NBER WORKING PAPER SERIES

NEIGHBORHOOD CHOICE AFTER COVID: THE ROLE OF RENTS, AMENITIES, AND WORK-FROM-HOME

Fernando V. Ferreira
Maisy Wong

Working Paper 29960
http://www.nber.org/papers/w29960

NATIONAL BUREAU OF ECONOMIC RESEARCH
1050 Massachusetts Avenue
Cambridge, MA 02138
April 2022

We are grateful for support from the Wharton Dean's Research Fund, the Research Sponsors Program of the Zell/Lurie Real Estate Center, and the Wharton Behavioral Lab. The views expressed herein are those of the authors and do not necessarily reflect the views of the National Bureau of Economic Research.

Neighborhood Choice After COVID: The Role of Rents, Amenities, and Work-From-Home
Fernando V. Ferreira and Maisy Wong
NBER Working Paper No. 29960
April 2022
JEL No. G0,H0,J0,R0

ABSTRACT

We investigate how neighborhood preferences and choices changed one year after the beginning of the COVID pandemic. We study a Neighborhood Choice Program that helped graduating students choose where to live by providing new information about rents and amenities. Using panel data on neighborhood rankings before and after information, we find that changes in rankings favor neighborhoods where social and professional network shares are higher by 2.2 percentage points, rents are lower by $432, and are 2.4 kilometers farther from the city center. Interestingly, we did not detect this movement away from downtowns when the program was offered prior to the pandemic. We then estimate a neighborhood choice model to recover MWTP for amenities both before and after the pandemic. Our estimates reveal that MWTP for network shares post COVID is markedly lower than prior to COVID. Finally, we perform counterfactuals to quantitatively assess how changes in preferences affect where people live, and find that weaker network preferences are most impactful, while heterogeneity by commute and work-from-home are less relevant.

Fernando V. Ferreira
Wharton Real Estate
University of Pennsylvania
430 Vance Hall
3733 Spruce Street
Philadelphia, PA 19104-6301
and NBER
fferreir@wharton.upenn.edu

Maisy Wong
Wharton Real Estate
University of Pennsylvania
434 Vance Hall
3733 Spruce Street
Philadelphia, PA 19104-6301
and NBER
maisy@wharton.upenn.edu

1 Introduction

Understanding household preferences for neighborhoods have long been of interest for researchers given their direct implications for the way households sort in the housing market and the subsequent amount of local economic activity, among others.[1] The COVID pandemic, however, has led many people to rethink where they want to live. Many experienced a negative shock to their social and professional interactions that changed the value of those activities, including the consequences for venues which facilitate those interactions, such as offices, hotels, restaurants, bars, etc. Moreover, spending more time at home led many people to desire to live in bigger residences, especially with the rising prevalence of work-from-home (WFH). More generally, the pandemic may have led individuals to rethink how much they value different neighborhood amenities. All those factors can have implications for where people live and the demand of residential real estate, potentially reshaping the landscape of the real economy.

There are several challenges to study how neighborhood choices were impacted by the COVID pandemic. The first is a measurement issue as there are not many datasets that link the multiple factors associated with household mobility during the pandemic, such as social and professional network opportunities, commuting, and work-from-home. Second, it is important to have a model to empirically assess how these multiple channels influence neighborhood choices, and compare estimates using data before and after COVID. Third, neighborhood choices are endogenous, riddled with unobserved quality and imperfect information, so it is necessary to implement appropriate research designs to deal with resulting omitted variable biases (Baum-Snow and Ferreira, 2015).

This paper addresses the challenges above in order to investigate household location decisions one year after the beginning of the COVID pandemic in the United States. We estimate a discrete choice model where households maximize utility by choosing from a set of neighborhoods that are differentiated along several dimensions. We provide new estimates of preferences for neighborhood rents and amenities, including how those preferences vary by individual preferences for work-from-home and for commuting. Our model allows us to quantitatively assess how changes in the valuation of these amenities ultimately matter for where households live. We benchmark our new preference estimates against pre

[1] Neighborhood demand is critical for the determination of housing asset prices and mortgage debt usage (Gupta, 2019; Baldauf, Garlappi, and Yannelis, 2020; Miller and Soo, 2021), the level of income and racial segregation and the quality of amenities such as schools (Bayer, Ferreira, and McMillan, 2007; Wong, 2013), and the determinants of place of work and place of residence (Bayer, Ross, and Topa, 2008).

COVID parameters, and characterize how the spatial distribution of residential economic activity can change because of post COVID changes in preferences.

The empirical work focuses on university graduates, who constitute a large fraction of new household moves.[2] In April 2021, we partnered with a large professional school in the United States to offer a Neighborhood Choice Program that provides information to help students choose where to live after graduation. This is an opportune time, about a month before graduation, giving us access to hundreds of potential movers at a time when they already have a job and are seriously searching for a new home. Since they know their employer, we circumvent the usual problem whereby the decision of where to live is jointly determined with the decision of where to work. Importantly, the school offered a similar Neighborhood Choice Program in 2019 as well, allowing us to quantify how neighborhood choices have changed before and after the pandemic.[3]

The 2021 program provided information on three important neighborhood factors: rental costs, access to social and professional network, and size of the rental units, which proxies for living and home office space. To complete the 7-minute online survey, students first indicated which MSA they desired to live in upon graduation. We then asked them to rank up to ten neighborhoods in that MSA, provided information about the three neighborhood factors, and asked them to re-rank again. In addition, the survey asked about the relevance of different factors in location decisions, such as how many days they expect to work-from-home, their commuting preferences, and if they are married with kids.

Our reduced form analysis utilizes panel data on rankings before and after information. Around 40% of the graduating cohort completed the survey.[4] We observe 310 students ranking around 20 neighborhoods in their desired MSA, totaling 6,941 potential individual-by-neighborhood rankings. Notably, many students only rank 5 neighborhoods out of a possible consideration set of 10. Accordingly, 71% of the potential rankings are associated

[2]There were almost 20 million individuals enrolled in college and 3.1 million enrolled in post college degrees in 2019 in the United States (National Center for Education Statistics, 2020), and every year a new cohort of graduates have to choose where to live. Recent research has highlighted the role of housing costs, amenities, social and professional networks in shaping the location choices of young adults (Moretti, 2013; Diamond, 2016) and the importance of friends in their decision-making (Bailey, Cao, Kuchler, Stroebel, and Wong, 2018).

[3]Ferreira and Wong (2020) used the 2019 data to develop a new method that identifies preferences in a generalized neighborhood choice model that allows for individual imperfect information around amenities, finding that information on rents and networks have persistent effects on mobility decisions, up to one year after graduation.

[4]This is relatively high considering we only provided information for 18 of the most popular MSA's in the United States. In the past decade, 66% of graduates lived in these MSA's right after graduation.

with neighborhoods that are never considered in the top 10. An additional 8% are associated with rankings that are always the same before and after information. Crucially, 21% of the choices reflect changes in rankings. Our descriptive analysis will focus on showing what happens to rents, network shares, and rental home size, for four types of neighborhoods (those switched in and out of the top ten consideration set, as well as always- and never-top neighborhoods). This variation in the data will motivate our identification strategy for the structural estimation of the neighhborhood choice model.

Neighborhoods that are always and never in the top ten consideration set are systematically different. Always-top neighborhoods are more expensive, have higher network shares, and have smaller size, relative to never-top neighborhoods. This likely reflects students' preferences for renting units close to the city center. That popular neighborhoods are centrally located, expensive, and have larger networks also underscores confounding due to unobserved neighborhood quality. Importantly, after the information, individuals are more likely to switch in neighborhoods where the average rent is lower by $432 and the network shares are higher by 2.2 percentage points (pp). These are large magnitudes relative to sample averages ($2500 and 5%, respectively), and are consistent with preferences to live in neighborhoods with low cost of living and high alumni network shares. The effects for rental size are small and insignificant in the full sample. The patterns above survive individual fixed effects, and are also robust to using other neighborhood preference outcomes after the online survey.[5]

When we interact switcher behavior with our measures of heterogeneity, we find stronger effects along a few dimensions. Households with children switch in places with higher square footage. Moreover, a high tolerance for long commutes leads to neighborhoods that are cheaper and with smaller networks (favoring the suburbs). Interestingly though, variation in expected days working-from-home, based on what their individual employers were planning to adopt, did not lead to any significant effect on the behavior of switchers. This is consistent with the hybrid model of WFH with workers only working from home part of the week (Ramani and Bloom, 2021). Fifty percent of our respondents expect to WFH at most 1 day per week and 25% expect to WFH 2 days per week.

Ultimately, what do the findings above mean for where individuals choose to live? Preferences for lower rent and larger units point towards the suburbs whilst the importance

[5]These survey estimates in the 2021 post COVID sample were corroborated by the search behavior of students who used a mapping tool provided by the school in order to help them during their neighborhood search from April to July of 2021. In Ferreira and Wong (2020), we corroborate the stated preference results using actual locations one year after graduation.

of network amenities point towards the city center.[6] We find that always-top choices are associated with neighborhoods that are 4.4 kilometers from the city center and never-top neighborhoods are 16.3 km away, indicating that city centers are still quite popular. However, post information, individuals switch in neighborhoods that are 8.3 km from the city center and switch out neighborhoods that are 5.9 km away - so at the margin, we find patterns consistent with preferences for neighborhoods just outside the city, where rents are lower and network shares are still moderate.

By comparison, using the 2019 data, we find that individuals still prefer cheap neighborhoods and high network shares but the relative valuations are not enough to favor the suburbs over the city center. The patterns for always-top and never-top neighborhoods remain similar in 2019 and 2021. However, switchers in 2019 did not care as much about lower rents (they switch in neighborhoods that are only $347 cheaper in 2019 relative to $432 in 2021) and they responded more to network shares (2.5 pp relative to 2.2). Together, the weaker rent effects coupled with the stronger network effects both work against the suburbs and can rationalize why we do not detect significant effects on distance in 2019.

Next, to quantify how the different channels influence neighborhood choices, we estimate a discrete choice model that recovers preferences for rents and neighborhood amenities (see, for example, Berry (1994), Berry et al. (1995), Bayer et al. (2007)). In the model individuals maximize utility by choosing neighborhoods among a set of consideration neighborhoods (based on rankings provided by each respondent). Our model also allows for heterogeneity in preferences - since individuals have different commuting tolerances and different work-from-home arrangements. Estimation follows two stages. In the first stage we recover heterogeneity in preferences using a rank ordered logit model. The first stage also produces a set of neighborhood fixed effects that are decomposed into mean preferences in a second stage.

