Measuring Competition in Spatial Retail

Paul B. Ellickson[1] Paul L.E. Grieco [2] Oleksii Khvastunov[2]

[1]University of Rochester

[2]Penn State University

October, 2018

Introduction

- We study spatial competition between modern retail platforms.
- Active (and contested) area of anti-trust enforcement.
- Our challenges
  - Observe only store revenues.
  - Don't see prices or assortments.
  - Many outlets, several formats. Overlapping geographies.
- Given this data, what can be said about spatial retail competition?

Agenda: Why should you care?

- Retail is big (globally)
- Modern retail systems are platform oligopolies
  - Market power/foreclosure are potential concerns
- Modern retail systems key source of productivity/welfare gains
  - Increasing evidence that gains are regressive, urban-centric
  - Atkin et al (2018), Lagakos (2016), Handbury (2013)
- Not yet clear how these firms compete (price, assortment, format)
- Interplay between demand and cost sides

Research Agenda: This Paper

- We propose a simple framework for linking store revenues to consumer (census tract-level) demographics
  - Spatial logit model of expenditure allocation/store choice by heterogeneous consumers
  - In lieu of prices, include chain fixed effects that vary with income
- Apply to merger screening problem
  - Light data and modeling requirements
  - Delivers rich (and sensible) substitution patterns that reflect the heterogeneity and spatial location of consumers
  - Yields highly localized measures of concentration (tract or store level HHIs) for merger analysis
  - Provides store and firm level diversion ratios as input to UPP/partial simulation

Model: Consumer's Choice Problem

- Extend Holmes' (2011) revenue model to include multiple firms.
  - Spatial logit model, aggregated to store-level data.
- Consumers allocate grocery expenditures across competing outlets within D miles of home, or choose outside good.
  - Consumers are heterogenous, differentiated by location and income.
  - Stores have characteristics x_s, including possible chain affiliation.
- We assume a representative household at every census tract, indexing consumers by their home tract t.
- Consumers are endowed with a location (t) and characteristics z_t (e.g. income, car) that affect their utility for groceries.
- Consumers' food budgets (including spending on outside good) are a fixed proportion α of income.
  - But wealthy consumers may spend more outside grocery channel.

Utility Framework: Nested Logit

- Individuals allocate budgets via DC-RUM over nearby stores, endowed with locations and characteristics.
- Each consumer makes continuum of purchasing decisions.
- For each unit of expenditure i, consumer t's utility for spending at store s is

$$u_{sti} = u_{st} + \varepsilon_{sti} = \tau_0 d_{st} + \tau_1 d_{st} z_t + \gamma_0 x_s + \gamma_1 x_s \otimes z_t + \varepsilon_{sti}.$$

- Note that u_{st} is a function of distance d_{st}, store characteristics x_s, and tract-level consumer demographics z_t.
  - Store characteristics include size, checkouts, and FTEs.
  - Also include **fixed effects** for all large chains (+ interact with **income**).
- Each purchase decision is subject to an iid shock ε_{sit}, distributed GEV with nesting structure on formats (described below).

Role of Outside Good

- We assume choice set includes all stores within $D = 10$ miles of home tract, plus outside option, $C_t = \{s : d_{ts} \leq D\} \cup 0$.
- Spending on the outside good is moderated by demographics z_t and tract characteristics w_t that control for alternative consumption options in the tract's proximity,

$$u_{0ti} = \lambda_0 w_t + \lambda_1 w_t \otimes z_t + \varepsilon_{0ti}.$$

  - w_t includes population density and household size.
- Note that consumer's income impacts spending via two pathways:
  1. their overall budget ($\alpha \cdot \text{inc}_t$), and
  2. their choice of store (including outside good).

Nesting Structure: Alternative Store Formats

- We are particularly interested in evaluating competition from new formats (e.g. clubs, supercenters and organics)
- To allow for stronger substitution within format, we group firms into K nests, with ε_{sti} correlated across stores in same nest.
  - By integrating over ε_{sti}, we obtain the share of the budget that consumers in tract t spend at store s as a function of the model's parameters, $\theta = (\tau, \gamma, \lambda, \beta, \mu)$, and observed covariates.
- Given nesting structure, share of spending at store s (as a fraction of all spending in tract t) can be decomposed as follows

$$p_{st}(\theta) \equiv \Pr(\iota_{ti} = s) = \Pr(\iota_{ti} \in C_{t,k(s)}) \Pr(\iota_{ti} = s | \iota_{ti} \in C_{t,k(s)}).$$

where $\Pr(\iota_{ti} \in C_{t,k(s)})$ is the probability of choosing any store in nest $C_{t,k(s)}$ and $\Pr(\iota_{ti} = s | \iota_{ti} \in C_{t,k(s)})$ is the probability of choosing a particular store, given that you are choosing it from nest $C_{t,k(s)}$.

