NBER WORKING PAPER SERIES

HOW SEGREGATED IS URBAN CONSUMPTION?

Donald R. Davis
Jonathan I. Dingel
Joan Monras
Eduardo Morales

Working Paper 23822
http://www.nber.org/papers/w23822

NATIONAL BUREAU OF ECONOMIC RESEARCH
1050 Massachusetts Avenue
Cambridge, MA 02138
September 2017

We thank Jesse Shapiro, four anonymous referees, Treb Allen, David Atkin, Pierre-Philippe Combes, Victor Couture, Thomas Covert, Alon Eizenberg, Ingrid Gould Ellen, Mogens Fosgerau, Marçal Garolera, Robin Gomila, Joshua Gottlieb, Jessie Handbury, Art O'Sullivan, Albert Saiz, and many seminar audiences for helpful comments. We thank Bowen Bao, Luis Costa, Amrit K. Daniel, David Henriquez, Yan Hu, Charlene Lee, Rachel Piontek, Anil Sindhwani, Ludwig Suarez, Shirley Yarin, and, especially, Kevin Dano, Ben Eckersley, Hadi Elzayn, and Benjamin Lee for research assistance. Thanks to the New York Police Department, and especially Gabriel Paez, for sharing geocoded crime data. Dingel thanks the Kathryn and Grant Swick Faculty Research Fund at the University of Chicago Booth School of Business for supporting this work. Monras thanks the Banque de France Sciences Po partnership. Morales thanks the University of Wisconsin-Madison and the Cowles Foundation at Yale University for their hospitality and support. Part of this work is supported by a public grant overseen by the French National Research Agency (ANR) as part of the “Investissements d’Avenir” program LIEPP (ANR-11-LABX-0091, ANR-11-IDEX-0005-02). The views expressed herein are those of the authors and do not necessarily reflect the views of the National Bureau of Economic Research.

How Segregated is Urban Consumption?
Donald R. Davis, Jonathan I. Dingel, Joan Monras, and Eduardo Morales
NBER Working Paper No. 23822
September 2017
JEL No. D12,J15,L83,R2

ABSTRACT

We provide measures of ethnic and racial segregation in urban consumption. Using Yelp reviews, we estimate how spatial and social frictions influence restaurant visits within New York City. Transit time plays a first-order role in consumption choices, so consumption segregation partly reflects residential segregation. Social frictions also have a large impact on restaurant choices: individuals are less likely to visit venues in neighborhoods demographically different from their own. While spatial and social frictions jointly produce significant levels of consumption segregation, we find that restaurant consumption in New York City is only about half as segregated as residences. Consumption segregation owes more to social than spatial frictions.

Donald R. Davis
Department of Economics
Columbia University
1004 International Affairs Building
420 West 118th St.
New York, NY 10027
and NBER
drd28@columbia.edu

Jonathan I. Dingel
Booth School of Business
University of Chicago
5807 South Woodlawn Avenue
Chicago, IL 60637
and NBER
jdingel@chicagobooth.edu

Joan Monras
CEMFI
Calle Casado del Alisal, 5
28014 Madrid
jm3364@gmail.com

Eduardo Morales
Department of Economics
Princeton University
310 Fisher Hall
Princeton, NJ 08544
and NBER
ecmorales@princeton.edu

1 Introduction

For half a century, the United States has prohibited ethnic and racial discrimination in housing, jobs, and education. Even so, segregation in each of these domains remains a stubborn feature of modern America (Hellerstein and Neumark, 2008; Boustan, 2011). Many studies have documented these facts and examined their consequences for socioeconomic outcomes (Massey and Denton, 1993; Cutler and Glaeser, 1997; Chetty et al., 2014).

Discrimination in consumption venues has also been prohibited for decades, yet racial and ethnic segregation in this domain has been studied much less. A major achievement of the civil rights movement was the Civil Rights Act of 1964, prohibiting discrimination based on race or ethnicity in public accommodations.[1] Jim Crow laws had segregated places where people meet socially in order to maintain segregation of intimate contact (Myrdal, 1944, p.588). In contemporary America, these shared spaces have the potential to form what Anderson (2011) calls "a cosmopolitan canopy," a place where a diversity of people may interact such that "a cognitive and cultural basis for trust is established that often leads to the emergence of more civil behavior."[2] Such social capital has potential consequences for many economic outcomes (Guiso et al., 2009; Smith, 2010). Gone are the days of whites-only lunch counters. Yet we do not know the degree to which consumption venues are integrated and serve as places where people of different backgrounds encounter each other in everyday life (O'Flaherty, 2015, p.236-237).

We cannot say a priori whether segregation along demographic lines is greater in consumption or residences. If spatial frictions (i.e., the costs of traversing the city) were infinite, then residential and consumption segregation would necessarily be identical. The fact that individuals may move around the city for consumption purposes could mitigate the effect of the spatial separation of residences if diverse consumers choose common destinations. Conversely, mobility may allow them to segregate even more if their choices diverge. Divergent choices could result from social frictions, such as aversion to consuming in areas with different population demographics or racial segregation of social networks that influence choices (e.g., family or friends). Consumers from different backgrounds may also choose different venues due to differences in tastes. The interactions of spatial frictions, social frictions, and heterogeneous tastes make the degree of consumption segregation an empirical question.

In this paper, we estimate a measure of consumption segregation along racial and ethnic lines for the residents of New York City and quantify the contributions of spatial and social frictions to it. We do so by estimating a discrete-choice model of restaurant visits using information on more than 18,000 consumption decisions by individuals living in New York City who use Yelp.com, a website where users review local businesses.[3] We use our estimated

[1] Public accommodations include hotels, restaurants, and entertainment venues.

[2] There is a large social psychology literature on the mechanisms that may reduce intergroup prejudice. Specifically, the literature on the "contact hypothesis" focuses on the potential for intergroup contact to reduce such prejudice (Allport, 1954). Study results in this literature have generally been consistent with the hypothesis that intergroup contact reduces intergroup prejudice (Pettigrew and Tropp, 2006), including some laboratory and field experiments (Paluck and Green, 2009).

[3] Estimating a model of consumption choices is necessary for both of our purposes. As discussed in Gentzkow et al. (2017), empirical settings in which individuals choose among many alternatives, such as ours, pose a small-sample-bias problem for estimating model-free measures of segregation. Furthermore, decomposing consumption segregation into spatial and social frictions requires separately identifying the

parameters to predict the consumption decisions of all New York City residents. We find that consumption choices are much less segregated than residential locations. Dissimilarity indices contrasting the consumption destinations of consumers of different demographic groups are about half the value of dissimilarity indices for residential locations. Both spatial and social frictions have quantitatively large influences on the geography of consumption, with social frictions contributing relatively more to consumption segregation along ethnic and racial lines.

Inferring spatial and social frictions from consumption behavior requires controlling for other determinants of consumers' choices. We exploit several advantages of our dataset, described in Section 2, to identify these frictions. First, we locate Yelp users' residences and workplaces, allowing us to measure spatial frictions that account for the fact that consumption may originate at home, work, or the commute between them and that both automobile and public transit may be used. Second, we combine data from Yelp and the United States Census Bureau to characterize user demographics, restaurant characteristics, and neighborhood demographics. This allows us to distinguish demographic differences in tastes from social frictions, and to measure the contributions of both individual-level homophily and demographic differences across neighborhoods to social frictions.

However, our dataset is not without limitations. First, Yelp users are not representative of the general population. In terms of observable characteristics, users in our estimation sample are more likely to be Asian and female. They live in neighborhoods with higher incomes and more residents between the ages of 21 and 39. To address possible biases in our measures of consumption segregation for the whole population of New York City arising from these compositional differences, we allow our estimates of preference parameters to vary with these observable individual characteristics. This approach addresses selection on these observable characteristics but it cannot alleviate other differences between Yelp users and the general population that we may not observe. Our results for the whole population of New York City residents thus necessarily predict the level of consumption segregation that prevails if everyone behaves as the observationally equivalent users in our estimation sample. Second, we have limited ability to distinguish between Hispanic and white users in our sample, so we cannot separately identify these two groups' preference parameters. However, we exploit tract-level information on demographics to capture social frictions between whites and Hispanics. Third, we observe every review written by Yelp users, not every restaurant visit. Identifying consumer preferences therefore involves assumptions on review-writing behavior, which we discuss in Section 3.

We model consumers' behavior using a conditional-logit specification in which a consumer's valuation of a restaurant may depend upon spatial frictions, social frictions, and a large set of observable characteristics of the consumer and the restaurant. All preference parameters are allowed to vary flexibly by race. Our estimation procedure makes use of the McFadden (1978) choice-set construction technique to address the computational burden arising from consumers choosing among the thousands of restaurants in New York City.

We present our parameter estimates in Section 4. Our quantification of consumers' aversion to incurring longer travel times reveals a first-order role for spatial frictions in determining the geography of consumption. Depending on the origin of the trip and the

relevance of each for consumers' choices.

mode of transport used, halving the minutes of travel time to a venue implies that a user would be two to nearly four times more likely to visit the venue from that origin by that mode. These spatial frictions will cause consumption patterns to partly inherit residential patterns of segregation.[4]

Consumption segregation also reflects the influence of demographic-related social frictions. These frictions make consumers' decisions depend upon the contrast between the residential demographics of the restaurant's location and either the residential demographics of the user's home location or the user's own racial or ethnic identity. All else equal, a user is more likely to visit a venue in a census tract that is more demographically similar to her home tract. Individuals are also more likely to visit restaurants in tracts with a larger share of residents of their own racial group. While consumption may be integrated de jure, these social frictions make consumption less integrated de facto.

Importantly, our estimates of both spatial and social frictions are obtained after controlling for race-specific tastes for observable features of restaurants and areas of the city. For example, we incorporate cuisine-category fixed effects. Thus, for example, our finding that Asian consumers are more likely to visit restaurants (of any type) located in neighborhoods with more Asian residents is conditional on the fact that Asian consumers are more likely to visit restaurants serving Asian cuisines. Similarly, we allow consumers' valuations of restaurant prices and ratings to depend on the income level of their home census tract and control for income differences when estimating social frictions associated with racial demographic differences. In robustness checks, we introduce restaurant fixed effects that vary by race and allow for correlation in consumer-specific preferences across restaurants of similar characteristics in nested-logit specifications. These specifications yield similar estimates of spatial and social frictions.

Our estimated model fits the data well. Race-specific preference parameters are key to capturing the level of consumption segregation that we observe in the estimation sample: a specification that assumes that many preference parameters are common across all consumers cannot replicate the in-sample isolation of consumers of different races. The specifications that introduce race-specific restaurant fixed effects yield only very modest improvements in fit. This is consistent with the fact that there is little segregation of Yelp reviewers between pairs of restaurants that are observationally equivalent.

Using our estimated model of the restaurant-visit decision, we compute measures of consumption segregation for the entire residential population of New York City in Section 5. Specifically, we characterize the ethnic and racial segregation of the predicted consumption choices using dissimilarity indices. A dissimilarity index describes the fraction of the population belonging to a group – Asian consumers, for example – that would have to alter their consumption choices in order to match the distribution of predicted restaurant choices made by the remainder of the population. Despite the magnitude of the estimated spatial and social frictions, consumption dissimilarity is notably lower than residential dissimilarity for all ethnic and racial groups.

To quantify the contribution of spatial frictions, social frictions, and demographic differences in tastes to consumption segregation, we re-compute the dissimilarity indices using

[4] This is reminiscent of the spatial mismatch hypothesis, due to Kain (1968), which posits that spatial frictions make residential segregation a source of workplace segregation.

the consumption decisions predicted by our estimated model when the coefficients capturing the corresponding friction are set to zero. Social frictions make a larger contribution to consumption segregation than spatial frictions. Eliminating spatial frictions entirely would reduce consumption dissimilarity indices by 10% to 20%, depending on the race or ethnicity. While eliminating spatial frictions could have freed individuals to further segregate their consumption, the predicted decrease implies that different races' shared valuations of choices are stronger than the divergent forces of social frictions and taste heterogeneity. In the absence of social frictions, consumption would be considerably more integrated, with dissimilarity indices at roughly two-thirds their current levels. In the absence of both spatial and social frictions, differences in preferences and income levels between ethnic and racial groups would predict consumption segregation that is still substantial but quite modest relative to residential segregation.

We also use our estimated model to examine the impact on consumption segregation of counterfactual changes in transportation policy and in the preference parameters determining the degree of social frictions. Consistent with the modest role of spatial frictions overall, major changes in transportation infrastructure have only small effects on consumption dissimilarity. Reductions in social frictions would integrate consumption.

Finally, we use our estimates to measure the welfare consequences of neighborhood change for incumbent residents in Section 6. Gentrification is associated with changes in both restaurants' and residents' characteristics that affect the value of consumption. We compute the change in welfare that the residents of a census tract in the middle of Harlem would experience if the surrounding census tracts were to exhibit the residential and restaurants characteristics of high-income, majority-white census tracts of the Upper East Side. We find a significant reduction in the value of their consumption opportunities. This is attributable to the increase in social frictions associated with the change in racial demographics. The change in restaurants' characteristics would have very modest effects on their welfare.

Our findings relate to a recent literature on the geography of urban consumption. Studies have documented cross-city variation in the tradable goods available for consumption (Handbury, 2013; Handbury and Weinstein, 2015), and geographic variation in the supply of non-tradables has been posited to shape the relative attractiveness of cities (Glaeser et al., 2001; Schiff, 2015). Waldfogel (2007) documents that restaurant entry in different cuisine categories is correlated with local demographic composition. This dimension of economic life has grown increasingly important in recent decades.[5] Prior studies of the geography of consumption within the city include Katz (2007), Houde (2012), Couture (2015), and Eizenberg et al. (2017). Relative to this prior work, we build a unique dataset that combines information on individuals' home and work locations, their demographics, and characteristics of the restaurants they patronize, and we use it to separately identify the effect of spatial and social frictions on consumer decisions.

We study urban consumption using online user-generated content, which is increasingly exploited by economists. Among others, Anderson and Magruder (2012) and Luca (2016a) examine Yelp's effects on restaurant outcomes. Edelman and Luca (2014) infer racial iden-

[5] US households' share of food spending devoted to food prepared away from home grew from less than 26% in 1970 to more than 43% in 2012 (USDA 2014). Analogously, while the number of daily commuting trips has stayed relatively constant for decades, trips for social/recreational purposes have steadily grown (Commuting in America III, 2006).

tities from profile photos to study discrimination on Airbnb.com. Caetano and Maheshri (2015) document consumption segregation by gender using Foursquare data. Luca (2016b) surveys this growing body of research on user-generated content and social media.

We contribute to the large literature on social and economic fragmentation related to demographic differences by measuring consumption segregation in restaurants. The prior literature has largely focused on residential segregation, though there are studies documenting the segregation of workplaces (Hellerstein and Neumark, 2008), students' friendship networks (Echenique and Fryer, 2007), and media consumption (George and Waldfogel, 2003; Oberholzer-Gee and Waldfogel, 2009). We study racial segregation of consumption in a setting in which consumers travel to consume and come face-to-face with each other. Everyday encounters between people of different backgrounds in shared public spaces may be a basis for building understanding and tolerance (Anderson, 2011), though consumption venues are also sometimes sites of racial and ethnic discrimination (Labaton, 1994; Lee, 2000; Ayres, 2001; Antecol and Cobb-Clark, 2008; Schreer et al., 2009). We provide the first quantification of the segregation of these consumption choices.

2 Data

We combine data from Yelp and other sources to estimate our model of the restaurant-visit decision and compute measures of consumption segregation. Section 2.1 describes the Yelp data and Section 2.2 describes the other sources of data we use. Section 2.3 presents evidence suggestive of the influence of spatial and social frictions on consumers' restaurant choices.

2.1 Yelp data

Yelp.com is a website where users review local businesses, primarily restaurants and retail stores (Yelp, 2013). It describes a venue in terms of its address, average rating, user reviews, and a wide variety of other characteristics. Yelp's coverage of restaurants is close to comprehensive (see Appendix B.1). The website is relevant for the general population: discontinuities in Yelp ratings have substantive effects on restaurants' revenues (Luca, 2016a) and reservation availability (Anderson and Magruder, 2012). In addition to assigning a rating of one to five stars, users write a review describing their personal experience with the business. Crucial for our purposes is that users sometimes disclose information in their reviews about their residential and work locations.[6]

We use data on Yelp users who reviewed a New York City (henceforth, NYC) restaurant venue between 2005 and 2011. As described in detail in Appendix B.2, we identify users' residential and work locations from the text of their reviews. Specifically, we first search the text of a large number of reviews for 26 key phrases related to location, such as "close to me," "block away," and "my apartment." Then we read the reviews containing these phrases to infer whether the user's home or work is proximate to the reviewed business. Finally, we estimate the residential and work locations of a user as the average of the latitude-longitude coordinates of the sets of venues identified as being close to this user's home and work

[6] Another example of using the information disclosed in reviews is detecting outbreaks of food poisoning unreported to NYC health authorities (Harrison et al., 2014).

locations, respectively. Restricting our sample to users whose reviews do not reveal a change in residence or workplace within NYC and whose home and work locations are in census tracts with demographic and income information, we obtain an estimation sample of 18,015 reviews written by 440 distinct users.[7]

Yelp users typically post a profile photo, which we use to infer a user's apparent gender and race.[8] Mayer and Puller (2008) compare measures of ethnicity and race inferred from photos with administrative data and finds a high degree of accuracy in partitioning subjects into three racial groups: Asian; black; and white or Hispanic. Consequently, when inferring each user's race from her profile picture, we limit ourselves to classifying individuals into these three groups.

Table 1 reports summary statistics for the 440 users included in our estimation sample and the broader NYC population. Sixty-one percent of estimation-sample users are female, and only 5% of users are of unidentified gender. We could not infer race for 26% of users. Asian users constitute 24% of the estimation sample, while white or Hispanic users are 42% of the sample. Asians are thus overrepresented in our sample, as Asian residents constitute only about 12% of the population of NYC. Although only 10% of the users with an inferred race were identified as black, these users wrote more than one thousand reviews.

Figure 1 depicts the home and work locations of the users in our estimation sample. Consistent with patterns in the broader population of NYC, these users' workplaces are concentrated in Manhattan below Fifty-Ninth Street, while their residences are more dispersed. The average user in our estimation sample lives in a census tract with median household income near \$75,600, which is typical of Manhattan but higher than NYC as a whole (\$56,300). The users in our estimation sample tend to live in census tracts with a share of the population between the ages of 21 and 39 (42%) that is higher than that of both Manhattan (37%) and NYC as a whole (30%). These patterns are consistent with statements that Yelp's global user base is younger and higher-income than the population as a whole (Yelp, 2013).

Asian and black users in our estimation sample are less residentially segregated than Asian and black residents of NYC as a whole. Table 1 reports "isolation indices" as defined in the canonical survey of Massey and Denton (1988), $\sum_k (pop_{gk}/pop_g) \cdot (pop_{gk}/pop_k)$, where pop_g is the population of group g, pop_k is the population of tract k, and pop_{gk} is the population of group g in tract k. These characterize the average group g (e.g., Asian) share of tract residents experienced by members of group g. White/Hispanic users in our estimation sample live in census tracts with a white/Hispanic share of residents that is typical of that experienced by white/Hispanic residents of NYC. By contrast, Asian and black users in our estimation sample live in census tracts that have lower Asian and black shares, respectively, than is typical. The isolation indices for Asian and black users in our estimation sample are about three-quarters their values for Manhattan residents and half their values for NYC

[7] We could use a considerably larger sample if we only required information on the user's home location. Estimates of models that do not require information on individuals' workplaces are quantitatively similar irrespective of whether we use the restricted sample in which both home and work locations are known or the larger sample of individuals for whom we only have information on home locations.

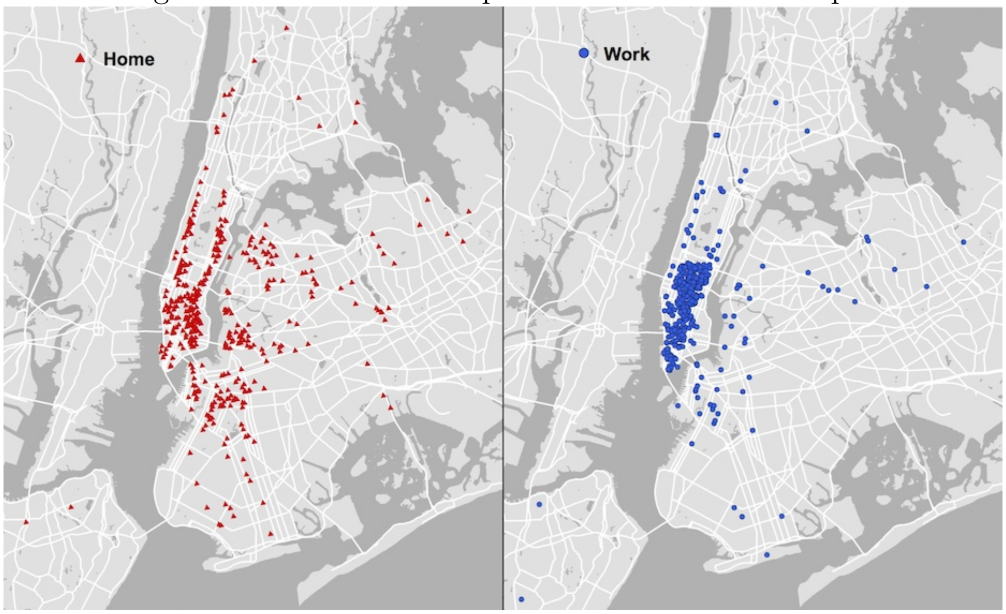
[8] While users may choose "male" or "female" for their gender on their Yelp profile, this information is not publicly displayed. Thus, we classify users based on their gender presentation in their profile photo. Users with profile photos for which we could not classify the gender (e.g., cartoon graphics, photos of animals) have both male and female dummy variables equal to zero.

Table 1: Estimation sample and NYC summary statistics

	Estimation sample Yelp users	Manhattan tracts	NYC tracts
User appearance / Tract demographics			
Female	.61	.525	.524
Male	.34	.475	.476
Asian	.24	.117	.125
Black	.08	.142	.245
White or Hispanic	.42	.716	.600
Hispanic		.232	.265
White		.484	.335
User race indeterminate	.26		
Home tract characteristics			
Median household income (thousands)	75.6	76.3	56.3
Age 21-39 residents share	.42	.374	.302
Asian isolation index	.197	.273	.325
Black isolation index	.28	.383	.569
White/Hispanic isolation index	.778	.786	.73
Observations	440	279	2110

NOTES: This table summarizes characteristics of the 440 Yelp users in our estimation sample and all census tracts in Manhattan and New York City. User demographics are inferred from Yelp profile photos. Tract demographics from 2010 Census of Population and tract incomes from 2007-2011 American Community Survey. Isolation indices as defined in Massey and Denton (1988).

Figure 1: Locations of Yelp users in estimation sample



Notes: This figure depicts the distribution of home and work locations of the 440 users in our estimation sample.

residents as a whole.

Figure A.1 displays all the restaurants reviewed by users in our estimation sample. These venues are concentrated in Manhattan below Fifty-Ninth Street, but our estimation sample contains venues in many parts of NYC. Table A.1 summarizes the distribution of reviews in terms of venues' prices, ratings, cuisine types, and boroughs for both our estimation sample and Yelp in its entirety. The users in our estimation sample exhibit review frequencies similar to those of the broader Yelp population.[9]

2.2 NYC transit, demographic, and crime data

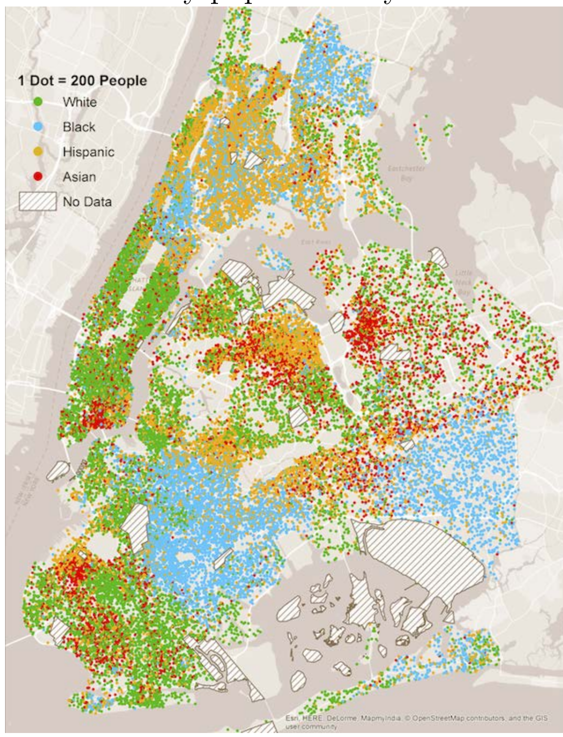
To measure spatial frictions, we use car and public-transit times between the centroids of census tracts from Google Maps. In addition to direct travel from home or work, we compute the additional transit time a user would incur by incorporating a visit to a venue as part of her commute. Denote the transit time from location x to location y by $time(x, y)$. For user i living in h_i and working in w_i, the travel time associated with visiting venue j in tract k_j from her commuting path p_i is computed as

$$time(p_i, j) = \frac{1}{2} \max \{time(h_i, k_j) + time(w_i, k_j) - time(h_i, w_i), 0\},$$

where the maximum operator imposes the triangle inequality on transit times.

[9]Within-user geographic dispersion in restaurants reviewed, as measured by the average distance between any two reviewed restaurants, and within-user demographic variation, as measured by the average Euclidean demographic distance (defined below) between any two reviewed restaurants' surrounding residents, are also very similar for users in our estimation sample and Yelp users as a whole. Users in our sample thus do not seem to be particularly "adventurous" relative to the typical Yelp user.

Figure 2: New York City population by race or ethnicity, 2010



Notes: This figure depicts the residential NYC population in terms of four demographic categories that cover 97% of the population. Each dot represents 200 people. Tract-level population data from the 2010 Census of Population.

To measure social frictions associated with racial and ethnic demographics, we use data from the 2010 Census of Population that describe each census tract's residential population in terms of five groups: Asian, black, Hispanic, white, and other.[10] These population counts are depicted in Figure 2.[11] Using these data, we measure ethnic and racial differences between two tracts as the Euclidean distance between the vectors containing the two tracts' five residential population shares. Specifically, defining $shares_{tract}$ as the five-element vector containing these population shares, the "Euclidean demographic distance" (henceforth, EDD) between origin and destination tracts is

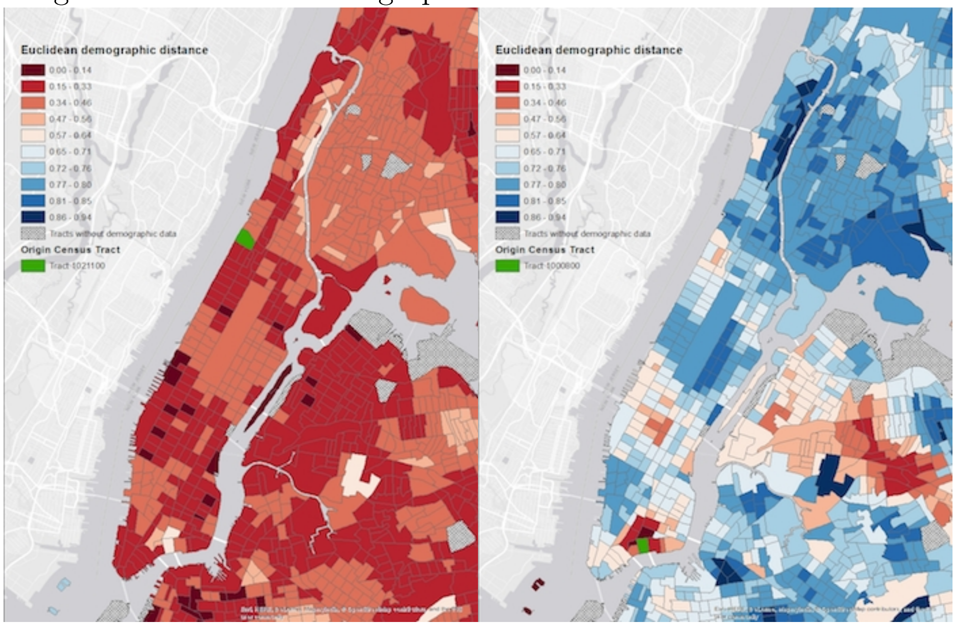
$$\|shares_{origin} - shares_{destination}\|/\sqrt{2},$$

where $\|\cdot\|$ indicates the L^2 norm. This measure ranges from zero to one. Figure 3 illustrates the EDD for two origin tracts and many destination tracts. The Morningside Heights origin in the left panel has a diverse population that is similar to most NYC tracts, and thus its EDD to most destinations is low. The Manhattan Chinatown origin in the right panel is

[10] To be precise, we use the population counts of non-Hispanic Asians, non-Hispanic blacks, all Hispanics, and non-Hispanic whites to respectively define the groups we call "Asian," "black," "Hispanic," and "white." The "other" group includes Native Americans, Hawaiians, other races, and mixed-race categories; it constitutes about 3% of the NYC population.

[11] This map was inspired by a *New York Times* project, "Mapping America: Every City, Every Block."

Figure 3: Euclidean demographic distances from two census tracts



NOTES: These maps depict Euclidean demographic distances from an origin tract to other NYC tracts. In the left panel, the origin tract is in Morningside Heights; in the right panel, Manhattan's Chinatown. Demographic data from 2010 Census of Population.

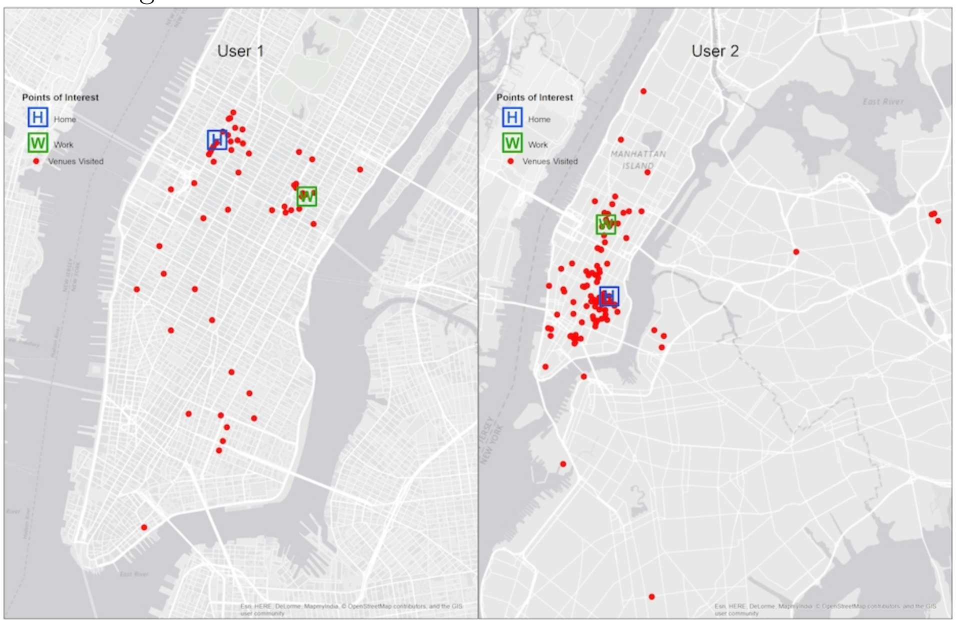
overwhelmingly Asian and thus quite demographically distant from most tracts, with the exception of the Flushing Chinatown in Queens.

To allow social frictions to depend not only on the demographic composition of the tract in which the restaurant is located but also on the surrounding demographic composition, we calculate the Echenique and Fryer (2007) spectral segregation index (henceforth, SSI) for the modal residential race or ethnicity in each census tract. This index measures the degree to which a census tract borders census tracts of the same residential demographic plurality, and the further degree to which those tracts themselves border tracts of the same plurality, ad infinitum. For example, in Figure 2, the black census tracts at the center of the cluster of blue dots in Queens, on the right edge of the map, will have higher SSI values than those at the edge of the cluster.

We also allow income and crime levels in a restaurant's tract to influence consumer choices. The data on median household incomes come from the 2007-2011 American Community Survey 5-Year Estimate. To measure crime rates by location, we compute tract-level robbery statistics for 2007-2011 using confidential, geocoded incident-level reports provided to us by the New York Police Department.[12] We use robberies as our crime measure because these are likely the most common and relevant threat to individuals visiting a restaurant. All tract-level characteristics are summarized in Tables 1 and A.2.

[12]Fewer than 3% of the Yelp reviews in our estimation sample were posted outside of 2007-2011.

Figure 4: Two users' locations and restaurant reviews



NOTES: These two maps display two users' home and work locations and the Yelp restaurant venues they reviewed. Red dots denote Yelp venues reviewed by this user. The "H" denotes the average coordinates of those venues identified as home locations in the text of this user's reviews. The "W" denotes the similarly defined work location.

2.3 Observed behavior and frictions

We present here some evidence suggesting the relevance of spatial and social frictions for restaurant choices.

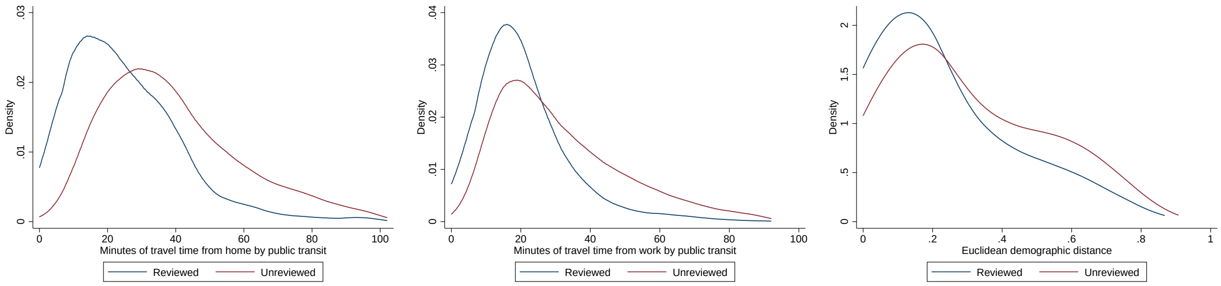
Individual users' reviews suggest that both proximity and venue characteristics influence user behavior. Figure 4 maps home, work, and restaurant-review locations for two individuals in our sample. The user in the left panel lives and works in midtown Manhattan. The other user works in midtown Manhattan and resides in a southeastern Manhattan development called Stuyvesant Town. Both users primarily review venues that are near their home or work locations. At the same time, both users visit more downtown venues than uptown venues, which may reflect differences in the quantity or quality of venues in these areas.

The choices made by all users in our estimation sample suggest the importance of spatial and social frictions. Figure 5 plots, for all users in our estimation sample, the density of three covariates for the set of venues reviewed by these users and for a random sample of venues that they did not review. The left panel depicts transit times from home, the center panel transit times from work, and the right panel the EDD between the home census tract and the venue's tract. The plots show that, unconditionally, Yelp users are more likely to review venues that are closer to their residential and workplace locations and located in tracts with demographics more similar to those of their home tract.[13]

While these density plots are consistent with the idea that spatial and social frictions

[13]The fact that both reviewed and unreviewed venues have shorter travel times from work than from home reflects the fact that most venues and workplaces are in Manhattan.

Figure 5: Travel times, demographic differences, and consumer choice



NOTES: These plots are kernel densities for three distributions of user-venue pairs: those venues chosen by users in our estimation sample and a random sample of venues not chosen by these users. The left panel plots the densities as functions of travel time from home by public transit; the center panel shows travel time from work by public transit; the right panel shows Euclidean demographic distances. Epanechnikov kernels with bandwidths of 3, 3 and 0.1, respectively.

influence consumers' decisions, they are only suggestive. Each plot neglects the influence of other frictions and restaurant characteristics on consumers' decisions. For example, given the city's significant residential segregation, transit times between census tracts are positively correlated with measures of demographic differences. Therefore, from Figure 5 alone, it is not possible to identify the relative influence of spatial and demographic distance on consumers' decisions.

3 Empirical approach

To measure the relative importance of tastes and spatial and social frictions in determining the restaurant choices of consumers of different races or ethnicities, we introduce a discrete choice model of restaurant visits. Section 3.1 describes the assumptions we impose on consumers' preferences. Since we observe restaurant reviews, not all restaurant visits, assumptions on the review-writing behavior of Yelp users are necessary for identification of their preferences. Section 3.2 introduces these assumptions. Section 3.3 describes the steps we follow to exploit the data introduced in Section 2 to estimate our model. In Section 3.4, we introduce several extensions that relax the key identifying assumptions in our baseline model.

3.1 Demand specification

Individuals decide whether to visit any venue and, if they do, which venue to visit. We index individuals by i, venues by j, and by t the occasions in which i needs to decide on whether to visit a venue. In our empirical application, we assume that the set J of potential venues that a consumer may visit is the set of all NYC restaurants listed on Yelp and located in a census tract for which information on residents' median income is available.[14] We denote

[14]Specifically, we restrict the set of restaurants to only those with price and rating information listed on Yelp in June 2011, when we collected our data. Appendix B.1 shows that Yelp's coverage of venues is nearly comprehensive. Section 2.2 shows that data on median income is available for nearly all census tracts.

the outside option of not visiting any venue by $j = 0$.

