Appendix From Gender Homophily and Segregation Within Neighborhoods

Gregorio Caetano and Vikram Maheshri

This appendix is divided into three sections. In Section A, we provide robustness checks for our analysis of sorting by gender. In Section B, we present the results of a full analysis of sorting by age instead of by gender. In Section C, we present more detailed venue summary statistics stratified by subcategory.[39]

A Robustness Checks: Sorting by Gender

A.1 Measurement Error

Although user-generated datasets offers much promise, they are accompanied by several potential concerns regarding measurement error. In this section, we present evidence that our main results are qualitatively robust to many reasonable forms of measurement error.

A.1.1 Checkins Are Not Representative of Venue Visits

Our primary concern is that the proportion of females that we observe in a venue may be systematically different from the proportion of females that actually visit the venue. We argue that this likely does not confound our analysis, and, in any case, we show empirically that our results are qualitatively robust to the extent that it does. Indeed, in the presence of such measurement error, our results should actually be understood as conservative estimates of the amount of sorting within neighborhoods and the effects of this sorting on neighborhood and venue diversity.

To fix ideas, let $\tilde{f}_{jk}$ and $\tilde{m}_{jk}$ represent the actual numbers of females and males who visit venue j in neighborhood k. We can write the relationships between the observed and actual variables as

$$f_{jk} = \gamma^{f}_{jk} \cdot \tilde{f}_{jk} \tag{9}$$

$$m_{jk} = \gamma^{m}_{jk} \cdot \tilde{m}_{jk} \tag{10}$$

[39] This appendix is available online at http://bit.ly/1KzNf2X.

where the γ_{jk} parameters represent gender and venue specific check-in rates. All observed variables previously defined in terms of f_{jk} and m_{jk} have an actual, unobserved counterpart denoted with a tilde.

When mismeasurement is not gender specific, i.e., $\gamma_{jk}^f = \gamma_{jk}^m$, the female shares of check-ins at venues are unchanged, so all of our results are unaffected. This is a particularly nice feature, as it ensures our results are robust to any basic form of measurement error due to the fact that not all venue customers use the Foursquare app. Moreover, if mismeasurement is gender specific, but the mismeasurement in the female share of venues is only neighborhood specific (i.e., $s_{jk} = \gamma_k^s \cdot \tilde{s}_{jk}$), then our estimates of neighborhood Theil indices and their geographic decompositions are unchanged. This ensures that our results are robust to neighborhood specific sources of measurement error such as those correlated to unobserved neighborhood amenities.

In general, measurement error may be not only gender and neighborhood specific but also venue specific. We check the sensitivity of our main results to a general form of measurement error by conducting a Monte Carlo simulation. Without loss of generality, we define $\omega_{jk} = \frac{\gamma_{jk}^m}{\gamma_{jk}^f}$ to be the relative oversampling of males in venue j. For each iteration l, we randomly draw ω_{kj}^l for each venue from a uniform distribution $[\underline{\omega}, \overline{\omega}]$. We then calculate the "true" values of $\tilde{s}_{jk}^l$, $\tilde{T}_k^l$ for that iteration. Using these "true" values, we can simulate the main results of the paper, and the variation of the results across iterations allows us to construct confidence intervals. Although ω_{jk}^l is randomly drawn, it is positively correlated to $\tilde{s}_{jk}^l$ by construction.[40]

We conduct the Monte Carlo simulation under three separate parameterizations to capture qualitatively different types of measurement errors. In the first parametrization, we set $\underline{\omega} = 0.5, \overline{\omega} = 1.5$, which allows males to check in up to 50% less or more frequently than females, though they check in at the same rate on average. In the second parametrization, we set $\underline{\omega} = 2, \overline{\omega} = 4$. This increases the measurement error in two ways: it assumes that on average males check in three times more than females do, and it allows for greater dispersion of γ_{jk} across venues. In the third parametrization, we set $\underline{\omega} = 1, \overline{\omega} = 5$ which further worsens measurement error by allowing for even greater dispersion of ω_{jk} across venues.[41]

[40]We also performed alternative Monte Carlo simulations where we allowed ω_{jk}^l to be positively (or negatively) correlated to s_{jk} instead and obtained qualitatively similar results.

[41]We also performed analogous Monte Carlo simulations assuming females check in more rather than less frequently than males on average and found analogous results.

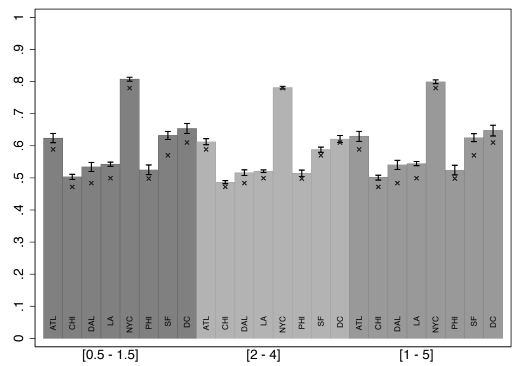
We report the Monte Carlo ($N = 500$) results for each of the main estimates of the paper in Figure 1 and in Table 1. Each panel in Figure 1 contains 24 bars, which represent the three different sets of parameters for each city in our sample. For each set of parameters, the bars represent the average estimate of that result across all 500 iterations. We also show 95% confidence intervals for these estimates along with the previously presented value of that result under the assumption of no measurement error denoted with an "x".

In the first panel of Figure 1, it is clear that our estimate of the fraction of the city sorting that occurs within census blocks is robust to various amounts of measurement error; if anything we underestimate the amount of sorting that occurs locally.[42] Even though the actual estimates under the assumption of no measurement error may fall outside of the confidence interval, they are qualitatively the same. A large proportion of sorting happens within blocks under all reasonable assumptions on measurement error. In second and third panels, we show how our regression results are affected by different kinds of measurement errors. If anything, measurement error leads to attenuation bias, mainly in $\hat{\beta}^V$. This is consistent with the results of our panel and IV identification strategies and suggests that our conclusion that $\beta^V < 0$ and $\beta^N > 0$ may be conservative. Overall, these simulations suggest that our results are generally robust to measurement error. Even though erroneously assuming away measurement error might lead us to estimate parameters that would fall outside of the true confidence intervals in some cases, our qualitative conclusions should not be affected even by very extreme forms of measurement error.

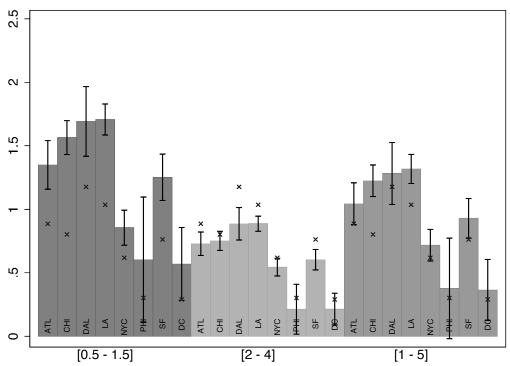
In Table 1, we present 95% confidence intervals for our estimates of γ_w^V and γ_w^N from analogous Monte Carlo simulations of row (4) of Table 4. Consistent with our main findings, our estimates of γ_{low}^V are robustly significantly different from zero at the 90% level.

[42]Because the measurement error that we introduce in the Monte Carlo simulation is correlated to the female share of venues, the across-neighborhood component of city sorting tends to be magnified more than the within-neighborhood component (see equation (2)).

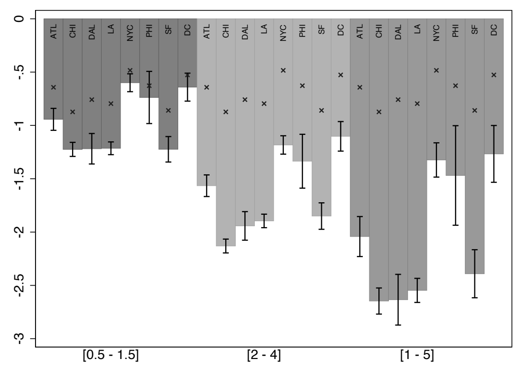
Figure 1: Robustness: Monte Carlo Results



(a) Proportion of City Sorting due to Within Census Blocks Sorting



(b) $\hat{\beta}^N$



(c) $\hat{\beta}^V$

Notes: Each panel presents Monte Carlo results for three different set of parameters $[\underline{\omega}, \overline{\omega}]$, which represent the interval of the uniform distribution from which ω_{jk} is drawn: $[0.5, 1.5]$, $[2, 4]$ and $[1, 5]$. The bars represent the estimates of the Monte Carlo with 95% confidence intervals, and "x" represents the estimates under the assumption of no measurement error, which are reported in the paper.

Table 1: Monte Carlo Estimates of Effects of Gender Diversity on Labor Force Participation Gaps

	Low Wage Jobs		Medium Wage Jobs		High Wage Jobs	
$[\underline{\omega}, \overline{\omega}]$:	γ^V_{low}	γ^N_{low}	$\gamma^V_{\text{med.}}$	$\gamma^N_{\text{med.}}$	γ^V_{high}	γ^N_{high}
$[0.5, 1.5]$	-0.011**	0.000	-0.002	0.001	-0.010	0.006
$[2, 4]$	-0.011***	0.002	0.002	-0.001	-0.016**	-0.009
$[1, 5]$	-0.009*	0.001	0.001	-0.000	-0.012	0.006

Notes: Low wage jobs pay less than $1,250 monthly, medium wage pay jobs pay between $1,250 and $3,333 monthly, and high wage pay jobs pay more than $3,333 monthly. All specifications include block group - wage group fixed effects, cubic B-spline (with as many knots as possible) for number of venues in block, numbers of female and male block visitors, numbers of female and male block residents, numbers of younger ($\leq$ 35) and older ($>$ 35) block residents (38 covariates for each group). ***: 1% significance level, **: 5% significance level, *: 10% significance level computed from Monte Carlo simulations with 500 draws.

A.1.2 Selected Venue Coverage

There may be some venues that do not experience any check-in activity during the sample period, so it is useful to consider the implications of this form of measurement error on our results. Given the vast size of our data set, the number of unobserved venues is likely to be small in the well traveled urban areas that comprise our sample. In Figure 2, we present heat maps of the density of venues in our sample for each of our eight cities. Borders correspond to census tracts, and more darkly shaded tracts contain more venues. In all of the sample cities, we find a concentration of venues in the central business district, and some reduction in venue density in more residential surrounding areas. This is anecdotally consistent with the structure of these cities and indicates the density of venues in our sample is spatially consistent with the density of venues in the overall population of venues.

Although we do not have information on unsampled venues by definition, we conjecture that unsampled venues would tend to be more similar to "barely sampled" venues (i.e., those that experience only a small number of check-ins) than to the more robustly sampled venues that comprise the bulk of our data set. This suggests an empirical robustness check that we can perform to see if hypothetically observing unsampled venues would dramatically alter our results. A venue is included in our sample if it experiences at least 10 check-ins over the one year sample period. As a robustness check, we increase this threshold in increments of 5 check-ins and replicate our entire analysis using these diminishing subsamples. If our results do not change much near the 10 check-in threshold, then it is reasonable to assume that the exclusion of unsampled venues would also have a small effect on our results.[43]

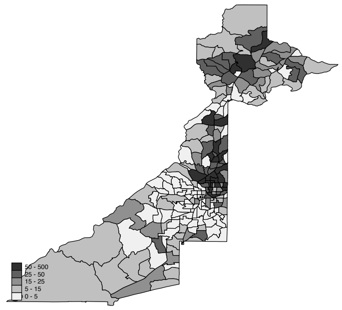
In the first two panels of Figure 3, we present our three main results – the fraction of sorting in each city due to sorting within census blocks and the estimates of β^V and β^N from our baseline regressions – replicated on subsamples with inclusion thresholds varying from 10 check-ins to 365 check-ins during our sample period. In the first panel, the fractions of sorting in each city that are due to sorting within blocks are quite flat near the 10 check-in threshold, which indicates that measurement error due to unsampled venues is not likely to affect our evidence of the intensity of homophily and highly local sorting. In the second panel, the estimates of β^V and β^N are also flat

[43] It is inadvisable to include venues that experience fewer than 10 check-ins in our sample because then we would be unable to obtain sufficiently fine estimates of the gender compositions of those venues.

near the 10 check-in threshold. To the extent that they trend away from zero as we include venues with fewer check-ins suggests that this form of measurement error attenuates our results. Hence, if anything our reported estimates are conservative.[44]

[44]The trends away from zero of our regression coefficient estimates as we include venues with fewer check-ins are consistent with our finding that the effects of venue variety on venue and neighborhood diversity are largest in neighborhoods with low levels of venue variety (b-spline specification). This serves as additional evidence that, if anything, our regression estimates are conservative.

