The Diffusion of Wal-Mart and Economies of Density

by

Thomas J. Holmes[1]

University of Minnesota

Federal Reserve Bank of Minneapolis

and National Bureau of Economic Research

November, 2005

Preliminary and Incomplete Draft for Workshop Presentation

[1] The views expressed herein are solely those of the author and do not represent the views of the Federal Reserve Banks of Minneapolis or the Federal Reserve System. The research presented here was funded by NSF grant SES 0136842. I thank Junichi Suzuki for excellent research assistance for this project. I thank Emek Basker for sharing data with me and I thank Ernest Berkas for help with the data.

1. Introduction

A retailer can often achieve cost savings by locating its stores close together. A dense networks of nearby stores facilities the logistics of deliveries and facilitates the sharing of infrastructure such as distribution centers. When stores are close together they are easier to manage and it is easier to reshuffle employees between stores. Stores located near each other can potentially save money on advertising. All such cost savings are *economies of density.*[2]

Understanding these benefits is of interest because they matter for determining policies towards mergers. To the extent that merger of nearby facilities into one company confers cost savings, these benefits potentially offset concerns about increased market power. Study of these benefits is also of interest for understanding firm behavior. To the extent these benefits matter, firms may have an incentive to preemptively build a large network of stores to grab a first-mover advantage. Finally, in recent years there has been a general interest in the industrial organization literature in network benefits of all kinds.[3] The network benefits of a dense network of stores is a potentially important efficiency and there has been little work on this topic.

Wal-Mart is the world's largest corporation in terms of sales. It is regarded as a company that excels in logistics. The goal of this paper is assess the importance of economies of density to Wal-Mart. Preliminary results suggest the benefits are significant.

Wal-Mart is notorious for being secretive about internal data—I am not going to get access to confidential data on its logistics costs, managerial costs, advertising, or any of the other cost components that depend upon economies of density. Instead, I draw inferences about the cost structure that Wal-Mart faces by examining its revealed preferences in its site-selection decisions. I study the time path of Wal-Mart's store openings, the diffusion of Wal-Mart. The idea underlying my approach is that alternative sites vary in quality. If economies of density were not important, Wal-Mart would go to the highest quality sites first and work its way down over time. The highest quality sites wouldn't necessarily be bunched together, so initial Wal-Mart stores would be scattered in different places. But

[2] There is a larger literautre on economies of density in electricity markets (e.g. Roberts (1986)) and transportation markets (e.g....)

[3] Gowrisankaran and Stavins, etc.

when economies of density matter, Wal-Mart might chose lower quality sites that are closer to its existing network, keeping the stores bunched together, putting off the higher quality sites until later when it can expand out to them.

The latter is what happened. Wal-Mart started with its first store near Bentonville, Arkansas, in 1962. The diffusion of store openings radiating out from this point was very gradual. It is very helpful to view a movie of the entire year-by-year diffusion posted on the web. Figure 1 shows the process over the years 1970-1980. Wal-Mart did not first grab the "low hanging fruit" in the most desirable location throughout the county and then come back for the "high hanging fruit," with fill-in stores. Desirable locations far from Bentonville had to wait to get their Wal-Marts.

I bring to the analysis a number of pieces of information about Wal-Mart's problem. I use store-level data from ACNeilsen and demographic data from the Census to estimate a model of demand for Wal-Mart at a rich level of geographic detail. I use this to estimate Wal-Mart's sales from alternative location configurations. I also incorporate information about other aspects of costs that can be measured, store-level labor costs, land costs, etc. The underlying principle I use here is to plug into the model the things that I can estimate, and back out the economies of density as a residual. Of course this leaves open the possibility that I have other things out.

Given the enormous number of different possible combinations of stores that can be opened, it is difficult to solve Wal-Mart's optimization problem. This makes conventional approaches used in the industrial organization literature infeasible. Instead, I follow a perturbation approach. I consider a set of selected deviations from what Wal-Mart actually did and determine the set of parameters consistent with this decision. The deviations take the form of a resequencing the date of store openings.

The paper contributes to the literature on entry and store location in retail. Related contributions include Breshahan and Reiss (1991), Toivanen and Waterson (2005), Andrews et al (2004).

In addition to contributing to the literature on economies of density, the paper also contributes to a new and growing literature about Wal-Mart itself (e..g., Basker (forthcoming), Stone (1995), Hausman and Leibtag (2005), Ghemawat, Mark, and Bradley (2004)). Wal-

Mart has had a huge impact on the economy. It has been argued that this one company contributed a non-negligible portion of the aggregate productivity grow in recent years. Wal-Mart is responsible for major changes in the structure of industry, of production, and in of labor markets. One good question is: what exactly is a Wal-Mart, why is it different from a K-Mart or a Sears? One thing that distinguishes Wal-Mart is its emphasis on logistics and distribution. (See, for example, Holmes (2001)). It is plausible that Wal-Mart's recognition of economies of density and its knowledge of how to exploit these economies distinguished it from K-Mart and Sears and is part of the secret of Wal-Mart's success.

2. Model

Consider a model of a retailer that I will call "Wal-Mart." At a particular point in time, Wal-Mart has as set of stores and consumers make buying decisions based on the location of the stores. I first describe consumer demand holding the set of Wal-Mart store locations as fixed. Next I describe the cost structure and the process through which Wal-Mart opens new stores.

2.1 Demand

We expect that consumers will tend to shop at the closest Wal-Mart to their home. Nonetheless, in some cases, a consumer might prefer a further Wal-Mart. For example, for a particular consumer, a further Wal-Mart might be more convenient for stopping on the way home from work. Since a consumer at a given location might potentially shop at several different Wal-Marts, we need a model of product differentiation across different Wal-Marts. To this end, I follow the common practice in the literature of taking a discrete choice approach to product differentiation. I specify a *nested logit* model and put the various Wal-Marts in a consumer's vicinity in one nest and put the *outside good* in a second nest.

Now for some notation. Consumers are located across L discrete locations indexed by ℓ. Suppose at a point in time Wal-Mart has J stores indexed by j, with each store in a unique location. For a given location ℓ, let $y_{\ell j}$ denote the distance in miles between location ℓ and store j . Let n_ℓ denote the *population* of location ℓ and let m_ℓ be the *population density* at ℓ.

Consider a particular consumer k at a particular location ℓ. Let B_ℓ denote the set of Wal-Marts in the *vicinity* of the consumer's home. (In the empirical work, this will be defined as the set of Wal-Mart's within 25 miles of the consumer's home.). The consumer has a dollar amount of spending λ that he or she allocates between the following discrete choices: the outside good (good 0) or one of the nearby Wal-Marts in B_ℓ (if B_ℓ is non-empty). The utility of the outside good 0 is

$$u_{k\ell 0} = f(m_\ell) + z_\ell \omega + \zeta_{k0} + (1-\sigma)\varepsilon_{k0}. \tag{1}$$

The first term is a function $f(\cdot)$ that depends upon the population density m_ℓ at consumer i's location. Assume $f'(m) > 0$; i.e., the outside option is better with more people around. This is a sensible assumption as we would expect there to be more substitutes for Wal-Mart in larger markets for the usual reasons. A richer model of demand would explicitly specify the alternative shopping options available to the consumer. In my empirical analysis this isn't feasible for me since I don't have detailed data on all various shopping options besides Wal-Mart a particular consumer might have. Instead I specify the reduced form relationship between $f(m_\ell)$ and population density.

The second term allows demand for the outside good to depend upon a vector of the average characteristics z_ℓ (average demographic characteristics and income) of consumers at location ℓ times a parameter vector ω. The final two terms, ζ_{k0} and ε_{k0}, are random taste parameters for the outside good that are specific to consumer k. The distributions for these draws are explained momentarily.