When visiting a venue, individuals must choose whether to visit it from home, work, or by deviating from their commuting path, and whether to travel via public transit or car. We index pairs of origin locations and transportation modes by l and assume that a trip to a venue may be of one of six types: from home via car ($l = hc$), from home via public transit ($l = hp$), from work via car ($l = wc$), from work via public transit ($l = wp$), from the commuting path via car ($l = pc$), or from the commuting path via public transit ($l = pp$). We denote the set of these six potential origin-mode pairs as $\mathcal{L} \equiv \{hc, hp, wc, wp, pc, pp\}$.

We allow preferences for restaurants, trip origin locations and transportation modes to differ across racial or ethnic groups. We index groups by g, which may take three values: white or Hispanic ($g = w$), Asian ($g = a$), and black ($g = b$). We denote the set of these three potential groups as $\mathcal{G} \equiv \{w, a, b\}$.

For an individual i belonging to the racial or ethnic group $g(i)$, we assume that her utility of visiting restaurant j in period t from origin-mode l is

$$U_{ijlt} = \gamma^1_{g(i)l} X^1_{ijl} + \gamma^2_{g(i)} X^2_{ij} + \beta^1_{g(i)} Z^1_j + \beta^2_{g(i)} Z^2_{ij} + \nu_{ijlt}, \tag{1}$$

where X^1_{ijl} measures the spatial frictions that i incurs when visiting j from l, the vector X^2_{ij} measures the social frictions that may affect the appeal that restaurant j has to individual i, and Z^1_j and Z^2_{ij} control for other observed venue and individual-venue-specific characteristics, respectively. The variable ν_{ijlt} is a scalar unobserved by the econometrician. We allow all preference parameters $(\gamma^1, \gamma^2, \beta^1, \beta^2)$ to vary across demographic groups g and, since the coefficient γ^1 is l-specific, we additionally allow the marginal disutility of a trip to flexibly depend on both its origin and the mode of transit.

Although our dataset describes users' home and work locations, it does not indicate the origin-mode l of each trip. We address this data limitation by assuming that consumers jointly optimize the restaurant they patronize and the origin-mode from which they do so, choosing thus the jl combination that maximizes their utility. Accordingly, defining a dummy variable d_{ijlt} that equals one if individual i travels to venue j from origin-mode l at period t, we assume that

$$d_{ijlt} = \mathbb{1}\{U_{ijlt} \geq U_{ij'l't}; \forall\ j' \in J,\ l' \in \mathcal{L}\}, \tag{2}$$

where $\mathbb{1}\{A\}$ is an indicator function that equals one if A is true. We also define a variable d_{ijt} that is one if individual i chooses venue j at period t, $d_{ijt} = \sum_{l \in \mathcal{L}} d_{ijlt}$, irrespective of the origin-mode of the trip.

In our benchmark specification, we assume that the vector of unobserved utilities for individual i at period t, $\nu_{it} = \{\nu_{ijlt}; \forall\ j \in J,\ l \in \mathcal{L}\}$, is independent across individuals and time periods and has a joint type I extreme value distribution: its cumulative distribution function is $F(\nu_{it}) = \exp(-\sum_{j \in J} \sum_{l \in \mathcal{L}} \exp(-\nu_{ijlt}))$. This distribution yields a conditional-logit discrete-choice model of restaurant visits.

3.2 Review-writing behavior

Let d^*_{ijt} be a dummy variable that equals one if individual i writes a review of restaurant j at time t. The fact that we observe reviews rather than restaurant visits (i.e., we observe d^*_{ijt} but not d_{ijt}) implies that estimating the preference parameters in equation (1) requires making assumptions on the review-writing behavior of Yelp users. We impose three assumptions. First, users do not review restaurants they have not visited. Second, they write at most one review per restaurant (independently of how many times they visit a restaurant). Third, conditional on having visited a restaurant and not having previously reviewed it, they write a review with a probability p^*_{it} that is independent of the restaurant's characteristics and the origin-mode of the trip.

3.3 Estimation procedure

We estimate the preference parameters in equation (1) using a maximum likelihood estimator. To derive the relevant likelihood function, we implement the following five steps.

Step 1: derive restaurant visit probability. According to the assumptions in Section 3.1, the probability that individual i visits venue j from origin-mode l at period t is

$$P(d_{ijlt} = 1|X_i, Z_i, J; (\gamma, \beta)) = \exp(V_{ijl}) \Big/ \sum_{j' \in J} (\sum_{l' \in \mathcal{L}} \exp(V_{ij'l'})), \tag{3}$$

where X_i, Z_i, γ, and β are vectors that collect their respective terms and[15]

$$V_{ijl} \equiv \gamma^1_{g(i)l} X^1_{ijl} + \gamma^2_{g(i)} X^2_{ij} + \beta^1_{g(i)} Z^1_j + \beta^2_{g(i)} Z^2_{ij}. \tag{4}$$

The probability that individual i visits venue j at period t is then,

$$\begin{aligned} P(d_{ijt} = 1|X_i, Z_i, J; (\gamma, \beta)) &= P(\sum_{l \in \mathcal{L}} d_{ijlt} = 1|X_i, Z_i, J; (\gamma, \beta)) \\ &= \sum_{l \in \mathcal{L}} P(d_{ijlt} = 1|X_i, Z_i, J; (\gamma, \beta)) \\ &= \sum_{l \in \mathcal{L}} \exp(V_{ijl}) \Big/ \sum_{j' \in J} (\sum_{l' \in \mathcal{L}} \exp(V_{ij'l'})). \end{aligned} \tag{5}$$

The first equality applies the definition of d_{ijt}, the second one takes into account that the mode-origin of a trip is unique (i.e. the joint probability that individual i visits restaurant j at period t from two different origins l and l' is zero), and the third equality uses equation (3).

Step 2: derive restaurant review probability. As shown in Appendix C.1, equation (5) and the review-writing model described in Section 3.2 imply that the probability of observing a review of venue j written by individual i at period t is

$$P(d^*_{ijt} = 1|X_i, Z_i, J_{it}, J'_{it}; (\gamma, \beta, p^*_{it})) = p^*_{it} \mathbb{1}\{j \neq 0, j \in J'_{it}\} P(d_{ijt} = 1|X_i, Z_i, J; (\gamma, \beta)), \tag{6}$$

where J'_{it} denotes the set of restaurants not previously reviewed by i; i.e. $J'_{it} \equiv \{j \in J : d^*_{ijt'} = 0 \text{ for all } t' < t \text{ and } j \neq 0\}$.[16] As shown in Appendix C.1, combining equations (5)

[15] Formally, $X_i \equiv \{(X^1_{ijl}, X^2_{ij}); \forall j \in J, l \in \mathcal{L}\}$, $Z_i \equiv \{(Z^1_j, Z^2_{ij}); \forall j \in J\}$, $\gamma \equiv \{(\gamma^1_{gl}, \gamma^2_g); \forall g \in \mathcal{G}, l \in \mathcal{L}\}$, and $\beta \equiv \{(\beta^1_g, \beta^2_g); \forall g \in \mathcal{G}\}$.

[16] As reflected in equation (6), if individuals were to review every restaurant they visit for the first time, $p^*_{it} = 1$, the probability of observing a review would equal either the probability in equation (5) (for any venue j that user i has not previously visited) or zero (for any previously visited venue).

and (6), we can derive the probability that individual i reviews restaurant j at period t conditional on i reviewing any restaurant at that period:

$$P(d^*_{ijt} = 1 | d^*_{it} = 1, X_i, Z_i, J'_{it}; (\gamma, \beta)) = \frac{\mathbb{1}\{j \neq 0, j \in J'_{it}\} \sum_{l \in \mathcal{L}} \exp(V_{ijl})}{\sum_{j \in J'_{it}} \sum_{l \in \mathcal{L}} \exp(V_{ij'l})}, \tag{7}$$

where $d^*_{it} = \sum_{j=1}^{J} d^*_{ijt}$ is a dummy variable that equals one if i writes a review at t.

Step 3: reduce choice set. The cardinality of the choice set J'_{it} makes it computationally burdensome to construct the denominator of the probability in equation (7). As J'_{it} equals the set of all restaurants in NYC, J, minus those reviewed by individual i prior to period t, the large dimensionality of J implies that the set J'_{it} will also be very large.

To address this dimensionality issue, we adapt the choice-set-reduction procedure from McFadden (1978) to our empirical setting. For every individual i and period t in which we observe a review written by i, we define a set S_{it} that is a subset of J'_{it}. We construct S_{it} by including the restaurant j for which $d^*_{ijt} = 1$ plus a random subset of all other alternatives in J'_{it}, selecting them from J'_{it} with equal probability. As all elements of S_{it} other than the actual choice of i at t are selected randomly, the set S_{it} itself is random. We denote by $\pi(S_{it} | d^*_{ijt} = 1, J'_{it})$ the probability of assigning the subset S_{it} to an individual i who reviewed venue j at t. Our sampling scheme implies that

$$\pi(S_{it} | d^*_{ijt} = 1, J'_{it}) = \begin{cases} \kappa_{it} & \text{if } j \in S_{it}, \\ 0 & \text{otherwise,} \end{cases} \tag{8}$$

where $\kappa_{it} \in (0, 1)$ is a constant determined by our choice of the number of venues in S_{it} and the number of venues in J'_{it}.

As we show in Appendix C.1, given equations (7) and (8), we can write the probability that i reviews restaurant j at period t conditional on a randomly drawn set S_{it} and that i writes a review at t as

$$P(d^*_{ijt} = 1 | d^*_{it} = 1, X_i, Z_i, S_{it}; (\gamma, \beta)) = \frac{\mathbb{1}\{j \in S_{it}\} \sum_{l \in \mathcal{L}} \exp(V_{ijl})}{\sum_{j' \in S_{it}} \sum_{l \in \mathcal{L}} \exp(V_{ij'l})}. \tag{9}$$

It is important to remark that, to be able to randomly draw a set S_{it} from the set of non-reviewed restaurants J'_{it}, one needs to observe *all* reviews previously written by user i. Building the probability in equation (9) will thus be infeasible in empirical applications in which, contrary to our case, researchers do not observe the complete history of reviews of the individuals in the sample. Only because we observe this complete history is it true that the set J'_{i1} corresponding to the first observed review of user i is identical to the set of all NYC restaurants, J.

Step 4: derive individual i-specific likelihood function. Using j_{it} to denote the restaurant reviewed by individual i at period t, the joint probability of observing an individual i writing the T_i reviews $\{j_{i1}, j_{i2}, \ldots, j_{iT_i}\}$ conditional on observing a review written by i in each of the periods $\{1, \ldots, T_i\}$ and on randomly drawing the sets $\{S_{i1}, S_{i2}, \ldots, S_{iT_i}\}$ is

$$\prod_{t=1}^{T_i} \frac{\mathbb{1}\{j \in S_{it}\} \sum_{l \in \mathcal{L}} \exp(V_{ijl})}{\sum_{j' \in S_{it}} \sum_{l \in \mathcal{L}} \exp(V_{ij'l})}. \tag{10}$$

This joint probability is simply the product of the corresponding marginal probabilities in equation (9). Intuitively, all the dynamic effects of a review written by individual i at period t are reflected in the subsequent choice sets $\{J'_{is}, \forall s > t\}$, and the effect of each of these choice sets is subsumed in the randomly selected subset $\{S_{is}, \forall s > t\}$. The mathematical derivation of the probability in equation (10) is in Appendix C.1.

Step 5: derive log-likelihood function. Given equation (10) and assuming that we observe a random sample $i = 1, \ldots, N$ of individuals from the population of interest, we can write the log-likelihood function as

$$\sum_{i=1}^{N} \sum_{t=1}^{T_i} \sum_{j \in S_{it}} \mathbb{1}\{d^*_{ijt} = 1\} \ln \left(\frac{\sum_{l \in \mathcal{L}} \exp(V_{ijl})}{\sum_{j' \in S_{it}} \sum_{l \in \mathcal{L}} \exp(V_{ij'l})} \right), \tag{11}$$

where V_{ijl} is the function of the parameter vector of interest (γ, β) defined in equation (4).

All the estimates presented in Section 4 are computed as the values of (γ, β) that maximize the function in equation (11). As proved in McFadden (1978), these maximands are consistent estimators of the preference parameters defined in equation (1).[17] Specifically, the fact that we maximize a likelihood function that conditions on the randomly chosen sets $\{S_{it}; \forall i, t\}$ does not affect the consistency of our estimator. However, the variance of this estimator decreases as we increase the cardinality of each set S_{it}. While our benchmark estimates use sets $\{S_{it}; \forall i, t\}$ with 20 restaurants each, Table A.3 shows that our conclusions are robust to using sets that include 50 or 100 restaurants.[18]

3.4 Discussion

The baseline model described in sections 3.1 and 3.2 embeds several key identifying assumptions. In this section, we discuss why we impose these assumptions, what they imply in our empirical context, and how we relax them in several extensions to this baseline model.

Absence of race-specific preferences for restaurants' unobserved characteristics. Except for the vector of idiosyncratic errors ν_{it}, the utility function in equation (1) assumes that consumers' preferences exclusively depend on observed restaurant characteristics (see Section 4 for a complete description of the covariates (X_i, Z_i) entering our demand model). However, individuals from different racial groups may have heterogeneous preferences for restaurants on the basis of characteristics that we do not observe. For example, group-g consumers may prefer a specific venue because this venue is frequently patronized by other customers belonging to the same group g. To explore the robustness of our baseline results to the presence of unobserved venue characteristics that determine consumers' preferences, we generalize the utility function in equation (1) to allow for race- and restaurant-specific

[17] Appendix C.2 presents simulation results that illustrate the asymptotic properties of our estimator. Section 4.4 illustrates its finite-sample properties.

[18] Using smaller simulated choice sets is computationally advantageous. For example, doubling their size roughly doubles the computation time while yielding only modestly smaller standard errors (see Table A.3).

unobserved effects:

$$U_{ijlt} = \gamma^1_{g(i)l} X^1_{ijl} + \gamma^2_{g(i)} X^2_{ij} + \beta^1_{g(i)} Z^1_j + \beta^2_{g(i)} Z^2_{ij} + \sum_{j' \in J} \alpha_{g(i)j'} \mathbb{1}\{j' = j\} + \nu_{ijlt}, \qquad (12)$$

where α_{gj} captures the group-g specific component of the utility of visiting restaurant j that is determined by unobserved characteristics.[19]

Independence of irrelevant alternatives. Our baseline model assumes that all unobserved determinants of the utility to individual i of patronizing venue j at period t from origin-mode l, as captured in the scalar ν_{ijlt}, are independent across venues, possible origins of the trip and modes of transport. To illustrate the robustness of our baseline estimates, we relax this independence assumption in multiple directions.

First, in Appendix C.3, we introduce an alternative model in which we assume that ν_{ijlt} does not vary with the origin-mode l; i.e. $\nu_{ijlt} = \nu_{ijl't}$ for any pair (l, l'). An implication of this alternative model is that, conditional on visiting a restaurant j, every individual i travels to it using the origin and mode of transport that maximizes the term $\gamma^1_{g(i)l} X^1_{ijl}$. As discussed in Section 4.1, the covariate X^1_{ijl} in our baseline model equals the (log) number of minutes that it takes individual i to reach venue j from the origin-model l; thus, assuming that $\nu_{ijlt} = \nu_{ijl't}$ and $\gamma^1_{g(i)l} = \gamma^1_{g(i)l'} < 0$ for any pair (l, l') is equivalent to assuming that users minimize travel time to each venue.

Second, in Section 4.5.2, we present estimates of nested-logit models that allow for correlation in the unobserved terms ν_{ijlt} across restaurants j and origin-modes l. Specifically, we allow the terms ν_{ijlt} to be correlated across restaurants that share a number of characteristics. Following Train et al. (1987), Appendix C.4 describes how we adapt the estimation procedure in Section 3.3 to these nested-logit demand models.

Absence of within-group unobserved parameter heterogeneity. One additional limitation of the conditional-logit model described in Section 3.1 is that it does not allow for within-group-g unobserved heterogeneity in the parameters capturing preferences for observable restaurant characteristics. The standard approach to do so is to assume that individual-specific preferences follow a known distribution in the population of interest. In our setting, estimating such a model is infeasible: unobserved heterogeneity in the vector (γ, β) makes the choice-set construction procedure in McFadden (1978) inapplicable and,

[19] Incorporating the set of fixed effects $\{\alpha_{gj}; \forall g \in \{a, b, w\}, j \in J\}$ significantly increases the dimensionality of the parameter vector we estimate. To explore computational costs, we have estimated this specification in two ways. First, a procedure similar to that in Section 3.3 but with two adjustments: (a) we add the term $\sum_{j' \in J} \alpha_{g(i)j'} \mathbb{1}\{j' = j\}$ to the expression for V_{ijl} in equation (4); (b) for each individual i and period t, we form the set S_{it} by drawing only from those restaurants visited by at least one sample user that belongs to the same group g as individual i. Importantly, as Train (2009, 3.7) discusses, while the procedure to sample restaurants described in (b) implies that our estimates of the restaurant fixed effects are biased, it does not affect the consistency of our estimates of the spatial and social frictions. Second, the Poisson approximation to the conditional-logit model described in Taddy (2015) and implemented in Gentzkow et al. (2017). In our setting, this approximation is exact if all individuals in the sample have equal expected utility from each restaurant trip. The results from these two estimation procedures are described in Section 4.5.1 and Appendix D.3, respectively.

therefore, requires estimating a likelihood function that, for each individual i and period t in our sample, depends on the actual choice set J'_{it}. This is computationally infeasible in a city with thousands of restaurants like NYC.[20]

While we do not allow for within-group g unobserved heterogeneity in preferences, the utility function in equation (1) does allow preferences to vary within groups with the observed characteristics of the home and work census tracts of each individual. Namely, X^2_{ij} and Z^2_{ij} contain interactions of individual i and restaurant j characteristics. For example, we allow users living in tracts of different income levels to value restaurants's prices and ratings differently.[21]

Non-representative sample. As described in Section 2, our estimation sample, while representative of the population of Yelp users, is not representative of the population of NYC at large. However, as long as the demand assumptions described in Section 3.1 hold, the lack of representativeness of our sample does not affect the consistency of our maximum likelihood estimator as an estimate of the preference parameters of NYC residents. The lack of representativeness of our sample may, however, affect the external validity of our estimates if the preference parameters (γ, β) are heterogeneous across individuals within each group g. One should bear in mind this limitation when interpreting the measures of consumption segregation and the counterfactuals presented in Section 5.

Exogeneity of home and work locations. Section 3.1 implicitly assumes that individuals' home and work locations are exogenously given. In practice, individuals choose where to live and work, and these locations may be determined as a function of restaurant characteristics. However, the endogenous location of home and work will not bias our estimates of the preference parameters (γ, β) if the distribution of the vector of unobserved characteristics affecting individuals' restaurant choices, ν_{it}, is independent of the characteristics determining the optimal selection of home and work location. Note that this is compatible with the vector of observed characteristics (X_i, Z_i) affecting users' endogenous home and work locations.[22]

Restaurant and origin independence of review-writing probabilities. As described in Section 3.2, our baseline model assumes that the probability that an individual writes a review about a visited restaurant does not depend on the restaurant itself nor on the origin of the trip. This allows the review-writing decision to depend upon the consumption experience

[20] Fox (2007) introduces a semiparametric estimation approach that similarly exploits the absence of unobserved heterogeneity in preference parameters to estimate discrete-choice models with large choice sets.

[21] Katz (2007) and Pakes (2010) show that there is a moment inequality approach that allows one to handle both large choice sets and unobserved heterogeneity in preferences for observed choice characteristics. We discuss in Appendix C.5 the relative advantages and disadvantages of this moment inequality approach for our particular application.

[22] In fact, if individuals' home and work locations are determined as a function of the expected utility of restaurant consumption, then our estimates are less likely to be biased the larger the set of characteristics that we explicitly control for through the vector (X_i, Z_i). Intuitively, the fewer the variables that are accounted for by the unobserved term ν_{it}, the more likely it is that this composite is independent of the characteristics determining each individual's choice of home and work location. A detailed discussion of this point is contained in Appendix C.6.

in a number of ways. First, our assumption allows arbitrary variation in the propensity to write a review across users and time. It can therefore account for the fact that users are more likely to contribute to an online platform when they are nearing a reputational reward, such as Yelp's "elite" status (Luca, 2016b). Second, our model is consistent with individuals being more likely to write reviews about dining experiences that surprised them, either negatively or positively. Surprises are, by definition, independent of the variables that are in the information set of consumers when deciding which restaurant venue to patronize and, therefore, independent of the consumers' restaurant choice.

In robustness checks, we address three possible violations of our assumption that the review-writing probability is independent of the patronage choice. First, one could claim that users are more likely to review restaurants with few prior reviews or that are not chain establishments well known by most consumers. As we show in Appendix C.7, one may control for characteristics that affect the review-writing probabilities of Yelp users by introducing them explicitly as covariates in our conditional-logit model. In Sections 4.5.3, we report estimates in which we control for a restaurant's total number of reviews and whether it belongs to a chain with more than eight NYC locations.

Second, it is possible that users are more likely to review restaurants that they want to signal they have patronized. The specifications in which we introduce race-specific restaurant fixed effects allow us to control for the possibility that, for example, individuals want to signal that they have visited a restaurant that is idiosyncratically popular with members of group g. We discuss these estimates in Section 4.5.1.

Finally, users may be more or less likely to review restaurants that they visited from a particular origin, such as a business lunch near their workplace. To address this possibility, we introduce race-mode-origin-specific fixed effects in specifications reported in Section 4.5.3.[23]

Identification via the extensive margin. We identify preference parameters using only information on the set of restaurants an individual reviews. Since Yelp users very rarely review a venue a second time, information on their visit frequency is unavailable. And since individuals only review some of the venues they visit, our information on visited venues is incomplete. However, if we employ the estimation procedure described in Section 3.3, neither of these features is an obstacle to consistently estimating individuals' preference parameters in our setting.

First, we are able to identify preference parameters without information on the intensive margin. Whenever we observe a restaurant review, this is the result of a consumer's choice to visit this venue rather than other venues, both those previously reviewed and unreviewed. Comparisons with other previously unreviewed venues are informative about consumers' preferences, while comparisons with previously reviewed venues would be misleading, since we cannot observe a review if a user visits an already-reviewed venue. Thus, there is a sample-selection process that must be addressed. Equations (7) and (10) show that, in the case of the conditional-logit model, accounting for such sample selection is equivalent to modifying the relevant choice set used to construct the conditional-logit probabilities. For each user i and period t, excluding previously reviewed restaurants from the set from

[23]This set of fixed effects also accounts for the possibility that users prefer certain origins or modes of transport for unobserved reasons.

which we draw the subset of restaurants S_{it} is thus sufficient to correctly identify preference parameters in our setting.

Second, we can identify the parameters of interest even though the information on the extensive margin of restaurants visited by each individual is imperfect. The key feature of our model that allows us to overcome this data limitation is the assumption that the probability that an individual i at period t reviews any restaurant visited and not previously reviewed is invariant across restaurants. This implies that two restaurants' relative probability of being reviewed is equal to their relative probability of being visited: restaurant reviews are informative about restaurant visits and can thus be used to identify preference parameters.

Lack of serial correlation in unobserved preferences. Our baseline model assumes that users' unobserved restaurant preferences (captured in the vector ν_{it}) are independent over time. As discussed in detail in Appendix C.8, if, contrary to our assumption, the preference shocks ν_{it} are serially correlated, the fact that we identify users' preferences from their reviews and that users do not review a restaurant twice will generate a selection bias in our estimates of consumers' preference parameters. Specifically, positive serial correlation would cause attenuation bias: upward bias in the estimates of coefficients on characteristics that consumers dislike (e.g., spatial and social frictions) and downward bias in the estimates of coefficients on characteristics that appeal to consumers (e.g., restaurants' rating). We illustrate the possible size of this bias through a simulation in Appendix C.8. To reduce this selection bias, we report estimates that use only the first half and first fifth of each user's reviews in Section 4.5.3.

4 Estimation results

This section reports the results of estimating discrete-choice models of the form described in Section 3 using the data introduced in Section 2. The models differ in the set of spatial and social frictions we incorporate. In Section 4.1, we introduce the elements of X^1_{ijl} to capture spatial frictions, while omitting X^2_{ij}. In Section 4.2, we additionally introduce the elements of X^2_{ij} to capture social frictions. In all cases, we include a set of venue and user-venue characteristics that may influence consumer demand. These correspond to the vectors Z^1_j and Z^2_{ij} in equation (1).

We measure spatial frictions as the impact that the time a consumer has to travel to visit a restaurant has on the probability of visiting this restaurant. As discussed in Section 3.1, we assume that a user making a trip to a Yelp venue may start that trip from her home location, work location, or during her commute. Similarly, she may travel via mass transit or automobile. The variable X^1_{ijl} is the log of the number of minutes it would take individual i to travel to restaurant j using the transportation mode and origin pair indexed by l. The disutility of travel time, $\gamma^1_{g(i)l}$, may vary with l because the direct pecuniary cost of an additional minute of travel time differs across modes of transportation (positive for taxis, zero for the subway). These coefficients may also vary due to heterogeneity across transportation modes or origin locations in non-pecuniary costs (e.g., cleanliness or convenience).

We measure social frictions as the effect of racial demographic differences on the probability of a consumer visiting an otherwise identical restaurant. The vector X^2_{ij} contains the Euclidean demographic distance and spectral segregation index measures introduced in Section 2, as well as their interaction. We also include the residential population share of each racial and ethnic group in the restaurant's tract (white share is the omitted category). Heterogeneity across Yelp users of different races in the estimated coefficients on these population shares would reveal social frictions.

The venue characteristics in Z_j^1 are the restaurant's price (captured by dummy variables corresponding to Yelp's four price categories) and Yelp rating, the log median household income of the tract in which the venue is located, 28 area dummies, and nine cuisine dummies.[24] The user-venue characteristics in Z_{ij}^2 are the restaurant's price and Yelp rating interacted with the user's home census tract's median household income, as well as the percent difference and absolute percent difference in median incomes between the home and venue tracts.

As discussed in sections 3.4, as we control for restaurant characteristics and area dummies in all specifications, our estimates of spatial frictions are not biased due to our sample of Yelp users living and working near generally attractive dining options. Similarly, incorporating both the income level of the restaurant's tract and the difference in incomes between the user's home tract and the restaurant's tract as controls implies that our estimates of social frictions cannot be attributed to tracts' racial demographics covarying with income levels.

All the specifications presented in the main text are estimated using a fixed set of randomly generated choice sets, $\{S_{it}, \forall i, t\}$, so that variation in the estimates across columns and tables is exclusively due to variation in the included covariates.[25]

4.1 Spatial frictions

We first estimate a specification in which home is the only origin and then sequentially introduce the work and commuting origins.

Table 2 presents estimates for nine models. In the first three columns, we find large, negative coefficients on travel times from home, with Asian users exhibiting somewhat more modest responses to spatial frictions.[26] Users are much less likely to visit venues that are far from their home via public transit or automobile.[27] The fact that the coefficients on travel time via automobile tend to be more negative than those on travel time via public transit is

[24] We aggregate NYC community districts to partition the city into 28 areas and aggregate Yelp cuisine types into 9 categories; see Appendix B.3.

[25] We have also re-estimated these specifications using different randomly sampled choice sets and in both cases we have obtained very similar estimates; see Table A.3.

[26] We find little heterogeneity in the coefficients on spatial frictions along other demographic dimensions. Unreported specifications in which spatial frictions vary with income or age yield economically small and statistically insignificant coefficients. Unreported specifications with gender-specific coefficients on spatial frictions, social frictions, and tastes yield very similar results.

[27] A potential concern could be selection bias related to the fact that we locate users based on reviews of Yelp venues located near their homes and workplaces. This is unlikely to have a large effect on our estimates. First, we locate users based on any Yelp review, not necessarily of restaurants. Thus, users in our estimation sample have not necessarily reviewed a restaurant near their home or work. Dropping the 5% of restaurant reviews used to locate users alters the coefficients of interest little. Second, estimating the travel-time coefficients separately for subsamples of users that differ in the number of Yelp reviews revealing their locational information yields similar estimates across these subsamples.

Table 2: Travel time

	(1) Asian	(2) black	(3) white/Hisp	(4) Asian	(5) black	(6) white/Hisp	(7) Asian	(8) black	(9) white/Hisp
Log travel time from home by public transit	-1.01^a	-1.40^a	-1.32^a	-1.04^a	-1.19^a	-1.24^a	-1.07^a	-.996^a	-1.15^a
	(.042)	(.061)	(.023)	(.046)	(.067)	(.030)	(.101)	(.119)	(.058)
Log travel time from home by car	-1.17^a	-2.06^a	-1.84^a	-1.17^a	-1.50^a	-1.50^a	-1.19^a	-1.24^a	-1.38^a
	(.043)	(.161)	(.048)	(.041)	(.092)	(.033)	(.086)	(.141)	(.059)
Log travel time from work by public transit				-1.38^a	-1.99^a	-1.88^a	-1.27^a	-2.16	-1.92^a
				(.085)	(.450)	(.106)	(.145)	(2.43)	(.298)
Log travel time from work by car				-1.65^a	-2.00^a	-2.01^a	-1.69^a	-2.02^a	-2.01^a
				(.078)	(.168)	(.062)	(.188)	(.584)	(.181)
Log travel time from commute by public transit							-.955^a	-.997^a	-1.11^a
							(.063)	(.098)	(.042)
Log travel time from commute by car							-1.08^a	-1.43^a	-1.46^a
							(.060)	(.171)	(.056)
Dummy for 2-dollar bin	.287^a	.696^a	.319^a	.327^a	.639^a	.313^a	.309^a	.645^a	.317^a
	(.086)	(.196)	(.083)	(.087)	(.198)	(.084)	(.087)	(.194)	(.082)
Dummy for 3-dollar bin	.148	-.216	-.074	.176	-.318	-.100	.175	-.283	-.075
	(.115)	(.345)	(.119)	(.116)	(.341)	(.121)	(.115)	(.334)	(.120)
Dummy for 4-dollar bin	.122	.095	-.386^c	.142	-.343	-.452^b	.086	-.313	-.398^c
	(.184)	(1.21)	(.217)	(.186)	(1.11)	(.222)	(.185)	(1.18)	(.219)
Yelp rating of restaurant	.511^a	.050	.316^a	.588^a	.021	.335^a	.583^a	.036	.335^a
	(.063)	(.138)	(.059)	(.064)	(.139)	(.060)	(.064)	(.137)	(.059)
African cuisine category	.268	-.099	.297	.294	-.090	.343	.271	-.046	.319
	(.296)	(.548)	(.261)	(.297)	(.547)	(.262)	(.297)	(.548)	(.259)
American cuisine category	.426^a	.533^a	.614^a	.420^a	.539^a	.624^a	.421^a	.542^a	.596^a
	(.054)	(.119)	(.051)	(.055)	(.119)	(.051)	(.054)	(.118)	(.050)
Asian cuisine category	.944^a	.157	.320^a	.948^a	.172	.328^a	.931^a	.201	.308^a
	(.054)	(.133)	(.055)	(.054)	(.134)	(.055)	(.054)	(.132)	(.054)
European cuisine category	.201^a	-.383^b	.236^a	.200^a	-.360^b	.250^a	.204^a	-.339^b	.247^a
	(.059)	(.155)	(.056)	(.060)	(.155)	(.057)	(.059)	(.153)	(.056)
Indian cuisine category	.373^a	-.386	.017	.375^a	-.527^c	-.006	.374^a	-.422	-.018
	(.091)	(.300)	(.098)	(.092)	(.306)	(.099)	(.091)	(.299)	(.097)
Latin American cuisine category	.496^a	.993^a	.694^a	.493^a	1.02^a	.711^a	.491^a	1.03^a	.699^a
	(.070)	(.136)	(.062)	(.070)	(.136)	(.062)	(.070)	(.134)	(.061)
Middle Eastern cuisine category	.242^b	.120	.212^b	.245^b	.092	.218^b	.264^a	.066	.204^b
	(.101)	(.250)	(.096)	(.101)	(.252)	(.096)	(.100)	(.250)	(.094)
Vegetarian/vegan cuisine category	.394^a	-.005	.625^a	.372^a	-.014	.635^a	.365^a	-.041	.596^a
	(.137)	(.410)	(.116)	(.139)	(.409)	(.117)	(.138)	(.408)	(.116)
2-dollar bin × home tract median income	.042^a	-.005	.047^a	.039^a	.003	.051^a	.041^a	-.002	.049^a
	(.011)	(.032)	(.010)	(.011)	(.032)	(.010)	(.011)	(.032)	(.009)
3-dollar bin × home tract median income	.087^a	.105^c	.087^a	.086^a	.120^b	.096^a	.086^a	.109^b	.089^a
	(.014)	(.055)	(.013)	(.014)	(.053)	(.013)	(.014)	(.052)	(.013)
4-dollar bin × home tract median income	.080^a	-.180	.103^a	.082^a	-.090	.115^a	.088^a	-.119	.105^a
	(.021)	(.231)	(.022)	(.022)	(.208)	(.023)	(.022)	(.224)	(.022)
Yelp rating × home tract median income	.019^b	.005	.020^a	.010	.009	.018^a	.010	.007	.017^a
	(.008)	(.023)	(.007)	(.008)	(.023)	(.007)	(.008)	(.023)	(.007)
Percent absolute difference in median incomes $(h_i - k_j)$	-.000	.681^a	-.035	-.141^a	.485^a	-.308^a	-.218^a	.469^a	-.350^a
	(.047)	(.117)	(.047)	(.046)	(.117)	(.046)	(.045)	(.114)	(.046)
Percent difference in median incomes $(k_j - h_i)$	-.381	.843	.575^b	-.226	1.29	.676^b	-.233	1.04	.791^a
	(.283)	(.826)	(.286)	(.291)	(.847)	(.298)	(.292)	(.826)	(.293)
Log median household income in k_j	.336	-.693	-.395	.141	-1.03	-.552^b	.119	-.869	-.694^a
	(.250)	(.731)	(.251)	(.257)	(.753)	(.263)	(.258)	(.733)	(.259)
Number of origin-mode points	2	2	2	4	4	4	6	6	6
Number of trips	6447	1079	6936	6447	1079	6936	6447	1079	6936

NOTES: Each column reports an estimated conditional-logit model of individuals' decisions to visit a Yelp venue. Standard errors in parentheses. Statistical significance denoted by a (1%), b (5%), c (10%). Unreported controls are 28 area dummies.

consistent with the fact that NYC public-transit fares are invariant to distance, while taxi fares are not.

In the next three columns, we introduce the work origin. This is potentially important for measuring consumption segregation.[28] In principle, the opportunity to consume starting from work could disconnect consumption segregation from residential segregation. In practice, our estimates imply that users start many trips from home. The coefficients on travel times from home are similar to their values in the first three columns.[29] The coefficients on travel times from work are roughly 30% to 40% larger than those on travel times from home, consistent with the opportunity cost of travel time from work being greater than that from home.

In the final three columns of Table 2, we account for the possibility that users travel to venues from their commuting path. The home and work coefficients are largely unchanged, and the commuting path travel times exhibit significant negative coefficients.[30]

Table 2 yields a very clear finding: spatial frictions play a first-order role in individuals' consumption choices within the city. Consumers are less likely to visit venues that, in terms of mass-transit and automobile travel time, are more distant from their home and work locations, as well as the commuting path between these. Consider two hypothetical restaurants, identical in their characteristics except for the number of minutes away from the user's optimal origin of the trip. The first restaurant is 15 minutes from the user's workplace by car; the second restaurant is 30 minutes away. The estimated coefficients in the last three columns of Table 2 imply that the user would be about four times as likely to visit the more proximate venue from work by car (e.g., for a black user, $2^{2.02} \approx 4.06$). Similarly, if the two restaurants were 15 and 30 minutes from the commuting path by public transit, the user would be about twice as likely to visit the more proximate venue from his or her commute by public transit.[31]

Finally, note that users' choices also depend upon restaurants' characteristics in predictable ways. Restaurants with higher ratings and lower prices are generally more attractive. However, restaurants in the $$ price category are attractive relative to $ restaurants, indicating that prices also reflect quality.[32] Users residing in census tracts with higher incomes exhibit more willingness to pay higher prices. Asian users' most preferred cuisine category is Asian cuisine, while white/Hispanic users' most preferred categories are Latin

[28] Further, including the workplace origin is important for our subsequent identification of social frictions. Compare Tables 3 and A.4.

[29] Since our coefficients' scale is determined by the standard deviation of the logit error ν_{ijlt}, comparisons of the levels of coefficients across columns implicitly assume that this standard deviation is constant. Conversely, ratios of estimates of the same specification are invariant to this normalization. The coefficients on travel times from home relative to the coefficients on restaurant characteristics (e.g., Yelp rating and cuisine categories) are similar across columns. Thus, if we were to normalize our estimates by the coefficient on any of these restaurant characteristics instead, we would also conclude that introducing the work origin does not significantly alter the estimates of the coefficients on travel time from home.

[30] The coefficients on spatial frictions are less precisely estimated for black users, due to the smaller number of observations. Upon controlling for social frictions in Table 3, all spatial-friction coefficients are statistically significant.