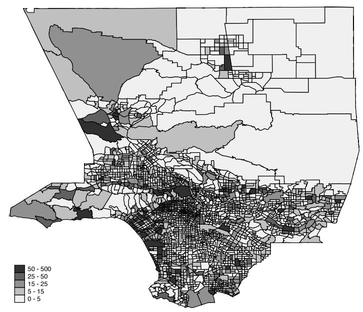
Figure 2: Venue Coverage Maps of Sample Cities



(a) Atlanta

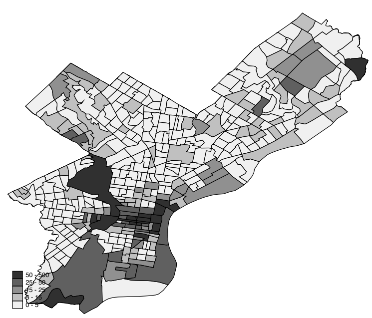
(b) Chicago

(c) Dallas



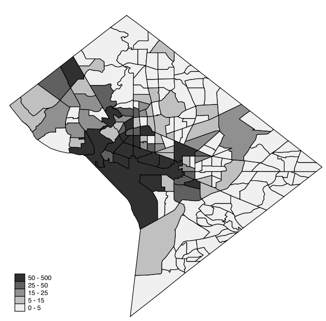
(d) Los Angeles

(e) New York City



(f) Philadelphia

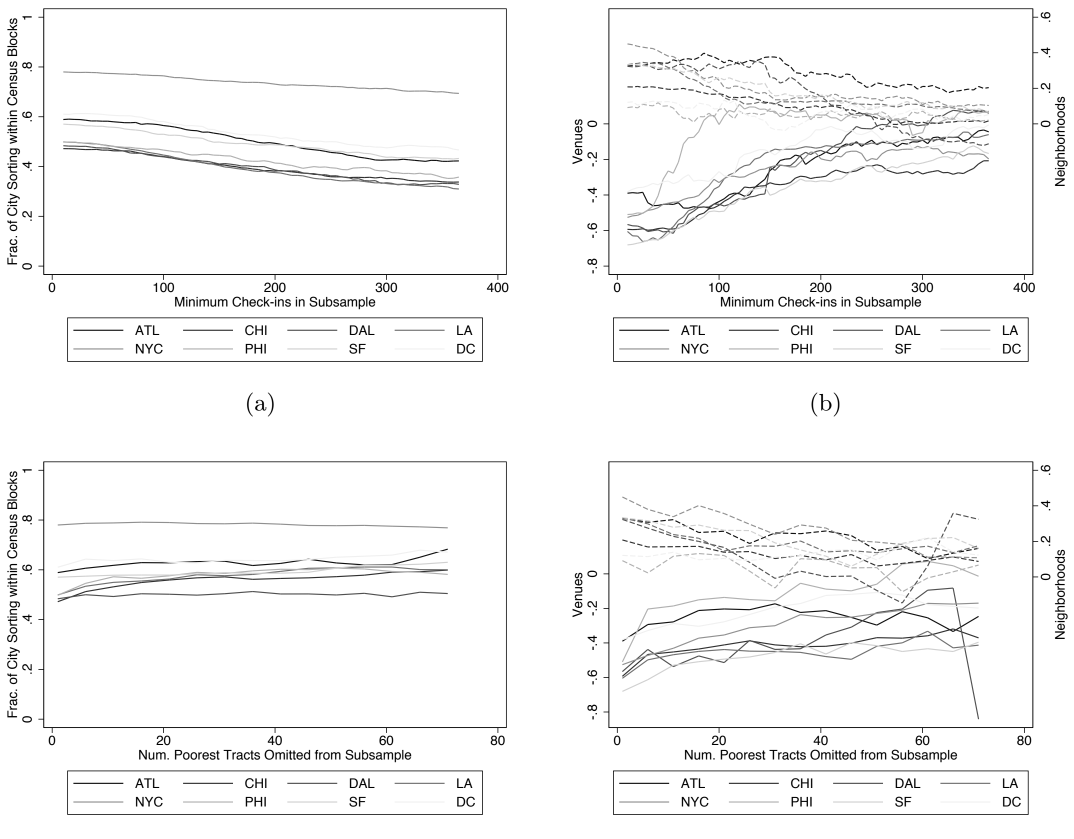
(g) San Francisco



(h) Washington, DC

Notes: Each map shows the number of venues in the sample overlaid on a map of all census tracts in the primary county of each sample city. Darker regions correspond to tracts with more venues.

Figure 3: Robustness: Selected Venue Coverage



(a) (b)

(c) (d)

Notes: Solid lines refer to the y-axis on the left, and dashed lines refer to the y-axis on the right. The measures in the first two panels are recalculated using subsamples that include only venues that experience at least a given number of check-ins during our sample period. The measures in the last two panels are recalculated using subsamples that include only venues in tracts with sufficiently high median income ranks according to the 2013 American Communities Survey.

Because checking in on Foursquare requires the use of a "smart" mobile device, Foursquare users likely tend to be wealthier, and hence they might disproportionately frequent more expensive venues. We can assess the extent to which the potential selection of venues in our sample due to this effect biases our results in a similar exercise to the one above. In each city, we rank all tracts by their median household income according to the 2013 American Community Survey. We incrementally eliminate all venues in the 5 poorest tracts, 10 poorest tracts, 15 poorest tracts, etc. and replicate our entire analysis using these diminishing subsamples. If our results do not change much as we are changing the poorest tracts of the sample, then it is reasonable to assume that any selection of the venues in our sample due to users being wealthier would also have a small effect on our results. In the third and fourth panels of Figure 3, we present the same three results replicated on subsamples with the 5 to 75 poorest tracts in each city omitted. The results are highly similar to their counterparts in the first two panels, which suggests that this form of measurement error does not qualitatively affect our main results.

In sum, these results allows us to conclude that the coverage of the venues in our sample is quite comprehensive, and to the extent that there may be selection in the sample then our reported results will be conservative.

A.1.3 Sampling Error: A Falsification Test

Consider the extreme situation in which all venues in neighborhood k have the same true female share $\tilde{s}_{jk}$ but we observe variation in s_{jk} across venues purely because of sampling error. Under this falsification exercise, how would our main results differ? To answer this, we simulate a counterfactual in which the individuals in a city sort across tracts, block groups and blocks according to the data, but they do not sort within blocks. This provides an intuitive falsification test of our interpretation of our main findings: if the block level Theil indices constructed under this counterfactual are similar to their analogs as constructed with our data, then our results should not be interpreted as evidence of local sorting.

We implement this test by randomly assigning individuals to venues in a particular block in proportion to the overall gender distribution that we observe in that block. If we observe venue i in block b with $f_{ib} + m_{ib}$ check-ins in our data, we recreate the gender composition of venue i by taking $f_{ib} + m_{ib}$ independent draws from a Bernoulli(p_b) distribution with replacement, where

$$p_b = \frac{\sum_{i \in b} f_{ib}}{\sum_{i \in b} f_{ib} + m_{ib}} \tag{11}$$

is the overall proportion of female check-ins in block b (i.e., across all venues). For each 1 that is drawn, we add a female to venue i, and for each 0 that is drawn, we add a male to venue i. The variation in the gender composition of venues within blocks in this simulated sample is fully attributable to measurement error.

Table 2: Placebo Tests: No Active Sorting Within Census Blocks

Placebo for:	Proportion of city-wide sorting due to sorting within:			$\hat{\beta}^V$	$\hat{\beta}^N$
	Tracts	Block Groups	Blocks		
Atlanta	0.73	0.59	0.03	0.00	0.39
Chicago	0.68	0.53	0.02	0.00	0.26
Dallas	0.62	0.46	0.03	0.00	0.33
Los Angeles	0.68	0.51	0.04	0.00	0.32
New York City	0.70	0.51	0.05	0.00	0.71
Philadelphia	0.72	0.60	0.03	0.00	0.26
San Francisco	0.65	0.54	0.04	0.00	0.44
Washington, DC	0.70	0.61	0.03	0.00	0.40

Notes: All results are calculated under the placebo assumption that individuals do not actively sort within census blocks. Bootstrapped standard errors for all entries in all cities are less than 0.005 and are omitted for clarity.

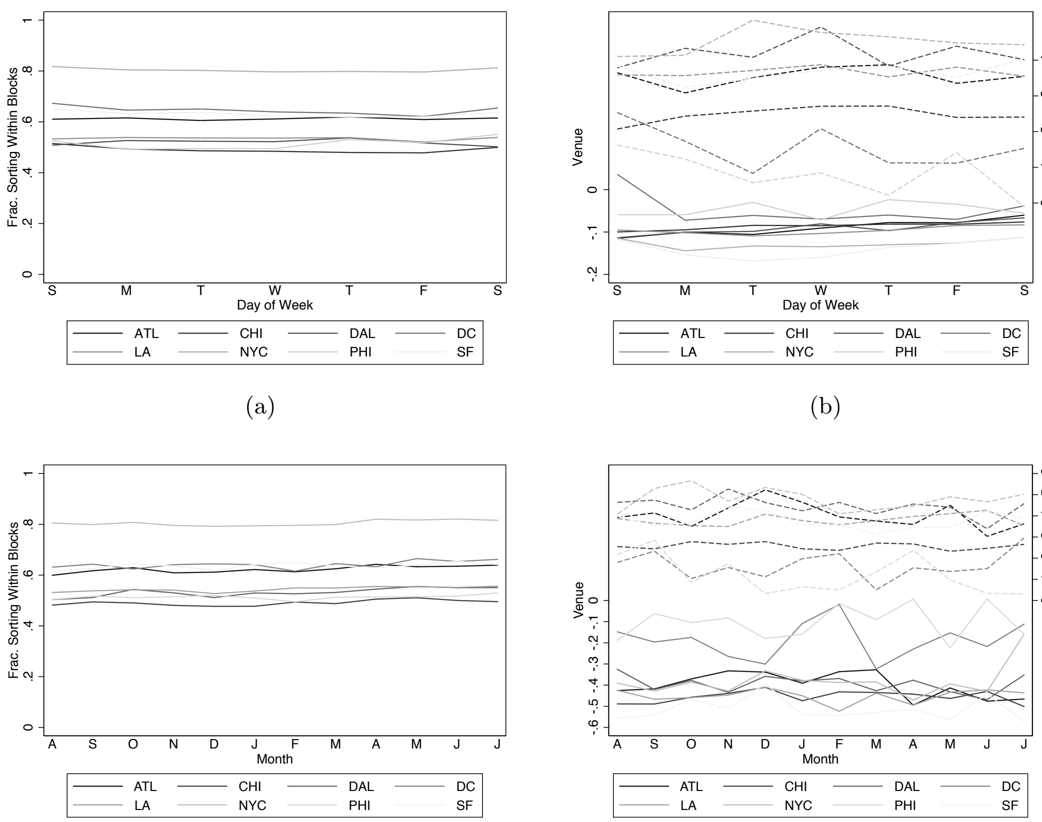
For each simulated sample of venues, we can re-estimate our results. We repeat this exercise 500 times and report the mean and standard deviation of these counterfactual results across all repetitions. In Table 2, we present the fraction of sorting within each city that is due to sorting within neighborhood types, and baseline estimates of β^V and β^Nunder this counterfactual assumption of no sorting within blocks. The results are as expected. The proportion of venue sorting within tracts and within block groups decreases slightly as expected, and the proportion of sorting within blocks is reduced to nearly zero, as such sorting can only be due to measurement error. In addition, our estimates of β^V decrease to zero as expected (with no sorting within blocks, venue diversity should be unaffected by venue variety) while our estimates of β^N remain positive and of the same order of magnitude as before, as sorting across neighborhoods is unchanged under the counterfactual.

From this exercise, we find that our main results differ completely from their counterfactual counterparts, which constitutes further evidence that our main results are not artifacts of measurement error.

A.2 Dynamic Misaggregation

Per the discussion in the data section, we aggregated check-ins in our sample annually to reduce any potential measurement error. However, if there are strong dynamic components to gender sorting, this aggregation could potentially obscure interesting longitudinal variation in venue sorting. For example, this could happen if venues varied in substitutability by season (e.g., people may not enjoy parks as much in the winter, especially in cold weather cities), or by day of the week (e.g., people may prefer downtown venues on weekdays). We replicate our analysis disaggregated by day of week and by month, and present the main results in Figure 4.

Figure 4: Robustness: Dynamic Aggregation



(a) (b) (c) (d)

Notes: Solid lines refer to the y-axis on the left, and dashed lines refer to the y-axis on the right. All measures are calculated by replicating the analysis by day of week (first two panels) or by month (last two panels).

Looking at the first and third panels, it is immediate that there is nearly zero dynamic variation in the fraction of sorting in each city due to sorting within blocks. There is markedly more dynamic variation in our estimates of β^V and β^N for each city, as depicted in the second and fourth panels (left and right axes, respectively). However, this variation does not follow any systematic trend, and we infer that $\beta^V < 0$ and $\beta^N > 0$ for all days of the week and months of the year. These exercises suggest that our main results are unaffected by our choice of annual aggregation.

