urban areas[10] within NY. The increase in the share of NYC homebuyers is even more pronounced in non-urban areas (Figure 2-b). In both urban and non-urban areas, the increase in NYC homebuyers mostly replaces the percentage of purchases by local residents. Figure 3 shows the changes in the number of housing purchases in the study region. Only New Jersey experiences decreased migration flow in most zip code areas[11], while all other states have higher inflows of NYC homebuyers.

[Insert Figure 3 Here]

[10] For simplicity purposes, we use urban areas to denote non-NYC urban areas in the rest of the paper.
[11] For graphing purpose, transactions at zip code level are projected to the ZIP Code Tabulation Areas (ZCTAs) level in figure 3. ZCTAs are generalized areal representations of United States Postal Service (USPS) ZIP Code service areas developed by the U.S. Census Bureau. There are 4,724 ZCTA areas in our study region.

4.2 Quantifying the Increase in Out-of-Town Property Purchases by NYC Homebuyers

Figure 4 (Panels A-C and E-G) presents the season-adjusted relative percentage changes in quarterly housing transactions (or β_{qt} in equation 2) in *non-urban* and *urban* areas by the origins of homebuyers. The baseline is the number of property transactions in Q1 2020. Figure 4B shows that, in the four quarters from Q3 2020 to Q2 2021, the season-adjusted property purchases in the non-urban areas by NYC homebuyers increase by 189%, 204%, 101%, and 53% compared to baseline, respectively. In contrast, property purchases by local residents decrease after the COVID-19 outbreak and then recover to the pre-COVID-19 level since the third quarter of 2020 (Figure 4A). Results for the urban areas (Figure 4F) show that purchases by NYC homebuyers increase by 30%, 84%, 50%, and 25% in the same period from Q3 2020 to Q2 2021, respectively, but the relative magnitudes of increase are smaller compared to non-urban (Figure 4B). Purchases by local and other out-of-town homebuyers first decrease at the beginning of the pandemic, then partially recover in late 2020 and early 2021 (Figure 4E and 4G).

[Insert Figure 4 Here]

Panels A, C, E, and G in Figure 4 show that demand from local and other homebuyers also fluctuated after the onset of COVID-19. These fluctuations may result from market frictions at the beginning of COVID-19 (e.g., lock-down, fear of in-person contact, etc.) and the eventual release of the pent-up demand. To isolate the difference between the demand of NYC homebuyers and the fluctuation in overall demand, we use transactions by local homebuyers as the control group for a DID analysis. The DID estimations in Figures 4D and 4H show that the relative increases in property purchases by NYC homebuyers in urban and non-urban regions reach 79% and 187% at the highest points, respectively, which is similar to the event analysis results.

4.3 Location Characteristics and the Change of NYC Homebuyers after the COVID-19 Pandemic

The surge of homebuyers from NYC after COVID-19 could be accompanied by changing demographic composition of homebuyers and different location preferences. We construct a linear model and regress the change in total property purchases by NYC homebuyers (ΔN) in each zip code area of the study region before (i.e., April 2018 to June 2019) and after (i.e., April 2020 to June 2021) the COVID-19 outbreak on a collection of location characteristics. Since the location characteristics are time-invariant in our data during the study window, the coefficients can be interpreted as how the importance of different pull factors changes after the pandemic. We first perform the regression analysis on the NY sample (as most of the property purchases by NYC homebuyers are within NY) followed by a separate analysis using the sample including adjacent states. The location characteristics (see Table A1 in Appendix) include the logarithmic form of the total population, the logarithmic form of distance to NYC, natural amenity rank, high-speed internet access per 1,000 households, urban amenity measured by the number of retail businesses and food service establishments per 1,000 population, race, education, and the difference in housing price indices at the first month of the two commensurate time windows (i.e., Jan. 2018 and Jan. 2020). Detailed variable definitions are described in Section 2. Figure 5 depicts estimation results using NY sample (hollow circles) and all study regions sample (solid circles), separated by sub-samples of Non-Urban, Urban, and All Regions, respectively.

