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

April, 2007

[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 0551062. I thank Junichi Suzuki, Julia Thornton, and David Molitor 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]

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. My results suggest the benefits are significant.

In choosing store locations, Wal-Mart faces a tradeoff. By concentrating stores in the same area, it enjoys economies of density. Offsetting this gain, however, is diminishing returns in store sales. As Wal-Mart adds more and more stores to a given area, the market areas of the stores begin to overlap and new stores cannibalize sales from existing stores. Because of these diminishing returns, if density economies were negligible, Wal-Mart would not concentrate its stores in one state before moving on to a new state. It would tend to scatter stores around the country, then go back later and fill in. In contrast, density economies are significant, we would expect Wal-Mart to fill out one area before moving to the next.

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. Wal-Mart did not scatter stores in desirable locations throughout the county and then come back for the "high hanging fruit," with fill-in stores. Locations far from Bentonville had to wait to get their Wal-Marts. The process repeated itself in 1989 when Wal-Mart introduced the "supercenter" format, which carries all of the regular merchandise of a Wal-Mart plus groceries. The first supercenters started in the center near Bentonville and again gradually radiated from the middle out.

It would be difficult to directly measure the economies of density that Wal-Mart enjoys.

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

Wal-Mart is notorious for being secretive—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.

But how store sales depend upon store locations is something that can I can estimate. I use store-level sales estimates from ACNeilsen and demographic data from the Census to estimate a model of demand for Wal-Mart at a rich level of geographic detail. Wal-Mart And I combine this with additional information about cannibalization that Wal-Mart itself releases in its annual reports.

Using my sales model, I determine that Wal-Mart has encountered significant diminishing returns in sales as it has piled up many stores in the same area. New stores cannibalize existing firms sales in the same area in a substantial way. From this I conclude that the economies of density must be substantial.

I write down a structural model of Wal-Mart and attempt to quantity parameters relating to density economies. 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 use 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. Using the procedure, I am able to determine a lower bound on the importance of density economies.

The paper contributes to the literature on entry and store location in retail. Related contributions include Bresnahan 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)), Neumark et al (2005), Jia (2005). 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 growth 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.

There are two kinds of Wal-Mart stores. A *regular* Wal-Mart sells general merchandise (e.g. clothes, hardware, toys, etc.) as well as some selected nonperishable food items (e.g. soda). A *supercenter* sells a full line of groceries in addition to general merchandise. One issue that has to be dealt with is how to model shopping decisions for these two classes of products. My approach is to assume consumers make two distinct shopping decisions: (1) where to shop for general merchandise and (2) where to shop for groceries. I recognize this goes against the basic logic of a supercenter to get consumers into a store to conveniently purchase both kinds of merchandise. As I discuss below, data limitations make it difficult for me to estimate this complementarity, so I zero it out. It would be nice to get this complementarity right, but I don't think it is a first-order issue for the question I am looking at.

Now for some notation. There is a set of possible store locations B. At a particular point in time, let $B^{wal} \subseteq B$ be the subset of locations for which there is a Wal-Mart store. I will refer to an element $j \in B^{wal}$ as store j. Some subset $B^{super} \subseteq B^{wal}$ also carry groceries, the supercenters.

There are L discrete locations indexed by ℓ where consumers live. In general, the number of locations will be large relative to the number of Wal-Marts open so each Wal-Mart will tend to draw from many locations. For a given location ℓ, let n_ℓ denote the *population* of location ℓ and let m_ℓ be the *population density* at ℓ. Let $y_{\ell j}$ denote the distance in miles between location ℓ and store location j. Define B_ℓ^{wal} to be the set of Wal-Marts in the *vicinity* of the consumer's location, which I take to be those within 25 miles.

$$B_l^{wal} = \left\{ j, j \in B^{wal} \text{ and } y_{\ell j} \leq 25 \right\}$$

Consider a particular consumer k at a particular location ℓ. I first explain how spending on general merchandise is allocated. The analysis of spending on groceries is similar. The consumer has a budget λ^{gen} for spending on general merchandise. The consumer makes a discrete choice between buying the outside good (good 0) or from one of the Wal-Marts in B_ℓ^{wal} (assuming B_ℓ^{wal} is non-empty). The utility of the outside good 0 is

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

The first term is a function $o(\cdot)$ that depends upon the population density m_ℓ at consumer i's location. Assume $o'(m) > 0$; 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 a 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 that a particular consumer might have. Instead I specify the reduced form relationship between $o(m_\ell)$ 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 from buying general merchandise at given Wal-Mart store $j \in B_\ell^{wal}$ is

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

The first term is the utility decrease from travelling to Wal-Mart store j that is a distance $y_{\ell j}$ from the consumer's home. The weight $\tau\left(m_\ell\right)$ the consumer places on distance can depend 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 a 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 general merchandise from some Wal-Mart is

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

for

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

and the probability of purchasing general merchandise at a particular Wal-Mart $j \in B_\ell^{wal}$, conditional on purchasing from some Wal-Mart is

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

The probability a consumer at ℓ buys his or her general merchandise at Wal-Mart j is

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

Total revenue from general merchandise sales of store j is

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

This equals the spending λ^{gen} 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.

I model spending on groceries exactly the same way. I assume the parameters are the same except for the spending λ^{groc} per consumer. Total revenue from groceries Rev_j^{groc} is calculated analogous to the above. Even thought the parameters are the same as with general merchandise, the probability a consumer shops at a given supercenter $j \in B^{super}$ will in general be different from the probability he or she shops there for general merchandise because the set of alternative supercenters in the vicinity of the consumer is in general different from the set of Wal-Marts in the vicinity (because all Wal-Marts are not supercenters)

2.2 Cost Structure and Openings of New Stores

This subsection describes the cost structure. It first specifies input requirements for merchandise, labor and miscellaneous inputs. These determine operating profit. It next specifies a form of the fixed cost. Finally, it specifies the form of the *density economies*, which will be the main target of the estimation.

2.2.1 Operating Profit

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

I assume labor requirements are proportionate to sales. Let ν_{Labor} be the amount of labor required for one unit of sales. Suppose the wage for retail labor at location ℓ is W_ℓ so that the wage bill is $W_\ell L$.

