The Diffusion of Wal-Mart and Economies of Density

Thomas J. Holmes*†

January 23, 2008

Abstract

The roll-out of Wal-Mart store openings followed a pattern that radiated from the center out with Wal-Mart maintaining high store density and a contiguous store network all along the way. This paper estimates the benefits of such a strategy to Wal-Mart, focusing on the savings in distribution costs afforded by a dense network of stores. The paper takes a revealed preference approach, inferring the magnitude of density economies by the extent of sales cannibalization from closely-packed stores that Wal-Mart is willing to sustain to achieve density economies. The model is dynamic with rich geographic detail on the locations of stores and distribution centers. Given the enormous number of possible combinations of store-opening sequences, it is difficult to directly solve Wal-Mart's problem, making conventional approaches infeasible. The moment inequality approach is used instead and it works well. The estimates show the benefits to Wal-Mart of high store density are substantial and likely extend significantly beyond savings in trucking costs.

*University of Minnesota, Federal Reserve Bank, and the National Bureau of Economic Research.

†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. I am grateful to for NSF Grant 0551062 for support of this research. I have benefited from the comments of many seminar participants. In particular, I thank Glenn Ellison, Gautam Gowrisankaran and Avi Goldfarb for their comments as discussants and Ariel Pakes for advice on how to think about this problem. I thank Junichi Suzuki, Julia Thornton, David Molitor, and Ernest Berkas for research assistance. I thank Emek Basker for sharing data.

1 Introduction

Wal-Mart opened its first store in 1962 and today there are over 3,000 Wal-Marts in the United States. The roll-out of stores illustrated in Figure 1 displays a striking pattern. (See also a movie of the roll-out posted at the web.[1]) Wal-Mart started in a relatively central spot in the country (near Bentonville, Arkansas) and store openings radiated from the inside out. Wal-Mart never jumped to some far off location to later fill in the area in between. With the exception of store number one at the very beginning, Wal-Mart always placed new stores close to where they already had store density.

This process was repeated in 1988 when Wal-Mart introduced the supercenter format. (See Figure 2.) With this format, Wal-Mart added a full-line grocery store alongside the general merchandise of a traditional Wal-Mart. Again, the diffusion of the supercenter format began at the center and radiated from the inside out.

This paper estimates the benefits of such a strategy to Wal-Mart, focusing on the logistic benefits afforded by a dense network of stores. Wal-Mart is vertically integrated into distribution: general merchandise is supplied by Wal-Mart's own regional distribution centers, groceries for supercenters through its own food distribution centers. Over 80 percent of what Wal-Mart sells goes through its distribution center network (as opposed to direct store delivery by manufacturers or wholesalers).[2] When stores are packed close together, it is easier to set up a distribution network that keeps stores close to a distribution center. And when stores are close to a distribution center, Wal-Mart can save on trucking costs. Moreover, such proximity allows Wal-Mart to respond quickly to demand shocks. The ability of Wal-Mart to quickly respond to demand shocks is widely considered to be a key aspect of the Wal-Mart model. (See Holmes (2001) and Ghemawat, Mark, and Bradley (2004).) Wal-Mart famously was able to restock it shelves with American flags on the very day of 9/11.

A challenge in estimating the benefits of density is that Wal-Mart is notorious for being secretive—I am not going to get access to confidential data on its logistics costs. So it is not possible to conduct a direct analysis relating costs to density. And even if Wal-Mart were to cooperate and make its data available, the benefits of being able to quickly respond to demand shocks might be difficult to quantify with standard accounting data. Instead, I pursue an indirect approach that exploits revealed preference. While density has a benefit,

[1] A video of Wal-Mart's store openings can be seen at www.econ.umn.edu/~holmes/research.html

[2] See Ghemawat, Mark, and Bradley (2004).

it also has a cost, and I am able to make some progress in pinning down the cost. By examining Wal-Mart's choice behavior of how it trades off the benefit (not observed) versus the cost (observed with some work), it is possible to draw inferences about how Wal-Mart values the benefits.

The cost of high store density is that when stores are close together their market areas overlap and new stores cannibalize sales from existing stores. The extent of such cannibalization is something I can estimate. For this purpose, I bring together store-level sales estimates from ACNeilsen and demographic data from the Census at a very fine level of geographic detail to estimate a model of demand in which consumers choose among all the Wal-Mart stores in the general area where they live. The demand model fits the data well and I am able to corroborate its implications for the extent of cannibalization with certain facts Wal-Mart discloses in its annual reports. Using my sales model, I determine that Wal-Mart has encountered significant diminishing returns in sales as it has packed stores close together in the same area.