Figure 5 (Column 3) shows that relative to a "Fair" amenity (rankings = 2), "Good" (ranking = 3) and "Excellent" (ranking = 4) amenity rankings in NY overall drive up ΔN by 0.11 and 1.02 standard deviations, respectively. In contrast, when urban amenity, measured by the number of retail and food service establishments per 1,000 population, increases by one standard deviation, ΔN decreases by 0.21 standard deviations. The increase in the attraction of natural amenity and the reduction in the attraction of urban amenity is consistent with existing evidence on how residents' location preference changes due to COVID-19 (Guglielminetti & Rondinelli, 2021; Liu & Su, 2021). In NY, when access to high-speed internet increases by one standard deviation, ΔN increases by 0.24 standard deviation. The increased importance of internet access may reflect the preference of WFH workers. Places in NY closer to NYC, with more population, less black population, more highly educated population, and less increase in housing price also attract more NYC homebuyers. Most of the above results for NY still hold for the entire study region (solid circles in Figure 4), though the effect of internet access is no longer statistically significant. The most notable differences between urban and non-urban areas are that the decrease in the attraction of urban amenities happens in urban areas, while the negative effect of distance to NYC is more pronounced in non-urban areas.

[Insert Figure 5 Here]

Besides WFH considerations, homebuyers are likely to avoid locations with worse COVID-19 conditions. To test this hypothesis, we regress the change in monthly property transactions at the zip code level relative to the corresponding month two years ago before COVID-19 on death rates and case rates. The summary statistics for the variables are reported in Appendix Table A2. Current death rates or case rates and those in three lagged months are included. The differencing in property transactions removes the influence of time-invariant location characteristics and seasonal fluctuations. We also include year-month fixed effects to control omitted variables common for all zip codes, such as the mortgage interest rate. County fixed-effects are included to control for county-specific trends in local conditions such as wage and employment. Finally, changes in the housing price index (ZHPI) and its lags are included to control for price effects.

In general, the analysis shows a significant correlation between COVID-19 status and the number of SFH transactions (Figure 6). The red circles in Figure 6A show that COVID-19 death rates have substantial deterrence effects on within-state housing purchases by NYC homebuyers. When the monthly death rate increases by one per 1,000 population, property purchases in NY by NYC homebuyers decrease by 1.18 standard deviations in the current month, and lagged effects are statistically significant up to two lagged months (i.e., t-2 and t-1). On the other hand, an increase in the monthly infection rate in t-2 and t-1 by one per 1,000 population decreases property purchases by NYC homebuyers by 0.009 and 0.016 standard deviations, respectively (Figure 6B). The effects of case rates are delayed relative to the effects of death rates, which is consistent with how the pandemic spreads. The Estimates based on the entire study region (Figure 6C and 6D) are qualitatively similar but smaller in magnitude than those from the NY sample[12]. One potential explanation is that within-state and interstate migrants may have different demographic compositions and location preferences.

[Insert Figure 6 Here]

COVID-19 death rates and infection rates may pick up pre-existing trends in different zip code areas. As a falsification test, we construct two additional datasets (ranging from 2016-2019 and 2014-2017, respectively) with the real COVID-19 death rates and case rates artificially moved to 24 months and 48 months earlier. We repeat the same panel analysis with the artificial COVID-19 data. The green triangles and blue squares in Figure 6 show that, when moved earlier by 2 and 4 years, COVID-19 death and infection rates have no statistically significant effects on property purchases by NYC homebuyers either within NY or in the entire study area.

5. Concluding Remarks

Since the start of the COVID-19 pandemic and the rise of WFH, researchers have hypothesized that there will be a significant redistribution of population and economic activities. While anecdotal evidence is plenty, rigorous empirical evaluations are mostly based on indirect evidence such as housing prices. This study shows that out-of-town property transactions is a convenient proxy for migration flows and provides direct evidence for the spatial redistribution of population after COVID-19. Using individual-level property transaction data, we study the change in property purchases by NYC residents in urban and non-urban areas in NY and five adjacent states. At the highest point, quarterly property purchases in urban and non-urban areas by NYC homebuyers increase by 79% and 187% relative to pre-COVID-19 levels, respectively. As a result, housing demand by NYC homebuyers makes up a significantly larger percentage of both the urban and non-urban housing markets in surrounding areas after the COVID-19 outbreak. We also find that higher natural amenities and access to high-speed internet have a higher

[12] We do not find evidence for the deterrence effects on cross-state housing purchases using individual samples of the other five study states. See Appendix Tables B5 and B6 for detailed regression results.

attraction to NYC homebuyers after COVID-19. In contrast, the attraction of urban amenities decreases. Furthermore, higher COVID-19 death and infection rates have significant deterrence effects on NYC homebuyers' decision-making.

The increase in the percentage of NYC homebuyers in the local markets is substantial, especially in non-urban areas. The magnitude of the increase suggests that spatial redistribution has significant impacts on the destination regions. NYC homebuyers will pay property taxes that fund local public goods, spend money on locally produced goods and services, and create spillover effects that improve the local productivity. They may also vote differently compared to locals, changing the political landscape. While this study takes the initial step in quantifying spatial redistribution after the pandemic, future research can expand to the national level and utilize other data sources, such as administrative and Census data, when they become available.