The utility of a given Wal-Mart store $j \in B_\ell$ is

$$u_{k\ell j} = -\tau(m_\ell)\, y_{\ell j} - x_j \gamma + \zeta_1 + (1-\sigma)\, \varepsilon_{kj}.$$

The first term is the utility decrease from travelling to the Wal-Mart j that is a distance $y_{\ell j}$ from the consumer's home. The weight $\tau(m_\ell)$ the consumer places on distance depends upon population density. This is another reduced form relationship; because of differences in the availability of substitutes induced by differences in population density, consumers in areas with high population density may respond differently with distance than consumers in low density areas. The second terms allows utility to depend upon other characteristics x_j

of Wal-Mart store j. In the empirical analysis, the store-specific characteristic that I will focus on is store age. In this way, it will be possible in the demand model for a new store to have less sales, everything else the same. This captures in a crude way that it takes a while for new store to ramp up sales. The final two terms are random utility components specific to store j.

As discussed in Wooldrige (2002), McFadden(1984) showed that under certain assumptions about the distribution of $(\zeta_{k0}, \zeta_{k1}, \varepsilon_{k0}, \varepsilon_{k1}, ...\varepsilon_{kJ})$ that I impose here, the probability a consumer at ℓ purchases from some Wal-Mart is

$$p_\ell^W = \frac{\left[\sum_{j \in B_\ell} \exp\left(\left(1-\sigma\right)\delta_{\ell j}\right)\right]^{\frac{1}{1-\sigma}}}{\left[\exp\left(\delta_{\ell 0}\right)\right] + \left[\sum_{j \in B_\ell} \exp\left(\left(1-\sigma\right)\delta_{\ell j}\right)\right]^{\frac{1}{1-\sigma}}} \tag{2}$$

for

$$\begin{aligned} \delta_{\ell 0} &\equiv f(m_\ell) + z_\ell \omega \\ \delta_{\ell j} &\equiv -\tau\left(m_\ell\right) y_{\ell j} - x_j \gamma, \end{aligned}$$

and the probability of purchasing at a particular store $j \in B_\ell$, conditional on purchasing from some Wal-Mart is

$$p_\ell^{j|W} = \frac{\exp((1-\sigma)\,\delta_{\ell j})}{\sum_{k \in B_\ell} \exp((1-\sigma)\,\delta_{\ell k})}. \tag{3}$$

The probability a consumer at ℓ shops at Wal-Mart j is

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

Total revenue of store j is

$$R_j = \sum_{\{\ell | j \in B_\ell\}} \lambda \times p_\ell^j \times n_\ell. \tag{4}$$

This equals the spending λ of a consumer times the probability a consumer at ℓ shops at j times the population n_ℓ at ℓ, aggregated over all locations in the vicinity of store j.

2.2 Cost Structure and Openings of New Stores

This subsection describes the cost structure. It first specifies input requirements for merchandise, labor, land, miscellaneous inputs. It next specifies an urbanization cost. Finally, it specifies the form of the *density economies*, which will be the main target of the estimation.

2.2.1 CGS, Labor, Land, and Miscellaneous Costs

Suppose the gross margin is μ, so that μR equals sales minus cost of goods sold.

Assume that the labor requirements *Labor* of store in a period depend upon the sales R at the store in a log linear fashion,

$$Labor = \nu_{Labor} R^{\nu_0},$$

for parameters ν_{Land} and ν_0.

Suppose the wage for retail labor at location ℓ is W_ℓ so that the wage bill is $W_\ell L$. Assume that wage at a location depends on population density

$$W_\ell = g_{Labor}(m_\ell).$$

Assume for now that land and building requirements are proportional to sales,

$$\begin{aligned} Land &= \nu_{Land} R \\ Bldg &= \upsilon_{Bldg} R \end{aligned} \tag{5}$$

(In later work I plan to allow for scale economies and a richer structure). Let P_{Land} and P_{Bldg} be the rental prices. Assume the land prices depend upon population density,

$$P_{Land} = g_{Land}(m_\ell).$$

Assume that building prices are the same everywhere; i.e.P_{Bldg} is a constant. I discuss this further below.

Miscellaneous costs have two parts, a fixed cost and a marginal cost. Assume the fixed miscellaneous is constant across stores. This means I can ignore it in the analysis since it will be independent of where stores are located. The second part is proportionate to R and is denominated in dollars,

$$C_{Misc} = \nu_{Misc} R.$$

Importantly, the cost ν_{Misc} is assumed to be constant across locations.

2.2.2 Urbanization Costs

The Wal-Mart store has a distinct format, a big box one-floor store with huge parking lot on a convenient interstate exit. This approach has obvious limitations in a big city. To capture this in the model, assume an urban fixed cost $C_{Urban}(m)$ that depends upon the population density m of a location. If Wal-Mart were to locate in an highly urbanized area, they would have to do things, like make a multi-store structure, that is not necessary in a less urbanized area. For example, there are reports that best Buy Buy expects to pay $200 per square foot in construction costs to enter the Los Angeles market which is four times their normal building cost of $50 per square foot.

Assume there is range of m, $m \leq \overline{m}$, where $C_{Urban}(m) = 0$ and it is only above the threshold $\overline{m}$ where the urbanization cost is positive. The idea here is that Dubuque, Iowa, the eight largest city in Iowa with a population of 62,220, is relatively similar to the small towns in Iowa, in terms of the applicability of the Wal-Mart model while Dubuque is very different from New York City. In other words, $m_{Dubuque} < \overline{m}$.

2.2.3 The Density Economies

I now specify the main target of this inquiry, *density economies.* There is a store-level profit term that is increased with a higher density of stores. This component is intended to capture a broad set of factors, including management. Certainly a significant component is logistics and distribution cost. A delivery struck may cost the same to operate whether full or half full. If two stores are near each other, the stores can be replenished on the same delivery run. Also included here are savings in marketing cost (advertising) by locating stores near each other.

Rather than develop a micro-model of distribution economies and route structures or micro-model of economies of management, I follow the literature on productivity spillovers and take a reduced-form approach. I assume a parametric form whereby cost savings "spill over" from one store to another. These spillovers won't give rise to any externalities, of course, since central headquarters will be making location decisions that internalize these benefits. The functional form for the spillover collected by store j is

$$s_j = \sum_k \exp(-\alpha y_{jk}), \tag{6}$$

where y_{jk} is the distance in miles from store j to store k. If store j is right next to another store j so that y_{jk} is approximately zero, the spillover collected by j from k is approximately one spillover unit. As the distance to store k is increased, the spillover decays at a rate $\alpha > 0$. A store collects a spillover of exactly one from itself, so the minimum value that s_j can take is one.

The density economy for store j depends upon the level of spillover s_j it collects as follows:

$$D(s_j) = -\phi \frac{1}{s_j}, \tag{7}$$

for $\phi \geq 0$. The density economy is strictly increasing in s_j as it becomes smaller in absolute value and thus less negative. This form has an intuitive interpretation. Suppose there is a single store. Then $s_1 = 1$ and we can interpret ϕ as a fixed cost. Suppose next there are two stores located right on top of each other. Then for each store $s_1 = s_2 = 2$, and the density benefit is $-\frac{1}{2}\phi - \frac{1}{2}\phi = -\phi$, so the "fixed cost" ϕ is the same as if there just one store.

2.2.4 Store Openings

Everything that has been discussed so far considers quantities for a particular time period, i.e., revenues or fixed cost. I now explain the dynamic aspects of the model.

Let B_t be the set of Wal-Mart store locations at period t. This consists of the stores B_{t-1} operating in the previous period as well as a set of B_t^{new} new stores opened in the current period, so $B_t = B_{t-1} + B_t^{new}$. This is a good assumption for Wal-Mart. It rarely exits a location once it opens a store.