A.3 Robustness Checks for $\hat{\beta}^V$ and $\hat{\beta}^N$ by City

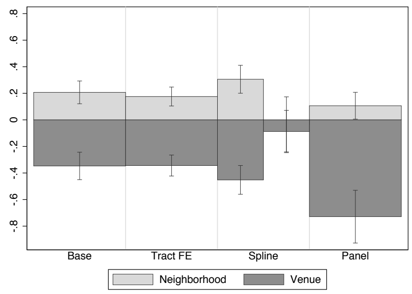
The following plots show estimates of β^V and β^N by city for the baseline specification (block group FEs), the specification replacing block group FEs for tract FEs, the b-spline specification and the panel data specification replacing block FEs and city-month FEs for block group FEs.[45]

The dark shaded bars represent $\hat{\beta}^V$, and the light shared bars represent $\hat{\beta}^N$. Venue variety is defined as the number of unique venue categories in the first column and the number of unique venue subcategories in the second column. The first bars correspond to baseline estimates. The second bars replace the block group fixed effects in the baseline estimates with tract fixed effects. The third set of bars correspond to estimates of the parameters specified as a linear b-spline with a knot at 3 three categories or subcategories. The fourth bars correspond to estimates where the dataset is disaggregated to a monthly panel and the block group fixed effects are replaced with block fixed effects. As can be seen, the results are similar to the ones reported in the paper.

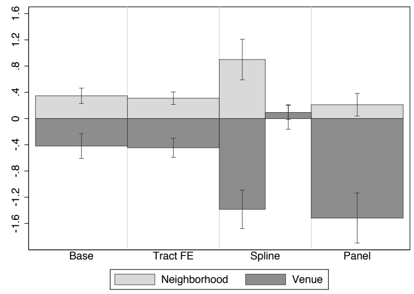
[45]We were unable to obtain IV estimates disaggregated by city due to their lack of precision.

Figure 5: $\hat{\beta}^V$ and $\hat{\beta}^N$: Alternative Identification Strategies For Gender Sorting by City (1 of 2)

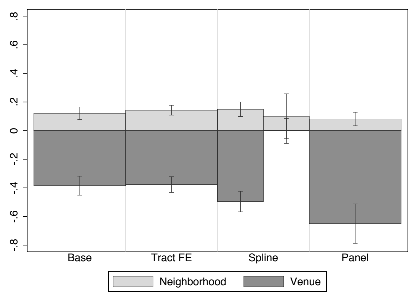
V_b = Category $\quad\quad$ V_b = Subcategory



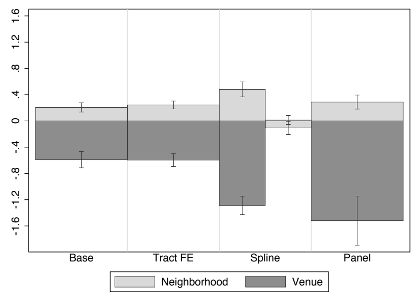
(a) Atlanta



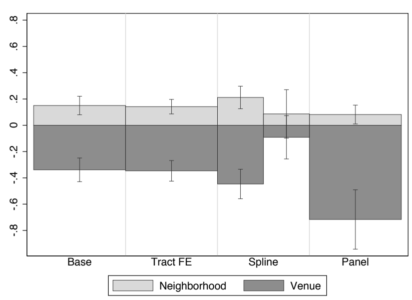
(b) Atlanta



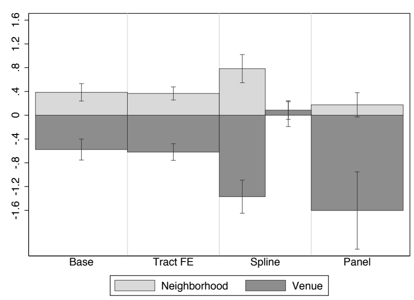
(c) Chicago



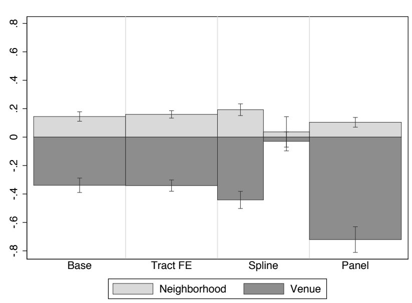
(d) Chicago



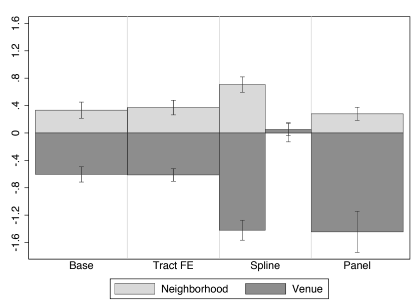
(e) Dallas



(f) Dallas



(g) Los Angeles

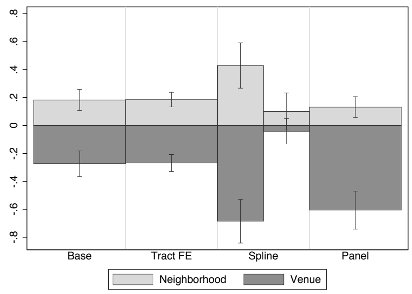


(h) Los Angeles

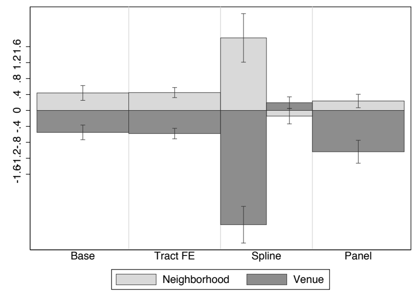
Note: See next page.

Figure 6: $\hat{\beta}^V$ and $\hat{\beta}^N$: Alternative Identification Strategies For Gender Sorting by City (2 of 2)

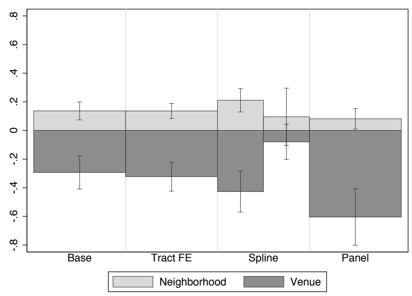
V_b = Category $\qquad\qquad$ V_b = Subcategory



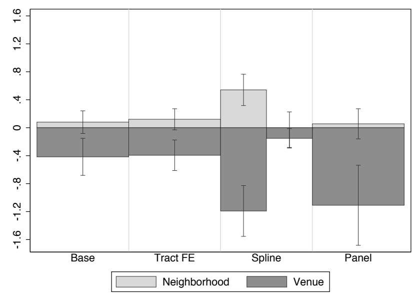
(a) New York City



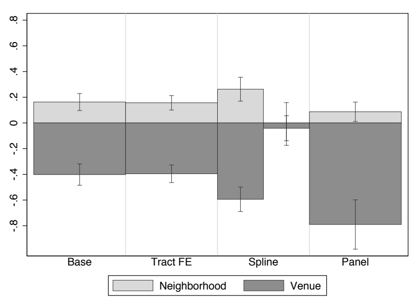
(b) New York City



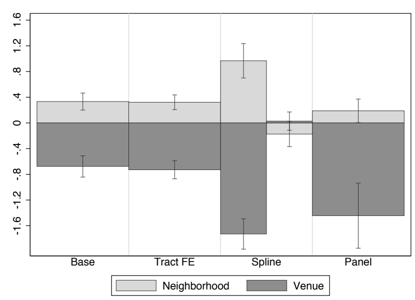
(c) Philadelphia



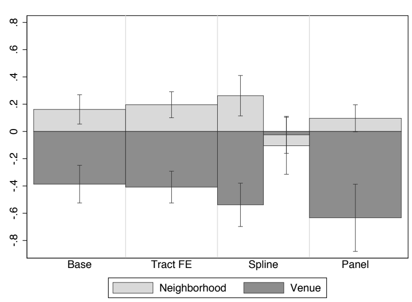
(d) Philadelphia



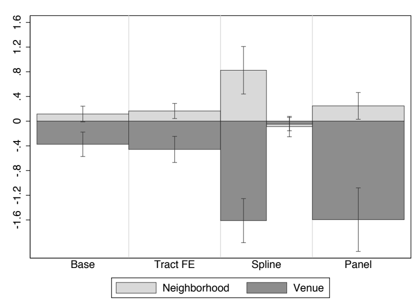
(e) San Francisco



(f) San Francisco



(g) Washington, DC



(h) Washington, DC

A.4 Gender Homophily and the Labor Force Participation Gap: Robustness Checks

In Table 3, we present results for 10 distinct specifications of equation (8) that feature varying controls (X_b), neighborhood fixed effects (α) and estimation subsamples to establish the robustness of our identification strategy. Each row corresponds to a different specification. Specification (4) corresponds to the preferred specification of Table 4 in the main text.

In specification (1), we specify only covariates from our Foursquare data in order to control for differences in venue density and check-in intensities of males and females across blocks. In specification (2), we control for the number of male and female residents on each block. Our estimates are essentially unchanged, which supports our assumption that there is no systematic sorting by gender across blocks within block groups. In specification (3), we specify all control variables from specification (2) flexibly using cubic B-splines.[46] Doing so substantially improves the fit of the model, and we estimate slightly larger effects of venue diversity on low wage employment gaps. Furthermore, specifying MR_b and FR_b flexibly ensures that our use of them in constructing GAP_{wbt} does not introduce spurious correlation into the regression. The fact that as we better control for resident characteristics, our estimates of γ^V_{low} become larger in magnitude suggests that unobservables due to residential sorting might, if anything, bias our estimate of γ^V_{low} downward. In specification (4), we similarly control for the number of young and old residents on each block and find no changes in our estimates, which further supports our assumption of no residential sorting by gender within block groups.[47] In principle, our identifying assumption allows for sorting across blocks along other dimensions as long as this sorting occurs uniformly by gender. To check for whether sorting along other dimensions might bias our results, in specification (5) we control for other demographic characteristics of workers (race, Hispanic ethnicity and level of education). Although the fit of the model slightly improves, all coefficient estimates are unchanged, as expected. As a final test for the presence of sorting within block groups, we re-estimate the regression with *tract*-wage group fixed effects instead of block group-wage group fixed effects in specification (6). Any bias due to sorting within block groups in the baseline specification should be further exacerbated in this specification

[46] We use as many knots as possible for each control variable: 7 equally spaced knots for the numbers of male and female visitors and residents (24 covariates) and 3 equally spaced knots for the number of venues (2 covariates).

[47] This is also suggestive that measurement error in GAP_{wbt} does not bias our results, as older residents are likely a better match for jobs in higher wage groups.

since sorting across block groups within the same tract is no longer controlled for. The fact that our estimates in specification (6) are smaller suggests that our baseline results might actually be conservative. This result corroborates our intuition when comparing the results from specifications (2) and (3).

Because of the linkages we found in Section 4 of the main text, one might worry that we have only identified the effect of venue variety rather than gender diversity in venues. For instance, a lack of venue variety might generate more venue diversity and less neighborhood diversity in gender, age, race, and other demographic characteristics, all of which could plausibly influence the gender gap in employment. To address this concern, we include block level measures of age diversity in neighborhoods and venues (i.e., age analogs of D_b^N and D_b^V) as controls in specification (7). Our estimates decrease only slightly, and these two variables do not improve the fit, which supports our claim that the effects we identify are due to gender sorting itself. Because we found that venue variety affects both venue and neighborhood diversity, we drop D_b^N from the model in specification (8). Our estimates are unchanged, which supports our claim that the effects we identify are mediated by gender diversity inside venues. In specification (9), we restrict our sample by dropping all blocks in which at least one resident works on his or her residential block in order to gauge whether our results are primarily driven by residents who work extremely close to home (e.g., workers who live above the store that they own). If anything, our estimates slightly increase, suggesting that this is not the case. In specification (10), we re-estimate the model with block group-*year*-wage group fixed effects. Our results are unchanged.

Because the adjusted R^2 is substantially reduced in specification (10), this suggests that pooling data from multiple years could be useful as it will only improve the precision of our estimates. Accordingly, in Table 4, we reestimate all 10 specifications on an augmented sample created by pooling LODES data on employment gaps from 2010-2013, which extends the sample period from two to four years. Although our explanatory variables do not vary over time, this should still increase our statistical power. Indeed, all standard errors decrease. Our main finding of $\hat{\gamma}_{\text{low}}^V < 0$ is unchanged, but we now obtain more precise zero-estimates for all effects on medium and high wage jobs.

