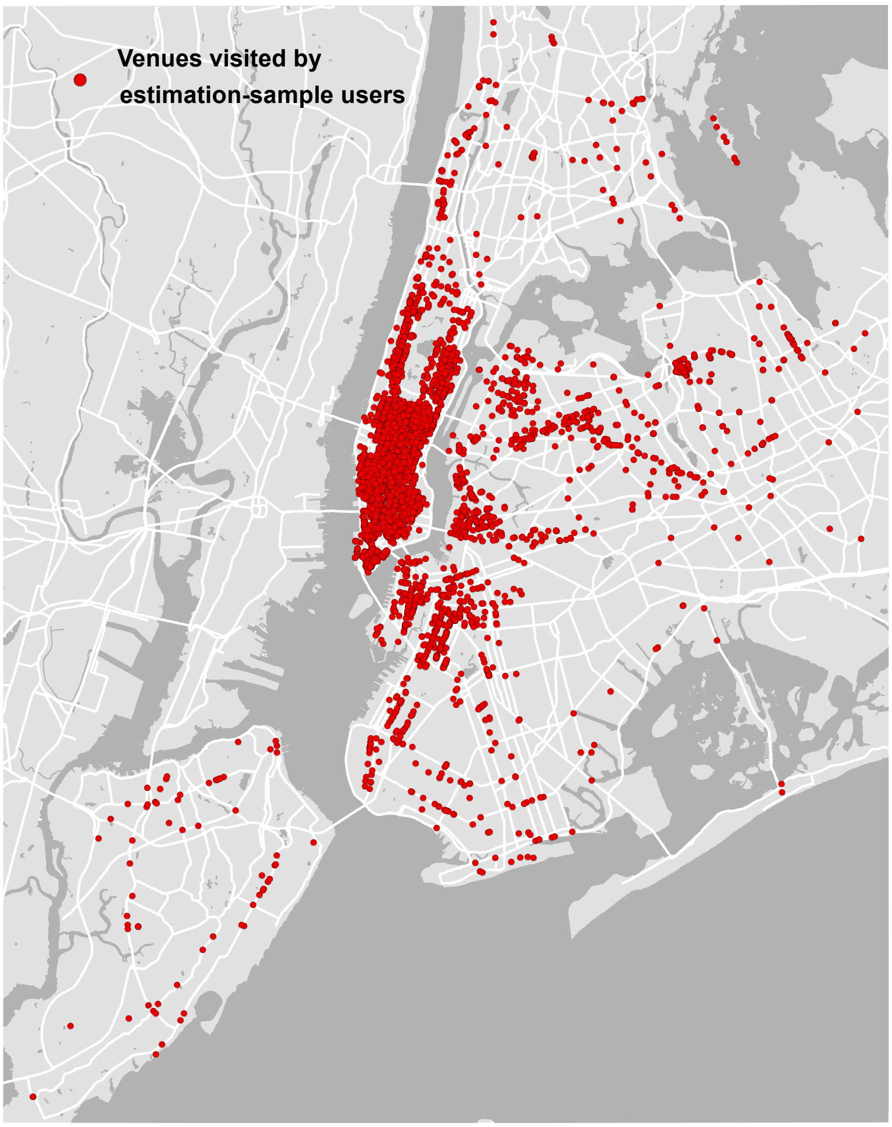
Online Appendix

Davis, Dingel, Monras, Morales - How Segregated is Urban Consumption?

March 2018

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.

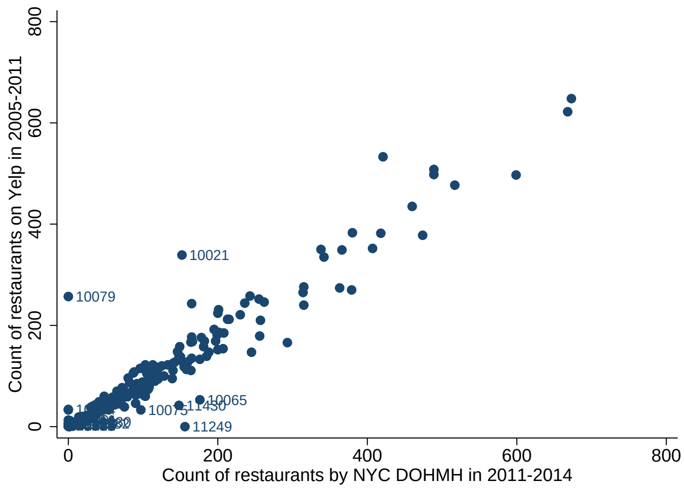
Table A.1: Venue review summary statistics

Restaurant characteristic	Share of all NYC Yelp reviews	Share of estimation-sample reviews: all races	Asian	black	white/Hispanic
Price of $	.227	.242	.233	.335	.240
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	.009	.008	.006	.023	.008
Rating of 2.5 stars	.036	.037	.032	.069	.035
Rating of 3 stars	.140	.144	.125	.204	.151
Rating of 3.5 stars	.360	.370	.358	.343	.379
Rating of 4 stars	.394	.386	.415	.308	.380
Rating of 4.5 stars	.056	.050	.060	.042	.044
Rating of 5 stars	.002	.001	.001	.001	.002
Cuisine: American	.332	.336	.282	.405	.367
Cuisine: Asian	.248	.257	.341	.177	.200
Cuisine: European	.181	.167	.165	.092	.180
Cuisine: Latin American	.090	.097	.072	.187	.106
Cuisine: No Category	.076	.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	.754	.803	.858	.582	.779
Located in Brooklyn	.170	.123	.064	.361	.147
Located in Queens	.065	.066	.074	.044	.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	.057	.068
Located in plurality black	.017	.015	.006	.089	.011
Located in plurality Hispanic	.046	.043	.030	.074	.052
Located in plurality white	.853	.854	.843	.780	.869

Within-reviewer across-review dispersion	Mean for: all races	Asian	black	white/Hispanic
Distance (km), non-located reviewers	4.19	4.27	5.11	4.06
Distance (km), estimation sample	4.18	4.07	5.13	4.21
EDD, non-located reviewers	.207	.221	.278	.191
EDD, estimation sample	.215	.225	.274	.206

NOTES: The upper panel summarizes the distribution of reviews across different venue characteristics for all Yelp reviews of NYC restaurants (column 1), our estimation sample (column 2), and by race within our estimation sample (columns 3–5). The lower panel summarizes the within-reviewer across-review dispersion in physical distance and Euclidean demographic distance for both the estimation sample and non-located reviewers for whom we inferred racial demographics.

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



NOTES: This plot compares the number of food establishments in each ZIP code reported in New York City Department of Health & Mental Hygiene inspections data for 2011-2014 to the number of restaurants listed in our Yelp data covering 2005-2011. See Appendix B.1 for notes on outliers.

Table A.2: NYC census tract summary statistics

Variable	Mean	Std. Dev.
Tract characteristics		
Population	3866	2114
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		
Percentage absolute difference in median household income	0.506	0.355
Percentage 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: Spatial frictions with home, work, and commuting-path origins

	(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 the decision 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.

Table A.4: Six-origin-mode specifications: Asian reviewers (part 1)

	(1) Main spec	(2) Choice 50	(3) Choice 100	(4) Half	(5) Fifth	(6) Droploca
Log travel time from home by public transit	-1.06^a	-1.04^a	-1.03^a	-1.07^a	-1.20^a	-.982^a
	(.107)	(.098)	(.096)	(.158)	(.291)	(.103)
Log travel time from home by car	-1.17^a	-1.15^a	-1.15^a	-1.25^a	-1.35^a	-1.08^a
	(.091)	(.084)	(.082)	(.152)	(.250)	(.088)
Log travel time from work by public transit	-1.24^a	-1.25^a	-1.26^a	-1.29^a	-1.41^a	-1.17^a
	(.149)	(.148)	(.149)	(.248)	(.413)	(.147)
Log travel time from work by car	-1.60^a	-1.61^a	-1.61^a	-1.58^a	-1.52^a	-1.48^a
	(.176)	(.174)	(.172)	(.261)	(.294)	(.168)
Log travel time from commute by public transit	-.943^a	-.938^a	-.931^a	-.919^a	-.995^a	-.876^a
	(.067)	(.064)	(.062)	(.094)	(.146)	(.066)
Log travel time from commute by car	-1.04^a	-1.05^a	-1.04^a	-1.05^a	-1.11^a	-.984^a
	(.061)	(.059)	(.058)	(.089)	(.130)	(.062)
Euclidean demographic distance between h_i and k_j	-1.00^a	-.957^a	-.924^a	-1.13^a	-.996^a	-.928^a
	(.121)	(.115)	(.113)	(.172)	(.279)	(.123)
Spectral segregation index of k_j	.150^a	.138^a	.138^a	.118^c	.195^c	.147^a
	(.051)	(.048)	(.048)	(.063)	(.102)	(.052)
EDD × SSI	-.149	-.137	-.138	-.174	-.203	-.151
	(.117)	(.108)	(.108)	(.160)	(.242)	(.118)
Share of tract population that is Asian	1.03^a	.973^a	.993^a	1.07^a	1.09^a	.989^a
	(.120)	(.114)	(.112)	(.168)	(.273)	(.122)
Share of tract population that is black	.220	.184	.237	.553	.728	.193
	(.319)	(.310)	(.306)	(.441)	(.730)	(.328)
Share of tract population that is Hispanic	-.251	-.141	-.112	-1.22^a	-1.05^c	-.276
	(.235)	(.228)	(.225)	(.365)	(.608)	(.238)
Share of tract population that is other	.059	.636	.729	-.009	5.33	-.289
	(2.07)	(1.98)	(1.95)	(2.90)	(4.62)	(2.15)
Dummy for 2-dollar bin	.375^a	.354^a	.338^a	.348^a	.406^b	.393^a
	(.087)	(.084)	(.082)	(.122)	(.199)	(.089)
Dummy for 3-dollar bin	.287^b	.235^b	.229^b	.179	.072	.310^a
	(.116)	(.112)	(.110)	(.165)	(.269)	(.118)
Dummy for 4-dollar bin	.220	.263	.254	.035	-.015	.248
	(.188)	(.180)	(.175)	(.265)	(.399)	(.190)
Yelp rating of restaurant	.579^a	.561^a	.536^a	.488^a	.415^a	.586^a
	(.064)	(.060)	(.059)	(.090)	(.146)	(.065)
African cuisine category	.280	.129	.358	.384	.453	.295
	(.299)	(.289)	(.286)	(.448)	(.757)	(.299)
American cuisine category	.432^a	.411^a	.428^a	.403^a	.430^a	.444^a
	(.054)	(.053)	(.052)	(.077)	(.127)	(.056)
Asian cuisine category	.886^a	.871^a	.875^a	.917^a	.996^a	.888^a
	(.054)	(.052)	(.052)	(.077)	(.126)	(.056)
European cuisine category	.195^a	.195^a	.193^a	.189^b	.116	.200^a
	(.059)	(.057)	(.057)	(.085)	(.140)	(.060)
Indian cuisine category	.370^a	.338^a	.328^a	.370^a	.358^c	.380^a
	(.091)	(.088)	(.087)	(.129)	(.213)	(.093)
Latin American cuisine category	.517^a	.495^a	.505^a	.487^a	.575^a	.497^a
	(.070)	(.068)	(.067)	(.101)	(.162)	(.072)
Middle Eastern cuisine category	.280^a	.304^a	.316^a	.352^b	.220	.293^a
	(.101)	(.097)	(.096)	(.140)	(.239)	(.103)
Vegetarian/vegan cuisine category	.392^a	.329^b	.389^a	.728^a	.833^a	.387^a
	(.138)	(.131)	(.129)	(.171)	(.267)	(.140)
2-dollar bin × home tract median income	.034^a	.034^a	.036^a	.029^c	.027	.036^a
	(.011)	(.010)	(.010)	(.015)	(.025)	(.011)
3-dollar bin × home tract median income	.075^a	.074^a	.076^a	.078^a	.103^a	.078^a
	(.014)	(.013)	(.013)	(.019)	(.031)	(.014)
4-dollar bin × home tract median income	.074^a	.064^a	.063^a	.087^a	.124^a	.076^a
	(.022)	(.021)	(.020)	(.031)	(.046)	(.022)
Yelp rating × home tract median income	.011	.010	.012^c	.017	.028	.011
	(.008)	(.007)	(.007)	(.011)	(.018)	(.008)
Percent absolute difference in median incomes ($h_i - k_j$)	-.062	-.061	-.063	-.061	-.081	-.053
	(.050)	(.048)	(.047)	(.071)	(.116)	(.051)
Percent difference in median incomes ($k_j - h_i$)	.114	.099	.060	.113	-.093	.055
	(.305)	(.292)	(.289)	(.421)	(.686)	(.310)
Log median household income in k_j	-.109	-.095	-.057	-.156	-.009	-.065
	(.267)	(.256)	(.254)	(.367)	(.603)	(.272)
Average annual robberies per resident in k_j	-3.41^a	-3.43^a	-3.31^a	-4.44^a	-7.81^a	-3.66^a
	(.676)	(.657)	(.650)	(.977)	(1.97)	(.714)
Number of trips	6447	6447	6447	3205	1258	6181

NOTES: Each column reports an estimated conditional-logit model of the decision to visit a Yelp venue. Column 1 shows specification from main text. Columns 2 and 3 show specifications in which randomly generated choice sets have 50 and 100 restaurants, respectively. Columns 4 and 5 show specifications in which observations are limited to the first half and first fifth of NYC restaurant reviews posted by each reviewer, respectively. Column 6 drops observations that are restaurant reviews containing locational information used to identify the residence or workplace of the reviewer. Standard errors in parentheses. Statistical significance denoted by a (1%), b (5%), c (10%). Unreported controls are 28 area dummies.

Table A.4: Six-origin-mode specifications: Asian reviewers (continued)

	(1) Main spec	(7) Locainfo1	(8) Locainfo2	(9) Late adopt	(10) Cuisine	(11) Cars	(12) Chains
Log travel time from home by public transit	-1.06^a	-.982^a	-1.11^a	-1.18^a	-1.06^a	-1.03^a	-1.08^a
	(.107)	(.154)	(.150)	(.161)	(.108)	(.113)	(.121)
Log travel time from home by car	-1.17^a	-1.19^a	-1.17^a	-1.25^a	-1.18^a	-1.15^a	-1.15^a
	(.091)	(.150)	(.119)	(.123)	(.093)	(.098)	(.097)
Log travel time from work by public transit	-1.24^a	-1.05^a	-1.44^a	-1.34^a	-1.24^a	-1.20^a	-1.35^a
	(.149)	(.164)	(.272)	(.206)	(.146)	(.150)	(.198)
Log travel time from work by car	-1.60^a	-1.55^a	-1.64^a	-1.71^a	-1.61^a	-1.53^a	-1.70^a
	(.176)	(.265)	(.236)	(.238)	(.176)	(.172)	(.216)
Log travel time from commute by public transit	-.943^a	-.838^a	-1.05^a	-1.07^a	-.943^a	-.920^a	-1.10^a
	(.067)	(.089)	(.104)	(.103)	(.067)	(.071)	(.103)
Log travel time from commute by car	-1.04^a	-1.04^a	-1.09^a	-1.13^a	-1.06^a	-1.02^a	-1.06^a
	(.061)	(.100)	(.085)	(.082)	(.063)	(.065)	(.069)
Euclidean demographic distance between h_i and k_j	-1.00^a	-.869^a	-.925^a	-.972^a	-1.03^a	-.965^a	-1.07^a
	(.121)	(.170)	(.184)	(.196)	(.122)	(.122)	(.134)
Spectral segregation index of k_j	.150^a	.100	.612^a	.442^a	.153^a	.142^a	.168^a
	(.051)	(.061)	(.131)	(.128)	(.051)	(.051)	(.055)
EDD × SSI	-.149	-.111	-.818^a	-.407	-.185	-.203^c	-.157
	(.117)	(.144)	(.245)	(.260)	(.123)	(.122)	(.131)
Share of tract population that is Asian	1.03^a	.800^a	1.20^a	.981^a	.890^a	.954^a	1.05^a
	(.120)	(.159)	(.188)	(.185)	(.125)	(.121)	(.136)
Share of tract population that is black	.220	.896^b	-.761	-.163	.091	-.073	.273
	(.319)	(.408)	(.523)	(.555)	(.327)	(.325)	(.364)
Share of tract population that is Hispanic	-.251	-1.03^a	.955^a	.159	-.320	-.416^c	.029
	(.235)	(.308)	(.370)	(.357)	(.240)	(.238)	(.259)
Share of tract population that is other	.059	-1.23	1.61	-.022	.425	.281	-4.89^b
	(2.07)	(2.78)	(3.19)	(2.99)	(2.07)	(2.07)	(2.48)
Dummy for 2-dollar bin	.375^a	.337^a	.621^a	.394^a	.244^a	.367^a	-.395^a
	(.087)	(.111)	(.155)	(.122)	(.089)	(.087)	(.096)
Dummy for 3-dollar bin	.287^b	.124	.826^a	.312^c	.124	.272^b	-.621^a
	(.116)	(.154)	(.197)	(.166)	(.120)	(.116)	(.127)
Dummy for 4-dollar bin	.220	.324	.326	.186	.143	.216	-.694^a
	(.188)	(.242)	(.321)	(.274)	(.193)	(.188)	(.206)
Yelp rating of restaurant	.579^a	.371^a	.920^a	.584^a	.575^a	.576^a	.074
	(.064)	(.082)	(.112)	(.092)	(.065)	(.064)	(.082)
African cuisine category	.280	.065	.442	.087		.287	-.118
	(.299)	(.471)	(.391)	(.530)		(.298)	(.305)
American cuisine category	.432^a	.577^a	.271^a	.477^a		.432^a	-.121^b
	(.054)	(.076)	(.078)	(.087)		(.054)	(.061)
Asian cuisine category	.886^a	.962^a	.811^a	.982^a		.884^a	.251^a
	(.054)	(.077)	(.077)	(.087)		(.054)	(.061)
European cuisine category	.195^a	.253^a	.131	.336^a		.192^a	-.130^b
	(.059)	(.084)	(.084)	(.093)		(.059)	(.065)
Indian cuisine category	.370^a	.411^a	.323^b	.447^a		.370^a	-.015
	(.091)	(.129)	(.130)	(.145)		(.091)	(.098)
Latin American cuisine category	.517^a	.586^a	.455^a	.607^a		.515^a	-.129^c
	(.070)	(.097)	(.102)	(.109)		(.070)	(.077)
Middle Eastern cuisine category	.280^a	.340^b	.221	.349^b		.271^a	-.184^c
	(.101)	(.142)	(.143)	(.156)		(.101)	(.111)
Vegetarian/vegan cuisine category	.392^a	.816^a	-.052	.324		.383^a	-.506^a
	(.138)	(.183)	(.209)	(.225)		(.138)	(.146)
2-dollar bin × home tract median income	.034^a	.032^b	.015	.021	.036^a	.035^a	.044^a
	(.011)	(.016)	(.017)	(.015)	(.011)	(.011)	(.012)
3-dollar bin × home tract median income	.075^a	.079^a	.032	.065^a	.073^a	.077^a	.081^a
	(.014)	(.020)	(.021)	(.019)	(.014)	(.014)	(.015)
4-dollar bin × home tract median income	.074^a	.044	.075^b	.067^b	.074^a	.075^a	.079^a
	(.022)	(.032)	(.034)	(.031)	(.022)	(.022)	(.024)
Yelp rating × home tract median income	.011	.040^a	-.025^b	.015	.009	.011	.013
	(.008)	(.012)	(.012)	(.011)	(.008)	(.008)	(.010)
Percent absolute difference in median incomes ($h_i - k_j$)	-.062	-.031	-.107	-.013	-.055	-.035	-.055
	(.050)	(.067)	(.083)	(.075)	(.051)	(.050)	(.055)
Percent difference in median incomes ($k_j - h_i$)	.114	-.354	.691	.293	.234	.390	.117
	(.305)	(.433)	(.443)	(.435)	(.307)	(.309)	(.334)
Log median household income in k_j	-.109	.327	-.620	-.296	-.231	-.411	-.135
	(.267)	(.381)	(.388)	(.379)	(.270)	(.272)	(.293)
Average annual robberies per resident in k_j	-3.41^a	-3.94^a	-2.79^a	-3.00^a	-2.78^a	-2.66^a	.167
	(.676)	(.921)	(1.00)	(.966)	(.677)	(.670)	(.709)
Establishment belongs to chain							.140
							(.114)
Log number of Yelp reviews							1.01^a
							(.015)
Difference in tracts' private vehicle ownership						-1.02^a	
						(.129)	
Number of trips	6447	3326	3121	2766	6447	6447	6447

Notes: Each column reports an estimated conditional-logit model of the decision to visit a Yelp venue. Column 1 shows specification from main text. Columns 7 and 8 split the estimation sample into reviewers with residences located using information contained in one or two reviews in column 2 and three or more reviews in column 3. Column 9 restricts the sample to Yelp reviewers with later-than-median dates of first review written. Column 10 introduces dummies for 39 more disaggregated cuisine categories. Column 11 controls for origin-destination differences in vehicle ownership. Column 12 controls for the number of Yelp reviews of each restaurant and its membership in a chain with more than eight NYC locations. Standard errors in parentheses. Statistical significance denoted by a (1%), b (5%), c (10%). Unreported controls are 28 area dummies.