Nesting Structure: Alternative Store Formats

- Given GEV structure, the share of expenditure on stores in $C_{t,k(s)}$ (e.g. any club store close to tract t) is

$$\Pr(\iota_{ti} \in C_{t,k(s)}) = \frac{\left(\sum_{q \in C_{t,k(s)}} e^{u_{qt}/\mu_{k(s)}}\right)^{\mu_{k(s)}}}{\sum_{v=0}^{K}\left(\sum_{q \in C_{t,v}} e^{u_{qt}/\mu_v}\right)^{\mu_v}}.$$

- The probability of choosing a particular store s from nest $C_{t,k(s)}$ (e.g. a Sam's Club near t) is then

$$\Pr\left(\iota_{ti} = s | \iota_{ti} \in C_{t,k(s)}\right) = \frac{e^{u_{st}/\mu_{k(s)}}}{\sum_{q \in C_{t,k(s)}} e^{u_{qt}/\mu_{k(s)}}}.$$

- Finally, the unconditional share is given by

$$p_{st}(\theta) = \frac{e^{u_{st}/\mu_{k(s)}}\left(\sum_{q \in C_{t,k(s)}} e^{u_{qt}/\mu_{k(s)}}\right)^{\mu_{k(s)}-1}}{\sum_{v=0}^{K}\left(\sum_{q \in C_{t,v}} e^{u_{qt}/\mu_v}\right)^{\mu_v}}.$$

Moving from Choices to Revenues

- We observe store-level revenues, so we must aggregate up to them.
- Predicted revenue at store s coming from tract t is given by

$$\hat{R}_{st}(\theta, \alpha) = \alpha \text{inc}_t \cdot n_t \cdot p_{st}(\theta),$$

  where inc_t is PC income and n_t is total population residing in tract t.
- We assume store s collects revenue from all tracts for which it's in choice set (i.e. all tracts within 10 miles of its location).
- Therefore, predicted total revenue for store s is

$$\hat{R}_s(\theta, \alpha) = \sum_{t \in L_s} R_{st}(\theta, \alpha),$$

  where $L_s = \{t : s \in C_t\} = \{t : d_{st} \leq D\}$ is the set of tracts for which store s is included in some consumer's choice set.

Estimation

- We estimate parameters by matching *model-generated revenue predictions* to the *store-level* revenues *observed in the data*.
- Assuming these observed revenues R_s are perturbed by a multiplicative shock,

$$R_s = e^{\eta_s} \hat{R}_s(\theta_0, \alpha_0),$$

  where (θ_0, α_0) are true parameters of the DGP and η_s is the shock.
- Assuming η_s is mean zero and independent of exogenous variables, parameters can be estimated via NLLS,

$$(\hat{\theta}, \hat{\alpha}) = \underset{\theta,\alpha}{\operatorname{argmin}} \sum_s \left(\log(\hat{R}_s(\theta, \alpha)) - \log(R_s)\right)^2.$$

identification

Data: Sources and Content

- Grocery data come from Trade Dimension's 2006 TDLinx dataset.
- Observe all grocery stores, supermarkets, supercenters and club stores earning at least 2 million in revenues.
  - Focus on stores (and consumers) located in 317 MSAs (dropping NYC).
- Data include revenues, store features (size, FTEs, and checkouts), and full ownership structure.
  - Note: we do not observe FTEs or checkouts for clubs.
- Demographic information comes from the 2010 US Census.
  - GeoLocation, per capita income, vehicle ownership and household size.

Data Summary: Store Characteristics

Table 1: Store Characteristics by Type of Chain

	Mean	St. Dev.	1st Quartile	Median	3rd Quartile
Small and Medium Grocery Chains					
39.02 % of all MSA stores, 18.16 % of MSA Revenue					
Store Size in 1000 sqft	22.32	16.45	11	18	30
Store Weekly Volume in 1000s	182.34	174.40	80	125	225
Full Time Employee Equivalents	45.73	44.61	22	33	55
Checkouts	6.63	4.11	4	6	8
Revenue Per Square Feet	9.71	9.82	5.65	7.56	10.36
Large Grocery Chains					
49.87 % of all MSA stores, 47.17 % of MSA Revenue					
Store Size in 1000 sqft	36.74	15.51	25	37	48
Store Weekly Volume in 1000s	370.45	219.45	200	350	500
Full Time Employee Equivalents	69.34	43.61	37	64	93
Checkouts	9.56	3.96	7	9	11
Revenue Per Square Feet	10.46	5.72	6.67	9.29	12.50
Supercenters					
7.06 % of all MSA stores, 17.88 % of MSA Revenue					
Store Size in 1000 sqft	64.18	9.68	60	68	70
Store Weekly Volume in 1000s	991.51	333.48	725	1,025	1,225
Full Time Employee Equivalents	337.52	123.81	278	342	408
Checkouts	27.97	6.27	25	30	32
Revenue Per Square Feet	15.29	4.20	12.50	15.48	18.12
Club Stores					
4.03 % of all MSA stores, 16.76 % of MSA Revenue					
Store Size in 1000 sqft	124.75	16.06	113	130	135
Store Weekly Volume in 1000s	1,627.90	742.22	1,125	1,500	1,975
Revenue Per Square Feet	12.96	5.54	8.86	11.84	15.53
All Stores					
24,117 stores in 317 MSAs					
Store Size in 1000 sqft	36.60	26.26	17	32	49
Store Weekly Volume in 1000s	391.65	412.74	125	250	500
Full Time Employee Equivalents	79.49	91.35	28	52	89
Checkouts	9.73	6.81	5	8	11
Revenue Per Square Feet	10.61	7.65	6.36	9.00	12.75

