

Determinants of the presence, density, and popularity of U.S. food retailers

Syed Fuad, Syed Badruddoza, Modhurima Dey Amin
Department of Agricultural and Applied Economics, Texas Tech University
Correspondence: Syed.M.Fuad@ttu.edu

BACKGROUND

- Diet choices and quality may be influenced/limited by food environment.
- Poor access to foods such as fruits/vegetables, whole grains, and low-fat dairy products could lead to poor diets and eventually to health conditions such as obesity or diabetes.
- We use machine learning and traditional econometric models to extract census tract-level features that predict food store location, growth, density, and popularity.
- Identifying determinants that predict food stores may help tease out proactive approaches to alter food environments by encouraging healthy store options in underserved areas.

DATA & METHODS

- Point of Interest (POI) data from anonymized cellphone GPS ‘pings’ to obtain raw counts of stores and visits, frequency of visits, and category of the store type.
- POI data are complemented with census tract level demographic, housing, and socioeconomic data from ACS.
- Specify four models based on the nature of the response variables:
 - Presence of stores: Boosted Logit on binary variable indicating whether a particular food retailer type exists in the census tract;
 - Density: Negative Binomial on count number of food retailers of a particular type in a census tract;
 - Popularity: Negative Binomial on count number of customers who visited the particular food store type in a census tract; and
 - Growth: Boosted OLS on percentage growth of food retailers of a particular type.

RESULTS

Table 1. Models detect food store presence with a prediction accuracy of 72%-92% out of the sample.

	Large retail	Small retail	Fast food	Full service
Accuracy (%)	72.15	92.04	77.90	74.14
95% CI (%)	(71.48, 72.82)	(91.63, 92.44)	(77.27, 78.52)	(73.48, 74.79)
Sensitivity (%)	75.55	48.57	65.38	77.84
Specificity (%)	61.63	92.13	80.45	64.07
Kappa (%)	33.00	2.03	36.85	38.93

Table 2. Prediction of store count growth tends to be weak

	Large retail growth	Small retail growth	Fast food growth	Full-service growth
St. Dev. NRMSE (%)	100.006	100.013	100.003	100.023