Table A.5: Six-origin-mode specifications: Black reviewers (part 1)

	(1) Main spec	(2) Choice 50	(3) Choice 100	(4) Half	(5) Fifth	(6) Droploca
Log travel time from home by public transit	-.938[a] (.127)	-.953[a] (.123)	-.920[a] (.119)	-1.08[a] (.199)	-1.08[a] (.351)	-.929[a] (.128)
Log travel time from home by car	-1.19[a] (.158)	-1.20[a] (.151)	-1.15[a] (.142)	-1.42[a] (.264)	-1.59[a] (.609)	-1.17[a] (.158)
Log travel time from work by public transit	-1.85[c] (1.11)	-2.02 (1.66)	-1.97 (1.47)	-22.5 (201588.1)	-1.45 (.994)	-1.87 (1.22)
Log travel time from work by car	-1.79[a] (.459)	-1.91[a] (.545)	-1.93[a] (.577)	-2.01[a] (.706)	-1.32[a] (.433)	-1.74[a] (.433)
Log travel time from commute by public transit	-.930[a] (.105)	-.937[a] (.099)	-.917[a] (.099)	-1.01[a] (.130)	-.950[a] (.222)	-.915[a] (.104)
Log travel time from commute by car	-1.32[a] (.177)	-1.31[a] (.157)	-1.28[a] (.156)	-1.42[a] (.208)	-1.44[a] (.392)	-1.29[a] (.171)
Euclidean demographic distance between h_i and k_j	-1.84[a] (.280)	-1.89[a] (.271)	-1.74[a] (.265)	-1.88[a] (.411)	-2.05[a] (.765)	-1.86[a] (.281)
Spectral segregation index of k_j	.075 (.093)	.177[b] (.086)	.125 (.083)	.125 (.158)	.392 (.574)	.074 (.092)
EDD × SSI	-.171 (.239)	-.369 (.256)	-.346 (.271)	-.154 (.373)	-.839 (1.19)	-.160 (.231)
Share of tract population that is Asian	.011 (.345)	-.015 (.333)	.028 (.330)	.354 (.489)	.839 (.833)	-.022 (.348)
Share of tract population that is black	1.08[a] (.399)	.876[b] (.391)	1.02[a] (.383)	1.59[a] (.596)	.966 (.984)	1.15[a] (.404)
Share of tract population that is Hispanic	.467 (.381)	.502 (.359)	.536 (.355)	1.10[b] (.557)	1.70[c] (.915)	.495 (.383)
Share of tract population that is other	3.56 (3.43)	5.61[c] (2.99)	5.53[c] (2.95)	5.49 (5.44)	11.3 (9.67)	2.76 (3.51)
Dummy for 2-dollar bin	.771[a] (.197)	.757[a] (.189)	.727[a] (.186)	.742[a] (.279)	1.09[b] (.473)	.789[a] (.200)
Dummy for 3-dollar bin	-.090 (.341)	.140 (.320)	.102 (.316)	1.14[c] (.581)	1.59[c] (.875)	-.042 (.345)
Dummy for 4-dollar bin	-.074 (1.22)	.092 (1.21)	-.108 (1.15)	.386 (2.59)	-24.0 (2737379.3)	-.060 (1.22)
Yelp rating of restaurant	.053 (.138)	.094 (.126)	.028 (.123)	.165 (.202)	.174 (.348)	.056 (.140)
African cuisine category	-.198 (.553)	-.009 (.534)	.037 (.522)	-.228 (.668)	-.674 (1.15)	-.193 (.553)
American cuisine category	.523[a] (.119)	.596[a] (.114)	.551[a] (.112)	.410[b] (.168)	.245 (.273)	.538[a] (.121)
Asian cuisine category	.255[c] (.134)	.279[b] (.129)	.253[b] (.127)	.190 (.190)	.215 (.302)	.283[b] (.136)
European cuisine category	-.326[b] (.154)	-.239 (.148)	-.283[c] (.147)	-.512[b] (.222)	-.410 (.348)	-.312[b] (.156)
Indian cuisine category	-.451 (.301)	-.562[c] (.292)	-.558[c] (.290)	-.706 (.458)	-.822 (.795)	-.420 (.301)
Latin American cuisine category	1.01[a] (.136)	1.01[a] (.129)	.922[a] (.126)	.836[a] (.193)	.816[a] (.301)	1.03[a] (.137)
Middle Eastern cuisine category	.104 (.251)	.089 (.243)	-.026 (.240)	-.310 (.401)	-.690 (.782)	.147 (.251)
Vegetarian/vegan cuisine category	.001 (.409)	-.204 (.400)	-.339 (.401)	-.399 (.617)	-.517 (1.06)	.035 (.410)
2-dollar bin × home tract median income	-.022 (.032)	-.028 (.031)	-.023 (.030)	.008 (.045)	-.040 (.077)	-.023 (.032)
3-dollar bin × home tract median income	.077 (.053)	.031 (.050)	.035 (.049)	-.175 (.107)	-.185 (.159)	.070 (.054)
4-dollar bin × home tract median income	-.167 (.234)	-.203 (.232)	-.163 (.217)	-.420 (.550)	-.954 (733056.1)	-.166 (.234)
Yelp rating × home tract median income	.008 (.023)	.007 (.021)	.014 (.020)	-.003 (.033)	.016 (.059)	.008 (.024)
Percent absolute difference in median incomes ($h_i - k_j$)	.850[a] (.126)	.837[a] (.122)	.789[a] (.119)	.801[a] (.182)	.613[b] (.307)	.867[a] (.127)
Percent difference in median incomes ($k_j - h_i$)	.619 (.853)	.015 (.821)	.397 (.821)	1.37 (1.22)	3.76[c] (2.06)	.562 (.857)
Log median household income in k_j	-.360 (.744)	.120 (.719)	-.170 (.719)	-1.02 (1.07)	-3.27[c] (1.78)	-.307 (.747)
Average annual robberies per resident in k_j	2.43[b] (1.20)	2.22[b] (1.10)	2.18[b] (1.08)	1.84 (1.76)	2.27 (3.14)	2.52[b] (1.20)
Number of trips	1079	1079	1079	533	205	1061

NOTES: See notes to Table A.4.

Table A.5: Six-origin-mode specifications: Black reviewers (continued)

	(1) Main spec	(7) Locainfo1	(8) Locainfo2	(9) Late adopt	(10) Cuisine	(11) Cars	(12) Chains
Log travel time from home by public transit	-.938^a	-.839^a	-1.02^a	-.919^a	-.962^a	-.907^a	-.989^a
	(.127)	(.156)	(.246)	(.188)	(.127)	(.126)	(.150)
Log travel time from home by car	-1.19^a	-.999^a	-1.22^a	-1.14^a	-1.22^a	-1.16^a	-1.20^a
	(.158)	(.171)	(.250)	(.208)	(.159)	(.159)	(.167)
Log travel time from work by public transit	-1.85^c	-1.67	-20.4	-16.6	-1.89	-1.84	-1.76^b
	(1.11)	(1.03)	(60876.6)	(40721.8)	(1.15)	(1.15)	(.893)
Log travel time from work by car	-1.79^a	-1.62^a	-1.81^c	-1.99^b	-1.83^a	-1.76^a	-1.63^a
	(.459)	(.463)	(.995)	(.929)	(.464)	(.459)	(.348)
Log travel time from commute by public transit	-.930^a	-.992^a	-.833^a	-.821^a	-.954^a	-.922^a	-.956^a
	(.105)	(.196)	(.125)	(.117)	(.104)	(.110)	(.115)
Log travel time from commute by car	-1.32^a	-1.76^a	-1.14^a	-1.64^a	-1.37^a	-1.32^a	-1.29^a
	(.177)	(.574)	(.200)	(.428)	(.182)	(.186)	(.172)
Euclidean demographic distance between h_i and k_j	-1.84^a	-2.16^a	-1.84^a	-1.92^a	-1.91^a	-1.86^a	-1.81^a
	(.280)	(.415)	(.500)	(.431)	(.283)	(.281)	(.303)
Spectral segregation index of k_j	.075	.063	.135	.277	.065	.044	.114
	(.093)	(.107)	(.609)	(.375)	(.091)	(.101)	(.104)
EDD × SSI	-.171	-.203	.024	-.284	-.165	-.112	-.162
	(.239)	(.283)	(1.14)	(.718)	(.230)	(.258)	(.266)
Share of tract population that is Asian	.011	.113	-.232	.312	.046	.030	-.233
	(.345)	(.442)	(.604)	(.440)	(.356)	(.347)	(.375)
Share of tract population that is black	1.08^a	.658	1.29^c	1.10^b	.707^c	.981^b	1.48^a
	(.399)	(.539)	(.693)	(.513)	(.415)	(.403)	(.432)
Share of tract population that is Hispanic	.467	-.190	.881	1.17^b	.310	.367	.842^b
	(.381)	(.550)	(.566)	(.468)	(.393)	(.384)	(.417)
Share of tract population that is other	3.56	4.81	.332	3.16	3.73	3.72	-.518
	(3.43)	(4.18)	(6.41)	(4.01)	(3.41)	(3.40)	(4.19)
Dummy for 2-dollar bin	.771^a	.419	.211	1.42^a	.655^a	.755^a	.049
	(.197)	(.287)	(.333)	(.292)	(.203)	(.197)	(.209)
Dummy for 3-dollar bin	-.090	-.277	-.794	.202	-.151	-.115	-.990^a
	(.341)	(.471)	(.582)	(.523)	(.346)	(.342)	(.359)
Dummy for 4-dollar bin	-.074	.161	-.850	-.623	-.113	-.081	-1.21
	(1.22)	(1.67)	(1.87)	(1.70)	(1.28)	(1.21)	(1.15)
Yelp rating of restaurant	.053	.086	-.121	.429^b	.042	.054	-.370^b
	(.138)	(.194)	(.226)	(.203)	(.140)	(.138)	(.163)
African cuisine category	-.198	-.922	1.43	.012		-.185	-.592
	(.553)	(.757)	(.913)	(.581)		(.552)	(.575)
American cuisine category	.523^a	.433^a	.608^a	.562^a		.521^a	-.064
	(.119)	(.153)	(.195)	(.149)		(.119)	(.128)
Asian cuisine category	.255^c	.210	.272	.293^c		.249^c	-.337^b
	(.134)	(.171)	(.222)	(.168)		(.134)	(.142)
European cuisine category	-.326^b	-.492^b	-.107	-.510^b		-.330^b	-.767^a
	(.154)	(.199)	(.250)	(.205)		(.155)	(.161)
Indian cuisine category	-.451	-.344	-.814	-.467		-.441	-.777^b
	(.301)	(.350)	(.618)	(.380)		(.301)	(.308)
Latin American cuisine category	1.01^a	1.03^a	.952^a	1.09^a		1.00^a	.384^a
	(.136)	(.174)	(.224)	(.168)		(.136)	(.143)
Middle Eastern cuisine category	.104	.089	.009	.116		.101	-.268
	(.251)	(.308)	(.439)	(.308)		(.251)	(.258)
Vegetarian/vegan cuisine category	.001	.286	-21.9	.123		-.009	-.819^c
	(.409)	(.427)	(41620.7)	(.538)		(.409)	(.420)
2-dollar bin × home tract median income	-.022	.104^c	-.005	-.150^a	-.027	-.019	-.030
	(.032)	(.054)	(.047)	(.051)	(.032)	(.032)	(.034)
3-dollar bin × home tract median income	.077	.169^b	.123	.031	.066	.082	.076
	(.053)	(.084)	(.079)	(.089)	(.053)	(.053)	(.056)
4-dollar bin × home tract median income	-.167	-.200	-.058	-.065	-.162	-.163	-.105
	(.234)	(.353)	(.314)	(.321)	(.243)	(.230)	(.215)
Yelp rating × home tract median income	.008	.013	.022	-.053	.011	.008	.004
	(.023)	(.036)	(.033)	(.036)	(.023)	(.023)	(.027)
Percent absolute difference in median incomes $(h_i - k_j)$	.850^a	.737^a	.789^a	.965^a	.929^a	.877^a	.815^a
	(.126)	(.167)	(.224)	(.162)	(.129)	(.127)	(.136)
Percent difference in median incomes $(k_j - h_i)$	.619	1.91	-.480	1.51	.746	.823	-.763
	(.853)	(1.23)	(1.35)	(1.23)	(.866)	(.857)	(.920)
Log median household income in k_j	-.360	-1.59	.628	-1.06	-.514	-.599	.855
	(.744)	(1.09)	(1.18)	(1.08)	(.755)	(.750)	(.802)
Average annual robberies per resident in k_j	2.43^b	1.96	4.15^b	2.94^b	2.91^b	2.63^b	5.18^a
	(1.20)	(1.50)	(2.09)	(1.47)	(1.22)	(1.20)	(1.27)
Establishment belongs to chain							.224
							(.198)
Log number of Yelp reviews							.873^a
							(.035)
Difference in tracts' private vehicle ownership						-.920^a	
						(.338)	
Number of trips	1079	670	409	707	1079	1079	1079

NOTES: See notes to Table A.4.

Table A.6: Six-origin-mode specifications: White/Hispanic reviewers (part 1)

	(1) Main spec	(2) Choice 50	(3) Choice 100	(4) Half	(5) Fifth	(6) Droploca
Log travel time from home by public transit	-1.13[a] (.059)	-1.10[a] (.055)	-1.11[a] (.053)	-1.15[a] (.080)	-1.22[a] (.135)	-1.08[a] (.059)
Log travel time from home by car	-1.36[a] (.060)	-1.33[a] (.056)	-1.34[a] (.054)	-1.38[a] (.082)	-1.39[a] (.121)	-1.29[a] (.059)
Log travel time from work by public transit	-1.87[a] (.287)	-1.88[a] (.288)	-1.84[a] (.260)	-1.95[a] (.448)	-2.16[b] (1.05)	-1.85[a] (.317)
Log travel time from work by car	-1.95[a] (.171)	-1.97[a] (.178)	-1.96[a] (.169)	-1.93[a] (.215)	-2.20[a] (.503)	-1.86[a] (.167)
Log travel time from commute by public transit	-1.10[a] (.044)	-1.09[a] (.043)	-1.08[a] (.041)	-1.11[a] (.058)	-1.13[a] (.087)	-1.04[a] (.044)
Log travel time from commute by car	-1.43[a] (.058)	-1.39[a] (.053)	-1.38[a] (.050)	-1.48[a] (.081)	-1.50[a] (.123)	-1.34[a] (.056)
Euclidean demographic distance between h_i and k_j	-1.19[a] (.130)	-1.16[a] (.125)	-1.14[a] (.121)	-1.31[a] (.185)	-1.30[a] (.306)	-1.17[a] (.133)
Spectral segregation index of k_j	.045[c] (.027)	.045[c] (.026)	.036 (.026)	-.015 (.052)	-.037 (.086)	.046[c] (.027)
EDD × SSI	-.068 (.083)	-.095 (.086)	-.096 (.085)	.065 (.125)	.023 (.225)	-.071 (.083)
Share of tract population that is Asian	.363[a] (.138)	.364[a] (.131)	.347[a] (.129)	.378[c] (.195)	.494 (.319)	.412[a] (.140)
Share of tract population that is black	.140 (.265)	.057 (.257)	.076 (.250)	.248 (.383)	.894 (.627)	.082 (.273)
Share of tract population that is Hispanic	.415[b] (.188)	.312[c] (.180)	.320[c] (.177)	.618[b] (.267)	.495 (.433)	.431[b] (.191)
Share of tract population that is other	.484 (1.99)	1.43 (1.84)	1.36 (1.81)	-2.07 (2.93)	-.853 (4.65)	.784 (2.02)
Dummy for 2-dollar bin	.355[a] (.083)	.372[a] (.080)	.387[a] (.078)	.399[a] (.117)	.695[a] (.193)	.375[a] (.085)
Dummy for 3-dollar bin	-.026 (.120)	-.068 (.115)	-.035 (.113)	.212 (.172)	.257 (.280)	-.007 (.122)
Dummy for 4-dollar bin	-.347 (.221)	-.249 (.211)	-.268 (.207)	-.228 (.312)	-.182 (.505)	-.304 (.223)
Yelp rating of restaurant	.344[a] (.059)	.343[a] (.056)	.344[a] (.055)	.155[c] (.084)	.294[b] (.138)	.354[a] (.061)
African cuisine category	.298 (.261)	.361 (.246)	.170 (.242)	.422 (.362)	.601 (.537)	.325 (.266)
American cuisine category	.591[a] (.050)	.596[a] (.048)	.604[a] (.048)	.595[a] (.072)	.582[a] (.117)	.602[a] (.051)
Asian cuisine category	.307[a] (.054)	.301[a] (.052)	.297[a] (.051)	.371[a] (.077)	.366[a] (.125)	.308[a] (.056)
European cuisine category	.235[a] (.056)	.225[a] (.053)	.208[a] (.053)	.207[b] (.080)	.160 (.131)	.241[a] (.057)
Indian cuisine category	-.039 (.097)	-.040 (.094)	-.065 (.092)	-.084 (.141)	-.042 (.232)	-.008 (.099)
Latin American cuisine category	.690[a] (.062)	.643[a] (.059)	.644[a] (.058)	.754[a] (.088)	.835[a] (.142)	.683[a] (.063)
Middle Eastern cuisine category	.203[b] (.094)	.279[a] (.090)	.265[a] (.089)	.321[b] (.131)	.453[b] (.207)	.206[b] (.096)
Vegetarian/vegan cuisine category	.587[a] (.116)	.603[a] (.109)	.621[a] (.107)	.682[a] (.164)	.847[a] (.256)	.629[a] (.116)
2-dollar bin × home tract median income	.042[a] (.009)	.040[a] (.009)	.039[a] (.009)	.036[a] (.013)	.008 (.022)	.043[a] (.010)
3-dollar bin × home tract median income	.081[a] (.013)	.083[a] (.012)	.079[a] (.012)	.051[a] (.018)	.047 (.029)	.083[a] (.013)
4-dollar bin × home tract median income	.095[a] (.023)	.082[a] (.022)	.081[a] (.021)	.089[a] (.032)	.093[c] (.050)	.096[a] (.023)
Yelp rating × home tract median income	.016[b] (.007)	.014[b] (.006)	.014[b] (.006)	.030[a] (.009)	.024 (.015)	.016[b] (.007)
Percent absolute difference in median incomes ($h_i - k_j$)	-.100[c] (.053)	-.113[b] (.050)	-.118[b] (.049)	-.074 (.075)	-.047 (.122)	-.108[b] (.054)
Percent difference in median incomes ($k_j - h_i$)	.719[b] (.300)	.787[a] (.286)	.811[a] (.282)	.582 (.417)	.389 (.687)	.767[b] (.307)
Log median household income in k_j	-.625[b] (.262)	-.696[a] (.250)	-.708[a] (.246)	-.516 (.363)	-.291 (.597)	-.649[b] (.268)
Average annual robberies per resident in k_j	-3.74[a] (.771)	-3.65[a] (.753)	-3.82[a] (.751)	-2.75[a] (1.04)	-1.30 (1.57)	-3.82[a] (.800)
Number of trips	6936	6936	6936	3431	1314	6592

NOTES: See notes to Table A.4.