additional tables

Data Summary: Census Tracts

Table 4: Census tracts: Demographic and choice set variation

	Mean	St. Dev.	1st Quartile	Median	3rd Quartile
Population	4,381.67	1,984.38	3,001	4,119	5,444
Average income	28.05	14.02	18.96	25.29	33.59
Population Density	2,862.98	3,013.04	846.48	2,043.98	3,733.49
Household size	2.43	0.59	2.11	2.38	2.69
Stores within 5 miles	20.19	19.70	6	15	28
Stores within 10 miles	59.52	58.57	16	41	84
Large chain within 5 miles	11.30	10.51	3	9	17
Large chain within 10 miles	33.82	31.99	9	25	50
Club stores within 5 miles	0.77	0.89	0	1	1
Club stores within 10 miles	2.33	2.11	1	2	4

Model: Parameter Estimates

Table 5: Parameter estimates.

	Baseline (1)	Multinomial Logit (2)	No Clubs (3)	No FTE/Checkouts (4)
Grocery Stores and Supercenters				
dist	−0.169	−0.197	−0.177	−0.177
	(0.001)	(0.001)	(0.001)	(0.001)
dist*log(inc)	−0.109	−0.144	−0.115	−0.109
	(0.002)	(0.003)	(0.002)	(0.002)
log(size)	0.151	0.207	0.153	0.399
	(0.002)	(0.003)	(0.002)	(0.002)
log(size)*log(inc)	0.131	0.173	0.107	0.273
	(0.008)	(0.010)	(0.007)	(0.005)
log(fte)	0.240	0.317	0.244	
	(0.002)	(0.002)	(0.002)	
log(fte)*log(inc)	−0.117	−0.150	−0.124	
	(0.007)	(0.009)	(0.006)	
log(chk)	0.217	0.299	0.222	
	(0.003)	(0.004)	(0.003)	
log(chk)*log(inc)	0.255	0.339	0.263	
	(0.012)	(0.014)	(0.010)	
Club Stores				
dist	−0.050	0.021		−0.051
	(0.008)	(0.006)		(0.007)
dist*log(inc)	−0.184	−0.297		−0.175
	(0.019)	(0.017)		(0.018)
log(size)	0.680	0.844		0.622
	(0.054)	(0.058)		(0.051)
log(size)*log(inc)	0.127	0.376		0.111
	(0.176)	(0.183)		(0.169)
Outside option				
hhsize	0.472	0.650	0.506	0.455
	(0.005)	(0.008)	(0.005)	(0.005)
hhsize*log(inc)	0.553	0.642	0.700	0.546
	(0.011)	(0.018)	(0.010)	(0.010)
log(density)	1.482	2.207	1.780	1.438
	(0.134)	(0.148)	(0.129)	(0.122)
log(density)^2	−0.130	−0.237	−0.226	−0.141
	(0.054)	(0.064)	(0.052)	(0.048)
$\mu_{grocery}$	0.737		0.746	0.723
	(0.020)		(0.021)	(0.018)
$\mu_{supercenters}$	0.752		0.773	0.642
	(0.056)		(0.055)	(0.052)
μ_{club}	0.785			0.762
	(0.104)			(0.099)
α	0.132	0.112	0.113	0.133
	(0.004)	(0.002)	(0.003)	(0.004)
R^2	0.840	0.836	0.812	0.807

Notes: All specifications include chain effects which vary with income. Standard errors in parentheses.

Parameter Estimates: Nesting Parameters, Budget and Fit

Table 5: Parameter estimates.

	Baseline (1)	Multinomial Logit (2)	No Clubs (3)	No FTE/Checkouts (4)
$\mu_{grocery}$	0.737		0.746	0.723
	(0.020)		(0.021)	(0.018)
$\mu_{supercenters}$	0.752		0.773	0.642
	(0.056)		(0.055)	(0.052)
μ_{club}	0.785			0.762
	(0.104)			(0.099)
α	0.132	0.112	0.113	0.133
	(0.004)	(0.002)	(0.003)	(0.004)
R^2	0.840	0.836	0.812	0.807

- FEs and slopes are reported in Appendix of paper.

Parameter Estimates: Outside Good

Table 5: Parameter estimates.

	Baseline (1)	Multinomial Logit (2)	No Clubs (3)	No FTE/Checkouts (4)
Outside option				
hhsize	0.472	0.650	0.506	0.455
	(0.005)	(0.008)	(0.005)	(0.005)
hhsize*log(inc)	0.553	0.642	0.700	0.546
	(0.011)	(0.018)	(0.010)	(0.010)
log(density)	1.482	2.207	1.780	1.438
	(0.134)	(0.148)	(0.129)	(0.122)
$\log(\text{density})^2$	−0.130	−0.237	−0.226	−0.141
	(0.054)	(0.064)	(0.052)	(0.048)

Parameter Estimates: Store Characteristics

Table 5: Parameter estimates.