Let J_t be the number of stores operating at t, the cardinality of B_t. Let N_t be the number of stores opened at t, i.e., the cardinality of B_t^{new}. I take N_t as exogenous in my analysis. Wal-Mart in its first years added only one or two stores a year. The number of new store openings has grown substantially over time, now sometimes several stores in one week. I am not going to make any attempt to model the growth rate at which Wal-Mart added stores. Presumably capital market considerations played an important role here. Rather I will take as given that Wal-Mart gets to add a certain number of new stores in each period and the question of interest is where Wal-Mart puts them. Formally, the number of new openings $\{N_1, N_2, ...N_T\}$ in periods $t = 1$ through the terminal period $t = T$ is taken as given.

I take B_T as exogenous in my analysis; I impose a terminal condition that Wal-Mart has a given set of stores in the last period. Thus I focus on the timing of openings at a given set of store locations, rather than consider every parcel of land in the U.S. as a potential Wal-Mart store site. This leads to a considerable simplification of the analysis. Let $A(B_0, B_T, N_1, N_2, ..N_{T-1})$ be the set of all policies that start with initial condition B_0, end with terminal condition B_T and open N_t new stores in period t. This is a finite set, though potentially quite large. Let policies in A be indexed by i.

One final issue is productivity growth. I allow for exogenous productivity growth of Wal-Mart at a rate of ρ_t per period. What I mean by this is that if Wal-Mart where to hold fixed the set of stores, from period $t-1$ to period t, then revenue at each store and all components of level costs would grow at a constant amount ρ_t, i.e.

$$\begin{aligned} R_{j,t} &= (1+\rho_t) R_{j,t-1} \\ C_{j,t} &= (1+\rho_t)\, C_{j,t-1}. \end{aligned}$$

This means the profit grows at a rate ρ_t, holding fixed the set of Wal-Mart's stores and demographic characteristics. As will be discussed later, the growth of sales per store of Wal-Mart has been remarkable. Part of this growth is due the gradual expansion of its product line, from hardware and variety items to eye glasses and tires later, to groceries today, and perhaps banking tomorrow. Rather than model this expansion of product variety directly, I take the process as occurring exogenously.

Given a discount factor β, the Wal-Mart's problem is

$$\max_{i \in A(B_0, B_T, N_1, N_2, ..N_{T-1})} \sum_{t=1}^{T} \beta^{t-1} \sum_{j \in B_{it}} \left[R_{ijt} - C_{ijt} + D_{ijt} \right] + \eta_i. \tag{8}$$

for

$$\begin{aligned} C_{ijt} &= (1-\mu)\, R_{ijt} + W_{jt} Labor_{ijt} + P_{Land,jt} Land_{jt} + P_{Bldg,t} Bldg_{jt} \\ &\quad + C_{Misc,ijt} + C_{Urban,ijt} \end{aligned}$$

A notable addition is the term η_i. This is a random component of discounted profit associated with decision i. Various alternative assumptions of on this error are discussed below.

3. Data and Some Facts

This section begins by explaining the basic data sources. It then discusses some facts about Wal-Marts expansion process.

3.1 Data

There are five main data elements used in the analysis. The first element is store-level data on sales and other store characteristics that I have obtained from a commercial source. The second element is information about the timing of store openings that has been cobbled together from various sources. The third element is demographic information from the Census. The fourth is land price data for Wal-Mart stores obtained from tax records. The fifth element is data on how retail wages vary with population density from the Census.

Data element one, store-level data variables such as sales, was obtained from *TradeDimensions*, a unit of ACNeilsen. This data provides estimates of average weekly store level sales for all Wal-Marts open at Feb. 2004, as well as the following additional store characteristics: employment, square footage of the store building, store location exact geographic coordinates and whether or not the store is a supercenter. (Supercenters sell perishable groceries like meat and vegetables in addition to the products carried by regular stores.) This data is the best available and is the primary source of market share data used in the retail industry. Ellickson (2004) is a recent user of this data for the supermarket industry.

Table 1 presents summary statistics of the TradeDimensions data for the 2,936 Wal-Marts in existence in the contiguous part of the United States as the end of 2003.[4] [5] (Alaska and Hawaii are excluded in all of the analysis.) As of the end of 2003, slightly over half of Wal-Mart's stores are supercenters. The average Wal-Mart racks up annual sales of $60

[4] I will refer to the TradeDimensions data as from 2003, even though it is for Feb 2004. I will think of this as the beginning of 2004, so the data is for 2003.

[5] The Wal-Mart Corporation has other types of stores that I exclude in the analysis. In particualr, I am excluding Sam's Club (a wholesale club) and Neighborhood Market stores, Wal-Marts recent entry into the pure grocery store segment.

million. The breakdown is \$42 million per regular store and \$76 million per supercenter. The average employment is 223 and the average square feet is approximately 150,000.

As part of its expansion process, Wal-Mart routinely tears down old stores and builds larger ones either on the same property or just down the road. However, it is an extremely rare event for Wal-Mart to shut down a store and exit a location. I estimate this has happened on the order of 30 times over a 42 year period in which Wal-Mart has opened 3,000 stores. Since it is negligible, I am going to ignore exit in the analysis and focus only on openings.

Every Wal-Mart store has a store number. Wal-Mart stores retain this number even when they are upgraded and relocated down the street, which makes it very convenient for keeping track of the stores. The first Wal-Mart store opened in Rogers, AR in 1962. This is Wal-Mart store #1. The next store opened in 1964 in Harrison, AR, store #2. Since numbers are assigned in sequential order, store number provides very good information about relative rankings of store opening dates. In the first version of this paper, I just pulled the information about store numbers and addresses from Wal-Mart's web site, and combined this with counts of stores by year from annual reports to come up with estimates of store age. This is a reasonably accurate dating system but it isn't perfect. A potential store site picks up a number in the planning process, but it might not be built right away, so store openings aren't perfectly sequenced with store number. Also, the original store #3 was closed, but this number was reused for a store that opened in 1989. This sort of thing is rare, but it does happen.

Fortunately, additional information is available to give a more precise dating method. Emek Basker (see Basker (forthcoming)) has assembled data on store openings from store # 1 in 1962 up to about the year 2001. and I use her data. In its annual reports up to 1978, Wal-Mart published a complete list of all its stores. Basker uses this information as well as analogous information from directories for later years up to the year 2001. I combined the Basker data with information from TradeDimensions and information from Wal-Marts web site about openings since 2001 to determine the year that each store opened. This is data element two. Table 2 reports the frequency distribution of opening year categories. In the 1970s, Wal-Mart added about 30 stores a year. Since that time it has averaged over 100 new stores a year.

The third data element, demographic information, comes from the three decennial censuses, 1980, 1990, 2000. The data is at the level of the *block group*, a geographic unit finer than the Census tract. Summary statistics are provided by Table 3. In 2000, there were 206,960 block groups with an average population of 1,350. The Census provides information about the geographic coordinates of the block group which I use extensively in the analysis. For each block group I determine all the block groups within a five mile radius and add up the population of these neighboring areas. This population within a five mile radius is the population density measure m I use in the analysis. With this measure, the average block group in 2000 had a population density of 219,000 people per five mile radius. The table also reports mean levels of per capita income, share old (65 or older), share young (21 or younger), and share black. The per capita income figure is in 2000 dollars for all the Census years using the CPI as the deflator.[6]

The fourth data element, data on land values for Wal-Mart stores, was obtained from county tax records. At this point, only data for stores in Minnesota and Iowa have been collected (more to follow). The data was obtained from the internet for those counties posting records. Through this method, I was able to obtain the assessed valuations for half of the stores in these states (50 stores in total). Counties in rural areas are less likely to post valuations on the internet for obvious fixed cost reasons. But this selection is not an issue in my analysis since I control for population density.

The fifth data element is average retail wage by county for the year 2000 from County Business Patterns. The variable is total payroll divided by number of retail employees. This wage information is cruder than some other possibilities in terms of its wage information, e.g. the PUMS data. However, its availability at the county level affords a richer geography than other sources.