Table A.6: Six-origin-mode specifications: White/Hispanic reviewers (continued)

	(1) Main spec	(7) Locainfo1	(8) Locainfo2	(9) Late adopt	(10) Cuisine	(11) Cars	(12) Chains
Log travel time from home by public transit	-1.13^a	-1.08^a	-1.16^a	-1.10^a	-1.13^a	-1.07^a	-1.16^a
	(.059)	(.078)	(.092)	(.081)	(.059)	(.060)	(.064)
Log travel time from home by car	-1.36^a	-1.35^a	-1.33^a	-1.37^a	-1.36^a	-1.32^a	-1.37^a
	(.060)	(.086)	(.084)	(.090)	(.060)	(.063)	(.063)
Log travel time from work by public transit	-1.87^a	-1.92^a	-1.80^a	-2.07^a	-1.86^a	-1.85^a	-1.98^a
	(.287)	(.464)	(.354)	(.493)	(.281)	(.296)	(.336)
Log travel time from work by car	-1.95^a	-2.08^a	-1.86^a	-2.13^a	-1.94^a	-1.86^a	-2.02^a
	(.171)	(.321)	(.216)	(.289)	(.169)	(.167)	(.191)
Log travel time from commute by public transit	-1.10^a	-1.01^a	-1.20^a	-1.15^a	-1.10^a	-1.06^a	-1.16^a
	(.044)	(.052)	(.083)	(.068)	(.044)	(.047)	(.051)
Log travel time from commute by car	-1.43^a	-1.47^a	-1.38^a	-1.57^a	-1.44^a	-1.42^a	-1.46^a
	(.058)	(.088)	(.079)	(.098)	(.059)	(.065)	(.063)
Euclidean demographic distance between h_i and k_j	-1.19^a	-1.12^a	-1.42^a	-1.50^a	-1.19^a	-1.17^a	-1.15^a
	(.130)	(.168)	(.218)	(.222)	(.131)	(.131)	(.142)
Spectral segregation index of k_j	.045^c	-.093	.047	-.019	.051^c	.042	.072^b
	(.027)	(.103)	(.030)	(.109)	(.027)	(.029)	(.029)
EDD × SSI	-.068	.117	.033	.116	-.076	-.057	-.072
	(.083)	(.196)	(.102)	(.209)	(.084)	(.089)	(.091)
Share of tract population that is Asian	.363^a	.359^b	.525^b	.761^a	.330^b	.307^b	.230
	(.138)	(.177)	(.228)	(.232)	(.142)	(.139)	(.151)
Share of tract population that is black	.140	.024	.476	.015	.112	-.120	.422
	(.265)	(.347)	(.415)	(.440)	(.270)	(.269)	(.293)
Share of tract population that is Hispanic	.415^b	.271	.721^b	.828^a	.377^b	.285	.890^a
	(.188)	(.245)	(.296)	(.311)	(.190)	(.189)	(.204)
Share of tract population that is other	.484	-.724	1.76	-1.69	.876	1.32	-3.87^c
	(1.99)	(2.62)	(3.12)	(3.32)	(1.99)	(1.98)	(2.30)
Dummy for 2-dollar bin	.355^a	.692^a	-.213	.680^a	.232^a	.347^a	-.385^a
	(.083)	(.106)	(.136)	(.143)	(.084)	(.083)	(.089)
Dummy for 3-dollar bin	-.026	.707^a	-1.14^a	.093	-.189	-.034	-.964^a
	(.120)	(.152)	(.205)	(.207)	(.123)	(.120)	(.129)
Dummy for 4-dollar bin	-.347	.312	-1.62^a	.210	-.451^b	-.340	-1.32^a
	(.221)	(.268)	(.404)	(.359)	(.224)	(.220)	(.237)
Yelp rating of restaurant	.344^a	.348^a	.351^a	.351^a	.356^a	.348^a	-.105
	(.059)	(.076)	(.096)	(.103)	(.061)	(.060)	(.074)
African cuisine category	.298	.114	.588	.468		.283	-.184
	(.261)	(.347)	(.398)	(.412)		(.262)	(.270)
American cuisine category	.591^a	.615^a	.561^a	.647^a		.591^a	.027
	(.050)	(.069)	(.074)	(.083)		(.050)	(.054)
Asian cuisine category	.307^a	.431^a	.142^c	.302^a		.304^a	-.277^a
	(.054)	(.073)	(.081)	(.090)		(.054)	(.058)
European cuisine category	.235^a	.285^a	.169^b	.365^a		.234^a	-.074
	(.056)	(.076)	(.082)	(.090)		(.056)	(.059)
Indian cuisine category	-.039	.167	-.333^b	-.126		-.034	-.446^a
	(.097)	(.125)	(.157)	(.167)		(.098)	(.103)
Latin American cuisine category	.690^a	.814^a	.520^a	.899^a		.684^a	.053
	(.062)	(.082)	(.094)	(.099)		(.062)	(.066)
Middle Eastern cuisine category	.203^b	.175	.237^c	.136		.199^b	-.244^b
	(.094)	(.129)	(.138)	(.159)		(.094)	(.101)
Vegetarian/vegan cuisine category	.587^a	.737^a	.377^b	.360^c		.584^a	-.297^b
	(.116)	(.149)	(.187)	(.216)		(.116)	(.123)
2-dollar bin × home tract median income	.042^a	-.005	.119^a	-.002	.041^a	.043^a	.046^a
	(.009)	(.012)	(.016)	(.016)	(.010)	(.009)	(.010)
3-dollar bin × home tract median income	.081^a	-.015	.213^a	.065^a	.080^a	.082^a	.087^a
	(.013)	(.017)	(.021)	(.022)	(.013)	(.013)	(.014)
4-dollar bin × home tract median income	.095^a	.023	.226^a	.039	.092^a	.096^a	.099^a
	(.023)	(.030)	(.038)	(.038)	(.023)	(.023)	(.024)
Yelp rating × home tract median income	.016^b	.018^b	.013	.016	.017^b	.016^b	.019^b
	(.007)	(.009)	(.010)	(.012)	(.007)	(.007)	(.008)
Percent absolute difference in median incomes ($h_i - k_j$)	-.100^c	-.036	-.178^b	-.031	-.103^c	-.061	-.137^b
	(.053)	(.068)	(.084)	(.086)	(.053)	(.053)	(.057)
Percent difference in median incomes ($k_j - h_i$)	.719^b	.840^b	.678	.998^b	.754^b	1.07^a	.632^b
	(.300)	(.377)	(.504)	(.481)	(.302)	(.304)	(.321)
Log median household income in k_j	-.625^b	-.778^b	-.515	-.906^b	-.660^b	-.998^a	-.554^b
	(.262)	(.326)	(.447)	(.418)	(.264)	(.266)	(.280)
Average annual robberies per resident in k_j	-3.74^a	-2.52^a	-6.63^a	-4.46^a	-3.24^a	-2.96^a	-.182
	(.771)	(.906)	(1.51)	(1.23)	(.771)	(.752)	(.784)
Establishment belongs to chain							-.020
							(.104)
Log number of Yelp reviews							.923^a
							(.014)
Difference in tracts' private vehicle ownership						-1.42^a	
						(.128)	
Number of trips	6936	3866	3070	2647	6936	6936	6936

NOTES: See notes to Table A.4.

Table A.7: Minimum-time specifications: Asian reviewers (part 1)

	(1) Mintime	(2) Choice 50	(3) Choice 100	(4) Half	(5) Fifth	(6) Droploca
Log minimum travel time	-.932^a	-.936^a	-.935^a	-.949^a	-1.03^a	-.859^a
	(.022)	(.020)	(.020)	(.031)	(.049)	(.022)
Euclidean demographic distance between h_i and k_j	-1.13^a	-1.08^a	-1.06^a	-1.25^a	-1.11^a	-1.04^a
	(.120)	(.114)	(.112)	(.171)	(.277)	(.122)
Spectral segregation index of k_j	.153^a	.143^a	.145^a	.116^c	.190^c	.151^a
	(.051)	(.048)	(.048)	(.062)	(.101)	(.052)
EDD × SSI	-.123	-.116	-.116	-.140	-.169	-.128
	(.115)	(.107)	(.107)	(.157)	(.238)	(.116)
Share of tract population that is Asian	1.08^a	1.02^a	1.05^a	1.12^a	1.15^a	1.04^a
	(.119)	(.113)	(.111)	(.166)	(.270)	(.121)
Share of tract population that is black	.328	.317	.373	.661	.857	.284
	(.317)	(.308)	(.305)	(.439)	(.726)	(.327)
Share of tract population that is Hispanic	-.215	-.090	-.061	-1.19^a	-.991^c	-.242
	(.233)	(.227)	(.224)	(.362)	(.602)	(.237)
Share of tract population that is other	.619	1.24	1.19	.447	5.95	.251
	(2.04)	(1.95)	(1.92)	(2.85)	(4.55)	(2.12)
Dummy for 2-dollar bin	.382^a	.377^a	.358^a	.360^a	.416^b	.401^a
	(.086)	(.083)	(.082)	(.121)	(.198)	(.089)
Dummy for 3-dollar bin	.313^a	.272^b	.257^b	.209	.131	.338^a
	(.116)	(.112)	(.110)	(.165)	(.269)	(.118)
Dummy for 4-dollar bin	.222	.283	.279	.041	.019	.251
	(.188)	(.179)	(.175)	(.264)	(.396)	(.189)
Yelp rating of restaurant	.570^a	.550^a	.531^a	.484^a	.404^a	.580^a
	(.064)	(.060)	(.059)	(.090)	(.144)	(.065)
African cuisine category	.281	.160	.353	.440	.538	.297
	(.298)	(.288)	(.286)	(.445)	(.754)	(.298)
American cuisine category	.435^a	.413^a	.428^a	.406^a	.438^a	.446^a
	(.054)	(.053)	(.052)	(.077)	(.127)	(.056)
Asian cuisine category	.884^a	.875^a	.876^a	.914^a	.994^a	.884^a
	(.054)	(.052)	(.052)	(.076)	(.126)	(.055)
European cuisine category	.199^a	.203^a	.200^a	.191^b	.115	.204^a
	(.059)	(.057)	(.057)	(.084)	(.140)	(.060)
Indian cuisine category	.390^a	.353^a	.344^a	.392^a	.389^c	.399^a
	(.091)	(.088)	(.087)	(.128)	(.212)	(.093)
Latin American cuisine category	.520^a	.496^a	.508^a	.489^a	.566^a	.500^a
	(.070)	(.067)	(.067)	(.100)	(.161)	(.072)
Middle Eastern cuisine category	.294^a	.307^a	.315^a	.371^a	.256	.303^a
	(.100)	(.097)	(.096)	(.139)	(.237)	(.102)
Vegetarian/vegan cuisine category	.396^a	.341^a	.392^a	.740^a	.847^a	.393^a
	(.137)	(.130)	(.129)	(.170)	(.265)	(.140)
2-dollar bin × home tract median income	.033^a	.031^a	.033^a	.027^c	.025	.034^a
	(.011)	(.010)	(.010)	(.015)	(.025)	(.011)
3-dollar bin × home tract median income	.071^a	.069^a	.073^a	.075^a	.096^a	.074^a
	(.014)	(.013)	(.013)	(.019)	(.031)	(.014)
4-dollar bin × home tract median income	.073^a	.061^a	.059^a	.087^a	.121^a	.075^a
	(.022)	(.021)	(.020)	(.030)	(.046)	(.022)
Yelp rating × home tract median income	.011	.011	.012^c	.016	.028	.011
	(.008)	(.007)	(.007)	(.011)	(.018)	(.008)
Percent absolute difference in median incomes $(h_i - k_j)$	-.117^b	-.124^a	-.128^a	-.123^c	-.159	-.105^b
	(.050)	(.048)	(.047)	(.071)	(.115)	(.051)
Percent difference in median incomes $(k_j - h_i)$	.173	.177	.139	.163	-.018	.112
	(.305)	(.292)	(.289)	(.421)	(.688)	(.310)
Log median household income in k_j	-.155	-.160	-.123	-.195	-.071	-.112
	(.267)	(.256)	(.254)	(.368)	(.605)	(.272)
Average annual robberies per resident in k_j	-3.05^a	-2.98^a	-2.86^a	-3.99^a	-7.39^a	-3.37^a
	(.672)	(.655)	(.650)	(.973)	(1.96)	(.711)
Number of trips	6447	6447	6447	3205	1258	6181

NOTES: Each column reports an estimated conditional-logit model of the decision to visit a Yelp venue. Column 1 shows specification from main text. Columns 2 and 3 show specifications in which randomly generated choice sets have 50 and 100 restaurants, respectively. Columns 4 and 5 show specifications in which observations are limited to the first half and first fifth of NYC restaurant reviews posted by each reviewer, respectively. Column 6 drops observations that are restaurant reviews containing locational information used to identify the residence or workplace of the reviewer. Standard errors in parentheses. Statistical significance denoted by a (1%), b (5%), c (10%). Unreported controls are 28 area dummies.

Table A.7: Minimum-time specifications: Asian reviewers (continued)

	(1) Mintime	(7) Locainfo1	(8) Locainfo2	(9) Late adopt	(10) Cuisine	(11) Cars	(12) Chains
Log minimum travel time	-.932^a	-.844^a	-1.03^a	-1.05^a	-.940^a	-.900^a	-.982^a
	(.022)	(.029)	(.034)	(.035)	(.022)	(.022)	(.024)
Euclidean demographic distance between h_i and k_j	-1.13^a	-1.02^a	-.983^a	-1.09^a	-1.16^a	-1.08^a	-1.21^a
	(.120)	(.168)	(.183)	(.195)	(.121)	(.121)	(.133)
Spectral segregation index of k_j	.153^a	.103^c	.617^a	.449^a	.156^a	.144^a	.172^a
	(.051)	(.061)	(.130)	(.126)	(.051)	(.051)	(.055)
EDD × SSI	-.123	-.087	-.791^a	-.397	-.157	-.184	-.141
	(.115)	(.142)	(.241)	(.258)	(.121)	(.121)	(.130)
Share of tract population that is Asian	1.08^a	.866^a	1.26^a	1.05^a	.952^a	.995^a	1.13^a
	(.119)	(.157)	(.187)	(.183)	(.124)	(.120)	(.135)
Share of tract population that is black	.328	.995^b	-.684	.020	.200	-.012	.401
	(.317)	(.404)	(.521)	(.549)	(.324)	(.323)	(.360)
Share of tract population that is Hispanic	-.215	-.983^a	.976^a	.175	-.290	-.411^c	.076
	(.233)	(.305)	(.368)	(.355)	(.238)	(.236)	(.257)
Share of tract population that is other	.619	-.616	2.11	.451	.971	.840	-4.46^c
	(2.04)	(2.72)	(3.16)	(2.91)	(2.03)	(2.03)	(2.45)
Dummy for 2-dollar bin	.382^a	.342^a	.647^a	.405^a	.251^a	.373^a	-.391^a
	(.086)	(.111)	(.154)	(.122)	(.089)	(.086)	(.096)
Dummy for 3-dollar bin	.313^a	.140	.863^a	.346^b	.149	.296^b	-.599^a
	(.116)	(.153)	(.196)	(.165)	(.119)	(.116)	(.127)
Dummy for 4-dollar bin	.222	.326	.330	.218	.148	.218	-.703^a
	(.188)	(.241)	(.320)	(.272)	(.192)	(.187)	(.205)
Yelp rating of restaurant	.570^a	.362^a	.912^a	.582^a	.565^a	.567^a	.071
	(.064)	(.082)	(.111)	(.091)	(.065)	(.063)	(.082)
African cuisine category	.281	.038	.465	.064		.290	-.124
	(.298)	(.471)	(.391)	(.531)		(.298)	(.305)
American cuisine category	.435^a	.574^a	.281^a	.479^a		.435^a	-.120^b
	(.054)	(.076)	(.078)	(.087)		(.054)	(.060)
Asian cuisine category	.884^a	.953^a	.814^a	.982^a		.882^a	.247^a
	(.054)	(.076)	(.077)	(.086)		(.054)	(.060)
European cuisine category	.199^a	.259^a	.137	.343^a		.196^a	-.129^b
	(.059)	(.084)	(.084)	(.093)		(.059)	(.065)
Indian cuisine category	.390^a	.411^a	.365^a	.484^a		.389^a	-.002
	(.091)	(.129)	(.128)	(.143)		(.091)	(.098)
Latin American cuisine category	.520^a	.586^a	.461^a	.604^a		.517^a	-.130^c
	(.070)	(.097)	(.101)	(.109)		(.070)	(.076)
Middle Eastern cuisine category	.294^a	.353^b	.231	.369^b		.284^a	-.170
	(.100)	(.141)	(.142)	(.154)		(.100)	(.110)
Vegetarian/vegan cuisine category	.396^a	.821^a	-.054	.330		.387^a	-.504^a
	(.137)	(.182)	(.208)	(.224)		(.137)	(.146)
2-dollar bin × home tract median income	.033^a	.030^c	.012	.019	.034^a	.034^a	.043^a
	(.011)	(.016)	(.017)	(.015)	(.011)	(.011)	(.012)
3-dollar bin × home tract median income	.071^a	.076^a	.027	.060^a	.069^a	.074^a	.077^a
	(.014)	(.020)	(.021)	(.019)	(.014)	(.014)	(.015)
4-dollar bin × home tract median income	.073^a	.043	.074^b	.062^b	.072^a	.074^a	.078^a
	(.022)	(.032)	(.034)	(.031)	(.022)	(.022)	(.024)
Yelp rating × home tract median income	.011	.040^a	-.024^b	.015	.009	.011	.012
	(.008)	(.012)	(.012)	(.011)	(.008)	(.008)	(.010)
Percent absolute difference in median incomes ($h_i - k_j$)	-.117^b	-.052	-.221^a	-.102	-.111^b	-.085^c	-.106^c
	(.050)	(.067)	(.083)	(.074)	(.050)	(.050)	(.055)
Percent difference in median incomes ($k_j - h_i$)	.173	-.282	.817^c	.301	.293	.488	.155
	(.305)	(.431)	(.443)	(.435)	(.307)	(.309)	(.334)
Log median household income in k_j	-.155	.270	-.741^c	-.305	-.280	-.504^c	-.167
	(.267)	(.379)	(.389)	(.379)	(.270)	(.272)	(.292)
Average annual robberies per resident in k_j	-3.05^a	-3.64^a	-2.40^b	-2.19^b	-2.43^a	-2.21^a	.473
	(.672)	(.918)	(.994)	(.952)	(.672)	(.664)	(.704)
Establishment belongs to chain							.147
							(.113)
Log number of Yelp reviews							1.01^a
							(.015)
Difference in tracts' private vehicle ownership						-1.17^a	
						(.127)	
Number of trips	6447	3326	3121	2766	6447	6447	6447

NOTES: Each column reports an estimated conditional-logit model of the decision to visit a Yelp venue. Column 1 shows specification from main text. Columns 7 and 8 split the estimation sample into reviewers with residences located using information contained in one or two reviews in column 2 and three or more reviews in column 3. Column 9 restricts the sample to Yelp reviewers with later-than-median dates of first review written. Column 10 introduces dummies for 39 more disaggregated cuisine categories. Column 11 controls for origin-destination differences in vehicle ownership. Column 12 controls for the number of Yelp reviews of each restaurant and its membership in a chain with more than eight NYC locations. Standard errors in parentheses. Statistical significance denoted by *a* (1%), *b* (5%), *c* (10%). Unreported controls are 28 area dummies.