	Baseline (1)	Multinomial Logit (2)	No Clubs (3)	No FTE/Checkouts (4)
Grocery Stores and Supercenters				
dist	−0.169	−0.197	−0.177	−0.177
	(0.001)	(0.001)	(0.001)	(0.001)
dist*log(inc)	−0.109	−0.144	−0.115	−0.109
	(0.002)	(0.003)	(0.002)	(0.002)
log(size)	0.151	0.207	0.153	0.399
	(0.002)	(0.003)	(0.002)	(0.002)
log(size)*log(inc)	0.131	0.173	0.107	0.273
	(0.008)	(0.010)	(0.007)	(0.005)
log(fte)	0.240	0.317	0.244	
	(0.002)	(0.002)	(0.002)	
log(fte)*log(inc)	−0.117	−0.150	−0.124	
	(0.007)	(0.009)	(0.006)	
log(chk)	0.217	0.299	0.222	
	(0.003)	(0.004)	(0.003)	
log(chk)*log(inc)	0.255	0.339	0.263	
	(0.012)	(0.014)	(0.010)	
Club Stores				
dist	−0.050	0.021		−0.051
	(0.008)	(0.006)		(0.007)
dist*log(inc)	−0.184	−0.297		−0.175
	(0.019)	(0.017)		(0.018)
log(size)	0.680	0.844		0.622
	(0.054)	(0.058)		(0.051)
log(size)*log(inc)	0.127	0.376		0.111
	(0.176)	(0.183)		(0.169)

Demographic Effects

- So what do the estimates imply about consumer tastes?
- Using the model, we can compute the revenue elasticity of each store with respect to distance or income.
  - To construct a measure of *chain-level* response, we aggregate up, weighting by revenue shares.
- The distance elasticity for revenue at store s from tract t is

$$\eta_{st} = \frac{\partial R_{st}}{\partial d_{st}} \frac{d_{st}}{R_{st}} = d_{st}(\tau_0 + \tau_1 z_t) \left(\frac{1}{\mu_{k(s)}} + \left(1 - \frac{1}{\mu_{k(s)}}\right) p_{st|k} - p_{st} \right),$$

  where $p_{st} = p_{st}(\theta)$ and $p_{st|k} = \Pr\left(\iota_{ti} = s | \iota_{ti} \in C_{t,k(s)}\right)$ are the relevant unconditional and conditional choice probabilities.
- The corresponding income elasticity is

$$\nu_{st} = 1 + \sum_{q \in C_t \setminus 0} (\tau_1 d_{qt} + \gamma_1 x_q) \left(1[s = q] \frac{1}{\mu_{k(s)}} + 1[q \in C_{t,k(s)}] \left(1 - \frac{1}{\mu_{k(s)}}\right) p_{qt|k} - p_{qt} \right) - \lambda_1 w_t p_{0t}.$$

Distance and Income Elasticities

Table 6: Distance and Income Elasticities Large Chains and Clubs

	Distance Elasticity	Income Elasticity
Small Chains	−1.075	0.416
Medium Chains	−1.092	0.683
Albertsons	−1.074	0.693
Aldi	−1.103	0.516
Bashas Markets	−1.090	0.662
Delhaize America (Food Lion)	−1.089	0.631
Fred Meyer	−1.116	0.851
Giant Eagle	−1.101	0.870
Giant Food	−1.218	0.514
Great A & P Tea Co.	−1.145	0.613
HE Butt	−0.972	0.779
Hannaford Bros	−1.032	0.521
Hy Vee Food Stores	−0.990	0.789
Ingles Markets	−1.070	0.657
Kroger	−1.095	0.662
Lone Star Funds (Bi-Lo)	−1.058	0.792
Publix	−1.122	0.773
Raleys	−1.005	0.481
Roundys	−1.078	0.491
Ruddick Corp (Harris Teeter)	−1.182	0.749
Safeway	−1.151	0.484
Save A Lot	−1.056	0.549
Save Mart	−0.867	0.502
Smart & Final	−1.071	0.281
Stater Bros	−1.015	0.410
Stop & Shop	−1.169	0.702
SuperValu	−1.145	0.563
Trader Joes	−1.158	0.253
Weis Markets	−1.083	0.630
Whole Foods	−1.197	0.525
Wild Oats	−1.145	0.449
Winn-Dixie	−1.031	0.731
Meijer	−0.966	0.506
Target	−1.126	0.620
Wal Mart	−0.874	0.741
BJs	−0.491	0.191
Costco	−0.585	0.509
Sam's Club	−0.386	0.413

Competitive Effects

- Since we don't observe prices, we can't calculate price elasticities.
  - But we can construct semi-elasticities for a Δ improvement in the (vertical) quality offered by a given chain.
    - The semi-elasticity for chain f wrt g is the percent decrease in revenue at f due to a Δ improvement in the chain FE for stores in g.
- Formally, the semi-elasticity is given by

$$\sigma_{f,g} = \frac{1}{R^f} \sum_{s \in F_f} \sum_{t \in L_s} R_{st} \sum_{q \in F_g \cap C_t} \left(1[s = q] \frac{1}{\mu_{k(s)}} + 1[q \in C_{t,k(s)}] \left(1 - \frac{1}{\mu_{k(s)}} \right) p_{qt|k(s)} - p_{qt} \right), \quad (1)$$

  where R^f is total revenue for chain f and F_f and F_g are the stores in chains f and g respectively.
- Recall that L_s is the set of tracts featuring store s in their choice set and C_t is the choice set of consumers in tract t.