Table 3: Effects of Diversity on the Labor Force Participation Gender Gap

	Specification	Low Wage Jobs γ^V_{low}	Low Wage Jobs γ^N_{low}	Medium Wage Jobs $\gamma^V_{\text{med.}}$	Medium Wage Jobs $\gamma^N_{\text{med.}}$	High Wage Jobs γ^V_{high}	High Wage Jobs γ^N_{high}	N	Adj. R^2
(1)	Controls from Foursquare Data[1]	-0.013*	-0.001	-0.006	0.008	-0.016	0.020	76,236	0.159
		(0.007)	(0.009)	(0.009)	(0.009)	(0.015)	(0.015)		
(2)	(1) + residential gender controls[2]	-0.013*	-0.002	-0.006	0.007	-0.015	0.017	76,236	0.188
		(0.007)	(0.009)	(0.009)	(0.009)	(0.014)	(0.014)		
(3)	(2) + flexibly specified controls [3]	-0.015**	-0.002	-0.003	0.001	-0.016	0.013	76,236	0.325
		(0.007)	(0.007)	(0.008)	(0.008)	(0.012)	(0.012)		
(4)	(3) + flexible residential age controls[4]	-0.015**	-0.003	-0.003	0.000	-0.015	0.013	76,236	0.325
		(0.007)	(0.007)	(0.008)	(0.008)	(0.012)	(0.012)		
(5)	(4) + workers' demographics characteristics[5]	-0.015**	-0.002	-0.004	0.001	-0.016	0.012	76,236	0.330
		(0.007)	(0.007)	(0.008)	(0.008)	(0.012)	(0.012)		
(6)	(4) + tract-wage group FEs instead of block group-wage group FEs	-0.010**	-0.003	-0.005	0.006	-0.006	0.003	76,236	0.310
		(0.005)	(0.005)	(0.006)	(0.006)	(0.009)	(0.009)		
(7)	(4) + D_b^V and D_b^N for age[6]	-0.012**	-0.001	-0.001	-0.003	-0.013	0.011	76,236	0.325
		(0.007)	(0.007)	(0.008)	(0.009)	(0.012)	(0.013)		
(8)	(4), drop D_b^N	-0.015**	–	-0.003	–	-0.014	–	76,236	0.325
		(0.007)		(0.008)		(0.011)			
(9)	(4), drop blocks where residents also work[7]	-0.017**	-0.005	-0.009	0.003	-0.016	0.009	63,146	0.298
		(0.008)	(0.008)	(0.009)	(0.009)	(0.013)	(0.014)		
(10)	(4) + block group-year-wage group FEs instead of block group-wage group FEs	-0.016*	-0.003	0.000	0.001	-0.015	0.012	76,236	0.262
		(0.008)	(0.009)	(0.009)	(0.010)	(0.014)	(0.015)		

Notes: Low wage jobs pay less than \$1,250 monthly, medium wage pay jobs pay between \$1,250 and \$3,333 monthly, and high wage pay jobs pay more than \$3,333 monthly. All specifications include block group - wage group fixed effects, with the exceptions of (6) and (10). Robust standard errors clustered at the block level are presented in parentheses.

[1] : number of venues in block, and numbers of female and male block visitors (3 covariates for each group); [2] : add to (1) the numbers of female and male block residents (2 additional covariates for each group);[3] : Cubic B-spline (with as many knots as possible) of all controls in (2) (26 covariates for each group). [4] : add to (3) cubic B-spline of numbers of younger (≤ 35) and older (> 35) block residents (12 additional covariates for each group). [5] : Add to (4) the numbers of block workers who are White, Black, other (non-White and non-Black), Hispanic, non-Hispanic, college graduates, college non-graduates (7 additional covariates for each group). [6] : Add to (4) D_b^{Vy} and D_b^{Ny}, which are the analogous measures of D_b^V and D_b^N based on the proportion of younger visitors (≤ 35), rather than based on the proportion of female visitors (2 additional covariates for each group). [7] : Drop observations from blocks with at least one resident who works in the same block. **: 5% significance level, *: 10% significance level.

Table 4: Effects of Diversity on the Labor Force Participation Gender Gap: $t \geq$2010

	Specification	Low Wage Jobs γ^V_{low}	Low Wage Jobs γ^N_{low}	Medium Wage Jobs $\gamma^V_{\text{med.}}$	Medium Wage Jobs $\gamma^N_{\text{med.}}$	High Wage Jobs γ^V_{high}	High Wage Jobs γ^N_{high}	N	Adj. R^2
(1)	Controls from Foursquare Data[1]	-0.014**	-0.004	-0.004	0.006	0.001	0.008	152,335	0.195
		(0.006)	(0.006)	(0.007)	(0.007)	(0.010)	(0.011)		
(2)	(1) + residential gender controls[2]	-0.014**	-0.004	-0.004	0.005	0.002	0.005	152,335	0.218
		(0.005)	(0.006)	(0.006)	(0.007)	(0.010)	(0.010)		
(3)	(2) + flexibly specified controls [3]	-0.014**	-0.007	0.001	-0.000	-0.002	0.005	152,335	0.328
		(0.005)	(0.005)	(0.005)	(0.006)	(0.008)	(0.008)		
(4)	(3) + flexible residential age controls[4]	-0.014**	-0.007	0.001	-0.000	-0.001	0.005	152,335	0.328
		(0.005)	(0.005)	(0.005)	(0.006)	(0.007)	(0.008)		
(5)	(4) + workers' demographics characteristics[5]	-0.013**	-0.006	0.000	-0.000	-0.002	0.004	152,335	0.334
		(0.005)	(0.005)	(0.005)	(0.006)	(0.007)	(0.008)		
(6)	(4) + tract-wage group FEs instead of block group-wage group FEs	-0.010**	-0.006	-0.002	0.006	0.002	0.000	152,335	0.303
		(0.004)	(0.004)	(0.004)	(0.005)	(0.006)	(0.006)		
(7)	(4) + D_b^V and D_b^N for age[6]	-0.011**	-0.004	0.003	-0.002	-0.001	0.003	152,335	0.328
		(0.005)	(0.005)	(0.006)	(0.006)	(0.008)	(0.009)		
(8)	(4), drop D_b^N	-0.014**	–	0.001	–	-0.000	–	152,335	0.328
		(0.005)		(0.005)		(0.007)			
(9)	(4), drop blocks where residents also work[7]	-0.014**	-0.007	-0.001	-0.000	-0.001	0.000	127,024	0.309
		(0.006)	(0.006)	(0.006)	(0.006)	(0.008)	(0.009)		
(10)	(4) + block group-year-wage group FEs instead of block group-wage group FEs	-0.013**	-0.007	0.003	0.000	-0.002	0.005	152,335	0.267
		(0.006)	(0.007)	(0.007)	(0.007)	(0.010)	(0.011)		

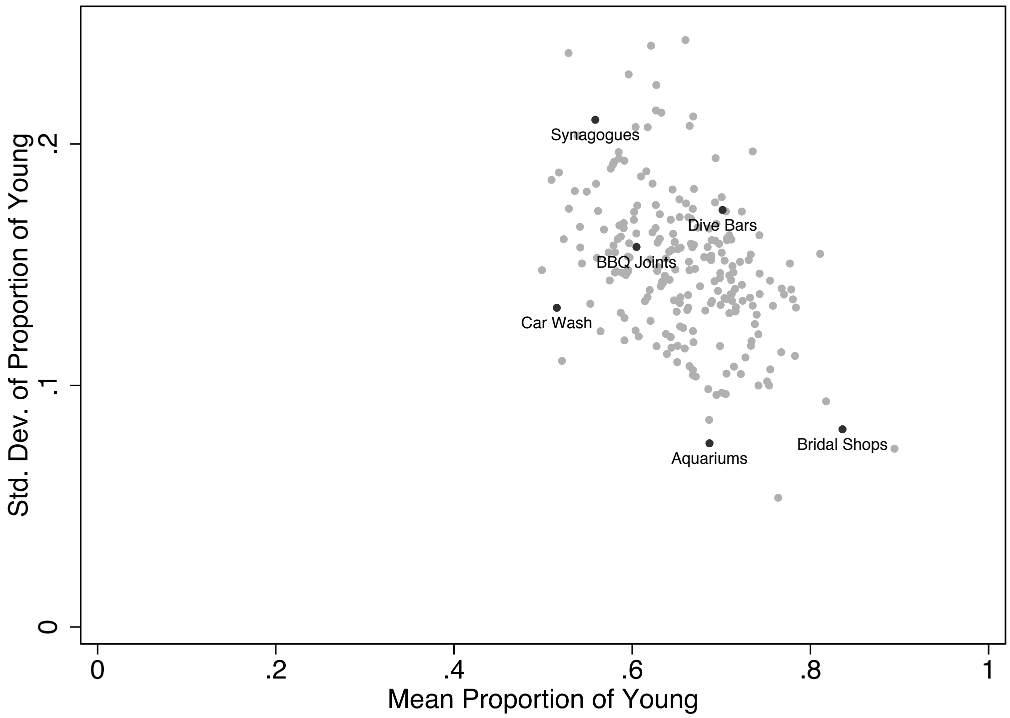
Notes: This Table is analogous to Table 3 in this appendix, however it is estimated on a subsample of LODES data from 2010-2013. Low wage jobs pay less than \$1,250 monthly, medium wage pay jobs pay between \$1,250 and \$3,333 monthly, and high wage pay jobs pay more than \$3,333 monthly. All specifications include block group - wage group fixed effects, with the exceptions of (6) and (10). Robust standard errors clustered at the block level are presented in parentheses.

[1] : number of venues in block, and numbers of female and male block visitors (3 covariates for each group); [2] : add to (1) the numbers of female and male block residents (2 additional covariates for each group);[3] : Cubic B-spline (with as many knots as possible) of all controls in (2) (26 covariates for each group). [4] : add to (3) cubic B-spline of numbers of younger ($\leq$ 35) and older ($>$ 35) block residents (12 additional covariates for each group). [5] : Add to (4) the numbers of block workers who are White, Black, other (non-White and non-Black), Hispanic, non-Hispanic, college graduates, college non-graduates (7 additional covariates for each group). [6] : Add to (4) D_b^{Vy} and D_b^{Ny}, which are the analogous measures of D_b^V and D_b^N based on the proportion of younger visitors ($\leq$ 35), rather than based on the proportion of female visitors (2 additional covariate for each group). [7] : Drop observations from blocks with at least one resident who works in the same block. **: 5% significance level, *: 10% significance level.

B Results of Analysis by Age

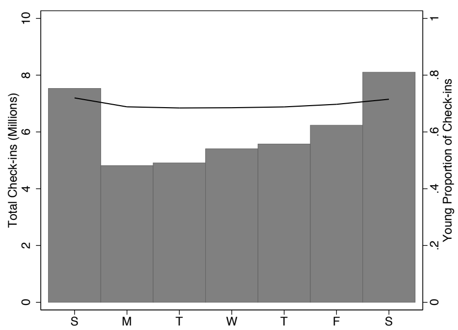
We replicate all tables and figures for sorting by age, including those reported in the previous section.

Figure 7: Proportion of Youth in Venues by Subcategory

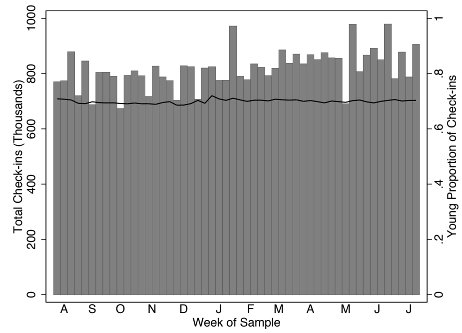


Note: This scatter plot pools venues from all cities in the sample. Each dot represents all of the venues within a subcategory.

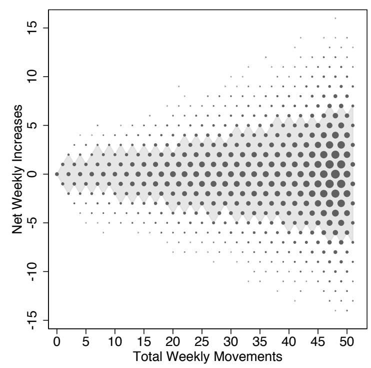
Figure 8: Check-ins and Age Composition Over Time



(a) Day-of-Week Variation



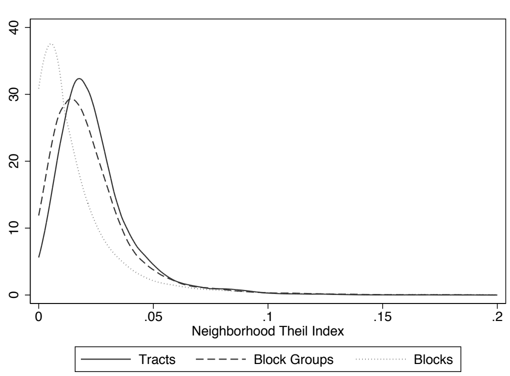
(b) Week-of-Year Variation



(c) Within Venue Variation

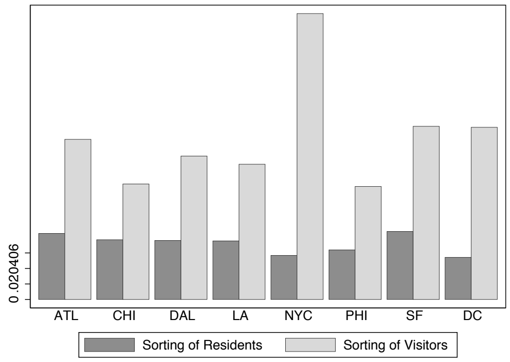
Notes: (a), (b): Bars represent total check-ins, lines represent age composition of aggregate check-ins. The 53rd week of the sample is omitted because it only contains a single day. (c): In this scatter plot of venues in our data, larger dots correspond to a greater numbers of venues. A venue experiences a weekly increase (decrease) in gender composition if the proportion of female check-ins rises (falls) by at least one percentage point.