Table A.8: Minimum-time specifications: Black reviewers (part 1)

	(1) Mintime	(2) Choice 50	(3) Choice 100	(4) Half	(5) Fifth	(6) Droploca
Log minimum travel time	-.957[a]	-.973[a]	-.950[a]	-1.10[a]	-1.03[a]	-.938[a]
	(.044)	(.042)	(.040)	(.067)	(.107)	(.045)
Euclidean demographic distance between h_i and k_j	-2.00[a]	-2.06[a]	-1.90[a]	-2.05[a]	-2.23[a]	-2.01[a]
	(.277)	(.270)	(.263)	(.407)	(.757)	(.278)
Spectral segregation index of k_j	.113	.192[b]	.161[b]	.137	.492	.112
	(.084)	(.079)	(.077)	(.153)	(.569)	(.084)
EDD × SSI	-.214	-.379	-.381	-.132	-.948	-.205
	(.224)	(.242)	(.252)	(.358)	(1.18)	(.217)
Share of tract population that is Asian	.017	-.003	.019	.331	.882	-.020
	(.344)	(.333)	(.329)	(.488)	(.830)	(.347)
Share of tract population that is black	1.13[a]	.911[b]	1.08[a]	1.65[a]	.992	1.21[a]
	(.396)	(.388)	(.379)	(.590)	(.977)	(.401)
Share of tract population that is Hispanic	.430	.497	.518	1.05[c]	1.58[c]	.450
	(.379)	(.357)	(.353)	(.554)	(.910)	(.381)
Share of tract population that is other	4.60	6.42[b]	6.22[b]	6.85	12.6	3.74
	(3.36)	(2.94)	(2.89)	(5.28)	(9.54)	(3.44)
Dummy for 2-dollar bin	.778[a]	.761[a]	.752[a]	.761[a]	1.13[b]	.797[a]
	(.195)	(.187)	(.185)	(.276)	(.465)	(.198)
Dummy for 3-dollar bin	-.061	.197	.164	1.20[b]	1.65[c]	-.008
	(.340)	(.320)	(.317)	(.582)	(.892)	(.345)
Dummy for 4-dollar bin	-.007	.188	.008	.316	-24.7	.007
	(1.22)	(1.22)	(1.15)	(2.57)	(4569027.9)	(1.21)
Yelp rating of restaurant	.062	.077	.037	.149	.136	.061
	(.136)	(.125)	(.122)	(.199)	(.341)	(.139)
African cuisine category	-.199	-.032	.039	-.217	-.632	-.193
	(.552)	(.533)	(.522)	(.667)	(1.14)	(.553)
American cuisine category	.532[a]	.599[a]	.546[a]	.446[a]	.269	.548[a]
	(.119)	(.114)	(.112)	(.168)	(.272)	(.120)
Asian cuisine category	.267[b]	.278[b]	.249[b]	.236	.239	.295[b]
	(.134)	(.128)	(.127)	(.189)	(.302)	(.135)
European cuisine category	-.312[b]	-.235	-.285[c]	-.483[b]	-.386	-.296[c]
	(.154)	(.148)	(.146)	(.222)	(.347)	(.155)
Indian cuisine category	-.416	-.557[c]	-.554[c]	-.655	-.699	-.383
	(.300)	(.292)	(.289)	(.457)	(.790)	(.301)
Latin American cuisine category	1.02[a]	1.01[a]	.920[a]	.867[a]	.853[a]	1.04[a]
	(.136)	(.128)	(.126)	(.193)	(.300)	(.137)
Middle Eastern cuisine category	.098	.083	-.030	-.344	-.744	.150
	(.250)	(.242)	(.239)	(.402)	(.788)	(.251)
Vegetarian/vegan cuisine category	.001	-.214	-.329	-.356	-.498	.032
	(.409)	(.400)	(.400)	(.618)	(1.07)	(.410)
2-dollar bin × home tract median income	-.025	-.031	-.028	.002	-.052	-.026
	(.031)	(.030)	(.030)	(.044)	(.075)	(.032)
3-dollar bin × home tract median income	.068	.017	.023	-.193[c]	-.203	.060
	(.053)	(.050)	(.049)	(.108)	(.163)	(.054)
4-dollar bin × home tract median income	-.180	-.225	-.185	-.405	-1.07	-.178
	(.232)	(.233)	(.219)	(.542)	(1268375.8)	(.231)
Yelp rating × home tract median income	.006	.009	.012	-.002	.020	.007
	(.023)	(.021)	(.020)	(.033)	(.058)	(.023)
Percent absolute difference in median incomes $(h_i - k_j)$	.843[a]	.816[a]	.769[a]	.804[a]	.623[b]	.859[a]
	(.125)	(.121)	(.118)	(.181)	(.304)	(.126)
Percent difference in median incomes $(k_j - h_i)$	.419	-.154	.173	1.08	3.68[c]	.369
	(.845)	(.814)	(.814)	(1.21)	(2.05)	(.849)
Log median household income in k_j	-.175	.278	.038	-.748	-3.19[c]	-.129
	(.737)	(.713)	(.713)	(1.06)	(1.77)	(.740)
Average annual robberies per resident in k_j	2.49[b]	2.26[b]	2.37[b]	2.21	2.74	2.57[b]
	(1.18)	(1.09)	(1.07)	(1.74)	(3.08)	(1.19)
Number of trips	1079	1079	1079	533	205	1061

NOTES: See notes to Table A.7.

Table A.8: Minimum-time specifications: Black reviewers (continued)

	(1) Mintime	(7) Locainfo1	(8) Locainfo2	(9) Late adopt	(10) Cuisine	(11) Cars	(12) Chains
Log minimum travel time	-.957^a	-.931^a	-.900^a	-.912^a	-.987^a	-.932^a	-.980^a
	(.044)	(.066)	(.068)	(.056)	(.045)	(.045)	(.047)
Euclidean demographic distance between h_i and k_j	-2.00^a	-2.30^a	-2.02^a	-2.04^a	-2.07^a	-2.03^a	-1.94^a
	(.277)	(.416)	(.491)	(.428)	(.280)	(.279)	(.300)
Spectral segregation index of k_j	.113	.106	.068	.355	.103	.082	.153
	(.084)	(.103)	(.605)	(.372)	(.083)	(.094)	(.095)
EDD × SSI	-.214	-.263	.186	-.395	-.206	-.152	-.211
	(.224)	(.282)	(1.12)	(.711)	(.213)	(.248)	(.251)
Share of tract population that is Asian	.017	.151	-.305	.344	.054	.043	-.205
	(.344)	(.439)	(.603)	(.437)	(.354)	(.345)	(.372)
Share of tract population that is black	1.13^a	.808	1.27^c	1.15^b	.785^c	1.03^b	1.57^a
	(.396)	(.533)	(.690)	(.507)	(.411)	(.400)	(.428)
Share of tract population that is Hispanic	.430	-.318	.890	1.09^b	.288	.315	.797^c
	(.379)	(.546)	(.561)	(.464)	(.391)	(.383)	(.414)
Share of tract population that is other	4.60	6.01	.841	3.94	4.62	4.74	.555
	(3.36)	(4.09)	(6.41)	(3.93)	(3.34)	(3.32)	(4.11)
Dummy for 2-dollar bin	.778^a	.439	.234	1.42^a	.669^a	.761^a	.043
	(.195)	(.283)	(.331)	(.289)	(.201)	(.195)	(.207)
Dummy for 3-dollar bin	-.061	-.215	-.722	.251	-.123	-.087	-.966^a
	(.340)	(.470)	(.577)	(.516)	(.345)	(.341)	(.359)
Dummy for 4-dollar bin	-.007	.278	-.806	-.572	-.056	-.017	-1.13
	(1.22)	(1.67)	(1.86)	(1.70)	(1.27)	(1.20)	(1.17)
Yelp rating of restaurant	.062	.097	-.117	.431^b	.047	.063	-.353^b
	(.136)	(.191)	(.226)	(.200)	(.139)	(.137)	(.161)
African cuisine category	-.199	-.890	1.33	-.023		-.181	-.622
	(.552)	(.755)	(.914)	(.580)		(.552)	(.577)
American cuisine category	.532^a	.457^a	.609^a	.575^a		.529^a	-.068
	(.119)	(.152)	(.194)	(.148)		(.119)	(.127)
Asian cuisine category	.267^b	.235	.275	.302^c		.261^c	-.336^b
	(.134)	(.170)	(.221)	(.167)		(.134)	(.141)
European cuisine category	-.312^b	-.457^b	-.114	-.507^b		-.315^b	-.768^a
	(.154)	(.198)	(.249)	(.205)		(.154)	(.161)
Indian cuisine category	-.416	-.274	-.823	-.430		-.404	-.763^b
	(.300)	(.347)	(.618)	(.378)		(.300)	(.307)
Latin American cuisine category	1.02^a	1.05^a	.956^a	1.10^a		1.01^a	.379^a
	(.136)	(.173)	(.223)	(.167)		(.136)	(.143)
Middle Eastern cuisine category	.098	.108	-.038	.131		.092	-.285
	(.250)	(.307)	(.437)	(.307)		(.250)	(.257)
Vegetarian/vegan cuisine category	.001	.287	-21.8	.106		-.014	-.822^b
	(.409)	(.427)	(38566.5)	(.539)		(.410)	(.419)
2-dollar bin × home tract median income	-.025	.096^c	-.010	-.150^a	-.031	-.022	-.032
	(.031)	(.053)	(.047)	(.051)	(.032)	(.032)	(.033)
3-dollar bin × home tract median income	.068	.147^c	.113	.017	.058	.073	.065
	(.053)	(.084)	(.079)	(.088)	(.053)	(.053)	(.056)
4-dollar bin × home tract median income	-.180	-.221	-.068	-.072	-.174	-.174	-.130
	(.232)	(.353)	(.313)	(.320)	(.240)	(.228)	(.218)
Yelp rating × home tract median income	.006	.010	.022	-.054	.009	.005	.000
	(.023)	(.035)	(.033)	(.035)	(.023)	(.023)	(.027)
Percent absolute difference in median incomes $(h_i - k_j)$	.843^a	.713^a	.819^a	.928^a	.918^a	.873^a	.815^a
	(.125)	(.166)	(.221)	(.160)	(.127)	(.126)	(.135)
Percent difference in median incomes $(k_j - h_i)$	.419	1.91	-.879	1.40	.565	.671	-.997
	(.845)	(1.22)	(1.35)	(1.22)	(.857)	(.849)	(.914)
Log median household income in k_j	-.175	-1.57	.987	-.967	-.344	-.465	1.07
	(.737)	(1.07)	(1.18)	(1.07)	(.748)	(.743)	(.797)
Average annual robberies per resident in k_j	2.49^b	2.09	4.22^b	3.11^b	2.96^b	2.70^b	5.32^a
	(1.18)	(1.46)	(2.10)	(1.44)	(1.21)	(1.18)	(1.25)
Establishment belongs to chain							.210
							(.197)
Log number of Yelp reviews							.881^a
							(.035)
Difference in tracts' private vehicle ownership						-1.10^a	
						(.335)	
Number of trips	1079	670	409	707	1079	1079	1079

NOTES: See notes to Table A.7.

Table A.9: Minimum-time specifications: White/Hispanic reviewers (part 1)

	(1) Mintime	(2) Choice 50	(3) Choice 100	(4) Half	(5) Fifth	(6) Droploca
Log minimum travel time	-1.19[a]	-1.17[a]	-1.18[a]	-1.22[a]	-1.27[a]	-1.11[a]
	(.019)	(.018)	(.017)	(.027)	(.044)	(.019)
Euclidean demographic distance between h_i and k_j	-1.39[a]	-1.36[a]	-1.35[a]	-1.49[a]	-1.47[a]	-1.36[a]
	(.130)	(.124)	(.121)	(.184)	(.304)	(.133)
Spectral segregation index of k_j	.045[c]	.042	.035	-.019	-.041	.046[c]
	(.027)	(.026)	(.026)	(.052)	(.085)	(.027)
EDD × SSI	-.045	-.065	-.072	.085	.052	-.053
	(.081)	(.084)	(.084)	(.123)	(.218)	(.082)
Share of tract population that is Asian	.485[a]	.492[a]	.489[a]	.492[b]	.577[c]	.531[a]
	(.136)	(.131)	(.128)	(.194)	(.317)	(.139)
Share of tract population that is black	.200	.149	.155	.306	.866	.131
	(.264)	(.256)	(.250)	(.382)	(.628)	(.272)
Share of tract population that is Hispanic	.477[b]	.371[b]	.398[b]	.652[b]	.489	.501[a]
	(.187)	(.179)	(.176)	(.266)	(.432)	(.190)
Share of tract population that is other	.315	1.16	1.18	-2.24	-.956	.685
	(1.97)	(1.83)	(1.80)	(2.91)	(4.62)	(2.00)
Dummy for 2-dollar bin	.362[a]	.370[a]	.398[a]	.415[a]	.708[a]	.386[a]
	(.082)	(.079)	(.078)	(.116)	(.191)	(.084)
Dummy for 3-dollar bin	.001	-.044	.003	.232	.283	.025
	(.120)	(.114)	(.113)	(.171)	(.278)	(.122)
Dummy for 4-dollar bin	-.289	-.220	-.224	-.159	-.110	-.243
	(.221)	(.212)	(.209)	(.311)	(.506)	(.223)
Yelp rating of restaurant	.338[a]	.330[a]	.328[a]	.158[c]	.312[b]	.349[a]
	(.059)	(.056)	(.054)	(.083)	(.137)	(.060)
African cuisine category	.289	.378	.218	.412	.667	.320
	(.259)	(.244)	(.239)	(.353)	(.511)	(.263)
American cuisine category	.583[a]	.593[a]	.607[a]	.577[a]	.584[a]	.594[a]
	(.050)	(.048)	(.047)	(.071)	(.116)	(.051)
Asian cuisine category	.301[a]	.297[a]	.299[a]	.355[a]	.359[a]	.302[a]
	(.054)	(.052)	(.051)	(.077)	(.125)	(.055)
European cuisine category	.233[a]	.231[a]	.217[a]	.196[b]	.156	.239[a]
	(.055)	(.053)	(.053)	(.080)	(.130)	(.057)
Indian cuisine category	-.032	-.034	-.057	-.074	-.019	-.001
	(.097)	(.093)	(.092)	(.139)	(.230)	(.098)
Latin American cuisine category	.685[a]	.640[a]	.643[a]	.737[a]	.846[a]	.675[a]
	(.061)	(.059)	(.058)	(.087)	(.141)	(.063)
Middle Eastern cuisine category	.187[b]	.278[a]	.263[a]	.298[b]	.438[b]	.194[b]
	(.094)	(.090)	(.089)	(.130)	(.205)	(.096)
Vegetarian/vegan cuisine category	.583[a]	.588[a]	.628[a]	.673[a]	.877[a]	.625[a]
	(.115)	(.109)	(.107)	(.163)	(.254)	(.116)
2-dollar bin × home tract median income	.040[a]	.039[a]	.037[a]	.033[b]	.006	.041[a]
	(.009)	(.009)	(.009)	(.013)	(.021)	(.010)
3-dollar bin × home tract median income	.075[a]	.079[a]	.074[a]	.046[b]	.041	.078[a]
	(.013)	(.012)	(.012)	(.018)	(.029)	(.013)
4-dollar bin × home tract median income	.086[a]	.075[a]	.074[a]	.079[b]	.082	.088[a]
	(.023)	(.022)	(.021)	(.032)	(.050)	(.023)
Yelp rating × home tract median income	.017[b]	.015[b]	.015[b]	.030[a]	.021	.016[b]
	(.007)	(.006)	(.006)	(.009)	(.015)	(.007)
Percent absolute difference in median incomes $(h_i - k_j)$	-.116[b]	-.142[a]	-.146[a]	-.098	-.066	-.119[b]
	(.052)	(.050)	(.049)	(.075)	(.121)	(.053)
Percent difference in median incomes $(k_j - h_i)$	.882[a]	.993[a]	1.04[a]	.754[c]	.526	.913[a]
	(.298)	(.286)	(.282)	(.416)	(.686)	(.306)
Log median household income in k_j	-.751[a]	-.858[a]	-.896[a]	-.650[c]	-.387	-.761[a]
	(.260)	(.250)	(.247)	(.361)	(.596)	(.267)
Average annual robberies per resident in k_j	-3.87[a]	-3.84[a]	-3.99[a]	-2.88[a]	-1.35	-3.95[a]
	(.768)	(.750)	(.747)	(1.04)	(1.57)	(.797)
Number of trips	6936	6936	6936	3431	1314	6592

NOTES: See notes to Table A.7.