The main identification problem is that rents and amenities are usually correlated with neighborhood variables that are not observed by the econometrician, such as quality of sidewalks, number of trees, etc. Our research design overcomes this endogeneity in amenities by utilizing panel data on neighborhood rankings and variation from the switchers. We rely on individuals' pre information rankings to construct a latent index for neighborhood quality. Our assumption is that pre information rankings remain relevant proxies of neighborhood quality after the information shock, which seems plausible since the rankings

[6] We define city center using the city hall for each MSA but our results are robust to using the centroid based on location data from previous cohorts.

occur within a span of 2 minutes (Wiswall and Zafar, 2014). After the information shock, switchers change rankings due to the new information provided by the survey, allowing us to properly estimate preferences for those variables. The individual level variation also allows us to control for heterogeneity in knowledge about neighborhood amenities (Ferreira and Wong, 2020).[7]

Identical models estimated with the 2019 and 2021 data show that the average household displays a similar marginal utility for rents, before and after COVID: -$478 in 2019 and -$454 in 2021. However, estimates for social and professional network became smaller in 2021 (0.2) when compared to 2019 (0.37). That resulted in marginal willingness to pay estimates for network of only $86 per month for a 1 percentage point increase in network shares in 2021, while in 2019 that same MWTP was $171. Next, we estimate an augmented model using the additional heterogeneity variables captured in the 2021 data, such as WFH and commute. Even though the estimated heterogeneity parameters for WFH and commute are not significantly different from zero, the resulting MWTP estimates for network becomes moderately higher ($109).

Next, we use the 2021 model to explore several counterfactuals that could affect neighborhood choices. First, we quantify the importance of the main change in preferences between 2019 and 2021, i.e., the decline in willingness to pay for networks. The 2021 baseline estimates show that the top choice for all individuals in the sample implies an average distance of 4.83 kilometers from the city center. Using a counterfactual simulation where we only revert the network preferences to 2019 levels, we find that, on average, individuals would locate closer to center city, at 4.17 km. That is a sizable displacement in location of residential activity, consistent with our reduced form findings above. We find similar results for a second counterfactual where neighborhood rents completely revert to 2019 levels (in addition to changing preferences for networks). Such changes are more pronounced when focusing on New York and San Francisco, the most preferred MSAs in our data. We also considered counterfactuals that shut down the importance of heterogeneity in preferences for WFH and commute. Not considering commute or WFH heterogeneity lead to small changes in distances of 0.28 km and 0.12 km, respectively.

We make three contributions in this paper. First, we use a novel dataset that links neighborhood choices with key factors including rents, networks, work-from-home, commuting,

[7]The validity of the latent quality index relies on relevance and excludability assumptions elaborated in Section 4. First, that individuals report pre information rankings that remain relevant proxies for the desirability of neighborhoods. Second, that changes in ranking after the information intervention reflect new information about rents and amenities.

and family status. We study young professionals who are actively searching for where to live before and after the pandemic. Second, we develop research designs that overcome omitted variable biases in neighborhood choice models, leveraging exogenous variation from the Neighborhood Choice Program and panel data with neighborhood preferences before and after information. Third, we estimate a model that allows us to quantify the relative importance of the different channels and explore counterfactuals.

Our findings add to a growing literature on the impacts of the pandemic and work-from-home on real estate markets. Ramani and Bloom (2021) find evidence of real estate demand moving away from dense central business districts towards lower density suburban zipcodes. Gupta et al. (ming) also find a pandemic-induced flattening of the rent gradient, in line with a reallocation away from the city center. Davis, Ghent, and Gregory (2022) study labor market implications by examining complementarity between WFH and work at the office and implications for productivity and city structure.[8]

Our work also adds to the literature on the importance of information and networks on housing markets. Gao, Sockin, and Xiong (2021) establish how learning about neighborhood conditions influence housing cycles, distorting migration into the neighborhood and the supply of capital and labor. Similarly, Kurlat and Stroebel (2015) show how neighborhood characteristics are a key source of information heterogeneity in determining asset values. There is also a related literature on different dimensions of imperfect information, with applications to asset markets (Kacperczyk, Van Nieuwerburgh, and Veldkamp, 2016), and to household mobility, such as Fujiwara, Morales, and Porcher (2019) and Kosar, Ransom, and Van der Klaauw (ming).

Our research is also part of a body of work that combines surveys, experiments, and structural estimation, such as Galiani, Murphy, and Pantano (2015) who use the moving to opportunity experiment to simulate the effect of changes in the subsidies, Bottan and Perez-Truglia (ming) who apply information surveys to medical students in order to understand the importance of relative income in city-level choices, and Bergman, Chan, and Kapor (2020) who study the role of imperfect information about school quality.

The paper proceeds as follows: We present the survey design in section 2 and descriptive analyses in section 3. In section 4, we describe and estimate the neighborhood choice model and perform counterfactuals. Section 5 concludes the paper.

[8] Bloom et al. (2015), Emanuel and Harrington (2021), and Barrero et al. (2021) examine effects of WFH on labor markets.

2 Information Survey and Data

2.1 Neighborhood Choice Program

We partnered with a large professional school in the East Coast to design a neighborhood choice program to help students choose where to live after graduation. In our discussions with students and administrators to understand how students chose neighborhoods, many acknowledged that this was a complex decision given the large number of neighborhood characteristics and the large number of neighborhoods to choose from. Four main issues surfaced in qualitative interviews. First, students mentioned anxiety about cost of living due to high housing costs in many cities. Second, students highly value access to the professional and social network of fellow students and alums, and wanted to preserve that network after graduation (Shue, 2013). Third, students had unequal access to information and relied on limited networks to obtain neighborhood information.[9] Finally, students believed in a social norm where the same-school network tended to live in neighborhoods with high cost of living, inadvertently leading students to choose expensive neighborhoods upon graduation. To address these concerns, we launched a Neighborhood Choice Program in 2019 to provide all students with information about cost of living and the same-school network shares, to help them choose neighborhoods. In 2021, we expanded the information to include the size of rental units in neighborhoods and added more questions related to WFH and commute.

In April 2021, we emailed all students in the graduating cohort to introduce the neighborhood choice program. April is the ideal timing because it is about a month before graduation and a majority of the students already have a job and know which city they want to move to - and at the same time most students had just started the process of searching for housing in their new destination. The program provided neighborhood information in two ways. Students would first access an online survey which provided information about the neighborhoods in their preferred metropolitan area in the United States and also asked basic questions about their neighborhood choices. Students were given a $25 Amazon gift card to encourage them to complete the survey. After completing the survey, students could access a mapping tool which provides the same information at a more granular spatial resolution for all metropolitan areas available in our data.

[9] In our initial survey, 88% of students report speaking to fewer than four contacts about their search process, 95% connected with fewer than four contacts online or through social media.

2.2 Neighborhood Information

Below, we describe how we defined neighborhoods, and explain the type of information provided to survey takers.

Neighborhood names. We begin by selecting the top 18 most popular labor markets (MSAs) in the United States, based on the current residence of all school cohorts who graduated from 2010 to 2019. Other MSAs with a small number of graduates were not used in order to preserve data confidentiality. We then split each MSA into a set of comprehensive yet parsimonious neighborhoods. As a baseline, we used shapefiles from Zillow which classifies the urban core into neighborhoods. In places without Zillow neighborhoods - usually suburban areas - each county would be a neighborhood. In some instances where we had to combine neighborhoods to reduce the total number of choices in each MSA, we joined adjacent neighborhoods with similar levels of college graduates, based on the census, and reported both names in the survey. To generate a list of neighborhood names that students would be familiar with, we relied on Google Trends data. In particular, when there were multiple ways to identify a location, we chose the most popular name according to Google Trends. Ultimately, we ended up with around 20 neighborhoods across all MSA's. Columns 1 and 2 of Table 1 lists the MSA's and the number of neighborhoods in each MSA.

Cost of Living. We focus on monthly rents instead of housing prices because most students occupy rented residences in their first few years after graduation. We obtained monthly rents from Zillow, which publicly provides a monthly rent index for an average home in each neighborhood. We chose to present information about the average 2020 rent using all months in order to mitigate outliers coming from solely using one or two months of data.

Same-School Network. To describe the same-school network comprehensively, we first obtained proprietary administrative data with the current street addresses of all recent graduates of the school. The school utilizes various sources to ensure that the addresses are current and accurate. We focused on the cohorts who graduated between 2010 and 2019 - this aggregation was required in order to preserve student privacy. Additionally, survey respondents do not have access to the total number of individuals living in individual neighborhoods. Instead, we present them with same-school network shares in neighborhood j in MSA m (N_{jm}), by dividing the total number of graduates currently living in neighborhood j by the total number of graduates living in the MSA.

Size of Rental Unit. There is no comprehensive data on the size of rental units. To provide our best guess of the average size of rental properties in each neighborhood, we begin with 2015-2019 data from the American Community Survey (ACS) which reports the average number of rooms and number of bedrooms for rental properties by zipcode. We then estimated a conversion factor from rooms to square feet using 2017 property tax records data from Corelogic. Specifically, for each MSA, we estimated a separate regression linking total square footage of a property to the number of rooms, the number of bedrooms, and a constant. We also checked that the results are robust to different specifications (logs and quadratics). Next, we use these estimated coefficients for each MSA to convert size (in number of rooms and bedrooms from the ACS) to area in square footage for the average rental property for each zipcode. Finally, we aggregate to the neighborhood level.

2.3 Survey Design

We designed the neighborhood choice survey to collect unique information about how individuals make choices before and after receiving information about neighborhoods. Each student would first choose a metropolitan area in which they were planning to live upon graduation. They had to choose among 18 MSA's or select the option "None of the above". In this latter case students self-reported the name of another city in the US or abroad - 29 students picked that option and we exclude them from the remaining analysis. Another 310 students chose one of the top MSA's. Column 3 of Table 1 shows that forty percent of these respondents selected New York, followed by San Francisco (16.9%), DC (7.9%), Philadelphia (7.1%), and Boston (6.4%).

Around 40% of the graduating cohort of students participated in the neighborhood choice program (relatively high compared to 66% of past cohorts who live in these top MSA's). Panel A of Table A1 in the Appendix reports the summary statistics of our estimation sample. The average age is 29, half of them are female, 17% are first-generation or part of an under-represented minority, 14% are married or have children, and 24% are international students (by citizenship). Survey takers are similar to non-respondents by age and marital/family status, 13 p.p. more likely to be female, 6 p.p. less likely to be first-generation or under-represented minority (URM) students, and 9 p.p. less likely to be international (likely because international students are less likely to want to live in the United States). Some of these differences reflect compositional effects associated with choosing the top MSA's.[10] Reassuringly, our results are similar with and without these demographic controls, suggesting that these compositional differences are not driving our main results.

To study neighborhood preferences, we ask students to rank up to ten neighborhoods in their preferred metropolitan area, allowing them to create their own neighborhood consideration sets. Only for these neighborhoods we then ask students to provide their best estimates of the same-school network shares, monthly rent and size for the average rental home. Next we displayed the program information about the rent, size, and network shares in *all* neighborhoods in that MSA. Here, each respondent would also see in a figure how this new information compared with her own unique estimates.

After presenting the information, we asked students to re-rank up to ten neighborhoods. This page looks identical to the pre information stage, except with the new information just so students would not need to scroll back in order to check the data. Students could choose from a menu of all the neighborhoods along with information about the rent, size, and network shares. We did not pre populate this page with their prior rankings so as not to prime their post information choices. Finally, at the end of the survey we also asked some questions about whether and why they thought the information influenced their neighborhood choices, and other factors related to WFH and their search processes. The Survey Appendix provides more details about the survey questions.