I write down a dynamic structural model of how Wal-Mart rolled out its stores over the period 1962-2005. The model is quite detailed and distinguishes the exact location of each individual store, the location of each distribution center, the type of store (regular Wal-Mart or supercenter) and the kind of distribution center (general merchandise or food). The model takes into account wage and land price differences across locations. The model takes into account that while there might be benefits of high *store density* to Wal-Mart, there also might be disadvantages of high *population density*—beyond high wages and land prices—as the Wal-Mart model might not work so well in very urban locations.

Given the enormous number of different possible combinations of store-opening sequences, it is difficult to directly solve Wal-Mart's optimization problem. This makes conventional approaches used in the industrial organization literature infeasible. Instead, I use the moment inequality approach outlined in Pakes, Porter, Ho, and Ishii (2006). I consider a set of selected deviations from what Wal-Mart actually did and determine the set of parameters consistent with this decision. The procedure works well and I am able to bound the importance of density economies. My estimates indicate that the benefits to Wal-Mart of high store density are substantial and likely extend significantly beyond savings in trucking costs.

An economy of density is a kind of economy of scale. Over the years various researchers have made distinctions among types of scale economies and noted the role of density. For the airline industry, Caves, Christensen, and Tretheway (1984) distinguish an economy of density from traditional economies of scale as arising when an airline increases the frequency of flights on a given route structure (as opposed to increasing the size the route structure holding fixed the flight frequency per route). The analogy here would be Wal-Mart expanding by adding more stores in the same markets it already serves (as opposed to expanding its geographic reach and keeping store density the same). Roberts (1986) makes an analogous distinction in the electric power industry. This paper differs from the existing empirical literature in three ways. First, there is a rich micro modeling with an explicit spatial structure. I don't have lumpy market units (e.g. a metro area) within which I count stores; rather I employ a continuous geography. Second, I explicitly model the channel of the density benefits through the distribution system, rather than have them be a "black box." Third, rather than directly relate costs to density, I use a revealed preference approach as explained above.

There is a large literature on entry and store location in retail.[3] There is actually a growing literature on Wal-Mart itself.[4] This paper is most closely related to the recent parallel work of Jia (2007).[5] Jia estimates density economies by examining the site selection problem of Wal-Mart as the outcome of a static game with K-Mart. Jia's paper features interesting oligopolistic interactions that my paper abstracts away from. My paper highlights (1) dynamics and (2) cannibalization of sales by nearby Wal-Marts that Jia's paper abstracts away from.

2 Model

A retailer (Wal-Mart) has a network of stores supported by a network of distribution centers. The model specifies how Wal-Mart's revenues and costs in a period depend on the configuration of stores and distribution centers that are open in the period. And the model specifies how the networks change over time.

There are two categories of merchandise: *general merchandise* (abbreviated by g) and *food* (abbreviated by f). There are two kinds of Wal-Mart stores. A *regular store* sells only general merchandise. A *supercenter* sells both general merchandise and food.

There are a finite set of locations in the economy. Locations are indexed by $\ell = 1, ...L$. Let $d_{\ell\ell'}$ denote the distance in miles between any given pair of locations ℓ and ℓ'. At any given

[3] See, for example, Bresnahan and Reiss (1991) and Toivanen and Waterson (2005).

[4] Recent papers on Wal-mart include: Basker (2005), Stone (1995), Hausman and Leibtag (2005), Ghemawat, Mark, and Bradley (2004)), and Neumark et al (2005), and Basker (2007).

[5] See also Andrews, Berry, and Jia (2004).

period t, a subset B_t^{Wal} of locations have a Wal-Mart. Of these, a subset $B_t^{Super} \subseteq B_t^{Wal}$ are supercenters and the rest are regular stores. In general the number of locations with Wal-Marts will be small relative to the total set of locations and a typical Wal-Mart will draw sales from many locations.

Sales revenues at a particular store depend upon the store's location and its proximity to other Wal-Marts. Let $R_{jt}^g(B_t^{Wal})$ be general merchandise sales revenue of store j at time t given the set of Wal-Mart stores open at time t. If store j is a supercenter, then its food sales $R_{jt}^f(B_t^{Super})$ analogously depend upon the configuration of supercenters. The model of consumer choice will be specified below in Section 4 from which this demand function will be derived. In this demand structure, Wal-Mart stores that are near each other will be regarded as substitutes by consumers. That is, increasing the number of nearby stores will decrease sales at a particular Wal-Mart.

I abstract from price variation and assume Wal-Mart sets constant prices across all stores and over time. In reality, prices are not always constant across Wal-Marts, but the company's Every Day Low Price (EDLP) policy makes this a better approximation for Wal-Mart than it would be for many retailers. Let μ denote the gross margin. Thus for store j at time t, $\mu R_{jt}^g(B_t^{Wal})$ is sales receipts less cost of goods sold for general merchandise.

In the analysis there will be three components of cost that will be relevant besides cost of goods sold: (1) distribution costs, (2) variable store costs, and (3) fixed costs at the store level. I describe each in turn.

