The Diffusion of Wal-Mart and Economies of Density

- Density Economies

- How important are they for Wal-Mart?

Idea

- If economies of density don't matter store locations will be scattered across the country
  - Retail locations vary in quality—best ones won't all be right next to each other.
  - Plus bad to put stores next to each as since they will cannibalize each others' sales.
- Revealed preference approach: If always pick stores next to other stores, infer density economies must matter.

Strategy

- Estimate quality of retail sites ignoring issue of density economies.

- Then look at choice behavior and back out density economies as a residual.

Develop model of store-level operating profit

- Includes a model of store-level sales
  - Demographic data at rich geographic detail
  - Takes into account competition with other stores through population density.
  - Takes into account cannibalization of sales from other Wal-Marts

Model

- Store locations j on the plain, B is set open at a given time.
- $Rev_j(B)$ is revenue of an open store j given B
- Gross margin μ
- Operating costs $C_j(Rev_j)$
- Operating profit $\pi_j(B) = \mu Rev_j - C_j$

Density Economies

- Spillover to store j

$$s_j = 1 - \frac{1}{\sum_{k \in B} \exp(-\alpha y_{jk})}$$

  where y_{kj} distance from j to k and $\alpha = .02$.

- Takes values on range $s_j = [0, 1)$.

- Density benefit is additive ϕs_j.

Examples

- One store. Distance from itself is $y_{11} = 0$.

$$s_1 = 1 - \frac{1}{\sum_k \exp(-\alpha y_{jk})} = 1 - \frac{1}{1} = 0$$

- Two stores, $y_{12} = 55$

$$s_1 = 1 - \frac{1}{e^0 + e^{-.02*55}} = 1 - \frac{1}{1 + 1.33} = .25$$

Other Costs

- Urbanization costs $c^{urban}(m_j)$, where m_j population density at location j.

Wal-Mart's Problem

- No exits, so B_t set of stores open weakly increases over time.
- Fix N_t the number of stores open at time t.
- Let r denote a particular "rollout" of stores
- Discount rate ρ (continuous time)
- Wal-Mart picks r to maximize

$$v(r) = \int_0^\infty e^{-\rho t} g_t \sum_{j \in B_t} \left[\pi_j(B_t) + \phi s_j(B_t) - c_j^{urban} \right] dt$$

  Subject to having N_t stores at time t.

A complicated, nonconvex problem. Dynamic (because store opening is permanent).

Particulars of Demand:

- Consumers distributed across discrete locations (blockgroups)
- $y_{j\ell}$ distance between store j and location ℓ.
- Consumer k at location ℓ has characteristics $z_{\ell,t}$ and total spending λ_t.

- Discrete choice *nested-logit* model
  - outside good: composite of retail alternatives
  - inside goods: all Wal-Marts within 25 miles of the consumer's home (in a nest)

- Specification of utilities for consumer k at ℓ

$$u_{k\ell 0} = o(m_\ell) + z_\ell \omega + \zeta_{k\ell 0} + (1-\sigma)\varepsilon_{k\ell 0}.$$

$$u_{k\ell j} = -\tau\left(m_\ell\right) y_{\ell j} + x_j \gamma + \zeta_{k1} + \left(1-\sigma\right)\varepsilon_{k\ell j}.$$

m_ℓ population density (population within 5 mile radius).

x_j store characteristics

- $$\begin{aligned} o(m) &= \omega_0 + \omega_1 \ln(m) + \omega_2 \left(\ln(m)\right)^2 \\ \tau(m) &= \tau_0 + \tau_1 \ln(m) \end{aligned}$$

Can use McFadden's formulas to derive:

- $p_\ell^W(m_\ell, z_\ell, y_\ell, x, \theta)$ share of consumers at ℓ who buy at some Wal-Mart

- $p_\ell^{j|W}(m_\ell, z_\ell, y_\ell, x, \theta)$ share of consumers buying at j conditioned upon buying at some Wal-Mart