The survey was designed to be short. The median student completed the survey in 7 minutes, spending one minute on the pre information ranking of neighborhoods, 97 seconds to estimate the rent, size, and network shares in their chosen neighborhoods, 47 seconds to read about the information for the full set of neighborhoods, and another one minute to re-rank neighborhoods.

The neighborhood choice program did not include a control group since the objective of the program was to promote equal access to information. Having a control group which did not receive the same information intervention would raise equity concerns.[11] In Ferreira and Wong (2020), we demonstrate that our findings are robust when we compare location decisions for participants in the program to students in the same cohort who did not participate. Nonetheless, this is not as relevant to our estimation because the key identifying variation in our model comes from comparing neighborhood rankings by the same

[10]For example, the international difference halves once we also consider the 29 respondents who indicated preferences for other cities.

[11]Given the nature of the information intervention, it would also be difficult to prevent treated students from sharing information with those assigned to the control group

individual before and after the information intervention.

Table A2 presents the neighborhood characteristics of all the 334 neighborhoods, as well as the 209 neighborhoods considered in the pre information rankings and the 198 neighborhoods considered post information. Interestingly, individuals only rank an average of 5 neighborhoods, even though they can rank up to 10 neighborhoods in each MSA. The considered neighborhoods have higher monthly rent ($3,200) relative to $2,500 for all neighborhoods. The ranked neighborhoods also have higher network shares (6.79 percent post information) compared to all the neighborhoods (5.09 percent). By contrast, rental units are smaller (808 sf post information) compared to 968 for all neighborhoods. This likely reflects the popularity of neighborhoods in city centers which tend to be smaller, more expensive, and have more alumni.

Mapping and Search. Upon completion of the survey, students were directed to a restricted-access mapping service with the information about rents, sizes, and network shares at an even more granular geographic resolution for all metropolitan areas available in our data. These interactive maps require students to click in neighborhood geographies in order to access the relevant information. The maps became permanently available in order to help students during their housing and neighborhood search, and we collected data on map clicks.

3 Reduced Form Results

3.1 Heterogeneity in Neighborhood Knowledge

Panel A of Figure 1 presents a kernel density of individuals' estimates of rents minus the neighborhood rents from Zillow. This figure includes 1,841 neighborhood-by-individual choices where we have rent estimates for neighborhoods considered in the pre period by 310 individuals. On average, individuals under-estimate monthly rent by $278 relative to the average monthly rent of $2,500. But there is a fair amount of heterogeneity: about 57% of the choices are underestimates and about 43% of the choices are over-estimates. Panel B presents heterogeneity in knowledge about network shares, with 72% of the estimates being over-estimates. On average, network shares are over-estimated by 7.6 percentage points (p.p.) - a large difference relative to a mean of 5 percent for all neighborhoods. For rental square footage (Panel C), individuals over-estimate it by 72 square feet (sf) on average, relative to a mean of 968 sf.

3.2 Rankings Before and After Information

Table 2 presents a cross-tabulation to compare characteristics of neighborhoods ranked in the top 10 before and after information. The sample includes 6,941 potential individual-by-neighborhood choices. The diagonal entries in Panel A report the average rent of neighborhoods that are always and never ranked top 10 before and after information. As expected, high quality neighborhoods that are popular and always-top are more expensive ($3,238) relative to the never-top neighborhoods that are cheaper ($2,785). 20% of the 6,941 potential choices are always ranked in the top 10 compared to 70% that are never considered. These persistent preferences for always- and never-top neighborhoods are quite heterogeneous across neighborhoods as well. Out of 334 potential neighborhoods, 182 are considered always-top and 318 are considered never-top by someone.

Next, the off-diagonal entries compare average rents for neighborhoods that were switched in and out of individuals' top 10 rankings after receiving information. Interestingly, neighborhoods that are switched in are cheaper ($2,937) relative to neighborhoods that are switched out ($3,315). Around 10% of the choices were switches made by around 200 students and more than 150 neighborhoods. There are more switches (21%) if we consider intensive margin changes in ranks, instead of extensive margin (top-10 or not).

Panel B reports analogous patterns for network shares, with always-top neighborhoods having high network shares (6.73%), never-top neighborhoods having low shares (3.67%). We also see individuals switch in neighborhoods with high shares (7.25%) and switch out neighborhoods with low shares (5.12%).

Panel C reports more nuanced patterns for the square footage of rental homes. The always-top neighborhoods are smaller (801 sf) and the never-top neighborhoods are larger (934 sf). This is consistent with the results for rents and network shares above since students in this school tend to favor neighborhoods in the city center that are more expensive and have smaller rental units. As we examine the switches, we begin to see evidence of preference for space, with individuals switching in neighborhoods with larger rental units (858 sf) and switching out neighborhoods with smaller units (825 sf).

Overall, the patterns in the cross-tabulation demonstrate the importance of having panel data of choices before and after information. There are persistent preferences for high quality neighborhoods from the always- and never-top choices. There are also important sources of variation from changes in rankings which will help to identify preferences for neighborhood amenities.

Next, Table 3 shows that the cross-tabulation patterns above survive regression analyses with rich controls and individual fixed effects. The full sample includes 6,941 choices made by 310 students in 18 MSA's. The dependent variables are monthly rent from Zillow (columns 1 and 2), same-school network shares from administrative data (columns 3 and 4), and square footage of rental homes (columns 5 and 6). The three key regressors include indicators that are 1 if neighborhood j is switched into the top 10 by individual i post information, as well as indicators for always-top and for never-top neighborhoods. The omitted group comprises neighborhoods that were switched out of the top 10. The odd columns have as controls MSA fixed effects and demographics (age, gender, an indicator each for first-generation/under-represented minority, married or having children, international). The even columns include 310 individual fixed effects and rely on changes in rankings within each individual. Standard errors are clustered at the individual level.

The average rent is lower by \$432 and network shares are higher by 2.2 percentage points for neighborhoods that are switched in post information relative to neighborhoods that are switched out. These magnitudes are large relative to the average for all neighborhoods (\$2500 for monthly rent and 5.1% for network shares). The results are similar with and without individual fixed effects. The estimate for square footage is positive but insignificant (16 sf relative to an average of 968 sf for all neighborhoods).

In addition, Table A3 in the Appendix report stronger reductions for rent (\$668), a smaller increase for network shares (1.4 p.p.) for New York and San Francisco only (Panel A), the two most popular cities which cover slightly more than half of the respondents. Notably, the results for square footage are now significant at the 1% level (34.4 sf). These are large relative to sample averages. For other cities (Panel B), we find more significant results for network shares (3.5 p.p. relative to a mean of 5.2%) whilst the effects for rent and square footage become insignificant.

Heterogeneity by work-from-home and commuting. Table 4 presents heterogeneity analysis across individuals. At the end of our survey, we asked students about their preferences for important neighborhood characteristics. Here, we interact the three key regressors (*SwitchIn*, *AlwaysTop*, *NeverTop*) with an indicator for types of individuals. We report only the *SwitchIn* indicator and its interaction, for simplicity.

Columns 1 to 3 examine the preferences for work from home. Nineteen percent indicated they expect to work from home 0 days, 31% reported 1 day, 25% expected 2 days, 13% reported 3 days, 12% indicated 4 or 5 days. We define a work from home indicator using 2 or more days (the median). Our results are similar using other cutoffs. We do not find significant patterns.

Moreover, we also asked students two questions about their commute preferences. First, we asked how much commute time they would prefer, from less than 15 minutes (8% of respondents), 15 to 30 minutes (47%), 30 to 60 minutes (40%), more than 60 minutes (5%). Second, we asked students how important commute time is when choosing neighborhoods. We constructed an index for commuting using z-scores.[12]

Columns 4 to 6 report that the results are stronger for students who are able to tolerate longer commutes. In particular, they switch in neighborhoods with even lower rent (-235 dollars in addition to -282 for non-commuters), lower network shares (-1.8 p.p. relative to 2.9), and larger rental units (41.5 sf relative to -5.1). These patterns are consistent with potential commuters being even more open to moving out of the city center in search of lower rents and larger units, at the expense of less access to their social and professional network.

Finally, we present heterogeneity by marital and family status. For the 14% of students who are married or have children, we find that they are more likely to switch in neighborhoods with homes larger by 62.7 sf. We do not detect significant differences by age, gender, and minority status.

Evidence from map clicks. Next, Table 5 shows that the patterns we uncover using stated preferences in our survey remain the same using map clicks to track preferences. Upon completion of the survey, respondents are directed to a proprietary map created to help them in their neighborhood search. The map displays more granular neighborhood information with smaller polygons that indicate the rent, network shares, and square footage of each polygon that they click on. We use these annonymized click patterns to define the top 10 choices using the number of clicks in neighborhood j by individual i. Reassuringly, we find similar patterns using map clicks, with individuals switching in neighborhoods that are cheaper (-$210) and have higher network shares (1.7 p.p.). The effect for square footage is positive (12.3) but insignificant. This echoes our earlier research where we document that the stated preferences in our survey are consistent with search behavior using map clicks and actual neighborhood choices upon graduation a year later (Ferreira and Wong, 2020).

[12] We standardized both variables and aligned the signs so that both are higher for students who can tolerate longer commutes. We took the mean of both z-scores and defined an indicator that is 1 if the average z-score is above 0.

3.3 Distance to City Center

So far, we have shown that after the information, students are more likely to switch in neighborhoods with lower rent and larger networks. There is evidence of students switching in neighborhoods with larger rental units, for students who can tolerate longer commutes and for families. These can present opposing forces on distance to the city center. Network preferences favor the city center whilst lower rent and larger units point towards the suburbs.

Table 6 considers what these patterns mean for where students choose to live. We define city center using the City Hall for each MSA. The average neighborhood is 14.7 kilometers from city hall (4.9 km if we restrict to all neighborhoods considered post information). Instead of City Hall, we also considered the centroid of the city based on the current address information of alumni in the past decade. Both distance metrics are very similar (the average neighborhood is 14.6 km away from the alums' centroid versus 14.7 using City Hall). Panel A presents the results using 2021 data for all MSA's (columns 1 and 2) and for New York and San Francisco only (columns 3 and 4).

Interestingly, we find that students are more likely to switch in neighborhoods that are 2.4 km away from the city center relative to the omitted group (neighborhoods that are switched out which are on average 8.3 km from the center). Moreover, it is also instructive to compare this with always-top choices (20% of all choices) which are associated with neighborhoods that are 1.5 kilometer closer to the city center and never-top choices (70%) which are 10.4 kilometer farther away from the center. Overall, this suggests that city centers are still favored, farflung neighborhoods are still less popular. On the margin, though, students are more likely to consider neighborhoods that are slightly out of the city center. These patterns are stronger for New York and San Francisco (switching in of 3.6 km relative to a mean of 11.5 km).

When we repeat the same using 2019 survey data before the pandemic, we do not find significant effects on distance for switches (Panel B). The estimates for *SwitchIn* are smaller in magnitudes (1.3 for all MSA's and 0.9 for New York and San Francisco) and statistically insignificant. The coefficients for always-top and never-top neighborhoods are similar in 2021 and 2019 (-1.4 for always-top and 10.7 for never-top).