Competitive Effects: Own and Cross Semi-Elasticities

Table 7: Competition Between Chains: Own and Cross Semi-elasticities

Chain	Own Semi-Elasticity	First Comp	Cross Semi-Elasticity	Second Comp	Cross Semi-Elasticity	Outside Cross Semi-Elasticity
Small Chains	1.112	Medium Chains	-0.104	Kroger	-0.082	-0.381
Medium Chains	1.002	Wal Mart	-0.097	Small Chains	-0.086	-0.324
Albertsons	1.162	Wal Mart	-0.133	Safeway	-0.110	-0.330
Aldi	1.360	Medium Chains	-0.178	Small Chains	-0.143	-0.323
Bashas Markets	1.026	Kroger	-0.241	Safeway	-0.146	-0.257
Delhaize America (Food Lion)	1.108	Wal Mart	-0.156	Medium Chains	-0.089	-0.331
Fred Meyer	1.078	Safeway	-0.198	SuperValu	-0.135	-0.329
Giant Eagle	1.104	Small Chains	-0.155	Medium Chains	-0.129	-0.332
Giant Food	1.099	Safeway	-0.116	Small Chains	-0.088	-0.451
Great A & P Tea Co.	1.256	Small Chains	-0.164	Kroger	-0.107	-0.385
HE Butt	0.710	Wal Mart	-0.163	Sam's Club	-0.062	-0.264
Hannaford Bros	0.890	Medium Chains	-0.165	SuperValu	-0.134	-0.319
Hy Vee Food Stores	0.948	Medium Chains	-0.194	Wal Mart	-0.170	-0.283
Ingles Markets	1.121	Wal Mart	-0.172	Lone Star Funds (Bi-Lo)	-0.123	-0.298
Kroger	0.956	Wal Mart	-0.112	Medium Chains	-0.076	-0.303
Lone Star Funds (Bi-Lo)	1.152	Wal Mart	-0.226	Delhaize America (Food Lion)	-0.105	-0.298
Publix	0.909	Wal Mart	-0.137	Winn-Dixie	-0.095	-0.305
Raleys	1.058	Safeway	-0.165	Small Chains	-0.088	-0.383
Roundys	1.060	Medium Chains	-0.153	SuperValu	-0.143	-0.405
Ruddick Corp (Harris Teeter)	1.161	Delhaize America (Food Lion)	-0.192	Medium Chains	-0.120	-0.361
Safeway	1.103	Kroger	-0.104	SuperValu	-0.084	-0.409
Save A Lot	1.297	Small Chains	-0.139	Medium Chains	-0.135	-0.310
Save Mart	1.041	Small Chains	-0.140	Safeway	-0.127	-0.378
Smart & Final	1.322	Kroger	-0.155	Safeway	-0.150	-0.422
Stater Bros	1.092	Kroger	-0.161	SuperValu	-0.131	-0.376
Stop & Shop	1.033	Medium Chains	-0.166	SuperValu	-0.130	-0.402
SuperValu	1.089	Medium Chains	-0.096	Small Chains	-0.095	-0.385
Trader Joes	1.305	Safeway	-0.151	Kroger	-0.116	-0.445
Weis Markets	1.203	Giant Food	-0.286	Small Chains	-0.144	-0.362
Whole Foods	1.323	Safeway	-0.119	Kroger	-0.102	-0.473
Wild Oats	1.286	Kroger	-0.180	Safeway	-0.102	-0.363
Winn-Dixie	1.119	Publix	-0.298	Wal Mart	-0.180	-0.300
Meijer	1.018	Kroger	-0.167	Wal Mart	-0.157	-0.299
Target	1.236	Wal Mart	-0.333	Sam's Club	-0.079	-0.344
Wal Mart	0.760	Kroger	-0.069	Sam's Club	-0.064	-0.270
BJs	1.156	Sam's Club	-0.125	Costco	-0.085	-0.380
Costco	0.920	Sam's Club	-0.096	Safeway	-0.057	-0.387
Sam's Club	0.958	Wal Mart	-0.121	Costco	-0.085	-0.315

Diversion Ratios

- To unpack Table 7, we compute diversion ratios (Shapiro, 1996).
- Usually, the diversion ratio from j to k is

$$D_{jk} = -\frac{\partial q_k}{\partial p_j} / \frac{\partial q_j}{\partial p_j}$$

  which measures the fraction of lost sales, in response to a price increase at j, that are captured by k.
  - Here, instead of price, we use "quality" (i.e. the FEs).
- In Table 7, ratio of column 4 to column 2 gives share of increased sales for column 1 firm that are drawn from its largest rival.
- Diversion to the outside good is the ratio of column 7 to 2.