Suppose the amount of land needed for a store is also proportionate to sales, ν_{Land}, let r_ℓ be the land rent at location ℓ, so land costs are

$$C_{land} = \nu_{Land} r_\ell R$$

Assume there are other costs that are proportionate to sales and are the same at all locations. This would include, for example, the cost of shelving and other aspects of the physical plant (these are assumed to be variable inputs).

$$C_{other} = \nu_{other} R$$

Operating profit equals gross margin less labor costs and miscellaneous costs,

$$\pi = \mu Rev - C_{labor} - C_{land} - C_{other}$$

2.2.2 Fixed Costs

There is also a fixed cost to operating at a particular location. I allow the fixed cost to depend upon population density $f(m)$.

A Wal-Mart store has a distinctive format, a big box, single story facility with a huge parking lot on a convenient interstate exit. This approach has obvious limitations in a big city. If Wal-Mart were to locate in an highly urbanized area, things would have to be done, like using multiple floors and a parking ramp, that would not be necessary in a less urbanized area. By allowing the fixed cost to depend upon population density I mean to capture this. It is also meant to capture other factors for why a Wal-Mart in an urban area would be less desirable than in less populated area.

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. By having stores close to distribution centers, Wal-Mart saves money on shipping but also can respond more quickly to demand shocks. Also included here are savings in marketing cost (advertising) by locating stores near each other.

I take into account two aspects of density, store density and distribution center density. I begin with a measure of the density of stores at a particular location ℓ. I sum up all the stores that are "near" this location. I use proportionate decay at the rate of α per mile. (I fix $\alpha = .02$ and then experiment with different values.) So store density for general merchandise at location ℓ is

$$Density_l^{gen} = \sum_{k \in B^{wal}} \exp(-\alpha y_{\ell k})$$

where $y_{\ell k}$ is the distance from store k to location ℓ. Supercenter density is defined in an analogous way. Each individual Wal-Mart store enjoys a cost savings that depends upon the store density at that location, we call this *density economies* and it takes the following function form

$$\begin{aligned} DensityEconomy_\ell^{gen} &= \phi_\ell^{gen} d_\ell^{gen}, \text{ for} \\ d_\ell^{gen} &\equiv \left[1 - \frac{1}{Density_\ell^{gen}}\right] \end{aligned}$$

The variable d_ℓ^{gen} is an index of density. This formulation has the following proprieties. If there is only a single store, then the index at the location is $d_\ell^{gen} = 0$. If there are an infinity of stores, the index equals $d_\ell^{gen} = 1$. I define density economies for groceries in the same way with a density index d^{groc} and a parameter ϕ^{groc} governing the magnitude of density benefits.

I take a different approach to modeling savings from distribution center density. Typically, a store will deal with one regional distribution center (RDC) for general merchandise, and if is a supercenter, it will deal with a single food distribution center (FDC). Let y_ℓ^{DC} and y_ℓ^{FDC} be the distance of location ℓ to the closest RDC and FDC Then suppose spillover for general merchandise is simply

$$DensityEconomy_\ell^{RDC} = \phi^{RDC} d_\ell^{RDC}.$$

for

$$d_\ell^{RDC} = -y_\ell^{RDC}.$$

So reducing the distance by one mile to the closest RDC increases profit at a store by ϕ^{RDC} dollars. Density economies for FDCs are defined in an analogous way.

2.2.4 Dynamics

Everything that has been discussed so far considers quantities for a particular time period. I now explain the dynamic aspects of the model. I assume Wal-Mart operates in a deterministic environment in discrete time and that Wal-Mart has perfect foresight. The general problem Wal-Mart faces is to determine for each period:

1. How many new Wal-Marts and how many new supercenters to open?
2. Where to put the new Wal-Marts and supercenters?
3. How many new distribution centers to open?
4. Where to put the new distribution centers?

In what follows I just focus on part 2 of Wal-Mart's problem. I condition the answers to 1, 2, and 4, on what Wal-Mart actually did, and solve Wal-Mart's problem of getting 2 right. Of course, if Wal-Mart's actual behavior solves the true problem of choosing 1 through 4, then it also solves the constrained problem of choosing 2, condition on 1, 3, and 4 being what Wal-Mart actually did.

Getting at part 1 of Wal-Mart's problem—how many new stores Wal-Mart opens in a given year—is far afield from the main issues of this paper. In its first few years, Wal-Mart added only one or two stores a year. The number of new store openings has grown substantially over time and in recent years they sometimes number several stores in one week. Presumably capital market considerations have played an important role here. This is an interesting issue, but not one I will have anything to say about with this project. A similar point can be made about part 3 of Wal-Mart problem concerning the number of distribution centers.

Getting at 4—where to put distribution centers—is closely related to the main issue of this paper. A useful avenue of future result would be to jointly study part 2 and part 4. For this paper, I condition on the choice of part 4.

Now for notation. Let B_t^{wal} be the set of Wal-Mart stores in period t. Assume that once a store is opened, it never shuts down. This assumption simplifies the analysis considerably and is not inconsistent with Wal-Mart's behavior (it rarely closes stores). Then we can write $B_t^{wal} = B_{t-1}^{wal} + A_t^{wal}$, where A_t^{wal} is the set of new stores opened in period t. Analogously, $B_t^{super} \subseteq B_t^{wal}$ is the set of supercenters at t. A supercenter is also an absorbing state, $B_t^{super} = B_{t-1}^{super} + A_t^{super}$, for A_t^{super} being new supercenter openings in period t. Note a supercenter can open two ways. It can be a new Wal-Mart store that opens as a supercenter as well. Or it can be a conversion of an existing Wal-Mart store. Finally, I take as fixed the openings of distribution centers.

Let N_t^{wal} and N_t^{super} be the number of new Wal-Marts and supercenters opened at t, i.e. the cardinality of the sets A_t^{wal} and A_t^{super}. Choosing these values was defined as part 1 of Wal-Mart's problem. These are taken as given here.