The insignifciant effects for the switches are also consistent with our reduced form findings on amenities for 2019. When we repeat our main descriptive analysis (Table 3), we still find qualitatively similar patterns but the magnitudes suggest weaker preferences for amenities associated with city centers. In particular, students still switched in cheaper neighborhoods but the rent effects are weaker ($347 relative to $432) and the network effects are larger (2.5 pp relative to 2.2). The weaker preference for cheap neighborhoods coupled with the stronger preference for network shares can explain why we detect no marginal effects away from the center in our distance analysis for 2019.

Below, we develop a model to properly account for these different channels that have offsetting effects on where households choose to live. This is difficult to parse out using a reduced form analysis only as the changes in discrete neighborhood rankings jointly affect all amenities. To address endogeneity problems when estimating neighborhood choice models, we develop a research design that leverages panel data in rankings and variation from the switchers. In the structural estimation, we will utilize all changes in rankings (21% of our choice data), not just top-10 or not.

4 Neighborhood Choice Model and Counterfactuals

In this section we estimate a neighborhood choice model to achieve two goals. First, recover demand preferences for neighborhood rents and amenities and second, use those parameters in counterfactual exercises to better understand how post COVID preferences impacted residential location decisions. In the model individuals choose neighborhoods among a set of considered neighborhoods, and they have heterogeneous preferences for amenities that are a function of individual characteristics. The model accounts for unobserved amenities, and also for heterogeneity in individual knowledge about amenities and neighborhoods. Upon estimation of this structural model, we proceed with running counterfactuals where we restore preferences for amenities to pre COVID parameters. We then recover how those different preferences lead to different neighborhood choices, focusing on the average distance between place of residence and center city.

Neighborhood Choice. We model the neighborhood location decision of each individual as a discrete choice, following the utility function specification of the random utility models originally developed by McFadden (1973, 1978) and Berry, Levinsohn, and Pakes (1995). The individual i's indirect utility from choosing neighborhood j among J_m neighborhoods in labor market area m is:

$$u_{ijm} = x^k_{jm}\beta^k_i + \xi_{ijm} + \varepsilon_{ijm} = V_{ijm} + \varepsilon_{ijm} \tag{1}$$

where the econometrician observes a limited number of neighborhood features x^k, β^k_i

is a vector of individual preferences for each neighborhood amenity, ξ_{ijm} is unobserved by the econometrician, and capture both unobserved neighborhood amenities (such as quality of sidewalks, number of trees, etc.) and also individual heterogeneity in knowledge about any amenity. Finally, ε_{ijm} is an i.i.d. Type-I extreme value error term that reflects i's idiosyncratic preference for neighborhood j in metropolitan area m. Preferences for each observed amenity are a function of the individual's own demographic attributes z_{id}:

$$\beta_i^k = \beta_o^k + \sum_{d=1}^{D} \beta_d^k z_{id} = \beta_o^k + \beta_{id}^k \tag{2}$$

Credible identification of β_i^k relies on dealing with the correlation between x_{jm}^k and ξ_{ijm}. We follow the research design proposed by Ferreira and Wong (2020) who use proxies for ξ_{ijm} by constructing latent quality indices $g(\widetilde{\xi}_{ijm})$, which we describe below.

Estimation. We follow the neighborhood choice literature and estimate equation 1 in two stages. In the first stage (equation 3), we estimate heterogeneous parameters β_{id} and a set of neighborhood fixed effects δ_{jm}. The key neighborhood amenities we observe include post information network share A, and post information monthly rent P.[13] The preference heterogeneity terms β_{id} for each of the neighborhood amenities follow equation 2 and are a function of observed demographic variables: age, gender, married and/or with children, first-generation or minority status, and citizenship status. In our final 2021 model we also include tolerance for commuting, and number of work-from-home days.

$$u_{ijm} = \delta_{jm} + P_{jm}\beta_{id}^P + A_{jm}\beta_{id}^A + X_{jm}\beta_{id}^X + \xi_{ijm} + \varepsilon_{ijm} \tag{3}$$

In the second stage (equation 4), we decompose the neighborhood fixed effects δ_{jm} to recover mean preferences β_o, and also the metropolitan area effects, μ_m. The unobserved errors in equation 4 include the unobserved average neighborhood quality ξ_{jm} and η_{jm} is an idiosyncratic error term:

$$\delta_{jm} = \mu_m + A_{jm}\beta_o^A + P_{jm}\beta_o^P + X_{jm}\beta_o^X + \xi_{jm} + \eta_{jm} \tag{4}$$

We estimate the neighborhood choice model above using post information data on rankings, network shares, and rents. We assume that individuals will fully update their knowledge of network shares and rents, after the information.

[13] We also include a set of 2010 Census characteristics X, including average income, the share of college graduates, and the non-White share. In a separate specification we include rents normalized by house size.

Latent quality index. Next, we introduce the pre information rankings to construct latent quality indices $g_1(\widetilde{\xi}_{ijm})$ and $g_2(\widetilde{\xi}_{jm})$ to respectively control for unobserved heterogeneity associated with the first stage (ξ_{ijm}) and the second stage (ξ_{jm}).[14] We establish in Ferreira and Wong (2020) that the latent quality index using pre information rankings is sufficient for identification if it exhausts all the information about how ξ_{ijm} influences neighborhood choices in equation 1 (Manski, 1988; Dahl, 2002; Berry and Haile, 2014). Intuitively, the pre information rankings reveal individual i's knowledge about the quality of neighborhood j. Since the time between rankings before and after information is two minutes, it is plausible that these pre information rankings remain relevant proxies of quality post information as well. Changes in rankings reflect new information received about networks and rents, as explained in the previous section.[15] Moreover, we document that our effects are not driven by compositional differences in demograhics associated with these switches.[16]

We estimate equation 3 using a rank-ordered logit model (Beggs et al., 1981) based on the post information neighborhood rankings in each individual's consideration set. For each individual i choosing metropolitan area m, our data reveals:

$$U_{ir_{i1}} > U_{ir_{i2}} > \ldots > U_{ir_{iL_i}} \tag{5}$$

where r_{il} denotes the neighborhood that received post information ranking l by individual i. Each individual could rank between two and ten neighborhoods, and we denote the last ranked neighborhood for each individual as L_i. For each market m, we normalize by setting to zero the utility for the two most expensive neighborhoods in the MSA (our conclusions remain similar if we use other normalizations). Given the extreme value assumption for ε_{ijm}, the probability of individual i choosing a ranking r_i is:

[14]To construct $g_1(\widetilde{\xi}_{ijm})$, we use six categorical variables based on whether each individual pre ranked a neighborhood in her consideration set as top 1, 2, 3, 4, 5, or from 6 to 10. We omit from the estimation a dummy for neighborhoods never ranked in the pre information data. Additionally, we include the average of these pre rank dummies and interact these averages with demographics. We include these heterogeneity terms in equation 3. For the second stage, we construct $g_2(\widetilde{\xi}_{jm})$ using six averages of the rank dummies in equation 4.

[15]The variation from switches is key. Since neighborhood choices are typically stable over time, post information rankings (the dependent variable we use to estimate our discrete choice model) would usually be collinear with pre information rankings (our latent quality index controls), if not for changes in rankings associated with the information shock.

[16]Table A4 in the Appendix compares demographic characteristics between switchers and non-switchers.

$$
\begin{aligned}
\pi_{ir_i} &= P[U_{ir_{i1}} > U_{ir_{i2}} > \ldots > U_{ir_{iL_i}}] &\quad (6)\\
&= \prod_{l=1}^{L_i-1} \frac{\exp(V_{ir_{il}})}{\sum_{h=l}^{L_i} \exp(V_{ir_{ih}})} &\quad (7)
\end{aligned}
$$

We rely on maximum likelihood to estimate the model. The log likelihood function is just the sum of the log of the individual probabilities across all individuals:

$$
\mathscr{L} = \sum_{i=1}^{I} \log \pi_{ir_i} \quad (8)
$$

This first step returns the heterogeneity in preference parameters and the neighborhood fixed effects that maximize the log likelihood function above, i.e., maximize the probability that each individual makes the correct rank ordering of neighborhoods. We estimate the second step using OLS.

Counterfactuals. The model above generates preference parameters for all neighborhood amenities, and can also be used to back out underlying probabilities of each individual choosing a neighborhood as ranking 1, 2, 3, etc. Focusing on neighborhoods chosen as ranking 1, we can write the estimated probability of choosing a neighborhood as:

$$
\hat{\pi_{i1}} = \frac{\exp(\hat{V_{ih}})}{\sum_{l=l}^{L} \exp(\hat{V_{il}})} \quad (9)
$$

The estimated probabilities for each individual sum to one, and every neighborhood in the individual consideration set will have a probability between 0 and 1. Multiplying those probabilities by the actual distance between each neighborhood and the downtown area leads to the average estimated distance from downtown for each individual. We can then average those distances across individuals, which correspond to our baseline counterfactual estimates. Standard errors for these simulated distances are based on Bootstrap samples.

The next set of counterfactual simulations provide new estimated probabilities (and distances to downtown) by substituting the 2021 preferences for neighborhood amenities with the same parameter estimated using our 2019 pre pandemic data. This exercise can be thought of as a partial equilibrium result, as prices and price preferences are not adjusted due to the changes in network share preferences. Stable price preferences are justified by the similar rent estimates reported both in 2019 and 2021. For prices themselves, we explore an alternative counterfactual where we also change prices to reflect 2019 pre pandemic data. Finally, we exploit another two counterfactuals, where we shutdown preferences for work-from-home and commuting.

Estimates. Table 7 shows mean estimated preferences for rents and network shares for models that use 2019 pre COVID pandemic and the post pandemic 2021 data. Models in columns 1 and 2 do not use heterogeneity measures of work-from-home and commuting. We find that while rent elasticities are similar before and after the pandemic (-\$478 in 2019 and -\$454 in 2021), preferences for network shares are smaller in 2021 (0.20) than in 2019 (0.37). Those results corroborate the reduced form effects in section 3, where we find a weaker effect for network shares in 2021 relative to 2019. We then combine rent and network estimates into a measure of willingness to pay. Our MWTP estimate for 2021 is only \$86 per month for a 1 percentage point increase in network shares (relative to a mean and SD of 5 percent), while in 2019 that same MWTP was \$171.[17]

Next, we augment our 2021 model estimates by also accounting for heterogeneity in WFH and commute preferences in the first stage (equation 3). For WFH, we add an indicator that is 1 if individual *i* expects to work from home 2 or more days (median). For commute, we have an indicator that is 1 if individual *i* can tolerate a maximum commute time more than 30 minutes (45% of the respondents). The utility estimates for rent remains similar (-\$459) and the network estimates are slightly larger (0.25), resulting in a higher MWTP (\$109), which still remains a fair amount smaller than the MWTP in 2019. While the model estimates heterogeneity in preferences for those parameters, they were not statistically different from zero, which is not a surprise given the limited number of individuals in the sample (Table A5 in the Appendix). We also estimated a second set of models using rent per square foot, as opposed to only rents, and the results are displayed in Table A6. Our conclusion remains similar with MWTP being more than twice in 2019 relative to 2021, but we lose statistical precision. This is expected since our information program was based on showing average rents not rents per square foot in our sample.