*NRMSE close to zero implies high predictive accuracy

Figure 1. Features predicting presence of food stores

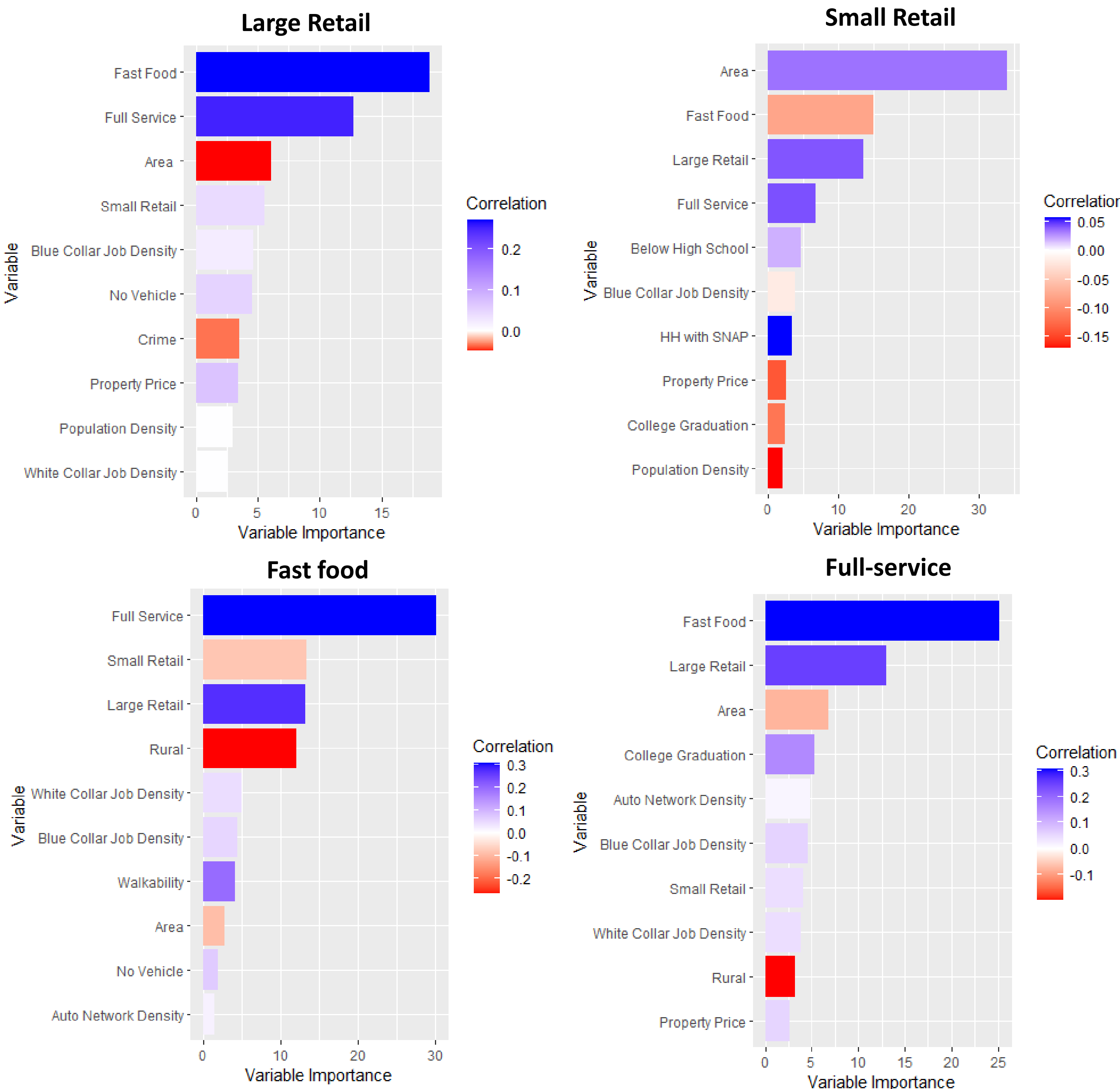


Figure 2. Low-income tracts less likely to have large retailers, fast food, full-service stores, but more likely to have small retailers.