Distribution Cost

Each store requires distribution services. General merchandise is supplied by a General Distribution Center (GDC) and food by a Food Distribution Center (FDC). For each store, these services are supplied by the closest distribution center. Let d_{jt}^g be the distance in miles from store j to the closest GDC at time t and analogously define d_{jt}^f. If store j is a supercenter, its distribution cost at time t is

$$Distribution\_Cost_{jt} = \tau d_{jt}^g + \tau d_{jt}^f.$$

where the parameter τ is the cost per mile per period per merchandise segment (general or food) of servicing this store. If j carries only general merchandise the cost is τd_{jt}^g.

Note that the distribution cost here is a fixed cost that does not depend upon the volume of store sales. This would be an appropriate assumption if Wal-Mart made a single delivery run from the distribution center to the store each day. The driver's time is a fixed cost and the implicit rental on the tractor is a fixed cost that must be incurred regardless of the size of the load. To keep a tight rein on inventory and to allow for quick response, Wal-Mart aims to have daily deliveries even for its smaller stores. So there clearly is an important fixed cost component to distribution. Undoubtedly there is a variable cost component as well, but for simplicity I abstract from it.

Variable Costs

The larger the sales volume at a store, the greater the number of workers needed to staff the checkout lines, the larger the parking lot, the larger the required shelf space, and the bigger the building. All of these costs are treated as variable in this analysis. It may seem odd to treat building size and shelving as a variable input. However, Wal-Mart very frequently updates and expands its stores. So in practice, store size is not a permanent decision that is made once and for all but is rather a decision made at multiple points over time. Treating store size as a variable input simplifies the analysis significantly.

Assume that the variable input requirements at store j are all proportionate to sales volume R_j,

$$\begin{aligned} Labor_j &= \nu^{Labor} R_j \\ Land_j &= \nu^{Land} R_j \\ Other_j &= \nu^{Other} R_j. \end{aligned}$$

Wages and land prices vary across locations and across time. So let w_{jt} and r_{jt} denote the wage and land rental rate that store j faces at time t. *Other* consists of everything left out so far that varies with sales, including the rental on structure and equipment in the store (the shelving, the cash registers, etc.) The other cost component of variable costs is assumed to be the same across locations and the price is normalized to one.

Fixed Cost

We might expect there to be a fixed cost with operating a store. To the extent the fixed cost is the same across locations, it will play no role in the analysis of where Wal-Mart places a given number of stores. We are only interested in the component of fixed cost that varies across locations.

From Wal-Mart's perspective, urban locations have some disadvantages compared to non-urban locations. These disadvantages go beyond higher land rents and higher wages that have already been taken into account above. The Wal-Mart model of a big box store at a convenient highway exit is not applicable in a very urban location. Moreover, Sam Walton was very concerned about the labor force available in urban locations, as he explained in his autobiography (Walton (1992)). We might expect, for example, that urban workers would be less interested in joining in on the trademark Wal-Mart cheer (Give me a "W"...). Urban locations are more susceptible to unions and Wal-Mart has been very up front about not wanting unions in its stores.

To capture potential disadvantages of urban locations, the fixed cost of operating store j is written as a function $f(m_j)$ of the population density m_j of the store's location. The functional form is quadratic in logs,

$$f(m_j) = \omega_0 + \omega_1 \ln(m_j) + \omega_2 \ln(m_j)^2. \tag{1}$$

A supercenter is actually two stores, a general merchandise and food store, so the fixed cost is paid twice. It will be with no loss of generality in our analysis to assume that the constant term $\omega_0 = 0$ since the only component of the fixed cost that will matter in the analysis is the part that varies across locations.

Dynamics

Everything that has been discussed so far considers quantities for a particular time period. I turn now the dynamic aspects of the model. Wal-Mart operates in a deterministic environment in discrete time where it 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?

The main focus of the paper is on part 2 of Wal-Mart's problem. The analysis conditions on the answers to 1, 2, and 4, in terms of what Wal-Mart actually did, and solves 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, conditioned 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 in this paper.

Problems 3 and 4 regarding distribution centers are closely related to the main issue of this paper. I will have something to say about this later in Section 7.

Now for more notation. To begin with, the discount factor each period is β. The period length is a year and discount factor is set to $\beta = .95$.

As defined earlier, B_t^{Wal} is the set of Wal-Mart stores in period t and B_t^{Super} is the subset of supercenters. 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 as it rarely closes stores.[6] 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, a supercenter is an absorbing state, $B_t^{Super} = B_{t-1}^{Super} + A_t^{Super}$, for A_t^{Super} the set of new supercenter openings in period t. 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.

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 is defined as part 1 of Wal-Mart's problem. These are taken as given here. Also taken as given is the location of distribution centers of each type and their opening dates (parts 3 and 4 of Wal-Mart's problem).