Figure 9: Densities of Age Theil Indices for Various Neighborhood Definitions



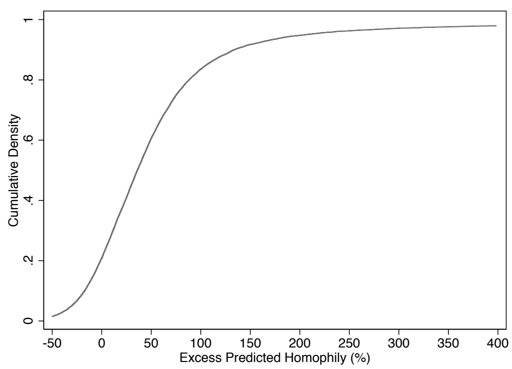
Notes: All densities are estimated using a bandwidth of 0.005 and an Epanechnikov kernel. For clarity, we present the density only for values of the domain less than 0.2; fewer than 1% of neighborhoods of any type have a Theil Index in excess of 0.2. Theil Indices are pooled across neighborhoods in all cities.

Figure 10: Sorting of Residents vs. Sorting of Visitors by Age



Note: "Sorting of Residents" is calculated as the Theil index of the gender composition of block residents from the 2010 Census. For comparability, "Sorting of Visitors" is calculated as the Theil index of the gender composition of check-ins in blocks. Bootstrapped standard errors for all estimates are below 0.005 and are omitted for clarity.

Figure 11: Excess Predicted Homophily in Venue Data



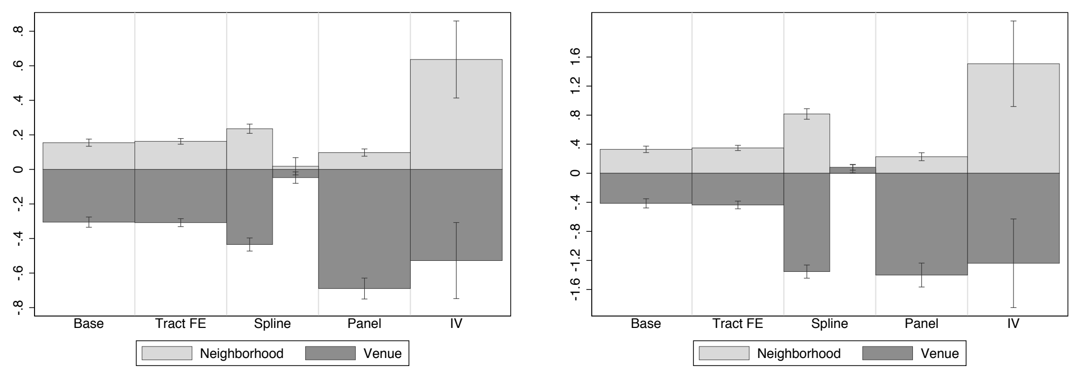
Note: In this figure, we present the empirical cumulative distribution of how much more likely we would predict that a youth would encounter another youth in a census block using venue level data than if we used residential data. Note that in some neighborhoods, we are more likely to predict homophily in residential data than in venue level data (negative excess predicted homophily). However, in most neighborhoods the opposite is true.

Table 5: Proportion of Within-Neighborhood Sorting By Age Due to Sorting Across Subcategories:

	City	Tracts	B. Groups	Blocks
Atlanta	0.14	0.75	0.82	0.91
Chicago	0.12	0.81	0.86	0.93
Dallas	0.14	0.79	0.83	0.92
Los Angeles	0.09	0.80	0.85	0.91
New York City	0.15	0.67	0.77	0.88
Philadelphia	0.16	0.79	0.83	0.93
San Francisco	0.16	0.71	0.77	0.91
Washington, DC	0.20	0.71	0.77	0.90

Note: Subcategories (225) are defined in the next section. Bootstrapped standard errors for all entries are less than 0.005 and are omitted for clarity.

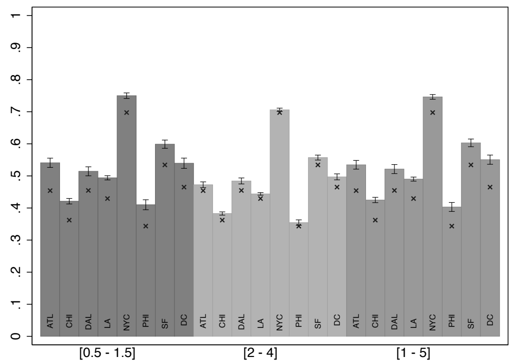
Figure 12: $\hat{\beta}^V$ and $\hat{\beta}^N$: Alternative Identification Strategies for Age Sorting



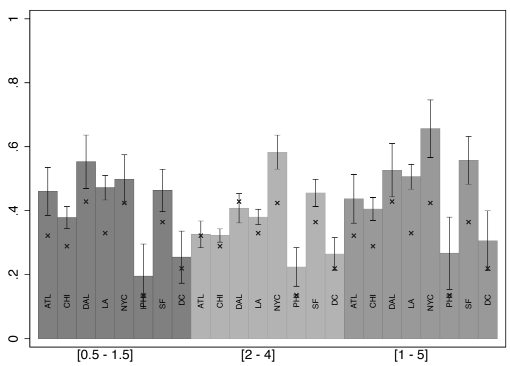
(a) V_b = Number of Categories (b) V_b= Number of Subcategories

Notes: The dark shaded bars represent $\hat{\beta}^V$, and the light shared bars represent $\hat{\beta}^N$. The first bars correspond to baseline estimates (block group FEs). The second bars replace the block group fixed effects in the baseline estimates with tract fixed effects. The third bars correspond to estimates from where the dataset is disaggregated to a monthly panel and the block group fixed effects are replaced with block fixed effects. The fourth bars correspond to 2SLS estimates of the baseline regressions with zoning instruments.

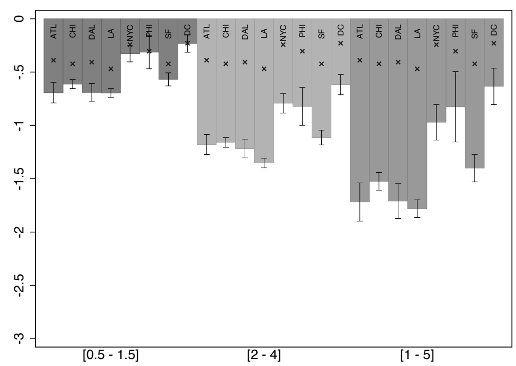
Figure 13: Robustness (Age): Monte Carlo Results



(a) Proportion of City Sorting due to Within Census Blocks Sorting



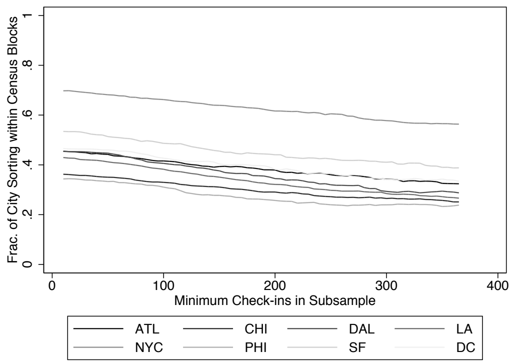
(b) $\hat{\beta}^N$



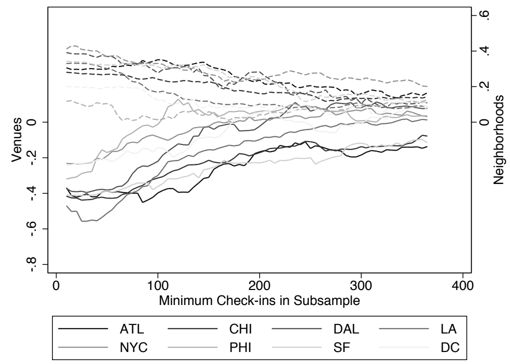
(c) $\hat{\beta}^V$

Notes: Each panel presents Monte Carlo results for three different set of parameters $[\underline{\omega}, \overline{\omega}]$, which represent the interval of the uniform distribution from which ω_{jk} is drawn: $[0.5, 1.5]$, $[2, 4]$ and $[1, 5]$. The bars represent the estimates of the Monte Carlo with 95% confidence intervals, and "x" represents the estimates under the assumption of no measurement error, which are reported in the paper.

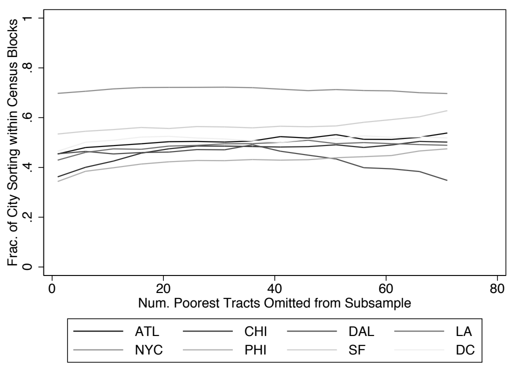
Figure 14: Robustness (Age): Selected Venue Coverage



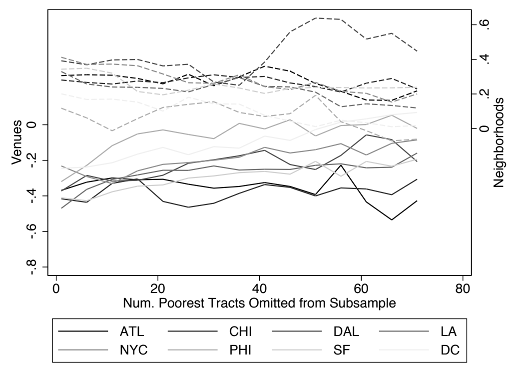
(a)



(b)



(c)



(d)

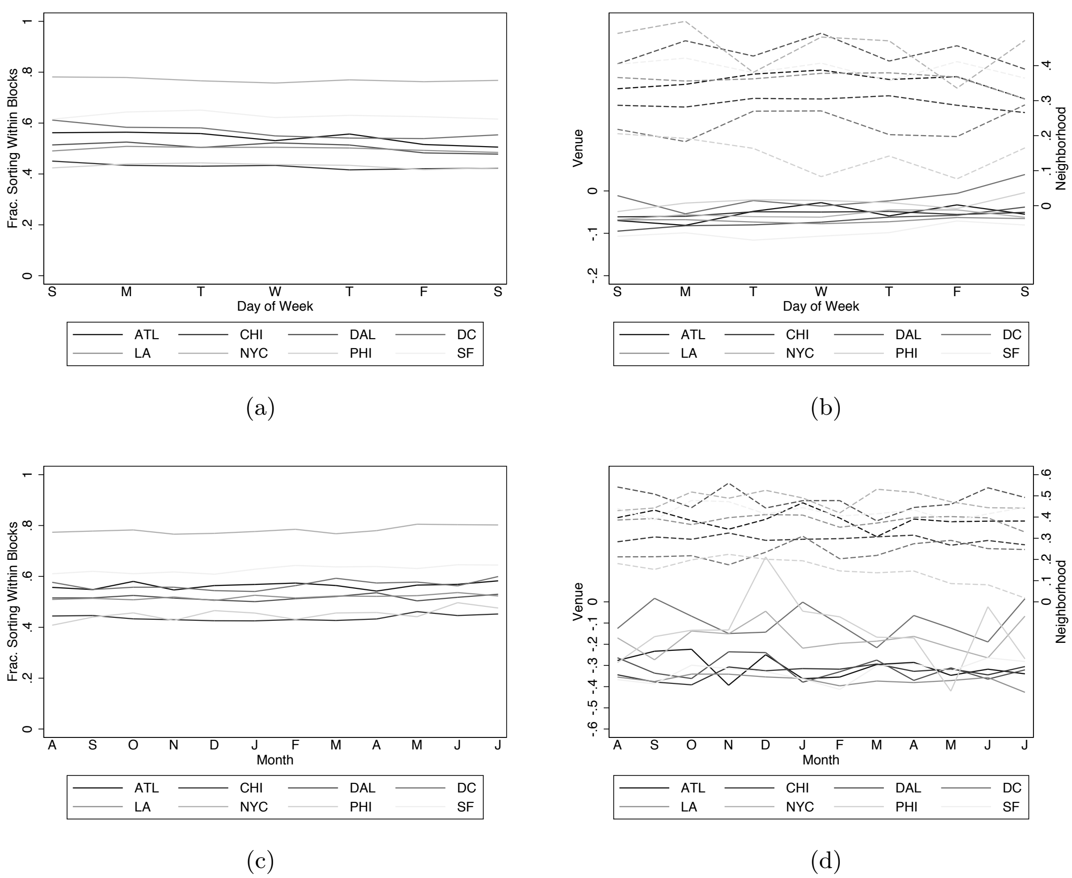
Notes: Solid lines refer to the y-axis on the left, and dashed lines refer to the y-axis on the right. The measures in the first two panels are recalculated using subsamples that include only venues that experience at least a given number of check-ins during our sample period. The measures in the last two panels are recalculated using subsamples that include only venues in tracts with sufficiently high median income ranks according to the 2013 American Communities Survey.