Diversion Ratios

Firm	Rival 1	DR1	Rival 2	DR2	DRO
Albertsons	Wal Mart	-0.11	Safeway	-0.09	-0.28
Aldi	Medium Chains	-0.13	Small Chains	-0.11	-0.24
Bashas Markets	Kroger	-0.23	Safeway	-0.14	-0.25
Delhaize America (Food Lion)	Wal Mart	-0.14	Medium Chains	-0.08	-0.30
Fred Meyer	Safeway	-0.18	SuperValu	-0.13	-0.31
Giant Food	Safeway	-0.11	Small Chains	-0.08	-0.41
HE Butt	Wal Mart	-0.23	Sam's Club	-0.09	-0.37
Hannaford Bros	Medium Chains	-0.19	SuperValu	-0.15	-0.36
Hy Vee Food Stores	Medium Chains	-0.20	Wal Mart	-0.18	-0.30
Kroger	Wal Mart	-0.12	Medium Chains	-0.08	-0.32
Bi-Lo	Wal Mart	-0.20	Food Lion	-0.09	-0.26
Publix	Wal Mart	-0.15	Winn-Dixie	-0.10	-0.34
Raleys	Safeway	-0.16	Small Chains	-0.08	-0.36
Harris Teeter	Food Lion	-0.17	Medium Chains	-0.10	-0.31
Safeway	Kroger	-0.09	SuperValu	-0.08	-0.37
Save A Lot	Small Chains	-0.11	Medium Chains	-0.10	-0.24
Stater Bros	Kroger	-0.15	SuperValu	-0.12	-0.34
Stop Shop	Medium Chains	-0.16	SuperValu	-0.13	-0.39
Trader Joes	Safeway	-0.12	Kroger	-0.09	-0.34
Whole Foods	Safeway	-0.09	Kroger	-0.08	-0.36
Wild Oats	Kroger	-0.14	Safeway	-0.08	-0.28
Winn-Dixie	Publix	-0.27	Wal Mart	-0.16	-0.27
Target	Wal Mart	-0.27	Sam's Club	-0.06	-0.28
Wal Mart	Kroger	-0.09	Sam's Club	-0.08	-0.36
BJs	Sam's Club	-0.11	Costco	-0.07	-0.33
Costco	Sam's Club	-0.10	Safeway	-0.06	-0.42
Sam's Club	Wal Mart	-0.13	Costco	-0.09	-0.33

Key Insights from Diversion Ratios

- Firms that are relatively isolated from competition:
  - Wal-Mart, Clubs, Safeway, Whole Foods.
- Firms that face the most competition:
  - Target, Winn-Dixie, Southern Chains.
- Firms that draw most from outside good:
  - Costco, Northeast chains.
- Firms that draw least from outside good:
  - Aldi, Save A Lot, Southern chains.
- Clubs belong in the choice set:
  - Clubs draw 20% from other clubs, 50% from other formats.

Merger Screening

- Merger analysis is one of the largest and most difficult areas of antitrust enforcement (Hosken and Tenn, 2016).
- Defining markets is especially controversial, since it can effectively determine the outcome *ex ante*.
  - Whole Foods/Wild Oats as PNOS, Office Depot/Staples as OSS
- To show how our model can be used to quickly "pre-screen" horizontal mergers, we consider two examples:
  1. The 2007 Whole Foods/Wild Oats merger, which the FTC contested.
  2. The 2016 Ahold/Delhaize merger, which was recently approved.
- Our model can reveal the true overlap between stores or firms, without taking a strong *ex ante* stance on market definition.
- Can also identify which consumers are most impacted and what stores should be divested (usual remedy) and to whom.

Merger Screening

- In particular, *for each census tract*, the model recovers the total revenue flowing from that tract to each store in its vicinity.
- We then construct tract-level HHIs to measure market concentration,

$$HHI_t = \sum_{f \in C_{t \setminus 0}} \left(100 \cdot \frac{p_{ft}}{1 - p_{0t}} \right)^2 .$$

  where $p_{ft} = \sum_{s \in F_f \cap C_t} p_{st}$ is chain f's total share from tract t.
- According to the 2010 Merger Guidelines, a market is considered
  1. highly concentrated if the HHI is over 2,500,
  2. moderately concentrated if the HHI is between 1,500 and 2,500, and
  3. un-concentrated (competitive) if the HHI is under 1,500.
- Focusing first on the industry as a whole, we compute these HHI's for every tract in all 317 MSAs.

Market Structure (Pre-Merger, 2006)

Table 8: **Firm concentration computed at the level of the tract**

Concentration	Number of Tracts	Income	Density	Mean Number of within 5/10 miles			
				All Stores	Large Chain Stores	Large Chains	Club Stores
Low (< 1500)	9,196	26.76	6212.01	43.42	20.27	5.35	1.20
				134.22	65.77	7.09	4.20
Moderate	22,749	30.85	3017.46	21.79	13.43	4.44	0.95
				64.28	39.85	6.11	2.83
High (> 2500)	21,423	25.65	1261.35	8.52	5.18	2.39	0.39
				22.41	13.69	3.53	0.99
Total	53,368	28.05	2862.98	20.19	11.30	3.77	0.77
				59.52	33.82	5.24	2.33

- Overall industry is quite concentrated (locally).