There is exogenous productivity growth of Wal-Mart according to a growth factor ρ_t in period t. If Wal-Mart were to hold fixed the set of stores and demographics also stayed the same, then from period 1 to period t revenue and all components of costs would grow by (an annualized) factor ρ_t. 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 initially hardware and variety items to food, drugs, eye glasses and tires, etc.. The part of growth due to food through the expansion into supercenters is explicitly accounted for here. But expansion into drugs, eye glasses, tires, etc., is not modeled explicitly. Instead this growth is implicitly picked up through the exogenous growth parameter ρ_t. The role ρ_t plays in Wal-Mart's problem is like a discount factor.

[6] Wal-Mart's annual reports disclose store closings that are on the order of two a year.

A policy choice of Wal-Mart is a vector $a = (A_1^{Wal}, A_1^{Super}, A_2^{Wal}, A_2^{Super}, ..., A_T^{Wal}, A_T^{Super})$ specifying the locations of the new stores opened in each period t. Define a choice vector a to be feasible if the number of store openings in period t under policy a equals what Wal-Mart actually did, i.e. N_t^{Wal} new stores in a period and N_t^{Super} supercenter openings. Wal-Mart's problem at time 0 is to pick a feasible a to maximize

$$\max_a \sum_{t=1}^{T} (\rho_t \beta)^{t-1} \left[\sum_{j \in B_t^{Wal}} \left[\pi_{jt}^g - f_{jt}^g - \tau d_{jt}^g \right] + \sum_{j \in B_t^{Super}} \left[\pi_{jt}^f - f_{jt}^f - \tau d_{jt}^f \right] \right], \tag{2}$$

where the operating profit for merchandise segment $k \in \{g, f\}$ at store j in time t is

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

and where d_{jt}^k is the distance to the closest segment k distribution center at time t.

3 The Data

There are five main data elements used in the analysis. The first element is store-level data on sales and other store characteristics. The second is opening dates for stores and distribution centers. The third is demographic information from the Census. The fourth element is data on wages and rents across locations. The fifth is other information about Wal-Mart from annual reports.

Data element one, store-level 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 of the end of 2005.[7] (Alaska and Hawaii are excluded in all of the analysis.) As of this point in time, almost two thirds of all Wal-Marts (1,980 out of 3,176) are supercenters carrying both general

[7] 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.

merchandise and food. The remaining 1,196 are *Regular Wal-Marts* that do not have a full selection of food. The average Wal-Mart has annual sales of $70 million. The breakdown is $47 million per regular Wal-Mart and $85 million per supercenter. The average employment is 255 employees.

The second data element is opening dates of the four types of Wal-Mart facilities. The table treats a supercenter as two different stores: a general merchandise store and a food store. There are two kinds of distribution centers, general (GDC) and food (FDC). Table 2 tabulates opening dates for the four types of facilities by decade. The appendix explains how this information was collected. Note that if a regular store is later converted to a supercenter, it has an opening date for its general merchandise store and a later opening date for its food store. This is called a *conversion.*

The third data element, demographic information, comes from three decennial censuses: 1980, 1990, and 2000. The data is at the level of the *block group*, a geographic unit finer than the Census tract. Summary statistics are provided in Table 3. In 2000, there were 206,960 block groups with an average population of 1,350. I use the geographic coordinates of each block group to draw a circle of radius five miles around each block group. I take the population within this five mile radius and use this as my population density measure. Table 3 reports that the mean density in 2000 across block groups equals 219,000 people within a 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.[8]

The fourth data element is information about local wages and local 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. It is difficult to obtain a consistent measure of land rents at a fine degree of geographic detail over a long period of time. To proxy land rents, I use information about residential property values from the 1980, 1990, and 2000 decennial censuses. For each Census year and each store location, I create an index of property values by adding up the total value of residential property within two miles of the store's location and scaling it so the units are in inflation adjusted dollars per acre. See the appendix for how the index is constructed. Interpolation is used to obtain values between Census years. The Census data is supplemented with property tax data on property valuations of actual Wal-Mart store locations in Iowa and Minnesota. As discussed in the appendix, there is a high correlation between the tax assessment property

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

valuations of a Wal-Mart site and the property value index.

The fifth data element is information from Wal-Mart's annual reports including information about aggregate sales for earlier years. I also use information provided in the "Management Discussion" section of the reports on the degree to which new stores cannibalize sales of existing stores. The specifics of this information are explained below when the information is incorporated into the estimation.

4 Estimates of Operating Profits

This section estimates the components of Wal-Mart's operating profits. Part 1 specifies the demand model and Part 2 estimates it. Part 3 treats various cost parameters. Part 4 explains how estimates for 2005 are extrapolated to other years.

4.1 Demand Specification

Conditioned on shopping at some Wal-Mart, presumably a consumer will tend to shop at the Wal-Mart closest to home. Nevertheless, for various reasons, some consumers will shop at other Wal-Marts. For example, a particular consumer may pass a Wal-Mart on the way to work and it may be more convenient to shop there than the Wal-Mart closest to home. To allow for substitution possibilities such as this, the various Wal-Marts in the vicinity of a consumer are assumed to be differentiated products. A discrete choice approach is employed, following common practice in the literature.