3.2 Facts about the Diffusion of Wal-Mart

Any discussion of the diffusion of Wal-Mart is best started by viewing on a map the year-by-year expansion of stores. Figure 1 shows the expansion process for years 1971-1980. The

[6] Per capita income is truncated from below at $5,000 in year 2000 dollars.

expansion process over the entire period 1962-2004 is posted on the web.

From inspection of this process it is clear that Wal-Mart diffusion path was from the inside out. Starting from Bentonville AR as the center, it gradually expanded its radius over time. There is one case of a jump where between 1980 and 1981 it filled in South Carolina, skipping most of Georgia. (But coming back to fill it in soon enough.) This is due so an external expansion when it bought Kuhn's Big K and added a large number of stores. The rest of the expansion process is smooth. External expansion such as what happened in 1981 is rare. (My comment refers to domestic expansion. Foreign expansion has frequently taken place though acquisition.)

Along its expansion path, Wal-Mart made choices along the way about priority locations. It is well known that it avoided very large cities, at least initially. Some evidence of Wal-Mart's priorities can be obtained by looking at where they are at now. Table 4 presents information on the average distances to the nearest Wal-Mart across block groups. Consider those block groups in the highest density category, 500 thousand or more within a 5 mile radius. Average distance to the nearest Wal-Mart, weighted by population, is 6.7 miles. If we look at the next lower density category, distance falls to 4.2 miles and then again it falls to 3.7. Thereafter distance increase as density falls. If we go all the way to extremely sparse locations, the average distance is 24 miles. Wal-Mart is known for preferring small towns. But as Table 4 makes clear, it is actually medium-sized towns that are the "sweet spot" for Wal-Mart.

The next column conditions on block groups that are within 25 miles of a Wal-Mart to start. This decreases average distance, of course, but the pattern remains the same. I condition in this way to make comparisons with earlier years. By conditioning in this way, we are restricting attention to Wal-Mart's market area and then we can look where it is putting its stores in its market area. The basic pattern is the same if we go back to 1990 or 1980, a U-shaped relationship. Interestingly, the "sweet spot" is changing. In 1980, block groups in the 10-20 density range used to be the closest to a Wal-Mart. In 1990 the sweet spot was 20-40. Now the broad range of 40-250 is the sweet spot.

4. First Stage Parameter Estimation

In the first stage, I estimate in pieces various parameters of the model. I take the pieces to the second stage analysis of the dynamic problem of Wal-Mart.

Part 1 of this section estimates the demand parameters. Part 2 estimates various cost parameters. I only have data from one year to estimate demand and costs. So Part 3 explains how I extrapolate to other years.

4.1 Demand Estimation

With a given vector θ of parameters from the demand model, we can plug in the demographic data and obtain predicted values of revenues $\hat{R}_j(\theta)$ for each store j from equation (4). Let ε_j be the difference between log actual sales and log predicted sales,

$$\varepsilon_j = \ln(R_j) - \ln(\hat{R}_j(\theta)).$$

I assume the discrepancy is normally distributed measurement error. I estimate the parameters with maximum likelihood.

Before going to the estimates I have to take care of two unresolved issues. The first is about specification. I need to specify the forms of the reduced form functions $f(m)$ and $\tau(m)$. Assume

$$\begin{aligned} f(m) &= \omega_0 + \omega_1 \ln(\underline{m}) + \omega_2 (\ln(\underline{m}))^2 \\ \tau(m) &= \tau_0 + \tau_1 \ln(\underline{m}) \end{aligned}$$

for

$$\underline{m} = \max\{1, m\},$$

for population density in thousands. (Thus the minimum value of $\ln(\underline{m})$ is zero.)

The second issue is what to do about supercenters. As can be seen in Table 1, supercenter sales are almost twice as large as regular store sales. What is going on here is clear: supercenters have a broader product line, so everything else the same we would expect supercenters to have larger sales. But this is not something that fits easily into the model just outlined. Even if I were able to put supercenters cleanly in the demand model, in my later analysis I would have the problem that I don't know the dates when a given supercenter was converted from a regular store, I only know store openings. (A large percentage of supercenter were once regular stores.) My product would be a lot simpler if Wal-Mart had never got into the supercenter format.

I finesse the supercenter issue in the following way. I imagine that for the consumer, shopping for groceries and shopping items found at a regular Wal-Mart are two separate things and the activities take place at separate shopping trips. (Of course this goes against one of the basic premises of the supercenter format.) A supercenter is then two distinct stores: a regular Wal-Mart combined with a grocery store. The demand model described above just applies for the regular Wal-Mart component of a supercenter. The predicted sales $\hat{R}_j$ for a store j that is a supercenter is only the predicted sales of the items in a regular store. If I observed a breakdown of sales for each supercenter into those items carried at regular-store items and those not carried, then my sales figure I would use in the estimation would just be the regular items component. However, this is unobserved for supercenters. My strategy then is to exclude the unobserved data in my likelihood function. But importantly, the supercenters remain in the choice set of consumers. So if a regular store is near a supercenter, it's sales will be lower, everything else the same.

Table 5 reports the demand estimates for three specifications. The specifications differ in the extent to which store age is used as a store characteristics. Specification 1 uses no store-age information. It fits the data reasonably well, with an R^2 of .674. Specification 2 adds a dummy variables for stores 2 years and older from brand new stores. The effect age is substantial, a mature store increases log sales by .25. Specification 3 breaks the mature category into four different groups. There is some effect of further increases in age. The effect increase from .24 for 3-5 to .319 for 6-10. But the differences are relatively small compared to the effect of just being 2 or above. And there is not much improvement in goodness of fit. I will use specification 2 for my baseline model of demand. An advantage of this specification for later use relative to specification 3 is that the impacts of a change in store location will not have a lagged effect 20 years down the line as is the case for Specification 3.

The parameters in Table 5 are difficult to interpret directly so I will look at how fitted values vary with the underlying determinants of demand. Table 6 examines how demand varies with distance to the closest Wal-Mart and population density. For the analysis, the demographic variables are set to their mean level from Table 3. There is assumed to be only one store within the vicinity of the consumer (i.e. within 25 miles) and the distance of this single Wal-Mart is varied in the table. Consider the first row, where distance is set to zero (the consumer is right-next door to a Wal-Mart) and population density is varied. As expected, there is a substantial negative effect of population density on demand. A rural consumer right next to a Wal-Mart shops there with a probability that is essentially one. With a population density of 40 this falls to .77 and up to 250 it falls to less than .25. In a large market there are many substitutes. Even a customer right next to a Wal-Mart is not likely to shop there. While *per capita* demand falls, overall demand overwhelmingly increases. A market that is 250 times as large as an isolated market may have a per capita demand that is only a fourth as large, but overall demand is over 50 times as large.

Next consider the effect of distance holding fixed population density. In a very rural area, increasing distance from 0 to 5 miles has only a small effect on demand. This is exactly what we would expect. Now raising the distance from 5 to 10 miles does have an appreciable effect, .971 to .596. In thinking about the reasonableness of this effect, it is worth noting the miles here are "as the crow flies," not driving distance. An increase of 5 to 10 could be the equivalent of a 10 to 20 mile increase in driving time. In that light, the change in demand from .971 to .596 seems highly plausible. Demand taper out at 15 miles and goes to zero at 20 miles.

Next consider the effect of distance in larger markets. The negative effect of distance begins much earlier in larger markets. For a market of size 250, an increase in distance from 0 to 5 miles reduces demand by on the order of 80 percent while the effect of distance in rural markets is miniscule. This is what we would expect.

Other demand characteristics are of note. It is possible to calculate consumer demand when there are multiple Wal-Marts in his or her area. At the mean characteristics, if a consumer is zero miles to one Wal-Mart and 2 miles to another, (and no others are in the area), the consumer goes to the one next door with probability .75 and the other with probability .25, conditioned upon shopping at one. So allowing for product differentiation among Wal-Mart, instead of just assuming consumers shop at the closest one, is important. But if the distance disadvantage of the further store is increased, demand for the further store drops off sharply.