I allow for exogenous productivity growth of Wal-Mart equal to a growth factor ρ_t in period t. What I mean by this is that if Wal-Mart where to hold fixed the set of stores and demographics also stayed the same, from period $t-1$ to period t, then revenue and all components of costs would grow at a factor ρ, i.e.

$$\begin{aligned} R_{j,t} &= \rho R_{j,t-1} \\ C_{j,t} &= \rho C_{j,t-1}. \end{aligned}$$

This means the profit grows by a factor ρ_t, holding fixed the set of Wal-Mart's stores and holding fixed demographics. 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. The part of this growth that is due to expansion to supercenters I explicitly account for. But the other part I do not model and take the process as occurring exogenously.

Let β be the discount factor. Let $a = (A_1^{wal}, A_1^{super}, A_2^{wal}, A_2^{super}, ..., A_T^{wal}, A_T^{super})$ be a vector specifying the new stores opened in each period t. Require this vector to be feasible so that the number of new openings in a given period is N_t^{wal} and N_t^{super}. Wal-Mart's problem at time t is

$$\max_{\alpha} \sum_{t=1}^{T} (\rho_t \beta)^{t-1} \begin{bmatrix} \sum_{j \in B_t^{wal}} \left[\pi_{jt}^{gen} - f_{jt}^{gen} + \phi^{gen} d_{jt}^{gen} + \phi^{RDC} d_{jt}^{RDC}\right] \\ + \sum_{j \in B_t^{super}} \left[\pi_{jt}^{groc} - f_{jt}^{groc} + \phi^{groc} s_{jt}^{groc} + \phi^{FDC} d_{jt}^{groc}\right] \end{bmatrix}. \tag{5}$$

for operating profit for merchandise line $k \in \{gen, groc\}$,

$$\pi_{jt}^{k} = \mu R_{jt}^{k} - w_{jt} Labor_{jt}^{k} - r_{jt} Land_{jt}^{k} - Other_{jt}^{k}.$$

3. The Data

This section begins by explaining the basic data sources. It then discusses some facts about Wal-Mart's 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 opening dates for stores, supercenters, general distribution centers and food distribution centers. The third element is demographic information from the Census. The four element is data on wages and rents by location.

Data element one, store-level data variables such as sales, was obtained from *TradeDimensions*, a unit of ACNeilsen. This data provides estimates of store-level sales for all Wal-Marts open as of the end of 2005. This data is the best available and is the primary source of market share data used in the retail industry. Ellickson (2007) is a recent user of this data for the supermarket industry.

Table 1 presents summary statistics of annual store-level sales and employment for the 3,176 Wal-Marts in existence in the contiguous part of the United States as the end of 2005.[3] (Alaska and Hawaii are excluded in all of the analysis.) Almost two thirds of all Wal-Marts (1,980 out of 3,176) are supercenters. The average Wal-Mart racks up annual sales of \$70 million. The breakdown is \$47 million per regular store and \$85 million per supercenter. The average employment is 255 full time equivalent employees.

[3] 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.

The second date element is opening dates of Wal-Mart facilities. Exact opening dates for Wal-Marts was obtained from a file posted on the web by Wal-Mart but subsequently withdrawn. Opening dates for supercenters was compiled by using data provided by Emek Basker for years before 1999 and for later years by store announcements posted on Wal-Mart's web site. Opening dates for regional distribution centers and food distribution centers was compiled from various sources such as Lexis-Nexis and the web. Table 2 tabulates opening dates for the four types of facilities by decade.

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 each 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.[4]

The fourth data element is information about local wages and rents. The wage measure is average retail wage by county from County Business Patterns. This is payroll divided by employment. I use annual data over the period 1977 to 2004. Measuring land rents is difficult. I proxy land rents using information about residential property values from the decennial census. For each Census year, I take each store location, I create an index of property value by adding up the total value of residential property within two miles of each store location. I have supplemented this information with data on property values of Wal-Mart properties for Iowa and Minnesota obtained from the web. As discussed in the (future) appendix, there is a high correlation of this index with land values of Wal-Mart stores in these states.

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

The fifth data element is information from Wal-Mart's annual reports including information about aggregate sales for earlier years. (The TradeDimensions provides only current data.) I also use information provided in the "Management Discussion" section of the report. In the 2004 Wal-Mart Annual Report, the following information was reported: "As we continue to add new stores in the United States, we do so with an understanding that additional stores may take sales away from existing units. We estimate that comparative store sales in fiscal year 2004, 2003, 2002 were negatively impacted by the opening of new stores by approximately 1%." This same paragraph was repeated in the 2006 annual report in regards to fiscal year 2005 and 2006. I use this information when estimating the demand model.

4. Estimates of Operating Profits

Section estimates various components of Wal-Mart's operating profits. The main work of the section, which is presented in Part 1, is to estimate the model of Wal-Mart's demand. Part 2 treats various other cost parameters. Part 3 discusses extrapolating 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 general merchandise sales $\hat{R}_j^{gen}(\theta)$ for each store j from equation (4) and predicted values of grocery sales $\hat{R}_j^{groc}(\theta)$.

The data has all commodity sales volume for each store. Call this R_j. I define general merchandise to include all items sold at a regular Wal-Mart. (So cases of Coke and Pepsi sold at regular Wal-Marts are considered general merchandise.) For regular stores, $R_j = R_j^{gen}$ by definition. For supercenters, I observe the sum of general merchandise and groceries, $R_j = R_j^{gen} + R_j^{groc}$.

Let η_j be the difference between log actual sales and log predicted sales for store j. For regular stores this is

$$\eta_j^{wal} = \ln(R_j) - \ln(\hat{R}ev_j^{gen}(\theta)).$$

For supercenters, this is

$$\eta_j^{super} = \ln(R_j) - \ln(\hat{R}_j^{gen}(\theta) + \hat{R}_j^{groc}(\theta)).$$

Assume the discrepancies η_j^{wal} and η_j^{super} are i.i.d. normally distributed measurement error.

To estimate the model, we need functional forms for $o(m)$ and $\tau(m)$. Assume

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

for

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

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

Two sets of parameter estimates are reported in Table 4. The first set, labeled MLE, is obtained from standard maximum likelihood. The second set imposes a constraint on the demand parameters that the statement about cannibalization in the annual report be exactly true. Recall that beginning with fiscal year 2002, Wal-Mart began reporting that the *cannibalization rate*—sales that existing stores would have received but did not because of entry of new stores—was approximately 1 percent. With the parameter estimates of the model and with the information about entry of new stores in each year, it is possible to calculate the cannibalization rate in the model and we do this in Table 5. The estimates are approximately 1, just as reported. If we round to one digit, it matches perfectly, except for 2005 where it is just over 1.5 so we round up to 2.