For a given location ℓ, let n_ℓ denote the *population* of location ℓ and let m_ℓ be the *population density* at ℓ. In the empirical work a location will be a Census Block Group and the population density measure will be the number of people residing within a five mile radius of the Block Group.

Let $y_{\ell j}$ denote the distance in miles between location ℓ and store location j. Define $\bar{B}_\ell^{Wal}$ to be the set of Wal-Marts in the *vicinity* of the consumer's location, defined to be those locations within 25 miles,

$$\bar{B}_\ell^{Wal} = \left\{ j, j \in B^{Wal} \text{ and } y_{\ell j} \leq 25 \right\}.$$

(The time subscript t is left out here because the time period is held fixed in this part of the paper.)

In the model, the shopping decision for general merchandise is separated from the shopping decision for food. In general, we expect that there would be some complementarity—once a consumer is in the store to buy a lawn mower, the consumer might also buy food for dinner. I ignore this issue in part because I don't have good data to get a handle on the issue. (In particular, I don't have a clean breakdown between the two segments in my store-level data.) I have no reason to believe that abstracting from such a complementarity biases my results in a particular direction.

I explain the purchase decision for general merchandise; the food purchase choice problem is analogous. Consider a consumer at a particular location ℓ. The consumer has a budget λ^g for spending on general merchandise. The consumer makes a discrete choice between buying general merchandise from the outside alternative (labeled $j = 0$) or from one of the Wal-Marts in $\bar{B}_\ell^{wal}$ (assuming $\bar{B}_\ell^{wal}$ is nonempty).

If the consumer chooses the outside alternative 0, utility is

$$u_0 = b(m_\ell) + z_\ell \alpha + \varepsilon_0. \tag{3}$$

The first term is a function $b(\cdot)$ that depends upon the population density m_ℓ at the consumer's location ℓ. Assume $b'(m) \geq 0$; i.e., the outside option is better in higher population density areas. 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. I don't have sufficient data to conduct such an analysis so instead specify the reduced-form relationship between $b(m_\ell)$ and population density.[9] The functional form used in the estimation is

$$b(m) = \alpha_0 + \alpha_1 \ln(\underline{m}) + \alpha_2 \left(\ln(\underline{m})\right)^2$$

where

$$\underline{m}_j = \max\left\{1, m_j\right\}. \tag{4}$$

The units of the density measure is thousands of people within a five mile radius. By truncating m at one, $\ln(m)$ is truncated at zero. All locations with less then one thousand people within five miles are grouped together.[10]

[9] One way that the recent empirical literature in industrial organization has been making progress is by estimating policy functions and equilibrium relalationships directly rather than the underlying structural parameters, e.g. Bajari, Benkard, and Levin (2007).

[10] This same truncation is applied throughout the paper.

The second term of (3) allows demand for the outside good to depend upon a vector z_ℓ of location characteristics that impact utility through the parameter vector α. In the empirical analysis, these characteristics will include the demographic characteristics of the block group and income.

The third term is a logit error term. Assume this is drawn i.i.d. across all consumers living in the block group ℓ.

This explains utility of the outside alternative. Now consider the utility from buying at a particular Wal-Mart $j \in \bar{B}_\ell^{Wal}$. It equals

$$u_{\ell j} = -h\left(m_\ell\right) y_{\ell j} + x_j \gamma + \varepsilon_j, \tag{5}$$

for $h(m)$ parameterized by

$$h(m) = \xi_0 + \xi_1 \ln(\underline{m}).$$

The first term of (5) is the disutility of commuting $y_{\ell j}$ miles to the store from the consumer's home. The coefficient $h\left(m_\ell\right)$ can be interpreted as a transportation cost per mile. The specification allows the transportation cost to depend upon population density. The second term of (5) allows utility to depend upon other characteristics x_j of Wal-Mart store j. The other store characteristic included in the empirical analysis is store age. In this way, it is 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 last term is the logit error ε_j. A consumer who finds store j a convenient place to stop on the way home from work can be interpreted as a consumer with a high value of ε_j.

Using the standard logit formulas, the probability that a consumer at location ℓ find's Wal-Mart j to be be the best option for the consumer's general merchandise needs is

$$p_{j\ell}^{g} = \frac{\exp(\delta_{j\ell})}{\left[\exp\left(\delta_{0\ell}\right)\right] + \sum_{k \in \bar{B}_\ell^{Wal}} \exp(\delta_{k\ell})}, \tag{6}$$

for

$$\begin{aligned} \delta_{0\ell} &\equiv b(m_\ell) + z_\ell \alpha \\ \delta_{j\ell} &\equiv -h\left(m_\ell\right) y_{\ell j} + x_j \gamma. \end{aligned}$$

The model's predicted general merchandise revenue for store j is

$$R_j^g = \sum_{\{\ell | j \in \bar{B}_\ell^g\}} \lambda^g \times p_{j\ell}^g \times n_\ell. \tag{7}$$

In words, there are n_ℓ consumers at location ℓ and a fraction $p_{j\ell}^g$ of them are shopping at j where they will each spend λ^g dollars.