Demand varies by demographic characteristics in interesting ways. Wal-Mart is an inferior good in that demand decreases in income. Demand is higher among whites and lower among younger people and older people.

4.2 Labor Costs

Regressing log of employment on log of sales for 2004, I obtain the labor requirements function,

$$\begin{array}{ccccccc} \ln Labor & = & 2.29 & + & .74 & \ln R \\ & & (.06) & & (.02) & \end{array}.$$

Next I obtain an estimate of the function $W(m)$ which specifies how the retail wage varies with population density. I project average county wage on a quartic equation in population density (the coefficients not reported here.) Table 7 shows how average actual wage and the fitted wage varies with population density. As is typical, measured wages increase in density. For the under 10 category, the wage is \$17,150 which increases to \$18,520 for the 10-40 category and even higher thereafter.

There are obvious measurement difficulties here. Pay divided by total employment is a crude measure of the wage since hours worked varies substantially across individuals, particularly in retail. However, since my labor input level is in employment, not hours, even if I could come upon hourly wage information I would have to get data on hours from Wal-Mart to use it and such data is not available..

Of course I am not taking into account differences in labor quality across locations either here. There is evidence in the urban economics literature that workers in larger cities are better quality (see Glaeser and Mare). Later I show that Wal-Mart could have earned substantially more revenues if it reordered its opening sequence and went to larger cities first as compared to smaller cities. To the extent Wal-Mart could have obtained higher quality workers from this perturbation, it means my results understate the density cost savings it achieved by doing what it did.

4.3 Land Costs

Wal-Marts typically use relatively large plots of land, on the order of 10, 15, to even 20 acres. To open a Wal-Mart with this size of a plot of land in Manhattan would cost a fortune. So to open a Wal-Mart in a very urban area would results in substantial increases in land rents compares to a less urban area. Nevertheless, *a priori*, it is not obvious that rents in medium size cities will be more than rents in small cities or rural areas. Wal-Mart tends to open its stores on the outskirts of town. In the standard urban theory, rents on the outskirts of town equal the agricultural land rent.

To examine this hypothesis, I use the land value data on the 50 Wal-Marts in Iowa and Minnesota that I have collected. I don't know acreage, so I make use of the fixed coefficient assumptions made in (5) and assume acreage is proportionate to building size. I then regress the log of land prices (assessed value divided by building square footage) on dummy variables by population density class. I also include state fixed effects as well as age of the store. The results are reported in Table 8. Comparing the "Under 10" density class with the 40-80 and 80 and above density classes I find significant differences in land prices. Plugging in the coefficient estimates, the predicted prices differ by factors of 2.6 and 3.4, respectively, from the "Under 10" group. But the differences between "Under 10" and "10-40" are negligible In the analysis I will treat the land prices for these groups the same.

4.4 Other Costs

In the analysis I set gross margin less nonlabor variable costs equal to

$$\mu - P_{Land} v_{Land} - P_{Bldg} \nu_{Bldg} - \nu_{Misc} = .17.$$

The price of land applies for locations with density ≤ 40. (Since locations with density ≥ 40 are not altered, pricing for such parcels is not needed). Note land and buildings are variable costs here because larger sales require more space.

Wal-Mart's gross margin over the years has ranged from .22 to .26 (from Wal-Marts annual reports.), so $\mu = .24$ is a sensible value. The mean ratio of assessed value of land and building to annual sales in my sample is .14. Converting this to rental values results in a figure on the order of .01 to .02 for the quantity $P_{Land} v_{Land} + P_{Bldg} \nu_{Bldg}$. Setting ν_{Misc} to be on the order of .05 to .06 is on the high side. There is much cost that takes place outside of the store. I have already discussed how I am taking store-level labor costs out of this. And there are is also that large profit margin to consider. Here I am being conservative and erring in the direction of understating variable profit. This works against the incentive to increase revenues by going to larger markets.

4.5 Extrapolation to Other Years

So far I have constructed a model of Wal-Marts demand and costs circa 2003, the year of the TradeDimensions data. I will need a demand and cost model for all the years that Wal-Mart was in business to study its diffusion path.

Growth in Wal-Mart on a per store basis is remarkable. We see from Table 1 that in 2003, average store sales (regular stores)was \$42.4 million. In 1972, average sales (in 2003 dollars) was only \$11.1. How can I take this into account.

I applied the following procedure. First, I took the exact demand model from 2003 and evaluated average sales per store in the prior years, given the configuration of stores for each of these prior years. The 2003 demand model evaluated at the store configuration for 1972 predicted an average store sales (in 2003 dollars) of \$31.4 million. So one third of the difference in average in average store size of 11.1 in 1972 and 42.4 in 2003 is due to the change in the average market size from the two periods. The rest of the difference is unexplained. I attribute this to productivity growth. I determine the average growth r_{1972} from 1972 to 2003 that would generate the sales difference of 11.1 to \$31.4. The annual growth in this case is approximately .04. Proceeding this way, I determined that the following simple series fit well. Growth before 1980 at $r = .04$, growth after 2000 at $r = .02$ and linearly interpolating for the 20 years in between.

This growth factor was applied to all the cost functions as well. The impact of this assumption is that if Wal-Mart keeps the same set of stores over a given time period, and demographics were held fixed, then revenue and costs increase by a proportionate amount, so profit increases by a proportionate amount.

The growth factor applies holding demographics fixed. But demographics changed over time and I take this into account as well. I use data from the 1980, 1990, and 2000, decennial censuses. For years before 1980, I use 1980, for years after 2000 I use 2000. For years in between I use a convex combination of the appropriate censuses as follows. For example, for 1984 I convexify by placing .6 weight on 1980 and .4 weight on 1990. I so this by assuming that only 60 percent of the people in the people from the given 1980 block group are still there and that 40 percent of the people form the 1990 block group are already there as of 1984. This procedure is clean, since I avoid the issue of having to link the block groups longitudinally over time, which would be very difficult to do. Given my continuous approach to the geography, there is no need to link block groups over time.

5. Stage Two: Density Economies

There remain the two parameters of density economies to estimate, the spillover decay parameter α and the multiplicative term ϕ. Given the high dimensionality of the choice problem, standard procedures are infeasible here. This leads me to consider a set of three alternative procedures based on perturbation methods. The approaches differ in the required assumption about the random profit term η_i associated with action i. The first two approaches have a well-developed econometric theory. The third approach is exploratory.

5.1 Methods

5.1.1 PPHI

Approach one is a moment inequality method. There has been much recent work in this literature including Andrews, Berry, and Jia (2004). I follow the treatment in Pakes, Porter, Ho, Ishii (2005) (hereafter PPHI).

For a given value of α, and a policy choice $i \in A$, let $h_i(\alpha)$ be

$$h_i(\alpha) = \sum_{t=1}^{T} \beta^{t-1} \sum_{j \in B_{it}} \frac{1}{s_{ijt}(\alpha)}.$$

This is the discounted value of the spillover inverse at all open stores. The present value of density economies for policy i is then $D_i = -\phi h_i(\alpha)$. The total value of the policy is

$$v_i = \pi_i - \phi h_i(\alpha) + \eta_i,$$

where π_i is the discounted value of operating profit and η_i is the random component of the value of policy i. Suppose we number policies so that $i = 0$ is the policy that Wal-Mart actually chose.