References

Althoff, L., Eckert, F., Ganapati, S., & Walsh, C. (2021). *The Geography of Remote Work* [Preprint]. Institute Working Paper.

Barrero, J. M., Bloom, N., Davis, S. J., & Meyer, B. H. (2021). COVID-19 Is a Persistent Reallocation Shock. *AEA Papers and Proceedings*, *111*, 287–291.

Bick, A., Blandin, A., Virginia Commonwealth University, & Mertens, K. (2021). Work from Home Before and After the COVID-19 Outbreak. *Federal Reserve Bank of Dallas, Working Papers*, *2020*(2017).

Brueckner, J., Kahn, M. E., & Lin, G. C. (2021). *A New Spatial Hedonic Equilibrium in the Emerging Work-from-Home Economy?* (Working Paper No. 28526; Working Paper Series). National Bureau of Economic Research. https://doi.org/10.3386/w28526

Couture, V., Dingel, J. I., Green, A., Handbury, J., & Williams, K. R. (2021). JUE Insight: Measuring movement and social contact with smartphone data: a real-time application to COVID-19. *Journal of Urban Economics*, 103328.

Davis, M., Ghent, A., & Gregory, J. (2021). *The Work-from-Home Technology Boon and its Consequences* (No. w28461; p. w28461). National Bureau of Economic Research.

Delventhal, M. J., Kwon, E., & Parkhomenko, A. (2021). JUE Insight: How do cities change when we work from home? *Journal of Urban Economics*, 103331.

D'Lima, W., Lopez, L. A., & Pradhan, A. (2020). COVID-19 and housing market effects: Evidence from US shutdown orders. *Available at SSRN 3647252*.

Francke, M., & Korevaar, M. (2021). Housing markets in a pandemic: Evidence from historical outbreaks. *Journal of Urban Economics*, *123*, 103333.

Guglielminetti, E., & Rondinelli, C. (2021). Consumption and saving patterns in Italy during Covid-19. *Bank of Italy Occasional Paper*, *620*.

Gupta, A., Mittal, V., Peeters, J., & Van Nieuwerburgh, S. (2021). *Flattening the curve: Pandemic-induced revaluation of urban real estate*. National Bureau of Economic Research.

Haslag, P. H., & Weagley, D. (2021). From LA to Boise: How Migration Has Changed During the COVID-19 Pandemic. *Available at SSRN 3808326*.

Heather Kelly & Rachel Lerman. (2020, June 1). The pandemic is making people reconsider city living, trading traffic for chickens. *Washington Post*. https://www.washingtonpost.com/technology/2020/06/01/city-relocate-pandemic/

Hill, A. (2021, February 8). "We're happier, calmer": Why young adults are moving out of big cities. *The Guardian*. https://www.theguardian.com/world/2021/feb/08/were-happier-calmer-why-young-adults-are-moving-out-of-big-cities

Leyk, S., Uhl, J. H., Connor, D. S., Braswell, A. E., Mietkiewicz, N., Balch, J. K., & Gutmann, M. P. (2020). Two centuries of settlement and urban development in the United States. *Science Advances*, *6*(23).

Liu, S., & Su, Y. (2021). The impact of the COVID-19 pandemic on the demand for density: Evidence from the U.S. housing market. *Economics Letters, 207*, 110010.

Low, S. A., Rahe, M., & Van Leuven, A. J. (2022). Has COVID-19 made rural areas more attractive places to live? Survey evidence from northwest Missouri. *Regional Science Policy & Practice*, *n/a*(n/a).

Molloy, R., Smith, C. L., & Wozniak, A. (2011). Internal migration in the United States. *Journal of Economic Perspectives*, *25*(3), 173–196.

Nolte, C., Boyle, K. J., Chaudhry, A. M., Clapp, C. M., Guignet, D., Hennighausen, H., Kushner, I., Liao, Y., Mamun, S., Pollack, A., Richardson, J., Sundquist, S., Swedberg, K., & Uhl, J. H. (2021). *Studying the Impacts of Environmental Amenities and Hazards with Nationwide Property Data: Best Data Practices for Interpretable and Reproducible Analyses* (SSRN Scholarly Paper ID 3900806). Social Science Research Network.

Ozimek, A. (2020). *Remote Workers on the Move* (SSRN Scholarly Paper ID 3790004). Social Science Research Network.

Quesnel, K. J., Agrawal, S., & Ajami, N. K. (2020). Diverse paradigms of residential development inform water use and drought-related conservation behavior. *Environmental Research Letters*, *15*(12), 124009.