Table A.9: Minimum-time specifications: White/Hispanic reviewers (continued)

	(1) Mintime	(7) Locainfo1	(8) Locainfo2	(9) Late adopt	(10) Cuisine	(11) Cars	(12) Chains
Log minimum travel time	-1.19^a	-1.15^a	-1.20^a	-1.26^a	-1.19^a	-1.14^a	-1.23^a
	(.019)	(.025)	(.029)	(.032)	(.019)	(.019)	(.021)
Euclidean demographic distance between h_i and k_j	-1.39^a	-1.30^a	-1.65^a	-1.67^a	-1.39^a	-1.37^a	-1.36^a
	(.130)	(.167)	(.217)	(.221)	(.130)	(.130)	(.141)
Spectral segregation index of k_j	.045^c	-.075	.044	-.011	.051^c	.039	.070^b
	(.027)	(.100)	(.030)	(.108)	(.027)	(.029)	(.029)
EDD × SSI	-.045	.121	.055	.134	-.052	-.028	-.050
	(.081)	(.190)	(.101)	(.204)	(.082)	(.088)	(.089)
Share of tract population that is Asian	.485^a	.481^a	.650^a	.853^a	.461^a	.424^a	.368^b
	(.136)	(.175)	(.226)	(.230)	(.140)	(.138)	(.150)
Share of tract population that is black	.200	.013	.635	.022	.170	-.101	.454
	(.264)	(.347)	(.411)	(.437)	(.269)	(.268)	(.292)
Share of tract population that is Hispanic	.477^b	.337	.769^a	.878^a	.449^b	.333^c	.965^a
	(.187)	(.244)	(.296)	(.309)	(.189)	(.188)	(.204)
Share of tract population that is other	.315	-.699	1.20	-1.90	.750	1.24	-4.01^c
	(1.97)	(2.58)	(3.12)	(3.27)	(1.98)	(1.96)	(2.27)
Dummy for 2-dollar bin	.362^a	.694^a	-.193	.680^a	.243^a	.352^a	-.377^a
	(.082)	(.105)	(.135)	(.141)	(.084)	(.082)	(.088)
Dummy for 3-dollar bin	.001	.718^a	-1.08^a	.095	-.158	-.010	-.944^a
	(.120)	(.152)	(.203)	(.206)	(.122)	(.120)	(.128)
Dummy for 4-dollar bin	-.289	.380	-1.56^a	.190	-.389^c	-.284	-1.27^a
	(.221)	(.269)	(.403)	(.362)	(.224)	(.220)	(.237)
Yelp rating of restaurant	.338^a	.349^a	.331^a	.361^a	.350^a	.345^a	-.113
	(.059)	(.076)	(.095)	(.102)	(.060)	(.059)	(.073)
African cuisine category	.289	.140	.541	.412		.273	-.219
	(.259)	(.339)	(.402)	(.413)		(.259)	(.268)
American cuisine category	.583^a	.610^a	.549^a	.642^a		.582^a	.018
	(.050)	(.068)	(.074)	(.082)		(.050)	(.054)
Asian cuisine category	.301^a	.429^a	.132	.307^a		.299^a	-.285^a
	(.054)	(.073)	(.081)	(.089)		(.054)	(.058)
European cuisine category	.233^a	.284^a	.165^b	.372^a		.232^a	-.076
	(.055)	(.075)	(.082)	(.090)		(.055)	(.059)
Indian cuisine category	-.032	.186	-.345^b	-.118		-.025	-.432^a
	(.097)	(.124)	(.156)	(.166)		(.097)	(.102)
Latin American cuisine category	.685^a	.809^a	.515^a	.896^a		.678^a	.050
	(.061)	(.082)	(.093)	(.098)		(.061)	(.065)
Middle Eastern cuisine category	.187^b	.165	.211	.130		.183^c	-.262^a
	(.094)	(.128)	(.138)	(.157)		(.094)	(.100)
Vegetarian/vegan cuisine category	.583^a	.750^a	.349^c	.359^c		.580^a	-.306^b
	(.115)	(.148)	(.186)	(.214)		(.115)	(.122)
2-dollar bin × home tract median income	.040^a	-.006	.115^a	-.003	.038^a	.041^a	.044^a
	(.009)	(.012)	(.016)	(.016)	(.009)	(.009)	(.010)
3-dollar bin × home tract median income	.075^a	-.018	.204^a	.061^a	.074^a	.077^a	.081^a
	(.013)	(.017)	(.021)	(.022)	(.013)	(.013)	(.014)
4-dollar bin × home tract median income	.086^a	.013	.217^a	.036	.083^a	.087^a	.089^a
	(.023)	(.030)	(.038)	(.038)	(.023)	(.023)	(.024)
Yelp rating × home tract median income	.017^b	.018^b	.014	.014	.017^b	.016^b	.020^b
	(.007)	(.009)	(.010)	(.012)	(.007)	(.007)	(.008)
Percent absolute difference in median incomes $(h_i - k_j)$	-.116^b	-.050	-.202^b	-.078	-.119^b	-.069	-.148^a
	(.052)	(.068)	(.083)	(.086)	(.053)	(.053)	(.057)
Percent difference in median incomes $(k_j - h_i)$	.882^a	.941^b	.972^c	1.18^b	.927^a	1.29^a	.785^b
	(.298)	(.376)	(.497)	(.480)	(.300)	(.302)	(.319)
Log median household income in k_j	-.751^a	-.842^a	-.774^c	-1.06^b	-.795^a	-1.19^a	-.673^b
	(.260)	(.325)	(.440)	(.417)	(.262)	(.265)	(.278)
Average annual robberies per resident in k_j	-3.87^a	-2.63^a	-6.82^a	-4.63^a	-3.35^a	-2.94^a	-.178
	(.768)	(.899)	(1.52)	(1.22)	(.768)	(.745)	(.778)
Establishment belongs to chain							.001
							(.103)
Log number of Yelp reviews							.923^a
							(.014)
Difference in tracts' private vehicle ownership						-1.62^a	
						(.128)	
Number of trips	6936	3866	3070	2647	6936	6936	6936

Notes: See notes to Table A.7.

Table A.10: Robustness of spatial frictions

	(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	-.489	-.614	-.989^{a}	-.214	-.714^{b}	-.944^{a}	-1.07^{a}	-1.12^{a}	-1.06^{a}
	(.335)	(.714)	(.184)	(.260)	(.308)	(.160)	(.140)	(.199)	(.084)
Log travel time from home by car	-.796^{b}	-2.29	-1.21^{a}	-1.05^{a}	-.834^{b}	-1.28^{a}	-1.21^{a}	-1.45^{a}	-1.32^{a}
	(.310)	(2.04)	(.195)	(.251)	(.354)	(.158)	(.129)	(.279)	(.093)
Log travel time from work by public transit	-1.14^{b}	-.425	-1.17^{c}	-1.68^{a}	.142	-1.84^{a}	-1.47^{a}	-23.1	-2.21^{a}
	(.476)	(1.64)	(.611)	(.351)	(2.12)	(.644)	(.274)	(12959563.5)	(.771)
Log travel time from work by car	-1.13^{b}	6.49^{c}	-1.41^{a}	-2.10^{a}	-1.16	-2.32^{a}	-1.82^{a}	-2.31^{b}	-1.82^{a}
	(.550)	(3.34)	(.415)	(.399)	(.773)	(.403)	(.305)	(1.12)	(.231)
Log travel time from commute by public transit	-.778^{a}	-2.07	-.881^{a}	-.740^{a}	-.758^{a}	-.812^{a}	-1.07^{a}	-1.31^{a}	-1.09^{a}
	(.248)	(1.46)	(.137)	(.272)	(.265)	(.109)	(.115)	(.297)	(.073)
Log travel time from commute by car	-.873^{a}	-.394	-1.15^{a}	-1.12^{a}	-1.83^{a}	-1.50^{a}	-1.21^{a}	-1.40^{a}	-1.42^{a}
	(.222)	(.709)	(.178)	(.184)	(.485)	(.155)	(.100)	(.211)	(.100)
Log travel time from home by public transit × age 21-39	-1.24	-1.10	-.344						
	(.795)	(1.95)	(.425)						
Log travel time from home by car × age 21-39	-.827	2.41	-.350						
	(.681)	(4.55)	(.448)						
Log travel time from work by public transit × age 21-39	-.229	-3.11	-1.74						
	(.988)	(4.61)	(1.74)						
Log travel time from work by car × age 21-39	-1.06	-29.7^{b}	-1.31						
	(1.24)	(14.8)	(1.05)						
Log travel time from commute by public transit × age 21-39	-.373	2.25	-.540						
	(.512)	(3.27)	(.330)						
Log travel time from commute by car × age 21-39	-.425	-2.63	-.688						
	(.461)	(2.00)	(.421)						
Log travel time from home by public transit × income				-.121^{b}	-.040	-.023			
				(.049)	(.051)	(.018)			
Log travel time from home by car × income				-.017	-.059	-.010			
				(.031)	(.058)	(.017)			
Log travel time from work by public transit × income				.053^{c}	-.523	-.003			
				(.031)	(.779)	(.065)			
Log travel time from work by car × income				.064^{c}	-.110	.040			
				(.038)	(.156)	(.033)			
Log travel time from commute by public transit × income				-.031	-.031	-.035^{a}			
				(.038)	(.043)	(.013)			
Log travel time from commute by car × income				.006	.073	.006			
				(.022)	(.051)	(.015)			
Log travel time from home by public transit × female							-.008	.263	-.124
							(.210)	(.255)	(.115)
Log travel time from home by car × female							.055	.402	-.073
							(.172)	(.332)	(.117)
Log travel time from work by public transit × female							.361	2.88	.503
							(.319)	(12959689.6)	(.811)
Log travel time from work by car × female							.338	.639	-.241
							(.368)	(1.24)	(.335)
Log travel time from commute by public transit × female							.190	.484	-.017
							(.142)	(.321)	(.091)
Log travel time from commute by car × female							.249^{b}	.159	-.029
							(.123)	(.309)	(.121)
Euclidean demographic distance between h_i and k_j	-.956^{a}	-1.88^{a}	-1.20^{a}	-.959^{a}	-1.85^{a}	-1.18^{a}	-1.01^{a}	-1.86^{a}	-1.20^{a}
	(.121)	(.285)	(.130)	(.121)	(.281)	(.130)	(.121)	(.282)	(.130)
Spectral segregation index of k_j	.153^{a}	.063	.043	.155^{a}	.074	.051^{c}	.150^{a}	.052	.046^{c}
	(.051)	(.099)	(.027)	(.051)	(.094)	(.027)	(.051)	(.103)	(.027)
EDD × SSI	-.163	-.133	-.074	-.169	-.174	-.094	-.154	-.141	-.066
	(.116)	(.249)	(.083)	(.116)	(.239)	(.083)	(.117)	(.255)	(.083)
Share of tract population that is Asian	1.02^{a}	.019	.365^{a}	1.02^{a}	.020	.360^{a}	1.04^{a}	-.001	.359^{a}
	(.120)	(.348)	(.138)	(.120)	(.345)	(.137)	(.120)	(.346)	(.138)
Share of tract population that is black	.189	1.03^{b}	.143	.195	1.08^{a}	.126	.245	1.12^{a}	.135
	(.318)	(.403)	(.265)	(.318)	(.399)	(.265)	(.319)	(.399)	(.265)
Share of tract population that is Hispanic	-.258	.429	.412^{b}	-.271	.471	.405^{b}	-.244	.447	.414^{b}
	(.234)	(.385)	(.188)	(.235)	(.381)	(.188)	(.235)	(.384)	(.188)
Share of tract population that is other	.094	3.51	.302	.273	3.52	.475	-.166	3.50	.531
	(2.07)	(3.45)	(1.99)	(2.07)	(3.44)	(1.98)	(2.08)	(3.47)	(1.98)
Number of trips	6447	1079	6936	6447	1079	6936	6447	1079	6936

NOTES: Each column reports an estimated conditional-logit model of the decision to visit a Yelp venue. Standard errors in parentheses. Statistical significance denoted by a (1%), b (5%), c (10%). For brevity, we do not report the following covariates: dollar-bin dummies, rating, cuisine-category dummies, interactions of dollar-bin dummies and rating with home tract median income, percent absolute difference in median incomes, percent difference in median incomes, log median household income in restaurant tract, average annual robberies per resident in restaurant tract, and 28 area dummies.

Table A.11: Estimates with home as only origin

	Home only sample			Estimation sample		
	(1)	(2)	(3)	(4)	(5)	(6)
	Asian	black	white/Hisp	Asian	black	white/Hisp
Log travel time from home by public transit	-1.01^a	-1.47^a	-1.45^a	-.966^a	-1.28^a	-1.29^a
	(.027)	(.087)	(.026)	(.043)	(.065)	(.024)
Log travel time from home by car	-1.33^a	-1.76^a	-1.76^a	-1.12^a	-1.88^a	-1.80^a
	(.043)	(.118)	(.034)	(.045)	(.167)	(.049)
Euclidean demographic distance between h_i and k_j	-.515^a	-1.05^a	-.617^a	-.788^a	-1.18^a	-.667^a
	(.108)	(.332)	(.118)	(.122)	(.294)	(.133)
Spectral segregation index of k_j	.014	-.130	.035	.146^a	.174^b	.056^b
	(.044)	(.230)	(.038)	(.051)	(.073)	(.027)
EDD × SSI	.083	.198	-.308^b	-.136	-.323	-.131
	(.077)	(.388)	(.133)	(.113)	(.210)	(.083)
Share of tract population that is Asian	.610^a	-1.05^b	.120	.972^a	-.061	.152
	(.109)	(.418)	(.126)	(.119)	(.341)	(.137)
Share of tract population that is black	-.439	.992^b	-.922^a	.419	.762^c	.062
	(.278)	(.386)	(.227)	(.312)	(.405)	(.265)
Share of tract population that is Hispanic	-.040	-.813^c	.302^c	-.395^c	.393	.165
	(.190)	(.463)	(.164)	(.233)	(.386)	(.188)
Share of tract population that is other	1.86	1.48	1.46	.272	3.19	1.22
	(1.89)	(3.09)	(1.72)	(2.04)	(3.48)	(1.98)
Dummy for 2-dollar bin	.528^a	.367	.519^a	.345^a	.786^a	.340^a
	(.068)	(.235)	(.077)	(.087)	(.198)	(.083)
Dummy for 3-dollar bin	.310^a	-.597^c	.159	.248^b	-.081	-.048
	(.097)	(.346)	(.108)	(.116)	(.348)	(.119)
Dummy for 4-dollar bin	.236	-.870	.054	.236	.235	-.368^c
	(.154)	(.648)	(.187)	(.186)	(1.24)	(.218)
Yelp rating of restaurant	.645^a	.095	.319^a	.510^a	.065	.324^a
	(.050)	(.149)	(.054)	(.063)	(.139)	(.060)
African cuisine category	-.161	1.19^a	.211	.276	-.202	.287
	(.339)	(.366)	(.240)	(.297)	(.553)	(.261)
American cuisine category	.481^a	.555^a	.630^a	.437^a	.513^a	.610^a
	(.052)	(.123)	(.047)	(.054)	(.119)	(.051)
Asian cuisine category	.951^a	.251^c	.338^a	.899^a	.198	.322^a
	(.052)	(.138)	(.050)	(.054)	(.134)	(.055)
European cuisine category	.264^a	-.046	.301^a	.195^a	-.361^b	.229^a
	(.057)	(.147)	(.051)	(.059)	(.155)	(.056)
Indian cuisine category	.426^a	-.328	.167^c	.372^a	-.409	.003
	(.087)	(.273)	(.088)	(.091)	(.301)	(.098)
Latin American cuisine category	.480^a	.882^a	.729^a	.525^a	.973^a	.690^a
	(.066)	(.145)	(.057)	(.070)	(.137)	(.062)
Middle Eastern cuisine category	.310^a	.573^b	.400^a	.262^a	.125	.209^b
	(.095)	(.222)	(.082)	(.101)	(.251)	(.096)
Vegetarian/vegan cuisine category	.437^a	.741^b	.902^a	.410^a	.014	.616^a
	(.127)	(.288)	(.098)	(.137)	(.410)	(.116)
2-dollar bin × home tract median income	.007	.108^a	.022^b	.038^a	-.020	.043^a
	(.008)	(.037)	(.009)	(.011)	(.032)	(.010)
3-dollar bin × home tract median income	.046^a	.266^a	.081^a	.079^a	.081	.082^a
	(.010)	(.049)	(.012)	(.014)	(.055)	(.013)
4-dollar bin × home tract median income	.064^a	.233^a	.078^a	.071^a	-.212	.098^a
	(.015)	(.086)	(.020)	(.022)	(.237)	(.022)
Yelp rating × home tract median income	-.010^c	.012	.018^a	.019^b	.006	.019^a
	(.005)	(.022)	(.006)	(.008)	(.023)	(.007)
Percent absolute difference in median incomes $(h_i - k_j)$	.098^b	-.030	-.021	.099^c	.908^a	.100^c
	(.045)	(.130)	(.047)	(.050)	(.128)	(.053)
Percent difference in median incomes $(k_j - h_i)$	.367^c	-.424	-.004	-.029	.549	.526^c
	(.222)	(.935)	(.286)	(.297)	(.862)	(.293)
Log median household income in k_j	-.392^b	.669	-.018	.112	-.327	-.365
	(.189)	(.826)	(.252)	(.259)	(.749)	(.255)
Average annual robberies per resident in k_j	-2.78^a	-.276	-4.47^a	-2.26^a	3.66^a	-2.50^a
	(.674)	(1.59)	(.821)	(.667)	(1.20)	(.755)
Number of trips	7415	1096	8961	6447	1079	6936

NOTES: Each column reports an estimated conditional-logit model of the decision to visit a Yelp venue. Columns 1-3 report parameter estimates for a sample of reviewers for whom we do not have workplace locational information. Columns 4-6 repeat the specifications in Table A.3 in which home is the only origin. Standard errors in parentheses. Statistical significance denoted by a (1%), b (5%), c (10%). Unreported controls are 28 area dummies.

Table A.12: Estimates with origin-mode-specific intercepts

	(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)	(2051.5)
Intercept for work by public transit	1.95[a]	2.21	-17.0
	(.552)	(4.85)	(7409.2)
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 $\times$ 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 trips	6447	1079	6936

NOTES: Each column reports an estimated conditional-logit model of the decision to visit a Yelp venue. This specification adds five origin-mode-specific intercepts to the specification in Table 2. Standard errors in parentheses. Statistical significance denoted by a (1%), b (5%), c (10%). For brevity, we do not report the following covariates: dollar-bin dummies, rating, cuisine-category dummies, interactions of dollar-bin dummies and rating with home tract median income, percent absolute difference in median incomes, percent difference in median incomes, log median household income in restaurant tract, average annual robberies per resident in restaurant tract, and 28 area dummies.

Table A.13: 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	.035^{a}
	(.006)
3-dollar bin × home tract median income	.080^{a}
	(.008)
4-dollar bin × home tract median income	.071^{a}
	(.014)
Yelp rating × 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 trips	18015

NOTES: Each column reports an estimated conditional-logit model of the decision 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%). Unreported controls are 28 area dummies.