We turn now to our counterfactual exercises to assess how these different factors ultimately affect neighborhood choices. Table 8 shows model-implied estimated distances from the city center for the standard model first (columns 1 and 2) followed by the richer

[17]The implied MWTP was calculated using $-\frac{\beta_o^A}{\beta_o^P}$, obtained from the second stage decomposition of the neighborhood fixed effects, equation 4.

model with WFH and commute preferences (columns 3 and 4). For each counterfactual, we compare the model-implied baseline distances (first column) to simulated distances from a counterfactual. Each row presents a different counterfactual. Here, we focus on the distances for the top-ranked neighborhoods (sample average of 5 km in 2021 and 4.8 km in 2019 for all MSA's). We report bootstrapped standard errors in parentheses.

In the first row, we show that post COVID changes in network preferences is quantitatively important. Our standard model (column 2 of Table 7) imply that top-ranked neighborhoods are on average 4.83 kilometers away from the city center. In our counterfactual, we restore the importance of network instead, by using preferences from our 2019 estimates (0.37 instead of 0.20), and keeping everything else the same. Doing so results in a shorter distance of 4.17 kilometers. This is a non-trivial displacement in location of residential choices and is explained by the fact that young professionals have larger networks closer to city center. Any increase in preferences for networks would result in those young professionals living closer to city center. This is likely conservative given our focus on top-ranked neighborhoods (which tend to be more persistent) when calculating the average distance. This spatial reallocation is even more pronounced for New York and San Francisco (4.91 km for the baseline model versus 3.91 km for the counterfactual with stronger network preferences).

While the counterfactual simulation above is a partial equilibrium result in nature - we are only simulating the behavior of a limited set of movers, that are unlikely to change market outcomes - we also run a second counterfactual where neighborhood rents completely revert to 2019 levels (in addition to changing preferences for networks). The counterfactual distance is only slightly higher (4.19 km instead of 4.17 km).

In the last two rows, we assess the importance of WFH and commute, respectively. Column 3 presents model-implied distances using preference estimates from the richer model (column 3 of Table 7). When we shut down WFH preference heterogeneity (row 3), the distance falls slightly to 4.65 km from 4.77 km. More strikingly, we find larger effects when we shut down commute heterogeneity (4.77 km for the baseline and 4.49 km for the counterfactual). It is intuitive that short commute preferences suggest a more pronounced movement away from the suburbs, in line with our reduced form findings of stronger effects by commuting preferences. WFH, on the other hand, does not reveal to be an important factor in the location decisions in our sample of graduates.

5 Conclusion

Where households choose to live has direct implications for the spatial distribution of economic activity and the value of real assets. The COVID-19 pandemic has potentially led many households to rethink how they value different types of amenities, especially potential movers such as the graduating students who participated in the 2021 Neighborhood Choice Program. We show that changes in rankings after information favor neighborhoods with larger networks, lower rents, and are farther from the city center. Interestingly, we do not detect movement away from cities for the 2019 cohort, consistent with the stronger preferences for networks and weaker preferences for cheap neighborhoods.

To quantitatively assess the importance of the different factors, we develop and estimate a neighborhood choice model using rich panel data in neighborhood rankings. Our research design utilizes pre information rankings as proxies to control for unobserved neighborhood quality, overcoming critical identification challenges associated with endogenous amenities. Our preference estimates are stable for marginal utility of rents but the marginal utility for networks weakens significantly from 2019 to 2021, implying a markedly lower MWTP for networks post pandemic. Accordingly, counterfactual exercises suggest that changes in network preferences leads to a sizable movement away from the city center to neighborhoods just outside. Heterogeneity for commuting and WFH are less relevant.

References

Bailey, M., R. Cao, T. Kuchler, J. Stroebel, and A. Wong (2018). Social Connectedness: Measurement, Determinants, and Effects. *Journal of Economic Perspectives 32*(3), 259–80.

Baldauf, M., L. Garlappi, and C. Yannelis (2020). Does Climate Change Affect Real Estate Prices? Only If You Believe in It. *The Review of Financial Studies 33*(3), 1256–1295.

Barrero, J. M., N. Bloom, and S. J. Davis (2021). Why Working from Home Will Stick. Technical report, National Bureau of Economic Research.

Baum-Snow, N. and F. Ferreira (2015). Causal Inference in Urban and Regional Economics. In G. Duranton, J. V. Henderson, and W. C. Strange (Eds.), *Handbook of Regional and Urban Economics Volume 5*, Chapter 1. Elsevier Science.

Bayer, P., F. Ferreira, and R. McMillan (2007). A Unified Framework for Measuring Preferences for Schools and Neighborhoods. *Journal of Political Economy 115*(4), 588–638.

Bayer, P., S. Ross, and G. Topa (2008). Place of Work and Place of Residence: Informal Hiring Networks and Labor Market Outcomes. *Journal of Political Economy 116*(6), 1150–1196.

Beggs, S., S. Cardell, and J. Hausman (1981). Assessing the Potential Demand for Electric Cars. *Journal of Econometrics 17*(1), 1–19.

Bergman, P., E. W. Chan, and A. Kapor (2020). Housing Search Frictions: Evidence from Detailed Search Data and a Field Experiment. Working Paper 27209, National Bureau of Economic Research.

Berry, S. T. (1994). Estimating Discrete-Choice Models of Product Differentiation. *The Rand Journal of Economics 25*(2), 242–262.

Berry, S. T. and P. A. Haile (2014). Identification in Differentiated Products Markets Using Market Level Data. *Econometrica 82*(5), 1749–1797.

Berry, S. T., J. Levinsohn, and A. Pakes (1995). Automobile Prices in Market Equilibrium. *Econometrica 63*(4), 841–890.

Bloom, N., J. Liang, J. Roberts, and Z. J. Ying (2015). Does Working from Home Work? Evidence from a Chinese Experiment. *The Quarterly Journal of Economics 130*(1), 165–218.

Bottan, N. L. and R. Perez-Truglia (Forthcoming). Choosing Your Pond: Location Choices and Relative Income. *The Review of Economics and Statistics*.

Dahl, G. B. (2002). Mobility and the Return to Education: Testing a Roy Model with Multiple Markets. *Econometrica 70*, 2367–2420.

Davis, M. A., A. C. Ghent, and J. Gregory (2022). The Working-from-Home Technology Boon and Its Consequences. Working Paper No. 28461, National Bureau of Economic Research.

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

Emanuel, N. and E. Harrington (2021). "Working" Remotely? Selection, Treatment and Market Revision of Remote Work. Working paper.

Ferreira, F. and M. Wong (2020). Estimating Preferences for Neighborhood Amenities Under Imperfect Information. Working Paper 28165, National Bureau of Economic Research.

Fujiwara, T., E. Morales, and C. Porcher (2019). A Revealed-Preference Approach to Measuring Information Frictions in Migration. Working paper.

Galiani, S., A. Murphy, and J. Pantano (2015). Estimating Neighborhood Choice Models: Lessons from a Housing Assistance Experiment. *American Economic Review 105*(11), 3385–3415.

Gao, Z., M. Sockin, and W. Xiong (2021). Learning about the Neighborhood. *The Review of Financial Studies 34*(9), 4323–4372.

Gupta, A. (2019). Foreclosure Contagion and the Neighborhood Spillover Effects of Mortgage Defaults. *The Journal of Finance 74*(5), 2249–2301.

Gupta, A., V. Mittal, J. Peeters, and S. Van Nieuwerburgh (Forthcoming). Flattening the Curve: Pandemic-Induced Revaluation of Urban Real Estate. *Journal of Financial Economics*.

Kacperczyk, M., S. Van Nieuwerburgh, and L. Veldkamp (2016). A Rational Theory of Mutual Funds' Attention Allocation. *Econometrica 84*(2), 571–626.

Kosar, G., T. Ransom, and W. Van der Klaauw (Forthcoming). Understanding Migration Aversion Using Elicited Counterfactual Choice Probabilities. *Journal of Econometrics*.

Kurlat, P. and J. Stroebel (2015). Testing for Information Asymmetries in Real Estate Markets. *The Review of Financial Studies 28*(8), 2429–2461.

Manski, C. F. (1988). Identification of Binary Response Models. *Journal of the American Statistical Association 83*(403), 729–738.

McFadden, D. (1973). Conditional Logit Analysis of Qualitative Choice Behavior. In P. Zarembka (Ed.), *Frontiers in Econometrics*. New York: Academy Press.

McFadden, D. (1978). Modelling the Choice of Residential Location. In A. Karlqvist (Ed.), *Spatial Interaction Theory and Planning Models*. New York: North-Holland.

Miller, S. and C. K. Soo (2021). Do Neighborhoods Affect the Credit Market Decisions of Low-Income Borrowers? Evidence from the Moving to Opportunity Experiment. *The Review of Financial Studies 34*(2), 827–863.

Moretti, E. (2013). *The New Geography of Jobs*. Mariner Books.

National Center for Education Statistics (2020). Digest of Education Statistics.

Ramani, A. and N. Bloom (2021). The Donut Effect of COVID-19 on Cities. Technical report, National Bureau of Economic Research. Working Paper No. 28876, National Bureau of Economic Research.

Shue, K. (2013). Executive Networks and Firm Policies: Evidence from the Random Assignment of MBA Peers. *Review of Financial Studies 26*(6), 1401–1442.

Wiswall, M. and B. Zafar (2014). Determinants of College Major Choice: Identification using an Information Experiment. *The Review of Economic Studies 82*(2), 791–824.

Wong, M. (2013). Estimating Ethnic Preferences Using Ethnic Housing Quotas in Singapore. *Review of Economic Studies 80*(3), 1178–1214.

Tables and Figures

Table 1: Number of Neighborhoods and Percent of Respondents by MSA

MSA	Number of Neighborhoods	Percent of Respondents
Atlanta	19	1.3
Austin	14	1
Boston	21	6.8
Chicago	22	1.9
Washington, DC	20	7.7
Dallas	18	1.3
Denver	19	1.9
Houston	15	1
Los Angeles	22	4.5
Miami	20	2.6
Minneapolis	17	.3
New York	25	39.4
Philadelphia	20	8.1
San Diego	20	.3
San Francisco	22	16.5
San Jose	18	1
Seattle	22	4.5
Total	334	100

Notes: Top 17 MSA's in the neighborhood choice program with the number of neighborhoods and percent of respondents for each MSA. We also offered Bridgeport but no respondents chose it.

Table 2: Comparing Neighborhoods Pre and Post Information

Panel A: Rent

		Pre: Yes	Pre: No
Post	Yes	Always Top 10 $3,238	Switch In $2,937
	No	Switch Out $3,315	Never Top 10 $2,785

Panel B: Network

		Pre: Yes	Pre: No
Post	Yes	Always Top 10 6.73%	Switch In 7.25%
	No	Switch Out 5.12%	Never Top 10 3.67%

Panel C: Square Footage

		Pre: Yes	Pre: No
Post	Yes	Always Top 10 801 sqft.	Switch In 858 sqft.
	No	Switch Out 825 sqft.	Never Top 10 934 sqft.