The main goal of this paper will be to obtain a lower bound on the importance of density economies. For this purpose, it is essential that I do not overstate the importance of cannibalization. For this reason, I estimate a constrained version of the model were I require the cannibalization rate for 2006 to exactly be one. The estimates are very similar. the cannibalization rates are recalculated for each of the earlier years. The rates for the constrained model are roughly two thirds of the unconstrained model. The constrained model will be the benchmark model for the paper.

A few remarks about the parameter estimates. The estimates of λ^{gen} and λ^{groc} are both about 1.7 in thousands of dollars per year. This is the spending budget per consumer that is allocated between the outside good and Wal-Mart. To get a sense of these numbers it is useful to report related Census figures. For 2005, per capita spending in the U.S. in general merchandise stores (NAICS 452) was 1.77 thousands of dollars and in food and beverage stores (NAICS 445) 1.75 thousands of dollars. It is remarkable that these numbers line up like this but I should say I wouldn't expect these numbers to be exactly the same as λ^{gen} and λ^{groc}. On one hand, the general merchandise category includes Saks Fifth Avenue which is not likely to be in the same spending budget with a Wal-Mart. One the other hand, the general merchandise category does not include the electronics giant Best Buy; a large portion of this merchandise would be in the same spending budget with Wal-Mart. Both of these categories are relatively small (electronics is less than a fifth of general merchandise sales) so perhaps it is not a surprise that my estimate of λ^{gen} is so close to U.S. per capita spending in this store category.

The remaining parameters in Table 4 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. The table reports the probability that a consumer shops at the Wal-Mart for his or her general merchandise. 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 50 this falls to .72 and at 250 it falls to only .22. 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 less than a quarter as large, but overall demand is almost 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. Raising the distance further from 5 to 10 miles does have an appreciable effect, .988 to .711. 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 .988 to .711 seems highly plausible. Going to 15 miles out, the probability drops substantially to .068 and at 20 miles is essentially zero.

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.

Demand varies by demographic characteristics in interesting ways. Wal-Mart is an inferior good in that demand decreases in income. Note the coefficient on the per capita income on the outside good is positive; this is makes Wal-Mart an inferior good. Demand is higher among whites and lower among younger people and older people.

Finally, the only store characteristic used in the model (besides location) is store ages. There is a dummy variable for stores at least two years old. Older stores have higher demand.

4.2 Labor Costs

I assume constant returns to scale. I assume the labor requirements for general merchandise sales are the same as for groceries. In my 2005 TradeDimensions data, on average there are 3.61 store employees per million dollars of annual sales. So I set $\nu_{Labor} = 3.61$.

I use property value information for selected Wal-Marts in Minnesota and Iowa along with the rent index described above to estimate land rends for each Wal-Mart location.

4.3 Other Costs

I need an estimate of variable profit per unit sales excluding variable labor costs. We begin with the gross margin, the percent of the price that is markup over the cost of goods sold. Wal-Mart's gross margin over the years has ranged from .22 to .26 (from Wal-Marts annual reports.), so if I set the gross margin to .24 that is a sensible place to start. I take out another .07 for nonlabor variable costs (more detail to come). This gives me variable profit (excluding variable labor costs) per dollar of sale equal to

$$\mu = .17.$$

4.4 Extrapolation to Other Years

So far I have constructed a model of Wal-Marts demand and cost's circa 2005, 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 2005, average store sales (regular stores)was \$47.5 million. In 1972, average sales (in 2005 dollars) was only \$11.1. How can I take this into account?

I applied the following procedure. First, I took the demand model from 2005 and evaluated average sales per store in the prior years, for the actual the configuration of stores for each of these prior years. The 2005 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 47.5 in 2005 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 2005 that would generate the sales difference of 11.1 to \$***. 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. Evidence on Diminishing Returns from Cannibalization of Existing Sales

In this section I evaluate whether Wal-Mart encountered diminishing returns as it concentrated stores in the same general area.

I proceed at the state level aggregating some of the smaller states. (For the purposes of this analysis, the New England states are treated as a single state. Maryland, Delaware and the District of Columbia are also aggregated.) For each state, I determine the opening date of the first Wal-Mart in the state as well as the opening state of the first supercenter. I then categorize Wal-Marts by the within-state age categories listed in Table 7. The first category, 1-2, are Wal-Marts opened in year one and two that Wal-Mart is in the state. The second category, 3-5, are Wal-Marts opened in year three to five, and so on. I categorize supercenters in the analogous way.

I define the incremental sales of a Wal-Mart j at its opening date as follows. I determine the set of stores open as of the exact opening date (to the day) of store j including store j. I calculate total Wal-Mart sales across all stores in this set. I use the 2005 demand model with no productivity adjustment. So this is what annual sales would be if we had store set as of the opening date of j but had the productivity term of 2005. Next I determine what sales are without store j and take the differences. This is 2005-demand-equivalent incremental sales. Using the estimates of labor requirements and such, I convert this to incremental operating profit (again a 2005 equivalent). I then calculate mean incremental sales and operating profit by within-state age category.

Table 7 shows while there are no diminishing returns within the first five years, they set in after five years and become substantial beyond ten. In the 11-15 year category, incremental sales average almost 3 million less (almost 10 percent) than in the first five years in the state. This translates into a half million less in operating profit, *per store*. For stores in the 21 and over category, the differences are very large, 8 million in sales and over 1 million in operating profit. Part B looks at supercenters and displays similar magnitudes of diminishing returns.

Some indication about the importance of cannibalization can be obtained from the *stand-alone operating profit* of each store. This would be a store's operating profit if it were the only Wal-Mart store. Incremental operating profit is bounded above by the stand-alone profit. The stand-alone operating profit for the 21 and above category is almost the same as for the 1-2 age category. So the one million dollar difference in operating profit is cannibalization.