Table A.14: 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		0.521	0.276	0.251	0.201	0.186
			[.262,.296]	[.238,.278]	[.185,.226]	[.171,.217]
black		0.653	0.322	0.284	0.230	0.214
			[.303,.365]	[.267,.334]	[.209,.278]	[.198,.263]
Hispanic		0.486	0.133	0.102	0.093	0.074
			[.122,.152]	[.095,.127]	[.084,.114]	[.068,.095]
white		0.636	0.182	0.144	0.090	0.068
			[.171,.201]	[.132,.163]	[.084,.108]	[.065,.088]
white or Hispanic		0.470	0.185	0.163	0.115	0.124
			[.173,.213]	[.156,.196]	[.108,.146]	[.116,.159]
Pairwise dissimilarity						
Asian	black	0.796	0.456	0.399	0.326	0.282
			[.432,.495]	[.372,.445]	[.296,.375]	[.257,.339]
Asian	Hispanic	0.584	0.252	0.240	0.177	0.176
			[.238,.275]	[.226,.269]	[.161,.203]	[.159,.210]
Asian	white	0.519	0.232	0.212	0.173	0.169
			[.221,.252]	[.202,.237]	[.158,.198]	[.153,.204]
black	Hispanic	0.558	0.299	0.263	0.227	0.212
			[.282,.347]	[.246,.312]	[.207,.273]	[.197,.261]
black	white	0.822	0.324	0.287	0.216	0.206
			[.307,.370]	[.269,.338]	[.198,.265]	[.192,.256]
Hispanic	white	0.658	0.157	0.113	0.089	0.025
			[.141,.174]	[.095,.133]	[.078,.097]	[.015,.035]

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 tract-level dissimilarity indices based on the coefficient estimates in columns 4-6 of Table 2. 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. Bootstrapped 95% confidence intervals from 496 draws reported in brackets.

Table A.15: Consumption segregation robustness checks

		(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)	(13)
Estimated	Asian	.315	.316	.317	.299	.326	.364	.299	.381	.344	.315	.326	.330	.322
	black	.352	.353	.354	.331	.366	.380	.353	.434	.396	.355	.371	.384	.348
	Hispanic	.142	.143	.143	.123	.144	.147	.131	.189	.164	.147	.146	.151	.155
	white	.190	.192	.190	.170	.201	.215	.174	.236	.213	.192	.192	.200	.200
	white or Hispanic	.205	.208	.206	.194	.219	.243	.205	.267	.241	.209	.214	.230	.211
No spatial friction	Asian	.290	.290	.291	.259	.296	.337	.271	.381	.324	.287	.297	.307	.290
	black	.322	.324	.324	.296	.329	.366	.323	.425	.376	.324	.347	.356	.322
	Hispanic	.114	.114	.114	.090	.120	.128	.110	.170	.139	.111	.121	.125	.114
	white	.153	.153	.153	.124	.158	.177	.135	.208	.172	.154	.157	.164	.153
	white or Hispanic	.189	.191	.189	.171	.196	.234	.191	.267	.225	.189	.201	.213	.189
No social friction	Asian	.245	.243	.250	.232	.247	.291	.243	.309	.269	.246	.241	.268	.245
	black	.273	.274	.278	.270	.275	.284	.276	.351	.314	.276	.292	.317	.269
	Hispanic	.106	.108	.106	.100	.105	.116	.097	.153	.121	.113	.109	.118	.110
	white	.112	.114	.113	.116	.112	.123	.109	.147	.124	.118	.114	.127	.112
	white or Hispanic	.150	.152	.151	.149	.156	.175	.153	.208	.182	.154	.158	.178	.156
Neither friction	Asian	.232	.231	.237	.211	.230	.277	.219	.323	.256	.228	.225	.256	.232
	black	.260	.262	.264	.252	.264	.284	.261	.359	.310	.263	.280	.302	.260
	Hispanic	.088	.090	.088	.079	.091	.104	.083	.143	.104	.085	.094	.100	.088
	white	.093	.095	.094	.090	.097	.114	.091	.141	.112	.100	.097	.107	.093
	white or Hispanic	.156	.159	.157	.146	.164	.193	.156	.230	.190	.156	.163	.179	.156

NOTES: This table reports dissimilarity indices computed using estimated preference parameters, as in the upper panel of Table 6. Column 1 uses the coefficients from columns four to six of Table 2, as in Table 6. Columns 2 and 3 use columns one to three and four to six of Table D.3, respectively. Column 4 uses the coefficients from the specification with origin-mode-specific intercepts reported in Table A.12. Columns 5 and 6 use coefficients from specifications in which observations are limited to the first half and first fifth of NYC restaurant reviews posted by each reviewer, respectively. Column 7 and 8 use coefficients from specifications in which the estimation sample is split into reviewers with residences located using information contained in one or two reviews in column 7 and three or more reviews in column 8. Column 9 uses the coefficients from a specification that restricts the sample to Yelp reviewers with later-than-median dates of first review written. Column 10 uses the coefficients from a specification that controls for origin-destination differences in vehicle ownership. Column 11 uses the coefficients from a specification that controls for the number of Yelp reviews of each restaurant and its membership in a chain with more than eight NYC locations. Column 12 uses the coefficients from a specification that introduces dummies for 39 more disaggregated cuisine categories. Column 13 uses the coefficients from columns four to six of Table 2 but constrains all trips to originate at home. See Tables A.4 through A.6 for these coefficients.

Table A.16: Residential and consumption segregation (minimum-time specification)

		Residential dissimilarity	Consumption dissimilarity			
			Estimated	No spatial	No social	Neither friction
		(1)	(2)	(3)	(4)	(5)
Dissimilarity index						
Asian		0.521	0.319	0.296	0.247	0.237
			[.308,.340]	[.286,.319]	[.234,.270]	[.224,.263]
black		0.653	0.356	0.328	0.268	0.262
			[.333,.398]	[.312,.376]	[.254,.313]	[.250,.311]
Hispanic		0.486	0.153	0.124	0.109	0.090
			[.144,.171]	[.116,.148]	[.101,.127]	[.084,.111]
white		0.636	0.202	0.163	0.116	0.095
			[.191,.220]	[.152,.181]	[.110,.132]	[.090,.114]
White or Hispanic		0.470	0.209	0.195	0.149	0.160
			[.199,.239]	[.185,.229]	[.141,.179]	[.151,.195]
Pairwise dissimilarity						
Asian	black	0.796	0.499	0.454	0.384	0.358
			[.477,.535]	[.436,.493]	[.365,.429]	[.341,.405]
Asian	Hispanic	0.584	0.295	0.281	0.223	0.222
			[.282,.317]	[.268,.307]	[.210,.248]	[.209,.250]
Asian	white	0.519	0.282	0.260	0.217	0.209
			[.269,.304]	[.248,.284]	[.204,.240]	[.196,.236]
black	Hispanic	0.558	0.332	0.306	0.258	0.252
			[.314,.375]	[.288,.357]	[.243,.303]	[.240,.302]
black	white	0.822	0.361	0.330	0.258	0.256
			[.340,.406]	[.315,.380]	[.244,.304]	[.243,.307]
Hispanic	white	0.658	0.179	0.131	0.105	0.038
			[.163,.195]	[.113,.149]	[.097,.116]	[.028,.049]

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 column 1 of Tables A.7 - A.9. 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. Bootstrapped 95% confidence intervals from 500 draws reported in brackets.

Table A.17: 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	0.098	0.114	0.107	0.123
Hispanic	0.066	0.339	0.332	0.372	0.360
black	0.023	0.047	0.054	0.062	0.070
white	0.810	0.517	0.501	0.458	0.447
		Community District 10: Central Harlem			
Asian	0.024	0.048	0.061	0.115	0.150
Hispanic	0.222	0.156	0.104	0.263	0.184
black	0.630	0.739	0.763	0.476	0.470
white	0.095	0.058	0.072	0.146	0.196
		Community District 11: East Harlem			
Asian	0.056	0.074	0.102	0.126	0.157
Hispanic	0.494	0.545	0.431	0.462	0.350
black	0.309	0.169	0.187	0.113	0.119
white	0.120	0.211	0.279	0.299	0.374

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 6.

Table A.18: Demographics of residents and consumers in lower Manhattan and west Brooklyn

		Consumption share			
	Residential share	Estimated	No Spatial	No Social	Neither
		Community District 2, Brooklyn: Williamsburg			
Asian	0.066	0.051	0.055	0.062	0.066
black	0.189	0.349	0.329	0.329	0.297
Hispanic	0.164	0.252	0.280	0.263	0.295
white	0.546	0.348	0.336	0.345	0.342
		Community District 3, Brooklyn: Bedford-Stuyvesant			
Asian	0.021	0.037	0.052	0.079	0.096
black	0.657	0.697	0.601	0.435	0.315
Hispanic	0.147	0.147	0.194	0.235	0.288
white	0.149	0.119	0.154	0.251	0.302
		Community District 1, Brooklyn: Greenpoint			
Asian	0.050	0.066	0.068	0.082	0.084
black	0.052	0.293	0.295	0.316	0.306
Hispanic	0.272	0.334	0.314	0.307	0.291
white	0.608	0.307	0.324	0.294	0.319
		Community District 3, Manhattan			
Asian	0.321	0.187	0.194	0.160	0.161
black	0.071	0.163	0.168	0.184	0.187
Hispanic	0.253	0.344	0.321	0.327	0.309
white	0.332	0.306	0.317	0.329	0.342

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 6.

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 individuals' 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.[1]

B.2 Yelp reviewers 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 reviewers in three rounds. In the first round, we examined all reviews written by a randomly selected 25% subsample of the 50,000 reviewers. In the second round, we selectively examined the remaining 75% of reviewers relying on the first round's lessons for successfully locating users. In the third round, we intensively examined reviews by a set of black reviewers in the

[1]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.

remaining 75%. The final dataset used for estimation contains only those reviewers 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 reviewers 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}.Of the 230,000 reviews analyzed in the first-round sample, 16,425 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 1,500 reviewers with a residential location, 575 reviewers with a workplace, and 450 reviewers with both home and work locations. Thus, we have locational information for nearly one-fifth of the Yelp users we examined. The median reviewer for which we obtained locational information had reviewed twenty venues in New York and New Jersey, while the median reviewer for which we obtained no information had reviewed five venues. Among 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 reviewers who appear in our estimation sample.

B.2.2 Yelp reviewers 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 reviewers, work locations for 38 reviewers, and both locations for 304 reviewers. After imposing the previously mentioned 1.5-mile proximity, non-mover, and tract-covariate-availability restrictions, this second round yielded 165 reviewers who appear in our estimation sample.

B.2.3 Yelp reviewers 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 reviewers 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 reviewers, work locations for 91 reviewers, and both locations for 51 reviewers. After imposing the previously mentioned 1.5-mile proximity and tract-covariate-availability restrictions, this third round yielded 31 reviewers (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 2,168 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 2,110 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 2,105) and Rikers Island (inmate population of 11,091).

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, Xu and Stults, 2014).

We aggregate New York City's 59 community boards to define 28 areas. 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 $\{j_{i1}, j_{i2}, \ldots, j_{iT_i}\}$ conditional on observing at least one review written by i in each of the periods $\{1, \ldots, T_i\}$ and on the randomly drawn sets $\{S_{i1}, \ldots, S_{iT_i}\}$ may be written as

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

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

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

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.[2] 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

[2] 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:

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

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	-0.94	-1.09	-1.02	-1.02
	(0.01)	(0.01)	(0.01)	(0.01)	(0.02)
rating	1.00	0.96	1.02	0.99	0.98
	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)
Choice set	All restaurants J	All restaurants J	Never reviewed J'_{iT_i}	Not previously reviewed J'_{it}	Not previously reviewed $S_{it} \in J'_{it}$
Observations	16000	7518	7518	7518	7518
Choice set size	1000	1000	i-specific	it-specific	20

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.

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.

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 13. 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 the first column of Tables A.7-A.9. The results are broadly similar to those in Table 2.

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 footnote 13 and, for every individual i, restaurant j and origin-mode 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 $i = 1, \dots, N$ 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 $\gamma^1_{li} = \gamma^1_{g(i)l} + \varepsilon^1_{li}$, $\gamma^2_i = \gamma^2_{g(i)} + \varepsilon^2_i$, $\beta^1_i = \beta^1_{g(i)} + \varepsilon^3_i$, $\beta^2_i = \beta^2_{g(i)} + \varepsilon^4_i$ 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 $\mathbb{E}_{i|g(i)}[\varepsilon^1_{li}] = \mathbb{E}_{i|g(i)}[\varepsilon^2_i] = \mathbb{E}_{i|g(i)}[\varepsilon^3_i] = \mathbb{E}_{i|g(i)}[\varepsilon^4_i] = 0$, where $\mathbb{E}_{i|g(i)}[\cdot]$ denotes the expectation across individuals in the population of interest belonging to race or ethnicity $g(i)$, 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)).[3]

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

[3] Dickstein and Morales (2015) 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 (2015) cannot handle the large choice sets that consumers face in our empirical application, it is not ideal for our setting.

like the time that it takes to travel to a venue or the average price of each restaurant.[4] 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, Sheu and Zahler (2015); Dickstein and Morales (2015); Wollman (2015)), 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.[5] 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

[4]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.

[5]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.

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 a constant, and $\mathcal{T}$ is the set of periods at which individual i will visit a restaurant.

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 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_j^1 or Z_{ij}^2 through the following function

$$P(d_{ijt}^{r*} | d_{ijt} = 1, X_i, Z_i, J, J_{it}'; (\gamma, \beta, p_{it}^*)) = p_{it}^* \mathbb{1}\{j \neq 0, j \in J_{it}'\} \exp(\zeta_{g(i)}^1 Z_j^1 + \zeta_{g(i)}^2 Z_{ij}^2). \quad \text{(C.8)}$$

In the case when $\zeta_g^1 = \zeta_g^2 = 0$ for every group g this function is identical to that assumed in the main text (see equation (6)). Conversely, if either ζ_g^1 or ζ_g^2 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:

$$P(d_{ijt}^* = 1 | X_i, Z_i, J, J_{it}'; (\gamma, \beta, p_{it}^*)) = p_{it}^* \mathbb{1}\{j \neq 0, j \in J_{it}'\}$$
$$\times \frac{\left(\sum_{l \in \mathcal{L}} \exp(\gamma_{g(i)l}^1 X_{ijl}^1 + \gamma_{g(i)}^2 X_{ij}^2 + (\beta_{g(i)}^1 + \zeta_{g(i)}^1) Z_j^1 + (\beta_{g(i)}^2 + \zeta_{g(i)}^2) Z_{ij}^2) \right)}{\sum_{j' \in J_t} \left(\sum_{l \in \mathcal{L}} \exp(\gamma_{g(i)l}^1 X_{ij'l}^1 + \gamma_{g(i)}^2 X_{ij'}^2 + (\beta_{g(i)}^1 + \zeta_{g(i)}^1) Z_{j'}^1 + (\beta_{g(i)}^2 + \zeta_{g(i)}^2) Z_{ij'}^2) \right)}, \quad \text{(C.9)}$$

This expression demonstrates that we cannot separately identify the parameter vectors $\beta_{g(i)}^1$ and $\beta_{g(i)}^2$ from the parameter vectors $\zeta_{g(i)}^1$and $\zeta_{g(i)}^2$. 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_j^1 and Z_{ij}^2. Furthermore, the probability in equation (C.9) is identical to that in equation (6) with the only exception that the parameter vectors

$$\tilde{\beta}_{g(i)}^1 = \beta_{g(i)}^1 + \zeta_{g(i)}^1 \qquad \text{and} \qquad \tilde{\beta}_{g(i)}^2 = \beta_{g(i)}^2 + \zeta_{g(i)}^2$$

take the place of the parameter vectors $\beta_{g(i)}^1$ and $\beta_{g(i)}^2$ 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_j^1 and Z_{ij}^2 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 individuals' unobserved restaurant preferences (captured in the vector ν_{it}) are independent over time, the fact that we identify reviewers' preferences from their Yelp reviews and 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 fact that we identify reviewers' 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 $\tilde{\nu}_{ijlt} \equiv a \times \omega_{ij} + \nu_{ijlt}$ and where a is a constant that governs the importance 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:

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

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:

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

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 usersǵ reviews, respectively.

where both the permanent shock ω_{ij} and the transitory shock ν_{ijt} have type I 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 one through three, the choice set contains 1000 restaurants and the standard deviation of the permanent shock has the same magnitude as the transitory shock, $a = 1$. In column one, we apply the to estimator the full sample; in column two, the first half of each user's reviews; in column three, 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 four through six, 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 seven through nine, 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 review venues whose other reviewers are also members of group g. The first summation is the review-weighted average of the share of a venue's reviewers who are members of g, using g members'

Table D.1: Isolation indices for various model specifications

Isolation index	Data	Model predictions Pooled	Race-specific	Nested 1	Nested 2	Minimum time
Asian	.087	[-.015, .018]	[.057, .089]	[.055, .089]	[.058, .09]	[.056, .087]
Black	.087	[-.011, .024]	[.042, .093]	[.044, .092]	[.045, .093]	[.042, .09]
White/Hispanic	.045	[-.02, .013]	[.025, .056]	[.025, .058]	[.025, .058]	[.024, .056]

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 computed from 500 generated samples of the same size. The pooled model is Table A.13. The race-specific model is columns four to six of Table 2. Nested 1 and Nested 2 are columns one to three and four to six of Table D.3, respectively. Minimum time is the first column of Tables A.7-A.9.

reviews 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 the average g exposure of non-members.[6]

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 (columns four through six of Table 2), two nested-logit specifications (Table D.3), and a minimum-travel-time specification (Tables A.7-A.9). 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 reviewers 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 individuals' 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 visits, even if the "true" value of $gap_p = 0$. We therefore compare the observed gap_p to the null hypothesis of a distribution in which every individual that visited one restaurant in a pair $p = \{j, j'\}$ such

[6]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 $\sum_j \frac{v_{gj}}{v_g} \cdot \left(\frac{v_{gj}}{v_j}\right)$.

that $X^1_{ijl} = X^1_{ij'l}$, $X^2_{ij} = X^2_{ij'}$, $Z^1_j = Z^1_{j'}$, and $Z^2_{ij} = Z^2_{ij'}$ is randomly assigned to one of the two restaurants within the pair. When performing this randomization, we condition on the total number of reviews observed for each of the restaurants in the pair. For example, if there are 20 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 4,569 venues in sets of venues that have the same tract-cuisine-price-rating quadruplet. There are 402 venues with between 10 and 40 reviews in sets of venues that have the same tract-cuisine-price-rating quadruplet. We photo-coded all the users who reviewed restaurants in a sample of 125 pairs of such venues. This yields the results depicted in Figure 6.

In addition, we compute the contribution of these pairs of restaurants to Gentzkow and Shapiro (2011) isolation indices for this set of restaurants. Define the within-pair difference

$$\left[\frac{v_{gj}}{v_g}\cdot\left(\frac{v_{gj}-1}{v_j-1}\right)-\frac{v_{\neg gj}}{v_{\neg g}}\cdot\left(\frac{v_{\neg gj}}{v_j-1}\right)\right]-\left[\frac{v_{gj'}}{v_g}\cdot\left(\frac{v_{gj'}-1}{v_{j'}-1}\right)-\frac{v_{\neg gj'}}{v_{\neg g}}\cdot\left(\frac{v_{\neg gj'}}{v_{j'}-1}\right)\right]\quad j,j'\in p$$

where v_g is the total visits by members of group g to all restaurants that belong to one of these pairs. Figure D.1 compares the distribution of these pair-level differences to those obtained under the null hypothesis. The mean of the difference for Asian consumers in the data is .0029 and under the null is .0028, .0027 and .0023 for black consumers, and .0028 and .0028 for white/Hispanic consumers. The p-values for the one-sided tests of equal means are .3331, .0698, and .4982, respectively.