Merger Screening

- We then look at how this structure would change under each merger.
- To do so, we examine how HHI changes at each tract in which both firms appear in choice set.
  - Mergers that raise HHI by > 100 points "often warrant scrutiny," while
  - Mergers that raise the HHI by > 200 points (and result in highly concentrated markets) "likely enhance market power."
- We use these criteria to identify merger "hot spots," where mergers either warrant scrutiny or enhance market power.
  - *Caveat emptor*: We are *not* solving for new equilibrium prices (or new entries, or exits, or re-positionings, ...).
- We also compute "store-level" HHIs that aggregate tracts in a store's catchment area, weighting each tract-level HHI by the tract's contribution to total store revenue.
- We then compare to a screen based on diversion ratios.

Impact of Whole Foods/Wild Oats Merger

Table 9: **Tract-level Impact of the Whole Foods/Wild Oats merger**

	Both Firms Present		Warrants Scrutiny		Enhance Market Power	
State	Number of Tracts	Population	Number of Tracts	Population	Number of Tracts	Population
AZ	411	1676.24	0	0	0	0
CA	1427	6353.02	0	0	0	0
CO	641	2643.28	12	54.97	0	0
CT	142	538.18	0	0	0	0
FL	245	1041.83	0	0	0	0
IA	7	22.56	0	0	0	0
IL	708	2908.50	0	0	0	0
IN	18	66.97	0	0	0	0
KS	126	493.86	0	0	0	0
KY	142	545.66	0	0	0	0
MA	451	1940.66	0	0	0	0
MO	301	1094.88	0	0	0	0
NE	178	609.34	0	0	0	0
NM	164	642.11	16	41.81	0	0
NV	373	1494.98	0	0	0	0
OH	138	562.23	0	0	0	0
OR	229	1042.37	0	0	0	0
TX	428	1958.68	0	0	0	0
WA	28	103.57	0	0	0	0
Total	6157	25738.92	28	96.78	0	0

Impact of Ahold/Delhaize Merger

Giant + Stop & Shop and Food Lion + Hannaford

Table 10: **Tract-level Impact of the Ahold/Delhaize merger**

	Both Firms Present		Warrants Scrutiny		Enhance Market Power	
State	Number of Tracts	Population	Number of Tracts	Population	Number of Tracts	Population
DC	58	194.13	0	0	0	0
DE	45	238.05	1	6.46	7	46.00
MA	974	4547.29	349	1729.96	131	684.34
MD	1214	4999.43	389	1785.68	150	672.94
NH	124	587.62	49	245.98	58	256.56
PA	76	361.57	9	47.43	17	91.67
RI	19	69.11	4	15.35	15	53.76
VA	577	2550.94	297	1365.45	111	514.96
WV	31	163.93	0	0	31	163.93
Total	3118	13712.08	1098	5196.30	520	2484.16

Store Level Analysis

Table 12: Store-Level Analysis of Potential Mergers

Chain	# of Competing Stores	Average # of Competitors	Diversion Ratios Div>.05	Div>0.1	Div>0.2	Concentration Warrants Scrutiny	Presumed Likely
Ahold	328	10.85	64	29	8	138	52
Delhaize	161	22.11	141	122	75	63	74
Whole Foods	69	2.92	1	0	0	2	0
Wild Oats	80	2.52	6	1	0	4	0

Notes: Each row contains information on the stores of a particular chain for whom the merger is relevant. # of Competing Stores is number of stores in the chain that compete in a tract where at least one store of the merging partner is present. Average # of Competitors is number of merger partner stores in the choice set of tracts that belong to the competing stores catchment area, L_s. "Warrant Scrutiny" and "Presumed Likely" indicate number of chain stores that would be classified as such according to the 2010 Merger Guidelines where HHI is calculated at the store level using (10).

Comparison

Table 13: Comparison of Store Level Merger Evaluation and Diversion Ratios

	Div<.05	.05<Div<0.1	.1<Div<0.2	.2<Div
Ahold				
No Concern	137	1	0	0
Warrants Scrutiny	112	18	7	1
Raise Concerns	15	16	14	7
Delhaize				
No Concern	11	2	10	1
Warrants Scrutiny	5	7	19	32
Raise Concerns	4	10	18	42

Impact of Including Club Stores on Analysis of A/D Merger

Table 14: **Effect of Excluding Club Stores on Evaluating the Ahold/Delhaize Merger**

	Without Club Stores			
With Club Stores	No Concern	Warrants Scrutiny	Presumed Likely	Total
No Concern	1,144	356	0	1,500
Warrants Scrutiny	1	426	671	1,098
Presumed Likely	0	2	518	520
Total	1,145	784	1,189	3,118

Conclusions

- We provide a simple framework for analyzing competition between multi-product retailers.
- The estimates from this model reveal how firms position themselves with respect to the income and travel costs of their customers.
- We use the model to evaluate two mergers, highlighting the importance of both careful market definition and including all relevant competitors.
- Future work will address how firms respond (re-optimize) to changes in market structure.