[31] Besides the impact that spatial frictions have for consumption segregation, our quantification of these frictions is an important input for models of consumption within the city (e.g., Allen et al. 2015). Similar to the finding that commuting flows are consistent with gravity (Monte et al., 2015), our estimates suggest gravity for consumption.

[32] Restaurants in the $ category are often fast-food venues.

American and vegetarian/vegan.

4.2 Social frictions

Understanding the role that social frictions play in determining consumption choices is key to our effort to measure consumption segregation. In this section, we use both tract-level and individual-level demographic information in order to distinguish between two potential phenomena. First, users of all demographics may be more likely to choose restaurants located in neighborhoods populated by races and ethnicities similar to those characteristic of the tract in which they reside, a role for environmental similarity. Second, individuals may be more likely to choose restaurants in places populated by individuals similar to themselves, homophily. An example of the first would be if a white individual living in Harlem, a heavily black and Hispanic neighborhood, is more likely to visit black and Hispanic neighborhoods than a white individual living elsewhere. An example of the second would be if a white individual living in Harlem is more likely to visit a restaurant in a white neighborhood than other residents of her home tract. Table 3 estimates these social frictions by introducing the vector of racial and ethnic demographic covariates, X^2_{ij}.

The negative coefficients on the Euclidean demographic distance (EDD) measure reveal a role for environmental similarity. Users are less likely to visit venues located in census tracts with demographics different from those of their home census tract. Since we control for transit times between tracts, this result cannot be attributed to the joint impact of residential segregation and disutility of travel. Similarly, our controls include income differences between tracts, so this result cannot be attributed to spatial differences in incomes predicting consumers' choices. Thus, the coefficient on EDD likely captures mechanisms linked to racial and ethnic differences. Users may have preferences regarding the residential demographics of the neighborhoods in which they consume. Alternatively, users may be more likely to visit restaurants located near the residences of friends and family members, with these relationships being more likely across members of the same race or ethnicity. Regardless of the mechanisms underlying the negative coefficient on EDD, these social frictions result in individuals of different races and ethnicities making different consumption choices.

The estimated coefficient on EDD implies an economically significant role for this social friction. Consider a user who contemplates visiting two venues that are identical except for their EDD from the home census tract, which differ by one standard deviation. Our estimates imply that an Asian user would be 25% more likely to visit the venue in the more demographically similar census tract.[33] A black user would be 51% more likely to visit the more similar tract. We can also express the economic significance of demographic differences as a trade-off between demographic distance and transit time. To hold constant an Asian user's utility of visiting a venue from home via public transit, a venue one standard deviation more demographically distant would have to be about 21% closer in terms of travel time.[34]

[33]Comparing two venues j and j' that are identical in every covariate except X^2_{ij}, $P(d_{ij} = 1|V_i)/P(d_{ij'} = 1|V_i) = \exp(\gamma^2_{g(i)}(X^2_{ij} - X^2_{ij'}))$. Table A.2 shows that the standard deviation of EDD across all pairs of census tracts in NYC is 0.226, so the coefficient of -1 in column one of Table 3 implies that a venue that has EDD 0.226 lower than an otherwise identical venue will be visited with 25% higher probability $(\exp(-1.00 \times (-0.226)) \approx 1.25)$.

[34]To hold U_{ijlt} constant, a change of ΔX^2_{ij} would be offset by the change $\Delta X^1_{ijl} = -\gamma^2_{g(i)}\Delta X^2_{ij}/\gamma^1_{g(i)l}$.

Table 3: Social frictions

	(1) Asian	(2) black	(3) white/Hisp
Log travel time from home by public transit	-1.06^a	-.938^a	-1.13^a
	(.107)	(.127)	(.059)
Log travel time from home by car	-1.17^a	-1.19^a	-1.36^a
	(.091)	(.158)	(.060)
Log travel time from work by public transit	-1.24^a	-1.85^c	-1.87^a
	(.149)	(1.11)	(.287)
Log travel time from work by car	-1.60^a	-1.79^a	-1.95^a
	(.176)	(.459)	(.171)
Log travel time from commute by public transit	-.943^a	-.930^a	-1.10^a
	(.067)	(.105)	(.044)
Log travel time from commute by car	-1.04^a	-1.32^a	-1.43^a
	(.061)	(.177)	(.058)
Euclidean demographic distance between h_i and k_j	-1.00^a	-1.84^a	-1.19^a
	(.121)	(.280)	(.130)
Spectral segregation index of k_j	.150^a	.075	.045^c
	(.051)	(.093)	(.027)
EDD × SSI	-.149	-.171	-.068
	(.117)	(.239)	(.083)
Share of tract population that is Asian	1.03^a	.011	.363^a
	(.120)	(.345)	(.138)
Share of tract population that is black	.220	1.08^a	.140
	(.319)	(.399)	(.265)
Share of tract population that is Hispanic	-.251	.467	.415^b
	(.235)	(.381)	(.188)
Share of tract population that is other	.059	3.56	.484
	(2.07)	(3.43)	(1.99)
Dummy for 2-dollar bin	.375^a	.771^a	.355^a
	(.087)	(.197)	(.083)
Dummy for 3-dollar bin	.287^b	-.090	-.026
	(.116)	(.341)	(.120)
Dummy for 4-dollar bin	.220	-.074	-.347
	(.188)	(1.22)	(.221)
Yelp rating of restaurant	.579^a	.053	.344^a
	(.064)	(.138)	(.059)
African cuisine category	.280	-.198	.298
	(.299)	(.553)	(.261)
American cuisine category	.432^a	.523^a	.591^a
	(.054)	(.119)	(.050)
Asian cuisine category	.886^a	.255^c	.307^a
	(.054)	(.134)	(.054)
European cuisine category	.195^a	-.326^b	.235^a
	(.059)	(.154)	(.056)
Indian cuisine category	.370^a	-.451	-.039
	(.091)	(.301)	(.097)
Latin American cuisine category	.517^a	1.01^a	.690^a
	(.070)	(.136)	(.062)
Middle Eastern cuisine category	.280^a	.104	.203^b
	(.101)	(.251)	(.094)
Vegetarian/vegan cuisine category	.392^a	.001	.587^a
	(.138)	(.409)	(.116)
2-dollar bin × home tract median income	.034^a	-.022	.042^a
	(.011)	(.032)	(.009)
3-dollar bin × home tract median income	.075^a	.077	.081^a
	(.014)	(.053)	(.013)
4-dollar bin × home tract median income	.074^a	-.167	.095^a
	(.022)	(.234)	(.023)
Yelp rating × home tract median income	.011	.008	.016^b
	(.008)	(.023)	(.007)
Percent absolute difference in median incomes ($h_i - k_j$)	-.062	.850^a	-.100^c
	(.050)	(.126)	(.053)
Percent difference in median incomes ($k_j - h_i$)	.114	.619	.719^b
	(.305)	(.853)	(.300)
Log median household income in k_j	-.109	-.360	-.625^b
	(.267)	(.744)	(.262)
Average annual robberies per resident in k_j	-3.41^a	2.43^b	-3.74^a
	(.676)	(1.20)	(.771)
Number of origin-mode points	6	6	6
Number of trips	6447	1079	6936

NOTES: Each column reports an estimated conditional-logit model of individuals' decisions to visit a Yelp venue. Standard errors in parentheses. Statistical significance denoted by *a* (1%), *b* (5%), *c* (10%). "EDD" is Euclidean demographic distance; "SSI" is spectral segregation index. Unreported controls are 28 area dummies.

For a black user, it would have to be 44% closer.

The consequences of demographic differences vary across users of different races due to differences in population sizes. Users of all races are more likely to visit restaurants that have lower values of the EDD covariate, and black users have the most negative coefficient on EDD. Yet the mean value of EDD for venues visited by black users (.34) is in fact greater than the mean value of EDD for venues *not* visited by white users (.29). This finding is consistent with the idea in Anderson (2015, p.10) that "white people typically avoid black space, but black people are required to navigate the white space as a condition of their existence."

Do demographic differences between census tracts matter more when the venue is located deep within a segregated area? To assess this, we use a spectral segregation index (SSI) that describes a tract's demographic isolation in terms of its racial or ethnic plurality. In Table 3, the coefficients on both SSI and the interaction of EDD and SSI are modest in magnitude. When EDD is close to one, their sum is close to zero. Thus, a restaurant in a tract near the edge of a racially or ethnically distinct area is about as likely to be visited as a tract with the same demographic differences located deep inside that area. Individuals' choices are therefore mostly predicted by the demographic composition of the area immediately surrounding the restaurant.

Users also exhibit homophily at the individual level. Users are more likely to visit venues located in tracts with a larger residential population of the same race as the user. Asian users are more likely to visit a restaurant in a tract with more Asian residents relative to all other residential racial and ethnic categories (white residents are the omitted category). Similarly, black users are more likely to visit a restaurant in a tract with more black residents. In the case of white or Hispanic users, homophily is less evident, perhaps due to the fact that we do not distinguish between whites and Hispanics when classifying users' profile photos. In column three, there is a positive coefficient on Hispanic population share, but there is also a positive coefficient on Asian residents relative to white residents.

These social frictions are distinct from the roles of income and crime. Since we incorporate both the income level of the restaurant's tract and the difference in incomes between the user's home tract and the restaurant's tract as controls, the demographic-differences results cannot be attributed to tracts' demographics covarying with income levels. We also control for users' responses to robberies per resident in the destination tract.[35]

Relative to Table 2, the coefficients on spatial frictions and cuisine categories in Table 3 are slightly attenuated toward zero. Residential segregation means that spatial frictions and social frictions are positively correlated, so spatial frictions will be overestimated if social frictions are omitted. Similarly, if restaurants in tracts with more Asian residents are more likely to serve Asian cuisine that appeals to Asian users, then the cuisine coefficients will be overestimated if social frictions are omitted. Comparing Tables 2 and 3 suggests that

Since the coefficient $\gamma^1_{g(i)hp}$ estimated in column one of Table 3 is -1.06, the change required to offset a one-standard-deviation increase in EDD is $-1.00 \times 0.226/1.06 \approx -0.21$.

[35] Fear of crime is still a possible explanation for the estimated responses to demographic differences to the extent that fears are not highly correlated with actual crime rates (Quillian and Pager, 2001). Crime rates, fear of crime, and residential segregation are interrelated phenomena (O'Flaherty and Sethi, 2015). Unreported specifications interacting users' home tracts' median incomes with actual robbery rates do not change our parameters of interest.

this occurs, but in the vast majority of cases the estimated coefficients differ by less than a standard error.[36]

This section has documented patterns in consumer behavior that tend to segregate consumption. We find roles for both environmental similarity and individual homophily. Regardless of the particular mechanisms underlying how demographic differences shape consumption in the city, our quantification indicates that these social frictions play an important role in shaping consumer behavior. These elements will contribute to our estimates of urban consumption segregation in Section 5.

4.3 Model fit

In this section, we discuss how well our estimated model fits the data.

4.3.1 In-sample isolation

We first compare our model's prediction of consumption segregation to that observed in the estimation sample. Following Gentzkow and Shapiro (2011), we compute isolation indices using "leave-out" means to address finite-sample bias. Denote the number of visits to venue j by members of racial group g by $v_{gj} = \sum_{i:g(i)=g} \sum_t d^*_{ijt}$, the number of visits to all venues by those members by $v_g = \sum_j v_{gj}$, the number of visits to venue j by users who are not members of group g by $v_{\neg g,j} = \sum_{i:g(i)\neq g} \sum_t d^*_{ijt}$, and the number of visits to venue j by all users by $v_j = \sum_g v_{gj}$. The Gentzkow and Shapiro (2011) "leave-out isolation index" that measures the extent to which members of group g disproportionately visit venues whose other visitors are also members of group g is

$$\hat{S}_g = \sum_j \frac{v_{gj}}{v_g} \cdot \left(\frac{v_{gj} - 1}{v_j - 1} \right) - \sum_j \frac{v_{\neg gj}}{v_{\neg g}} \cdot \left(\frac{v_{\neg gj}}{v_j - 1} \right).$$

To generate a model-predicted value of $\hat{S}_g$ that is comparable to that in the data, we simulate model-predicted visits to restaurants for the set of observations in the estimation sample. Our estimated model predicts that each user will visit a venue with a probability given by equation (5). One draw for each observation from this probability distribution generates one simulated value of $\hat{S}_g$. We simulate the model 500 times to obtain a distribution of $\hat{S}_g$ values. The value observed in the estimation sample and the 90% confidence interval for simulated values are presented in Table 4.

The actual data exhibit values of $\hat{S}_g$ within the 90% simulated confidence intervals. Thus, our estimated model matches observed consumption segregation well. Appendix D shows that allowing preference parameters to vary across races is key to matching the observed

[36] Recall that all our coefficients are normalized relative to the standard deviation of the logit error ν_{ijlt}, which need not be identical across specifications. The change in the coefficients on spatial frictions and cuisine categories across models should be compared to the change in a coefficient that is plausibly not biased by the omission of social frictions, such as the Yelp rating of the restaurant. Comparing Tables 2 and 3, the coefficient on Yelp rating is slightly attenuated for Asian users and actually larger for black and white/Hispanic users. Thus, the attenuation of coefficients on spatial frictions and cuisine categories in Table 3 is not solely attributable to a change in the standard deviation of ν_{ijlt} across models.

Table 4: Model fit: Isolation indices

	Estimation sample	Model predictions
Asian isolation index	.087	[.056, .089]
Black isolation index	.087	[.043, .089]
White/Hispanic isolation index	.045	[.025, .057]

NOTES: The reported leave-out isolation indices are the value for the estimation sample and the 90% confidence interval for model-predicted outcomes for generated samples of the same size. Isolation indices as defined in Gentzkow and Shapiro (2011).

consumption segregation: more restrictive specifications that pool preference parameters across races underpredict the isolation indices observed in the data.

4.3.2 Schelling-style segregation

Since we cannot observe the complete racial and ethnic composition of the patrons of every restaurant in NYC, our baseline specification assumes that consumer preferences do not depend on this restaurant characteristic.[37] Thus, our baseline model predicts that two restaurants with the same observable characteristics (tract, cuisine, price, and rating) will exhibit the same racial composition of patrons. This limits the scope for social frictions by ruling out Schelling-style outcomes in which high degrees of segregation arise due to preferences for same-race co-patrons.

To examine the plausibility of this assumption, we collect information on the racial composition of all Yelp reviewers for 119 pairs of restaurants that are identical in terms of their cuisine category, price category, Yelp rating, and census tract.[38] Define the "race gap" within a pair of restaurants p to be the Euclidean distance

$$gap_p \equiv \|share_j - share_{j'}\|/\sqrt{2},$$

with $share_j$ being a three-element vector of the fraction of users reviewing restaurant j who are Asian, black, and Hispanic/white.[39] We compare the observed distribution of gap_p for the 119 pairs of observationally equivalent restaurants to the one that would arise if, consistent with our model, individuals were randomly assigned to one of the two restaurants within each pair.

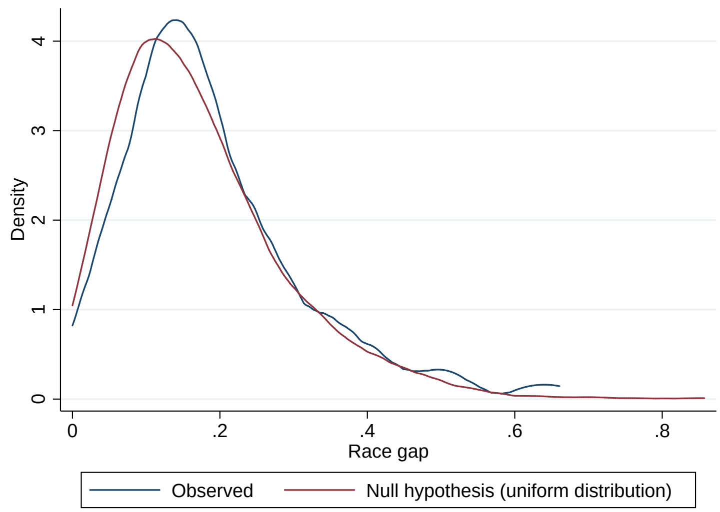
Figure 6 depicts the distribution of gap_p for both the data and the random draws. The mean of the race gap for the observed data is 0.189. The mean for the random distribution is 0.172. The p-value for the one-sided test of equal means is 0.06. More broadly, the two distributions seem sufficiently similar that the omission of the demographic composition of patrons from the observable covariates we use to predict restaurant visits seems unlikely to radically alter our results.

[37] When we introduce race-specific restaurant fixed effects, we implicitly allow the racial composition of a restaurant's patrons to enter each racial group's valuation of that restaurant.

[38] See Appendix D for details. For reasons of feasibility, we photo-coded users who reviewed restaurants that belonged to a tract-cuisine-price-rating pair and had between 10 and 40 reviews.

[39] We drop users whose race is not determined based on their photos from these computations.

Figure 6: Racial gap between pairs of observationally equivalent restaurants



If consumption were segregated within sets of restaurants that our model treats as observationally equivalent, one might worry that our model would underpredict the true degree of consumption segregation. Our examination suggests that this is not the case. Conditional on the observable covariates that we employ to predict consumption segregation, Yelp reviews do not exhibit much further racial segregation.

4.4 Parametric bootstrap

To examine the finite-sample behavior of our estimator, we perform a parametric bootstrap. Using the estimated model reported in Table 3, we simulate 500 samples of the same size as our estimation sample. We then estimate our model on each of these generated samples, obtaining a distribution of our estimator under the estimated data-generating process. Appendix D.5 reports the results in detail.

Generally, the parametric bootstrap shows that the estimator performs well. For the covariates describing social frictions and restaurant characteristics, the bootstrapped distributions are close to normal, their means are close to the point estimates we computed on the original sample, and their standard deviations are very similar to our estimated (asymptotically valid) standard errors. For the spatial-friction covariates, the bootstrapped samples occasionally produce estimates that are extreme outliers. This seems due to the fact that we identify these six spatial-friction parameters exclusively from restaurant-reviewing outcomes $d^*_{ij} = \sum_l d^*_{ijl}$, without actually observing the origin-mode-level outcomes d^*_{ijl}, and that transit times from the same origin are highly collinear.[40] When we assume that the error term ν_{ijlt} and the disutility of travel do not vary across origin-modes l – implying that there is a single spatial-friction parameter to estimate and that consumers visit each restaurant via the origin-mode pair with the minimum travel time (see Appendix C.3) – the standard error

[40]For example, the correlation between travel time from work by car and from work by public transit exceeds 0.9 for all three racial groups.

we compute for our estimator of the disutility caused by this spatial friction is very similar to the bootstrapped one.

4.5 Robustness checks

4.5.1 Restaurant fixed effects

It is feasible to introduce race-specific restaurant fixed effects into our model. However, as we cannot identify these fixed effects for restaurants that are not visited by users in the estimation sample, this generalized model cannot be used to compute city-wide measures of consumption segregation and counterfactuals.[41] We therefore employ the specification with race-specific restaurant fixed effects only to examine the robustness of the coefficients on observable user-restaurant covariates reported in Table 3 and to assess the relative fit of these specifications.

The results suggest that our baseline specification in Table 3 is sufficient to capture the relevant variation in consumers' choices. Table 5 reports the result of estimating the specification with restaurant fixed effects. The estimated coefficients on our measures of spatial and social frictions are similar to those reported in Table 3. Table 6 reports the result of a likelihood-ratio test comparing the fit of the restaurant-fixed-effects specification to the specifications in Table 6 that use only observable characteristics. For Asian and black users, we do not reject the hypothesis that the observables specification fits the data as well as the fixed-effects specification. For white/Hispanic users, the fixed-effects specification is superior in terms of fit, but the estimated coefficients are largely consistent with those in Table 3.

4.5.2 Nested-logit specification

We relax the independence-of-irrelevant-alternatives property of the conditional-logit model of Section 3.1 by specifying a nested-logit structure. Appendix C.4 derives this estimator in detail. We define nests by two schemes: (a) restaurants of the same cuisine category, Yelp rating, and area, and (b) restaurants of the same cuisine category, price category, and census tract. For this exercise, we define 39 cuisine categories (more disaggregated than the nine categories shown in Tables 2 and 3, and employ the 28 areas, four price categories, and nine Yelp ratings (from 1 star to 5 stars) described in Section 2. These two schemes group the 10,945 restaurants into 3,064 and 7,622 nests, respectively. Table D.3 reports the estimates. The estimated coefficients on spatial frictions, social frictions, and restaurant characteristics are all similar to the values reported in Table 3. The within-nest-correlation parameter λ is generally near one, consistent with the conditional-logit assumption we imposed in Section 3.1. Table D.3 reports the p-values for likelihood ratio tests of the null hypothesis that the true model is a conditional logit versus the alternative hypothesis that the true model is

[41] To extrapolate beyond the restaurants reviewed by users in the estimation sample, one must employ observable characteristics. Therefore, to compute the magnitudes reported in sections 5 and 6, a fixed-effects specification offers little advantage relative to one employing observed characteristics. Furthermore, estimating models with restaurant fixed effects is computationally costly: estimation takes days rather than minutes. This is true in our setting whether we estimate our fixed effects directly or assume the approximation in Taddy (2015). We report the results of the Taddy (2015) estimation procedure in Appendix D.3.

Table 5: Restaurant fixed effects

	(1) Asian	(2) black	(3) white/Hisp
Log travel time from home by public transit	-1.08[a]	-1.01[a]	-1.23[a]
	(0.109)	(0.143)	(0.062)
Log travel time from home by car	-1.18[a]	-1.23[a]	-1.48[a]
	(0.091)	(0.161)	(0.063)
Log travel time from work by public transit	-1.29[a]	-1.64[a]	-1.81[a]
	(0.152)	(0.562)	(0.207)
Log travel time from work by car	-1.73[a]	-1.78[a]	-2.07[a]
	(0.200)	(0.394)	(0.170)
Log travel time from commute by public transit	-1.00[a]	-1.00[a]	-1.15[a]
	(0.074)	(0.113)	(0.041)
Log travel time from commute by car	-1.12[a]	-1.34[a]	-1.57[a]
	(0.068)	(0.170)	(0.062)
Euclidean demographic distance between h_i and k_j	-0.78[a]	-2.38[a]	-1.28[a]
	(0.132)	(0.321)	(0.145)
EDD × spectral segregation index	-0.39[b]	-0.28	0.01
	(0.177)	(0.355)	(0.083)
2-dollar bin × home tract median income	0.04[a]	-0.01	0.04[a]
	(0.011)	(0.034)	(0.010)
3-dollar bin × home tract median income	0.06[a]	0.07	0.08[a]
	(0.014)	(0.056)	(0.013)
4-dollar bin × home tract median income	0.06[b]	-0.16	0.09[a]
	(0.022)	(0.208)	(0.023)
Yelp rating × home tract median income	0.02[c]	0.02	0.02[a]
	(0.010)	(0.026)	(0.009)
Percent difference in median incomes ($k_j - h_i$)	-0.17[a]	1.36[a]	-0.14[b]
	(0.057)	(0.178)	(0.060)
Percent absolute difference in median incomes ($h_i - k_j$)	0.02	0.00	0.45
	(0.360)	(1.025)	(0.377)
Number of trips	6447	1079	6936
Number of fixed effects	2867	892	3497

NOTES: Each column reports an estimated conditional-logit model of individuals' decisions to visit a Yelp venue. Standard errors in parentheses. Statistical significance denoted by *a* (1%), *b* (5%), *c* (10%). The unreported covariates are restaurant fixed effects.

Table 6: Observables versus restaurant fixed effects

Sample	Log likelihood values: Observables	Log likelihood values: Restaurant fixed effects	χ^2 test statistic	p-value
Asian users	-55339.28	-50243.65	10191.26	1.00
Black users	-9198.20	-7152.89	4090.62	1.00
White/Hispanic users	-59213.50	-53507.34	11412.32[a]	0.00

NOTES: The observables specifications include 60 covariates, while the fixed-effect specifications include 14 covariates plus the 10,945 restaurant fixed effects. Thus, the χ^2 test has 10,899 degrees of freedom. Statistical significance denoted by *a* (1%), *b* (5%), *c* (10%).

the corresponding nested-logit model. While the tests formally reject the hypothesis that $\lambda = 1$ for three of the six specifications, these nested-logit specifications yield very similar coefficients on observable covariates, and deliver very similar predictions of the in-sample isolation measures (see Table D.1). Since estimating these nested-logit models comes at considerably greater computational cost, we employ the conditional-logit specification when computing consumption segregation and counterfactual outcomes.

4.5.3 Additional observable characteristics and sample restrictions

In Appendix A, we implement several robustness checks by restricting the set of observations or introducing additional covariates.

To address concerns that the error term ν_{ijlt} may exhibit serial correlation (discussed in Section 3.4), Table A.7 restricts the sample to either the first half or the first fifth of reviews written by each user. These restrictions increase the standard errors and cause one coefficient to be unidentified in the resulting small sample of black users. However, they do not systematically increase the absolute values of coefficients, suggesting that there is not substantial attenuation bias caused by serial correlation in the unobserved preference shocks ν_{ijlt}. For a more detailed discussion, see Appendix C.8.

To address concerns that users may be more or less likely to review restaurants visited from a particular origin (see Section 3.4), Table A.8 introduces origin-mode-specific intercepts. This yields broadly similar coefficients on log travel times, though the accompanying standard errors are considerably larger. The coefficients on social frictions are modestly attenuated.

To address concerns that the decision to write a review may depend on certain restaurant characteristics (see Section 3.4), Table A.9 controls for two additional covariates: the restaurant's number of Yelp reviews and whether it belongs to a restaurant chain. While users in our estimation sample are more likely to review more-reviewed restaurants and less likely to review establishments belonging to a restaurant chain with more than eight NYC locations, the coefficients of interest change little.

To address concerns that early adopters of Yelp may be less sensitive to spatial and social frictions, Table A.6 restricts the sample to users who joined Yelp later. The estimated coefficients are broadly similar to those in Table 3, without any systematic increase or decrease. In an unreported specification, we find that our results change little when controlling for tract-level differences in private vehicle ownership rates. Finally, Table A.5 shows that controlling for 39 cuisine categories (rather than 9) slightly attenuates our point estimates of homophily and makes the environmental similarity coefficients slightly more negative.

In summary, the results in Table 3 are broadly unchanged by restricting the estimation sample or introducing additional covariates.

5 Consumption segregation

In this section, we use data on the demographic composition of all census tracts in NYC and the estimates presented in Table 3 to compute NYC-wide measures of segregation in consumption for Asian, black, Hispanic, and white consumers. In Section 5.1, we define the measure of segregation we use. Section 5.2 presents our estimates of consumption segregation and examines the contributions of spatial and social frictions to them. In Section 5.3, we illustrate the mechanisms underlying the city-wide results by focusing on the consumption patterns observed in particular neighborhoods within the city. Finally, in Section 5.4 we examine a number of counterfactual experiments.

5.1 Dissimilarity indices

To measure consumption segregation, we use the "dissimilarity index" commonly employed in the literature on residential segregation. For each group $\mathfrak{g}$, we compute

$$\text{Dissimilarity}(\mathfrak{g}) = \frac{1}{2}\sum_{j \in J} |\Pr(d_{ij} = 1|\mathfrak{g}(i) = \mathfrak{g}) - \Pr(d_{ij} = 1|\mathfrak{g}(i) \neq \mathfrak{g})|, \tag{13}$$

where $\mathfrak{g}$ is *Asian*, *black*, *Hispanic*, *white*, or *other*.[42] This index sums, across all restaurants, the absolute difference between the probability that a randomly selected individual belonging to group $\mathfrak{g}$ visits a restaurant and the probability that a randomly selected individual who does not belong to group $\mathfrak{g}$ visits the same restaurant. The higher the value of this index, the larger the differences in consumption choices. Think of the probability $\Pr(d_{ij} = 1|\mathfrak{g}(i) = \mathfrak{g})$ as the fraction of individuals of group $\mathfrak{g}$ that would visit restaurant j if they were all going out to dine; the dissimilarity index in equation (13) indicates the share of individuals in group $\mathfrak{g}$ that would need to alter their consumption choices in order to match the distribution of predicted restaurant choices made by the remainder of the population. This measure can also be computed to compare the consumption choices of any two groups, $\mathfrak{g}_1$ and $\mathfrak{g}_2$.[43] This dissimilarity index is invariant to the size of the groups being compared.

If we were to observe all visits to restaurants for a sufficiently large representative sample of NYC residents, we could estimate $\Pr(d_{ij} = 1|\mathfrak{g}(i) = \mathfrak{g})$ directly by its sample analogue. To our knowledge, such a large and representative dataset does not exist.[44] We therefore apply the parameters estimated in Section 4 to the broader population of NYC in order to consistently estimate the probability $\Pr(d_{ij} = 1|\mathfrak{g}(i) = \mathfrak{g})$ for all restaurants j and races

[42] As discussed in Section 3, we estimate parameters for three racial groups indexed by g. However, we compute dissimilarity indices for five racial groups indexed by $\mathfrak{g}$. We compute $\Pr(d_{ij} = 1|\mathfrak{g}(i) = \mathfrak{g})$ for Hispanic, white, and other using $\mathfrak{g}$-specific data on residential locations and the estimated white/Hispanic ($g = w$) preference parameters.

[43] The bilateral dissimilarity between any two groups $\mathfrak{g}_1$ and $\mathfrak{g}_2$ is computed as: $\text{Dissimilarity}(\mathfrak{g}_1, \mathfrak{g}_2) = (1/2)\sum_{j \in J} |P(d_{ij} = 1|\mathfrak{g}(i) = \mathfrak{g}_1) - P(d_{ij} = 1|\mathfrak{g}(i) = \mathfrak{g}_2)|$.

[44] As explained by Gentzkow et al. (2017), computing consistent measures of segregation that do not depend on estimates of a behavioral model requires using a sample that is large relative to the dimensionality of the choice set faced by individuals. In our context, this sample would have to be large relative to the number of restaurants in NYC. The advantage of the behavioral model introduced in Section 4 is that it expresses the probabilities $\Pr(d_{ij} = 1|\mathfrak{g}(i) = \mathfrak{g})$ as a function of a relatively small number of estimated parameters.

$\mathfrak{g}$.[45] Appendix E details how we construct consistent estimates of these probabilities.[46] A further advantage of relying on the demand model described in Section 3 as the basis for computing our measures of consumption segregation is that it can also be used to quantify the contributions of spatial and social frictions to our overall measure of consumption dissimilarity.

5.2 City-wide consumption segregation

Table 7 reports residential and consumption dissimilarity indices for each demographic group. The upper panel reports overall dissimilarity indices for each group; the lower panel reports dissimilarity indices across specific pairs of groups. We compute residential dissimilarity at the level of census tracts and consumption dissimilarity at the level of restaurant venues.[47]

Column one provides the dissimilarity index for residential segregation while column two provides the analogous index for consumption segregation. Comparing these columns, all groups are significantly more integrated in their consumption than in their residences. Across demographic groups, black and Asian users exhibit the highest values of consumption dissimilarity, but this is in part due to the fact that we assign white and Hispanic consumers the same preference parameters due to our inability to differentiate between white and Hispanic Yelp users in their photos. To the extent that white and Hispanic consumers differ in their preferences, we will underestimate consumption segregation for these two groups. In the lower panel, the largest pairwise dissimilarities are found between Asian and black consumers. Black and white consumers' choices are also dissimilar, while Hispanic-white and black-Hispanic consumption choices are more integrated.[48]

The fact that all groups are significantly more integrated in their consumption than in their residences is not a necessary consequence of the model assumptions imposed in Section 3. While the willingness of residents to travel outside of their home census tract

[45] Intuitively, equation (5) expresses the probability that i visits j as a function of the observed characteristics of i and j, reducing the number of parameters to estimate from the number of restaurants times the number of demographic groups to the dimensionality of the vector $(\{\gamma_l^1; \forall l\}, \gamma^2, \beta^1, \beta^2)$. Furthermore, estimating $\Pr(d_{ij} = 1 | \mathfrak{g}(i) = \mathfrak{g})$ for all restaurants j and racial groups $\mathfrak{g}$ using estimates of behavioral parameters allows us to obtain consistent estimates of the dissimilarity index in equation (13) even if our sample is not a random sample from the population of interest, as long as it is random conditional on the observed characteristics of individuals and restaurants.

[46] Given the estimates presented in Section 4 and the data described in Section 2, the only additional information required to compute $\Pr(d_{ij} = 1 | \mathfrak{g}(i) = \mathfrak{g})$ for all NYC residents is the joint distribution of home and work locations across census tracts for each demographic group. We obtain this information from tract-level commuting flows.

[47] The choice of spatial unit is non-trivial when computing dissimilarity indices, as Echenique and Fryer (2007) stress. Restaurants are a natural spatial unit within which interactions may occur, while tract-level results may be sensitive to how the Census Bureau chose to partition the city. To facilitate comparison of measures of residential and consumption segregation, Table A.12 reports consumption dissimilarity indices computed at the level of census tracts. The resulting indices are broadly similar to those computed at the level of restaurant venues.

[48] Despite their common preference parameters, white and Hispanic consumers still exhibit pairwise dissimilarity due to observable differences in residential locations that generate differences in the social-friction covariates employing tract-level racial and ethnic demographics (e.g., EDD) and other covariates employing tract-level characteristics (e.g. the interaction of price dummies and median household income in the user's home tract).

Table 7: Residential and consumption segregation

		Residential dissimilarity	Consumption dissimilarity			
			Estimated	No spatial	No social	Neither friction
		(1)	(2)	(3)	(4)	(5)
Dissimilarity index						
Asian		.521	.315	.290	.244	.232
Hispanic		.486	.143	.114	.106	.088
black		.653	.352	.322	.273	.260
white		.636	.191	.153	.112	.093
Pairwise dissimilarity						
Asian	Hispanic	.584	.288	.273	.220	.217
Asian	black	.796	.494	.448	.387	.357
Asian	white	.519	.279	.255	.212	.203
Hispanic	white	.658	.161	.115	.096	.037
black	Hispanic	.558	.328	.297	.262	.250
black	white	.822	.354	.324	.262	.255

NOTES: This table reports dissimilarity indices. The upper panel reports the index for each demographic group's residential/consumption locations compared to members of all other demographic groups. The lower panel reports the index for residential/consumption locations between each pair of demographic groups. The demographic group "other" is included in computations but not reported. Column 1 reports indices based on tracts' residential populations. The remaining columns report venue-level dissimilarity indices based on the coefficient estimates in Table 3. Column 2 uses the estimated coefficients. Column 3 sets the coefficients on travel-time covariates to zero. Column 4 sets the coefficients on demographic-difference covariates to zero. Column 5 sets the coefficients on travel-time and demographic-difference covariates to zero.

to consume may tend to reduce consumption segregation relative to residential segregation, demographic differences in cuisine tastes or demographically-linked social frictions could cause consumption segregation to exceed residential segregation. Our numbers show that social frictions and heterogeneity in tastes do not overturn the integrating effect of consumers' mobility.

Across ethnic and racial groups, residential dissimilarity is much greater than consumption dissimilarity. The ratio of residential dissimilarity to consumption dissimilarity is 3.4 for Hispanics, 3.4 for whites, 1.9 for blacks, and 1.7 for Asians. NYC's levels of residential dissimilarity are similar to the nationwide average level of dissimilarity for black residents in 2010, while the levels of dissimilarity in consumption implied by our estimates are comparable to the levels of black residential dissimilarity observed in America's most integrated metropolitan areas (Glaeser and Vigdor, 2012).[49] The 20 to 45 percentage-point difference

[49] Glaeser and Vigdor (2012) do not report dissimilarity indices for Asians, Hispanics, or whites. A conventional interpretation of residential dissimilarity indices is that indices below 0.3 are "low," those between 0.3 and 0.6 indicate "moderate" segregation, and dissimilarity in excess of 0.6 is "high" (Massey and Denton, 1993, p. 20). We thus use these as benchmarks for gauging the magnitude of consumption dissimilarity. There is no implication that rankings of cities by residential dissimilarity indices will have the same ranking by consumption dissimilarity.

between consumption segregation and residential segregation of NYC residents is one to two times the largest declines in black residential dissimilarity observed from 1970 to 2010 across US metropolitan areas (Glaeser and Vigdor, 2012). At the median historical rate of decline, residential segregation would have to continue its decline for nearly a century to reach levels comparable to our estimated levels of consumption segregation.

In order to measure the contributions of spatial and social frictions to consumption segregation, we again use the estimates in Table 3 and recompute the dissimilarity indices that arise from setting some of the estimated coefficients to zero. This calculation holds both the set of restaurants and their characteristics fixed and, thus, should not be interpreted as capturing the total effect of eliminating spatial or social frictions (which would likely generate supply responses). In computing the dissimilarity indices reported in column three of Table 7, the coefficients on travel-time covariates are set to zero, eliminating the role of spatial frictions.[50] In column four, the coefficients on the demographic-differences covariates are set to zero, eliminating the role of social frictions.[51] In column five, both these sets of coefficients are set to zero, so that consumers in different groups exhibit different predicted consumption behavior only due to: differences in their residential income levels and income-linked valuations of venues' prices and ratings; race-specific valuations of restaurants' cuisines, prices, and ratings; race-specific responses to robberies per resident; and race-specific area fixed effects.[52]

Comparing columns two and three, the elimination of spatial frictions has a mild integrating effect, causing consumption dissimilarity to fall by an average of three percentage points. That is, users from different demographic groups value restaurant destinations sufficiently similarly that eliminating spatial frictions would, all else equal, make their choices modestly more integrated. Due to spatial frictions, consumption segregation at least partly inherits the pattern of residential segregation. Comparing columns two and four, eliminating the roles of environmental dissimilarity and homophily (i.e., social frictions), reduces consumption dissimilarity by an average of 6.6 percentage points, or more than twice the effect of eliminating spatial frictions. If consumer behavior did not respond to differences between the residential demographics of the restaurant tract and both the consumer's individual identity and home tract demographics, predicted consumption behavior would be much more integrated.[53] Eliminating both spatial and social frictions would reduce consumption dissimilarity by about one third on average.