One issue is that there may be other cost differences across store locations that are not being accounted for in the simple crosstabulation in Table 7. In particular, as argued earlier, I expect fixed costs to be higher in higher population areas. Table 8 runs a regression to control for population density. I use a quadratic in logs. I also include state fixed effects. The idea is to hold fixed state and population and determine how incremental profit varies in a within-state age category. In the regression, within-state age 1-2 is the excluded category. Adding these controls makes little difference. For example, the difference between the 11-15 group and the 1-2 group is .63 in the regression and 3.55 - 2.95 = .60 in the raw data. For 16-20, the analogous differences are .76 and .69. The differences in the regression are highly statistically significant.

Obviously, initial store density is higher for stores that open later within a state. Table 8 reports means of the incremental density index at the year of opening. For the calculations, I assume $\alpha = .02$ (more on this later). I use (**??**) to calculate the density index for each store before and after a particular store j opens. Even for stores that are new in a state, the mean incremental spillover is .8, well above zero. This happens because the new stores in a state are getting density benefits from stores in adjacent states.

Define the incremental distribution center distance for store j to be the distance of store j from the closest distribution center at the date of store j's opening. On average, the first store in a state is quite far from a distribution center, 343 miles. As we move up the age category, there is a substantial decrease. Stores that open in a state where Wal-Mart has been there for 20 years or more are, at opening, within 90.1 miles of a distribution center. The same pattern occurs with supercenters.

6. Bounding Density Economies

In this section I use the information in Wal-Mart choice behavior to obtain bounds on Wal-Mart's cost parameters.

6.1 Method

The α parameter governs the shape of the spillover function while the ϕ parameter determines the weight placed on spillover. There is something of a tradeoff between the α parameter and the ϕ parameter in making density matter in Wal-Mart's behavior. My approach is to fix α and then estimate ϕ. Then I discuss how my choice of α matters in answering questions about the overall importance of density economies to Wal-Mart. For the discussion here, $\alpha = .02$ is assumed throughout.

I parameterize the fixed cost function to depend upon population density. This is motivated by my earlier discussion. Suppose

$$f(m) = \omega_0 + \omega_1 \ln(m) + \omega_2 \ln(m)^2.$$

I assume, a priori, that $\omega_1 \geq 0$ and $\omega_2 \leq 0$ so the relationship between fixed cost and density is weakly increasing and weakly concave. The parameter ω_0 makes no difference in the analysis so henceforth I normalize it to $\omega_0 = 0$.

The premise of this analysis is that sales are measured exactly with the estimate model but that there is measurement error in the wage as well as rent. Let $\varepsilon^{wage}_{j,t}$ and $\varepsilon^{rent}_{j,t}$ be the measurement error for these two variables. Then actual operating profit of store j in time t given openings at that time is

$$\pi^{gen}_{j,t} = (\mu - \nu_{other}) R^{gen}_{j,t} - w_{jt}\nu_{labor} R^{gen}_{j,t} - r_{jt}\nu_{land} R^{gen}_{jt}.$$

The observed operating profit is

$$\tilde{\pi}^{gen}_{j,t} = (\mu - \nu_{other}) R^{gen}_{j,t} - \left(w_{jt} + \varepsilon^{wage}_{jt}\right) \nu_{labor} R^{gen}_{j,t} - \left(r_{jt} + \varepsilon^{rent}_{jt}\right) \nu_{land} R^{gen}_{jt}$$

and the difference is

$$-\varepsilon_{jt}^{wage}\nu_{labor}R_{j,t}^{gen}-\varepsilon_{jt}^{rent}\nu_{land}R_{j,t}^{gen}. \tag{6}$$

Assume the measurement error is i.i.d. across store locations and has expected value zero.

As above, let a denote a particular choice of Wal-Mart, a particular feasible solution to problem (5). Let a_0 denote the choice Wal-Mart actually made. For each policy a, using the estimates above let $\Pi^{gen}(a)$ be estimated present value of operating profits at date $t=1$ under this policy. Let $d^{gen}(a)$ be the present discounted value at the initial date $t=1$ of the density index aggregated across all store in each period. Finally, let $F_1^k(a)$ be the present value of linear component of fixed cost and $F_2^k(a)$ be the present value of the quadratic component of fixed cost under a, $k\in\{gen,groc\}$. Analogously, define $d^{groc}(a)$, $d^{RDC}(a)$ and $d^{FDC}(a)$. Let $v(a,\theta)$ be the total discounted present value given action a and parameter vector $\theta=\left\{\phi^{gen},\phi^{groc},\phi^{RDC},\phi^{FDC},\omega_1,\omega_2,\zeta\right\}$

$$\begin{aligned}
& v(a,\theta)\\
= \; & \Pi^{gen}(a)+\phi^{gen}d^{gen}(a)+\phi^{RDC}d^{RDC}(a)-\omega_1F_1(a)-\omega_2F_2(a)\\
& +\Pi^{groc}(a)+\phi^{groc}d^{groc}(a)+\phi^{FDC}d^{FDC}(a)-\zeta_\omega\left(\omega_1F_1(a)-\omega_2F_2(a)\right)
\end{aligned}$$

Now the chosen policy a_0 solves problem (5). Hence at the true parameter θ,

$$v(a_0,\theta)\geq v(a,\theta),\text{ for all } a\neq a_0$$

Or

$$\Delta v(a,\theta)\geq 0,$$

for

$$\Delta v(a,\theta)=v(a_0,\theta)-v(a,\theta).$$

Given an alternative policy a and a parameter vector θ, we observe

$$\Delta\tilde{v}(a,\theta)=\Delta v(a,\theta)+\varepsilon_a,$$

Where a is the present value of the difference in measurement error (6) between the actual policy and the alternative a. Now $E[\varepsilon_a]=0$.

Like Bajari and Fox (2005) and Fox (2005), I consider pairwise resequencing. I consider only deviations a in which I reorder a pair of stores. For example, store #1 actually opened 1962 and #2 opened 1964. I consider the deviation where store #2 opens in 1962, store #1 in 1964, everything else the same. Let A^{pair} be the set of pairwise deviations. Consider two different alternatives a and a'. If there is no overlap in the two stores resequenced for a and a', then $E[\varepsilon_a \varepsilon_{a'}] = 0$. As there are a large number of stores, the likelihood of overlap is small.