Identification

Overall approach

- Exploit geographic variation in revenues, locations & demographics.
  - Assume (ϵ_{its}, η_s) independent of store location & size, as well as consumers' locations & incomes.
    - Consumers take store locations as given
    - Perceptions of store pricing, quality & assortment formed at chain (not store) level.
  - Control for endogeneity of overall policies using chain fixed effects.
    - Reasonable if prices and assortments mostly set at chain level.
    - Evidence from IRI and Nielsen data suggests they are.

Identification

Key parameters

- α identified by varying total number of stores across 'identical' markets and seeing change in total revenue across all stores.
- Given α, utility parameters identified by varying characteristics of stores and consumers, then observing resulting changes in share of total expenditure (within catchment area L_s) captured by each store.
  - Varying distance between a tract and store changes share of expenditures at that store relative to others in the tract's choice set.
  - Change will be reflected in store's revenue relative to others in same choice set, all of which are observed.
- Nesting parameters identified through variation in number and location of stores within versus across nests.

back to estimation

Data Summary: Chain Characteristics

Table 2: Chain Characteristics by Type

	Mean	St. Dev.	1st Quartile	Median	3rd Quartile
Medium Grocery Chains					
13.91 % of all MSA stores, 9.92 % of MSA Revenue					
Number of Stores	24.50	20.03	12	17	28
Number of MSA operating	4.83	5.68	1	3	6
Large Grocery Chains					
49.87 % of all MSA stores, , 47.17 % of MSA Revenue					
Number of Stores	400.93	451.08	125	189.50	510
Number of MSA operating	34.70	36.41	12	17	46
Supercenters					
7.06 % of all MSA stores, 17.88 % of MSA Revenue					
Number of Stores	568	707.54	159	160	1,385
Number of MSA operating	107	121.74	26	48	247
Club Stores					
4.03 % of all MSA stores, 16.76 % of MSA Revenue					
Number of Stores	324.33	209.32	122	311	540
Number of MSA operating	113.67	97.44	36	82	223

Data Summary: Large Chains

Table 3: Characteristics of Large Chains

	# Stores	# MSAs	Stores/MSA	Rev.	Rev. /sqft	Size
Large Grocery Chains						
Albertsons	510	71	7.18	357.94	6.75	54.16
Aldi	615	108	5.69	77.05	6.15	12.84
Bashas Markets	134	6	22.33	257.72	8.90	32.02
Delhaize America (Food Lion)	949	55	17.25	178.73	6.28	28.69
Fred Meyer	101	12	8.42	740.10	13.42	55.23
Giant Eagle	140	11	12.73	579.29	12.82	46.70
Giant Food	292	14	20.86	568.60	15.42	37.96
Great A & P Tea Co.	161	11	14.64	341.02	9.98	34.97
HE Butt	227	16	14.19	813.44	16.40	51.01
Hannaford Bros	108	9	12	528.47	12.61	42.05
Hy Vee Food Stores	102	15	6.80	513.48	11.59	45.82
Ingles Markets	112	11	10.18	205.27	5.02	41.59
Kroger	1,973	107	18.44	463.42	10.95	42.40
Lone Star Funds (Bi-Lo)	238	21	11.33	225.29	6.03	37.38
Publix	845	36	23.47	419.70	11.07	38.81
Raleys	127	12	10.58	428.15	9.91	43.60
Roundys	125	10	12.50	496.60	12.03	41.91
Ruddick Corp (Harris Teeter)	138	17	8.12	407.79	11.25	36.56
Safeway	1,339	46	29.11	424.96	11.98	37.33
Save A Lot	715	163	4.39	114.98	8.49	14.49
Save Mart	118	13	9.08	385.81	10.18	37.84
Smart & Final	217	29	7.48	147.03	10.14	15.18
Stater Bros	162	3	54	388.27	16.10	24.22
Stop & Shop	312	17	18.35	563.78	12.18	47.31
SuperValu	1,194	58	20.59	460.74	9.51	49.02
Trader Joes	236	37	6.38	302.22	32.66	9.45
Weis Markets	120	12	10	242.58	6.62	37.22
Whole Foods	159	47	3.38	511.79	21.12	26.99
Wild Oats	108	38	2.84	185.28	9.29	20.71
Winn-Dixie	451	36	12.53	250.78	5.54	46.27
Supercenters						
Meijer	159	26	6.12	826.10	14.11	59.56
Target	160	48	3.33	526.25	8.79	60.66
Wal Mart	1,385	247	5.61	1,064.24	16.18	65.12
Club Stores						
BJs	122	35	3.49	797.95	7.59	104.47
Costco	311	82	3.79	2,259.49	18.17	123.50
Sam's Club	540	223	2.42	1,451.67	11.17	130.05
Total	720.18	115.27	8.63	688.70	11.71	56.61

back to main table