The relative contributions of spatial and social frictions are strikingly consistent across demographic groups. For each group, social frictions make a notably greater contribution to the observed level of consumption dissimilarity than spatial frictions.[54] Moreover, the

[50]This is a *ceteris paribus* exercise. In reality, spatial frictions and social frictions may not be entirely independent. For example, if social frictions reflect segregated friendship networks and users visiting restaurants near their friends' residences, then the elimination of spatial frictions would eliminate this component of social frictions.

[51]Specifically, we set the coefficients on all the "EDD", "SSI", and "share" covariates reported in Table 3 to zero.

[52]For example, Asian users are, all else equal, more likely to visit Asian and Indian restaurants, and users residing in higher-income tracts are more likely to visit restaurants with higher prices.

[53]This decomposition holds residential segregation fixed. However, one could expect that residential segregation would be different in the absence of social frictions.

[54]The relative importance of spatial and social frictions for consumption dissimilarity depends on both

relative overall levels of estimated consumption dissimilarity appear to reflect dissimilarity attributable to demographic differences in tastes. In the absence of both spatial and social frictions, black consumers would exhibit the greatest consumption dissimilarity, and Hispanic consumers the least, just as they do in the estimated levels in column two.

The pairwise dissimilarity indices reported in the lower panel of Table 7 reflect the rich set of covariates incorporated in our behavioral model. For example, while Asian-black and black-white residential dissimilarity indices are similar, consumption dissimilarity is notably greater between Asian and black consumers than between black and white consumers. This partly reflects the magnitude of the Asian-black interactions in Table 3. Column five of Table 7 reveals that it also reflects divergent choices due to differences in income levels and tastes.

These results are robust to a number of the alternative estimating assumptions discussed in Section 4.5.3. In particular, we obtain similar results when restricting the estimation sample to users who joined Yelp later (Table A.13) and introducing origin-mode-specific intercepts (Table A.14). In addition, constraining all trips to start at home yields broadly similar results (Table A.15).

5.3 Illustrative examples

We illustrate the mechanisms behind the results in Table 7 in two parts of the city. First, we examine three neighborhoods in Manhattan: the Upper East Side, Central Harlem, and East (Spanish) Harlem (respectively Manhattan community districts 8, 10, and 11). The fact that each of these neighborhoods is residentially segregated, with a distinct demographic majority, makes the general process at work easy to visualize in Figure 7.[55]

Panel A in Figure 7 captures residential segregation. Each dot represents 5% of the tract population. The Upper East Side of Manhattan stretches from Fifty-Ninth Street to Ninety-Sixth Street, Central Park to the East River. While NYC is only 33% white, the Upper East Side is 81% white. If we restrict attention to the fourteen census tracts between Third Avenue and Central Park, the median tract is 92% white. Among nearly 61,000 residents of these tracts, only 726 were black. In short, this is a highly segregated area of the city.

Central Harlem comprises Fifth Avenue to Eighth Avenue, Central Park North (110th Street) to the Harlem River. This is the storied center of urban black America. While its black population share has fallen from its 1990 level of 88%, it remains 63% black. The next largest group is Hispanics at 22%, with modest levels of whites (10%) and Asians (2%). While Central Harlem's residential population is becoming more diverse, panel A of Figure 7 makes it evident that it remains a highly segregated area.

East (Spanish) Harlem stretches from Ninety-Sixth Street to the Harlem River, and Fifth Avenue to the East River. The Hispanic fraction of the population has remained roughly

the estimated parameters and the values of the covariates capturing these frictions across all NYC residents. In terms of the correlation between spatial frictions and demographic differences, NYC is not an outlier. For Census places with population greater than 100,000, the distribution of the correlation of Euclidean demographic distance and physical distance across census tracts in the same place has mean 0.27 and standard deviation 0.18. For NYC, this correlation is 0.14.

[55]Table A.16 reports shares of residents and predicted consumers by race for the three community districts depicted in these maps.

Figure 7: Residential and consumption segregation in three Manhattan communities

A: Residential B: Estimated C: No spatial D: No social E: Neither friction

NOTES: These maps depict community districts 8 (Upper East Side), 10 (Central Harlem), and 11 (East Harlem) in Manhattan. The five maps correspond to the five scenarios reported in Table 7. Each dot represents five percent of the tract's residential population or predicted restaurant visitors. Asian residents or consumers are represented by red dots, black by blue, Hispanic by orange, and white by green. In different tracts, dots represent a different number of people, so the maps depict variation in shares, not levels.

constant in the last twenty years at about 50%. There is a large, even if declining, black population in East Harlem, at roughly 30 percent, located most densely where East Harlem abuts Central Harlem to the west and in the more northerly areas of the district. Asians and whites are present in small but growing numbers.

Panel B of Figure 7 shows the degree of consumption segregation within these areas. Each dot represents five percent of the visits to that tract. Two features jump out from this panel. The first is that predicted consumption in panel B is strikingly less segregated than residences in panel A. This is consistent with a comparison of columns one and two of Table 7. The boundaries between black and Hispanic consumers are porous, consistent with the black-Hispanic interactions in Table 3 and the pairwise dissimilarity index in Table 7. Asian consumers are more prevalent in the Upper East Side, for example, than Central Harlem, reflecting both a shorter distance to Asian residential population centers and smaller social frictions between Asians and whites than between Asians and the two other groups. The second feature is that, nonetheless, there remains a very high level of segregation.

As summarized in Table A.16, black consumers dominate consumption in Central Harlem; Hispanics in East Harlem; and whites in the Upper East Side. Segregation of consumption is much less than residences but still strong.

The following three panels in Figure 7 follow the last three columns of Table 7 by illustrating the degree of consumption segregation for the respective cases in which spatial, social, or both types of frictions are set to zero when constructing predicted consumption patterns. The third panel is based on our estimates in which we wholly eliminate spatial frictions, effectively making every restaurant in the city instantly available to any resident of the city. Comparing panel C to B, there is a diminution of the degree of consumption segregation. Yet the change seems modest, consistent with a comparison of columns two and three of Table 7.

When we move to panel D of Figure 7, we allow spatial frictions to again be at their estimated level but now set social frictions to zero. Visually, comparing panels B and D in Figure 7, we see a large decline in the degree of consumption segregation in each of these neighborhoods. It is important to recognize that this did not need to be true. Residential segregation plus spatial frictions to consumption could have been enough to maintain very high levels of consumption segregation; as we have observed, they just happen not to do so. Social frictions matter a great deal for consumption integration, consistent with column four of Table 7.

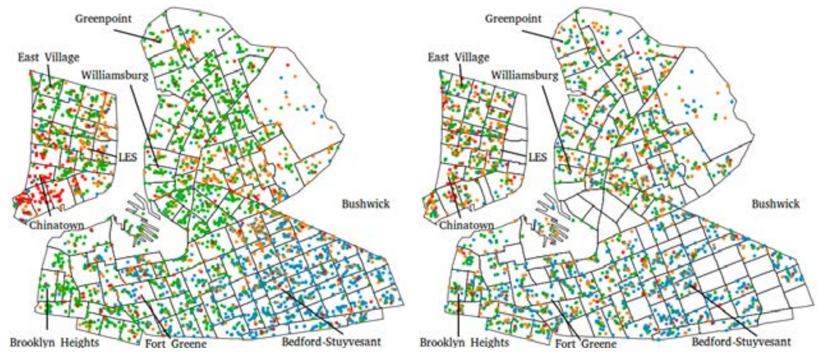
Finally, we consider the elimination of both spatial and social frictions by setting both sets of coefficients to zero. The fifth column of Table 7 tells us that there is some further modest decline in the degree of consumption segregation relative to the no-social-frictions case. However, it is sufficiently modest that it does not stand out clearly to the eye in panel E of Figure 7 in any of the three neighborhoods. This again suggests that spatial frictions explain a small share of consumption segregation.

Next, we examine another part of New York City with considerable, but less extreme, segregation. Figure 8 depicts neighborhoods near the portion of the East River separating the Lower East Side of Manhattan from Brooklyn.[56] Panel A show strong patterns of residential segregation. Manhattan community district 3 includes Chinatown (with predominantly Asian residents), the East Village (with predominantly white residents in the northern and western portions) and the Lower East Side (with predominantly Hispanic residents). Across the river, Brooklyn community district 1 is home to concentrations of white residents in Greenpoint and Williamsburg and concentrations of Hispanic residents, especially in the areas leading out to (mostly Hispanic) Bushwick. Brooklyn community district 2 includes the mostly white Brooklyn Heights as well as Fort Greene, with a mixture of black, Hispanic, and white residents. Finally, Bedford Stuyvesant (Brooklyn community district 3) is a traditionally black area that now has white residents in the area near Fort Greene and Hispanic residents in areas proximate to Williamsburg and Bushwick.

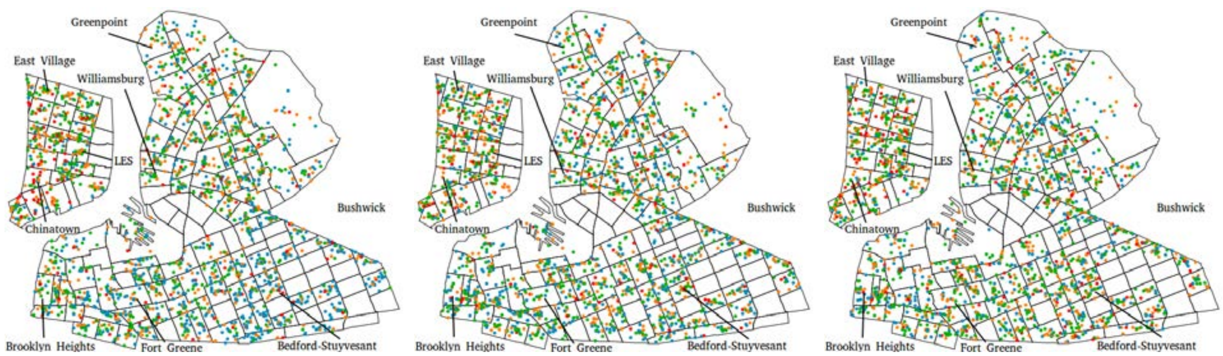
Panel B of Figure 8 shows that estimated consumption is strikingly more integrated. Chinatown shows considerable inflows of white and Hispanic consumers, who are proximate residents, but more modest inflows of black consumers, who are more remote. The East Village and Lower East Side are also notably more integrated, again with only modest numbers

[56] Table A.17 reports shares of residents and predicted consumers by race for the four community districts depicted in these maps.

Figure 8: Residential and consumption segregation in lower Manhattan and west Brooklyn



A: Residential B: Estimated



C: No spatial D: No social E: Neither

NOTES: These maps depict community district 3 in Manhattan and 1, 2, 3, and 4 in Brooklyn. Each dot represents 20% of the tract's residential population or predicted restaurant visitors. See notes to Figure 7.

of black consumers. Greenpoint and Williamsburg host predominantly white consumers who are augmented by Hispanic consumers residing, presumably, in Williamsburg and Bushwick. Brooklyn Heights and Fort Greene are similarly host to mostly white consumers augmented by black and Hispanic consumers. Finally, consumption in Bedford-Stuyvesant is notably more racially integrated than its residences.

Panels C, D, and E of Figure 8 reaffirm our finding that consumption segregation is driven more by social frictions than spatial frictions. To the eye, panel C, which sets spatial frictions to zero, is nearly identical to panel B. For example, the areas of Chinatown and Bedford-Stuyvesant continue to be dominated by Asian and black consumers, respectively. By contrast, panel D, which sets social frictions to zero, noticeably differs from panel B. The consumption enclaves of Chinatown and Bedford-Stuyvesant don't entirely disappear, but they are much more integrated. Indeed, setting both frictions to zero, as in panel E, does little to integrate consumption more visibly than the removal of social frictions alone.

5.4 Counterfactuals

We next study how changes in transportation infrastructure and in the level of social frictions may affect the level of consumption segregation we observe. Major changes in transportation policy have only modest effects on consumption dissimilarity, consistent with the small contribution of spatial frictions to overall consumption segregation. Conversely, reductions in social frictions (due, for example, to an improvement in racial tolerance) would integrate consumption to a greater degree.

5.4.1 Transportation policy and technology

Table 7 shows that the complete elimination of spatial frictions would only modestly reduce consumption dissimilarity indices. This bounds the conceivable impact of driverless cars or other "frictionless" technologies on consumption segregation. How might more immediately feasible transportation projects affect consumption segregation? We consider two counterfactuals relevant to NYC policymakers, noting that interventions may reduce or increase spatial frictions. First, we forecast the effect of the new Second Avenue Subway on consumption segregation. Second, we look at the effects of a general slowdown in NYC transit.

The Second Avenue Subway is an ongoing multi-billion-dollar expansion of the NYC subway system. When completed, the line will stretch from 125th Street in East Harlem all the way down the East Side to Hanover Square in the Financial District. We compute counterfactual travel times for this transportation infrastructure improvement and forecast the effect on consumption segregation.[57] The results are reported in Table 8. The Second Avenue subway line has almost no effect on consumption segregation. This is the joint effect of the fact that spatial frictions play a relatively small role in determining consumption segregation and that the Second Avenue line is inferred to have relatively small effect on travel times for the majority of the city's residents.

Second, we study consumption segregation when automobiles and public transit are 20% slower. In 2014, NYC lowered the speed limit within the city from 30 mph to 25 mph.[58] MTA subway speeds are also down about 20% in the last few years.[59] The effects on consumption segregation are heterogeneous, with increases for Asian, Hispanic, and white consumers, and a decrease for black consumers. This notwithstanding, all of the magnitudes are small. Our conclusion is that even quite substantial interventions in the transport sphere are unlikely to have a major impact on the integration of consumption.

5.4.2 Social frictions

Next we examine the effects of a decline in demographic-linked social frictions on consumption segregation. At first glance, this may appear to be an odd policy exercise, since social frictions likely reflect a variety of factors, such as tastes and social networks, that are not immediately under policymakers' control. Yet this does not really differentiate it from many other policy exercises in which we distinguish between targets (e.g., vehicle speed) and instruments (e.g., speed limits, policing, etc.). There are government policy initiatives at the

[57] Appendix E details how we construct the counterfactual transit times. Our computation only captures the value of new subway connections in the network graph and does not assign any value to benefits such as alleviating overcrowding, which is a major motivation for the Second Avenue project.

[58] Caroline Bankoff, "Things to Know About NYC's New 25-Miles-Per-Hour Speed Limit", *New York*, 7 November 2014.

[59] See Hamilton Nolan, "How the MTA Is Like an Alcoholic", *Fusion*, 15 June 2017, and Emma G. Fitzsimmons, "Subway's Slide in Performance Leaves Straphangers Fuming", *New York Times*, 12 Feb 2017.

Table 8: Counterfactual consumption dissimilarity

		Estimated	2nd Ave	Slowdown	Social change
		(1)	(2)	(3)	(4)
Dissimilarity index					
Asian		.315	.315	.318	.303
black		.352	.352	.349	.326
Hispanic		.143	.143	.146	.127
white		.190	.191	.192	.154
Pairwise dissimilarity					
Asian	black	.494	.494	.493	.469
Asian	Hispanic	.288	.288	.295	.276
Asian	white	.279	.278	.280	.268
black	Hispanic	.328	.328	.325	.307
black	white	.354	.354	.352	.318
Hispanic	white	.161	.161	.164	.128

NOTES: See notes to Table 7. This table reports venue-level dissimilarity indices based on the coefficient estimates in Table 3. Column 1 uses the estimated coefficients. Columns 2-4 implement counterfactuals.

federal, state, and municipal levels that aim to encourage understanding and prevent tensions between different demographic groups.[60] In short, government policies often aim to affect some elements that may contribute to what we term social frictions.

The counterfactual that we examine is the reduction of the magnitude of social frictions by 50 percent.[61] The results are reported in the fourth column of Table 8. The direct effect of reducing social frictions on consumption segregation is modest relative to the overall level of consumption segregation, in spite of the fact that these effects are one to two orders of magnitude greater than the impact of the transportation projects. The residual consumption segregation reflects taste differences that may well attenuate over time with reductions in other dimensions of segregation, but are unlikely to disappear over any short or even medium horizon.

[60] At the federal level, the Community Relations Service within the Department of Justice "is the Department's 'Peacemaker' for community conflicts and tensions arising from differences of race, color, national origin, gender, gender identity, sexual orientation, religion and disability." It "facilitates the development of viable, mutual understandings and solutions to the community's challenges." In New York City, the Commission on Human Rights has a dual mandate, to educate the public about legal protections for a variety of groups and "encouraging understanding and respect among New York City's many communities."

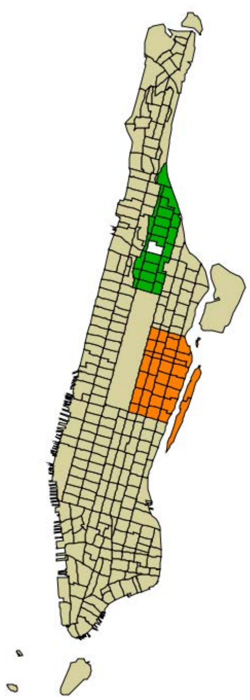
[61] Our choice of this magnitude is admittedly arbitrary and ambitious. Halving social frictions might be an implausibly large change. On one side, 42% of Americans were worried "a great deal" about race relations in 2017, a peak since Gallup began asking the question in 2001. On the other, race relations improved considerably over recent decades in a number of dimensions. To mention a single measure, approval of black-white marriage rose from 4% in 1958 to 87% in 2011. The realm of conceivable changes in social frictions is quite broad, and we are not aware of evidence disciplining the conceivable magnitude of policy-induced changes in social frictions.

6 Gentrification

Public concerns about gentrification are primarily about displacement of one group by another. Nominally this is an issue of social class, but in practice it also has a racial component – neighborhoods shifting composition from less affluent minorities toward more affluent whites. Since tastes for consumption venues differ by class and race, there is a second concern that the commercial mix may also shift away from the set of venues well suited to poor and minorities to those favored by affluent whites. This amounts to a concern about the potential for gentrification to erode the benefits of living in a neighborhood for those who stay behind. In New York City, the presence of rent-stabilized apartments and New York City Housing Authority buildings, with low turnover rates, makes this question salient.

We employ our estimates to quantify the welfare impact of gentrification for incumbent residents by examining the consequences of changes in resident and restaurant composition. The gentrification scenario we study is depicted in Figure 9. We select one low-income, majority-black census tract in Harlem (the white polygon) and compute the change in black residents' expected utility if the surrounding census tracts containing (in green) were to exhibit the residential and restaurants characteristics of high-income, majority-white census tracts of the Upper East Side (in orange). The changes in restaurant and residential characteristics are summarized in the table accompanying Figure 9.

Figure 9: Harlem gentrification scenario



Change in	Mean	Std. Dev.
Share Asian	0.055	0.038
Share black	-0.559	0.158
Share Hispanic	-0.123	0.06
Share white	0.639	0.178
Robberies per resident	-0.005	0.003
Spectral segregation index	-1.112	0.379
Yelp rating	0.009	1.087
Price ($ to $$$$)	0.509	0.991
Median household income	77.19	53.795
Euclidean demographic distance	1.24	0.828
N	108	

Notes: We compute the change in black residents' expected utility in the white polygon if the surrounding green tracts were to exhibit the characteristics of the orange tracts. The table reports the changes in these characteristics.

As a result of this gentrification, black residents of the unchanged Harlem census tract experience a 1.1% decrease in the expected utility of patronizing restaurants. This welfare

Table 9: Welfare losses due to gentrification of surrounding Harlem neighborhoods

Welfare change	Initial visit share	Change in value of characteristics ($\gamma\Delta\bar{X}_i, \beta\Delta\bar{Z}_i$)		
		Social frictions	Restaurant traits	Other traits
-1.3%	.113	-3.048	-.088	.689

NOTES: Welfare loss is $100\times\frac{U_i'-U_i}{U_i}$ for incumbent black residents. Initial visit share is $\sum_{j\in\mathcal{J}^G} P_{ij}$. Social frictions include EDD, SSI, EDD×SSI, and racial and ethnic population shares of k_j. "Restaurant traits" include price, rating, cuisine category, and price and rating interacted with median household income. "Other traits" include destination income, differences in incomes, and robberies per resident.

loss can be decomposed using a simple approximation of the utility change that we derive in Appendix E.3:

$$U_i' - U_i \approx \left[\sum_{j\in\mathcal{J}^G} P_{ij}\right] \times \left[\exp(\gamma_g^2\Delta\bar{X}_{ij}^2 + \beta_g^1\Delta\bar{Z}_j^1 + \beta_g^2\Delta\bar{Z}_{ij}^2) - 1\right],$$

where $\mathcal{J}^G$ is set of restaurants that change due to gentrification, P_{ij} is the probability that an incumbent resident would visit restaurant j prior to gentrification, and $(\Delta\bar{X}_{ij}^2, \Delta\bar{Z}_j^1, \Delta\bar{Z}_{ij}^2)$ are the average changes in restaurant and residential characteristics due to gentrification. Table 9 reports the results. The 108 restaurants in the gentrifying area account for 12.4% of predicted visits by incumbent residents prior to gentrification, so changes in the characteristics of these restaurants and residents could have large welfare effects. The welfare loss we compute is attributable to increases in social frictions associated with the shift of the surrounding tracts from mostly black residents to mostly white residents. This is partially offset by the accompanying increase in income levels, which make these tracts more attractive to incumbent consumers, ceteris paribus. The changes in restaurants' price and cuisines have very small effects on welfare.

This exercise illustrates potential welfare costs of gentrification to incumbent residents beyond increases in housing rents they may experience. In this Harlem example, as well as a Brooklyn example reported in Appendix E.3, the consumption value of the location for incumbent residents falls by a modest but non-trivial amount. This decline is not due to changes in restaurants' characteristics but increased social frictions associated with changes in surrounding neighborhoods' racial demographics.

7 Conclusions

How segregated is urban consumption? While racial discrimination has been outlawed for decades, the current level of integration in consumption venues has been unmeasured. Consumption segregation is an important component of social segregation. It may also be a lost opportunity for casual contact that can promote greater tolerance and understanding. Our paper develops an approach to provide measures of urban consumption segregation.

We use a novel data source to describe restaurant consumption in NYC and exploit properties of the conditional-logit discrete-choice model to tractably identify how consumers value venues' and locations' characteristics. Our dataset allows us to characterize how consumption in the city depends on travel times, demographic differences, crime rates, restaurant characteristics, and user characteristics. We then use our estimates to compute measures of consumption segregation for NYC residents. While NYC is distinctive in terms of its population density and diversity of restaurants, we believe that our paper's results for the largest city in the United States are both interesting in their own right and establish a basis for studying consumption segregation in other settings.

Both spatial and social frictions influence consumption choices. Consistent with theories of spatial competition, spatial frictions play a large role in determining the spatial distribution of consumption within the city. Our estimates show that measures of travel time, from both home and work by both public transit and by car, are relevant for predicting the restaurants patronized by NYC Yelp users. Across origin-mode pairs, halving the minutes of travel time to a venue would imply that the user would be two to nearly four times more likely to visit the venue from that origin by that mode.

Social frictions are suggested by the finding that consumers are less likely to visit restaurants in neighborhoods with different residential demographics. A venue in a location one standard deviation more demographically distant from a user's home location is 25% to 50% less likely to be visited. Users are also more likely to visit restaurants in neighborhoods where a larger fraction of the residents share the user's race. These social frictions are asymmetric, in the sense that the negative effects of tract-level demographic differences are larger for black consumers, yet black consumers experience larger tract-level demographic differences during their average restaurant visit.

While our estimation approach exploits data on the decision to eat at restaurants across NYC, the consequences of spatial and social frictions are likely to apply to a much broader scope of life in the city. These would include both the broader scope of consumption of non-tradable services, from bars to retailers, and the vast array of non-market activities that cause residents to traverse the city. For example, in our gentrification exercises, the reduction in the value of restaurant consumption for incumbent residents is much more attributable to increases in social frictions than changes in restaurant characteristics. These demographic changes presumably have consequences for many non-restaurant dimensions of urban life.

We use our estimates to characterize predicted consumption segregation for the city's population. While spatial frictions, social frictions, and demographic differences in tastes cause dissimilarity in consumption choices across racial and ethnic groups, dissimilarity indices for consumption are considerably lower than the dissimilarity indices for residential locations. Life in NYC is less segregated than one might infer from looking at residential segregation alone. Our analysis of these patterns reveals that social frictions contribute more to consumption segregation than spatial frictions. A consequence of this finding is that improved transportation linkages within the city would only modestly integrate consumption further, given existing residential patterns. Eliminating social frictions would result in substantially more integrated consumption.

References

Allen, Treb, Costas Arkolakis, and Xiangliang Li, "Optimal City Structure," July 2015.

mimeo.

Allport, Gordon W., *The Nature of Prejudice*, Addison-Wesley, 1954.

Anderson, Elijah, *The Cosmopolitan Canopy: Race and Civility in Everyday Life*, W.W. Norton & Company, 2011.

_ , ""The White Space"," *Sociology of Race and Ethnicity*, 2015, *1* (1), 10–21.

Anderson, Michael L. and Jeremy R. Magruder, "Learning from the Crowd: Regression Discontinuity Estimates of the Effects of an Online Review Database," *Economic Journal*, 2012, *122* (563), 957–989.

Andrews, Donald W. K. and Gustavo Soares, "Inference for Parameters Defined by Moment Inequalities Using Generalized Moment Selection," *Econometrica*, 2010, *78* (1), 119–157.

Antecol, Heather and Deborah A Cobb-Clark, "Racial and Ethnic Discrimination in Local Consumer Markets: Exploiting the Army's Procedures for Matching Personnel to Duty Locations," *Journal of Urban Economics*, 2008, *64* (2), 496–509.

Ayres, Ian, *Pervasive Prejudice? Unconventional Evidence of Race and Gender Discrimination*, University of Chicago Press, 2001.

Boustan, Leah, "Racial Residential Segregation in American Cities," in Gerrit-Jan Knaap Nancy Brooks, Kieran Donaghy, ed., *The Oxford Handbook of Urban Economics and Planning*, Oxford University Press, 2011.

Caetano, Gregorio and Vikram Maheshri, "Homophily and Sorting Within Neighborhoods," September 2015.

Chetty, Raj, Nathaniel Hendren, Patrick Kline, and Emmanuel Saez, "Where is the Land of Opportunity? The Geography of Intergenerational Mobility in the United States," *The Quarterly Journal of Economics*, 2014, *129* (4), 1553–1623.

Couture, Victor, "Valuing the Consumption Benefits of Urban Density," May 2015. mimeo.

Cutler, David M. and Edward L. Glaeser, "Are Ghettos Good or Bad?," *The Quarterly Journal of Economics*, 1997, *112* (3), 827–872.

Dickstein, Michael and Eduardo Morales, "What do Exporters Know?," December 2016. mimeo.

Echenique, Federico and Roland G. Fryer, "A Measure of Segregation Based on Social Interactions," *The Quarterly Journal of Economics*, 2007, *122* (2), 441–485.

Edelman, Benjamin and Michael Luca, "Digital Discrimination: The Case of Airbnb.com," January 2014. Harvard Business School Working Paper 14-054.

Eizenberg, Alon, "Upstream Innovation and Product Variety in the U.S. Home PC Market," *The Review of Economic Studies*, 2014, *81* (3), 1003–1045.

—, Saul Lach, and Merav Yiftach, "Retail Prices in a City: An Empirical Analysis," July 2017. mimeo.

Fox, Jeremy T., "Semiparametric Estimation of Multinomial Discrete-Choice Models Using a Subset of Choices," *RAND Journal of Economics*, December 2007, *38* (4), 1002–1019.

Gentzkow, Matthew and Jesse M. Shapiro, "Ideological Segregation Online and Offline," *The Quarterly Journal of Economics*, 2011, *126* (4), 1799–1839.

—, —, and Matt Taddy, "Measuring Polarization in High-dimensional Data: Method and Application to Congressional Speech," May 2017. mimeo.

George, Lisa and Joel Waldfogel, "Who Affects Whom in Daily Newspaper Markets?," *Journal of Political Economy*, 2003, *111* (4), 765–784.

Glaeser, Edward L. and Jacob L. Vigdor, "The End of the Segregated Century: Racial Separation in America's Neighborhoods, 1890-2010," Civic Report 66, Manhattan Instititute January 2012.

—, Jed Kolko, and Albert Saiz, "Consumer City," *Journal of Economic Geography*, January 2001, *1* (1), 27–50.

Guiso, Luigi, Paola Sapienza, and Luigi Zingales, "Cultural Biases in Economic Exchange?," *The Quarterly Journal of Economics*, 2009, *124* (3), 1095–1131.

Handbury, Jessie, "Are Poor Cities Cheap for Everyone? Non-Homotheticity and the Cost of Living Across U.S. Cities," July 2013. mimeo.

— and David E. Weinstein, "Goods Prices and Availability in Cities," *The Review of Economic Studies*, 2015, *82* (1), 258–296.

Harrison, Cassandra, Mohip Jorder, Henri Stern, Faina Stavinsky, Vasudha Reddy, Heather Hanson, HaeNa Waechter, Luther Lowe, Luis Gravano, and Sharon Balter, "Using Online Reviews by Restaurant Patrons to Identify Unreported Cases of Foodborne Illness—New York City, 2012–2013," *Morbidity and Mortality Weekly Report*, May 2014, *63* (20), 441–445.

Hellerstein, Judith K. and David Neumark, "Workplace Segregation in the United States: Race, Ethnicity, and Skill," *The Review of Economics and Statistics*, August 2008, *90* (3), 459–477.

Holmes, Thomas J., "The Diffusion of Wal-Mart and Economies of Density," *Econometrica*, 2011, *79* (1), 253–302.

Houde, Jean-François, "Spatial Differentiation and Vertical Mergers in Retail Markets for Gasoline," *American Economic Review*, 2012, *102* (5), 2147–82.

Kain, John F., "Housing Segregation, Negro Employment, and Metropolitan Decentralization," *The Quarterly Journal of Economics*, 1968, *82* (2), 175–197.

Katz, Michael, "Supermarkets and Zoning Laws," 2007. Harvard PhD dissertation.

Labaton, Stephen, "Denny's Restaurants to Pay $54 Million in Race Bias Suits," *New York Times*, May 25 1994.

Lee, Jennifer, "The Salience of Race in Everyday Life: Black Customers' Shopping Experiences in Black and White Neighborhoods," *Work and Occupations*, 2000, *27*, 353–376.

Logan, John R., Zengwang Xu, and Brian J. Stults, "Interpolating U.S. Decennial Census Tract Data from as Early as 1970 to 2010: A Longitudinal Tract Database," *The Professional Geographer*, 2014, *66* (3), 412–420.

Luca, Michael, "Reviews, Reputation, and Revenue: The Case of Yelp.com," March 2016. Harvard Business School Working Paper No. 12-016.

_ , "User-Generated Content and Social Media," in David Stromberg Simon P. Anderson, Joel Waldfogel, ed., *Handbook of Media Economics*, Vol. 1B, North Holland, 2016.

Massey, Douglas S. and Nancy A. Denton, "The Dimensions of Residential Segregation," *Social Forces*, 1988, *67* (2), 281–315.

_ and _ , *American Apartheid: Segregation and the Making of the Underclass*, Harvard University Press, 1993.

Mayer, Adalbert and Steven L. Puller, "The old boy (and girl) network: Social network formation on university campuses," *Journal of Public Economics*, February 2008, *92* (1-2), 329–347.

McFadden, Daniel, "Modelling the Choice of Residential Location," in A. Karlqvist, L. Lundqvist, F. Snickars, and J. Weibull, eds., *Spatial Interaction Theory and Planning Models*, Amsterdam: North Holland, 1978, pp. 75–96.

Monte, Ferdinando, Stephen J. Redding, and Esteban Rossi-Hansberg, "Commuting, Migration and Local Employment Elasticities," November 2015. NBER Working Paper 21706.

Morales, Eduardo, Gloria Sheu, and Andrés Zahler, "Extended Gravity," June 2017. mimeo.

Myrdal, Gunnar, *An American Dilemma: The Negro Problem and Modern Democracy*, Harper and Brothers, 1944.

Oberholzer-Gee, Felix and Joel Waldfogel, "Media Markets and Localism: Does Local News en Español Boost Hispanic Voter Turnout?," *American Economic Review*, December 2009, *99* (5), 2120–28.

O'Flaherty, Brendan, *The Economics of Race in the United States*, Harvard University Press, June 2015.

_ and Rajiv Sethi, "Urban Crime," in J. Vernon Henderson Gilles Duranton and William C. Strange, eds., *Handbook of Regional and Urban Economics*, Vol. 5, Elsevier, 2015, chapter 23, pp. 1519 – 1621.

Pakes, Ariel, "Alternative Models for Moment Inequalities," *Econometrica*, November 2010, *78* (6), 1783–1822.

Paluck, Elizabeth Levy and Donald P. Green, "Prejudice Reduction: What Works? A Review and Assessment of Research and Practice," *Annual Review of Psychology*, 2009, *60* (1), 339–367. PMID: 18851685.

Pettigrew, Thomas F. and Linda R. Tropp, "A meta-analytic test of intergroup contact theory," *Journal of Personality and Social Psychology*, 2006, *90* (5), 751–783.

Quillian, Lincoln and Devah Pager, "Black Neighbors, Higher Crime? The Role of Racial Stereotypes in Evaluations of Neighborhood Crime," *American Journal of Sociology*, 2001, *107* (3), pp. 717–767.

Schiff, Nathan, "Cities and Product Variety: Evidence from Restaurants," *Journal of Economic Geography*, 2015, *15* (6), 1085–1123.

Schreer, George E., Saundra Smith, and Kirsten Thomas, ""Shopping While Black": Examining Racial Discrimination in a Retail Setting," *Journal of Applied Social Psychology*, 2009, *39* (6), 1432–1444.

Smith, Sandra Susan, "Race and Trust," *Annual Review of Sociology*, 2010, *36* (1), 453–475.

Taddy, Matt, "Distributed Multinomial Regression," *Ann. Appl. Stat.*, 09 2015, *9* (3), 1394–1414.

Train, Kenneth E., *Discrete Choice Methods with Simulation*, New York: Cambridge University Press, 2009.

_ , Daniel L. McFadden, and Moshe Ben-Akiva, "The Demand for Local Telephant Service: a Fully Discrete Model of Residential Calling Patterns and Service Choices," *The Rand Journal of Economics*, 1987, *18* (1), 109–123.

Waldfogel, Joel, *The Tyranny of the Market: Why You Can't Always Get What You Want*, Harvard University Press, October 2007.

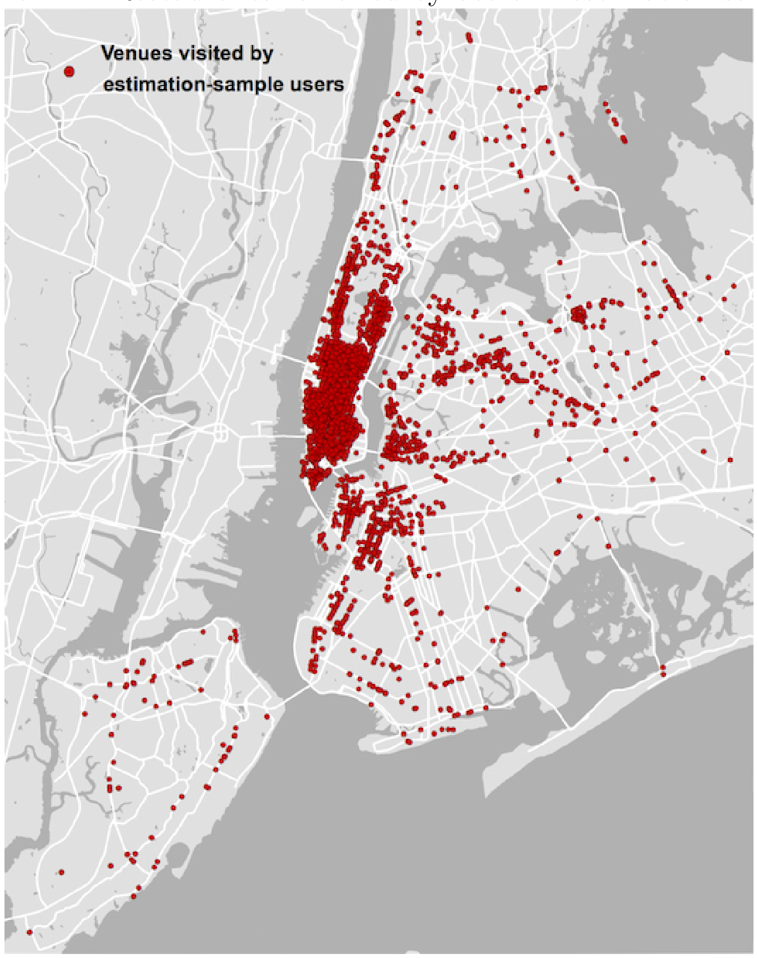
Wollman, Thomas, "Trucks Without Bailouts: Equilibrium Product Characteristics for Commercial Vehicles," May 2016. mimeo.