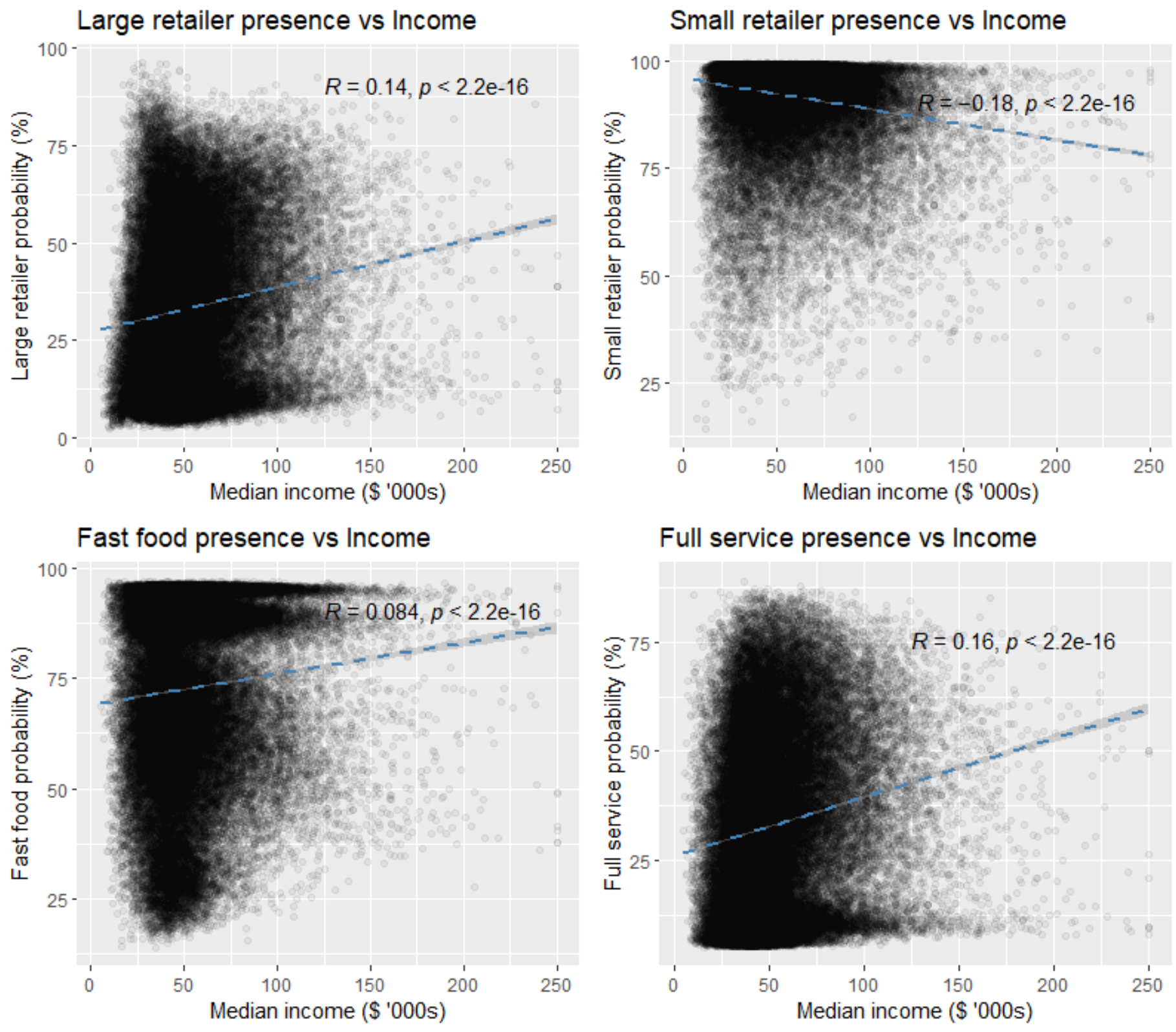


Figure 3. Non-white tracts less likely to have large and small retailers, full-service stores, but more likely to have fast food stores.