Table 6: Placebo Tests: No Active Age Sorting Within Census Blocks

Placebo for:	Proportion of city-wide sorting due to sorting within:			$\hat{\beta}^V$	$\hat{\beta}^N$
	Tracts	Block Groups	Blocks		
Atlanta	0.54	0.41	0.02	0.00	0.38
Chicago	0.60	0.44	0.02	0.00	0.25
Dallas	0.59	0.44	0.03	0.00	0.40
Los Angeles	0.59	0.44	0.03	0.00	0.35
New York City	0.60	0.42	0.04	0.00	0.48
Philadelphia	0.53	0.42	0.02	0.00	0.15
San Francisco	0.67	0.54	0.03	0.00	0.37
Washington, DC	0.65	0.53	0.02	0.00	0.19

Notes: All results are calculated under the placebo assumption that individuals do not actively sort within census blocks. Bootstrapped standard errors for all entries in all cities are less than 0.005 and are omitted for clarity.

Figure 15: Robustness (Age): Dynamic Aggregation

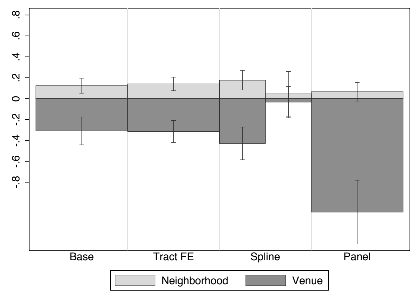


(a) (b) (c) (d)

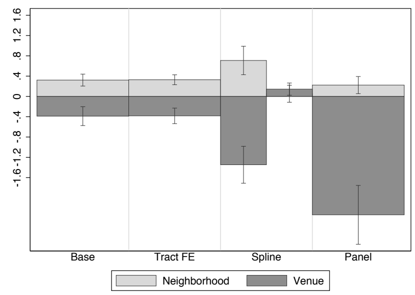
Notes: Solid lines refer to the y-axis on the left, and dashed lines refer to the y-axis on the right. All measures are calculated by replicating the analysis by day of week (first two panels) or by month (last two panels).

Figure 16: $\hat{\beta}^V$ and $\hat{\beta}^N$: Alternative Identification Strategies For Age Sorting by City (1 of 2)

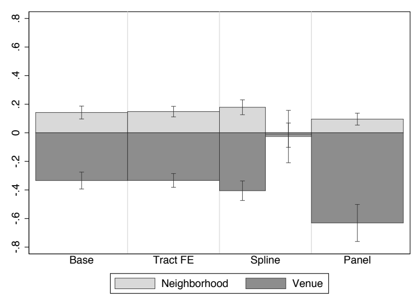
V_b = Category V_b = Subcategory



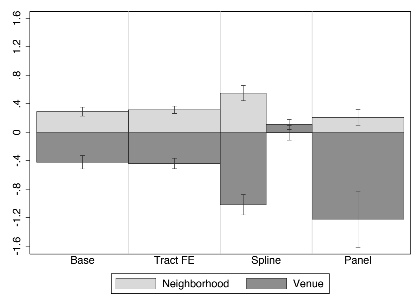
(a) Atlanta



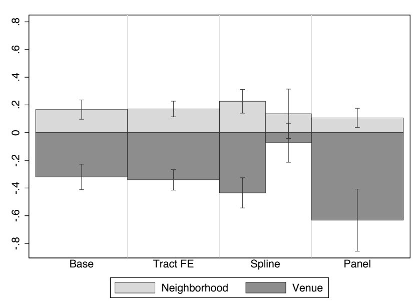
(b) Atlanta



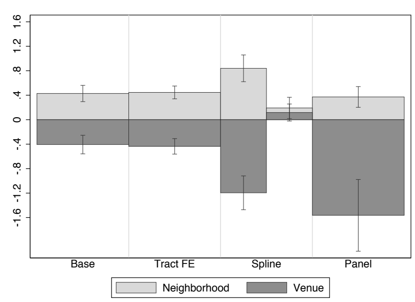
(c) Chicago



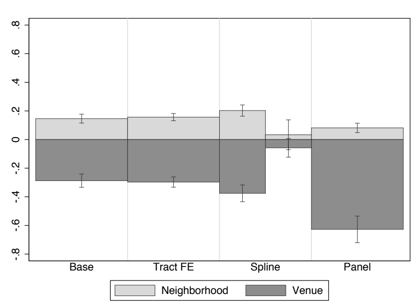
(d) Chicago



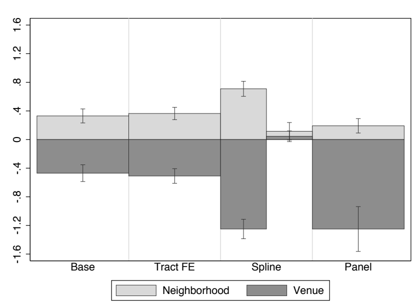
(e) Dallas



(f) Dallas



(g) Los Angeles

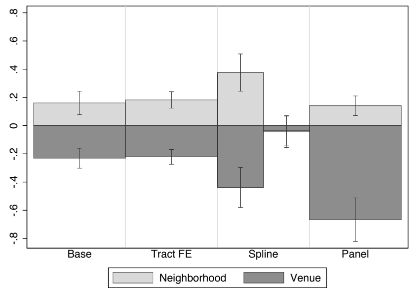


(h) Los Angeles

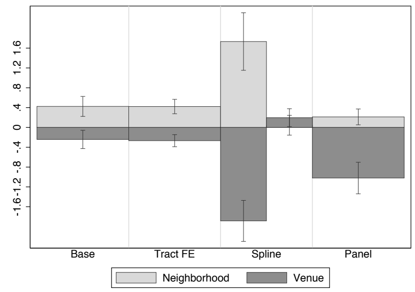
Note: See next page.

Figure 17: $\hat{\beta}^V$ and $\hat{\beta}^N$: Alternative Identification Strategies For Age Sorting by City (2 of 2)

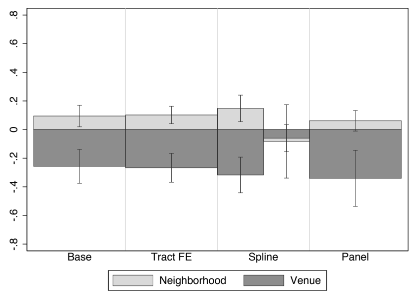
V_b = Category $\quad\quad$ V_b = Subcategory



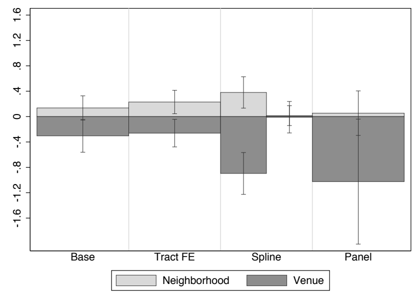
(a) New York City



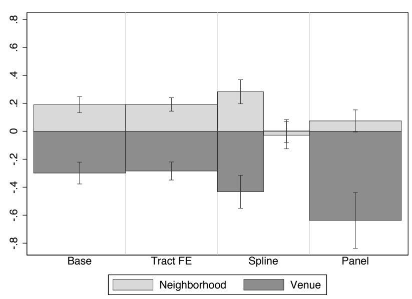
(b) New York City



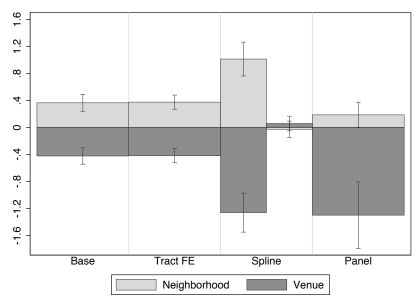
(c) Philadelphia



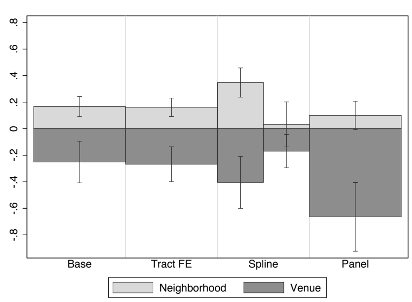
(d) Philadelphia



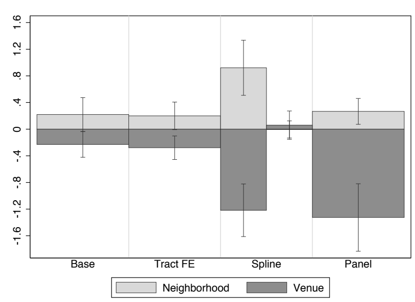
(e) San Francisco



(f) San Francisco



(g) Washington, DC



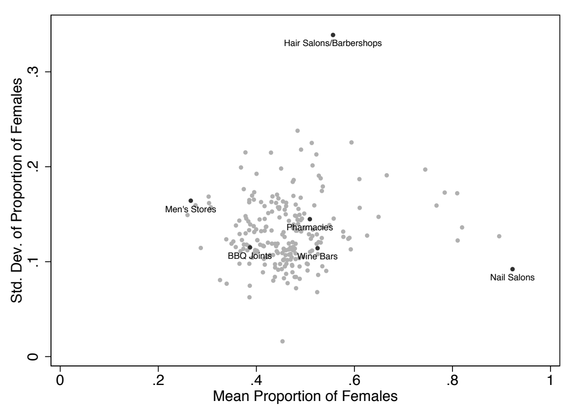
(h) Washington, DC

Notes: The dark shaded bars represent $\hat{\beta}^V$, and the light shared bars represent $\hat{\beta}^N$. Venue variety is defined as the number of unique venue categories in the first column and the number of unique venue subcategories in the second column. The first bars correspond to baseline estimates. The second bars replace the block group fixed effects in the baseline estimates with tract fixed effects. The third set of bars correspond to estimates of the parameters specified as a linear b-spline with a knot at 3 three categories or subcategories. The fourth bars correspond to estimates where the dataset is disaggregated to a monthly panel and the block group fixed effects are replaced with block fixed effects.

C Summary Statistics by Subcategory

The 9 categories of venues are further classified into 225 narrow subcategories. Foursquare users very actively check into even surprising types of venues such as *Banks*, *Cemeteries*, *Pharmacies, Synagogues*, and *Dog Runs.* In Figure 18, we present a scatter plot of the mean and standard deviation of the gender composition of venues for each subcategory throughout our entire sample. In general, the pattern of gender compositions of venues across subcategories looks intuitive and reasonable. For example, *Men's Stores*, not surprisingly, cater to mostly men, and this is fairly consistent across stores; conversely, *Nail Salons* cater to mostly women across all stores. *Hair Salons/Barbershops* cater to a mixed customer base in the aggregate; however the high standard deviation of the gender composition of these venues suggests that they may serve very different clientele – either predominantly male or predominantly female. In contrast, *Wine Bars*, which exhibit a similarly mixed clientele in the aggregate seem to also exhibit this mixed gender composition at the venue level. Although there are some small differences in the gender compositions of subcategories for different cities, their relative means and standard deviations tend to be stable.

Figure 18: Proportion of Females in Venues by Subcategory



Note: This scatter plot pools venues from all cities in the sample. Each dot represents all of the venues within a subcategory.

Full summary statistics disaggregated by subcategory can be found in the table below.