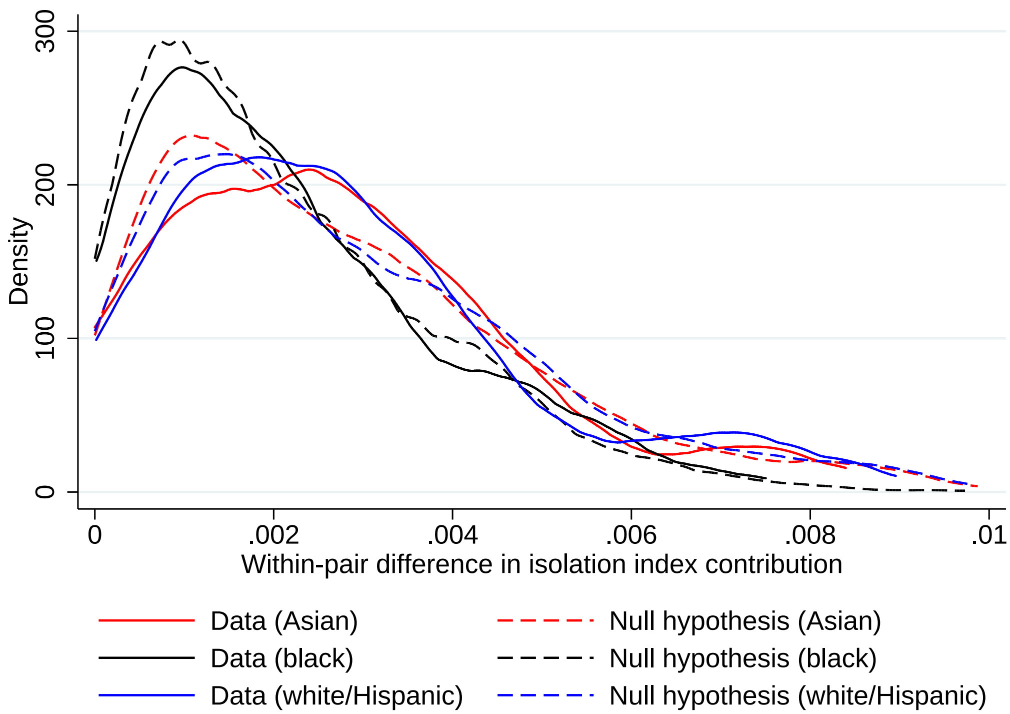
D.3 Restaurant fixed effects

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, this assumption implies that the expected utility of a restaurant trip is equal across all individuals within a racial group. Table D.2 reports estimates. They are generally congruent with the results in Tables 2 and 4. 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 4 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 reviewers, and in

Figure D.1: Isolation-index elements for each race in data and under null



NOTES: These kernel densities depict the distribution of differences in pairs of restaurants' contributions to the Gentzkow and Shapiro (2011) isolation index using leave-out means and 125 pairs of restaurants that are identical in terms of their cuisine category, price category, Yelp rating, and census tract. See appendix D.2 for details. The null hypothesis, in line with our model, is that individuals are randomly assigned to one of the two restaurants within each pair.

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

	(1) Asian	(2) black	(3) white/Hisp
Log travel time from home by public transit	-.995[a]	-.765[a]	-.982[a]
	(.104)	(.093)	(.047)
Log travel time from home by car	-1.04[a]	-.908[a]	-1.18[a]
	(.082)	(.100)	(.049)
Log travel time from work by public transit	-1.11[a]	-1.70[a]	-1.50[a]
	(.128)	(.585)	(.144)
Log travel time from work by car	-1.35[a]	-1.61[a]	-1.58[a]
	(.135)	(.316)	(.103)
Log travel time from commute by public transit	-.930[a]	-1.06[a]	-1.08[a]
	(.080)	(.163)	(.050)
Log travel time from commute by car	-.903[a]	-1.69[a]	-1.45[a]
	(.061)	(.338)	(.067)
Euclidean demographic distance between h_i and k_j	-.464[a]	-1.29[a]	-.831[a]
	(.087)	(.199)	(.103)
EDD $\times$ SSI	-.725[a]	-1.06[a]	-.084
	(.154)	(.374)	(.090)
2-dollar bin $\times$ home tract median income	.034[a]	-.025	.041[a]
	(.010)	(.028)	(.009)
3-dollar bin $\times$ home tract median income	.067[a]	.027	.078[a]
	(.013)	(.048)	(.012)
4-dollar bin $\times$ home tract median income	.058[a]	-.145	.082[a]
	(.020)	(.208)	(.021)
Yelp rating $\times$ home tract median income	.006	-.015	.009[b]
	(.005)	(.014)	(.004)
Percent absolute difference in median incomes $(h_i - k_j)$	-.242[a]	.929[a]	-.299[a]
	(.049)	(.142)	(.051)
Percent difference in median incomes $(k_j - h_i)$	.488[a]	-.304	.855[a]
	(.124)	(.337)	(.114)

NOTES: Each column reports an estimated conditional-logit model of individuals' decisions to visit a Yelp venue. Estimates computed per Taddy (2015) making the assumption that expected utility of a restaurant trip is equal across all individuals within a racial group. Standard errors in parentheses. Statistical significance denoted by *a* (1%), *b* (5%), *c* (10%). The unreported covariates are restaurant fixed effects.

favor of the cuisine-rating-area nesting in the case of white/Hispanic reviewers. 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 2.

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 2 to generate 500 samples of observations (equal in size to our estimation sample). 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.2 and D.3 depict these for the main specifications reported in Table 2; Figures D.4 and D.5 for the minimum-travel-time specifications in the first column of Tables A.7-A.9. In each figure, the solid red line shows the bootstrapped distribution of estimates and the dashed blue line depicts the asymptotic distribution. Figures D.2 and D.4 depict the distributions for the coefficients on our key social-friction covariates and two restaurants characteristics. For these coefficients, the bootstrap distribution is very close to the asymptotic distribution. For the spatial frictions depicted in Figure D.3, 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.5 shows that these outlying estimates do not arise if we assume that reviewers visit restaurants using the minimum-time origin-mode pair available to them rather than optimizing over the six origin-mode pairs. This suggests that the mismatch between finite-sample and the asymptotic 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}.

We further employ these 500 estimated parameter vectors to assess model fit. We do so in two ways. First, we use the average parameter vector over the 500 bootstrapped values to compute confidence intervals for isolation indices akin to those in Table 3. This confidence interval summarizes the distribution of isolation indices generated by the parameter values and the distribution of the unobserved preference shocks (the term ν_{ijlt} in equation (1)) across the individuals, restaurants, origin-mode pairs and periods in our sample. The confidence intervals for isolation indices predicted by the average of the bootstrapped parameters reported in Table D.4 are very similar to those in Table 3. Second, we compute the isolation-index confidence intervals for each of the 500 estimated parameter vectors. Figure D.6 shows that the distributions of the endpoints of these 90% confidence intervals are nearly centered around the data-generating process's values for these endpoints.

We employ these bootstrapped distributions of parameter estimates to compute confidence intervals for the dissimilarity indices reported in Table 6. Figure D.7 depicts the distributions of dissimilarity indices resulting from the bootstrapped distributions of parameter estimates. The gap between the estimated dissimilarity index and the mean of the bootstrapped distribution is typically less than a standard deviation, as reported in Table D.5.

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	1.09^a (.014)	1.01^a (.034)	1.12^a (.013)	1.09^a (.018)	.876^a (.055)	.983^a (.019)
Log travel time from home by public transit	-1.09^a (.113)	-.941^a (.128)	-1.16^a (.062)	-1.06^a (.109)	-.924^a (.119)	-1.13^a (.059)
Log travel time from home by car	-1.21^a (.098)	-1.19^a (.160)	-1.42^a (.067)	-1.17^a (.094)	-1.15^a (.142)	-1.35^a (.059)
Log travel time from work by public transit	-1.31^a (.165)	-25.4 (792989.8)	-2.03^a (.344)	-1.29^a (.163)	-1.69^c (.895)	-1.85^a (.282)
Log travel time from work by car	-1.72^a (.206)	-1.77^a (.425)	-2.13^a (.212)	-1.65^a (.190)	-1.66^a (.392)	-1.93^a (.169)
Log travel time from commute by public transit	-.981^a (.075)	-.928^a (.103)	-1.14^a (.049)	-.955^a (.073)	-.920^a (.095)	-1.09^a (.043)
Log travel time from commute by car	-1.08^a (.067)	-1.32^a (.178)	-1.50^a (.067)	-1.04^a (.064)	-1.31^a (.166)	-1.43^a (.057)
Euclidean demographic distance between h_i and k_j	-1.03^a (.125)	-1.84^a (.281)	-1.21^a (.137)	-.995^a (.121)	-1.84^a (.279)	-1.20^a (.130)
Spectral segregation index of k_j	.154^a (.053)	.075 (.093)	.037 (.030)	.146^a (.051)	.078 (.092)	.045^c (.027)
EDD × SSI	-.158 (.121)	-.174 (.240)	-.062 (.087)	-.156 (.118)	-.165 (.232)	-.066 (.083)
Share of tract population that is Asian	1.11^a (.127)	.015 (.347)	.385^a (.147)	.988^a (.120)	.054 (.346)	.368^a (.138)
Share of tract population that is black	.222 (.337)	1.09^a (.403)	.159 (.283)	.258 (.319)	1.03^a (.398)	.130 (.265)
Share of tract population that is Hispanic	-.260 (.248)	.472 (.384)	.392^c (.201)	-.167 (.236)	.406 (.382)	.402^b (.188)
Share of tract population that is other	.336 (2.19)	3.60 (3.44)	.426 (2.14)	.008 (2.08)	3.69 (3.40)	.500 (1.98)
Dummy for 2-dollar bin	.405^a (.092)	.776^a (.199)	.388^a (.088)	.377^a (.087)	.766^a (.197)	.354^a (.083)
Dummy for 3-dollar bin	.329^a (.122)	-.090 (.342)	-.027 (.128)	.323^a (.117)	-.125 (.341)	-.032 (.120)
Dummy for 4-dollar bin	.267 (.198)	-.067 (1.23)	-.339 (.237)	.262 (.189)	-.124 (1.22)	-.355 (.221)
Yelp rating of restaurant	.579^a (.066)	.052 (.138)	.334^a (.062)	.594^a (.065)	.055 (.134)	.343^a (.059)
African cuisine category	.384 (.300)	-.192 (.553)	.427 (.263)	.336 (.299)	-.212 (.552)	.290 (.262)
American cuisine category	.438^a (.055)	.525^a (.120)	.606^a (.051)	.433^a (.055)	.515^a (.119)	.590^a (.050)
Asian cuisine category	.881^a (.055)	.256^c (.135)	.308^a (.055)	.871^a (.055)	.261^c (.134)	.309^a (.054)
European cuisine category	.163^a (.060)	-.326^b (.154)	.205^a (.056)	.174^a (.060)	-.311^b (.154)	.239^a (.056)
Indian cuisine category	.378^a (.092)	-.449 (.301)	-.031 (.098)	.332^a (.092)	-.426 (.301)	-.032 (.098)
Latin American cuisine category	.556^a (.070)	1.01^a (.137)	.727^a (.062)	.536^a (.070)	.991^a (.136)	.687^a (.062)
Middle Eastern cuisine category	.313^a (.101)	.108 (.251)	.248^a (.095)	.297^a (.101)	.072 (.251)	.199^b (.094)
Vegetarian/vegan cuisine category	.307^b (.147)	-.004 (.412)	.519^a (.124)	.341^b (.143)	.061 (.388)	.593^a (.115)
2-dollar bin × home tract median income	.035^a (.011)	-.022 (.032)	.045^a (.010)	.034^a (.011)	-.023 (.032)	.042^a (.009)
3-dollar bin × home tract median income	.077^a (.014)	.078 (.054)	.086^a (.014)	.075^a (.014)	.075 (.053)	.081^a (.013)
4-dollar bin × home tract median income	.075^a (.023)	-.169 (.236)	.099^a (.024)	.074^a (.022)	-.167 (.233)	.095^a (.023)
Yelp rating × home tract median income	.012 (.008)	.008 (.023)	.018^a (.007)	.011 (.008)	.008 (.022)	.016^b (.007)
Percent absolute difference in median incomes ($h_i - k_j$)	-.061 (.052)	.855^a (.128)	-.102^c (.056)	-.051 (.050)	.834^a (.126)	-.101^c (.053)
Percent difference in median incomes ($k_j - h_i$)	.131 (.321)	.622 (.857)	.766^b (.321)	.206 (.306)	.605 (.853)	.706^b (.300)
Log median household income in k_j	-.104 (.282)	-.362 (.747)	-.638^b (.280)	-.155 (.268)	-.371 (.744)	-.619^b (.262)
Average annual robberies per resident in k_j	-3.72^a (.727)	2.43^b (1.21)	-4.26^a (.850)	-3.38^a (.675)	2.56^b (1.20)	-3.74^a (.771)
χ^2 test p-value	.000	.852	.000	.000	.022	.371
Number of trips	6447	1079	6936	6447	1079	6936

NOTES: Each column reports an estimated nested-logit model of the decision 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 entry in columns four to six of Table 2.

Figure D.2: Parametric bootstrap: Social frictions and restaurant characteristics

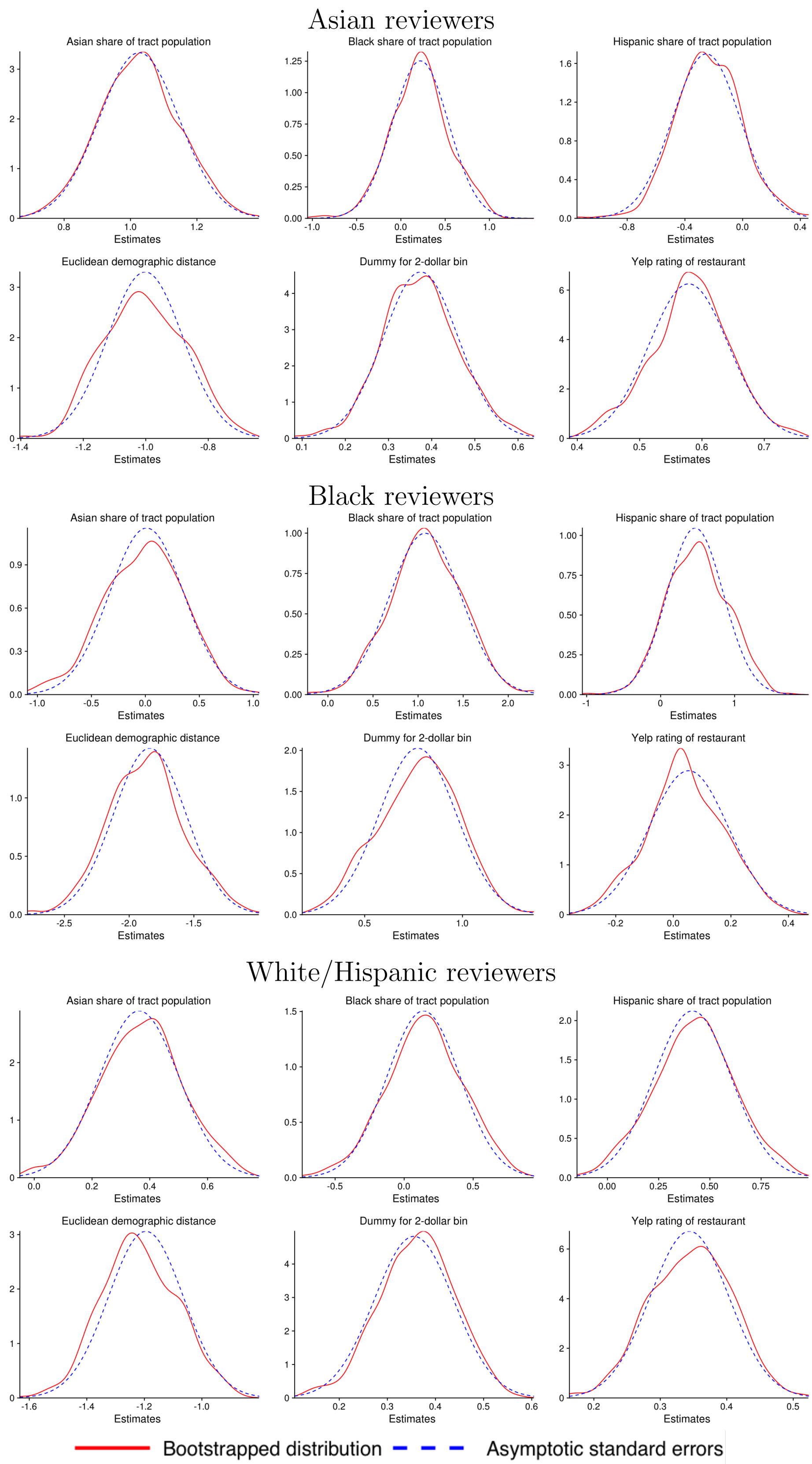


Figure D.3: Parametric bootstrap: Spatial frictions

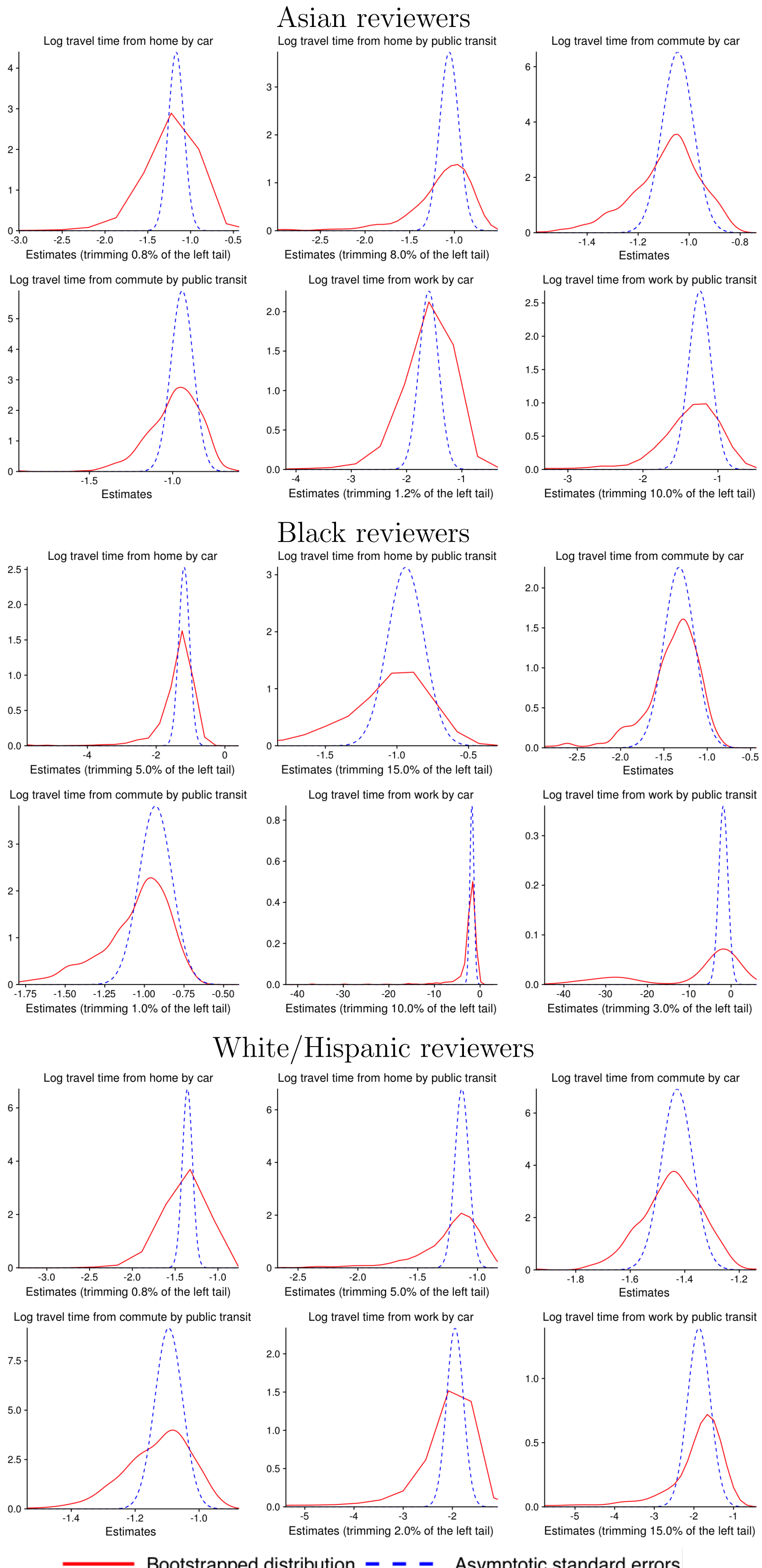


Figure D.4: Parametric bootstrap: Social frictions and restaurant characteristics in minimum-time specification