Ramani, A., & Bloom, N. (2021). *The Donut Effect of COVID-19 on Cities* (SSRN Scholarly Paper ID 3850758). Social Science Research Network.

Rosenthal, S. S., Strange, W. C., & Urrego, J. A. (2021). JUE insight: Are city centers losing their appeal? Commercial real estate, urban spatial structure, and COVID-19. *Journal of Urban Economics*, 103381.

Rupasingha, A., Liu, Y., & Partridge, M. (2015). Rural bound: Determinants of metro to non-metro migration in the United States. *American Journal of Agricultural Economics*, *97*(3), 680–700.

Shen, X., Liu, P., Qiu, Y. L., Patwardhan, A., & Vaishnav, P. (2021). Estimation of change in house sales prices in the United States after heat pump adoption. *Nature Energy*, *6*(1), 30–37.

Figures

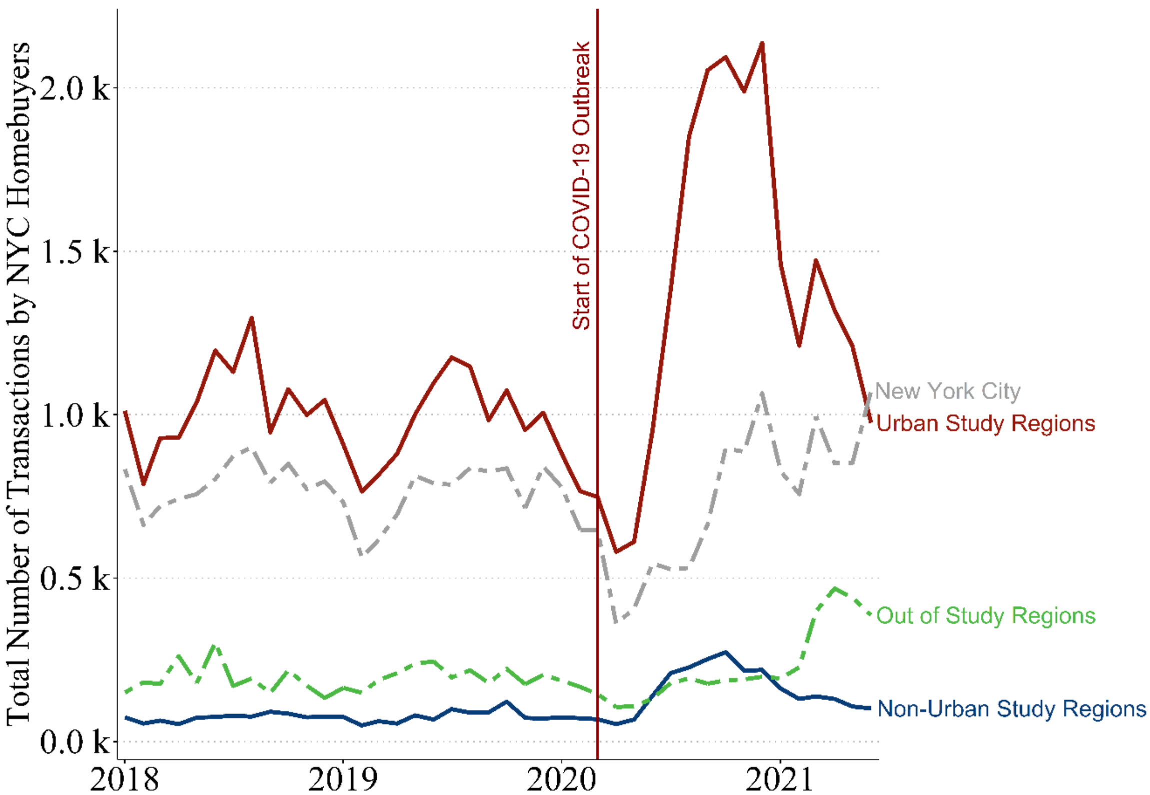

Figure 1. The trend of monthly transactions by NYC homebuyers in different regions.