Following PPHI, allow for the possibility of optimization error by the decision marker. Here, I let δ parameterize the absolute level of allowable error. Let $A_{PPHI} \subset A$ be a subset of the set of feasible actions. Let the random profit realization for each action i take the form

$$\eta_i = -\omega h_i(\alpha).$$

where ω has zero mean. The optimality of action 0, with allowable optimization error δ, implies the following inequalities

$$\begin{aligned}
& v_0 - v_1 + \delta \\
= \; & [\pi_0 - \phi h_0(\alpha)] - [\pi_i - \phi h_i(\alpha)] + \eta_0 - \eta_i + \delta \\
= \; & \pi_0 - \pi_i + \delta - (h_0(\alpha) - h_i(\alpha))(\phi + \omega) \geq 0.
\end{aligned}$$

Dividing through by $h_0(\alpha) - h_i(\alpha)$ yields the following moment inequalities:

$$\begin{aligned}
\Delta_i(\alpha, \phi) &= \frac{\pi_0 - \pi_i + \delta}{h_0(\alpha) - h_i(\alpha)} - \phi - \omega \geq 0, \text{ if } h_0(\alpha) - h_i(\alpha) > 0 \\
&= \frac{\pi_0 - \pi_i + \delta}{h_0(\alpha) - h_i(\alpha)} - \phi - \omega \leq 0, \text{ if } h_0(\alpha) - h_i(\alpha) < 0
\end{aligned} \tag{9}$$

The method is to find the set of ϕ and α that satisfy (9) for all $i \in A_{PPHI}$. In the event that no (ϕ, α) pair satisfy all the inequalities, I pick the pair that minimizes the sum of the squared deviations.

While the structure of this problem maps directly into the PPHI setup, it should be noted that there is only a single error draw here, the variable ω. There is only a single decision made here. The results in PPHI for constructing confidence intervals require multiple error draws and so do not apply here.

The final step is to construct a set deviations A_{PPHI}. I limit the set of alternatives considered in two ways. First, I consider deviations that differ from what actually happened over a limited time period and for limited geography. In particular, I set a t_{low} and t_{high} so that stores opening before t_{low} or after t_{high} are unchanged. I set a population density cutoff

$\hat{m}$ so that if $m_j > \hat{m}$, the opening date is left unchanged. I don't reallocate stores opening in high population density areas so that I can avoid having to determine urbanization costs. Let $A(t_{low}, t_{high}, \hat{m}) \subset A$ be the set of alternatives that differ from actual choice 0 only over this restricted set.

Second, I consider policies that have something to recommend them, i.e. they trade-off operating profits and density economies. In an ideal world I would solve the following problem for each ϕ and α,

$$i^*(\alpha, \phi) = \max_{i \in A(t_{low}, t_{high}, \hat{m})} v_i = \pi_i - \phi h_i(\alpha).$$

This is the solution with no random terms. I can't solve this problem because of the size of A. Instead, I consider the following algorithm . I start at the beginning with a pool of stores to be resequenced. I chose as my next store the one with the largest increase in incremental profit, operating profit less density-related costs. After I add this store, I go back to the pool and repeat. Call this the one-step-ahead algorithm. Let $i^{**}(\alpha, \phi)$ be the policy derived from this algorithm. I take a grid of α and ϕ and calculate $i^{**}(\alpha, \phi)$ for this finite set of discrete points. I define A_{PPHI} be be the set of alternative generated by varying α and ϕ in this finite set.

5.1.2 Bajari and Fox

Approach two is a pair-wise comparison, maximum score type method developed in Fox (2005) and Bajari and Fox (2005). To apply the method, I need to assume that the η_i are i.i.d. There is reason to be concerned with this assumption since policies that are similar to other policies would presumably have similar random terms and this would violate the i.i.d. assumption. While this error assumption is difficult to swallow, the procedure is easy to implement. So for comparison purposes I implement the procedure..

To implement the Bajari and Fox procedure, I make pair-wise comparisons. I select pairs of stores and switch opening dates. Let $A_{BF}(t_{low}, t_{high}, \hat{m})$ be the set of alternative policies that differ with the actual choice 0 in the opening dates of exactly two stores. Moreover, only stores opening in the time interval $[t_{low}, t_{high}]$ are switched and with $m_j < \hat{m}$. Finally, I restrict attention to stores with opening dates within two to four years apart.

For any (ϕ, α), let

$$\tilde{\Delta}_i(\phi, \alpha) = \pi_0 - \pi_i - \phi \left(h_0(\alpha) - h_i(\alpha) \right).$$

The estimator picks the (ϕ, α) that maximizes the number of times $\tilde{\Delta}_i(\phi, \alpha) > 0$ for $i \in A_{BF}(t_{low}, t_{high}, \hat{m})$.

5.1.3 Store-Level Randomness Approach

In the third approach, I assume that each store location j has a random profit component ε_{jt} at time t. For a given store j, the random component varies over time deterministically by the exogenous growth factor, $\varepsilon_{j,t} = \varepsilon_{j,t-1}(1 + r_t)$. The initial draws $\varepsilon_{j,0}$ are i.i.d. normal variables across stores with variance σ_ε^2. The random component is begins to flow to the firm the date when it is first opened. The variable η_i is the present value of store-level random components associated with a particular policy i.

The error structure here is appealing. Two policies that are approximately the same (e.g. differ only in the timing of two stores that open a year apart) will have η_i values that will be approximately the same.

Unfortunately, while maximum likelihood estimation is conceptually straightforward here, it is not feasible, as the ensuing discussion will make clear. I consider a alternative approach that is motivated by my interest in MLE. It is different from MLE in three ways. As I explain what I do, I will explain how I can hope to get something closer to MLE in each of these three respects, so this might be a fruitful avenue to pursue.

First, in what I do, I consider only actions that vary from the actual policy, policy 0, in periods $t \in \{t_{low}, ..., t_{high}\}$, and for stores with population densities below the thresholds $m_j < \hat{m}$. Consider a problem were we take as given the timing of store choices outside this time interval and above this population threshold. I assume that the firm solves the constrained problem, picking policy $i \in A(t_{low}, t_{high}, \hat{m})$, given the draws $\varepsilon_{j,t}$ for the stores that can be interchanged. MLE is consistent if the firm is in fact subject to these constraints. If the firm is solving the more general problem, MLE would in general not be consistent. Take this as a caveat in what I have done so far; in future work I intend to expand the choice set of the firm; this should diminish the significance of the issue.

So for the rest of this section, assume that the firm's problem is

$$\max_{i \in A(t_{low}, t_{high}, \hat{m})} v_i = \pi_i - \phi h_i(\alpha) + \eta_i. \tag{10}$$

For η_i the weighted sum of normal variables, being the present value of the store-level random terms associated with a given policy i. Let η be a vector of draws for each policy. Our interest is estimating ϕ and α.

Second, in what I do, I consider a constrained version of problem (10). As before, let $i^*(\phi, \alpha, \eta)$ be the policy that solves (10), for a given vector η. Let H be a set of vectors η. Let this set include the case of $\eta = 0$ as well as a set of K draws of η for a given σ_ε^2. Let $A^*(t_{low}, t_{high}, \hat{m}, H)$ be the set of i in $A(t_{low}, t_{high}, \hat{m})$ for which there is a (ϕ, α) and a $\eta \in H$, such that $i = i^*(\phi, \alpha, \eta)$. Consider the problem

$$\max_{i \in A^*(t_{low}, t_{high}, \hat{m}, H)} v_i = \pi_i - \phi h_i(\alpha) + \eta_i. \tag{11}$$

Let (ϕ^*, α^*) be the maximum likelihood estimates given the actual choice is 0 and given the firm solves the above constrained problem (and for a given σ_ε^2 and a given H). This is much easier to calculate than the maximum likelihood estimate of a firm solving (10), call this $(\phi_{MLE}, \alpha_{MLE})$. Furthermore, a case can be made that it approximates the MLE estimates. For small σ_ε^2, it is immediate that

$$\underset{\sigma_\varepsilon^2 \to 0}{plim} (\phi^*, \alpha^*) = (\phi_{MLE}, \alpha_{MLE}).$$