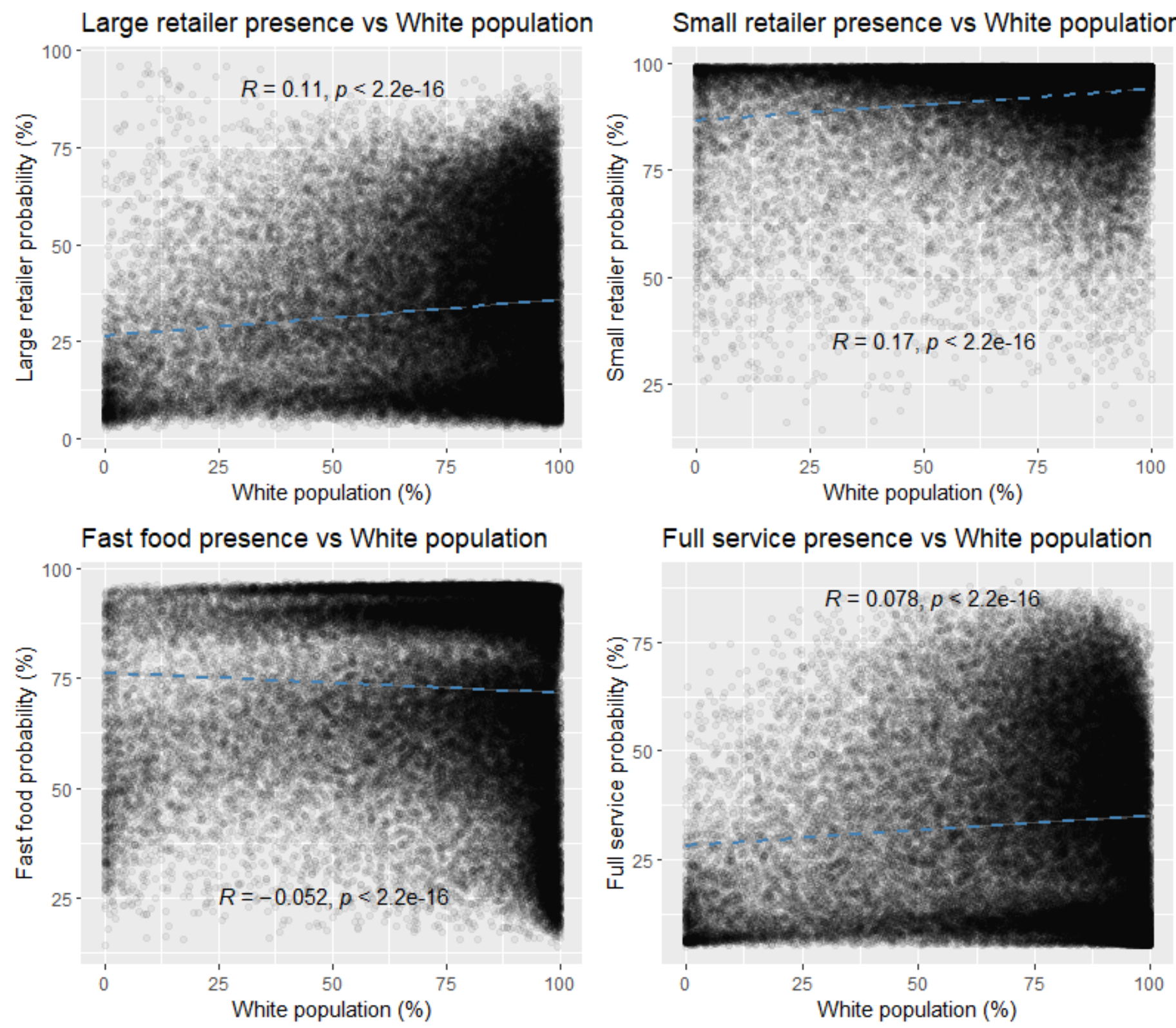


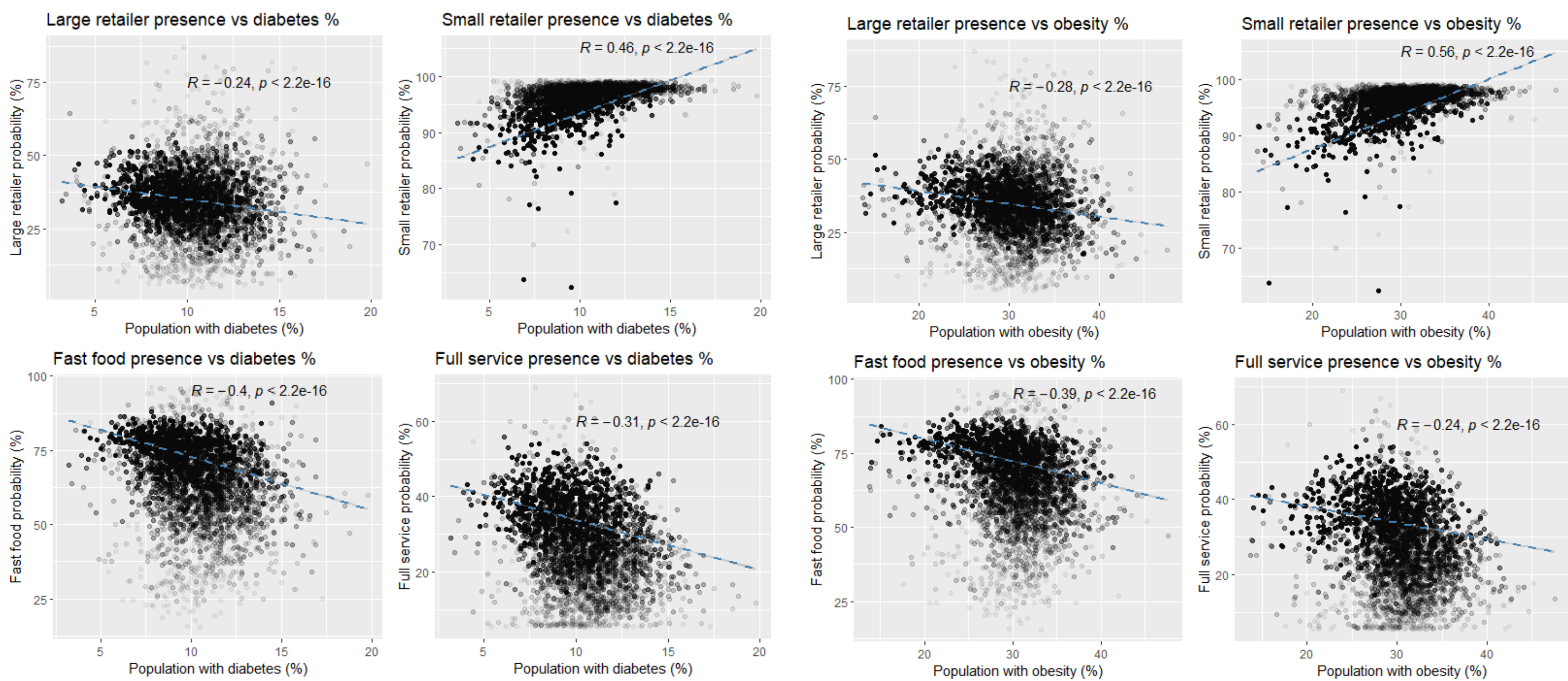
Table 3. Factors predicting store density and popularity