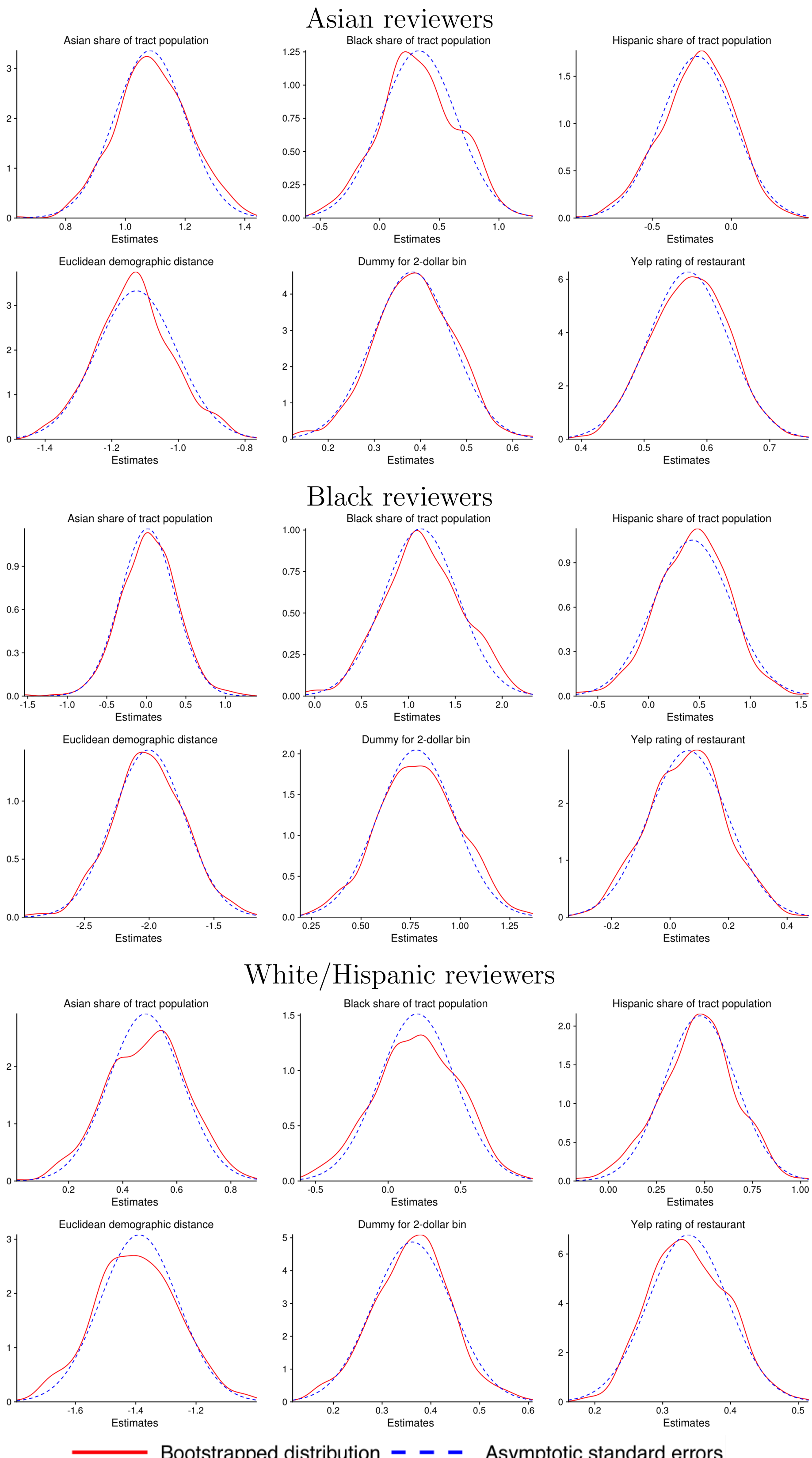


Table D.4: Parametric bootstrap: Isolation index confidence intervals

	Estimation sample	Bootstrap average confidence interval
Asian isolation index	.087	[.057 , .091]
Black isolation index	.087	[.042 , .088]
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 predicted outcomes for generated samples of the same size using the average of the bootstrapped distribution of parameters. Isolation indices as defined in Gentzkow and Shapiro (2011).

Figure D.5: Parametric bootstrap: Spatial frictions in minimum-time specification

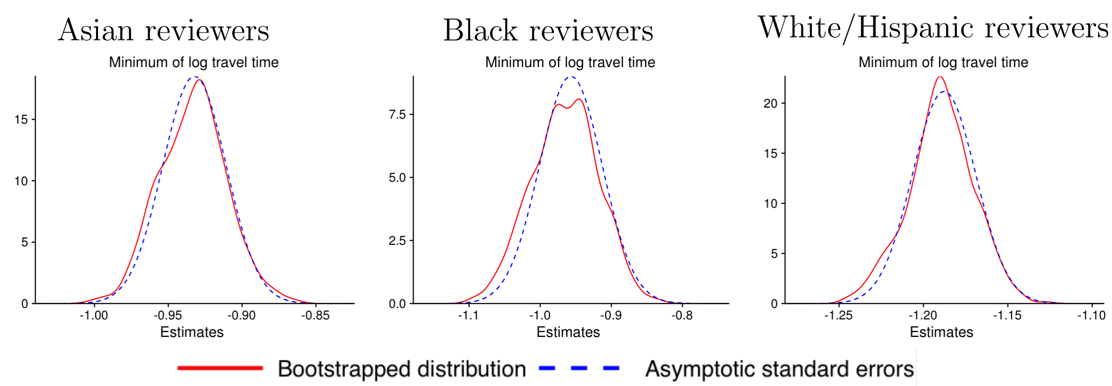
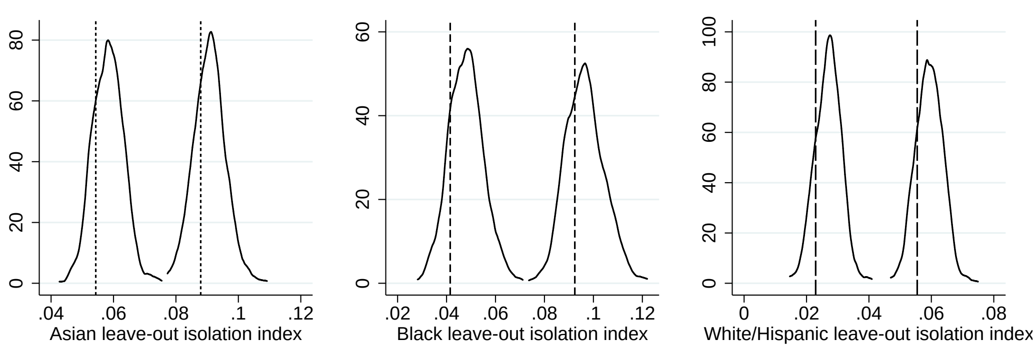
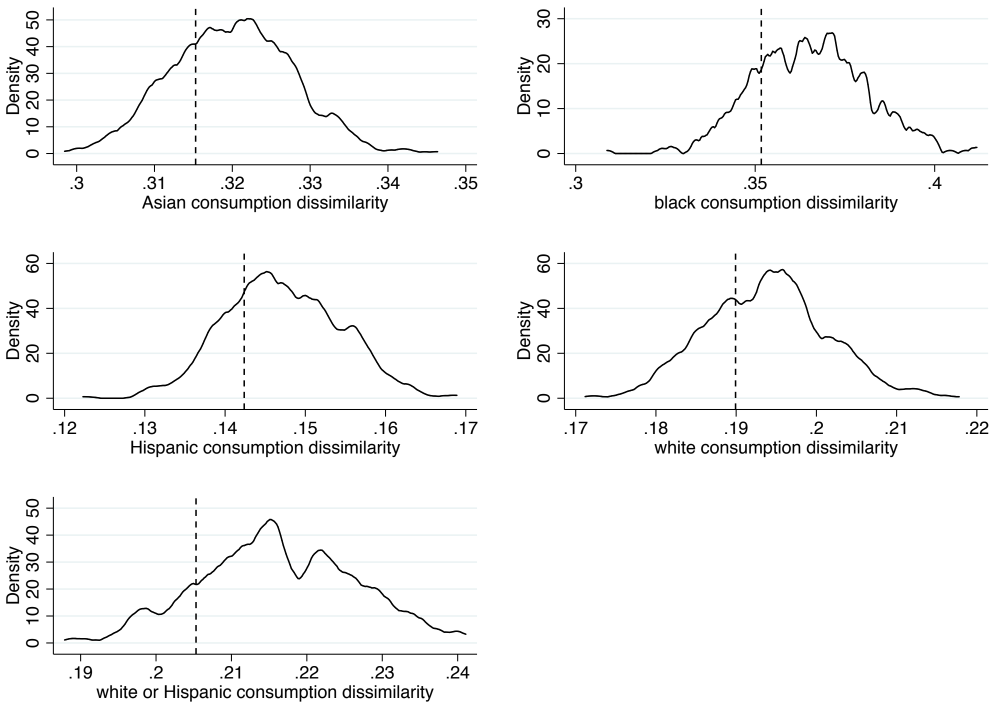


Figure D.6: Parametric bootstrap: Isolation index confidence intervals



NOTES: Each panel depicts the distribution of the endpoints of the 90% confidence interval for the isolation index obtained from the 496 bootstrap draws. The vertical lines depict the endpoints of the 90% confidence interval reported in Table 3.

Figure D.7: Dissimilarity indices for bootstrapped distribution of estimates



NOTES: Epanechnikov kernel with bandwidth of .001. Vertical lines denote estimated values reported in Table 6, upper panel, column 2.

Table D.5: Bootstrapped dissimilarity indices

Race	Estimate	Difference	Standard deviation
Asian	.315	-.005	.008
black	.352	-.014	.016
Hispanic	.142	-.005	.007
white	.19	-.004	.008
whiteorHispanic	.205	-.011	.01

NOTES: First column reports estimate from Table 6. Second column reports estimate minus the average of the bootstrapped distribution of dissimilarity indices. The third column reports the standard deviation of the bootstrapped distribution.

E Consumption segregation 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.[7] Using the definition of conditional and marginal probabilities, we can write

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

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 $P(d_{ij} = 1, \mathfrak{g}_i = \mathfrak{g})$ as a function of $P(d_{ij} = 1|h_i = h, w_i = w, \mathfrak{g}_i = \mathfrak{g})$, 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

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

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 probability distributions, we can write

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

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

$$\begin{aligned}
&P(d_{ij} = 1, \mathfrak{g}_i = \mathfrak{g}, h_i = h, w_i = w) = \\
&\quad = P(d_{ij} = 1|\mathfrak{g}_i = \mathfrak{g}, h_i = h, w_i = w) P(\mathfrak{g}_i = \mathfrak{g}|h_i = h) P(w_i = w|h_i = h) P(h_i = h).
\end{aligned} \tag{E.3}$$

Using the demand model in Section 3 and, specifically, the functional-form assumption in equation (5), we can write the probability $P(d_{ij} = 1|\mathfrak{g}_i = \mathfrak{g}, h_i = h, w_i = w)$ as a function of the parameter estimates presented in Section 4. The probabilities $P(\mathfrak{g}_i = \mathfrak{g}|h_i = h)$, $P(w_i = w|h_i = h)$ and $P(h_i = h)$ may all be computed using data from the US Census Bureau. Specifically, $P(\mathfrak{g}_i = \mathfrak{g}|h_i = h)$ 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

[7] 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.

overall population that lives in census tract h.[8] Combining equations (E.1), (E.2) and (E.3) we can write

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

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.1 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.

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.1 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

[8]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. 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.1: Second Avenue Subway expansion

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

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 two of Table 7.

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),

$$\begin{aligned} U_i &= \ln\left(\sum_j \sum_l \exp\left(\gamma X_{ijl} + \beta Z_{ij}\right)\right) + c, \\ U'_i &= \ln\left(\sum_j \sum_l \exp\left(\gamma X'_{ijl} + \beta Z'_{ij}\right)\right) + c, \end{aligned} \tag{E.4}$$

where c is an arbitrary constant. Taking the difference between these two utilities yields

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

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 tract that experiences gentrification, so there should be no correlation between initial P_{ij} and changes in venue and tract characteristics.

To state these welfare losses in terms of equivalent transit-time increases, we compute the percentage increase in travel time from home (via both automobile and public transit) that, if applied to all restaurants in the set $\mathcal{J}_i^G$, would generate a change in welfare equal to $U_i' - U_i$. Denote this percentage increase for individual i by Δ_i and the concomitant percentage increase in the commuting transit time by Δ_i'. An increase in the log minutes of travel time from home by Δ_i for all restaurants in $\mathcal{J}_i^G$ would generate a new welfare level equal to U_i' if and only if Δ_i is such that

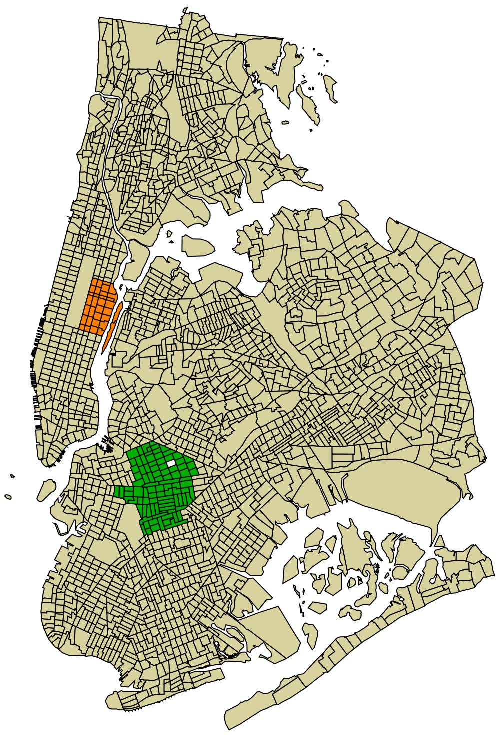
$$
\begin{aligned}
U_i' = \ln\Big(&\sum_{j\in\mathcal{J}_i^G}\sum_{l\in\{hp,hc\}} \exp(\gamma_{g(i)l}^1(X_{ijl}^1 + \Delta_i) + \gamma_{g(i)}^2 X_{ij}^2 + \beta_{g(i)}^1 Z_j^1 + \beta_{g(i)}^2 Z_{ij}^2) + \\
&\sum_{j\in\mathcal{J}_i^G}\sum_{l\in\{pp,pc\}} \exp(\gamma_{g(i)l}^1(X_{ijl}^1 + \Delta_i') + \gamma_{g(i)}^2 X_{ij}^2 + \beta_{g(i)}^1 Z_j^1 + \beta_{g(i)}^2 Z_{ij}^2) + \\
&\sum_{j\in\mathcal{J}_i^G}\sum_{l\in\{wp,wc\}} \exp(\gamma_{g(i)l}^1 X_{ijl}^1 + \gamma_{g(i)}^2 X_{ij}^2 + \beta_{g(i)}^1 Z_j^1 + \beta_{g(i)}^2 Z_{ij}^2) + \\
&\sum_{j\notin\mathcal{J}_i^G}\sum_{l\in\mathcal{L}} \exp(\gamma_{g(i)l}^1 X_{ijl}^1 + \gamma_{g(i)}^2 X_{ij}^2 + \beta_{g(i)}^1 Z_j^1 + \beta_{g(i)}^2 Z_{ij}^2)\Big) + c
\end{aligned}
$$

where c is the constant in expression (E.4) and X_{ijl}^1, X_{ij}^2, Z_j^1, and Z_{ij}^2 denote the initial (observed) values of the corresponding covariates.

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.2. 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.2: Bedford-Stuyvesant gentrification scenario



Change in	Mean	Std. Dev.
Asian residential share	0.061	0.048
Black residential share	-0.472	0.274
Hispanic residential share	-0.058	0.111
White residential share	0.477	0.327
Robberies per resident	-0.004	0.003
Spectral segregation index	-0.3	0.306
Yelp rating	0.014	1.04
Price ($ to $$$$)	0.792	0.935
Median household income (thousands)	54.407	51.432
Euclidean demographic distance	0.462	0.278
Number of restaurants	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.1 summarizes the decomposition of the resulting welfare loss, akin to Table 8 in the main text. Restaurants in the gentrifying area account for 20% of predicted visits by incumbent residents prior to gentrification. The change in restaurant and neighborhood characteristics is equivalent to the 221 restaurants becoming more than twice as far away in terms of transit times from home. 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.1: Welfare losses due to gentrification of surrounding Bedford-Stuyvesant tracts

Transit time increase equal to welfare loss	Initial visit share	Change in value of characteristics ($\gamma\Delta\bar{X}_i, \Delta\bar{Z}_i$) Social frictions	Restaurant traits	Other traits
160%	.203	-1.43	-.149	.545

NOTES: Welfare loss is expressed as the percentage increase in transit times from home that would be equivalent to the welfare loss associated with the covariate changes due to gentrification. See appendix E.3 for details. Initial visit share is $\sum_{j\in\mathcal{J}^G} P_{ij}$. Social frictions are EDD, SSI, EDD×SSI, and racial and ethnic population shares of k_j. Restaurant traits are price, rating, cuisine category, and price and rating interacted with median household income. Other traits are destination income, differences in incomes, and robberies per resident.

References

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

Dickstein, Michael, and Eduardo Morales. 2015. "What do Exporters Know?" mimeo.

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

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

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

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

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

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

Morales, Eduardo, Gloria Sheu, and Andrés Zahler. 2015. "Extended Gravity." mimeo.

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

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

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

Wollman, Thomas. 2015. "Trucks Without Bailouts: Equilibrium Product Characteristics for Commercial Vehicles." mimeo.