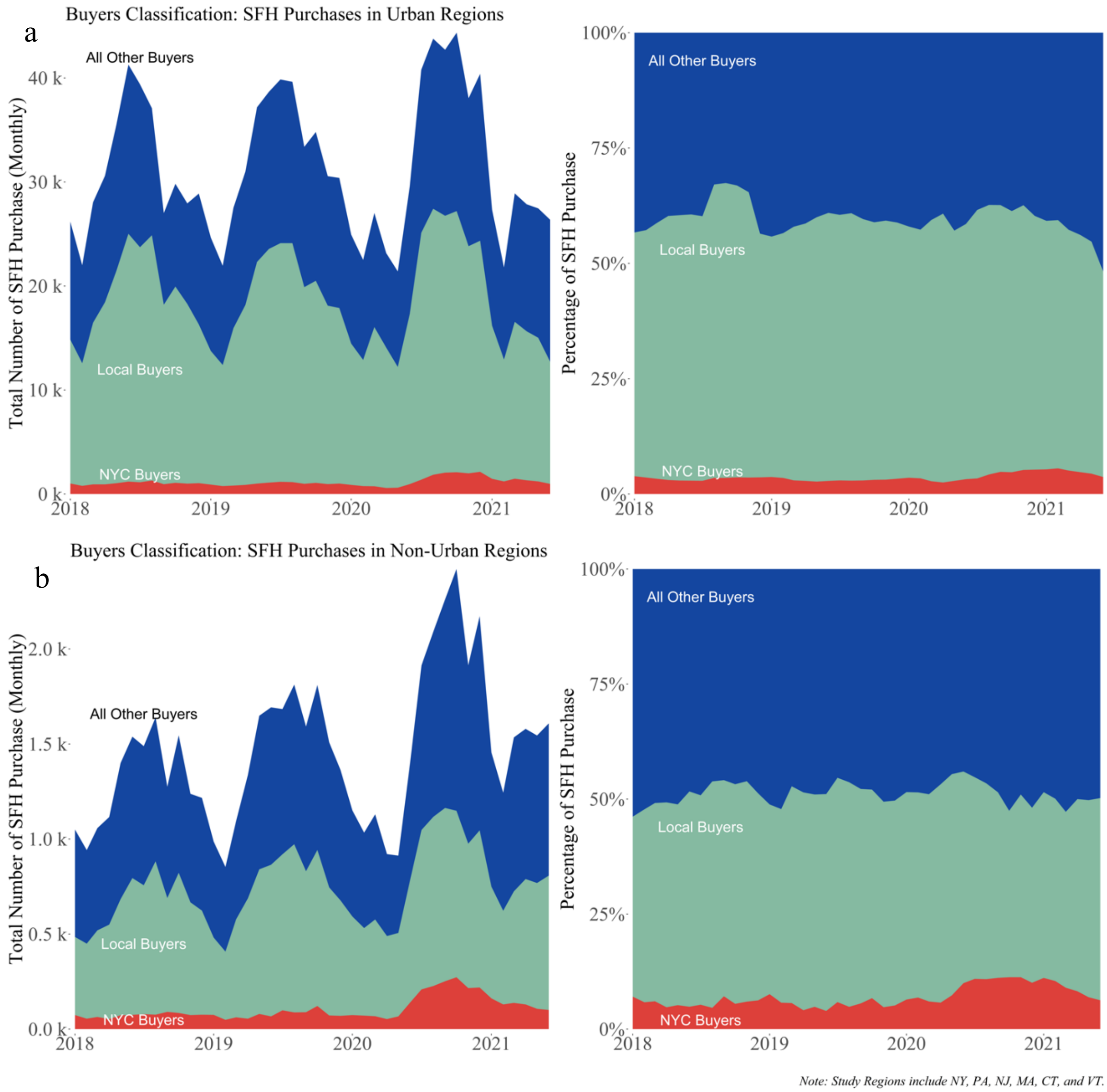

Figure 2. Housing transactions in non-NYC Urban and Non-Urban areas of study regions by NYC, local, and other homebuyers between January 2018 to June 2021.