		Proportion of Females			Proportion of Youth				
Category	Subcategory	μ	σ	$p_{75} - p_{25}$	μ	σ	$p_{75} - p_{25}$	Venues	Check-ins
Bars	Bars	0.44	0.11	0.12	0.73	0.15	0.19	2114	2,379,893
Bars	Beer Gardens	0.42	0.10	0.14	0.71	0.16	0.22	104	163,102
Bars	Breweries	0.39	0.10	0.13	0.63	0.16	0.21	127	193,472
Bars	Cocktail Bars	0.47	0.10	0.12	0.72	0.15	0.18	322	368,735
Bars	Dive Bars	0.40	0.11	0.12	0.70	0.17	0.22	360	271,981
Bars	Gastropubs	0.46	0.08	0.09	0.71	0.11	0.12	212	332,930
Bars	Hookah Bars	0.44	0.11	0.16	0.89	0.07	0.09	104	48,287
Bars	Hotel Bars	0.42	0.12	0.14	0.58	0.16	0.23	243	118,877
Bars	Karaoke Bars	0.47	0.12	0.13	0.77	0.14	0.20	179	108,432
Bars	Lounges	0.46	0.13	0.14	0.70	0.18	0.22	530	394,834
Bars	Nightclubs	0.43	0.13	0.16	0.78	0.14	0.16	463	380,946
Bars	Other Nightlife	0.48	0.16	0.20	0.67	0.18	0.24	69	31,027
Bars	Pubs	0.44	0.09	0.10	0.71	0.15	0.16	517	693,208
Bars	Sake Bars	0.47	0.10	0.10	0.72	0.14	0.18	19	12,780
Bars	Speakeasies	0.47	0.11	0.13	0.75	0.14	0.16	90	100,912
Bars	Sports Bars	0.43	0.11	0.13	0.71	0.16	0.23	404	545,632
Bars	Strip Clubs	0.30	0.17	0.22	0.65	0.18	0.22	145	38,122
Bars	Whisky Bars	0.43	0.08	0.10	0.74	0.13	0.12	45	75,513
Bars	Wine Bars	0.52	0.11	0.13	0.66	0.15	0.17	347	243,684
Bars	Wineries	0.51	0.13	0.17	0.60	0.16	0.23	47	14,709
Cafes	Cafeterias	0.37	0.20	0.36	0.60	0.23	0.36	41	17,696
Cafes	Cafes	0.48	0.14	0.18	0.69	0.17	0.21	1280	699,162
Cafes	Coffee Shops	0.48	0.14	0.18	0.67	0.16	0.22	3162	3,160,881
Entertainment	Aquariums	0.54	0.09	0.12	0.69	0.08	0.11	17	35,862
Entertainment	Arcades	0.44	0.17	0.21	0.71	0.16	0.18	78	38,214
Entertainment	Art Galleries	0.49	0.15	0.17	0.69	0.16	0.18	233	41,570
Entertainment	Art Museums	0.47	0.11	0.12	0.65	0.12	0.10	133	296,219

		Proportion of Females			Proportion of Youth				
Category	Subcategory	μ	σ	$p_{75} - p_{25}$	μ	σ	$p_{75} - p_{25}$	Venues	Check-ins
Entertainment	Bowling Alleys	0.44	0.12	0.15	0.63	0.14	0.20	144	106,817
Entertainment	Casinos	0.34	0.12	0.18	0.59	0.12	0.18	12	15,262
Entertainment	Comedy Clubs	0.47	0.12	0.12	0.74	0.12	0.14	115	87,982
Entertainment	Concert Halls	0.46	0.11	0.12	0.67	0.17	0.23	116	166,265
Entertainment	General Entertainment	0.46	0.15	0.19	0.65	0.16	0.21	822	367,113
Entertainment	Historic Sites	0.39	0.14	0.18	0.59	0.17	0.21	154	135,739
Entertainment	History Museums	0.45	0.13	0.14	0.57	0.14	0.16	165	116,896
Entertainment	Indie Movie Theaters	0.47	0.11	0.11	0.65	0.13	0.16	92	85,124
Entertainment	Indie Theaters	0.48	0.13	0.15	0.70	0.15	0.18	70	36,177
Entertainment	Jazz Clubs	0.42	0.10	0.11	0.60	0.12	0.20	80	55,955
Entertainment	Movie Theaters	0.45	0.10	0.11	0.67	0.12	0.14	180	275,574
Entertainment	Multiplexes	0.49	0.08	0.08	0.69	0.09	0.11	106	432,959
Entertainment	Museums	0.47	0.11	0.12	0.61	0.12	0.17	174	152,003
Entertainment	Music Venues	0.42	0.12	0.13	0.69	0.16	0.21	295	371,347
Entertainment	Performing Arts Venues	0.49	0.14	0.14	0.67	0.16	0.23	189	149,840
Entertainment	Piano Bars	0.52	0.07	0.09	0.74	0.15	0.32	15	14,778
Entertainment	Pool Halls	0.42	0.16	0.18	0.74	0.20	0.22	39	23,962
Entertainment	Public Art	0.39	0.17	0.23	0.59	0.19	0.29	52	48,019
Entertainment	Racetracks	0.40	0.17	0.21	0.54	0.18	0.19	49	21,480
Entertainment	Rock Clubs	0.45	0.09	0.10	0.73	0.12	0.15	106	172,387
Entertainment	Science Museums	0.46	0.10	0.09	0.63	0.12	0.17	74	153,446
Entertainment	Stadiums	0.40	0.11	0.12	0.59	0.13	0.17	31	252,616
Entertainment	Theaters	0.48	0.13	0.15	0.66	0.15	0.19	482	333,304
Entertainment	Water Parks	0.49	0.15	0.17	0.59	0.19	0.25	9	2,936

		Proportion of Females			Proportion of Youth				
Category	Subcategory	μ	σ	$p_{75} - p_{25}$	μ	σ	$p_{75} - p_{25}$	Venues	Check-ins
Entertainment	Zoos	0.47	0.09	0.10	0.52	0.11	0.13	157	90,529
Food	African Restaurants	0.50	0.13	0.16	0.68	0.15	0.25	25	7,774
Food	American Restaurants	0.47	0.11	0.13	0.63	0.15	0.20	2444	1,845,916
Food	Argentinian Restaurants	0.44	0.08	0.11	0.65	0.11	0.12	37	14,020
Food	Asian Restaurants	0.46	0.11	0.14	0.70	0.14	0.19	779	357,845
Food	Australian Restaurants	0.53	0.11	0.11	0.82	0.09	0.13	14	14,714
Food	Bakeries	0.53	0.12	0.14	0.68	0.14	0.18	910	555,341
Food	BBQ Joints	0.39	0.12	0.16	0.60	0.16	0.23	453	297,253
Food	Brazilian Restaurants	0.41	0.09	0.11	0.67	0.10	0.15	66	40,859
Food	Breakfast Spots	0.48	0.11	0.13	0.64	0.14	0.19	661	405,581
Food	Burger Joints	0.40	0.11	0.13	0.65	0.14	0.19	1283	913,425
Food	Burrito Places	0.38	0.11	0.14	0.71	0.15	0.20	175	117,843
Food	Caribbean Restaurants	0.48	0.12	0.17	0.64	0.12	0.16	93	32,959
Food	Cuban Restaurants	0.46	0.10	0.14	0.66	0.14	0.15	106	86,381
Food	Cupcake Shops	0.59	0.13	0.15	0.71	0.14	0.14	146	112,410
Food	Delis / Bodegas	0.38	0.15	0.20	0.67	0.17	0.24	824	343,274
Food	Dim Sum Restaurants	0.46	0.08	0.09	0.71	0.10	0.14	79	62,700
Food	Diners	0.42	0.10	0.13	0.63	0.15	0.20	654	447,090

		Proportion of Females			Proportion of Youth				
Category	Subcategory	μ	σ	$p_{75} - p_{25}$	μ	σ	$p_{75} - p_{25}$	Venues	Check-ins
Food	Donut Shops	0.39	0.14	0.18	0.64	0.17	0.23	180	91,355
Food	Eastern European Restaurants	0.46	0.10	0.12	0.69	0.10	0.16	65	28,409
Food	Ethiopian Restaurants	0.47	0.10	0.11	0.74	0.10	0.12	50	15,807
Food	Falafel Restaurants	0.38	0.12	0.15	0.73	0.15	0.17	93	44,572
Food	Fast Food Restaurants	0.42	0.14	0.18	0.64	0.15	0.19	2422	647,960
Food	Filipino Restaurants	0.46	0.10	0.10	0.67	0.11	0.13	16	11,933
Food	Food Trucks	0.40	0.13	0.17	0.72	0.14	0.17	62 4	210,581
Food	French Restaurants	0.50	0.09	0.12	0.65	0.13	0.16	514	330,423
Food	Fried Chicken Joints	0.38	0.12	0.15	0.62	0.14	0.20	339	80,861
Food	German Restaurants	0.39	0.06	0.07	0.66	0.11	0.13	60	102,679
Food	Greek Restaurants	0.45	0.11	0.17	0.65	0.14	0.18	233	99,018
Food	Hot Dog Joints	0.37	0.11	0.14	0.63	0.14	0.19	272	172,987
Food	Ice Cream Shops	0.53	0.12	0.15	0.69	0.15	0.21	750	409,897
Food	Indian Restaurants	0.39	0.10	0.14	0.68	0.13	0.17	525	208,403
Food	Italian Restaurants	0.50	0.11	0.13	0.63	0.14	0.18	1828	897,952

		Proportion of Females			Proportion of Youth				
Category	Subcategory	μ	σ	$p_{75} - p_{25}$	μ	σ	$p_{75} - p_{25}$	Venues	Check-ins
Food	Japanese Restaurants	0.47	0.11	0.14	0.69	0.14	0.19	779	335,856
Food	Juice Bars	0.52	0.13	0.16	0.74	0.13	0.17	322	159,643
Food	Korean Restaurants	0.48	0.10	0.13	0.78	0.11	0.14	417	235,319
Food	Latin American Restaurants	0.46	0.11	0.14	0.71	0.14	0.17	207	100,725
Food	Malaysian Restaurants	0.48	0.12	0.11	0.75	0.10	0.14	22	14,836
Food	Mediterranean Restaurants	0.44	0.12	0.17	0.69	0.13	0.18	371	177,166
Food	Mexican Restaurants	0.44	0.12	0.16	0.64	0.16	0.22	2361	1,301,614
Food	Middle Eastern Restaurants	0.41	0.12	0.16	0.70	0.13	0.18	236	73,273
Food	Mongolian Restaurants	0.48	0.11	0.15	0.66	0.12	0.22	9	2,290
Food	Moroccan Restaurants	0.47	0.11	0.11	0.73	0.11	0.14	28	5,678
Food	New American Restaurants	0.49	0.09	0.12	0.66	0.13	0.17	366	393,351
Food	Peruvian Restaurants	0.48	0.08	0.07	0.70	0.10	0.13	21	22,046
Food	Pizza Places	0.40	0.12	0.16	0.69	0.15	0.20	1993	841,333
Food	Portuguese Restaurants	0.53	0.09	0.15	0.76	0.05	0.09	6	7,238

		Proportion of Females			Proportion of Youth				
Category	Subcategory	μ	σ	$p_{75} - p_{25}$	μ	σ	$p_{75} - p_{25}$	Venues	Check-ins
Food	Ramen / Noodle House	0.46	0.09	0.10	0.75	0.11	0.13	228	203,034
Food	Salad Shop	0.51	0.11	0.16	0.74	0.13	0.17	184	151,994
Food	Sandwich Places	0.38	0.13	0.17	0.70	0.15	0.19	2265	888,057
Food	Scandinavian Restaurants	0.48	0.08	0.11	0.70	0.10	0.18	20	18,902
Food	Seafood Restaurants	0.47	0.10	0.11	0.62	0.13	0.18	550	403,785
Food	Soup Places	0.51	0.11	0.16	0.75	0.10	0.15	63	45,932
Food	South American Restaurants	0.44	0.09	0.12	0.69	0.10	0.13	63	18,152
Food	Southern / Soul Food Restaurants	0.48	0.11	0.12	0.60	0.15	0.18	172	143,823
Food	Spanish Restaurants	0.49	0.10	0.13	0.67	0.12	0.15	76	38,048
Food	Steakhouses	0.44	0.10	0.12	0.56	0.12	0.16	458	309,250
Food	Sushi Restaurants	0.50	0.10	0.12	0.71	0.14	0.18	1207	512,255
Food	Tapas Restaurants	0.53	0.10	0.11	0.72	0.13	0.14	139	119,025
Food	Tea Rooms	0.58	0.13	0.15	0.78	0.15	0.17	253	187,002
Food	Thai Restaurants	0.46	0.11	0.13	0.71	0.13	0.17	815	316,483
Food	Turkish Restaurants	0.46	0.10	0.14	0.72	0.10	0.10	27	13,034
Food	Vegetarian / Vegan Restaurants	0.52	0.12	0.13	0.70	0.12	0.13	329	194,472