Next consider a for fixed σ_ε^2 what happens for large K, again, the number of draws of η in H used for constructing A^*. Though I don't have a formal proof at this point, it appears obvious that

$$\underset{K \to \infty}{plim} (\phi^*, \alpha^*) = (\phi_{MLE}, \alpha_{MLE}).$$

Next to determine what happens for large σ_ε^2 and small K, I turn to Monte Carlo evidence. I consider an example where initially there are 5 stores, and the firm has to sequence 12 stores, choosing 6 in the first period and adding 6 in the last period. The actual value of $\phi = 5$ in the example. I set $K = 15$, the number of draws of η. Given this H, there are 37 different policies in A^*, compared with 924 policies in A for problem (10). I set a high level of σ_ε^2, so that ϕ_{MLE} is a noisy estimate. For the purpose of the exercise, I am not interested in how well ϕ_{MLE} tracks ϕ, but rather how well ϕ^* tracks ϕ_{MLE}. Figure 2 illustrates ϕ_{MLE} and ϕ^* for 500 random draws of η. The constrained estimate ϕ^* tracks ϕ_{MLE} well. The correlation is .983. It shouldn't be that surprising that ϕ^* would track ϕ_{MLE}. Essentially, I am throwing out from consideration policies that would tend to low payoffs, for any ϕ and α. These are unlikely to be chosen, so throwing them out doesn't make much difference.

Third, when I go to the data I don't use $i^*(\phi, \alpha, \eta)$ to construct the deviations for a given η and a grid of ϕ and α. Rather, I use $i^{**}(\phi, \alpha, \eta)$, defined earlier as the policy solving the 1-step-ahead algorithm. This is obviously an imperfect approximation. In future work, I will generalize the 1-step-ahead algorithm to an N-step-ahead algorithm. Even looking ahead five stores, the dimensionality of the problem is greatly reduced compared to the actual problem (10) and should be feasible.

A problem with the N-step ahead algorithm is that when ϕ is large, it might miss out on cases where it would be optimal to make a big jump leaving a current agglomeration to find a new agglomeration. I expect that the N-step ahead algorithm will work well in finding desirable local changes.

5.2 Time Period Considered

I set $\hat{m} = 20$. I consider two different deviation time intervals. (1) 1971-1980, where 200 stores are in the pool to be reallocated and (2) 1982-1989, with 377 stores in the initial reallocation pool.[7]

For illustration purposes, Table 9 reports π_i and $-h_i(.01)$ and for $\iota^{**}(0, .01)$ and $\iota^{**}(\infty, .01)$, i.e. the solutions to the 1-step ahead problem when density related costs are irrelevant ($\phi = 0$) or infinite ($\phi = \infty$). The values are discounted to the first period of the deviation. The present values include flows up to the period $t_{end} + 1$. I go one period after to take account of the lag in demand.

Start with the first time interval, 1971-1980. For policy 0, the actual policy chosen, discounted operating profit over the interval (in 2003 dollars) is \$1,413. Next consider the solution to the one-step ahead problem with $\phi = 0$. For this case, spillovers don't matter; discounted operating profit increases by 107 million dollars. This comes at a cost.

[7] I leave out 1981 because that is the year of the Kuhn's K acquisition.

Normalized density-related benefits decrease from -17 to -22.

The second deviation is constructed to maximize density economies. Here I solve the 1-step ahead problem with $\phi = \infty$. This raises (normalized) density economies from -17 to -15.

5.3 Estimates

Estimates at this point are tentative. Before presenting the results, it is worth nothing that there are two ways to increase the importance of density economies. For fixed α, increasing ϕ does this. For fixed ϕ, increasing α does this, since this decreases the spillover received by each store and thus increases the inverse of the spillover. As I consider different specifications, there seems to be a trade-off between α and ϕ in the estimates. In some specifications, density economies are big because α is high, but ϕ is low; in others it is reversed.

My first set of estimates fix $\alpha = .01$ and apply the PPHI method to the extreme deviations presented in Table 9. Using the data from 1970-1980, my estimate is $\hat{\phi} = 17.4$. For the later period, the estimate doubles to 31.7. While different, the estimates are within the same order or magnitude.

In applying the PPHI to the larger set of deviations A^{**} (generated by a grid of (ϕ, α) and $i^{**}(\alpha, \phi)$) I estimate α as well as ϕ. For the early period, I estimate $\alpha = .047$. With this high level of α comes a low level of ϕ.

Using the Bajari-Fox method, the estimates are roughly in the same range for 1971-1980. But for the later period, the estimate of ϕ goes to the corner.

Discussion of magnitudes to be completed.

REFERENCES

Andrews, Donald W.K., Steven Berry, and Panle Jia, (2004) “Confidence Regions for Parameters in Discrete Games with Multiple Equilibria, with an Application to Discount Chain Store Location,” Yale University working paper.

Bajari, Pat, Benkard, L., and J. Levin, “Estimating Dynamic Models of Imperfect Competition, ” Stanford University Working Paper, 2004.

Bajari, Patrick, and Jeremy Fox (2005), “Complementarities and Collusion in an FCC Spectrum Auction,” University of Chicago Working Paper

Basker, Emek,“Job Creation or Destruction? Labor-Market Effects of Wal-Mart Expansion” *forthcoming Review of Economics and Statistics*.

Bresnahan and Reiss, "Entry and Competition in Concentrated Markets" Journal of Political Economy v99, n5 (Oct 1991):977-1009.

Ellickson, Paul E., "Supermarkets as a Natural Oligopoly, " Duke working paper. (2004)

Fox, Jeremy, (2005), “Semi and Nonparametric Estimation of Multinomial Discrete Choice Models Using a Subset of Choices,” University of Chicago Working Paper

Ghemawat, Pankaj, Ken A. Mark, Stephen P. Bradley,(2004) “Wal-Mart Stores in 2003,” Harvard Business School Case Study 9-704-430.

Hausman, Jerry and Ephraim Leiptag, (2005) "Consumer Benefits from Increased Competition in Shopping Outlets: Measuring the Effect of Wal-Mart,” MIT working paper.

Holmes, Thomas J., "Barcodes lead to Frequent Deliveries and Superstores," Rand Journal of Economics Vol. 32, No. 4, Winter 2001, pp 708-725.

Holmes, Thomas J., (2005) "The Location of Sales Offices and the Attraction of Cities," , *Journal of Political Economy*

Ledyard, Margert (2004), "Smaller Schools or Longer Bus Rides? Returns to Scale and School Choice " working paper.

Pakes, Ariel, Jack Porter, Kate Ho, and Joy Ishii, “Moment Inequalities and Their Application,” Harvard working paper, April, 2005.

Roberts, Mark J. “Economies of Density and Size in the Production and Delivery of Electric Power Mark, *Land Economics*, Vol. 62, No. 4. (Nov., 1986), pp. 378-387.

Stone, Kenneth E., “Impact of Wal-Mart Stores on Iowa Communities: 1983-93,” Economic Development Review, Spring 1995, p. 60-69.

Toivanen, Otto and Michael Waterson (2005), "Market Structure and Entry: Where’s the Beef?" *Rand Journal of Economics,* Vol 36, No. 3,

Vance, Sandra S. and Roy Scott, *Wal-Mart: A History of Sam Walton’s Retail Phenomenon*, Simon and Schuster Macmillan: New York, 1994.