I follow recent work on partially identified sets (Manksi (2002)) and construct moment inequalities (Pakes, Porter, Ho, Ishii (PPHI)) in which the measurement error is averaged out. Let i index subsets of A^{pair} that will be defined in way that is independent of the measurement error. Suppose there are K instruments indexed by k where the instrument $z_k(a)$ is nonnegative and uncorrelated with the measurement error. For each i and k define

$$m_{ik}(\theta) = E\left[\Delta\tilde{v}(a,\theta) z_k(a)\right] \text{ for } a \in A_i^{pair}.$$

At the true θ,

$$m_{ik}(\theta) \geq 0, \text{ for all } i \text{ and } k. \tag{7}$$

Three classes of subsets are constructed.

1. Start with the set of stores opened 10 or more years after the first store in their state. For each such store j, find the set of stores, indexed by j', such that store j' opens three or more years after store j in a different state. Furthermore, require that store j' be opened within four years of the first store in j''s state. This is set A_1^{pair}. Below I will call these *farther sooner* deviations.

2. Start with the list of stores opened within five years of the first store in the state. For each such store j, find the set of stores, indexed by j', such that store j' opens three or more years after store j in a different state. Furthermore, require that store j' be opened 10 or more years after the first store in j''s state. Finally, require that the first store in j''s state is before the first store in j's state. Flipping the opening order, this is set A_2^{pair}

3. Define two population density groupings for store locations. For example, let group 1 be locations with less than 15,000 in population within five miles and let group 2 be locations with 15,000 to 40,000. Take store locations in the same state opening at different dates where one location is in density group 1 and the other is in density group 2. Flip the order of store openings. Do this for various different pairs of density groupings.

The purpose of set 1 is to provide information about a lower bound on the density economy parameters ϕ^{gen} and ϕ^{groc}. In the deviation, instead of adding yet another store in a state where Wal-Mart has been for over 10 years, I open early a store that would have been one of the early stores in a another state. The alternative location is one which has not yet been hit with diminishing returns. Set 2 is the opposite. The third category defines pairwise perturbations based on population density that are intended to provide information about the parameters ω_1 and ω_2 that govern how the fixed cost varies with population density.

Given the subsets A_i^{pair}, I take further subsets based on year founded and based on whether the deviation involves a new Wal-Mart store opening or a conversion of a Wal-Mart store into a supercenter.

The following variables are used for the instruments z_k. (i) A vector of ones. (ii) ΔF_1, ΔF_2, $-\Delta F_1$ and $-\Delta F_2$ plus constants so all are nonnegative. (iii) $\Delta d^{gen}, -\Delta d^{gen}$, Δd^{groc}, and $-\Delta d^{groc}$ plus constants so all are nonnegative. The instruments must be nonnegative to preserve the inequalities.

Let $\hat{\Theta}$ be the set of θ that satisfies the above for all i and k in the sample. I use standard linear programming techniques to characterize the set $\hat{\Theta}$. To simplify, I assume

$$\begin{aligned} \phi^{groc} &= \zeta_\phi \phi^{gen} \\ \phi^{FDC} &= \zeta_\phi \phi^{RDC}. \end{aligned}$$

Fix a particular ϕ^{RDC}. I determine bounds on ϕ^{gen}. For a given ζ_ϕ and ζ_ω, this is a standard linear program. I do a grid search over ϕ^{RDC}, ζ_ϕ and ζ_ω and for each use linear programming to solve the problem of minimizing ϕ^{gen} subject to the constraints.

6.2 Estimates

Consider the set of *farther-sooner* deviations defined above by A_1^{pair}. In each of these, Wal-Mart is delaying the opening of a store in it existing network and opening sooner a store farther out from its network. Given the parameters defining the deviation (e.g. with start with the set of stores opened 10 years after the first store in the state, etc.), there are 239,698 such deviations Wal-Mart could have considered involving reordering of store opening dates and 5,110 involving reordering of supercenter conversion dates. (There are less for conversions because supercenters have not been around as long.) Calculating the sales model is time consuming so I take a random sample to estimate the means. Table 9 show the sample means for the two types of deviations. For example, the mean difference in the present value of profit from doing the actual policy relative to the deviation is -1.28 million for the Wal-Mart store openings (the row labeled "General"). The table listed the mean value for all the variable that enter into $\Delta v(a, \theta)$.

By doing the actual policy, Wal-Mart lost, on average per deviation (i.e. per store), 1.28 million dollars relative to what it could have achieved in operating profit from the deviations. But by doing the actual policy instead of the it gained in the density measures. It gained .82 in the d^{gen} index of density for general merchandise and it gained 5.90 in the d^{RDC} density measure, where the units here are in hundreds of present value year-miles. The table also shows the difference in the present value of the log population density terms used in the fixed cost.

I begin the discussion of the estimated bounds with a simple case. Suppose we zero the dependence of the fixed cost on density, $\omega_1 = \omega_2 = 0$. And assume that $\phi^{gen} = \phi^{groc}$ and $\phi^{RDC} = \phi^{FDC}$. Consider the subset of the farther-sooner deviations where we change Wal-Mart store openings. Taking the means from Table 9 and following (7), it must be that

$$mean_{a \in A_1^{pair}}(\Delta v) = -1.28 + .82\phi^{gen} + 5.9\phi^{RDC} \geq 0$$

Now if $\phi^{RDC} = 0$ where to hold, this moment inequality implies that

$$\phi^{gen} \geq \frac{1.28}{.82} = 1.56.$$

If we increase ϕ^{RDC}, the bound on ϕ^{gen} decreases. The first row in Table 10 reports the bound on ϕ^{gen} from this moment given ϕ^{RDC}. The bound on ϕ^{gen} goes to zero at approximately $\phi^{RDC} = .22$.

I will discuss the interpretation of the units of ϕ^{gen}. But the interpretation of $\phi^{RDC} = .22$ is immediate. At this level, the cost savings is \$220,000 from having the regional distribution center closer to a given store by 100 miles and everything else the same. Ballpark figures for what trucking services charge to deliver trailers is on the order of \$2 a mile or \$800 for a 200 mile round trip. So the \$220,000 figure is on the order of \$200,000/800=275 trips a year between the distribution center and the store. Just to put this figure in perspective, suppose $\phi^{gen} = \phi^{groc} = .22$ every store was moved 100 miles from the closest regional distribution center and food distribution center. Then we multiply by 5,000 (3000 stores, 2000 of the also selling groceries) to get 1.2 billion, which is 10 percent of Wal-Mart's 2006 net income.