- Share of spending at store j

$$p_\ell^j = p_\ell^{j|W} \times p_\ell^W.$$

- Total revenue of store j is

$$R_j(\theta) = \sum_{\{\ell | j \in B_\ell\}} \lambda \times p_\ell^j \times n_\ell.$$

- Observed revenue $\tilde{R}_t$, so measurement error

$$\varepsilon_j^{measure} = \ln(\tilde{R}_j) - \ln(R_j(\theta)).$$

where $\varepsilon_j^{measure}$ is normally distributed

- Estimation: restrict attention to regular stores
  - Treat supercenters as an option for consumers
  - View supercenters as a combination of regular stores and grocery stores
  - So maximize likelihood of the sales figures for regular stores
  -

Other components of Operating Cost

- Labor requirements function for each store
- Take into account wages vary by city size (use County Business Patterns data)

Adjustment for other years

- Estimate demand for 2005
- Assume proportional growth of all revenues and costs
- Choose g_t to fit aggregate sales figures for earlier years..

Data Element 1: Store-Level Data for 2003
Source: TradeDimensions (ACNeilsen)

Store Type	N	Mean Sales ($Millions/Year	Employment	Bldg Size (1,000 sq ft.)
All	2,936	59.6	223.4	143.1
Regular	1,457	42.4	112.2	98.6
SuperCenter	1,479	76.5	332.8	186.9

Data Element 2: Store opening dates (from Wal-Mart). When relocated down the street, date is opening of original store (store keeps same number).

Data Element 3: Demographic Information by Block Group
Source: Census 1980, 1990, 2000

	1980	1990	2000
N	269,738	222,764	206,960
Mean population (1,000)	0.83	1.11	1.35
Mean Density (1,000 in 5 mile radius)	165.3	198.44	219.48
Mean Per Capita Income (Thousands of 2000 dollars)	14.73	18.56	21.27
Share old (65 and up)	0.12	0.14	0.13
Share yound (21 and below)	0.35	0.31	0.31
Share Black	0.1	0.13	0.13

First State Estimates
Step 1: Demand

Parameter	Definition	No Age	Limited Age	Detailed Age
λ	scaling parameter	29.742	29.057	18.702
		(.055)	(.057)	(.057)
ρ	Correlation parameter	.781	.767	.959
		(.055)	(.057)	(.057)
τ_0	Constant	.616	.621	.464
		(.054)	(.056)	(.031)
τ_1	population density within 5 miles	-.046	-.049	-.001
		(.047)	(.048)	(.016)
ω	Constant	-7.769	-7.586	-10.517
		(.055)	(.057)	(.057)
	lnmaxc(neig5)	1.503	1.605	2.596
		(.054)	(.056)	(.058)
	lnmaxc(neig5)^2	-.027	-.037	-.140
		(.043)	(.045)	(.010)
	Pcitrun	.023	.021	.018
		(.045)	(.046)	(.004)
	Blackshr	.928	.909	.841
		(.055)	(.057)	(.057)
	Youngshr	1.241	.881	.633
		(.055)	(.057)	(.057)
	Oldshr	1.369	1.158	1.288
		(.055)	(.057)	(.057)

Continued

γ	store age 3-dummy		.246 (.057)	
	store age 3-5 dummy			.240 (.062)
	store age 6-10 dummy			.319 (.060)
	store age 11-20 dummy			.340 (.057)
	store age 20-dummy			.225 (.057)
σ^2	measurement error	.092 (.055)	.090 (.057)	.090 (.003)
N		1457	1457	1457
SSE		134.746	131.039	130.554
R^2		.674	.683	.684

A Look at Demand
Distance and Density
(Assume only one Wal-Mart within 25 miles)

Distance	Population Density (1,000 per 5 mile radius)						
(miles)	1	5	10	20	40	100	250
0	.999	.984	.957	.893	.766	.499	.244
1	.997	.973	.930	.839	.678	.402	.185
2	.995	.954	.890	.765	.576	.312	.138
3	.991	.923	.829	.669	.467	.234	.102
4	.984	.875	.745	.558	.361	.171	.074
5	.971	.803	.637	.440	.267	.122	.053
10	.596	.213	.122	.069	.039	.019	.010
15	.062	.018	.011	.007	.004	.003	.002
20	.003	.001	.001	.001	.000	.000	.000
25	.000	.000	.000	.000	.000	.000	.000