		Proportion of Females			Proportion of Youth				
Category	Subcategory	μ	σ	$p_{75} - p_{25}$	μ	σ	$p_{75} - p_{25}$	Venues	Check-ins
Food	Vietnamese Restaurants	0.45	0.10	0.12	0.73	0.12	0.15	392	173,671
Food	Wings Joints	0.43	0.11	0.14	0.72	0.13	0.17	226	139,954
Food	Yogurt	0.59	0.11	0.13	0.77	0.11	0.13	73	37,867
Gyms	Baseball Fields	0.41	0.14	0.17	0.58	0.20	0.26	121	25,378
Gyms	Baseball Courts	0.39	0.16	0.22	0.62	0.21	0.27	38	7,981
Gyms	Dance Studios	0.67	0.19	0.28	0.78	0.14	0.18	85	45,293
Gyms	Golf Courses	0.31	0.16	0.24	0.56	0.17	0.24	295	83,751
Gyms	Gyms	0.49	0.22	0.27	0.69	0.19	0.27	640	984,359
Gyms	Martial Arts Dojos	0.48	0.24	0.12	0.67	0.21	0.28	18	13,145
Gyms	Skate Parks	0.30	0.16	0.22	0.63	0.17	0.19	25	4,948
Gyms	Skating Rinks	0.45	0.16	0.18	0.59	0.17	0.21	65	33,786
Gyms	Soccer Fields	0.40	0.13	0.19	0.62	0.24	0.29	40	10,549
Gyms	Tennis Courts	0.42	0.13	0.15	0.60	0.16	0.18	45	15,974
Gyms	Tracks	0.48	0.14	0.16	0.67	0.16	0.25	27	34,870
Gyms	Yoga Studios	0.77	0.16	0.16	0.72	0.17	0.25	226	154,706
Hotels	Bed & Breakfasts	0.39	0.07	0.10	0.60	0.17	0.28	10	1,083
Hotels	Hotels Pools	0.43	0.15	0.20	0.60	0.15	0.23	24	4,453
Hotels	Hotels	0.40	0.11	0.12	0.59	0.15	0.16	1637	2,203,596
Hotels	Motels	0.36	0.12	0.15	0.65	0.16	0.19	111	22,408
Hotels	Resorts	0.39	0.17	0.16	0.55	0.18	0.35	16	9,713
Outdoors	Beaches	0.45	0.15	0.17	0.58	0.16	0.19	187	120,023
Outdoors	Cemeteries	0.50	0.15	0.23	0.52	0.16	0.23	88	26,666
Outdoors	Cities	0.49	0.14	0.18	0.60	0.15	0.21	255	567,323
Outdoors	Dog Runs	0.48	0.19	0.23	0.61	0.19	0.26	167	81,925
Outdoors	Farms	0.48	0.14	0.27	0.57	0.16	0.17	28	6,522

		Proportion of Females			Proportion of Youth				
Category	Subcategory	μ	σ	$p_{75} - p_{25}$	μ	σ	$p_{75} - p_{25}$	Venues	Check-ins
Outdoors	Fields	0.42	0.15	0.16	0.60	0.21	0.29	72	27,714
Outdoors	Gardens	0.47	0.14	0.14	0.57	0.16	0.20	131	63,519
Outdoors	Harbors / Marinas	0.43	0.12	0.16	0.59	0.17	0.23	175	98,020
Outdoors	Lakes	0.42	0.14	0.18	0.52	0.19	0.29	93	61,825
Outdoors	Monuments / Landmarks	0.37	0.12	0.14	0.55	0.13	0.19	110	233,864
Outdoors	Mountains	0.46	0.16	0.26	0.58	0.15	0.23	15	2,798
Outdoors	Neighborhoods	0.45	0.16	0.19	0.61	0.17	0.23	558	1,069,034
Outdoors	Other Great Outdoors	0.45	0.17	0.22	0.58	0.19	0.27	377	189,642
Outdoors	Parks	0.43	0.17	0.22	0.58	0.19	0.26	1329	1,176,399
Outdoors	Playgrounds	0.45	0.16	0.19	0.53	0.17	0.23	348	74,294
Outdoors	Plazas	0.39	0.16	0.21	0.58	0.19	0.24	317	532,148
Outdoors	Pools	0.47	0.18	0.22	0.63	0.21	0.30	132	33,801
Outdoors	Rivers	0.41	0.12	0.16	0.56	0.15	0.22	16	12,557
Outdoors	Scenic Lookouts	0.41	0.16	0.17	0.56	0.18	0.23	247	158,037
Outdoors	Sculpture Gardens	0.37	0.18	0.23	0.51	0.19	0.27	145	76,976
Outdoors	Ski Areas	0.45	0.02	0.03	0.61	0.13	0.27	3	1,792
Outdoors	Vineyards	0.61	0.16	0.22	0.63	0.22	0.32	2	181
Shops/Services	Accessories Stores	0.52	0.21	0.30	0.70	0.14	0.17	189	29,879
Shops/Services	Arts & Crafts Stores	0.65	0.15	0.21	0.66	0.13	0.19	295	89,275
Shops/Services	Automotive Shops	0.39	0.13	0.17	0.58	0.16	0.22	838	129,367

		Proportion of Females			Proportion of Youth				
Category	Subcategory	μ	σ	$p_{75} - p_{25}$	μ	σ	$p_{75} - p_{25}$	Venues	Check-ins
Shops/Services	Banks	0.43	0.16	0.23	0.65	0.18	0.26	1546	328,903
Shops/Services	Bike Shops	0.35	0.12	0.14	0.65	0.17	0.23	199	34,493
Shops/Services	Board Shops	0.37	0.12	0.18	0.71	0.17	0.26	60	8,419
Shops/Services	Bookstores	0.44	0.16	0.20	0.65	0.16	0.19	339	193,775
Shops/Services	Boutiques	0.59	0.23	0.38	0.74	0.14	0.19	487	108,881
Shops/Services	Bridal Shops	0.90	0.13	0.08	0.84	0.08	0.10	56	10,276
Shops/Services	Butchers	0.37	0.10	0.12	0.54	0.17	0.19	41	13,392
Shops/Services	Camera Stores	0.35	0.12	0.14	0.67	0.1	0.12	26	9,480
Shops/Services	Candy Stores	0.53	0.12	0.15	0.65	0.15	0.15	146	67,379
Shops/Services	Car Dealerships	0.38	0.14	0.20	0.59	0.15	0.19	182	35,351
Shops/Services	Car Wash	0.40	0.11	0.15	0.52	0.13	0.17	92	23,370
Shops/Services	Cheese Shops	0.50	0.10	0.11	0.66	0.12	0.19	32	23,550
Shops/Services	Clothing Stores	0.53	0.19	0.26	0.73	0.14	0.17	1498	644,653
Shops/Services	Cosmetics Shops	0.81	0.17	0.19	0.72	0.14	0.18	684	156,441
Shops/Services	Department Stores	0.59	0.12	0.15	0.64	0.11	0.14	757	997,837
Shops/Services	Design Studios	0.46	0.16	0.19	0.68	0.16	0.19	221	42,758
Shops/Services	Drugstores / Pharmacies	0.51	0.14	0.20	0.62	0.16	0.24	1525	577,272
Shops/Services	Electronics Stores	0.36	0.14	0.17	0.64	0.16	0.19	508	381,047
Shops/Services	Farmers Markets	0.51	0.12	0.16	0.58	0.15	0.20	199	147,509
Shops/Services	Financial or Legal Services	0.38	0.22	0.22	0.63	0.21	0.30	26	15,416
Shops/Services	Flea Markets	0.53	0.14	0.16	0.69	0.13	0.17	91	30,724
Shops/Services	Flower Shops	0.49	0.17	0.24	0.62	0.19	0.24	52	8,407
Shops/Services	Food & Drink Shops	0.46	0.15	0.21	0.62	0.18	0.28	122	40,589

		Proportion of Females			Proportion of Youth				
Category	Subcategory	μ	σ	$p_{75} - p_{25}$	μ	σ	$p_{75} - p_{25}$	Venues	Check-ins
Shops/Services	Food Courts	0.42	0.15	0.22	0.67	0.17	0.20	168	81,914
Shops/Services	Gaming Cafes	0.28	0.16	0.19	0.66	0.24	0.19	14	5,084
Shops/Services	Garden Centers	0.52	0.16	0.19	0.54	0.16	0.26	16	3,338
Shops/Services	Gift Shops	0.53	0.17	0.24	0.64	0.15	0.20	384	74,114
Shops/Services	Gourmet Shops	0.51	0.13	0.13	0.65	0.16	0.22	140	121,390
Shops/Services	Grocery Stores	0.49	0.13	0.17	0.62	0.15	0.21	1903	1,749,221
Shops/Services	Gyms or Fitness Centers	0.51	0.23	0.27	0.69	0.18	0.23	463	1,005,051
Shops/Services	Hardware Stores	0.37	0.11	0.12	0.54	0.15	0.22	321	142,828
Shops/Services	Health Food Stores	0.43	0.21	0.36	0.70	0.14	0.20	20	2,558
Shops/Services	Hobby Shops	0.40	0.19	0.31	0.65	0.16	0.21	69	22,991
Shops/Services	Jewelry Stores	0.61	0.19	0.27	0.69	0.15	0.23	178	39,630
Shops/Services	Kids Stores	0.63	0.13	0.21	0.59	0.13	0.15	59	8,974
Shops/Services	Lingerie Stores	0.81	0.12	0.14	0.76	0.13	0.14	145	46,643
Shops/Services	Liquor Stores	0.38	0.15	0.19	0.66	0.17	0.23	381	109,533
Shops/Services	Malls	0.49	0.16	0.17	0.63	0.17	0.19	302	779,691
Shops/Services	Markets	0.48	0.07	0.13	0.64	0.12	0.20	19	67,220
Shops/Services	Men's Stores	0.27	0.16	0.17	0.70	0.15	0.19	244	50,261
Shops/Services	Miscellaneous Shops	0.54	0.18	0.25	0.63	0.16	0.20	750	213,273
Shops/Services	Mobile Phone Shops	0.40	0.12	0.16	0.67	0.15	0.20	148	22,030
Shops/Services	Motorcycle Shops	0.33	0.08	0.13	0.50	0.15	0.14	10	1,718
Shops/Services	Music Stores	0.37	0.13	0.13	0.65	0.12	0.18	62	14,940
Shops/Services	Nail Salons	0.92	0.09	0.06	0.77	0.14	0.18	163	32,808
Shops/Services	Newsstands	0.37	0.14	0.14	0.53	0.24	0.37	24	3,743

		Proportion of Females			Proportion of Youth				
Category	Subcategory	μ	σ	$p_{75} - p_{25}$	μ	σ	$p_{75} - p_{25}$	Venues	Check-ins
Shops/Services	Optical Shops	0.52	0.13	0.19	0.70	0.16	0.19	76	14,180
Shops/Services	Paper / Office Supplies Stores	0.48	0.15	0.18	0.60	0.17	0.25	365	78,112
Shops/Services	Pet Stores	0.58	0.13	0.17	0.59	0.16	0.22	362	99,574
Shops/Services	Record Shops	0.34	0.08	0.09	0.64	0.12	0.14	123	49,108
Shops/Services	Salons / Barbershops	0.56	0.34	0.67	0.71	0.16	0.21	860	187,652
Shops/Services	Shoe Stores	0.51	0.20	0.32	0.71	0.14	0.18	500	117,650
Shops/Services	Smoke Shops	0.26	0.15	0.23	0.54	0.20	0.35	73	19,403
Shops/Services	Spas / Massages	0.78	0.17	0.23	0.71	0.14	0.19	526	114,510
Shops/Services	Sporting Goods Shops	0.41	0.13	0.15	0.62	0.14	0.18	375	155,592
Shops/Services	Tanning Salons	0.74	0.20	0.29	0.81	0.15	0.22	80	18,677
Shops/Services	Tattoo Parlors	0.55	0.14	0.17	0.74	0.16	0.17	138	19,843
Shops/Services	Thrift / Vintage Stores	0.56	0.15	0.19	0.69	0.16	0.23	319	62,505
Shops/Services	Toy / Game Stores	0.48	0.15	0.20	0.60	0.15	0.20	204	112,059
Shops/Services	Video Game Stores	0.29	0.11	0.13	0.71	0.15	0.19	147	27,089
Shops/Services	Video Stores	0.45	0.20	0.23	0.66	0.21	0.23	48	8,928
Shops/Services	Wine Shops	0.46	0.13	0.16	0.66	0.18	0.23	191	56,436
Shops/Services	Women's Stores	0.82	0.14	0.15	0.78	0.13	0.16	358	80,636
Spiritual	Churches	0.48	0.17	0.23	0.58	0.19	0.28	671	264,923
Spiritual	Synagogues	0.53	0.19	0.28	0.56	0.21	0.30	43	12,120
Spiritual	Temples	0.44	0.16	0.17	0.63	0.17	0.21	31	8,202