Yelp, "Yelp.com Welcomes 100 Million Unique Visitors in January 2013," February 2013. http://officialblog.yelp.com/2013/02/yelpcom-welcomes-100-million-unique-visitors-in-january-2013.html.

Online Appendix (not for publication)

A Appendix Figures and Tables

Figure A.1: Restaurants reviewed by users in estimation sample



Notes: This map depicts the locations of 5363 Yelp restaurant venues reviewed by users in our estimation sample. Each dot represents a venue.

Figure A.2: Venue counts by ZIP code, Yelp vs NYC DOHMH

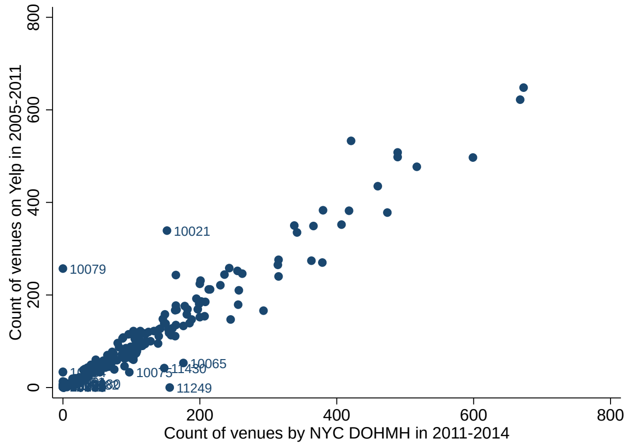


Table A.1: Venue review summary statistics

Restaurant characteristic	Share of all Yelp reviews	Share of estimation-sample reviews: all users	Asian	black	white or Hispanic
Price of $	.227	.242	.233	.334	.241
Price of $$	.567	.563	.536	.575	.580
Price of $$$	.161	.158	.181	.084	.148
Price of $$$$	.045	.037	.050	.006	.032
Rating of 1 stars	.001	.001	.000	.004	.000
Rating of 1.5 stars	.002	.002	.002	.007	.002
Rating of 2 stars	.011	.008	.006	.023	.008
Rating of 2.5 stars	.038	.037	.032	.068	.035
Rating of 3 stars	.141	.144	.125	.204	.150
Rating of 3.5 stars	.358	.369	.358	.344	.379
Rating of 4 stars	.390	.386	.415	.307	.380
Rating of 4.5 stars	.056	.050	.061	.042	.044
Rating of 5 stars	.003	.001	.001	.001	.002
Cuisine: American	.331	.336	.282	.406	.366
Cuisine: Asian	.248	.256	.340	.177	.200
Cuisine: European	.179	.166	.165	.092	.179
Cuisine: Latin American	.090	.098	.072	.187	.106
Cuisine: No Category	.079	.080	.075	.095	.082
Cuisine: Indian	.026	.024	.030	.013	.023
Cuisine: Middle Eastern	.026	.022	.023	.020	.025
Cuisine: Veggie	.016	.014	.011	.006	.016
Cuisine: African	.003	.003	.002	.004	.003
Located in Manhattan	.749	.802	.857	.581	.778
Located in Brooklyn	.170	.123	.063	.360	.147
Located in Queens	.068	.066	.076	.045	.061
Located in Bronx	.008	.005	.003	.006	.005
Located in Staten Island	.005	.004	.001	.008	.008
Located in plurality Asian	.084	.088	.121	.056	.069
Located in plurality black	.018	.015	.006	.089	.011
Located in plurality Hispanic	.047	.043	.031	.074	.052
Located in plurality white	.851	.854	.842	.781	.869

NOTES: This table summarizes the distribution of reviews across different venue characteristics for all Yelp users as a whole (column 1), our estimation sample (column 2), and by race within our estimation sample (columns 3–5).

Table A.2: NYC census tract summary statistics

Variable	Mean	Std. Dev.
Tract characteristics		
Population	3866	2115
Spectral segregation index for tract's plurality	0.914	2.394
Robberies per resident, 2007-2011 annual average	0.003	0.009
Tract-pair characteristics		
Percent absolute difference in median household income	0.506	0.355
Percent difference in median household income	0	0.618
Euclidean demographic distance between tracts	0.455	0.226
Travel time by public transport in minutes	72.436	30.319
Travel time by automobile in minutes	24.937	10.589

NOTES: The upper panel describes 2,110 NYC census tracts for which an estimate of median household income is available. The lower panel describes 4,452,012 pairs of 2010 NYC census tracts for which estimates of median household income and travel times are available. Data on incomes from 2007-2011 American Community Survey, demographics from 2010 Census of Population, robberies from NYPD, and travel times from Google Maps.

Table A.3: Varying choice set sizes

	(1)	(2)	(3)	(4)	(5)	(6)
	Asian		black		white/Hisp	
Log travel time from home by public transit	-1.04^a	-1.03^a	-.953^a	-.920^a	-1.10^a	-1.11^a
	(.098)	(.096)	(.123)	(.119)	(.055)	(.053)
Log travel time from home by car	-1.15^a	-1.15^a	-1.20^a	-1.15^a	-1.33^a	-1.34^a
	(.084)	(.082)	(.151)	(.142)	(.056)	(.054)
Log travel time from work by public transit	-1.25^a	-1.26^a	-2.02	-1.97	-1.88^a	-1.84^a
	(.148)	(.149)	(1.66)	(1.47)	(.288)	(.260)
Log travel time from work by car	-1.61^a	-1.61^a	-1.91^a	-1.93^a	-1.97^a	-1.96^a
	(.174)	(.172)	(.545)	(.577)	(.178)	(.169)
Log travel time from commute by public transit	-.938^a	-.931^a	-.937^a	-.917^a	-1.09^a	-1.08^a
	(.064)	(.062)	(.099)	(.099)	(.043)	(.041)
Log travel time from commute by car	-1.05^a	-1.04^a	-1.31^a	-1.28^a	-1.39^a	-1.38^a
	(.059)	(.058)	(.157)	(.156)	(.053)	(.050)
Euclidean demographic distance between h_i and k_j	-.957^a	-.924^a	-1.89^a	-1.74^a	-1.16^a	-1.14^a
	(.115)	(.113)	(.271)	(.265)	(.125)	(.121)
Spectral segregation index of k_j	.138^a	.138^a	.177^b	.125	.045^c	.036
	(.048)	(.048)	(.086)	(.083)	(.026)	(.026)
EDD × SSI	-.137	-.138	-.369	-.346	-.095	-.096
	(.108)	(.108)	(.256)	(.271)	(.086)	(.085)
Share of tract population that is Asian	.973^a	.993^a	-.015	.028	.364^a	.347^a
	(.114)	(.112)	(.333)	(.330)	(.131)	(.129)
Share of tract population that is black	.184	.237	.876^b	1.02^a	.057	.076
	(.310)	(.306)	(.391)	(.383)	(.257)	(.250)
Share of tract population that is Hispanic	-.141	-.112	.502	.536	.312^c	.320^c
	(.228)	(.225)	(.359)	(.355)	(.180)	(.177)
Share of tract population that is other	.636	.729	5.61^c	5.53^c	1.43	1.36
	(1.98)	(1.95)	(2.99)	(2.95)	(1.84)	(1.81)
Dummy for 2-dollar bin	.354^a	.338^a	.757^a	.727^a	.372^a	.387^a
	(.084)	(.082)	(.189)	(.186)	(.080)	(.078)
Dummy for 3-dollar bin	.235^b	.229^b	.140	.102	-.068	-.035
	(.112)	(.110)	(.320)	(.316)	(.115)	(.113)
Dummy for 4-dollar bin	.263	.254	.092	-.108	-.249	-.268
	(.180)	(.175)	(1.21)	(1.15)	(.211)	(.207)
Yelp rating of restaurant	.561^a	.536^a	.094	.028	.343^a	.344^a
	(.060)	(.059)	(.126)	(.123)	(.056)	(.055)
African cuisine category	.129	.358	-.009	.037	.361	.170
	(.289)	(.286)	(.534)	(.522)	(.246)	(.242)
American cuisine category	.411^a	.428^a	.596^a	.551^a	.596^a	.604^a
	(.053)	(.052)	(.114)	(.112)	(.048)	(.048)
Asian cuisine category	.871^a	.875^a	.279^b	.253^b	.301^a	.297^a
	(.052)	(.052)	(.129)	(.127)	(.052)	(.051)
European cuisine category	.195^a	.193^a	-.239	-.283^c	.225^a	.208^a
	(.057)	(.057)	(.148)	(.147)	(.053)	(.053)
Indian cuisine category	.338^a	.328^a	-.562^c	-.558^c	-.040	-.065
	(.088)	(.087)	(.292)	(.290)	(.094)	(.092)
Latin American cuisine category	.495^a	.505^a	1.01^a	.922^a	.643^a	.644^a
	(.068)	(.067)	(.129)	(.126)	(.059)	(.058)
Middle Eastern cuisine category	.304^a	.316^a	.089	-.026	.279^a	.265^a
	(.097)	(.096)	(.243)	(.240)	(.090)	(.089)
Vegetarian/vegan cuisine category	.329^b	.389^a	-.204	-.339	.603^a	.621^a
	(.131)	(.129)	(.400)	(.401)	(.109)	(.107)
2-dollar bin × home tract median income	.034^a	.036^a	-.028	-.023	.040^a	.039^a
	(.010)	(.010)	(.031)	(.030)	(.009)	(.009)
3-dollar bin × home tract median income	.074^a	.076^a	.031	.035	.083^a	.079^a
	(.013)	(.013)	(.050)	(.049)	(.012)	(.012)
4-dollar bin × home tract median income	.064^a	.063^a	-.203	-.163	.082^a	.081^a
	(.021)	(.020)	(.232)	(.217)	(.022)	(.021)
Yelp rating × home tract median income	.010	.012^c	.007	.014	.014^b	.014^b
	(.007)	(.007)	(.021)	(.020)	(.006)	(.006)
Percent absolute difference in median incomes $(h_i - k_j)$	-.061	-.063	.837^a	.789^a	-.113^b	-.118^b
	(.048)	(.047)	(.122)	(.119)	(.050)	(.049)
Percent difference in median incomes $(k_j - h_i)$	.099	.060	.015	.397	.787^a	.811^a
	(.292)	(.289)	(.821)	(.821)	(.286)	(.282)
Log median household income in k_j	-.095	-.057	.120	-.170	-.696^a	-.708^a
	(.256)	(.254)	(.719)	(.719)	(.250)	(.246)
Average annual robberies per resident in k_j	-3.43^a	-3.31^a	2.22^b	2.18^b	-3.65^a	-3.82^a
	(.657)	(.650)	(1.10)	(1.08)	(.753)	(.751)
Number of origin-mode points	6	6	6	6	6	6
Number of venues in choice set	50	100	50	100	50	100
Number of trips	6447	6447	1079	1079	6936	6936

NOTES: Each column reports an estimated conditional-logit model of individuals' decisions to visit a Yelp venue, identical to the specifications in Table 3. Standard errors in parentheses. Statistical significance denoted by a (1%), b (5%), c (10%).

Table A.4: Estimates with home as only origin

	(1) Asian	(2) black	(3) white/Hisp
Log travel time from home by public transit	-.966[a]	-1.28[a]	-1.29[a]
	(.043)	(.065)	(.024)
Log travel time from home by car	-1.12[a]	-1.88[a]	-1.80[a]
	(.045)	(.167)	(.049)
Euclidean demographic distance between h_i and k_j	-.788[a]	-1.18[a]	-.667[a]
	(.122)	(.294)	(.133)
Spectral segregation index of k_j	.146[a]	.174[b]	.056[b]
	(.051)	(.073)	(.027)
EDD $\times$ SSI	-.136	-.323	-.131
	(.113)	(.210)	(.083)
Share of tract population that is Asian	.972[a]	-.061	.152
	(.119)	(.341)	(.137)
Share of tract population that is black	.419	.762[c]	.062
	(.312)	(.405)	(.265)
Share of tract population that is Hispanic	-.395[c]	.393	.165
	(.233)	(.386)	(.188)
Share of tract population that is other	.272	3.19	1.22
	(2.04)	(3.48)	(1.98)
Number of origin-mode points	2	2	2
Number of trips	6447	1079	6936

NOTES: Each column reports an estimated conditional-logit model of individuals' decisions to visit a Yelp venue. Standard errors in parentheses. Statistical significance denoted by a (1%), b (5%), c (10%). All estimates in this table restrict the origin-mode pairs to $l \in \{hc, hp\}$. Unreported controls include venue price and rating interacted with home tract income, log median household income in tract of venue, percentage difference and percentage absolute difference in income levels between home and venue tract, 28 area dummies, 9 cuisine dummies, and robberies per resident.

Table A.5: Estimates with more disaggregated cuisine categories

	(1) Asian	(2) black	(3) white/Hisp
Log travel time from home by public transit	-1.06^a	-.962^a	-1.13^a
	(.108)	(.127)	(.059)
Log travel time from home by car	-1.18^a	-1.22^a	-1.36^a
	(.093)	(.159)	(.060)
Log travel time from work by public transit	-1.24^a	-1.89	-1.86^a
	(.146)	(1.15)	(.281)
Log travel time from work by car	-1.61^a	-1.83^a	-1.94^a
	(.176)	(.464)	(.169)
Log travel time from commute by public transit	-.943^a	-.954^a	-1.10^a
	(.067)	(.104)	(.044)
Log travel time from commute by car	-1.06^a	-1.37^a	-1.44^a
	(.063)	(.182)	(.059)
Euclidean demographic distance between h_i and k_j	-1.03^a	-1.91^a	-1.19^a
	(.122)	(.283)	(.131)
Spectral segregation index of k_j	.153^a	.065	.051^c
	(.051)	(.091)	(.027)
EDD × SSI	-.185	-.165	-.076
	(.123)	(.230)	(.084)
Share of tract population that is Asian	.890^a	.046	.330^b
	(.125)	(.356)	(.142)
Share of tract population that is black	.091	.707^c	.112
	(.327)	(.415)	(.270)
Share of tract population that is Hispanic	-.320	.310	.377^b
	(.240)	(.393)	(.190)
Share of tract population that is other	.425	3.73	.876
	(2.07)	(3.41)	(1.99)
Dummy for 2-dollar bin	.244^a	.655^a	.232^a
	(.089)	(.203)	(.084)
Dummy for 3-dollar bin	.124	-.151	-.189
	(.120)	(.346)	(.123)
Dummy for 4-dollar bin	.143	-.113	-.451^b
	(.193)	(1.28)	(.224)
Yelp rating of restaurant	.575^a	.042	.356^a
	(.065)	(.140)	(.061)
2-dollar bin × home tract median income	.036^a	-.027	.041^a
	(.011)	(.032)	(.010)
3-dollar bin × home tract median income	.073^a	.066	.080^a
	(.014)	(.053)	(.013)
4-dollar bin × home tract median income	.074^a	-.162	.092^a
	(.022)	(.243)	(.023)
Yelp rating × home tract median income	.009	.011	.017^b
	(.008)	(.023)	(.007)
Percent absolute difference in median incomes ($h_i - k_j$)	-.055	.929^a	-.103^c
	(.051)	(.129)	(.053)
Percent difference in median incomes ($k_j - h_i$)	.234	.746	.754^b
	(.307)	(.866)	(.302)
Log median household income in k_j	-.231	-.514	-.660^b
	(.270)	(.755)	(.264)
Average annual robberies per resident in k_j	-2.78^a	2.91^b	-3.24^a
	(.677)	(1.22)	(.771)
Number of origin-mode points	6	6	6
Number of trips	6447	1079	6936

NOTES: Each column reports an estimated conditional-logit model of individuals' decisions to visit a Yelp venue. Standard errors in parentheses. Statistical significance denoted by a (1%), b (5%), c (10%). The unreported controls are 28 area dummies and 39 cuisine category dummies.

Table A.6: Estimates for the late-adopters subsample

	(1) Asian	(2) black	(3) white/Hisp
Log travel time from home by public transit	-1.18^a	-.919^a	-1.10^a
	(.161)	(.188)	(.081)
Log travel time from home by car	-1.25^a	-1.14^a	-1.37^a
	(.123)	(.208)	(.090)
Log travel time from work by public transit	-1.34^a	-16.6	-2.07^a
	(.206)	(40721.8)	(.493)
Log travel time from work by car	-1.71^a	-1.99^b	-2.13^a
	(.238)	(.929)	(.289)
Log travel time from commute by public transit	-1.07^a	-.821^a	-1.15^a
	(.103)	(.117)	(.068)
Log travel time from commute by car	-1.13^a	-1.64^a	-1.57^a
	(.082)	(.428)	(.098)
Euclidean demographic distance between h_i and k_j	-.972^a	-1.92^a	-1.50^a
	(.196)	(.431)	(.222)
Spectral segregation index of k_j	.442^a	.277	-.019
	(.128)	(.375)	(.109)
EDD × SSI	-.407	-.284	.116
	(.260)	(.718)	(.209)
Share of tract population that is Asian	.981^a	.312	.761^a
	(.185)	(.440)	(.232)
Share of tract population that is black	-.163	1.10^b	.015
	(.555)	(.513)	(.440)
Share of tract population that is Hispanic	.159	1.17^b	.828^a
	(.357)	(.468)	(.311)
Share of tract population that is other	-.022	3.16	-1.69
	(2.99)	(4.01)	(3.32)
Dummy for 2-dollar bin	.394^a	1.42^a	.680^a
	(.122)	(.292)	(.143)
Dummy for 3-dollar bin	.312^c	.202	.093
	(.166)	(.523)	(.207)
Dummy for 4-dollar bin	.186	-.623	.210
	(.274)	(1.70)	(.359)
Yelp rating of restaurant	.584^a	.429^b	.351^a
	(.092)	(.203)	(.103)
African cuisine category	.087	.012	.468
	(.530)	(.581)	(.412)
American cuisine category	.477^a	.562^a	.647^a
	(.087)	(.149)	(.083)
Asian cuisine category	.982^a	.293^c	.302^a
	(.087)	(.168)	(.090)
European cuisine category	.336^a	-.510^b	.365^a
	(.093)	(.205)	(.090)
Indian cuisine category	.447^a	-.467	-.126
	(.145)	(.380)	(.167)
Latin American cuisine category	.607^a	1.09^a	.899^a
	(.109)	(.168)	(.099)
Middle Eastern cuisine category	.349^b	.116	.136
	(.156)	(.308)	(.159)
Vegetarian/vegan cuisine category	.324	.123	.360^c
	(.225)	(.538)	(.216)
2-dollar bin × home tract median income	.021	-.150^a	-.002
	(.015)	(.051)	(.016)
3-dollar bin × home tract median income	.065^a	.031	.065^a
	(.019)	(.089)	(.022)
4-dollar bin × home tract median income	.067^b	-.065	.039
	(.031)	(.321)	(.038)
Yelp rating × home tract median income	.015	-.053	.016
	(.011)	(.036)	(.012)
Percent absolute difference in median incomes ($h_i - k_j$)	-.013	.965^a	-.031
	(.075)	(.162)	(.086)
Percent difference in median incomes ($k_j - h_i$)	.293	1.51	.998^b
	(.435)	(1.23)	(.481)
Log median household income in k_j	-.296	-1.06	-.906^b
	(.379)	(1.08)	(.418)
Average annual robberies per resident in k_j	-3.00^a	2.94^b	-4.46^a
	(.966)	(1.47)	(1.23)
Number of origin-mode points	6	6	6
Number of trips	2766	707	2647

NOTES: Each column reports an estimated conditional-logit model of individuals' decisions to visit a Yelp venue. The estimation sample is restricted to Yelp users with later-than-median dates of first review written. Standard errors in parentheses. Statistical significance denoted by a (1%), b (5%), c (10%). The unreported controls are 28 area dummies.

Table A.7: Estimates using only users' early reviews

	(1)	(2)	(3)	(4)	(5)	(6)
	Asian		black		white/Hisp	
Log travel time from home by public transit	-1.07^a	-1.20^a	-1.08^a	-1.08^a	-1.15^a	-1.22^a
	(.158)	(.291)	(.199)	(.351)	(.080)	(.135)
Log travel time from home by car	-1.25^a	-1.35^a	-1.42^a	-1.59^a	-1.38^a	-1.39^a
	(.152)	(.250)	(.264)	(.609)	(.082)	(.122)
Log travel time from work by public transit	-1.29^a	-1.41^a	-17.53	-1.45	-1.95^a	-15.35
	(.248)	(.413)	(12969.00)	(.994)	(.448)	(4257.70)
Log travel time from work by car	-1.58^a	-1.52^a	-2.01^a	-1.32^a	-1.93^a	-2.16^a
	(.261)	(.294)	(.706)	(.433)	(.215)	(.457)
Log travel time from commute by public transit	-.919^a	-.995^a	-1.01^a	-.950^a	-1.11^a	-1.13^a
	(.094)	(.146)	(.130)	(.222)	(.058)	(.086)
Log travel time from commute by car	-1.05^a	-1.11^a	-1.42^a	-1.44^a	-1.48^a	-1.50^a
	(.089)	(.130)	(.208)	(.392)	(.081)	(.122)
Euclidean demographic distance between h_i and k_j	-1.13^a	-.996^a	-1.88^a	-2.05^a	-1.31^a	-1.31^a
	(.172)	(.279)	(.411)	(.765)	(.185)	(.306)
Spectral segregation index of k_j	.118^c	.195^c	.125	.392	-.015	-.037
	(.063)	(.102)	(.158)	(.574)	(.052)	(.086)
EDD × SSI	-.174	-.203	-.154	-.839	.065	.023
	(.160)	(.242)	(.373)	(1.19)	(.125)	(.225)
Share of tract population that is Asian	1.07^a	1.09^a	.354	.839	.378^c	.496
	(.168)	(.273)	(.489)	(.833)	(.195)	(.319)
Share of tract population that is black	.553	.728	1.59^a	.966	.248	.892
	(.441)	(.730)	(.596)	(.984)	(.383)	(.627)
Share of tract population that is Hispanic	-1.22^a	-1.05^c	1.10^b	1.70^c	.618^b	.494
	(.365)	(.608)	(.557)	(.915)	(.267)	(.433)
Share of tract population that is other	-.009	5.33	5.49	11.27	-2.07	-.819
	(2.90)	(4.62)	(5.44)	(9.67)	(2.93)	(4.65)
Dummy for 2-dollar bin	.348^a	.406^b	.742^a	1.09^b	.399^a	.695^a
	(.122)	(.199)	(.279)	(.473)	(.117)	(.193)
Dummy for 3-dollar bin	.179	.072	1.14^c	1.59^c	.212	.258
	(.165)	(.269)	(.581)	(.875)	(.172)	(.280)
Dummy for 4-dollar bin	.035	-.015	.386	-12.98	-.228	-.179
	(.265)	(.399)	(2.59)	(16381.00)	(.312)	(.505)
Yelp rating of restaurant	.488^a	.415^a	.166	.174	.155^c	.294^b
	(.090)	(.146)	(.202)	(.348)	(.084)	(.138)
African cuisine category	.384	.453	-.228	-.674	.422	.602
	(.448)	(.757)	(.668)	(1.15)	(.362)	(.537)
American cuisine category	.403^a	.430^a	.410^b	.245	.595^a	.583^a
	(.077)	(.127)	(.168)	(.273)	(.072)	(.117)
Asian cuisine category	.917^a	.996^a	.190	.215	.371^a	.366^a
	(.077)	(.126)	(.190)	(.302)	(.077)	(.125)
European cuisine category	.189^b	.115	-.512^b	-.410	.207^b	.160
	(.085)	(.140)	(.222)	(.348)	(.080)	(.131)
Indian cuisine category	.370^a	.358^c	-.706	-.822	-.084	-.041
	(.129)	(.213)	(.458)	(.795)	(.141)	(.232)
Latin American cuisine category	.487^a	.575^a	.836^a	.816^a	.754^a	.836^a
	(.101)	(.162)	(.193)	(.301)	(.088)	(.142)
Middle Eastern cuisine category	.352^b	.220	-.310	-.690	.321^b	.453^b
	(.140)	(.239)	(.401)	(.782)	(.131)	(.207)
Vegetarian/vegan cuisine category	.728^a	.833^a	-.399	-.517	.682^a	.848^a
	(.171)	(.267)	(.617)	(1.06)	(.164)	(.256)
2-dollar bin × home tract median income (10k USD)	.029^c	.027	.008	-.040	.036^a	.008
	(.015)	(.025)	(.045)	(.077)	(.013)	(.022)
3-dollar bin × home tract median income (10k USD)	.078^a	.103^a	-.175	-.185	.051^a	.047
	(.019)	(.031)	(.107)	(.159)	(.018)	(.029)
4-dollar bin × home tract median income (10k USD)	.087^a	.124^a	-.420	-1.20	.089^a	.093^c
	(.031)	(.046)	(.550)	(4760.20)	(.032)	(.050)
Yelp rating of restaurant × home tract median income	.017	.028	-.003	.016	.030^a	.024
	(.011)	(.018)	(.033)	(.059)	(.009)	(.015)
Percent absolute difference in median incomes ($h_i - k_j$)	-.061	-.081	.801^a	.613^b	-.074	-.047
	(.071)	(.116)	(.182)	(.307)	(.075)	(.122)
Percent difference in median incomes ($k_j - h_i$)	.113	-.093	1.37	3.76^c	.582	.393
	(.421)	(.686)	(1.22)	(2.06)	(.417)	(.687)
Log median household income in k_j	-.156	-.009	-1.02	-3.27^c	-.516	-.295
	(.367)	(.603)	(1.07)	(1.78)	(.363)	(.597)
Average annual robberies per resident in k_j	-4.44^a	-7.81^a	1.84	2.27	-2.75^a	-1.27
	(.977)	(1.97)	(1.76)	(3.14)	(1.04)	(1.57)
Number of origin-mode points	6	6	6	6	6	6
Number of trips	3205	1258	533	205	3431	1314

NOTES: Each column reports an estimated conditional-logit model of individuals' decisions to visit a Yelp venue, identical to the specifications in Table 3. Standard errors in parentheses. Statistical significance denoted by a (1%), b (5%), c (10%). The unreported controls are 28 area dummies. Observations are limited to the first half (odd-numbered columns) or first fifth (even) of NYC restaurant reviews posted by each user.

Table A.8: Origin-mode-specific-intercepts estimates

	(1) Asian	(2) black	(3) white/Hisp
Log travel time from home by public transit	-1.05[a]	-1.47[a]	-1.38[a]
	(.206)	(.118)	(.052)
Intercept for home by public transit	.665	6.34[c]	3.75[a]
	(2.14)	(3.31)	(.389)
Log travel time from home by car	-1.31[a]	-1.97[a]	-1.89[a]
	(.044)	(.185)	(.048)
Intercept for home by car	3.49[a]	5.96[c]	4.46[a]
	(.505)	(3.36)	(.351)
Log travel time from work by public transit	-1.17[a]	-1.55	-1.52
	(.120)	(.970)	(2065.5)
Intercept for work by public transit	1.95[a]	2.21	-17.0
	(.552)	(4.85)	(7443.5)
Log travel time from work by car	-1.48[a]	-1.63[a]	-1.69[a]
	(.090)	(.183)	(.058)
Intercept for work by car	2.62[a]	4.46	2.95[a]
	(.551)	(3.37)	(.349)
Log travel time from commute by public transit	-.973[a]	-1.08	-1.08[a]
	(.134)	(.988)	(.135)
Intercept for commute by public transit	.674	-.246	.100
	(.589)	(6.45)	(.567)
Log travel time from commute by car	-1.17[a]	-1.42	-1.62[a]
	(.195)	(1.11)	(.241)
Euclidean demographic distance between h_i and k_j	-.807[a]	-1.33[a]	-.704[a]
	(.124)	(.299)	(.135)
Spectral segregation index of k_j	.151[a]	.082	.062[b]
	(.051)	(.094)	(.027)
EDD × SSI	-.153	-.169	-.109
	(.116)	(.240)	(.084)
Share of tract population that is Asian	.989[a]	.016	.204
	(.120)	(.346)	(.139)
Share of tract population that is black	.176	.903[b]	.067
	(.318)	(.411)	(.267)
Share of tract population that is Hispanic	-.309	.503	.257
	(.235)	(.389)	(.190)
Share of tract population that is other	.082	2.62	1.64
	(2.08)	(3.57)	(2.00)
Number of origin-mode points	6	6	6
Number of trips	6447	1079	6936

NOTES: Each column reports an estimated conditional-logit model of individuals' decisions to visit a Yelp venue. This specification adds five origin-mode-specific intercepts to the specification in Table 3. Standard errors in parentheses. Statistical significance denoted by a (1%), b (5%), c (10%). Only spatial and social friction coefficients are reported for brevity.

Table A.9: Estimates controlling for number of reviews and chain establishments

	(1) Asian	(2) black	(3) white/Hisp
Log travel time from home by public transit	-1.09^a	-.977^a	-1.16^a
	(.115)	(.139)	(.062)
Log travel time from home by car	-1.18^a	-1.20^a	-1.37^a
	(.094)	(.159)	(.061)
Log travel time from work by public transit	-1.29^a	-1.74^b	-1.95^a
	(.165)	(.774)	(.317)
Log travel time from work by car	-1.67^a	-1.69^a	-1.98^a
	(.195)	(.372)	(.174)
Log travel time from commute by public transit	-1.02^a	-.964^a	-1.14^a
	(.082)	(.112)	(.047)
Log travel time from commute by car	-1.07^a	-1.34^a	-1.46^a
	(.065)	(.179)	(.060)
Euclidean demographic distance between h_i and k_j	-1.08^a	-1.79^a	-1.13^a
	(.128)	(.289)	(.136)
Spectral segregation index of k_j	.140^a	.072	.052^c
	(.052)	(.097)	(.028)
EDD × SSI	-.121	-.163	-.072
	(.118)	(.252)	(.085)
Share of tract population that is Asian	1.06^a	-.412	.162
	(.129)	(.374)	(.146)
Share of tract population that is black	.356	1.07^a	.203
	(.334)	(.411)	(.274)
Share of tract population that is Hispanic	-.049	.516	.506^a
	(.247)	(.392)	(.194)
Share of tract population that is other	-2.64	3.94	-.554
	(2.27)	(3.44)	(2.09)
Dummy for 2-dollar bin	.145	.645^a	.158^c
	(.093)	(.203)	(.086)
Dummy for 3-dollar bin	.020	-.335	-.291^b
	(.123)	(.356)	(.125)
Dummy for 4-dollar bin	-.257	-1.21	-.782^a
	(.206)	(1.20)	(.235)
Yelp rating of restaurant	.304^a	-.143	.141^b
	(.069)	(.143)	(.063)
African cuisine category	.412	-.146	.333
	(.300)	(.560)	(.263)
American cuisine category	.245^a	.362^a	.414^a
	(.058)	(.123)	(.052)
Asian cuisine category	.755^a	.179	.193^a
	(.058)	(.138)	(.056)
European cuisine category	.196^a	-.396^b	.230^a
	(.063)	(.160)	(.058)
Indian cuisine category	.467^a	-.358	.007
	(.094)	(.303)	(.099)
Latin American cuisine category	.406^a	.893^a	.596^a
	(.074)	(.140)	(.063)
Middle Eastern cuisine category	.233^b	.165	.130
	(.107)	(.255)	(.098)
Vegetarian/vegan cuisine category	.247^c	-.081	.412^a
	(.144)	(.416)	(.119)
Number of Yelp reviews	.005^a	.005^a	.004^a
	(.000)	(.000)	(.000)
Establishment belongs to chain	-.437^a	-.226	-.511^a
	(.110)	(.192)	(.100)
Number of origin-mode points	6	6	6
Number of trips	6447	1079	6936

NOTES: Each column reports an estimated conditional-logit model of individuals' decisions to visit a Yelp venue. This specification adds the number of Yelp reviews and a chain indicator as covariates to the specifications in Table 3. Standard errors in parentheses. Statistical significance denoted by a (1%), b (5%), c (10%). The unreported controls are venue price and rating interacted with home tract income, log median household income in tract of venue, percentage difference and percentage absolute difference in income levels between home and venue tract, robberies per resident, and 28 area dummies.

Table A.10: Estimates for minimum-travel-time specification

	(1) Asian	(2) black	(3) white/Hisp
Log minimum travel time	-.932^a	-.957^a	-1.19^a
	(.022)	(.044)	(.019)
Euclidean demographic distance between h_i and k_j	-1.13^a	-2.00^a	-1.39^a
	(.120)	(.277)	(.130)
Spectral segregation index of k_j	.153^a	.113	.045^c
	(.051)	(.084)	(.027)
EDD × SSI	-.123	-.214	-.045
	(.115)	(.224)	(.081)
Share of tract population that is Asian	1.08^a	.017	.485^a
	(.119)	(.344)	(.136)
Share of tract population that is black	.328	1.13^a	.200
	(.317)	(.396)	(.264)
Share of tract population that is Hispanic	-.215	.430	.477^b
	(.233)	(.379)	(.187)
Share of tract population that is other	.619	4.60	.315
	(2.04)	(3.36)	(1.97)
Dummy for 2-dollar bin	.382^a	.778^a	.362^a
	(.086)	(.195)	(.082)
Dummy for 3-dollar bin	.313^a	-.061	.001
	(.116)	(.340)	(.120)
Dummy for 4-dollar bin	.222	-.007	-.289
	(.188)	(1.22)	(.221)
Yelp rating of restaurant	.570^a	.062	.338^a
	(.064)	(.136)	(.059)
African cuisine category	.281	-.199	.289
	(.298)	(.552)	(.259)
American cuisine category	.435^a	.532^a	.583^a
	(.054)	(.119)	(.050)
Asian cuisine category	.884^a	.267^b	.301^a
	(.054)	(.134)	(.054)
European cuisine category	.199^a	-.312^b	.233^a
	(.059)	(.154)	(.055)
Indian cuisine category	.390^a	-.416	-.032
	(.091)	(.300)	(.097)
Latin American cuisine category	.520^a	1.02^a	.685^a
	(.070)	(.136)	(.061)
Middle Eastern cuisine category	.294^a	.098	.187^b
	(.100)	(.250)	(.094)
Vegetarian/vegan cuisine category	.396^a	.001	.583^a
	(.137)	(.409)	(.115)
2-dollar bin × home tract median income	.033^a	-.025	.040^a
	(.011)	(.031)	(.009)
3-dollar bin × home tract median income	.071^a	.068	.075^a
	(.014)	(.053)	(.013)
4-dollar bin × home tract median income	.073^a	-.180	.086^a
	(.022)	(.232)	(.023)
Yelp rating × home tract median income	.011	.006	.017^b
	(.008)	(.023)	(.007)
Percent absolute difference in median incomes ($h_i - k_j$)	-.117^b	.843^a	-.116^b
	(.050)	(.125)	(.052)
Percent difference in median incomes ($k_j - h_i$)	.173	.419	.882^a
	(.305)	(.845)	(.298)
Log median household income in k_j	-.155	-.175	-.751^a
	(.267)	(.737)	(.260)
Average annual robberies per resident in k_j	-3.05^a	2.49^b	-3.87^a
	(.672)	(1.18)	(.768)
Number of origin-mode points	1	1	1
Number of trips	6447	1079	6936

NOTES: Each column reports an estimated conditional-logit model of individuals' decisions to visit a Yelp venue under the assumption that $\nu_{ijlt} = \nu_{ijt}$ and $\gamma_l^1 = \gamma^1$. Standard errors in parentheses. Statistical significance denoted by a (1%), b (5%), c (10%). The unreported controls are 28 area dummies.

Table A.11: Estimates employing only tract-level demographics

	(1)
Log travel time from home by public transit	-1.14[a] (.045)
Log travel time from home by car	-1.26[a] (.037)
Log travel time from work by public transit	-1.64[a] (.126)
Log travel time from work by car	-1.82[a] (.098)
Log travel time from commute by public transit	-1.06[a] (.030)
Log travel time from commute by car	-1.28[a] (.033)
Euclidean demographic distance between h_i and k_j	-1.37[a] (.075)
Spectral segregation index of k_j	.072[a] (.022)
EDD × SSI	-.109[c] (.060)
Share of tract population that is Asian	.791[a] (.078)
Share of tract population that is black	.530[a] (.156)
Share of tract population that is Hispanic	.206[c] (.121)
Share of tract population that is other	-.997 (1.20)
Dummy for 2-dollar bin	.405[a] (.051)
Dummy for 3-dollar bin	.104 (.073)
Dummy for 4-dollar bin	.003 (.126)
Yelp rating of restaurant	.399[a] (.037)
African cuisine category	.367[b] (.152)
American cuisine category	.528[a] (.031)
Asian cuisine category	.563[a] (.033)
European cuisine category	.174[a] (.035)
Indian cuisine category	.058 (.059)
Latin American cuisine category	.663[a] (.039)
Middle Eastern cuisine category	.170[a] (.060)
Vegetarian/vegan cuisine category	.525[a] (.075)
2-dollar bin × home tract median income (10k USD)	.035[a] (.006)
3-dollar bin × home tract median income (10k USD)	.080[a] (.008)
4-dollar bin × home tract median income (10k USD)	.071[a] (.014)
Yelp rating of restaurant × home tract median income	.013[a] (.004)
Percent absolute difference in median incomes ($h_i - k_j$)	.027 (.031)
Percent difference in median incomes ($k_j - h_i$)	.260 (.185)
Log median household income in k_j	-.196 (.162)
Average annual robberies per resident in k_j	-3.66[a] (.428)
Number of origin-mode points	6
Number of trips	18015

NOTES: This table reports an estimated conditional-logit model of individuals' decisions to visit a Yelp venue. This specification uses no information on user-level racial demographics. Standard errors in parentheses. Statistical significance denoted by *a* (1%), *b* (5%), *c* (10%). The unreported controls are 28 area dummies.