	Large retail		Small retail		Fast food		Full service	
	Density	Popularity	Density	Popularity	Density	Popularity	Density	Popularity
Food stores: Location, density and popularity of store type is consistently and significantly related to the same features of other store types*								
Large Retail	-	-	13.84	0.02	30.11	0.05	9.49	0.03
Small Retail	8.03	0.01	-	-	14.97	0.03	8.60	0.03
Fast Food	7.53	0.04	7.95	0.03	-	-	16.66	0.05
Full Service	-3.37	0.00	0.84	0.01	12.96	0.02	-	-
Wealth: Indicators of wealth strongly associated with healthy food stores								
Poverty Rate	-1.27	-1.77	-1.16	-1.60	0.42	0.09	-0.57	-0.67
HH with SNAP	0.07	0.12	0.09	0.08	-0.01	0.02	-0.04	0.00
Unemployment	-0.92	-1.22	0.89	1.28	0.21	0.6	-1.10	-1.81
Below HS	-0.33	-0.25	-0.07	-0.25	-0.15	0.22	-0.13	-0.03
Property Value	0.04	0.01	-0.02	-0.06	-0.01	-0.02	-0.04	-0.03
Location and Access: Healthy food stores located in small, urban tracts; greater access increases foot traffic to all store types								
Pub. Transport	-1.05	-2.14	-0.80	-1.16	-0.64	-0.56	-4.46	-3.87
No Vehicle	0.03	0.04	0.04	0.05	0.05	0.06	0.04	0.04
Land area	-0.07	-0.49	-2.96	-2.00	-0.91	-1.16	-1.16	-1.05
Rural population	-0.42	-0.91	0.03	0.10	-0.75	-1.08	-0.79	-1.10
Walkability	1.78	2.18	0.32	-1.06	1.74	2.28	1.14	1.93
Transit Density	0.81	0.85	0.01	1.15	0.08	0.71	0.11	4.10
Race and Ethnicity: Less healthy food stores tend to be located in Hispanic tracts								
Black	-0.41	-0.55	0.03	0.07	-0.32	-0.31	-0.19	-0.42
Hispanic	-0.46	-0.05	0.08	0.37	0.33	0.25	0.11	0.26
Asian	-0.26	-0.62	0.10	0.59	0.51	0.57	0.60	0.37
Sales tax positively related to small food stores; crime generally negatively related								
Sales Tax	-3.52	-5.43	0.43	2.23	2.79	5.23	0.82	3.23
Crime	-0.50	-0.20	-0.07	0.25	-0.10	-0.04	0.55	0.82

*All estimates significant at 1% level.

**Numbers represent % change in dependent variable with a unit change in independent variable

Figure 4. Small retailer presence has a higher association with diabetes and obesity rates



CONCLUSION

- Models detect food store presence with a prediction accuracy of 72%-92% out of the sample.
- Non-white and low-income neighborhoods have lower access to healthful food options.
- Nearby businesses are often complementary and helpful in location selection and driving customer traffic to other stores; small stores benefit more from traffic spillover from large retailers.
- Reliance on public transportation decreases foot traffic to food stores of all types; more walkable tracts positively related to store counts and visits.