Spending on food is modeled the same way. The parameters are the same except for the spending λ^f per consumer. The formula for food revenue R_j^f at store j is analogous to (7). Even though the parameters for food are the same as for general merchandise, the probability $p_{j\ell}^f$ a consumer at ℓ shops at j for food will differ from the probability $p_{j\ell}^g$ the consumer shops for general merchandise. This follows because the set of alternatives for food $\bar{B}_\ell^f$ is in general different from the set of alternatives $\bar{B}_\ell^g$ for general merchandise.

4.2 Demand Estimation

Given a vector θ of parameters from the demand model, we can plug in the demographic data and obtain predicted values of general merchandise sales $\hat{R}_j^g(\theta)$ for each store j from equation (7) and predicted values of food sales $\hat{R}_j^f(\theta)$.

The data has *all commodity sales volume* for each store. Call this R_j. General merchandise is all items sold at a regular Wal-Mart. So for regular stores, $R_j = R_j^g$, by definition. For supercenters, all commodity sales volume includes general merchandise and food, $R_j = R_j^g + R_j^f$.

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

$$\eta_j^{Wal} = \ln(R_j) - \ln(\hat{R}_j^g(\theta)).$$

For supercenters, this is

$$\eta_j^{Super} = \ln(R_j) - \ln(\hat{R}_j^g(\theta) + \hat{R}_j^f(\theta)).$$

Assume the discrepancies η_j^{Wal} and η_j^{Super} are i.i.d. normally distributed measurement error. (The store-level sales figures in the TradeDimensions data are estimates so certainly measurement error is an issue.) The model is estimated using maximum likelihood and the coefficients are reported in Table 4 in the column labeled "Unconstrained Model."

The extent that new stores cannibalize the sales of existing stores will make a big difference in the subsequent analysis. So our first order of business is to assess how well the demand model is doing in getting this right. Fortunately, Wal-Mart has provided information that is helpful in this regard. Wal-Mart's annual report for 2004 disclosed (Wal-Mart Stores, Inc. (2004, p. 20)),

> "As we continue to add new stores domestically, 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 with regards to fiscal year 2005 and 2006. This information is summarized in Table 5.[11]

To define the model analog of cannibalization, for each vector θ of model parameters, first calculate what sales would have been in a particular year to preexisting stores if no new stores had opened in the year and if there were no new supercenter conversions. Next calculate predicted sales to pre-existing stores when the new store openings and supercenter conversions for particular year take place. Define the percentage difference to be the *cannibalization rate* for that year. This is the model analog of what Wal-Mart is disclosing.

Table 5 reports the cannibalization rate for various years using the estimated demand model. The parameter vector is the same across years. What varies over time are the new stores, the set of pre-existing stores and the demographic variables.[12] The demand model—estimated entirely off of the 2005 cross-section store-level sales data—does a very good fitting the cannibalization rates reported by Wal-Mart. For the years that Wal-Mart disclosed that the rate was "approximately one", the estimated rates range from .67 to 1.43. It is interesting to note the sharp increase in the estimated cannibalization rate beginning in 2002. Evidently, Wal-Mart reached some kind of saturation point in 2001. Given the pattern in Table 5, it is understandable that Wal-Mart has felt the need to disclose the extent of cannibalization in recent years.

In what follows, the estimated upper bound on the degree of density economies will be closely connected to the degree of cannibalization. The more cannibalization Wal-Mart is willing to tolerate, the higher the inferred density economies. The estimated cannibalization

[11] Wal-Marts fiscal year ends January 31. So the fiscal year corresponds (approximately) to the previous calendar year. For example, the 2006 fiscal year, began February 1, 2005. In this paper, I aggregate years like Wal-Mart (February through January), but I use 2005 to refer to the year begining February 2005.

[12] To obtain demographic characteristics between Census years, I interpolate as discussed below.

rates of 1.38, 1.43, and 1.27 for 2003, 2004, 2005 are certainly "approximately one" but one may worry that these rates are on the high end of what would be consistent with Wal-Mart's reports. To explore this issue further, I estimate a second demand model where the cannibalization rate for 2005 is constrained to be exactly one. The estimates are reported in the last column of Table 4. The goodness of fit under the constraint is close to the unconstrained model, although a likelihood ratio test leads to a rejection of the constraint. In the interests of being conservative in my estimate of an upper bound on density economies, I will use the constrained model throughout as the baseline model.