Table A.12: Tract-level residential and consumption segregation

		Residential dissimilarity	Consumption dissimilarity			
			Estimated	No spatial	No social	Neither friction
		(1)	(2)	(3)	(4)	(5)
Dissimilarity index						
Asian		.521	.275	.251	.201	.186
Hispanic		.486	.134	.102	.094	.074
black		.653	.322	.284	.229	.214
white		.636	.183	.144	.090	.068
Pairwise dissimilarity						
Asian	Hispanic	.584	.251	.240	.176	.176
Asian	black	.796	.455	.399	.325	.282
Asian	white	.519	.232	.212	.173	.169
Hispanic	white	.658	.158	.113	.090	.025
black	Hispanic	.558	.300	.264	.227	.212
black	white	.822	.323	.287	.215	.207

NOTES: See notes to Table 7. In this table, columns 2 through 5 report tract-level dissimilarity indices based on the coefficient estimates in Table 3 .

Table A.13: Late-adopter estimates: Residential and consumption segregation

		Residential dissimilarity	Consumption dissimilarity			
			Estimated	No spatial	No social	Neither friction
		(1)	(2)	(3)	(4)	(5)
Dissimilarity index						
Asian		.521	.344	.324	.269	.256
Hispanic		.486	.165	.139	.122	.104
black		.653	.397	.376	.314	.310
white		.636	.213	.172	.124	.112
Pairwise dissimilarity						
Asian	Hispanic	.584	.316	.306	.241	.238
Asian	black	.796	.544	.512	.432	.408
Asian	white	.519	.309	.282	.239	.225
Hispanic	white	.658	.179	.129	.095	.031
black	Hispanic	.558	.378	.352	.308	.300
black	white	.822	.396	.374	.299	.304

NOTES: See notes to Table 7. In this table, columns 2 through 5 report venue-level dissimilarity indices based on the coefficient estimates in Table A.6.

Table A.14: Origin-mode-specific intercepts: Residential and consumption segregation

		Residential dissimilarity	Consumption dissimilarity					
			Est (bias)	Est	No spatial	No soc (bias)	No soc	Neither friction
		(1)	(2)	(3)	(4)	(5)	(6)	(7)
Dissimilarity index								
Asian		.521	.286	.299	.259	.227	.232	.211
Hispanic		.486	.12	.123	.09	.101	.1	.079
black		.653	.335	.331	.296	.278	.269	.252
white		.636	.161	.17	.124	.113	.116	.09
Pairwise dissimilarity								
Asian	Hispanic	.584	.256	.273	.237	.202	.212	.195
Asian	black	.796	.462	.464	.408	.38	.373	.339
Asian	white	.519	.252	.265	.228	.191	.2	.181
Hispanic	white	.658	.117	.129	.07	.087	.089	.033
black	Hispanic	.558	.31	.302	.271	.263	.253	.239
black	white	.822	.334	.336	.295	.273	.266	.248

NOTES: See notes to Table 7. In this table, columns 2 through 7 report venue-level dissimilarity indices based on the coefficient estimates in Table A.8. Columns 2-3 use the estimated coefficients. Columns 2 and 5 sets the origin-specific intercepts to zero. Column 4 sets the coefficients on travel-time covariates to zero. Columns 5-6 set the coefficients on demographic-difference covariates to zero. Column 7 sets the coefficients on travel-time and demographic-difference covariates to zero.

Table A.15: Home as only origin: Residential and consumption segregation

		Residential dissimilarity	Consumption dissimilarity			
			Estimated	No spatial	No social	Neither friction
		(1)	(2)	(3)	(4)	(5)
Dissimilarity index						
Asian		.521	.322	.290	.245	.232
Hispanic		.486	.156	.114	.110	.088
black		.653	.348	.322	.269	.260
white		.636	.200	.153	.112	.093
Pairwise dissimilarity						
Asian	Hispanic	.584	.306	.273	.228	.217
Asian	black	.796	.488	.448	.374	.357
Asian	white	.519	.288	.255	.218	.203
Hispanic	white	.658	.177	.115	.094	.037
black	Hispanic	.558	.326	.297	.263	.250
black	white	.822	.356	.324	.261	.255

NOTES: See notes to Table 7. In this table, columns 2 through 5 report venue-level dissimilarity indices based on the coefficient estimates in Table 3 but constraining all trips to originate at home.

Table A.16: Demographics of residents and consumers in three Manhattan communities

	Residential share	Consumption share			
		Estimated	No Spatial	No Social	Neither
		Community District 8: Upper East Side			
Asian	0.080	0.095	0.110	0.104	0.119
Hispanic	0.066	0.329	0.320	0.361	0.347
black	0.023	0.045	0.052	0.061	0.068
white	0.810	0.500	0.484	0.442	0.431
		Community District 10: Central Harlem			
Asian	0.024	0.047	0.061	0.113	0.148
Hispanic	0.222	0.152	0.103	0.256	0.181
black	0.630	0.735	0.756	0.473	0.462
white	0.095	0.057	0.071	0.142	0.193
		Community District 11: East Harlem			
Asian	0.056	0.072	0.099	0.123	0.152
Hispanic	0.494	0.528	0.417	0.448	0.339
black	0.309	0.166	0.182	0.111	0.115
white	0.120	0.205	0.270	0.289	0.363

NOTES: This table reports the share of residents and model-predicted restaurant visitors by race in the three community districts illustrated in Figure 7. The five columns correspond to the five scenarios reported in the five columns of Table 7.

Table A.17: Demographics of residents and consumers in lower Manhattan and west Brooklyn

	Residential share	Consumption share			
		Estimated	No Spatial	No Social	Neither
		Community District 1, Brooklyn: Greenpoint			
Asian	0.050	0.064	0.066	0.080	0.082
black	0.052	0.284	0.287	0.306	0.298
Hispanic	0.272	0.324	0.305	0.298	0.283
white	0.608	0.299	0.315	0.287	0.311
		Community District 2, Brooklyn: Williamsburg			
Asian	0.066	0.050	0.054	0.060	0.064
black	0.189	0.338	0.320	0.319	0.289
Hispanic	0.164	0.244	0.272	0.254	0.286
white	0.546	0.340	0.327	0.337	0.332
		Community District 3, Brooklyn: Bedford-Stuyvesant			
Asian	0.021	0.036	0.051	0.076	0.093
black	0.657	0.684	0.590	0.423	0.307
Hispanic	0.147	0.144	0.191	0.229	0.280
white	0.149	0.117	0.151	0.244	0.294
		Community District 3, Manhattan			
Asian	0.321	0.182	0.188	0.155	0.157
black	0.071	0.158	0.163	0.179	0.182
Hispanic	0.253	0.333	0.312	0.318	0.300
white	0.332	0.298	0.307	0.320	0.333

NOTES: This table reports the share of residents and model-predicted restaurant visitors by race in the four community districts illustrated in Figure 8. The five columns correspond to the five scenarios reported in the five columns of Table 7.

B Data

B.1 Yelp venue data

Assigning venues to census tracts. Yelp describes venues' locations by their street addresses. First, we translate these addresses to latitude-longitude coordinates. We determine the latitude and longitude of each venue by a combination of methods. We match the venue addresses to a point using the address locators provided by the New York City Department of Urban Planning and StreetMap North America. For venues with an incorrect ZIP code, we use the borough in the text of the venue's address. For venues not matched using these address locators, we used an alternative address located via the online geocoding service FindLatitudeAndLongitude. For the addresses that cannot be matched using Esri's GIS software or the online service, we find the coordinates using GoogleMaps on a case-by-case basis. Second, after determining venues' coordinates, we assign each venue to a census tract based on a point-in-polygon matching strategy.

Assigning venues to cuisine types. We create nine cuisine dummies by aggregating Yelp cuisine classifications into the following categories: African, American, Asian, European, Indian, Latin American, Middle Eastern, vegetarian. The omitted cuisine category includes all restaurants with "unassigned" cuisine types, which includes venues whose cuisine is listed as "restaurant" on Yelp.

Set of venues included in the sample. The Yelp venues included in our estimation sample as possible elements of users' choice sets meet three criteria. First, they had been reviewed at least once as of 2011. Second, they had both a rating and price listed on Yelp as of 2011. Third, they are located in a census tract for which Census data on its median household income is available.

As one means of validating Yelp's venue coverage, we compare our count of Yelp restaurants by ZIP code to the number of establishments reported in health inspections data by the New York City Department of Health & Mental Hygiene (DOHMH). The DOHMH data report inspection results for 2011-2014, while our Yelp venue data, downloaded in 2011, covers venues reviewed between 2005 and 2011. Despite this temporal mismatch, the two data sources report similar venue counts at the ZIP-code level, as shown in Figure A.2.[62]

B.2 Yelp users data

We started with the roughly 50,000 Yelp users who reviewed a venue in the five boroughs of New York City prior to 14 June 2011. We collected locational information on these users in three rounds. In the first round, we examined all reviews written by a randomly selected 25% subsample of the 50,000 users. In the second round, we selectively examined the remaining 75% of users relying on the first round's lessons for successfully locating users. In the third round, we intensively examined reviews by a set of black users in the remaining 75%. The

[62] Most of the outliers are attributable to the temporal mismatch. The 10021 ZIP code was split into three in 2007, creating 10065 and 10075 (Sam Roberts, "An Elite ZIP Code Becomes Harder to Crack", *New York Times*, 21 March 2007). A similar story explains ZIP codes 11211 and 11249 (Joe Coscarelli, "Williamsburg Hipsters Robbed of Prestigious 11211 Zip Code", *Village Voice*, 2 June 2011). 11430 is JFK Airport. The 10079 ZIP code does not exist; it appears to be a placeholder on Yelp.

final dataset used for estimation contains only those users whose set of home locations is made up of venues all within 1.5 miles of each other and similarly for the set of work locations.

B.2.1 Yelp users data: First round

Between 1 January 2005 and 14 June 2011, users in the 25% sample analyzed in the first round wrote about 230,000 reviews of venues in New York and New Jersey. To identify residential and workplace locations, we examined the text of reviews that contain at least one of 26 key phrases. Those key phrases are ten home-related phrases {I live, my apt, my apartment, my building, my neighborhood, my house, my place, my hood, my block, laundr}, seven work-related phrases {I work, coworker, colleague, lunch break, my office, my work, my job}, and nine phrases related to both {my local, delivery, block away, block from m, blocks from m, close to me, close to my, minutes from m, street from m}.

16,425 of the 230,000 reviews analyzed in the first-round sample contain at least one of these phrases. Reading the text shows that twenty-one percent of these flagged reviews identify a user's home location and eleven percent of them identify a workplace. Reviews containing multiple home-related phrases identify a user's home location in 54% of cases; reviews with multiple work-related phrases yield a work location 45% of the time.

This process identified about 1500 users with a residential location, 575 users with a workplace, and 450 users with both home and work locations. Thus, we have locational information for nearly one-fifth of the Yelp users we examined. The median user for which we obtained locational information had reviewed twenty venues in New York and New Jersey, while the median for which we obtained no information had reviewed five venues. Amongst users with more than ten reviews of NY/NJ venues, we obtained locational information for about 40%.

We identified individuals who changed their residential and workplace locations via two means. First, we recorded any moves identified in the text of reviews containing the 26 key phrases above. Second, we reviewed the text of reviews containing at least one of four key phrases: {we moved, I moved, moving into, moving here}. When this search yielded reviews in which a user reveals that she has recently moved, we eliminated such user from our sample.

This first round yielded 241 users who appear in our estimation sample.

B.2.2 Yelp users data: Second round

In the second round, with the remaining 75% of users, we limited our examination to reviews that were likely to yield both home and work locations for a user. We investigated the text of 6,426 reviews of venues in New York City written by 569 users with at least one review containing two home-related phrases and at least one review containing two work-related phrases. In this round, we did not examine reviews in which the only key phrase was "delivery". We used workers on Amazon's Mechanical Turk marketplace to classify the text. This work was performed in triplicate, and we only use observations with unanimous responses.

This process investigated 569 users and identified home locations for 173 users, work locations for 38 users, and both locations for 304 users. After imposing the previously mentioned 1.5-mile proximity, non-mover, and tract-covariate-availability restrictions, this second round yielded 165 users who appear in our estimation sample.

B.2.3 Yelp users data: Third round

In the third round, we limited our examination to reviews that were written by a set of users identified as black or Hispanic based on their profile photos. We applied the first-round procedure for locating individuals by examining all the reviews written by these users, and we included three users who moved within NYC during the estimation sample by including only their reviews written prior to the move date revealed by the text of their reviews.

This process investigated 275 users and identified home locations for 133 users, work locations for 91 users, and both locations for 51 users. After imposing the previously mentioned 1.5-mile proximity and tract-covariate-availability restrictions, this third round yielded 31 users (20 of whom are black) who appear in our estimation sample.

B.3 NYC geographic and demographic data

Our data on census tracts' geographic areas and populations come from the 2010 Census of Population (Series G001 and P5). By 2010 Census definitions, there are 2168 tracts in New York City, of which 288 are in Manhattan.

The 2007-2011 American Community Survey 5-Year Estimate provides estimates of median household income (Series B19013) for 2110 of these tracts, for which summary statistics are provided in Table 1. Of the nine Manhattan tracts without median household income estimates, seven have a population below 25 persons, one is Inwood Hill Park (population 161), and one is Randall's Island (population 1648). More than 90% of the NYC tracts without median household income estimates have populations below 200 persons, the notable non-Manhattan exceptions being Bush Terminal (population 2105) and Rikers Island (inmate population of 11091).

Tract's historical demographic characteristics come from the Longitudinal Tract Data Base, which maps prior Census years' population counts to the 2010 geographic definitions (Logan et al., 2014).

We aggregate New York City's 59 community boards to define 28 areas, allowing us to use area dummies to control for unobservable characteristics when estimating spatial and social frictions. Each of Manhattan's 12 community districts constitute an area. We aggregate community districts to define 8 areas in Brooklyn (1, {2,6}, {3,8,9}, {4,5}, {7,11,12,13}, 10, {14,15}, {16,17,18}) and 6 in Queens ({1,2}, {3,4}, {5,6}, {7,8,11}, {9,10,14}, {12,13}). The boroughs of the Bronx and Staten Island each constitute one area. We assign each census tract to one of these areas; tracts split across areas are assigned to the area with the largest share of tract land area.

B.4 NYC crime data

We compute tract-level robbery statistics using confidential, geocoded incident-level reports provided by the New York Police Department (NYPD). We aggregate robbery incidents to the census-tract level; we assign each incident to a census tract based on a point-in-polygon matching strategy using ESRI's ArcMap software. We compute the average annual robberies over 2007-2011 for each census tract.

C Econometrics

Here we provide additional detail to the content of Section 3.

C.1 Estimation procedure: details

In this section, we present additional details on the content of Section 3.3.

Deriving equation (6). Given the assumptions in Section 3.2, it holds that

$$\begin{aligned}
&P(d^*_{ijt} = 1|X_i, Z_i, J, J'_{it}; (\gamma, \beta, p^*_{it})) = \\
&P(d^*_{ijt} = 1|d_{ijt} = 1, X_i, Z_i, J, J'_{it}; (\gamma, \beta, p^*_{it}))P(d_{ijt} = 1|X_i, Z_i, J, J'_{it}; (\gamma, \beta, p^*_{it}))+ \\
&P(d^*_{ijt} = 1|d_{ijt} = 0, X_i, Z_i, J, J'_{it}; (\gamma, \beta, p^*_{it}))P(d_{ijt} = 0|X_i, Z_i, J, J'_{it}; (\gamma, \beta, p^*_{it})) = \\
&P(d^*_{ijt} = 1|d_{ijt} = 1, X_i, Z_i, J, J'_{it}; (\gamma, \beta, p^*_{it}))P(d_{ijt} = 1|X_i, Z_i, J, J'_{it}; (\gamma, \beta, p^*_{it})) = \\
&P(d^*_{ijt} = 1|d_{ijt} = 1, J'_{it}; p^*_{it})P(d_{ijt} = 1|X_i, Z_i, J, J'_{it}; (\gamma, \beta, p^*_{it})) = \\
&P(d^*_{ijt} = 1|d_{ijt} = 1, J'_{it}; p^*_{it})P(d_{ijt} = 1|X_i, Z_i, J; (\gamma, \beta)) = \\
&p^*_{it}\mathbb{1}\{j \neq 0, j \in J'_{it}\}P(d_{ijt} = 1|X_i, Z_i, J; (\gamma, \beta)),
\end{aligned}$$

where the final expression is identical to that in equation (6) in the main text. The first equality in this derivation rewrites the probability that we observe a review of j by i at t as the sum of the probability that individual i writes such a review and visited j at t and the probability that individual i writes such a review without visiting j at t. The second equality imposes the assumption that individuals only write reviews about restaurants they actually visit. The third equality imposes the assumption that the probability that individual i writes a review about a restaurant j is independent of the vector of restaurant characteristics X_i and Z_i, and of the set of restaurants J. The fourth equality imposes an implication of equation (5): the probability that individual i visits restaurant j is independent of the previous reviews written by i and of the likelihood that individual i writes a review at period t, p^*_{it}. The last equality imposes the assumptions that the probability that individual i writes a review about a choice j at t is equal to zero when j is the outside option or was previously reviewed by i and is otherwise equal to an individual-time-specific constant, p^*_{it}.

Deriving equation (7). Let's denote the probability that we observe a review by individual i on a restaurant j at period t conditional on i writing a review at period t as $P(d^*_{ijt} = 1|d^*_{it} = 1, X_i, Z_i, J, J'_{it}; (\gamma, \beta, p^*_{it}))$. Using Bayes' Rule,

$$P(d^*_{ijt} = 1|d^*_{it} = 1, X_i, Z_i, J, J'_{it}; (\gamma, \beta, p^*_{it})) = \frac{P(d^*_{ijt} = 1, d^*_{it} = 1|X_i, Z_i, J, J'_{it}; (\gamma, \beta, p^*_{it}))}{P(d^*_{it} = 1|X_i, Z_i, J, J'_{it}; (\gamma, \beta, p^*_{it}))}.$$

The joint probability of observing a review about a restaurant j and observing a review about any restaurant is equal to the probability of observing a review about restaurant j.

Mathematically,

$$
\begin{aligned}
P(d^*_{ijt} = 1, d^*_{it} = 1 | X_i, Z_i, J, J'_{it}; (\gamma, \beta, p^*_{it})) = \\
P(d^*_{it} = 1 | d^*_{ijt} = 1, X_i, Z_i, J, J'_{it}; (\gamma, \beta, p^*_{it})) P(d^*_{ijt} = 1 | X_i, Z_i, J, J'_{it}; (\gamma, \beta, p^*_{it})) = \\
1 \times P(d^*_{ijt} = 1 | X_i, Z_i, J, J'_{it}; (\gamma, \beta, p^*_{it})) = \\
P(d^*_{ijt} = 1 | X_i, Z_i, J, J'_{it}; (\gamma, \beta, p^*_{it})).
\end{aligned}
$$

Analogously, the probability that we observe a review written by i at t is equal to the sum of the probabilities that i writes a review about each of the possible restaurants j in J,

$$
P(d^*_{it} = 1 | X_i, Z_i, J, J'_{it}; (\gamma, \beta, p^*_{it})) = \sum_{j' \in J} P(d^*_{ij't} = 1 | X_i, Z_i, J, J'_{it}; (\gamma, \beta, p^*_{it})).
$$

Therefore, we can write the probability that i writes a review about a restaurant j at period t conditional on observing a review (about any restaurant) written by i at t as

$$
P(d^*_{ijt} = 1 | d^*_{it} = 1, X_i, Z_i, J, J'_{it}; (\gamma, \beta, p^*_{it})) = \frac{P(d^*_{ijt} = 1 | X_i, Z_i, J, J'_{it}; (\gamma, \beta, p^*_{it}))}{\sum_{j' \in J} P(d^*_{ij't} = 1 | X_i, Z_i, J, J'_{it}; (\gamma, \beta, p^*_{it}))}.
$$

Applying the result in equation (6), we can rewrite this probability as

$$
\begin{aligned}
P(d^*_{ijt} = 1 | d^*_{it} = 1, X_i, Z_i, J, J'_{it}; (\gamma, \beta, p^*_{it})) = \\
\frac{p^*_{it} \mathbb{1}\{j \neq 0, j \in J'_{it}\} P(d_{ijt} = 1 | X_i, Z_i, J; (\gamma, \beta))}{\sum_{j' \in J} p^*_{it} \mathbb{1}\{j \neq 0, j' \in J'_{it}\} P(d_{ij't} = 1 | X_i, Z_i, J; (\gamma, \beta))} = \\
\frac{\mathbb{1}\{j \neq 0, j \in J'_{it}\} P(d_{ijt} = 1 | X_i, Z_i, J; (\gamma, \beta))}{\sum_{j' \in J} \mathbb{1}\{j \neq 0, j' \in J'_{it}\} P(d_{ij't} = 1 | X_i, Z_i, J; (\gamma, \beta))} = \\
\frac{\mathbb{1}\{j \neq 0, j \in J'_{it}\} P(d_{ijt} = 1 | X_i, Z_i, J; (\gamma, \beta))}{\sum_{j' \in J'_{it}} P(d_{ij't} = 1 | X_i, Z_i, J; (\gamma, \beta))},
\end{aligned}
$$

and, therefore,

$$
P(d^*_{ijt} = 1 | d^*_{it} = 1, X_i, Z_i, J, J'_{it}; (\gamma, \beta, p^*_{it})) = P(d^*_{ijt} = 1 | d^*_{it} = 1, X_i, Z_i, J, J'_{it}; (\gamma, \beta)).
$$

Applying the result in equation (5), we can additionally rewrite this probability as:

$$
P(d^*_{ijt} = 1 | d^*_{it} = 1, X_i, Z_i, J, J'_{it}; (\gamma, \beta)) = \frac{\mathbb{1}\{j \neq 0, j \in J'_{it}\} \sum_{l \in \mathcal{L}} \exp(V_{ijl})}{\sum_{j' \in J'_{it}} \sum_{l \in \mathcal{L}} \exp(V_{ij'l})},
$$

which is identical to that in equation (7) in the main text. Note that, once we condition on the set of non-reviewed restaurants J'_{it}, this probability does not depend on the complete set of restaurants J; therefore,

$$
P(d^*_{ijt} = 1 | d^*_{it} = 1, X_i, Z_i, J, J'_{it}; (\gamma, \beta)) = P(d^*_{ijt} = 1 | d^*_{it} = 1, X_i, Z_i, J'_{it}; (\gamma, \beta)).
$$

Deriving equation (9). The conditional probability of an individual i writing a review about venue j at period t, given a randomly drawn set S_{it} and that i wrote a review (about some restaurant) at period t, is:

$$\begin{aligned}
&P(d^*_{ijt} = 1 | d^*_{it} = 1, X_i, Z_i, S_{it}, J'_{it}; (\gamma, \beta)) = \\
&\frac{P(S_{it} | d^*_{ijt} = 1, d^*_{it} = 1, X_i, Z_i, S_{it}, J'_{it}; (\gamma, \beta)) P(d^*_{ijt} = 1 | d^*_{it} = 1, X_i, Z_i, J'_{it}; (\gamma, \beta))}{\sum_{j' \in J'_{it}} P(S_{it} | d^*_{ijt} = 1, d^*_{it} = 1, X_i, Z_i, S_{it}, J'_{it}; (\gamma, \beta)) P(d^*_{ijt} = 1 | d^*_{it} = 1, X_i, Z_i, J'_{it}; (\gamma, \beta))} = \\
&\frac{\pi(S_{it} | d^*_{ijt} = 1, J'_{it}) P(d^*_{ijt} = 1 | d^*_{it} = 1, X_i, Z_i, J'_{it}; (\gamma, \beta))}{\sum_{j' \in J'_{it}} \pi(S_{it} | d^*_{ijt} = 1, J'_{it}) P(d^*_{ijt} = 1 | d^*_{it} = 1, X_i, Z_i, J'_{it}; (\gamma, \beta))} = \\
&\frac{\pi(S_{it} | d^*_{ijt} = 1, J'_{it}) P(d^*_{ijt} = 1 | d^*_{it} = 1, X_i, Z_i, J'_{it}; (\gamma, \beta))}{\sum_{j' \in S_{it}} \pi(S_{it} | d^*_{ijt} = 1, J'_{it}) P(d^*_{ijt} = 1 | d^*_{it} = 1, X_i, Z_i, J'_{it}; (\gamma, \beta))} = \\
&\frac{\kappa_{it} \mathbb{1}\{j \in S_{it}\} P(d^*_{ijt} = 1 | d^*_{it} = 1, X_i, Z_i, J'_{it}; (\gamma, \beta))}{\sum_{j' \in S_{it}} \kappa_{it} P(d^*_{ijt} = 1 | d^*_{it} = 1, X_i, Z_i, J'_{it}; (\gamma, \beta))} = \\
&\frac{\mathbb{1}\{j \in S_{it}\} P(d^*_{ijt} = 1 | d^*_{it} = 1, X_i, Z_i, J'_{it}; (\gamma, \beta))}{\sum_{j' \in S_{it}} P(d^*_{ijt} = 1 | d^*_{it} = 1, X_i, Z_i, J'_{it}; (\gamma, \beta))},
\end{aligned}$$

The first equality comes by applying Bayes' rule. The second equality accounts for the fact that, once we condition on the observed review of individual i at period t, our procedure to draw the samples of venues S_{it} does not depend on any of the observed characteristics affecting the utility function U_{ijlt} in equation (1). Finally, the third, fourth and fifth equalities are implied by equation (8). Combining the last expression above and equation (7), we obtain that, for every $j \in S_{it}$

$$\begin{aligned}
P(d^*_{ijt} = 1 | d^*_{it} = 1, X_i, Z_i, S_{it}, J'_{it}; (\gamma, \beta)) &= \frac{\mathbb{1}\{j \in S_{it}\} p^*_{it} \mathbb{1}\{j \neq 0, j \notin d^*_{it}\} \frac{\sum_{l \in \mathcal{L}} \exp(V_{ijl})}{\sum_{j''' \in J'_{it}} \sum_{l \in \mathcal{L}} \exp(V_{ij'''l})}}{\sum_{j' \in S_{it}} p^*_{it} \mathbb{1}\{j' \neq 0, j' \notin d^*_{it}\} \frac{\sum_{l \in \mathcal{L}} \exp(V_{ij'l})}{\sum_{j'' \in J'_{it}} \sum_{l \in \mathcal{L}} \exp(V_{ij''l})}} \\
&= \frac{\mathbb{1}\{j \in S_{it}\} \mathbb{1}\{j \neq 0, j \in J'_{it}\} \sum_{l \in \mathcal{L}} \exp(V_{ijl})}{\sum_{j' \in S_{it}} \mathbb{1}\{j' \neq 0, j' \in J'_{it}\} \sum_{l \in \mathcal{L}} \exp(V_{ij'l})} \\
&= \frac{\mathbb{1}\{j \in S_{it}\} \sum_{l \in \mathcal{L}} \exp(V_{ijl})}{\sum_{j' \in S_{it}} \sum_{l \in \mathcal{L}} \exp(V_{ij'l})}
\end{aligned}$$

where the second equality cancels any term appearing both in the numerator and the denominator and the third equality takes into account that $S_{it} \in J'_{it}$, and, therefore, $\mathbb{1}\{j \neq 0, j \in J'_{it}\} = 1$ for all elements of the set S_{it}.

Deriving equation (10). The probability that an individual i reviews the restaurants $\{j_{i1}, j_{i2}, \dots, j_{iT_i}\}$ conditional on observing at least one review written by i in each of the periods $\{1, \dots, T_i\}$ and on the randomly drawn sets $\{S_{i1}, \dots, S_{iT_i}\}$ may be written as

$$P(d^*_{ij_{i1}1} = 1, \dots, d^*_{ij_{iT_i}T_i} = 1 | d^*_{i1} = 1, \dots, d^*_{iT_i} = 1, S_{i1}, \dots, S_{iT_i}, X_i, Z_i, J; (\gamma, \beta)).$$

Using the relationship between joint and conditional probabilities, we can rewrite this joint probability as

$$
\begin{aligned}
P(d^*_{ij_{i1}1} = 1, \ldots, d^*_{ij_{iT_i}T_i} = 1 | d^*_{i1} = 1, \ldots, d^*_{iT_i} = 1, S_{i1}, \ldots, S_{iT_i}, X_i, Z_i, J; (\gamma, \beta)) = \\
P(d^*_{ij_{iT_i}T_i} = 1 | d^*_{i1} = 1, \ldots, d^*_{iT_i} = 1, X_i, Z_i, J, S_{i1}, \ldots, S_{iT_i}, d^*_{ij_11} = 1, \ldots, d^*_{ij_{T_i-1}T_i-1} = 1; (\gamma, \beta) \times \\
P(d^*_{ij_{i1}1} = 1, \ldots, d^*_{ij_{iT_i-1}T_i-1} = 1 | d^*_{i1} = 1, \ldots, d^*_{iT_i} = 1, X_i, Z_i, J, S_{i1}, \ldots, S_{iT_i}; (\gamma, \beta)) = \\
P(d^*_{ij_{iT_i}T_i} = 1 | d^*_{iT_i} = 1, X_i, Z_i, J, S_{iT_i}, d^*_{ij_11} = 1, \ldots, d^*_{ij_{iT_i-1}T_i-1} = 1; (\gamma, \beta)) \times \\
P(d^*_{ij_{i1}1} = 1, \ldots, d^*_{ij_{iT_i-1}T_i-1} = 1 | d^*_{i1} = 1, \ldots, d^*_{iT_i-1} = 1, X_i, Z_i, J, S_{i1}, \ldots, S_{iT_i-1}; (\gamma, \beta)) = \\
P(d^*_{ij_{iT_i}T_i} = 1 | d^*_{iT_i} = 1, X_i, Z_i, S_{T_i}; (\gamma, \beta)) \times \\
P(d^*_{ij_{i1}1} = 1, \ldots, d^*_{ij_{T_i-1}T_i-1} = 1 | d^*_{i1} = 1, \ldots, d^*_{iT_i-1} = 1, X_i, Z_i, J, S_{i1}, \ldots, S_{iT_i-1}; (\gamma, \beta)).
\end{aligned}
$$

The second equality takes into account that, conditional on $(d^*_{ij_{i1}1} = 1, \ldots, d^*_{ij_{iT_i-1}T_i-1} = 1)$, neither the vector of dummies $(d^*_{i1} = 1, \ldots, d^*_{iT_i-1} = 1)$ nor the vector of random sets $(S_{i1}, \ldots, S_{iT_i-1})$ provide any information on the actual restaurant reviewed at period T_i. The third equality takes into account that all the information on review probabilities at period T_i contained in the specific vector of past reviews $(d^*_{ij_11} = 1, \ldots, d^*_{ij_{T_i-1}T_i-1} = 1)$ is already contained in the randomly drawn set S_{iT_i}. Therefore, we can rewrite the joint probability that we observe an individual i reviewing the restaurants $\{j_{i1}, j_{i2}, \ldots, j_{iT_i}\}$ as the product of the probability that we observe the review j_{iT_i} conditional on the set S_{iT_i} and the joint probability that we observe individual i reviewing the other $T_i - 1$ restaurants, $\{j_{i1}, j_{i2}, \ldots, j_{iT_i-1}\}$. Analogously, we can also rewrite this joint probability as the product of the probability that we observe i reviewing j_{iT_i-1} conditional on the set S_{iT_i-1} and the joint probability that we observe i reviewing the remaining $T_i - 2$ restaurants, $\{j_{i1}, j_{i2}, \ldots, j_{iT_i-2}\}$. Therefore, iterating these steps T_i times, we obtain the expression in equation (10).

C.2 Estimation procedure: simulation

In this section, we simulate data from simple variants of the model described in sections 3.1 and 3.2 for the purpose of illustrating the asymptotic properties of the estimator described in Section 3.3.

We generate data for 400 individuals with identical preference parameters (γ, β) who each make 40 choices, for a total of 16,000 trips. Each individual is located at a randomly drawn origin, from which they have one transport mode to reach 1,000 restaurants with randomly drawn locations and ratings.[63] Individual i's utility from choosing restaurant j at period t is $U_{ijt} = -\ln distance_{ij} + rating_j + \nu_{ijt}$.Therefore, using the notation introduced in Section 3.1, $\mathcal{L}$ is a singleton, the cardinality of the set J is 1,000, there is a single demographic group g, and the vector of preference parameters is $\{\gamma^1_l = \gamma^1 = -1, \gamma^2 = 0, \beta^1 = 1, \beta^2 = 0\}$. Consistent with our model, the terms $\{\nu_{ijt}, \forall i, j, t\}$ are assumed to follow *iid* logistic distributions. Conditional on visiting a restaurant, every individual i writes a review with probability 0.5 if she did not previously review the venue, and 0 otherwise. Therefore, using

[63]Specifically, each individual and restaurant has a location that is randomly drawn according to $latitude \sim U(40.75, 41.75)$ and $longitude \sim U(-74.25, -73.25)$. Restaurant ratings are drawn from $rating \sim U(1, 5)$. All draws are independent of each other.

the notation introduced in Section 3.2, $p^*_{it} = 0.5$ for all i and t. In our randomly generated sample, we observe 7,521 reviews.

The first column in Table C.1 reports estimates that maximize a likelihood function that (a) uses information on the restaurant visited for all 40 trips by all 400 sampled individuals and (b) includes the entire choice set J containing all $1,000$ restaurants. Specifically, we compute the estimates in column one by maximizing the log-likelihood function

$$LL_1 = \sum_{i=1}^{400}\sum_{t=1}^{40}\sum_{j\in J} \mathbb{1}\{d_{ijt} = 1\} \ln\left(\frac{\exp(\gamma^1 ln(distance_{ij}) + \beta^1 rating_j)}{\sum_{j'\in J}\exp(\gamma^1 ln(distance_{ij'}) + \beta^1 rating_{j'})}\right)$$

Not surprisingly, we obtain estimates of the impact of $\ln(distance)$ and rating on individuals' utility that are very close to their true values of -1 and 1. In our empirical application, we cannot estimate preference parameters this way because we do not observe every restaurant visit.

In columns two through five, we infer visits using only information from reviews, as in our empirical application. Columns two and three illustrate the consequences of two possible mistakes that a researcher might make when using information on reviews rather than actual visits. Specifically, column two illustrates the consequences of not taking into account that individuals do not review restaurants that they have previously reviewed; its estimates maximize the log-likelihood function

$$LL_2 = \sum_{i=1}^{400}\sum_{t=1}^{40}\sum_{j\in J} \mathbb{1}\{d^*_{ijt} = 1\} \ln\left(\frac{\exp(\gamma^1 ln(distance_{ij}) + \beta^1 rating_j)}{\sum_{j'\in J}\exp(\gamma^1 ln(distance_{ij'}) + \beta^1 rating_{j'})}\right).$$

Column three illustrates the consequences of over-correcting and assigning to each individual a choice set that excludes all venues ever reviewed during the sample period; its estimates maximize the log-likelihood function

$$LL_3 = \sum_{i=1}^{400}\sum_{t=1}^{40}\sum_{j\in J} \mathbb{1}\{d^*_{ijt} = 1\} \ln\left(\frac{\exp(\gamma^1 ln(distance_{ij}) + \beta^1 rating_j)}{\sum_{j'\in J_{iT_i}}\exp(\gamma^1 ln(distance_{ij'}) + \beta^1 rating_{j'})}\right).$$

In column two the estimates are too small in absolute value, and in column three they are too large. Column four shows that one can consistently estimate preference parameters using only reviews. Key to the estimator's consistency is that we assign to each individual i at period t a choice set that excludes those restaurants reviewed by i prior to t:

$$LL_4 = \sum_{i=1}^{400}\sum_{t=1}^{40}\sum_{j\in J'_{it}} \mathbb{1}\{d^*_{ijt} = 1\} \ln\left(\frac{\exp(\gamma^1 ln(distance_{ij}) + \beta^1 rating_j)}{\sum_{j'\in J'_{it}}\exp(\gamma^1 ln(distance_{ij'}) + \beta^1 rating_{j'})}\right).$$

The estimates in column four are very close to the true parameter vector, in line with the mathematical proof and discussion in Section 3.3. Finally, column five shows that the estimator using information on only a subset S_{it} of the choice set J'_{it} also consistently estimates the preference parameters:

$$LL_5 = \sum_{i=1}^{400}\sum_{t=1}^{40}\sum_{j\in S_{it}} \mathbb{1}\{d^*_{ijt} = 1\} \ln\left(\frac{\exp(\gamma^1 log(distance_{ij}) + \beta^1 rating_j)}{\sum_{j'\in S_{it}}\exp(\gamma^1 log(distance_{ij'}) + \beta^1 rating_{j'})}\right),$$

with $S_{it} \subset J'_{it}$ and drawn randomly according to the probability distribution in equation (8). The results in column five are implied by the content of Section 3.3.