Notes: Panel A reports the average Zillow rent of four types of neighborhoods in the cross-tabulation. The diagonal cells include neighborhoods that are always and never ranked top ten by a given individual pre and post information. The off-diagonal cells report average Zillow rent for neighborhoods that are switched in or out of the top ten. Panels B and C repeat the same for network shares and rental home size, respectively.

Table 3: Rent, Network, and Size Before and After Information

Dependent variable:	**Rent**		**Network**		**Square Footage**	
	(1)	(2)	(3)	(4)	(5)	(6)
Switch in	-397***	-432***	2.0***	2.2***	15.1	16.0
	(70)	(76)	(0.4)	(0.5)	(11.9)	(12.5)
Always	-100*	-133**	1.6***	1.7***	-29.5***	-29.0***
	(57)	(67)	(0.2)	(0.3)	(6.2)	(7.1)
Never	-591***	-656***	-1.5***	-1.6***	102.6***	112.2***
	(53)	(61)	(0.2)	(0.2)	(5.7)	(6.4)
N	6,941	6,941	6,941	6,941	6,941	6,941
R-squared	0.42	0.42	0.13	0.14	0.34	0.35
MSA FE	Y	N	Y	N	Y	N
Demographics	Y	N	Y	N	Y	N
Individual FE	N	Y	N	Y	N	Y

* 0.10 ** 0.05 *** 0.01

Notes: OLS regressions including the full set of 6,941 neighborhood-by-individual level choices. The dependent variables are monthly Zillow rent in dollars (columns 1 and 2), network shares (columns 3 and 4), and square footage of rental homes (columns 5 and 6). The key regressors include an indicator each for neighborhoods that are always and never ranked in the top ten consideration set by a given individual before and after information, as well as a third indicator for neighborhoods that were switched into the top ten. The omitted group includes neighborhoods switched out of the top ten. Odd columns have MSA fixed effects and demographic controls, including age, and a dummy for female, married or with children, under-represented minority and first-generation, international. Even columns include 310 individual fixed effects. Standard errors are clustered by individuals.

Table 4: Heterogeneity Analysis for Switched Neighborhoods

HTE:	Work from home			Commute			Married with Kids		
Dependent variable:	**Rent**	**Network**	**Square Footage**	**Rent**	**Network**	**Square Footage**	**Rent**	**Network**	**Square Footage**
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Switch in	-427***	2.4***	9.3	-282***	2.9***	-5.1	-427***	1.9***	6.5
	(95)	(0.6)	(17.4)	(91)	(0.6)	(16.4))	(74)	(0.4)	(12.5)
Switch in * HTE	60	-0.7	11.8	-235*	-1.8**	41.5*	147	0.6	62.7*
	(139)	(0.8)	(23.8)	(134)	(0.8)	(23.7))	(213)	(1.4)	(36.3)
N	6,941	6,941	6,941	6,941	6,941	6,941	6,941	6,941	6,941
R-squared	0.42	0.14	0.34	0.42	0.14	0.34	0.42	0.14	0.34
MSA FE	Y	Y	Y	Y	Y	Y	Y	Y	Y
Demographics	Y	Y	Y	Y	Y	Y	Y	Y	Y

* 0.10 ** 0.05 *** 0.01

Notes: Heterogeneity analysis extending OLS regressions in odd columns of Table 3 by adding interactions of various individual types with the three key regressors (*SwitchIn*, *Always*, and *Never*). For simplicity, we report the coefficients for *SwitchIn* and its interaction only. Columns 1 to 3 examine work-from-home using an indicator that is 1 for individuals who expect to work 2 or more days at home. Columns 4 to 6 explore commuting using an index constructed using z-scores so that higher corresponds to tolerance for longer commutes (see Footnote 12). Finally, we examine family status using an indicator that is 1 for individuals who are married or have children. Standard errors are clustered by individuals.

Table 5: Robustness using Map Clicks

Dependent variable:	**Rent**		**Network**		**Square Footage**	
	(1)	(2)	(3)	(4)	(5)	(6)
Switch in	-190**	-210**	1.4***	1.7***	13.8	12.3
	(86)	(102)	(0.4)	(0.4)	(12.7)	(14.9)
Always	-32	-32	2.2***	2.4***	-49.7***	-55.3***
	(76)	(86)	(0.3)	(0.4)	(10.4)	(11.5)
Never	-546***	-582***	-1.9***	-1.9***	109.7***	116.0***
	(72)	(78)	(0.3)	(0.3)	(8.4)	(9.1)
N	2,394	2,394	2,394	2,394	2,394	2,394
R-squared	0.41	0.41	0.15	0.16	0.35	0.37
MSA FE	Y	N	Y	N	Y	N
Demographics	Y	N	Y	N	Y	N
Individual FE	N	Y	N	Y	N	Y

* 0.10 ** 0.05 *** 0.01

Notes: Repeats Table 3 but using the number of map clicks in each neighborhood to define top ten neighborhoods. Standard errors are clustered by individual.

Table 6: Implications for Distance from City Center

Dependent variable:	**Distance from City Center**			
Sample:	**All**		**New York/ San Francisco**	
	(1)	(2)	(3)	(4)
Panel A: 2021				
Switch in	2.2**	2.4**	3.6***	3.6***
	(1.0)	(1.1)	(1.1)	(1.2)
Always	-1.7***	-1.5**	-0.6	-0.4
	(0.6)	(0.6)	(0.6)	(0.7)
Never	9.5***	10.4***	10.0***	10.9***
	(0.6)	(0.6)	(0.6)	(0.7)
N	6,941	6,941	4,172	4,172
R-squared	0.15	0.16	0.10	0.11
Panel B: 2019				
Switch in	1.1	1.3	0.7	0.9
	(0.8)	(0.9)	(0.9)	(1.0)
Always	-1.6***	-1.4***	-1.2**	-1.1
	(0.5)	(0.6)	(0.6)	(0.7)
Never	9.8***	10.7***	9.7***	10.5***
	(0.5)	(0.6)	(0.7)	(0.7)
N	6,993	6,993	4,761	4,761
R-squared	0.14	0.15	0.11	0.12
MSA FE	Y	N	Y	N
Demographics	Y	N	Y	N
Individual FE	N	Y	N	Y

* 0.10 ** 0.05 *** 0.01

Notes: Repeats Table 3 with distance from the city center (City Hall) as the dependent variable. Panel A uses the 2021 data whilst Panel B uses data from the 2019 Neighborhood Choice Program. Columns 1 and 2 include all MSA's and columns 3 and 4 include New York and San Francisco only. Standard errors are clustered by individual.

Table 7: Neighborhood Choice Model Estimates

Cohort Year:	**2019**	**2021**	**2021**
	(1)	(2)	(3)
Network	0.37***	0.20**	0.25***
	(0.09)	(0.08)	(0.08)
Rent	-478***	-454***	-459***
	(156)	(146)	(156)
Implied MWTP	**171****	**86****	**109****
	(71)	(43)	(47)
MSA FE	Y	Y	Y
Census Characteristics	Y	Y	Y
Latent Quality Indices	Y	Y	Y
Work-from-Home, Commute	N	N	Y

Notes: Mean preference estimates of the neighborhood choice model using neighborhood-by-individual choices post information. The first stage involves a rank-ordered Logit model using post information neighborhood rankings, neighborhood fixed effects, neighborhood amenities (rent, network shares, as well as Census characteristics - average income, share college graduates, average non-White share). The second stage represents a decomposition of the neighborhood fixed effects from the first stage, including MSA fixed effects. All columns include MSA fixed effects and the latent quality indices constructed using pre information rankings (six indicators for neighborhoods ranked 1, 2, 3, 4, 5, above 6, respectively), their interactions with demographics. Columns 1 and 2 use 2019 and 2021 data respectively for a standard model with five demographic attributes (age, gender, minority status, family status, international). Column 3 repeats the 2021 estimation and also adds two dummies for work-from-home (2 days or more) and for tolerance of long commutes (more than 30 minutes). For simplicity, we only report marginal utility estimates for networks and rent (and the implied MWTP for networks) from the second stage decomposition of neighborhood fixed effects. Standard errors calculated using the Delta method.

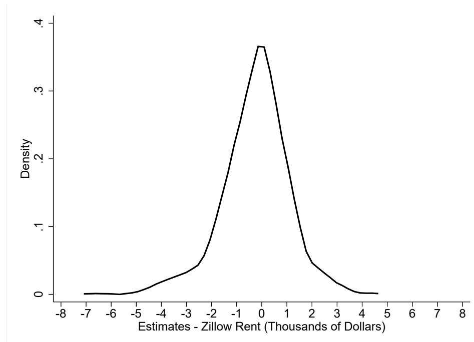
Table 8: Model-Implied and Counterfactual Distances from City Center

	Standard Model		**Adding WFH, Commute**	
	Baseline	**Counterfactual**	**Baseline**	**Counterfactual**
	(1)	(2)	(3)	(4)
Restoring network preference to 2019 estimate	4.83	4.17	4.77	4.02
	(0.42)	(0.55)	(0.42)	(0.66)
Restoring network preference and neighborhood rents to 2019	4.83	4.19	4.77	4.04
	(0.42)	(0.55)	(0.42)	(0.66)
Shutting down heterogeneity by work-from-home			4.77	4.65
			(0.42)	(0.68)
Shutting down heterogeneity by commuting			4.77	4.49
			(0.42)	(0.68)

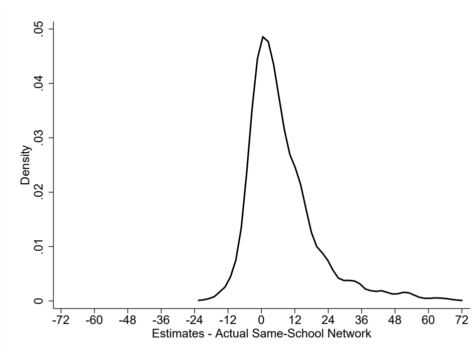
Notes: This table compares distance to the city center for the top-ranked neighborhood, as implied from our model estimates (odd columns) and counterfactuals (even columns). Given our preference estimates, we can calculate the probability that individual i ranks neighborhood j as the favorite. We use these probability weights to calculate the average distance of the top-ranked neighborhood. Columns 1 and 2 correspond to the standard model with five types (Table 7 , column 2) and columns 3 and 4 correspond to the extended model adding heterogeneity by work-from-home and commute (Table 7 , column 3). Each row explores a different counterfactual and the implied average distance under this counterfactual (reported in columns 2 and 4). Bootstrapped standard errors in parentheses.