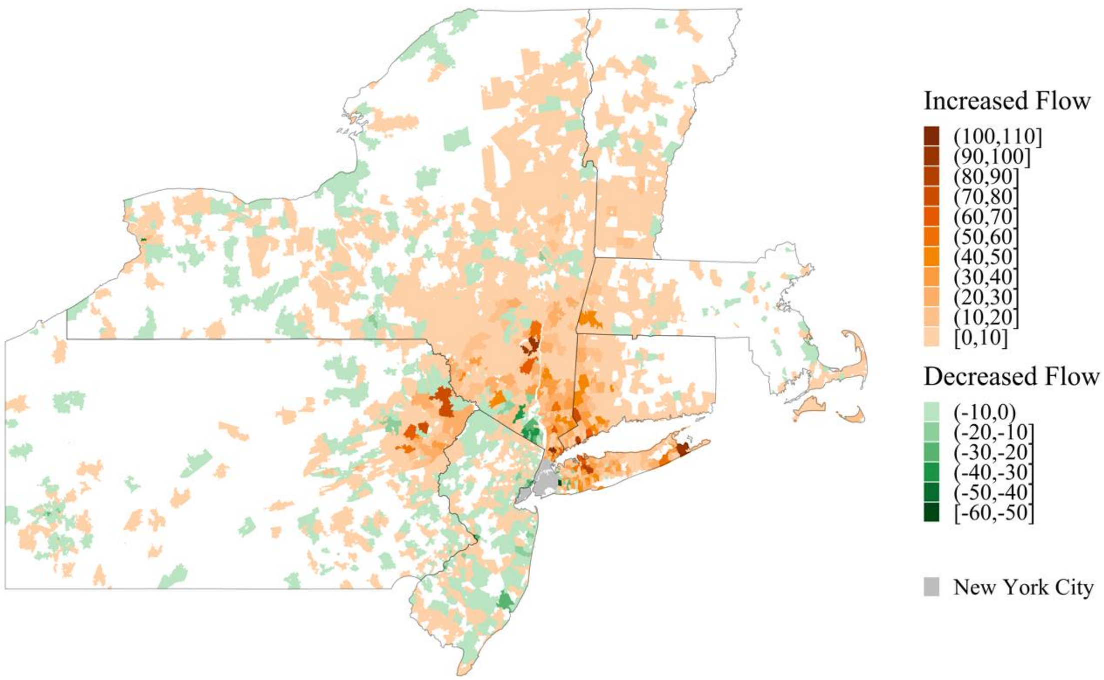

Figure 3. The change in the number of housing purchases by NYC homebuyers in the study area before (April 2018-June 2019) and after (April 2020-June 2021) COVID-19 started.