Table C.1: Choice sets and consistency in simulated data

Dummy for:	(1) Visit	(2) Review	(3) Review	(4) Review	(5) Review
ln(distance)	-1.01	-0.93	-1.09	-1.01	-1.01
	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)
rating	1.00	0.98	1.04	1.01	1.00
	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)
Choice set drawn from:	All restaurants	All restaurants	Never reviewed	Not previously reviewed	Not previously reviewed
Choice set:	J	J	J'_{iT_i}	J'_{it}	$S_{it} \in J'_{it}$
Choice set size:	$1,000$	$1,000$	i-specific	it-specific	20
Num. Obs.	16,000	7,521	7,521	7,521	7,521

Notes: Standard errors in parentheses. All five columns use information on the 400 users that form the randomly generated population of interest. Each of these users makes 40 choices. For each trip, each user writes a review about visited restaurant with probability 0.5.

C.3 A minimum-travel-time specification

Alternatives assumptions about the distribution of the vector of idiosyncratic terms ν_{it} and travel-time disutilities $\gamma^1_{g(i)l}$ yield a behavioral model in which individuals always select the fastest travel time. Consider a model in which the utility to individual i of visiting restaurant j at period t using origin-mode l may be represented as

$$U_{ijlt} = \gamma^1_{g(i)l} X^1_{ijl} + \gamma^2_{g(i)} X^2_{ij} + \beta^1_{g(i)} Z_j + \beta^2_{g(i)} Z_{ij} + \nu_{ijt},$$

where the variables X^1_{ijl}, X^2_{ij}, Z_j, and Z_{ij} are described in Section 3.1, and ν_{ijt} is an unobserved individual-restaurant-period specific characteristic. The difference between this demand model and that described in Section 3.1 is that the unobserved component ν does not vary across origin-mode l. This implies that, conditional on visiting a restaurant j at period t, an individual i uses the origin-mode l that maximizes the term $\gamma^1_{g(i)l} X^1_{ijl}$. Specifically, this means that the decision over the mode of transport that an individual i uses to visit a restaurant j is only a function of the parameter vector $\gamma^1_{g(i)} = \{\gamma^1_{g(i)l}, l \in \mathcal{L}\}$ and observed covariates. Denote the utility that individual i obtains from visiting restaurant j in period t if she uses the optimal origin-mode pair as

$$\begin{aligned} U_{ijt} &= \max_{l \in \mathcal{L}} \{U_{ijlt}\} = \max_{l \in \mathcal{L}} \{\gamma^1_{g(i)l} X^1_{ijl} + \gamma^2_{g(i)} X^2_{ij} + \beta^1_{g(i)} Z_j + \beta^2_{g(i)} Z_{ij} + \nu_{ijt}\} \\ &= \max_{l \in \mathcal{L}} \{\gamma^1_{g(i)l} X^1_{ijl}\} + \gamma^2_{g(i)} X^2_{ij} + \beta^1_{g(i)} Z_j + \beta^2_{g(i)} Z_{ij} + \nu_{ijt}. \end{aligned}$$

If $\gamma^1_{g(i)l} = \gamma^1_{g(i)} < 0$, i.e. the disutility of travel does not vary across origins or modes of transport and individuals dislike spending time to visit restaurants, then

$$U_{ijt} = \gamma^1_{g(i)} \min_{l \in \mathcal{L}} \{X^1_{ijl}\} + \gamma^2_{g(i)} X^2_{ij} + \beta^1_{g(i)} Z_j + \beta^2_{g(i)} Z_{ij} + \nu_{ijt}.$$

If we additionally assume that the vector $\nu_{it} = \{\nu_{ijt}; \forall j \in \mathcal{J}\}$ is independent across individuals and time periods and has a joint type I extreme value distribution, then the probability that individual i decides to visit restaurant j in period t will be

$$P(d_{ijt} = 1 | X_i, Z_i; (\gamma, \beta)) = \frac{\exp(\gamma^1_{g(i)} \min_{l \in \mathcal{L}} \{X^1_{ijl}\} + \gamma^2_{g(i)} X^2_{ij} + \beta^1_{g(i)} Z_j + \beta^2_{g(i)} Z_{ij})}{\sum_{j' \in \mathcal{J}} \exp(\gamma^1_{g(i)} \min_{l \in \mathcal{L}} \{X^1_{ij'l}\} + \gamma^2_{g(i)} X^2_{ij'} + \beta^1_{g(i)} Z_{j'} + \beta^2_{g(i)} Z_{ij'})},$$

where the vectors X_i Z_i, γ and β are defined in footnote 15. Given this probability and following the same steps described in Section 3.3, we derive the following log-likelihood function

$$\sum_{i=1}^{N} \sum_{t=1}^{T_i} \sum_{j \in S_{it}} \mathbb{1}\{d^*_{ijt} = 1\} \ln \left(\frac{\exp(V_{ij})}{\sum_{j' \in S_{it}} \exp(V_{ij'})} \right),$$

with

$$V_{ij} \equiv \gamma^1_{g(i)} \min_{l \in \mathcal{L}} \{X^1_{ijl}\} + \gamma^2_{g(i)} X^2_{ij} + \beta^1_{g(i)} Z_j + \beta^2_{g(i)} Z_{ij},$$

for every individual i and restaurant j. We report the estimates of such a model in Table A.10. The results are broadly similar to those in Table 3.

C.4 Nested logit

Step 1: derive restaurant visit probability. Let the set of restaurants J be partitioned into K non-overlapping subsets (nests) denoted $B_1, B_2, \ldots, B_K$. Assume that the probability an individual i visits restaurant j belonging to nest B_k from origin-mode l at period t is:

$$P(d_{ijt} = 1 | X_i, Z_i, J; (\gamma, \beta, \lambda)) = \frac{\left(\sum_{l \in \mathcal{L}} \exp(V_{ijl}/\lambda_{g(i)})\right) \left(\sum_{j' \in B_k} \sum_{l \in \mathcal{L}} \exp(V_{ij'l}/\lambda_{g(i)k})\right)^{\lambda_{g(i)k}-1}}{\sum_{k'=1}^{K} \left(\sum_{j' \in B_{k'}} \sum_{l \in \mathcal{L}} \exp(V_{ijl}/\lambda_{g(i)k'})\right)^{\lambda_{g(i)k'}}}, \tag{C.1}$$

where X_i, Z_i, γ, and β are defined in equation (15) and, for every individual i, restaurant j and mode-origin l, V_{ijl} is defined in equation (4). When $\lambda_{g(i)k} = 1$ for all k, indicating no correlation among the unobserved components of utility for alternatives within a nest, the choice probabilities become those in our baseline model (see equation (3)). In our application, we assume that this correlation parameter is common across all nests, such that $\lambda_{g(i)k} = \lambda_{g(i)k'}$ for any pair of nests (k, k').

Step 2: derive restaurant review probability. Given the review-writing model described in Section 3.2, the probability of observing a review of venue j written by individual i at period t, conditional on individual i writing a review at t, is

$$P(d^*_{ijt} = 1 | d^*_{it} = 1, X_i, Z_i, J'_{it}; (\gamma, \beta, \lambda)) =$$
$$\frac{\mathbb{1}\{j \neq 0, j \in J'_{it}\}\big(\sum_l \exp(V_{ijl}/\lambda_{g(i)})\big)\big(\sum_{j'' \in B(j)} \sum_{l \in \mathcal{L}} \exp(V_{ij''l}/\lambda_{g(i)})\big)^{\lambda_{g(i)-1}}}{\sum_{j' \in J'_{it}} \big\{\big(\sum_{l \in \mathcal{L}} \exp(V_{ij'l}/\lambda_{g(i)})\big)\big(\sum_{j'' \in B(j')} \sum_{l \in \mathcal{L}} \exp(V_{ij''l}/\lambda_{g(i)})\big)\big\}}$$

where $d^*_{it} = \sum_{j=1}^J d^*_{ijt}$ is a dummy variable that equals one if i writes a review at period t, and $B(j)$ denotes the nest to which restaurant j belongs. Defining

$$I_{ij}(\gamma_{g(i)}, \beta_{g(i)}, \lambda_{g(i)}) = (\lambda_{g(i)} - 1) \ln\Big(\sum_{j'' \in B(j)} \sum_{l \in \mathcal{L}} \exp(V_{ij''l}/\lambda_{g(i)})\Big),$$

we can rewrite the probability that i writes a review about a restaurant j (not previously reviewed) at period t conditional on observing a review (about any restaurant) written by i at t as

$$P(d^*_{ijt} = 1 | d^*_{it} = 1, X_i, Z_i, J'_{it}; (\gamma, \beta)) =$$
$$\frac{\big(\sum_{l \in \mathcal{L}} \exp((V_{ijl}/\lambda_{g(i)})\big) \exp(I_{ij}(\gamma_{g(i)}, \beta_{g(i)}, \lambda_{g(i)}))}{\sum_{j' \in J'_{it}} \big(\sum_{l \in \mathcal{L}} \exp((V_{ij'l}/\lambda_{g(i)})\big) \exp(I_{ij'}(\gamma_{g(i)}, \beta_{g(i)}, \lambda_{g(i)}))} =$$
$$\frac{\big(\sum_{l \in \mathcal{L}} \exp((V_{ijl}/\lambda_{g(i)}) + I_{ij}(\gamma_{g(i)}, \beta_{g(i)}, \lambda_{g(i)})\big)}{\sum_{j' \in J'_{it}} \big(\sum_{l \in \mathcal{L}} \exp((V_{ij'l}/\lambda_{g(i)}) + I_{ij'}(\gamma_{g(i)}, \beta_{g(i)}, \lambda_{g(i)})\big)}. \tag{C.2}$$

Step 3: reduce choice set. The cardinality of the choice set J'_{it} makes it computationally burdensome to construct the denominator of the probability in equation (C.2). To address this dimensionality issue, for every individual i and period t in which we observe a review written by i, we randomly draw a choice set following the procedure described in Section 3.3. The probability of randomly drawing each set S_{it} is thus that in equation (8). Given equations (8) and (C.2), we can write the probability that i reviews restaurant j at period t conditional on a randomly drawn set S_{it} and that i writes a review at t as

$$P(d^*_{ijt} = 1 | d^*_{it} = 1, X_i, Z_i, S_{it}; (\gamma, \beta)) =$$
$$\frac{\mathbb{1}\{j \in S_{it}\}\big(\sum_{l \in \mathcal{L}} \exp((V_{ijl}/\lambda_{g(i)}) + I_{ij}(\gamma_{g(i)}, \beta_{g(i)}, \lambda_{g(i)})\big)}{\sum_{j' \in S_{it}} \big(\sum_{l \in \mathcal{L}} \exp((V_{ij'l}/\lambda_{g(i)}) + I_{ij'}(\gamma_{g(i)}, \beta_{g(i)}, \lambda_{g(i)})\big)}. \tag{C.3}$$

Step 4: derive individual i-specific likelihood function. Using j_{it} to denote the restaurant reviewed by individual i at period t, the joint probability of observing an individual i writing the T_i reviews $\{j_{i1}, j_{i2}, \ldots, j_{iT_i}\}$ conditional on observing a review written by i in each of the periods $\{1, \ldots, T_i\}$ and on randomly drawing the sets $\{S_{i1}, S_{i2}, \ldots, S_{iT_i}\}$ is

$$\prod_{t=1}^{T_i} \frac{\mathbb{1}\{j \in S_{it}\}\big(\sum_{l \in \mathcal{L}} \exp((V_{ijl}/\lambda_{g(i)}) + I_{ij}(\gamma_{g(i)}, \beta_{g(i)}, \lambda_{g(i)})\big)}{\sum_{j' \in S_{it}} \big(\sum_{l \in \mathcal{L}} \exp((V_{ij'l}/\lambda_{g(i)}) + I_{ij'}(\gamma_{g(i)}, \beta_{g(i)}, \lambda_{g(i)})\big)}. \tag{C.4}$$

Step 5: derive log-likelihood function. Given equation (C.4) and assuming that we observe a random sample $i = 1, \ldots, N$ of individuals from the population of interest, we can write the log-likelihood function as

$$\sum_{i=1}^{N} \sum_{t=1}^{T_i} \sum_{j \in S_{it}} \mathbb{1}\{d^*_{ijt} = 1\} \ln \left(\frac{\left(\sum_{l \in \mathcal{L}} \exp((V_{ijl}/\lambda_{g(i)}) + I_{ij}(\gamma_{g(i)}, \beta_{g(i)}, \lambda_{g(i)})\right)}{\sum_{j' \in S_{it}} \left(\sum_{l \in \mathcal{L}} \exp((V_{ij'l}/\lambda_{g(i)}) + I_{ij'}(\gamma_{g(i)}, \beta_{g(i)}, \lambda_{g(i)})\right)} \right). \quad \text{(C.5)}$$

C.5 Moment inequalities

Katz (2007) and Pakes (2010) introduce an estimation approach that uses moment inequalities to handle both large, potentially unobserved choice sets and unobserved heterogeneity in the individuals' preferences for some observed choice characteristics. Applied to our setting, in this approach, the utility for an individual i of visiting venue j in period t from origin-mode l may be written as:

$$U_{ijlt} = \gamma^1_{li} \mathbb{E}[X^1_{ijl} | \mathcal{I}_{it}] + \gamma^2_i \mathbb{E}[X^2_{ij} | \mathcal{I}_{it}] + \beta^1_i \mathbb{E}[Z^1_j | \mathcal{I}_{it}] + \beta^2_i \mathbb{E}[Z^2_{ij} | \mathcal{I}_{it}], \quad \text{(C.6)}$$

where $\gamma^1_{li} = \gamma^1_l + \varepsilon^1_{li}$, $\gamma^2_i = \gamma^2 + \varepsilon^2_i$, $\beta^1_i = \beta^1 + \varepsilon^3_i$, $\beta^2_i = \beta^2 + \varepsilon^4_i$ and $\mathcal{I}_{it}$ denotes the information set of individual i at the time of deciding which restaurant to visit at period t. Under the assumption that $\mathbb{E}_i[\varepsilon^1_{li}] = \mathbb{E}_i[\varepsilon^2_i] = \mathbb{E}_i[\varepsilon^3_i] = \mathbb{E}_i[\varepsilon^4_i] = 0$, where $\mathbb{E}_i[\cdot]$ denotes the expectation across individuals in the population of interest, Katz (2007) and Pakes (2010) show how to derive moment inequalities that bound the average preference parameters (γ, β). The behavioral model in equation (C.6) differs from that in equation (1) in that: (a) allows consumers to have imperfect information about the characteristics of the different restaurants at the time of deciding which restaurant to visit; (b) allows individuals to differ in their preferences for the different observed characteristics affecting U_{ijlt}; (c) assumes that there is no additional individual-restaurant-origin-mode specific characteristics that affects individual choices and is unobserved to the econometrician (i.e. assumes away the logit shock $\{\nu_{ijlt}; l \in \mathcal{L}, j \in J\}$ included in equation (1)).[64]

There are three reasons why we opt for the demand model described in Section 3.1 instead of the model in equation (C.6). First, the observed restaurant and locational characteristics affecting U_{ijlt} are publicly available through Yelp.com, Google Maps, and SocialExplorer.com, so it is unlikely that individuals make large mistakes when forecasting variables like the time that it takes to travel to a venue or the average price of each restaurant.[65]

[64] Dickstein and Morales (2016) show how to estimate a binary choice model in which consumers may have imperfect information about observable choice characteristics and their choices may be affected by individual-choice specific unobserved shocks. However, as the estimator introduced in Dickstein and Morales (2016) cannot handle the large choice sets that consumers face in our empirical application, it is not ideal for our setting.

[65] While the NYPD only started making incident-level crime maps available on its website in December 2013, precinct-level crime statistics have been available on the NYPD website since 2003 and updated weekly. During our study period of 2007-2011, local newspapers like the *New York Times* produced incident-level maps based on felony reports.

Second, while using moment inequalities to estimate and perform inference on bounds on a small set of parameters is computationally straightforward (e.g. Holmes (2011); Eizenberg (2014); Morales et al. (2017); Dickstein and Morales (2016); Wollman (2016)), doing so for the set of parameters that we estimate in some of our specifications (i.e. those accounting simultaneously for spatial and social frictions) is computationally unrealistic.[66] Third, even if we are controlling for a large set of observed restaurant characteristics, it is likely that there are still multiple unobservable factors (e.g. is the restaurant child-friendly? do the other people in my party like the restaurant? do I feel like eating at a French restaurant today?) that may vary across individual-restaurant-occasions triplets and that are likely to be important in determining individuals' restaurant picks. The demand model in Section 3.1 accounts for all these different factors through the unobserved preference shocks $\{\nu_{ijlt}; l \in \mathcal{L}, j \in J_t\}$; conversely, the behavioral model in equation (C.6) assumes these factors away.

C.6 Endogenous home and work locations

The statistical model described in Section 3.1 implicitly assumes that individuals' home and work locations are exogenously determined. However, in practice, individuals optimally choose where to live and work. Consequently, the home and work locations of every individual in our sample may be endogenously determined as a function of the characteristics of the restaurants that they may consider visiting. In this section, we allow home and work locations to be endogenously determined and discuss the assumptions that we would need to impose so that the endogenous choice of home and work locations does not bias the estimates of the preference parameters (γ, β) obtained using the estimation approach in Section 3.3.

Assume that, in some period 0, individuals choose their home and work locations by maximizing a utility function that is a weighted average of: (a) the expected utility of visiting restaurants in future periods, and (b) a function of the characteristics of the home and work locations that have intrinsic value independently of their properties as sites from where to launch consumption.

In order to compute the expected utility of future restaurant visits, we need to make an assumption on the content of agents' information sets at the time at which they decide on their home and work locations. Using the notation in Section 3.1, we assume that, at the time of deciding on where to live and work, every individual i knows the value of the vector $\{(X^1_{ijl}, X^2_{ij}, Z^1_j, Z^2_{ij}); l \in \mathcal{L}, j \in J\}$. Individual i also knows the distribution –but ignores the realizations– of the preference shocks $\{\nu_{ijlt}; l \in \mathcal{L}, j \in J\}$ corresponding to any period t subsequent to that when the decision on the location of home and work is taken. Under this assumption, the expected utility for individual i of visiting restaurants from a particular home and work location (h, w) is

$$E_{ihw} = \sum_{t \in \mathcal{T}} \log \Big(\sum_{j \in J} \sum_{l \in \mathcal{L}} \exp(\gamma^1_{g(i)l} X^1_{ijl} + \gamma^2_{g(i)} X^2_{ij} + \beta^1_{g(i)} Z^1_j + \beta^2_{g(i)} Z^2_{ij}) \Big) + c,$$

where c is the Euler's constant, and $\mathcal{T}$ is the set of periods at which individual i will visit a restaurant.

[66] Applying the standard inference procedure to compute confidence sets Andrews and Soares (2010) for the large number of characteristics included in our exercise would be computationally prohibitive.

For every possible pair of home and work locations (h, w), we denote the vector of characteristics defining their intrinsic value (independently of their properties as locations from where to launch consumption) for individual i as Y_{ihw}. If we denote as α the weight that individuals assign to the expected utility that they will obtain from visiting restaurants, we can write the utility for an individual i of living in location h and working in location w as

$$W_{ihw} = (1-\alpha)\omega Y_{ihw} + \alpha E_{ihw},$$

where ω is a parameter vector of identical dimensions as Y_{ihw} that determines the impact of each of these characteristics on the utility for individual i of establishing her residence in location h and her workplace in location w. An individual i lives in location h^{opt} and works in location w^{opt} if

$$(h^{opt}, w^{opt}) = \arg \max_{w \in \mathcal{W}, h \in \mathcal{H}} \{(1-\alpha)\omega Y_{ihw} + \alpha E_{ihw}\}, \tag{C.7}$$

where $\mathcal{W}$ denotes the set of all possible work locations and $\mathcal{H}$ denotes the set of all possible home locations.

If we were to extend the model of restaurant choice in Section 3 to account for the endogenous selection of home and work locations, the relevant probability entering the likelihood function for our sample would be the probability that an individual i chooses to visit restaurant j at period t conditional on having chosen to live in location h^{opt} and work in location w^{opt}:

$$P(d_{ijt} = 1 | V_{it}, h_i = h^{opt}, w_i = w^{opt}; \alpha) = \sum_{l \in \mathcal{L}} P(d_{ijlt} = 1 | V_{it}, h_i = h^{opt}, w_i = w^{opt}; \alpha),$$

and

$$P(d_{ijlt} = 1 | V_{it}, h_i = h^{opt}, w_i = w^{opt}; \alpha) =$$
$$\int_{\nu_{it}} \mathbb{1}\{V_{ijl} + \nu_{ijlt} > V_{ij'l} + \nu_{ij'l't}; j' \in J_t, l \in \mathcal{L}\} f(\nu_{it} | V_{it}, h_i = h^{opt}, w_i = w^{opt}; \alpha) d\nu_{it},$$

where $\nu_{it} = \{\nu_{ijlt}; j \in J_t, l \in \mathcal{L}\}$ and $f(\nu_{it}|V_i, h_i = h^{opt}, w_i = w^{opt}; \alpha)$ denotes the density function of the vector ν_{it} conditional on the vector of observed characteristics determining the utility of restaurant visits, $V_{it} \equiv \{V_{ijlt}; l \in \mathcal{L}, j \in J\}$, and conditional on the observed house and work locations h_i and w_i being the optimal choices of individual i. Using equation (C.7), we can rewrite this density function as:

$$f\big(\nu_{it} | V_{it}, (h^{opt}, w^{opt}) = \arg \max_{w \in \mathcal{W}, h \in \mathcal{H}} \{(1-\alpha)\omega Y_{ihw} + \alpha E_{ihw}\}\big).$$

This representation of the density function clearly shows that we can recover the choice probability in equation (5) in the main text as long as we assume that the distribution of the vector of unobserved restaurant characteristics affecting individuals' restaurant choices, ν_{it}, verifies two conditions: (a) it is independent of the vector of characteristics determining the optimal selection of home and work location, V_{it} and $\{Y_{ihw}, h \in \mathcal{H}, w \in \mathcal{W}\}$; (b) it is distributed type I extreme value. The model in Section 3 already imposes that the distribution of the vector ν_{it} is type I extreme value and independent of V_{it}. Therefore, under the model for home and work location described above, allowing individuals to optimally determine their home and work will not bias the estimates described in Section 3 as long as we impose the additional restriction that ν_{it} is independent of the vector $\{Y_{ihw}, w \in \mathcal{W}, h \in \mathcal{H}\}$ conditional on V_{it}.

C.7 Review writing

Here we discuss the potential bias that would affect our estimates of the parameter vector (γ, β) if the probability that a user reviews a restaurant depends on some characteristic of this restaurant that is observed by the econometrician and that is correlated with some covariate in the vector (X_i, Z_i). Assume that the probability that individual i reviews restaurant j upon visiting j at period t depends on some of the restaurant characteristics included in Z^1_j or Z^2_{ij} through the following function

$$P(d^{r*}_{ijt}|d_{ijt} = 1, X_i, Z_i, J, J'_{it}; (\gamma, \beta, p^*_{it})) = p^*_{it}\mathbb{1}\{j \neq 0, j \in J'_{it}\} \exp(\zeta^1_{g(i)} Z^1_j + \zeta^2_{g(i)} Z^2_{ij}). \quad \text{(C.8)}$$

In the case when $\zeta^1_g = \zeta^2_g = 0$ for every group g this function is identical to that assumed in the main text (see equation 6). Conversely, if either ζ^1_g or ζ^2_g differ from zero for some group g, we can write the probability that we observe a review at period t written by a user i about restaurant j as:

$$P(d^*_{ijt} = 1|X_i, Z_i, J, J'_{it}; (\gamma, \beta, p^*_{it})) = p^*_{it}\mathbb{1}\{j \neq 0, j \in J'_{it}\}$$
$$\times \frac{\left(\sum_{l \in \mathcal{L}} \exp(\gamma^1_{g(i)l} X^1_{ijl} + \gamma^2_{g(i)} X^2_{ij} + (\beta^1_{g(i)} + \zeta^1_{g(i)}) Z^1_j + (\beta^2_{g(i)} + \zeta^2_{g(i)}) Z^2_{ij})\right)}{\sum_{j' \in J_t} \left(\sum_{l \in \mathcal{L}} \exp(\gamma^1_{g(i)l} X^1_{ij'l} + \gamma^2_{g(i)} X^2_{ij'} + (\beta^1_{g(i)} + \zeta^1_{g(i)}) Z^1_{j'} + (\beta^2_{g(i)} + \zeta^2_{g(i)}) Z^2_{ij'})\right)}, \quad \text{(C.9)}$$

This expression demonstrates that we cannot separately identify the parameter vectors $\beta^1_{g(i)}$ and $\beta^2_{g(i)}$ from the parameter vectors $\zeta^1_{g(i)}$and $\zeta^2_{g(i)}$. However, the estimates of γ are not affected by the fact that the probability of writing a review depends on the vectors of restaurant characteristics Z^1_j and Z^2_{ij}. Furthermore, the probability in equation (C.9) is identical to that in equation (6) with the only exception that the parameter vectors $\tilde{\beta_{g(i)}}^1 = \beta^1_{g(i)} + \zeta^1_{g(i)}$ and $\tilde{\beta_{g(i)}}^2 = \beta^2_{g(i)} + \zeta^2_{g(i)}$ take the place of the parameter vectors $\beta^1_{g(i)}$ and $\beta^2_{g(i)}$ in the main text. Therefore, following the same steps indicated in the main text, one can derive a likelihood function that identifies the parameter vector $(\gamma, \tilde{\beta})$. This means that an expression for the probability of writing a restaurant review as in equation (C.8) does not prevent us from obtaining consistent estimates of the parameters capturing the spatial frictions, γ^1, and the parameters capturing the social frictions, γ^2. However, the coefficients on the restaurant characteristics in the vectors Z^1_j and Z^2_{ij} account both for their effect on consumers' propensity to visit and their effect on visitors' propensity to write reviews: separately identifying these two effects is not possible.

C.8 Serial correlation in unobserved preferences

As we discuss in Section 3.3, as long as users' unobserved restaurant preferences (captured in the vector ν_{it}) are independent over time, the facts that we identify users' preferences from their Yelp reviews and Yelp users very rarely review any restaurant a second time do not prevent us from identifying consumers' preference parameters. The key to identifying preference parameters in this case is to compare the restaurant chosen by each individual i in each time period t to the set of restaurants this user has not previously reviewed.

If, contrary to our assumption, the preference shocks ν_{it} are correlated over time, the facts that we identify users' preferences from their Yelp reviews and that users do not generally review a restaurant twice can generate a selection bias in our estimates of consumers' preference parameters. To understand this bias, assume for simplicity that the unobserved preference shocks affecting individual i's utility of visiting restaurant j through origin-mode l at period t are the sum of a permanent, origin-mode-independent term ω_{ij} and the serially uncorrelated term ν_{ijlt} already incorporated in our baseline model. In this case, we can rewrite the utility function in equation (1) as

$$U_{ijlt} = \gamma^1_{g(i)l} X^1_{ijl} + \gamma^2_{g(i)} X^2_{ij} + \beta^1_{g(i)} Z^1_j + \beta^2_{g(i)} Z^2_{ij} + \tilde{\nu}_{ijlt}, \tag{C.10}$$

with $\tilde{\nu}_{ijt} \equiv a \times \omega_{ij} + \nu_{ijlt}$ and where a is a constant that governs the magnitude of the standard deviation of the permanent shock relative to the transitory component and is equal to zero in our baseline model.

Conditional on observable characteristics, user i will more often visit restaurants with higher values of the preference shocks $\tilde{\nu}_{ijlt}$. Given the assumptions on review-writing behavior in Section 3.2, user i is therefore more likely to review those restaurants earlier. Consequently, even if the unobserved preference shocks $\tilde{\nu}_{ijlt}$ are uncorrelated with all observed restaurant characteristics across all restaurants in NYC (i.e. across all restaurants in the set J defined in Section 3.1), this correlation will be non-zero for the subset of restaurants not previously reviewed by consumer i at any period t (i.e. for the restaurants included in J'_{it}). Specifically, for the subset of restaurants not previously reviewed, $\tilde{\nu}_{ijlt}$ will be negatively correlated with the part of the utility function in equation (C.10) that is a function of observable characteristics and parameters to estimate:

$$\gamma^1_{g(i)l} X^1_{ijl} + \gamma^2_{g(i)} X^2_{ij} + \beta^1_{g(i)} Z^1_j + \beta^2_{g(i)} Z^2_{ij}.$$

Consequently, if the preference shocks $\tilde{\nu}_{ijlt}$ are correlated over time, we should expect our estimation procedure to yield an upward bias in coefficients that are negative and a downward bias in coefficients that are positive. In other terms, we should expect an attenuation bias in all our estimates. However, by the same logic, an estimator employing only a subset of users' earlier reviews would suffer this selection bias less.

To verify this reasoning about the nature of the attenuation bias caused by serial correlation in unobserved preferences and the improvement associated with restricting attention to users' earlier reviews, we have simulated the following data-generating process:

$$U_{ijt} = 1.0 \times \text{rating}_j - 1.0 \times \text{distance}_{ij} + a \times \omega_{ij} + \nu_{ijt}$$

where both the permanent shock ω_{ij} and the transitory shock ν_{ijt} have type-1 extreme value distributions. In Table C.2, we apply our estimator to three samples of data generated from this process. Each sample includes 100 users making 40 trips. In columns 1-3, the choice

Table C.2: Estimation in the presence of permanent ω_{ij} shocks

	$a=1, J=1000$			$a=0.5, J=1000$			$a=1, J=11000$		
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Rating	0.84	0.85	0.89	0.97	0.99	1.01	0.85	0.85	0.92
	(0.023)	(0.034)	(0.056)	(0.020)	(0.029)	(0.047)	(0.021)	(0.030)	(0.051)
Distance	-0.82	-0.86	-0.86	-1.00	-1.00	-0.99	-0.87	-0.91	-0.90
	(0.025)	(0.035)	(0.055)	(0.020)	(0.027)	(0.042)	(0.022)	(0.030)	(0.049)
Sample share	1	1/2	1/5	1	1/2	1/5	1	1/2	1/5
Reviews	2203	1080	404	3408	1679	640	2738	1340	509

NOTES: Each triplet of columns reports three estimates applied to subsets of one draw from the data-generating process $U_{ijt} = 1.0 \times \text{rating}_j - 1.0 \times \text{distance}_{ij} + a \times \omega_{ij} + \nu_{ijt}$. In each draw, there are 100 users who make 40 trips to restaurants. Since users do not review restaurants they have previously visited, there are fewer than 4,000 reviews. In the second and third columns of each triplet, the estimation sample is restricted to the first half and first fifth of each user's reviews, respectively.

set contains 1000 restaurants and the standard deviation of the permanent shock has the same magnitude as the transitory shock, $a = 1$. In column 1, we apply the to estimator the full sample; in column 2, the first half of each user's reviews; in column 3, the first fifth. As expected, the estimates suffer attenuation bias, and this selection bias is reduced as we restrict the sample to earlier reviews. In columns 4-6, the permanent shock has half the magnitude of the *iid* shock, $a = 0.5$. For this magnitude, the attenuation bias is immaterial and restricting the sample simply increases the standard errors. In columns 7-9, the choice set contains 11,000 restaurants, and the selection bias is not as severe as in columns 1-3.

D Model fit

This appendix details the model-fit results discussed in Section 4.3.

D.1 Isolation indices

Gentzkow and Shapiro (2011) define an isolation index for racial group g as

$$S_g = \sum_j \frac{v_{gj}}{v_g} \cdot \left(\frac{v_{gj}}{v_j}\right) - \sum_j \frac{v_{\neg gj}}{v_{\neg g}} \cdot \left(\frac{v_{\neg gj}}{v_j}\right).$$

This index measure the extent to which members of group g disproportionately visit venues whose other visitors are also members of group g. The first summation is the visit-weighted average of the share of a venue's reviewers who are members of g, using g members' visits as the weights. The second summation uses reviews by users who do not belong to racial group g. S_g is therefore the difference between the average g exposure of members of g and

Table D.1: Isolation indices for various model specifications

	Data	Model predictions				
		Pooled	Race-specific	Nested 1	Nested 2	Minimum time
Asian	.087	[.035, .064]	[.056, .089]	[.057, .088]	[.057, .089]	[.054, .088]
black	.087	[.029, .073]	[.043, .089]	[.045, .094]	[.043, .093]	[.044, .093]
white/Hispanic	.045	[.014, .045]	[.025, .057]	[.024, .057]	[.026, .059]	[.024, .057]

NOTES: The reported leave-out isolation indices $\hat{S}_g$, as defined in Gentzkow and Shapiro (2011), are the values for the estimation sample and the 90% confidence interval for model-predicted outcomes for generated samples of the same size. The pooled model imposes common coefficients across all three racial groups for the spatial-friction, price, rating, income, and area-dummy coefficients. The race-specific model is Table 3. "Nested 1" and "Nested 2" are columns 1-3 and 4-6 of Table D.3, respectively. "Minimum time" is Table A.10.

the average g exposure of non-members.[67]

The sample analogue of this measure suffers a finite-sample upward bias, so we follow Gentzkow and Shapiro (2011) and compute the sample analogue using leave-out means, $\hat{S}_g$, as defined in Section 4.3. Table D.1 reports the values of $\hat{S}_g$ for the estimation sample data and the 90% confidence intervals for a model with pooled coefficients, our preferred race-specific model (Table 3), two nested-logit specifications (Table D.3), and a minimum-travel-time specification (Table A.10. The pooled model imposes common coefficients across all three racial groups for the spatial-friction, price, rating, income, and area-dummy coefficients. It generates lower values of $\hat{S}_g$ than observed in the data. The race-specific model, nested-logit specifications, and minimum-travel-time specification all perform similarly in matching the data.

D.2 Schelling-style segregation

We define a pairing p to be a set $p = \{j, j'\}$ such that $X^1_{ijl} = X^1_{ij'l}$, $X^2_{ij} = X^2_{ij'}$, $Z^1_j = Z^1_{j'}$, and $Z^2_{ij} = Z^2_{ij'}$. In practice, this means that we are comparing two restaurants with the same cuisine category, price, and Yelp rating that are located in the same census tract. If the two restaurants have identical shares of users of each race, then $gap_p = 0$. If there is zero overlap in the racial composition of the two restaurants, then $gap_p = 1$. Our estimator presumes that $gap_p = 0$. A Schelling-style model in which users' consumption choices depend upon the endogenous racial composition of venues' patrons might predict $gap_p = 1$.

The sample analogue of this gap measure suffers a finite-sample bias: it will typically be greater than zero when we observe a small number of restaurant visitors' races, even if the "true" value of $gap_p = 0$. We therefore compare the observed gap_p to the null hypothesis of a uniform distribution in which individuals are randomly assigned to j or j' within the pair. The randomization generates a finite-sample draw in which v_j and $v_{j'}$ are equal to their observed values and the assignment is drawn from a process with the property that $\Pr(g(i) = g | d^*_{ij} = 1) = \Pr(g(i) = g | d^*_{ij'} = 1)$ for $j, j' \in p$. For example, if there are 20

[67] Note that S_g as defined by Gentzkow and Shapiro (2011) is distinct from the "isolation index" in Massey and Denton (1988), which is simply $\sum_j \frac{v_{gj}}{v_g} \cdot \left(\frac{v_{gj}}{v_j}\right)$.

reviews of j and 40 reviews of j', the 60 reviews are randomly allocated between the two venues so that j is randomly assigned 20 reviews and the remaining 40 are assigned to j'.

In our data, there are 4569 venues in sets of venues that have the same tract-cuisine-price-rating quadruplet. There are 402 venues with 10-40 reviews in sets of venues that have the same tract-cuisine-price-rating quadruplet and have at least two venues with 10-40 reviews. We photo-coded all the users who reviewed restaurants in a sample of 119 pairs of such venues. This yields the results depicted in Figure 6.

D.3 Restaurant fixed effects

We estimate specifications with restaurant fixed effects employing two different procedures. The main text reports the results for an estimation procedure that employs sampling, as described in Section 3.3. In this appendix, we report the results of a procedure due to Taddy (2015). This specification incorporates restaurant fixed effects in a computationally feasible manner by making a structural assumption about the denominator of the logit probability expression. In our setting, the assumption is that the expected utility of a restaurant trip is equal across all users within a racial group. Table D.2 reports estimates. They are generally congruent with the results in Tables 3 and 5. Since we believe that the structural assumption of equal expected utilities is unlikely to hold true in reality, we prefer the estimates reported in Table 5 in the main text.

D.4 Nested logit

Table D.3 reports the estimates obtaining from applying the nested-logit estimator derived in Section C.4 to two nesting schemes: (a) restaurants of the same disaggregated cuisine category, Yelp rating, and area, and (b) restaurants of the same disaggregated cuisine category, price category, and census tract. For the former, the value of the nest-correlation parameter λ lies between 1.00 and 1.12; for the latter, 0.88 and 1.09. These values are near the conditional-logit benchmark of $\lambda = 1$. A likelihood ratio test formally rejects the conditional-logit model in favor of both nested specifications for the sample of Asian users, and in favor of the cuisine-rating-area nesting in the case of white/Hispanic users. However, these nested specifications offer little benefit of improved in-sample isolation fit, as shown in Table D.1 and come at considerably greater computational cost. These specifications take hours or days to estimate, rather than the few minutes required to estimate Table 3.