Figure 1: Heterogeneity in Knowledge of Amenities

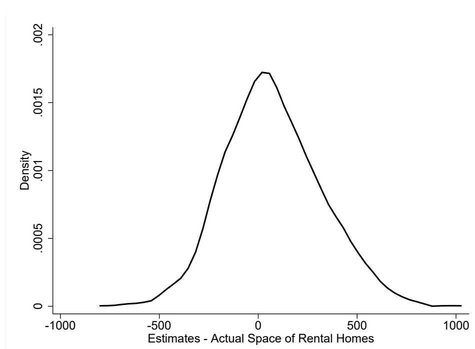
Panel A: Rent



Panel B: Network



Panel C: Square Footage



Notes: Panel A presents kernel density estimates of the difference of individuals' best guesses of neighborhood rent and Zillow rent for each neighborhood ranked in the pre information period. Panels B and C repeat the same for network shares and square footage of average rental home.

6 Online Appendix Tables

Table A1: Characteristics of Respondents and Full Sampling Frame

	All	Respondents	Non-respondents	Difference	p-value
Panel A: Survey respondents in top 18 MSA's					
Female	0.46	0.54	0.41	0.13***	0.00
Age	29.63	29.41	29.78	-0.36**	0.02
Married/Kids	0.13	0.14	0.12	0.01	0.55
First-gen/URM	0.21	0.17	0.23	-0.06**	0.04
International	0.29	0.24	0.33	-0.09***	0.01
N	794	310	484	794	794
Panel B: Survey respondents choosing all cities					
Female	0.46	0.52	0.41	0.11***	0.00
Age	29.63	29.48	29.75	-0.27*	0.08
Married/Kids	0.13	0.14	0.12	0.02	0.39
First-gen/URM	0.21	0.17	0.24	-0.06**	0.03
International	0.29	0.27	0.31	-0.04	0.28
N	794	339	455	794	794

* 0.10 ** 0.05 *** 0.01

Notes: Panels A and B show how the demographics for survey respondents compare to the full student population of 794 students. Panel A includes the 310 students in our primary estimation sample, i.e. those who chose the top MSA's in our program. Panel B includes 339 students who responded to the survey, including 29 who chose other cities not in the Neighborhood Choice Program. The five demographic characteristics include an indicator for females, age, an indicator for married individuals or those who have children, an indicator for first-generation or under-represented minorities, an indicator for international students who are not U.S. citizens.

Table A2: Summary Statistics for All and Considered Neighborhoods

	All			Considered (pre)			Considered (post)		
	N	Mean	SD	N	Mean	SD	N	Mean	SD
Network	334	5.09	5.02	209	6.32	4.16	198	6.79	4.30
Rent (in Dollars)	334	2502.17	925.86	209	3257.68	1136.46	198	3201.89	1070.04
Space (in square feet)	334	968.22	154.88	209	806.74	118.98	198	807.56	118.46
Rent per square foot	334	2.69	1.20	209	4.15	1.56	198	4.07	1.48
Distance from city center	334	14.71	14.68	209	4.79	7.33	198	4.90	7.14
Income (in Thousands)	334	80.91	25.25	209	95.60	30.56	198	94.56	28.13
Bachelor′s Degree+	334	0.46	0.19	209	0.64	0.18	198	0.64	0.17
Minority Share	334	0.37	0.18	209	0.38	0.15	198	0.37	0.15

Notes: Summary statistics for amenities in all (334) neighborhoods in the Neighborhood Choice Program, as well as the 209 (198) neighborhoods considered pre (post) information. The statistics for considered neighborhoods are weighted by number of respondents. The neighborhood amenities include the network share, average monthly rent from Zillow, average size of rental homes in square feet, average income, share of population with a college degree or more, non-White share. The latter three are from the 2010 Census. We also include the distance from the city center (City Hall for each MSA).

Table A3: Rent, Network, and Size of Neighborhoods by MSA

Dependent variable:	**Rent**		**Network**		**Square Footage**	
	(1)	(2)	(3)	(4)	(5)	(6)
Panel A: New York + San Francisco						
Switch in	-630***	-668***	1.2***	1.4***	34.2***	34.4***
	(91)	(98)	(0.4)	(0.4)	(11.6)	(12.1)
N	4,172	4,172	4,172	4,172	4,172	4,172
R-squared	0.18	0.19	0.11	0.12	0.21	0.23
Panel B: Other MSAs						
Switch in	33	31	3.3***	3.5***	-13.1	-11.1
	(77)	(85)	(0.8)	(0.9)	(23.0)	(24.9)
N	2,769	2,769	2,769	2,769	2,769	2,769
R-squared	0.58	0.58	0.15	0.16	0.32	0.33
MSA FE	Y	N	Y	N	Y	N
Demographics	Y	N	Y	N	Y	N
Individual FE	N	Y	N	Y	N	Y

* 0.10 ** 0.05 *** 0.01

Notes: Repeats Table 3 but Panel A restricts the sample to New York and San Francisco only and Panel B include other MSAs. Standard errors are clustered by individual.

Table A4: Compositional Differences for Switchers

	Always in top 10	Switch in	Switch out	Switch in - Switch out
Female	0.00	0.12***	0.00	0.12**
	[0.88]	[0.00]	[0.94]	[0.02]
Age	-0.07	0.07	-0.03	0.10
	[0.33]	[0.61]	[0.84]	[0.58]
Married/Kids	0.00	0.01	0.05	-0.04
	[0.97]	[0.70]	[0.11]	[0.24]
First-gen/URM	0.01	0.02	0.02	0.00
	[0.70]	[0.47]	[0.38]	[0.99]
International	-0.07***	0.01	0.05	-0.04
	[0.00]	[0.70]	[0.14]	[0.45]
Working from home	0.04	-0.06	0.05	-0.11
	[0.43]	[0.59]	[0.66]	[0.46]
Commute	0.07***	0.13**	0.02	0.11
	[0.01]	[0.03]	[0.78]	[0.17]
N	6941	6941	6941	6941

* 0.10 ** 0.05 *** 0.01

Notes: The analysis is similar as the odd columns in Table 3 but the dependent variables are now the demographic controls and individual characteristics. We include MSA fixed effects and standard errors are clustered by individuals.

Table A5: First Stage Heterogeneity Coefficients for Rent and Networks

	2021	
	Rent	**Network**
	(1)	(2)
Age	-6	0.01
	(172)	(0.13)
Married/Kids	-288	0.06
	(267)	(0.16)
International	-98	0.06
	(192)	(0.16)
First-gen/URM	-29	-0.03
	(173)	(0.15)
Female	-60	-0.06
	(134)	(0.13)
High WFH	195	0.01
	(137)	(0.13)
High Commute	-67	0.03
	(579)	(0.42)

Notes: We present first stage heterogeneity coefficients that are interacted with rent and network shares, for the 2021 full model including WFH and commuting (column 3 of Table 7). Bootstrapped standard errors in parentheses.

Table A6: Neighborhood Choice Model Estimates

Cohort Year:	**2019**	**2021**
	(1)	(2)
Network	0.36***	0.17**
	(0.10)	(0.08)
Rent Per Square Foot	-0.27*	-0.27*
	(0.15)	(0.15)
Implied MWTP	**0.38**	**0.18**
	(0.24)	(0.12)
MSA FE	Y	Y
Census Characteristics	Y	Y
Latent Quality Indices	Y	Y

Notes: We repeat our standard model using 2019 and 2021 data (columns 1 and 2 in Table 7), except we now replace rent (in dollars) with rent per square foot.

Table A7: Counterfactual Distances from City Center for New York and San Francisco

	Standard Model		**Adding Work From Home, Commute**	
	Baseline	**Counterfactual**	**Baseline**	**Counterfactual**
	(1)	(2)	(3)	(4)
Restoring network preference to 2019 estimate	4.91	3.91	4.83	3.87
	(0.63)	(0.62)	(0.64)	(0.63)
Restoring network preference and neighborhood rents to 2019	4.91	3.93	4.83	3.89
	(0.63)	(0.61)	(0.64)	(0.64)
Shutting down heterogeneity by work-from-home			4.83	4.99
			(0.64)	(0.65)
Shutting down heterogeneity by commuting			4.83	4.58
			(0.64)	(0.78)

Notes: Bootstrapped standard errors in parentheses.