Table 1
Summary Statistics: TradeDimensions Data

Store Type	Variable	N	Mean	Std. Dev	Min	Max
All	Sales ($millions/year)	2,936	59.6	29.2	5.2	170.3
Regular Store	Sales ($millions/year)	1,457	42.4	19.3	5.2	122.2
SuperCenter	Sales ($millions/year)	1,479	76.5	27.4	13.0	170.3
All	Employment	2,936	223.4	132.6	31.0	801.0
Regular Store	Employment	1,457	112.2	38.2	47.0	410.0
SuperCenter	Employment	1,479	332.8	96.5	31.0	801.0
All	Building (1,000 sq feet)	2,936	143.1	54.7	30.0	286.0
Regular Store	Building (1,000 sq feet)	1,457	98.6	33.3	30.0	219.0
SuperCenter	Building (1,000 sq feet)	1,479	186.9	31.1	76.0	286.0

Table 2
Distribution of Wal-Mart Stores by Year Open

Period Open	Frequency	Cumlative
1962-1970	25	25
1971-1980	277	302
1981-1990	1,236	1,538
1991-2000	1,080	2,618
2001-2003	318	2,936

Table 3
Summary Statistics for Census Block Groups

	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

Table 4
Mean Distance To Nearest Wal-Mart across Census Blockgroups
By Density and Year

		All Blkogroups	Conditioned upon blockgroups being with 25 miles of Wal-Mart at given Year		
Population Density (1,000 in 5 mile radius)	Population Percentile of Category	2003	2003	1990	1980
Under 1	1.3	24.2	15.1	17.2	17.2
1-5	9.6	16.7	13.2	16.6	16.6
5-10	16.1	11.3	9.9	15.0	14.2
10-20	24.0	7.2	6.6	13.6	12.2
20-40	33.2	5.1	4.8	12.4	12.3
40-100	50.4	4.0	3.9	13.3	15.5
100-250	76.2	3.7	3.6	15.7	18.4
250-500	90.2	4.2	4.2	16.7	19.8
500 and above	100.0	6.9	6.9	21.1	19.2

Table 7
Average Retail Wages and Population Density

Density Category (1,000 in 5 mile radius)	Actual Wage	Fitted Wage
Under 10	17.15	17.06
10-40	18.52	18.55
40-100	19.70	20.06
100-250	21.61	21.32
250-and up	22.76	22.88

Source: County Business Patterns 2000 and author's calcultions.

Table 8
Land Price Regression
Dependent Variable: Log of Estimated Land Price
(Excluded density group is 0-10)

Constant	2.09
	(.29)
Population Density 10-40	-.04
	(.27)
Population Density 40-80	.96
	(.28)
Population Density 80 and above	1.23
	(.27)
Store Age	.02
	(.02)
Iowa Dummy	-.64
	(.23)
N	43
R^2	.63

Table 9
Discounted Values over Deviation Intervals
(Millions of 2003 dollars)

Interval 1: 1971-1980

	Revenue	Operating Profit	Normalized (ϕ=1) Density Economy (α = .01)
Actual Policy	14,965	1,413	-17
$i^{**}(\varphi=0)$	15,950	1,519	-22
$i^{**}(\varphi=\infty)$	14,653	1,380	-15

Interval 2: 1982-1989

	Revenue	Operating Profit	Normalized (ϕ=1) Density-Economy (α = .01)
Actual Policy	133,577	12,673	-84
$i^{**}(\varphi=0)$	136,666	13,004	-89
$i^{**}(\varphi=\infty)$	133,661	12,683	-79

Table 10
Parameter Estimates for Density-Economies

	Interval 1 Estimates 1971-1980		Interval 2 Estimates 1982-1989	
Method	ϕ	α	ϕ	α
PPHI Table 9 Deviations (α fixed)	17.4	.010	31.7	.010
PPHI (A** Deviations)	2.0	.047	18.6	.025
Bajari-Fox	12.2	.019	∞	.023
Store-Level Randomness Approach	31.0	.008		

Table 11
Distribution of Spillover and Inverse Spillover at Terminal Periods

	N	Mean	Std. Dev	Minimum	Maximum
1980					
s	302	46.13	13.32	7.50	65.96
1/s	302	.024	.012	.015	.133
1989					
s	1320	60.65	21.90	2.07	91.73
1/s	1320	.023	.029	.011	.483

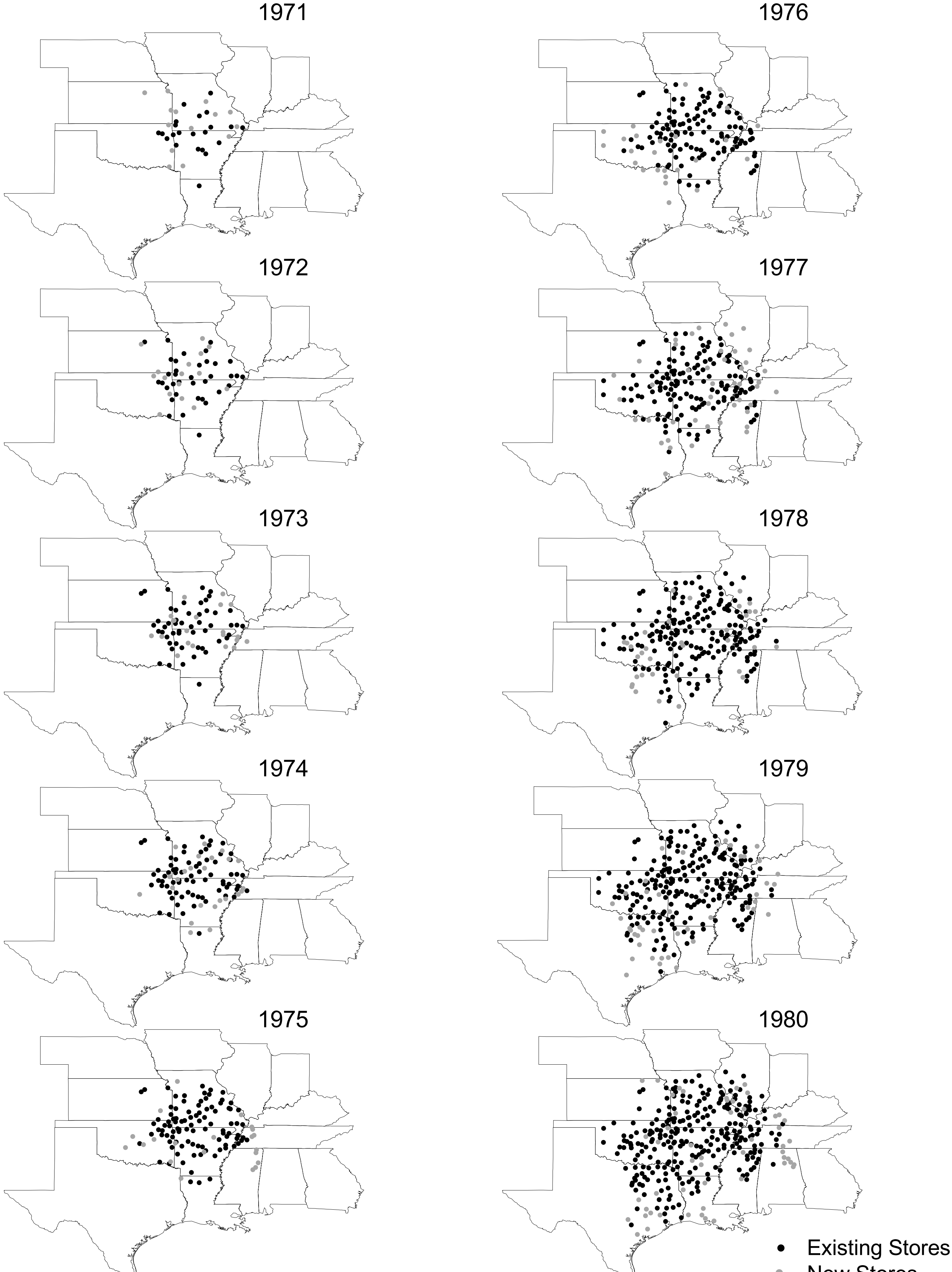

Fig 1: Dispersion of Wal-Mart Stores

Figure 2
Monte Carlo Results
Comparison of phi_MLE and phi_star