D.5 Parametric bootstrap

To assess the finite-sample properties of our estimator under the assumed data-generating process, we use the observed covariates and estimated parameters reported in Table 3 to generate 500 samples of observations (equal in size to our estimation data). Estimating our model on these generated samples yields a distribution of estimates that we can compare to the normal distribution associated with our asymptotic standard errors.

Figures D.3 and D.4 depict these for the main specifications reported in Table 3; Figures D.5 and D.6 for the minimum-travel-time specifications in Table A.10. In each figure, the solid red line shows the bootstrapped distribution of estimates and the dashed blue line

Table D.2: Restaurant fixed effects, estimated by Taddy (2015) procedure

	(1) Asian	(2) black	(3) white/Hispanic
Log travel time from home by public transit	-0.96	-0.70	-0.95
	(0.105)	(0.097)	(0.036)
Log travel time from home by car	-1.01	-0.85	-1.14
	(0.083)	(0.108)	(0.049)
Log travel time from work by public transit	-1.10	-1.68	-1.48
	(0.022)	(0.661)	(0.150)
Log travel time from work by car	-1.30	-1.59	-1.55
	(0.131)	(0.364)	(0.107)
Log travel time from commute by public transit	-0.89	-0.99	-1.05
	(0.079)	(0.178)	(0.051)
Log travel time from commute by car	-0.86	-1.47	-1.39
	(0.059)	(0.306)	(0.066)
Euclidean demographic distance between h_i and k_j	-0.42	-1.19	-0.77
	(0.087)	(0.199)	(0.102)
EDD $\times$ SSI	-0.72	-0.97	-0.09
	(0.155)	(0.364)	(0.090)
Percent difference in median incomes $(k_j - h_i)$	-0.23	0.87	-0.28
	(0.049)	(0.142)	(0.050)
Percent absolute difference in median incomes $(h_i - k_j)$	0.48	-0.16	0.83
	(0.124)	(0.336)	(0.073)
2-dollar bin $\times$ home tract median income (10k USD)	0.03	-0.02	0.04
	(0.010)	(0.029)	(0.009)
3-dollar bin $\times$ home tract median income (10k USD)	0.06	0.04	0.08
	(0.013)	(0.048)	(0.012)
4-dollar bin $\times$ home tract median income (10k USD)	0.05	-0.13	0.08
	(0.020)	(0.206)	(0.021)
Yelp rating $\times$ home tract median income (10k USD)	0.01	-0.01	0.01
	(0.005)	(0.014)	(0.003)

NOTES: Standard errors in parentheses.

Table D.3: Nested logit

	(1)	(2)	(3)	(4)	(5)	(6)
	Area-cuisine-rating nests			Tract-cuisine-price nests		
	Asian	black	white/Hisp	Asian	black	white/Hisp
λ	1.091^a	1.007^a	1.119^a	1.086^a	0.876^a	0.983^a
	(0.002)	(0.006)	(0.002)	(0.002)	(0.010)	(0.003)
Log travel time from home by public transit	-1.089^a	-0.939^a	-1.161^a	-1.060^a	-0.924^a	-1.130^a
	(0.015)	(0.022)	(0.010)	(0.014)	(0.021)	(0.010)
Log travel time from home by car	-1.210^a	-1.191^a	-1.420^a	-1.175^a	-1.152^a	-1.353^a
	(0.013)	(0.028)	(0.011)	(0.012)	(0.025)	(0.010)
Log travel time from work by public transit	-1.305^a	-1.861^a	-2.033^a	-1.289^a	-1.689^a	-1.855^a
	(0.021)	(0.197)	(0.056)	(0.021)	(0.157)	(0.046)
Log travel time from work by car	-1.723^a	-1.803^a	-2.126^a	-1.650^a	-1.665^a	-1.929^a
	(0.027)	(0.082)	(0.035)	(0.024)	(0.068)	(0.027)
Log travel time from commute by public transit	-0.981^a	-0.931^a	-1.136^a	-0.955^a	-0.920^a	-1.092^a
	(0.010)	(0.018)	(0.008)	(0.009)	(0.017)	(0.007)
Log travel time from commute by car	-1.077^a	-1.327^a	-1.497^a	-1.041^a	-1.313^a	-1.427^a
	(0.009)	(0.031)	(0.011)	(0.008)	(0.029)	(0.009)
Euclidean demographic distance between h_i and k_j	-1.034^a	-1.845^a	-1.210^a	-0.995^a	-1.839^a	-1.197^a
	(0.016)	(0.049)	(0.022)	(0.016)	(0.049)	(0.021)
Spectral segregation index of k_j	0.154^a	0.075^a	0.037^a	0.146^a	0.078^a	0.045^a
	(0.007)	(0.016)	(0.005)	(0.007)	(0.016)	(0.004)
EDD $\times$ SSI	-0.158^a	-0.172^a	-0.062^a	-0.156^a	-0.165^a	-0.066^a
	(0.016)	(0.042)	(0.014)	(0.015)	(0.041)	(0.013)
Share of tract population that is Asian	1.114^a	0.011	0.385^a	0.988^a	0.054	0.368^a
	(0.016)	(0.061)	(0.024)	(0.016)	(0.060)	(0.022)
Share of tract population that is black	0.222^a	1.087^a	0.159^a	0.258^a	1.034^a	0.130^a
	(0.043)	(0.070)	(0.046)	(0.041)	(0.070)	(0.043)
Share of tract population that is Hispanic	-0.260^a	0.471^a	0.392^a	-0.167^a	0.406^a	0.402^a
	(0.032)	(0.067)	(0.033)	(0.030)	(0.067)	(0.031)
Share of tract population that is other	0.336	3.572^a	0.426	0.008	3.685^a	0.500
	(0.282)	(0.602)	(0.348)	(0.268)	(0.595)	(0.323)
Dummy for 2-dollar bin	0.405^a	0.775^a	0.388^a	0.377^a	0.766^a	0.354^a
	(0.012)	(0.035)	(0.014)	(0.011)	(0.034)	(0.013)
Dummy for 3-dollar bin	0.329^a	-0.090	-0.027	0.323^a	-0.125^b	-0.032
	(0.016)	(0.060)	(0.021)	(0.015)	(0.060)	(0.020)
Dummy for 4-dollar bin	0.267^a	-0.068	-0.339^a	0.262^a	-0.124	-0.355^a
	(0.026)	(0.215)	(0.039)	(0.024)	(0.213)	(0.036)
Yelp rating of restaurant	0.579^a	0.052^b	0.334^a	0.594^a	0.055^b	0.343^a
	(0.008)	(0.024)	(0.010)	(0.008)	(0.023)	(0.010)
African cuisine category	0.384^a	-0.192^b	0.427^a	0.336^a	-0.212^b	0.290^a
	(0.039)	(0.097)	(0.043)	(0.038)	(0.096)	(0.043)
American cuisine category	0.438^a	0.525^a	0.606^a	0.433^a	0.515^a	0.590^a
	(0.007)	(0.021)	(0.008)	(0.007)	(0.021)	(0.008)
Asian cuisine category	0.881^a	0.256^a	0.308^a	0.871^a	0.261^a	0.309^a
	(0.007)	(0.024)	(0.009)	(0.007)	(0.023)	(0.009)
European cuisine category	0.163^a	-0.326^a	0.205^a	0.174^a	-0.311^a	0.239^a
	(0.008)	(0.027)	(0.009)	(0.008)	(0.027)	(0.009)
Indian cuisine category	0.378^a	-0.450^a	-0.031^c	0.332^a	-0.426^a	-0.032^b
	(0.012)	(0.053)	(0.016)	(0.012)	(0.053)	(0.016)
Latin American cuisine category	0.556^a	1.008^a	0.727^a	0.536^a	0.991^a	0.687^a
	(0.009)	(0.024)	(0.010)	(0.009)	(0.024)	(0.010)
Middle Eastern cuisine category	0.313^a	0.107^b	0.248^a	0.297^a	0.072^c	0.199^a
	(0.013)	(0.044)	(0.015)	(0.013)	(0.044)	(0.015)
Vegetarian/vegan cuisine category	0.307^a	-0.004	0.519^a	0.341^a	0.061	0.593^a
	(0.019)	(0.072)	(0.020)	(0.018)	(0.068)	(0.019)
2-dollar bin $\times$ home tract median income	0.035^a	-0.022^a	0.045^a	0.034^a	-0.023^a	0.042^a
	(0.001)	(0.006)	(0.002)	(0.001)	(0.006)	(0.002)
3-dollar bin $\times$ home tract median income	0.077^a	0.078^a	0.086^a	0.075^a	0.075^a	0.081^a
	(0.002)	(0.009)	(0.002)	(0.002)	(0.009)	(0.002)
4-dollar bin $\times$ home tract median income	0.075^a	-0.169^a	0.099^a	0.074^a	-0.167^a	0.095^a
	(0.003)	(0.041)	(0.004)	(0.003)	(0.041)	(0.004)
Yelp rating $\times$ home tract median income	0.012^a	0.008^b	0.018^a	0.011^a	0.008^c	0.016^a
	(0.001)	(0.004)	(0.001)	(0.001)	(0.004)	(0.001)
Percent absolute difference in median incomes ($h_i - k_j$)	0.131^a	0.623^a	0.766^a	0.206^a	0.605^a	0.706^a
	(0.041)	(0.150)	(0.052)	(0.039)	(0.149)	(0.049)
Percent difference in median incomes ($k_j - h_i$)	-0.061^a	0.854^a	-0.102^a	-0.051^a	0.834^a	-0.101^a
	(0.007)	(0.022)	(0.009)	(0.006)	(0.022)	(0.009)
Log median household income in k_j	-0.104^a	-0.362^a	-0.638^a	-0.155^a	-0.371^a	-0.619^a
	(0.036)	(0.131)	(0.046)	(0.034)	(0.130)	(0.043)
Average annual robberies per resident in k_j	-3.725^a	2.441^a	-4.261^a	-3.381^a	2.558^a	-3.740^a
	(0.094)	(0.211)	(0.138)	(0.087)	(0.210)	(0.126)
χ^2 test p-value	0.00	0.85	0.00	0.00	0.02	0.37
Number of trips	6447	1079	6936	6447	1079	6936

NOTES: Each column reports an estimated nested-logit model of individuals' decisions to visit a Yelp venue. Standard errors in parentheses. Statistical significance denoted by a (1%), b (5%), c (10%). Unreported controls are 28 area dummies. The χ^2 test is a likelihood-ratio test for each specification relative to its corresponding column in Table 3.

depicts the asymptotic distribution. Figures D.3 and D.5 depict the distributions for the coefficients on our key social-friction covariates and two restaurants characteristics. For these coefficients, our estimator exhibits finite-sample behavior under the assumed data-generating process that is very close to the asymptotic distribution. For the spatial frictions depicted in Figure D.4, our estimator occasionally generates extreme outlying negative coefficients, which we omit from the plots but are evident from the missing mass in the bootstrapped density. Figure D.6 shows that these outlying estimates do not arise if we assume that users visit restaurants using the minimum-time origin-mode pair available to them rather than optimizing over six the origin-mode pairs. This suggests that the finite-sample behavior of the estimated spatial-frictions coefficients is attributable to the fact that we infer these six spatial-friction parameters exclusively from restaurant-reviewing outcomes $d^*_{ij} = \sum_l d^*_{ijl}$, without actually observing the origin-mode-level outcomes d^*_{ijl}.

Figure D.3: Parametric bootstrap: Social frictions and restaurant characteristics

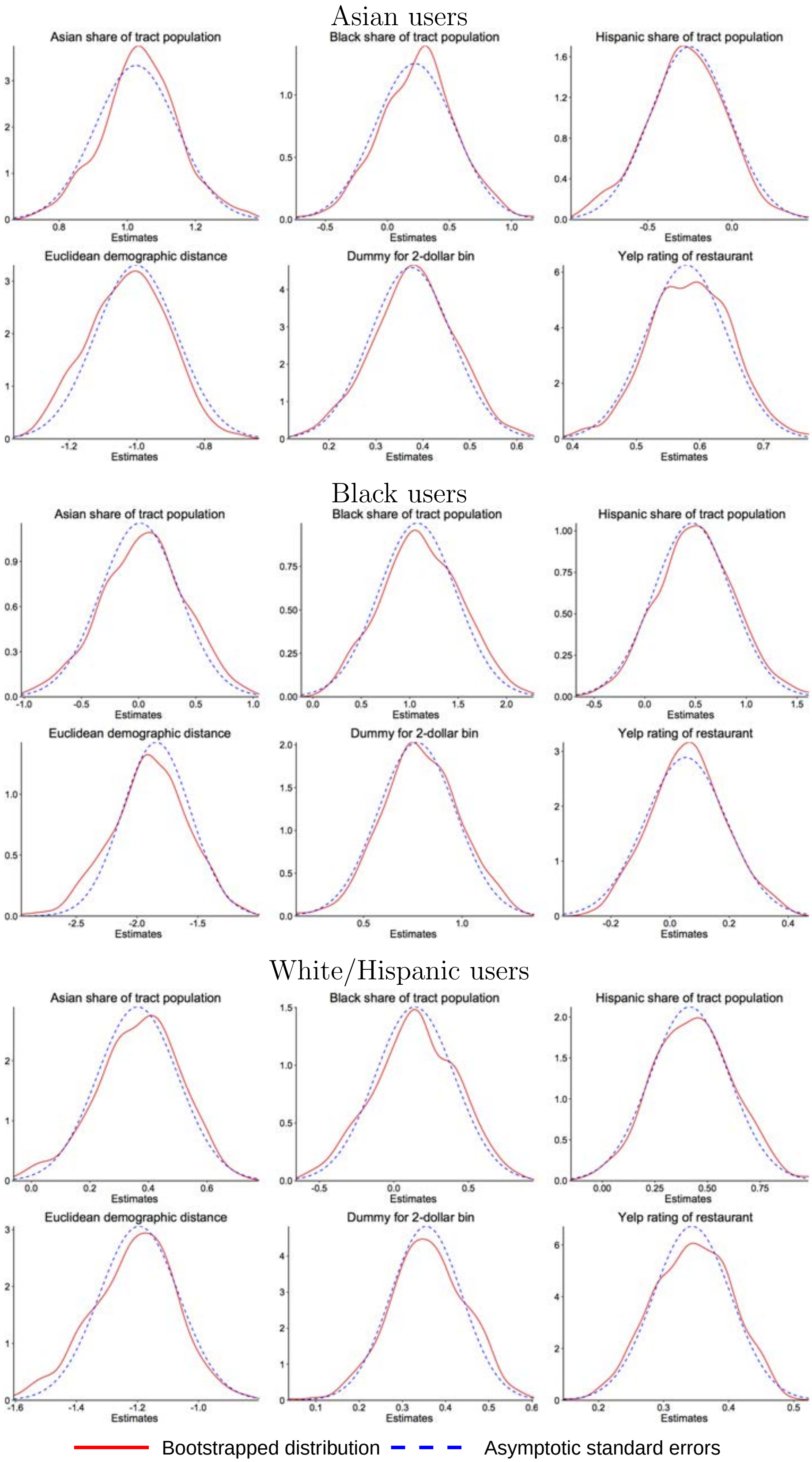


Figure D.4: Parametric bootstrap: Spatial frictions

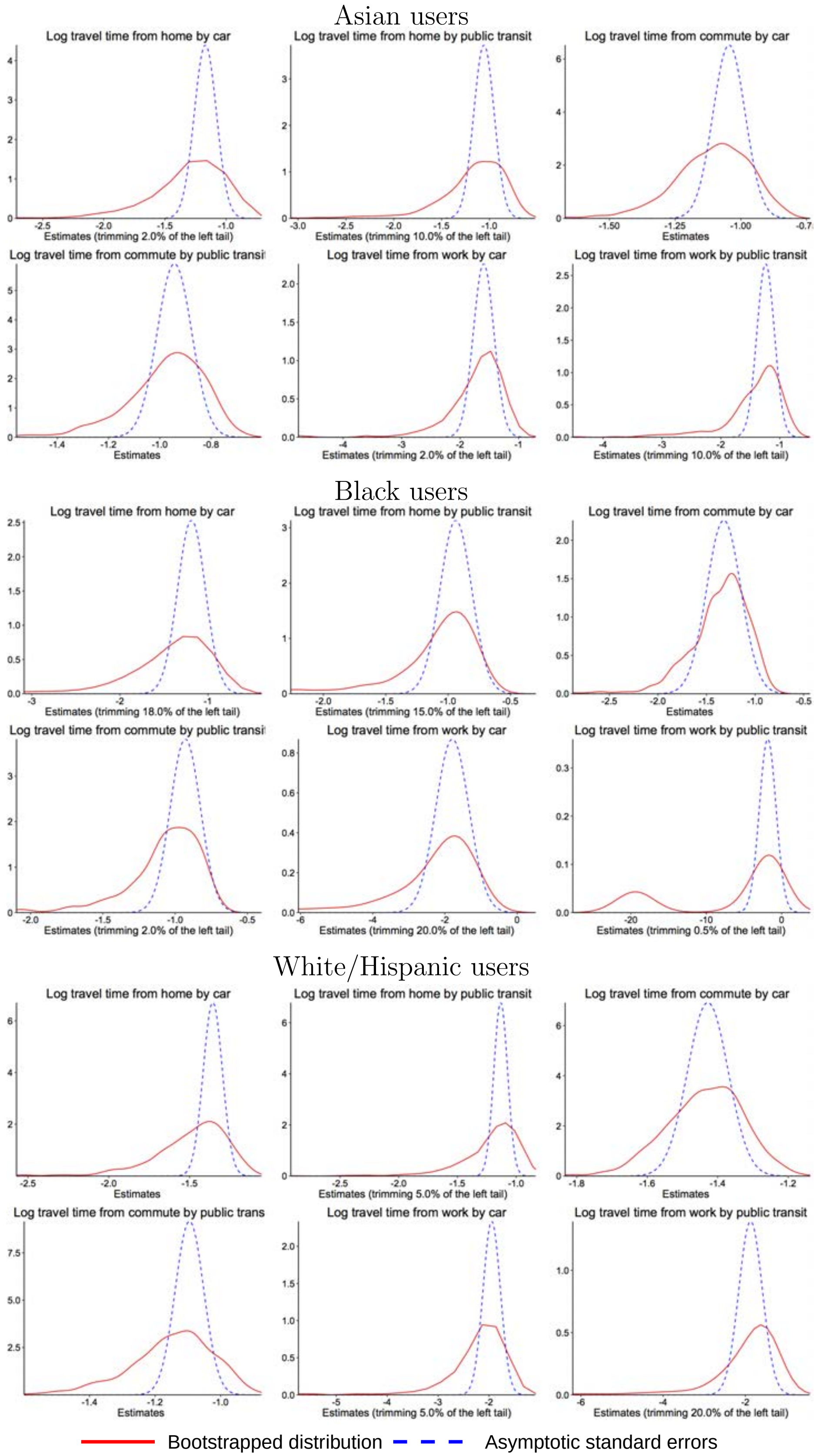


Figure D.5: Parametric bootstrap: Social frictions and restaurant characteristics in minimum-time specification

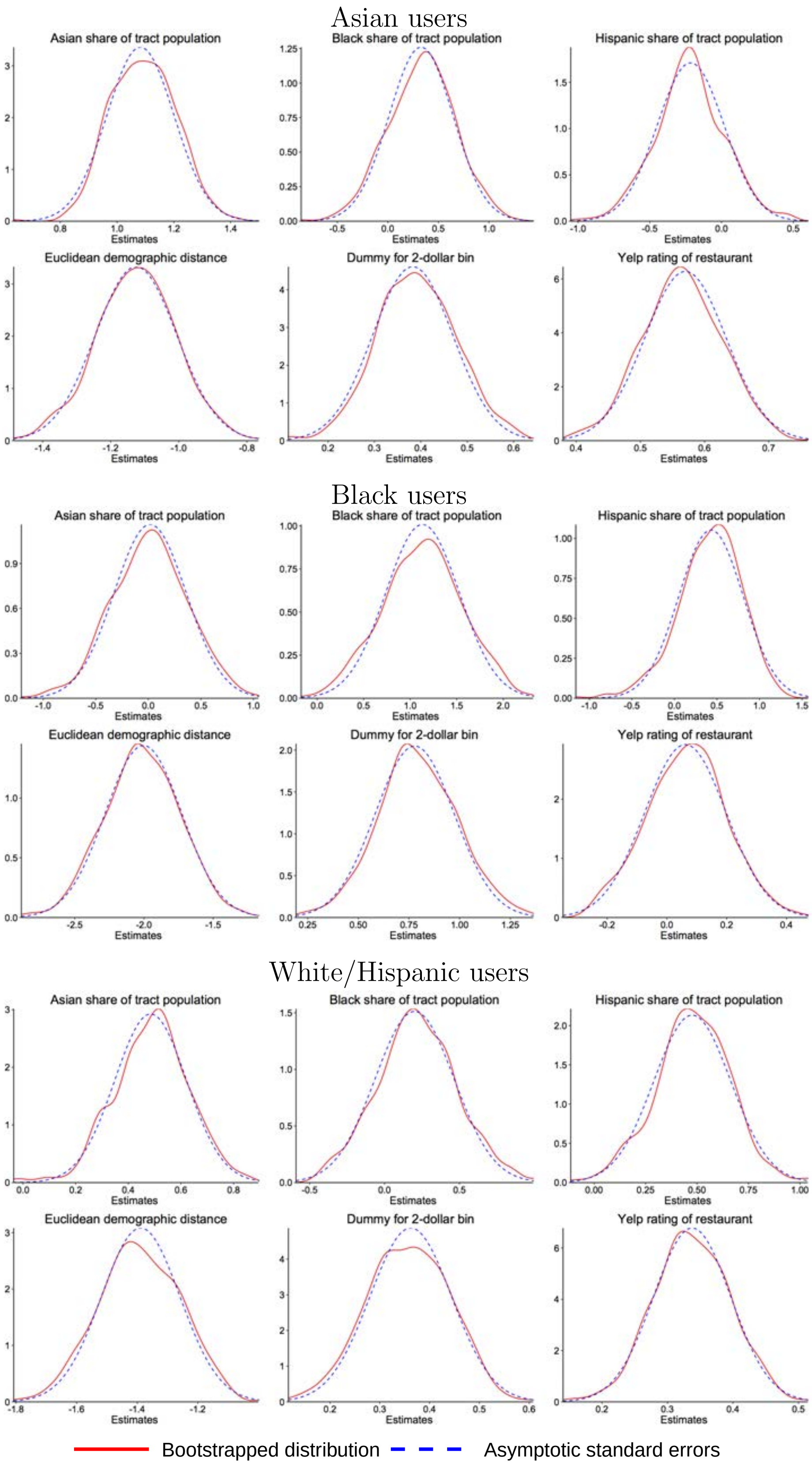
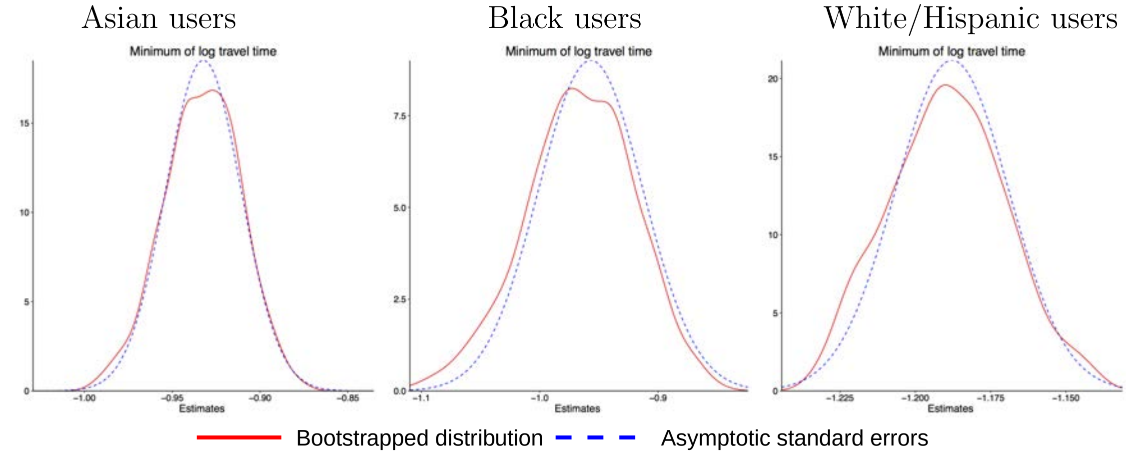


Figure D.6: Parametric bootstrap: Spatial frictions in minimum-time specification



E Consumption segregation estimates and counterfactuals

E.1 Measuring consumption dissimilarity

Here we show how to use the demand model in Section 3, the estimates reported in Section 4, and data on the joint distribution of home census tracts, work census tracts and race to compute dissimilarity indices.[68] Using the definition of conditional and marginal probabilities, we can write

$$P(d_{ij} = 1 | r_i = r) = \frac{P(d_{ij} = 1, r_i = r)}{P(r_i = r)} = \frac{P(d_{ij} = 1, r_i = r)}{\sum_{j \in J} P(d_{ij} = 1, r_i = r)}. \tag{E.1}$$

The estimates reported in Section 4 express the probability that an individual i visits a restaurant j as a function of the home and work location and race of individual i. Therefore, in order to exploit these estimates, we need to rewrite $P(d_{ij} = 1, r_i = r)$ as a function of $P(d_{ij} = 1 | h_i = h, w_i = w, r_i = r)$, for every possible h and w, where h_i indicates the home census tract of individual i and w_i indicates her home tract. We do so implementing the following steps.

Using the definition of a marginal probability distribution, we obtain

$$P(d_{ij} = 1, r_i = r) = \sum_h \sum_w P(d_{ij} = 1, r_i = r, h_i = h, w_i = w), \tag{E.2}$$

where $\sum_h$ denotes a sum over all possible home census tracts and $\sum_w$ denotes a sum over all possible work census tracts. Finally, using the relationship between joint and conditional

[68] We use the Census of Population to obtain information on the share of residents living in each census tract that belong to each of five races or ethnicities: Asian, black, Hispanics, whites, and others. Using commuting data from LEHD Origin-Destination Employment Statistics (LODES), we identify the five most common NYC workplace tracts associated with a given NYC residential tract. Assuming that the share of commutes to each destination tract for a given home tract does not vary across ethnicities/races (since the LODES data does not identify this demographic information), we recover the joint distribution of home and work census tracts by race.

probability distributions, we can write

$$\begin{aligned}
&P(d_{ij}=1,r_i=r,h_i=h,w_i=w)= \\
&\quad =P(d_{ij}=1|r_i=r,h_i=h,w_i=w)P(r_i=r,h_i=h,w_i=w) \\
&\quad =P(d_{ij}=1|r_i=r,h_i=h,w_i=w)P(r_i=r|h_i=h,w_i=w)P(h_i=h,w_i=w) \\
&\quad =P(d_{ij}=1|r_i=r,h_i=h,w_i=w)P(r_i=r|h_i=h,w_i=w)P(w_i=w|h_i=h)P(h_i=h).
\end{aligned}$$

Finally, assuming that all individuals living in the same census tract h have the same probability of commuting to any other census tract w independently of their race, we can conclude that

$$\begin{aligned}
&P(d_{ij}=1,r_i=r,h_i=h,w_i=w)= \\
&\quad =P(d_{ij}=1|r_i=r,h_i=h,w_i=w)P(r_i=r|h_i=h)P(w_i=w|h_i=h)P(h_i=h). \quad \text{(E.3)}
\end{aligned}$$

Using the demand model in Section 3 and, specifically, the functional-form assumption in equation (5), we can write the probability $P(d_{ij}=1|r_i=r,h_i=h,w_i=w)$ as a function of the parameter estimates presented in Section 4. The probabilities $P(r_i=r|h_i=h)$, $P(w_i=w|h_i=h)$ and $P(h_i=h)$ may all be computed using data from the US Census Bureau. Specifically, $P(r_i=r|h_i=h)$ denotes the fraction of residents in census tract h that belong to race or ethnicity r; $P(w_i=w|h_i=h)$ denotes the fraction of residents in census tract h that commute to census tract w; and $P(h_i=h)$ is simply the fraction of the overall population that lives in census tract h.[69] Combining equations (E.1), (E.2) and (E.3) we can write

$$P(d_{ij}=1|r_i=r)=$$
$$\frac{\sum_h\sum_w P(d_{ij}=1|r_i=r,h_i=h,w_i=w)P(r_i=r|h_i=h)P(w_i=w|h_i=h)P(h_i=h)}{\sum_{j\in J}\sum_h\sum_w P(d_{ij}=1|r_i=r,h_i=h,w_i=w)P(r_i=r|h_i=h)P(w_i=w|h_i=h)P(h_i=h)}.$$

E.2 Second Avenue Subway counterfactual

This section details how we compute transit times for the counterfactual scenario in which the Second Avenue Subway is available as a means of public transit. In short, we treat the NYC subway system as a graph and the Second Ave Subway expansion as the addition of new nodes and edges to that graph. We use Dijkstra's algorithm to compute the fastest routes between nodes of the graph with and without the subway expansion. We then modify the current Google Maps transit times by the speed improvement attributable to the addition of the Second Avenue Subway.

Figure E.7 depicts the Second Avenue Subway addition to the NYC subway system, which extends the existing Q line and introduces a new T line. The first phase, running from 96th Street to 57th Street, opened in early 2017, six years after the end of our estimation sample. Additional phases (some not yet funded) plan to extend Q line farther north and introduce many new T line stations along Second Avenue.

[69] Our assumption that $P(w_i=w|h_i=h)$ is independent of the ethnicity/race of i is necessitated by data constraints. Our results are robust in the sense that any downward bias in estimated consumption dissimilarity due to this assumption is very small. Appendix Table A.15 reports consumption dissimilarity indices under the constraint that all consumption trips originate at individuals' residential locations. The resulting dissimilarity indices are similar and show a similar contribution of spatial frictions to that segregation.

Figure E.7: Second Avenue Subway expansion

NOTES: This figure shows the planned Second Avenue Subway expansion. Source: Wikimedia.

We compute the change in transit times implied by this entire expansion. We use GTFS data from transitfeeds.com describing the current system of subway stations and average transit times between stations connected by subway lines. We introduce the new subway stations depicted in Figure E.7 and assume that the transit times between them equal the times between similar stations on A line on the west side of Manhattan. We use Dijkstra's algorithm to compute the fastest path between any two stations in the network, for both the existing subway network and the network enlarged by the Second Avenue expansion. We compute transit between pairs of census tracts under both scenarios by assigning the two nearest subway stations to each census tract (based on tract centroids) and assuming a walking speed of 5 kilometers per hour. While these computations abstract from the NYC bus system and do not account for congestion, we find that the computed transit times between tracts for the current network align well with the transit times from Google Maps that we employ in estimation. We therefore employ the difference in transit times to construct the counterfactual transit times.

The census tracts with the largest predicted improvements in average transit times are those in Manhattan along Second Avenue that are directly served by the new subway stops. However, there are also substantial gains for census tracts in Queens that are near the F and R lines and census tracts in Brooklyn that are near the B, F, and Q lines. These gains reflect improved connections to many Manhattan destinations due to the denser subway network and new potential transfers between lines. We add these transit-time improvements to the current Google Maps transit times and recompute predicted visits for all tracts in order to produce column 2 of Table 8.

E.3 Gentrification exercise

This section details the computation of the welfare losses reported in the Harlem gentrification exercise in Section 6 and reports similar results for a Bedford-Stuyvesant gentrification exercise.

E.3.1 Procedure

We model gentrification as a process by which X_{ij} and Z_{ij} becomes X'_{ij} and Z'_{ij} for a gentrifying set of restaurants $\mathcal{J}^G$. For j not in $\mathcal{J}^G$, $X_{ij} = X'_{ij}$ and $Z_{ij} = Z'_{ij}$. Starting from the expression for expected utility in our demand system (see Train 2009, Ch. 3),

$$
\begin{aligned}
U_i &= \ln\left(\sum_j \sum_l \exp\left(\gamma X_{ijl} + \beta Z_{ij}\right)\right) + c \\
U'_i &= \ln\left(\sum_j \sum_l \exp\left(\gamma X'_{ijl} + \beta Z'_{ij}\right)\right) + c
\end{aligned}
$$

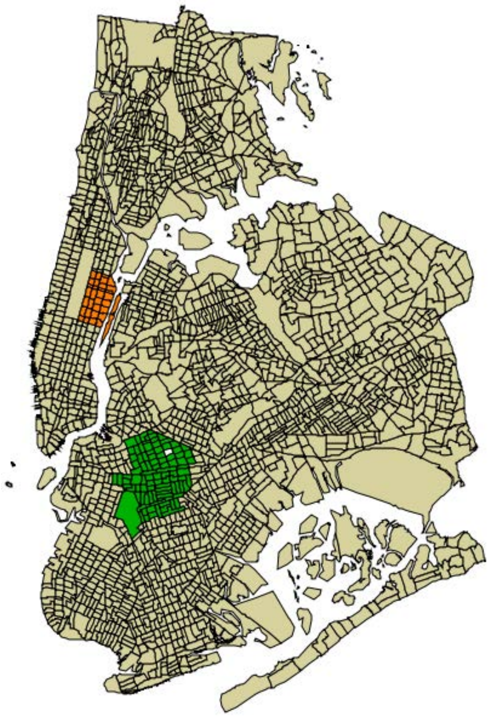
$$
\begin{aligned}
U'_i - U_i &= \ln\left(1 + \frac{\sum_{j\in\mathcal{J}^G}\sum_l \Delta \exp\left(\beta X_{ijl}\right)}{\sum_j \sum_l \exp\left(\beta X_{ijl}\right)}\right) \\
&\approx \frac{\sum_{j\in\mathcal{J}^G}\sum_l \Delta \exp\left(\beta X_{ijl}\right)}{\sum_j \sum_l \exp\left(\beta X_{ijl}\right)} \\
&= \sum_{j\in\mathcal{J}^G}\sum_l P_{ijl}\left[\exp(\beta \Delta X_{ijl}) - 1\right] \\
&\approx \left[\sum_{j\in\mathcal{J}^G} P_{ij}\right] \times \left[\exp(\gamma_g^2 \Delta \bar{X}_{ij}^2 + \beta_g^1 \Delta \bar{Z}_j^1 + \beta_g^2 \Delta \bar{Z}_{ij}^2) - 1\right],
\end{aligned}
$$

where c is Euler's constant. The first approximation employs the fact that $\ln(1+x)$ is approximately x when x is near zero. The second approximation employs the fact that we are randomly assigning the characteristics of Upper East Side tracts and venues to tracts and venues surrounding the incumbent Harlem tract, so there should be no correlation between initial P_{ij} and changes in venue and tract characteristics.

E.3.2 Bedford-Stuyvesant exercise

In this section, we report a gentrification exercise for an area of Brooklyn, akin to the gentrification exercise reported in Section 6. We modify the restaurant and tract characteristics of venues surrounding a low-income, majority-black tract in the Bedford-Stuyvesant neighborhood of Brooklyn, as depicted by the white polygon in Figure E.8. We compute the change in black residents' expected utility if the surrounding census tracts containing 221 Yelp restaurants (green polygons) were to exhibit the residential and restaurants characteristics of high-income, majority-white census tracts of the Upper East Side (orange polygons).

Figure E.8: Bedford-Stuyvesant gentrification exercise



Change in	Mean	Std. Dev.
Share Asian	0.041	0.054
Share black	-0.477	0.249
Share Hispanic	-0.055	0.106
Share white	0.502	0.267
Robberies per resident	-0.004	0.003
Spectral segregation index	-0.3	0.306
Yelp rating	-0.072	1.03
Price ($ to $$$$)	0.692	0.866
Median household income	68.121	51.652
Euclidean demographic distance	1.376	1.229
N	221	

NOTES: We compute the change in black residents' expected utility in the white polygon if the surrounding green tracts were to exhibit the characteristics of the orange tracts. The table reports the changes in these characteristics.

Table E.4 summarizes the decomposition of the resulting welfare loss, akin to Table 9 in the main text. Restaurants in the gentrifying area account for 23.7% of predicted visits by incumbent residents prior to gentrification. Again, the welfare loss we compute is attributable to increases in social frictions associated with the shift of the surrounding tracts from mostly black residents to mostly white residents. This partially offset by the increase in neighborhood incomes. The changes in restaurants' prices, ratings, and cuisines are immaterial. Thus, the results are very similar to those reported for the Harlem gentrification exercise in Section 6.

Table E.4: Welfare losses due to gentrification of surrounding Bedford-Stuyvesant tracts

		Change in value of characteristics $(\gamma\Delta\bar{X}_i, \beta\Delta\bar{Z}_i)$		
Welfare change	Initial visit share	Social frictions	Restaurant traits	Other traits
-2.7%	.203	-3.13	-.069	.585

NOTES: Welfare loss is $100 \times \frac{U_i' - U_i}{U_i}$. Initial visit share is $\sum_{j \in \mathcal{J}^G} P_{ij}$. Social frictions include EDD, SSI, EDD×SSI, and racial and ethnic population shares of k_j. "Restaurant traits" include price, rating, cuisine category, and price and rating interacted with median household income. "Other traits" include destination income, differences in incomes, and robberies per resident.