Survey Appendix

Figure A1: Pre Information Choice Set

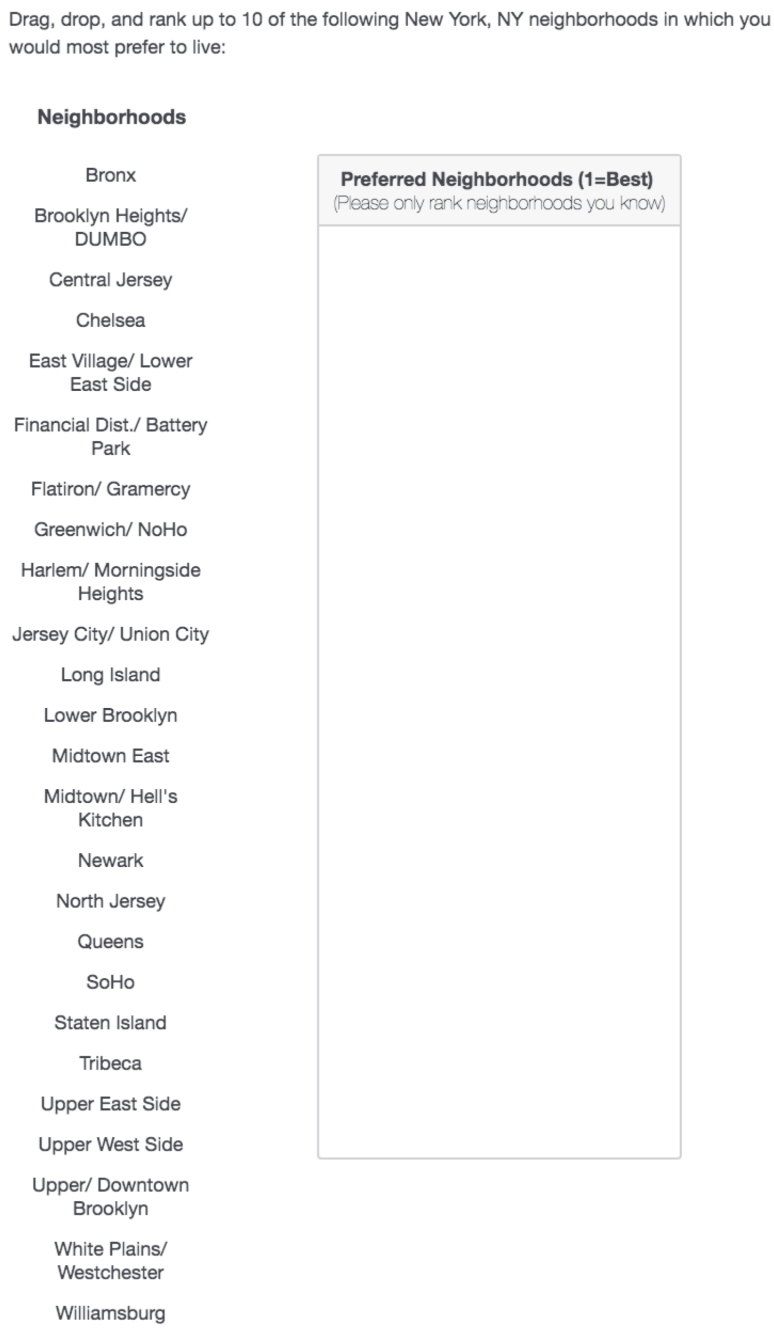
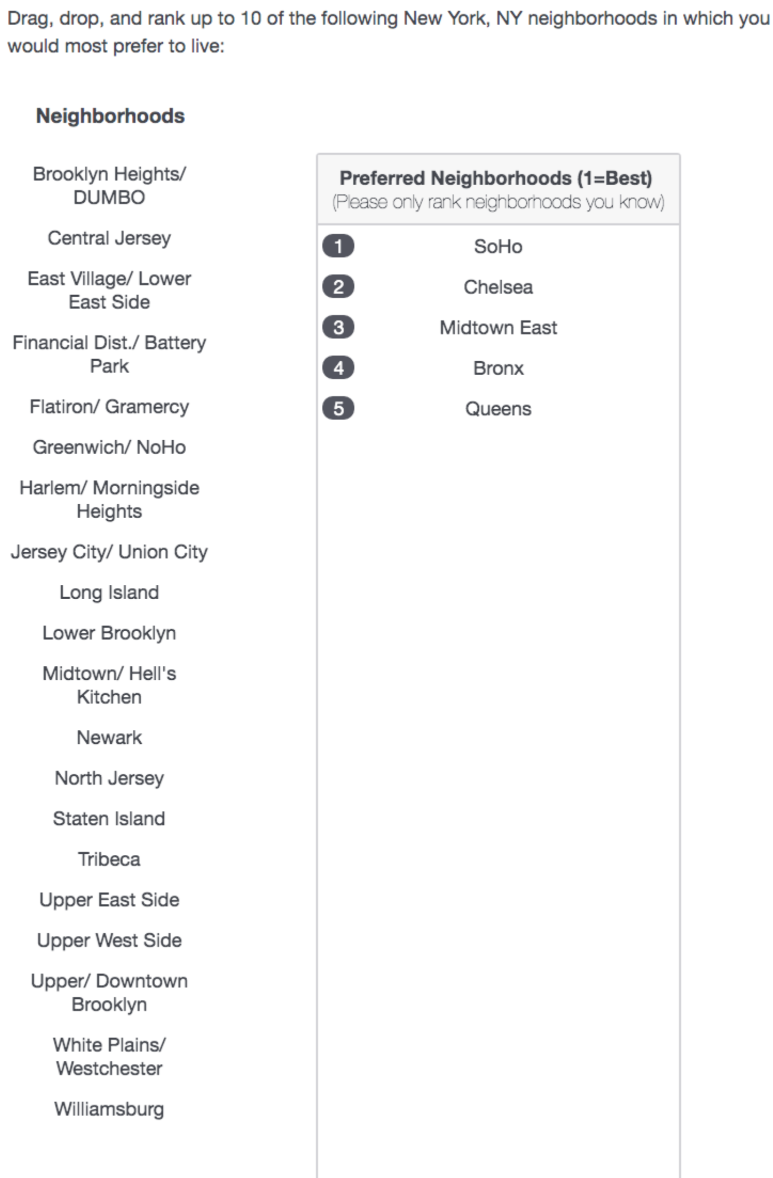


Figure A2: Pre Information Ranking of Neighborhoods



Drag, drop, and rank up to 10 of the following New York, NY neighborhoods in which you would most prefer to live:

Neighborhoods

Brooklyn Heights/ DUMBO
Central Jersey
East Village/ Lower East Side
Financial Dist./ Battery Park
Flatiron/ Gramercy
Greenwich/ NoHo
Harlem/ Morningside Heights
Jersey City/ Union City
Long Island
Lower Brooklyn
Midtown/ Hell's Kitchen
Newark
North Jersey
Staten Island
Tribeca
Upper East Side
Upper West Side
Upper/ Downtown Brooklyn
White Plains/ Westchester
Williamsburg

Preferred Neighborhoods (1=Best)
(Please only rank neighborhoods you know)

1 SoHo
2 Chelsea
3 Midtown East
4 Bronx
5 Queens

Figure A3: Estimates of Monthly Rent in Considered Neighborhoods

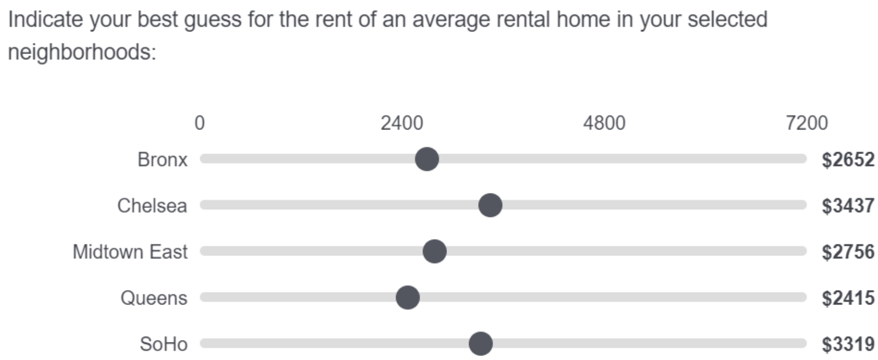


Figure A4: Estimates of Same-School Network Shares in Considered Neighborhoods

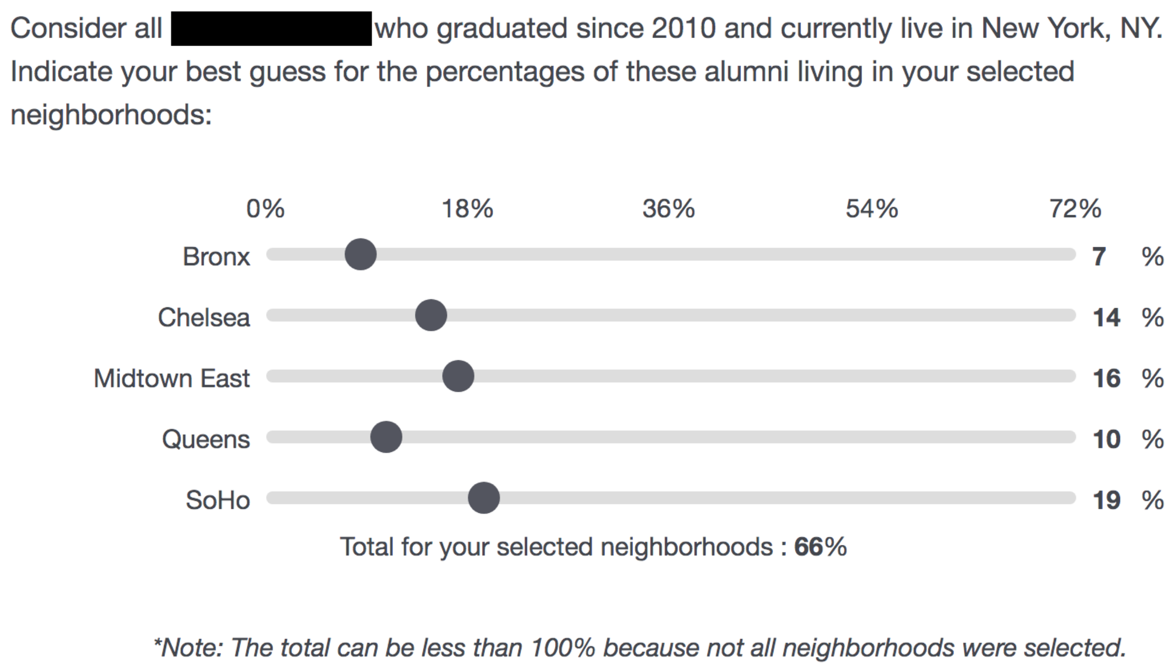
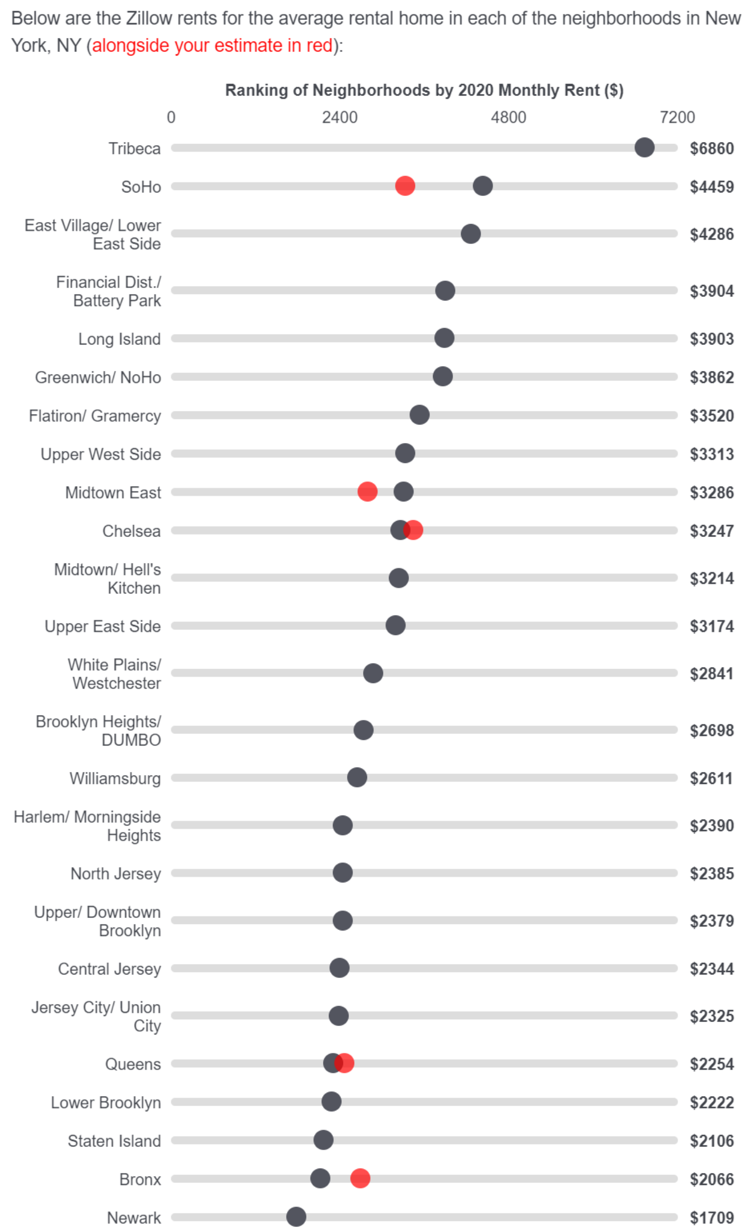


Figure A5: Monthly Zillow Rent (Alongside Pre Information Estimates) in All Neighborhoods



Notes: We also presented an analogous figure for size and network shares but suppressed it here due to the proprietary nature of the data.

Figure A6: Post Information Ranking of Neighborhoods

Notes: Survey repondents saw a full schedule of all neighborhoods, as well as the Zillow rent, size, and network shares (suppressed here due to the proprietary nature of the data).