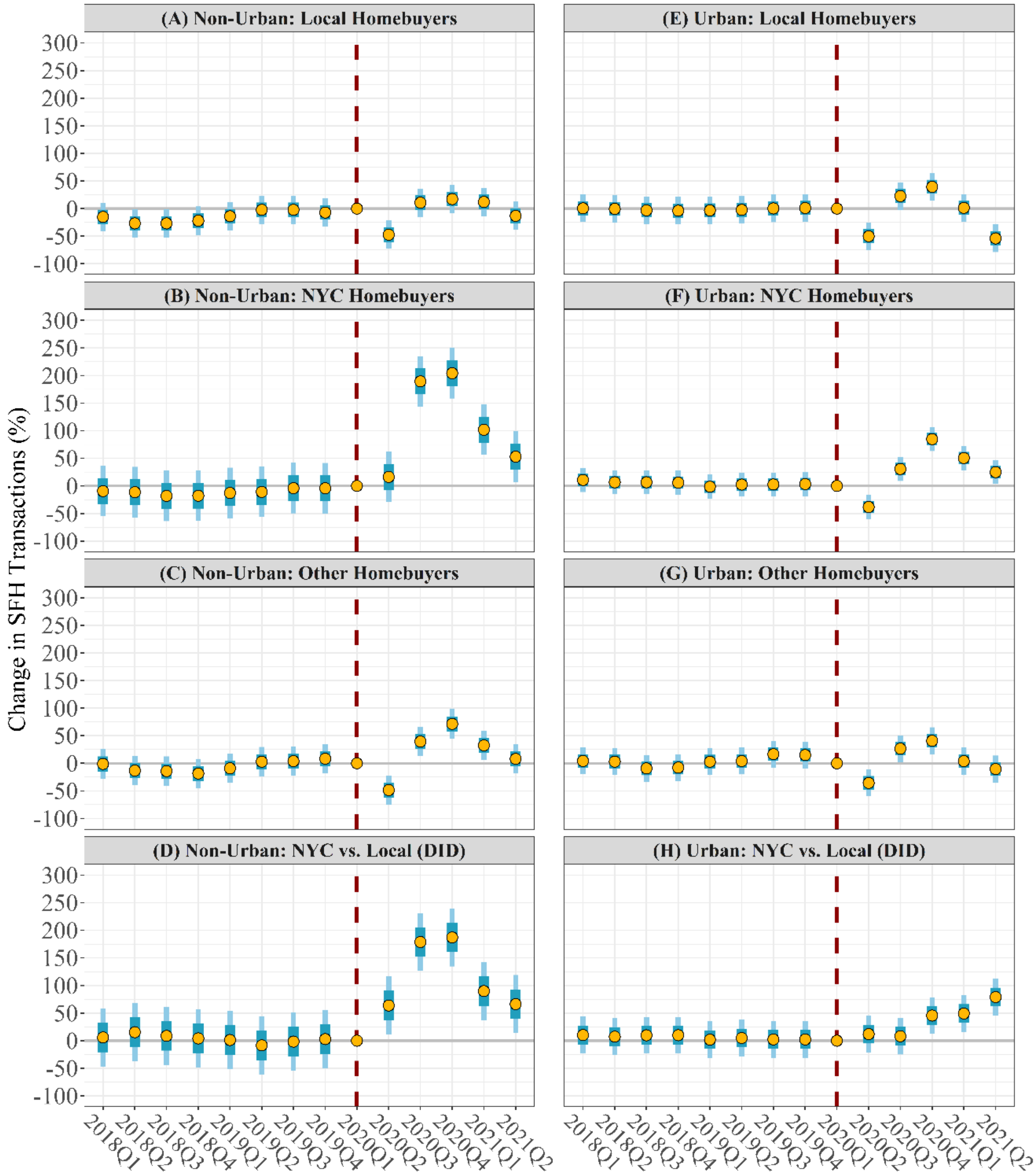

Figure 4. Event study and DID estimations of the changes in season-adjusted housing purchases in urban and non-urban areas

Notes: The study area includes non-urban areas in NY, PA, NJ, MA, CT, and VT. Panels A to D are event study results (the estimations of β_{qt} in equation 2) and DID results (the estimations of γ_{qt} in equation 3) in non-urban regions, while Panel E to H are those in urban regions. The coefficients represent the increase of season-adjusted property transactions as a percentage of housing transactions in Q1 2020. Standard deviations (boxes) and 95% confidence intervals (whiskers) are plotted with coefficient estimates.

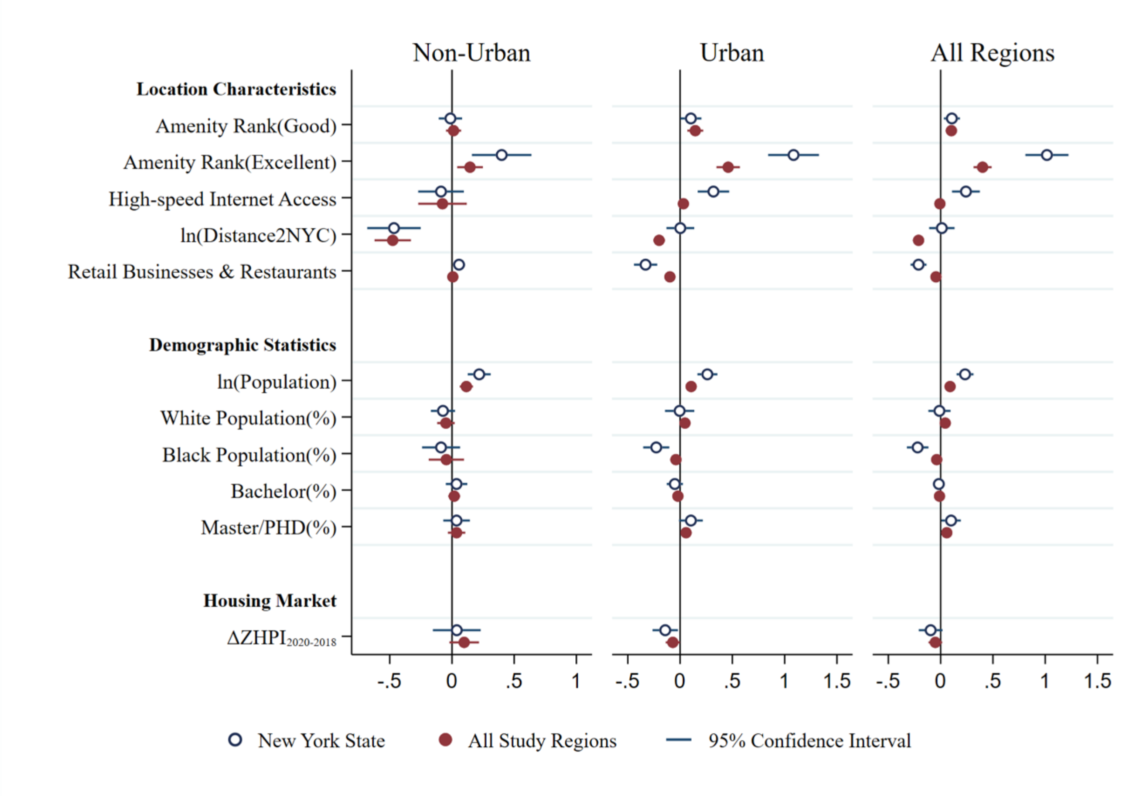

Figure 5. The coefficient plot of cross-sectional regressions for fixed location characteristics

Notes: This figure presents estimation results of equation (4) based on samples of non-urban zip code areas, urban zip code areas, and all zip code areas in NY (hollow circle) and in the entire study regions (solid circle). The dependent variable is the standardized change in the number of transactions by NYC homebuyers between two periods (ΔN), April 2018 to June 2019 and April 2020 to June 2021. Coefficients of standardized covariates are shown for continuous variables (all except Amenity Rank). Whiskers represent 95% confidence intervals. Estimation results using non-standardized and standardized covariates are in Appendix Tables B1 and B2.