A Look at Demand
Consistency with Wal-Mart's Reports about Cannibalization

	Cannibalization Percent	
Fiscal Year	Wal-Mart's Report	My Estimate
1999		0.69
2000		0.95
2001		0.61
2002	1.00	0.73
2003	1.00	1.41
2004	1.00	1.48
2005	1.00	1.55
2006	1.00	1.35

Stage 2

- Remaining parameters
  - ϕ, weight on spillover ($\alpha = .02$ is fixed)
  - Urban cost parameters
- Consider perturbation approaches

Estimated Diminishing Returns
General Merchandise

Years in State	Marginal Operating Profit		Mean marginal spillover	Distance to Closest DC (one year after opening)	N
	Mean	Median			
0-2	3.88	4.01	0.79	352.4	340
2-5	3.99	4.11	0.95	185.6	474
5-10	3.79	3.85	0.98	127.0	569
10-15	3.35	3.33	1.00	98.6	325
15-20	2.90	2.88	1.00	81.8	195
20-21	2.43	2.41	1.00	71.8	79

Estimated Diminishing Returns
Supercenters

Years in State	Marginal Operating Profit		Mean marginal spillover	Distance to Closest DC (one year after opening)	N
	Mean	Median			
0-2	4.08	4.13	0.71	270.0	233
2-5	4.06	4.18	0.92	148.9	485
5-10	3.83	3.85	0.99	103.8	769
10-15	3.26	3.26	1.00	82.3	432
15-20	2.73	2.66	1.00	66.5	67

Approach 1 (bad not doing anymore!): Deviate to Maximize Operating Profit

- Hold fixed number of stores, resequence (so don't need f_t)
- Change only cities with population density less than 20 so don't need to know parameters of urbanization cost
- Assume $\varepsilon_j = 0$, so don't need that either.
- For a given deviation
  - $\Delta\pi$:difference in PV of operating profit
  - Δs :difference in PV of spillovers

- Optimality of r^* implies

$$\Delta\pi + \phi\Delta s \geq 0$$

- Concern: measurement error $\tilde{\pi}_j = \pi_j + \varepsilon_j^{measure}$, we see $\tilde{\pi}_j$, firm acts on π_j.

Approach 1: Resequence to Maximize Operating Profit
(Present Value in Millions of 2003 Dollars)

Interval 1: 1971-1980

	Revenue	Operating Profit	Spillovers
Actual Policy	14,965	1,413	0
Deviation	15,950	1,519	15
Gain from Actual	−985	−106	15

Interval 2: 1982-1989

	Revenue	Operating Profit	Spillovers
Actual Policy	133,577	12,673	0
Deviation	136,665	13,004	15
Gain from Actual	−3,088	−331	15

Approach 2: Pairwise Resequencing

- Let k denote a pairwise resequencing. We have

$$\Delta\pi^k + \phi\Delta s^k \geq 0$$

  See

$$\Delta\tilde{\pi}^k + \phi\Delta s^k$$

  not necessarily positive because of classical measurement error.

- Follow ideas of Pakes, Porter, Ho, Ishii. Take averages to create moment inequalities with weights that can depend upon choices.

Approach 2: Pairwise Resequencings
1970-1980
Stores in Small or Medium Cities

	Number of resequencings	Mean $\Delta\pi$	Mean Δs	Implied Bound on φ
All Stores in Small or Medium Cities	22,433	-.057	.188	0.30 (lower)
Older Store Closer to Bentonville	16,745	-.197	.258	0.76 (lower)
And Older store in Smaller Town	8,170	-.992	.266	3.73 (lower)
Spillovers Lower in Original	5,293	.390	-.045	8.67 (upper)