A few remarks about the remaining parameter estimates. Recall that λ^g and λ^f are spending per consumer in the general merchandise and food categories. The estimates can be compared to aggregate statistics. For 2005, per capita spending in the U.S. was 1.8 in general merchandise stores (NAICS 452) and 1.8 in food and beverage stores (NAICS 445) (in thousands of dollars). The aggregate statistics match well the model estimates ($\lambda^g = 1.9$ and $\lambda^f = 1.9$ in the constrained model, $\lambda^g = 1.7$, $\lambda^f = 1.6$ in the unconstrained model), though we would not expect them to match exactly.[13]

The parameter estimates reveal that, as hypothesized, the outside good is better in more dense areas and that utility decreases in distance travelled to a Wal-Mart. To get a sense of the magnitudes, Table 6 examines how predicted demand in a block group varies with population density and distance to the closest Wal-Mart. (The table is constructed with the constrained model but things look very similar with the unconstrained model.) The table reports the probability that a consumer shops at Wal-Mart for general merchandise. For the analysis, the demographic variables in Table 3 are set to their mean values. There is assumed to be only one Wal-Mart (two or more years old) in the vicinity of the consumer (i.e. within 25 miles) and the distance to this Wal-Mart is varied. Consider the first row, where distance is set to zero (the consumer lives right next door to a Wal-Mart) and population density is varied. The negative effect of population density on demand is substantial. A rural consumer right next to a Wal-Mart shops there with a probability that is essentially one. At a population density of 50,000 this falls to .72 and falls to only .24 at 250,000. 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

[13] 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. On 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 (the electronics sector is less than a fifth of the general merchandise sector) so perhaps it is not a surprise that my estimate of λ^g is so close to U.S. per capita spending in this category.

increases in large markets. A market that is 250 times as large as an isolated market may have a per capita demand only a fifth as large, but overall demand is 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 has an appreciable effect, .98 to .57, but still much demand remains. Contrast this with higher density areas. At a population density of 250,000 an increase in distance from 0 to 5 miles reduces demand on the order of 80 percent while the effect of such a change in a very rural market would be 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. (This is the same thing as saying the coefficient on per capita income for the outside good is positive.) Demand for Wal-Marts is lower among blacks and young people.

The only store characteristic used in the demand model (besides location) is store age. This is captured with a dummy variable for stores that have been open two or more years. This variable enters positively in demand. So everything else the same, older stores attract more sales.

4.3 Variable Costs at the Store Level

In the description of the model, the required labor input at the store level was assumed to be proportionate to sales. To get a sense of the plausibility of this assumption, Figure 3 provides a scatter plot of square footage of each store along with sales per employee of each store in 2005. Also plotted are the fitted values of a locally-weighted regression. At the bottom end of the size distribution, there is evidence of increasing returns. But things flatten out and there is roughly constant returns over most of the store size distribution. The weighted average over all stores is $277 thousand dollars in sales per employee. Equivalently, there are 3.61 store employees per million dollars of annual sales. I use this as the estimate of the fixed labor coefficient, $\nu_{Labor} = 3.61$. To covert this into a cost of labor at a particular store, the coefficient is multiplied by average retail wage (annual payroll per worker) in the county where the store is located. Table 7 reports information about the distribution of labor costs across the 2005 set of Wal-Mart stores. The median store faces a labor cost of $20,700 per worker. Given $\nu_{Labor} = 3.61$, this translates into a labor cost of $3.61 \times 20,700$ per million in sales or equivalently 7.5 percent of sales. The highest labor costs can be found at stores in

San Jose, California where wages are almost twice as high as they are for the median store.

An issue that needs to be raised about the County Business Patterns wage data is measurement error. Dividing annual payroll by employment is a crude way to measure labor costs because it doesn't take into account potential variations in hours per worker (e.g. part-time versus full-time) or potential variations in labor quality. The empirical procedure used below explicitly takes into account measurement error.

Turning now to land costs, the appendix describes the construction of a *property value index* for each store through the use of Census data. As discussed in the appendix, this index along with property tax data for 46 Wal-Mart locations in Minnesota and Iowa are used to estimate a *land-value/sales ratio* for each store. The distributions of this index and ratio are reported in Table 7. Perhaps not surprisingly, the most expensive location is estimated to be the Wal-Mart store in Silicon Valley (in Mountain View, California) where the ratio of the land value for the store relative to store sales is estimated to be 65 percent. The rental cost of the land, including any taxes that vary with land value, is assumed to be 20 percent of the land value. So for the median store from Table 7 (the Wal-Mart in Cleburne, Texas), this implies annual land costs of about half a percent of sales ($.5 \approx .2 \times 2.4$). It is important to emphasize that this rental cost is for the land, not structures. (Half of a percent of sales would be a very low number for the combined rent on land and structures and equipment.) The rents on structures and equipment are separated out because these should be approximately the same across locations, as least as compared to variations across stores in land rents. The cost of cinderblocks for walls, steel beams for roofing, shelving, cash registers, asphalt for parking lots, etc., are all assumed to be the same across locations.