If we do the same exercise with conversion deviations, the lower bound is not as tight, it is on the order of one third as high.

The discussion so far is meant to be illustrative. It assumes that fixed costs do not depend upon population density, $\omega_1 = \omega_2 = 0$, but this is inconsistent with prior knowledge about Wal-Mart choice behavior. Furthermore, it does not use the information contained in additional moments.

The second set of estimates in Table 10 use the *basic moments*. For these the only instrument is the vector of ones; I do not include the interactions. For these estimates I fix ϕ^{RDC} and take a grid over $\zeta_\phi \in [0,1]$ and $\zeta_\omega \in [0,1]$ and then minimize ϕ^{gen} subject to the constraints. It turned out that the minima were obtained at $\zeta_\phi = 1$ and $\zeta_\omega = 1$ so the density parameters and the fixed cost parameters to produce the minimum are the same for general merchandise and groceries.

In freeing up the parameters on the fixed cost, the estimated lower bound on ϕ^{gen} decreases. Why this happens can be seen in Table 9. Note that for the general perturbations, the mean of $\Delta F_1^{gen} = -.74$. This is saying that on average the stores opened earlier in the actual path are rural areas compared to the deviation. From the other perturbations we determine that but some weight on this in fixed cost savings. But then we can push down the weight on the density benefits. Like the first estimates, we see a tradeoff here between ϕ^{gen} and ϕ^{RDC}. Of we put no weight on being close to distribution centers we have to put some weight on being close to stores. If we want to put no weight on being close to store, we have to put some weight on being close to the distribution centers.

Note that we also condition on the opening year, separating out the period 1988-2006 when Wal-Mart starting having supercenters from the earlier period. The bound is significantly tighter in the later period.

The last set of estimates use the full set of interactions. As can be seen from just using the farther-sooner moment, one can't distinguish between store density benefits and distribution center density benefits. But through use of the information in the interactions (e.g. multiplying through by Δd^{gen} plus a constant, Δd^{RDC} plus a constant, etc.) we can say more. Adding these additional moments has two qualitative impacts. First, the estimated bounds increase. They cannot go down, of course, since more constraints are being added.

Second and more interestingly, there is no longer a tradeoff between ϕ^{gen} and ϕ^{RDC}. I conclude from this that distance to the regional distribution center and the food distribution center is not the only thing that matters for density economies. There are advantages besides these two factors from having stores close. On explanation of these other benefits is that Wal-Mart's get deliveries from other sources besides these two distribution channels. First, Wal-Mart has other kinds of distribution (e.g. clothes come to the store a different way a different way). Plus there are direct store deliveries from manufacturers that Wal-Mart has negotiated with Proctor and Gamble. Wal-Mart reports that a typical supercenter receives 85 different delivery trucks a week. In addition to these logistic benefits of a dense store network, there are other benefits mentioned in the introduction.

The preferred bound uses the full set of instruments. The bound for the supercenter era is tightest, so I focus on this one, $\phi^{gen} \geq .85$. To interpret this number, consider figure 11. This lists the average store density index d^{gen} for selected states for Wal-Mart as of 2006. The lowest is in North Dakota at .50. The highest is New Jersey, at .978. California is at .945. With $\phi^{gen} \geq .85$, the difference in density economies between New Jersey and California has a lower bound of .85*(.978-.945)=.0281 or $28,100 per store per year (double that if a supercenter). The difference between New Jersey and Washington State is $85,000 per store, or $570 per store employee.

References

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., "Does Sutton Apply to Supermarkets?" Forthcoming in the Rand Journal of Economics. Spring, 2007

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*

Jia, Panle, "What Happens When Wal-Mart Comes to Town: An Empirical Analysis of the Discount Retailing Industry" Yale University Manuscript, 2005

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

Manski, C., (1989), "Anatomy of the Selection Problem," *Journal of Human Resources* 24, 343-60.

Neumark, David, Junfu Zhang, and Stephen M. Ciccarella, Jr. "The Effects of Wal-Mart on Local Labor Markets" NBER Working Paper No. W11782 November 2005

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
2005 (excluding Alaska and Hawii)

Store Type	Variable	N	Mean	Std. Dev	Min	Max
All	Sales ($millions/year)	3,176	70.5	30.0	9.1	166.4
Regular Store	Sales ($millions/year)	1,196	47.0	20.0	9.1	133.9
SuperCenter	Sales ($millions/year)	1,980	84.7	25.9	20.8	166.4
All	Employment	3,176	254.9	127.3	31.0	812.0
Regular Store	Employment	1,196	123.5	40.1	57.0	410.0
SuperCenter	Employment	1,980	333.8	91.5	31.0	812.0

Table 2
Distribution of Wal-Mart Facility Opens by Decade and Opening Type

	Wal-Marts		Supercenters		Regional Distribution Centers		Food Distribution Centers	
Decade Open	Opened this decade	Cumlative	Opened this decade	Cumlative	Opened this decade	Cumlative	Opened this decade	Cumlative
1962-1969	15	15	0	0	1	1	0	0
1970-1979	243	258	0	0	1	2	0	0
1980-1989	1,082	1,340	4	4	8	10	0	0
1990-1999	1,130	2,470	679	683	18	28	9	9
2000-2005	706	3,176	1,297	1,980	14	42	25	34

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
Parameter Estimates for Demand Model