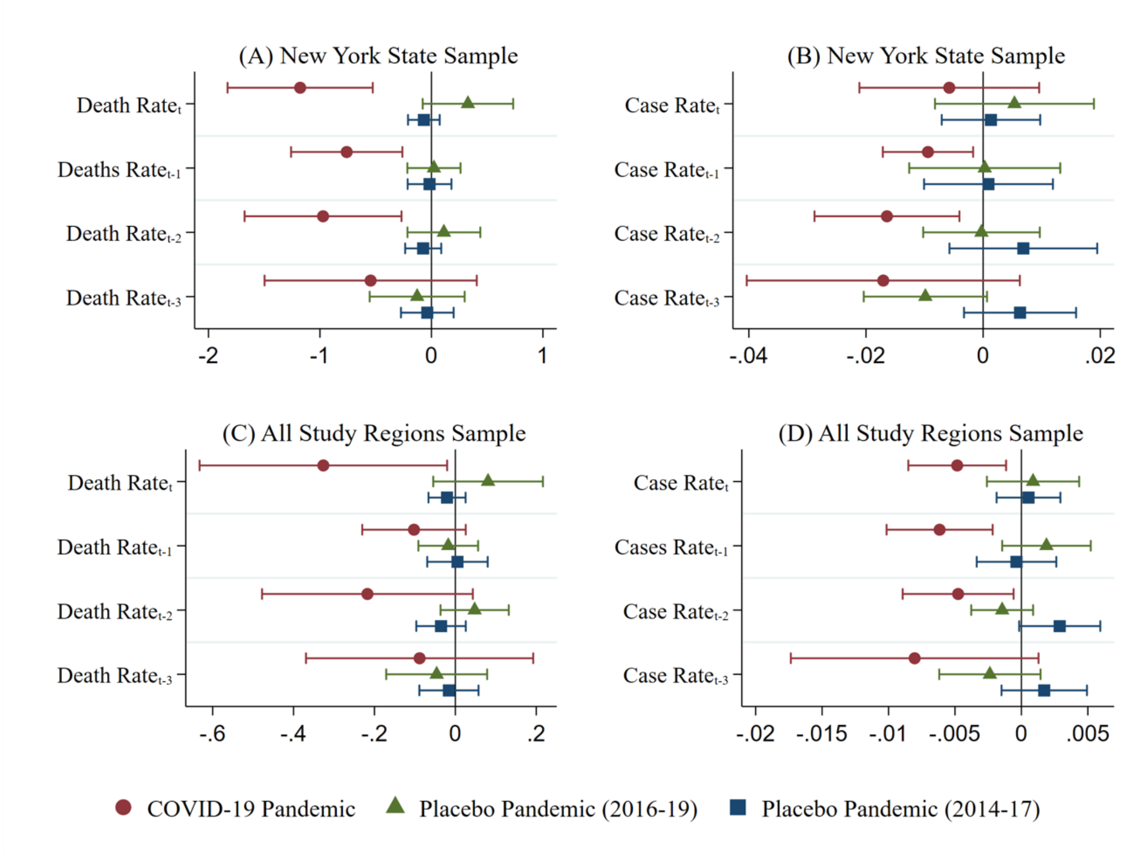

Figure 6. The Coefficient plot of panel regressions for pandemic death and case rates

Note: The figure presents estimation results for equation (5) based on the sample of NY (Panel A and B) and the sample of all study regions (Panel C and D), respectively. The dependent variable is the difference in the number of housing transactions in each included zip code area from a post-COVID-19 month t and a corresponding month $(t-24)$ two years earlier before COVID-19. Pandemic conditions are measured by death rates in panels A-C and by case rates in panels B-D. Red circles are results based on actual data associated with the COVID-19 pandemic between 2018-2021. Green triangles and blue squares are the results of two models that assume the COVID-19 outbreak occurs at 24 months and 48 months earlier. That is, the pandemic is assumed to occur in the years 2018 and 2016, respectively. The dependent variable and all covariates, except death and case rates, are standardized. Detailed regression results are in Appendix Tables B3 and B4.