Parameter	Definition	MLE Model	Constrained MLE Model
λ^{gen}	General Merchandise Spending per person (annual in $1,000)	1.731	1.736
		(.008)	(.070)
λ^{groc}	Grocery spending per person (annual in $1,000)	1.697	1.698
		(.004)	(.076)
ρ	correlation parameter	.962	1.065
		(.026)	(.023)
τ_0	constant	.626	.702
		(.034)	(.036)
τ_1	population density within 5 miles	-.046	-.053
		(.007)	(.007)
ω	Outside good valuation parameters		
	constant	-7.896	-8.675
		(.374)	(.610)
	ln(mbar)	1.859	2.095
		(.102)	(.168)
	ln(mbar)^2	-.062	-.078
		(.008)	(.015)
	Per Capita Income	.015	.014
		(.003)	(.003)
	Share of block group black	.341	.320
		(.077)	(.082)
	Share of block group young	1.090	1.147
		(.420)	(.479)
	Share of block group old	.580	.475
		(.335)	(.389)
γ	Store-specific parameters		
	store age 2+ dummy	.177	.205
		(.023)	(.024)
σ^2	measurement error	.065	.065
		(.002)	(.002)
N		3176	3176
SSE		205.030	206.065
R^2		.755	0.754
ln (L)		-155.081	-163.074

Table 5
Cannibalization Rates
From Model and As Reported in Annual Reports

Fiscal Year	From Annual Reports	MLE Model	Constrained MLE Model
1999	n.a.	0.69	0.44
2000	n.a.	0.95	0.65
2001	n.a.	0.61	0.37
2002	1	0.73	0.49
2003	1	1.41	0.93
2004	1	1.48	1.06
2005	1	1.55	1.10
2006	1	1.35	1.00*

*Cannibalization Rate is imposed to equal 1.00.

Table 6
Comparative Statics with Demand Model

Distance	Population Density (thousands of people within a 5 mile radius)						
(miles)	1	5	10	20	50	100	250
0	1.000	.991	.971	.912	.715	.482	.220
1	.999	.984	.949	.858	.605	.371	.158
2	.999	.971	.913	.779	.483	.272	.111
3	.997	.948	.854	.671	.363	.191	.076
4	.994	.907	.767	.543	.258	.130	.052
5	.988	.841	.648	.408	.175	.086	.035
10	.711	.195	.092	.044	.018	.009	.005
15	.068	.011	.006	.003	.002	.001	.001
20	.002	.001	.000	.000	.000	.000	.000
25	.000	.000	.000	.000	.000	.000	.000

Table 7
(All evaluated at 2005 Demand Equivalents)

Part A: General Merchandise (New Wal-Marts including supercenters)

Within-State Age	N	Incremental Sales	Incremental Operating Profit	Stand-alone Operating Profit	Incremental Store Density Index	Incremental Distribution Center Distance
1-2	288	38.35	3.55	3.62	0.82	343.26
3-5	614	39.98	3.55	3.70	0.96	202.04
6-10	939	38.04	3.39	3.64	0.98	160.68
11-15	642	36.75	2.95	3.36	0.99	142.10
16-20	383	33.48	2.86	3.47	1.00	113.66
21 and above	310	29.95	2.44	3.56	1.00	90.19

Part B: Groceries (New supercenters)

Within-State Age	N	Incremental Sales	Incremental Operating Profit	Stand-alone Operating Profit	Incremental Supercenter Density Index	Incremental Distribution Center Distance
1-2	202	42.30	3.86	3.93	0.73	252.90
3-5	484	42.71	3.97	4.13	0.93	171.17
6-10	775	41.00	3.63	3.97	0.99	113.52
11-15	452	36.70	3.19	3.84	1.00	95.32
16-20	67	29.69	2.71	3.42	1.00	93.95

Table 8
Incremental Operating Profit Regression
2005 Demand Equivalents
Includes State Fixed Effect

	General Merchandise	Groceries
Within-State Age Category		
3-5	-0.04	-0.11
	(.05)	(.06)
6-10	-0.31	-0.62
	(.05)	(.07)
11-15	-0.63	-1.13
	(.06)	(.08)
16-20	-0.76	-1.40
	(.06)	(.12)
21 plus	-1.33	
	(.07)	
log population density	5.80	6.23
	(.19)	(.31)
(log population density)^2	-0.26	-0.27
	(.01)	(.01)
R^2	.52	.50
N	3176	1986

Table 9
Farther Sooner Deviations
Weighted Mean Changes
(All years)

Facility Perturbations	Number of Deviations	Sample Size	$\Delta\Pi^{gen}$ ($million)	Δd^{gen}	Δd^{RDC} (100s of year miles)	ΔF_1^{gen}	ΔF_2^{gen}	$\Delta\Pi^{groc}$ ($million)	Δd^{groc}	Δd^{FDC} (100s of year miles)	ΔF_1^{groc}	ΔF_2^{groc}
General	239,698	15,000	-1.28	0.82	5.90	-0.74	-4.63	0.00	0.00	0.00	0.00	0.00
Grocery	5,110	3,625	0.00	0.00	0.00	0.00	0.00	-0.13	0.23	1.17	-0.01	0.07

Table 10
Estimates of Lower Bound on φ^{gen}

Moments		Merchandise Included		Time Period	φ^{RDC}						
	Number	General	Grocery		0.00	.01	.02	.05	.10	.15	.20
Farther Sooner Deviations and $\omega_1 = 0$ and $\omega_2 = 0$	1	yes	no	All Years	1.56	1.50	1.42	1.20	.84	.48	.12
	1	no	yes	All Years	.57	.51	.46	.31	.07	.00	.00
Basic	16	yes	yes	All Years	.59	.52	.46	.25	.00	.00	.00
	16	yes	yes	1962-1988	.11	.04	.00	.00	.00	.00	.00
	8	yes	yes	1988-2006	.79	.73	.66	.46	.12	.00	.00
Basic plus Interactions	272	yes	yes	All Years	.66	.66	.66	.67	.68	1.20	3.16
	272	yes	yes	1962-1988	.36	.36	.36	.36	.36	.36	.35
	136	yes	yes	1988-2006	.85	.85	.86	.86	.87	1.20	3.17
	136	yes	no	1988-2006	.71	.71	.71	.70	.70	.70	.70
	136	no	yes	1988-2006	.28	.29	.29	.29	.30	1.20	3.17

Table 11
Mean Store Density Index
Selected States

State	Rank (lowest to highest)	Mean Store Density Index
ND	1	.50
MT	2	.53
WY	3	.66
SD	4	.74
ID	5	.78
WA	10	.879
CA	20	.945
AL	30	.964
DE	40	.973
NJ	50	.978