So I now turn to those aspects of variable costs that are the same across locations. I begin with cost of goods sold. Wal-Mart's gross margin over the years has ranged from .22 to .26. (See Wal-Marts annual reports.) To be consistent with this, the gross margin is set equal to $\mu = .24$.

Wal-Mart has reported over the years *operating selling, general and administrative expenses* that are in the range of 16 to 18 percent of sales. Included in this is the store-level labor cost discussed above that is on the order of 7 percent of sales and has already been taken account of. Also included in this cost is the cost of running the distribution system, the fixed cost of running central administration and other costs I don't want to include as variable costs. I set the residual variable cost parameter $\nu^{other} = .07$. Netting this out of the gross margin μ yields a net margin $\mu - \nu^{other} = .17$. In the analysis, the breakdown between μ and ν^{other} is irrelevant, only the difference.

The analysis so far has explained how to calculate the operating profit of store j in 2005 as

$$\pi_{j,2005} = \left(\mu - \nu^{other}\right)\left(R^{g}_{j,2005} + R^{f}_{j,2005}\right) - LaborCost_{j,2005} - Landrent_{j,2005} \tag{8}$$

where the sales revenue comes from the 2005 demand model and labor cost and land rent are explained above. The next step is to extrapolate this model to earlier years.

4.4 Extrapolation to Other Years

We have a demand model for 2005 in hand but need models for earlier years. To get them, assume demand in earlier years is the same as in 2005 except for the multiplicative scaling factor ρ_t introduced above in the definition of Wal-Mart's problem (2). For example, the 2005 demand model with no rescaling predicts that, at the 1971 store set and 1971 demographic variables, average sales per store (in 2005 dollars) is $31.5 million. Actual sales per store (in 2005 dollars) for 1971 is $7.4 million. The scale factor for 1971 adjusts demand proportionately so that the model exactly matches aggregate 1971 sales. Over the 1971 to 2005 period, this corresponds to a compound annual real growth rate of 4.4 percent. Wal-Mart significantly widened the range of products it sold over this period (to include tires, eyeglasses, etc.). The growth factor is meant to capture this. The growth factor calculated in this manner has levelled off in recent years to around one percent a year. Wal-Mart has also been expanding by converting regular stores to supercenters. This expansion is captured explicitly in the model rather than indirectly though exogenous growth.

Demographics change over time and this is taken into account. For 1980, 1990, and 2000, I use the decennial census for that year.[14] For years in between I use a convex combination of the censuses. For example, for 1984 there is .6 weight on 1980 and .4 weight on 1990, meaning 60 percent of the people in each 1980 block group are assumed to be still around as potential Wal-Mart customers and 40 percent each 1990 block group consumers have already arrived as of 1984. This procedure keeps the geography clean, since the issue of how to link block groups over time is avoided. (Block group definitions do not stay constant across census years, but this poses no problem whatsoever with the continuous approach to geography taken here.)

[14] I use 1980 for years before 1980 and 2000 for years after 2000.

5 Preliminary Evidence of a Tradeoff

This section provides some preliminary evidence of an economically significant tradeoff to Wal-Mart. Namely, the benefits of increased economies of density come at the cost of cannibalization of existing stores. This section puts to work the demand model and other components of operating profits compiled above.

Consider some Wal-Mart store j that first opens in time t. Define the *incremental sales* $R_{j,t}^{k,inc}$ of store j to be what the store adds to total Wal-Mart sales in segment $k \in \{g, f\}$ in its opening year t, relative to what sales would otherwise be across all other stores open that year. The incremental sales of store #1 opening in 1962 equals $R_{j,1962}^{k}$ that year. For a later store j, however, the incremental sales are in general less than store j's sales, $R_{j,t}^{k,inc} \leq R_{j,t}^{k}$, because some part of the sales may be diverted from other stores. Using the demand model, we can calculate $R_{j,t}^{k,inc}$ for each store.

Table 8 reports that the average annual incremental sales in general merchandise across all Wal-Mart stores in the year the stores opened is \$36.3 million (in 2005 dollars throughout). Analogously, average incremental sales in food from new supercenters is \$40.2 million. (Note conversions of existing Wal-Marts to supercenters count as a store openings here.) To make things comparable across years, the 2005 demand model is applied to the store configurations and demographics of the earlier years with no multiplicative scale adjustment ρ_t. In an analogous manner, we can use (8), to determine the *incremental operating profit* of each store at the time it opens. The average incremental profit in general merchandise from a new Wal-Mart is \$3.1 million and in food from new supercenters is \$3.6 million. Finally, we can ask how far a store is from the closest distribution center in the year it is opened. On average a new Wal-Mart is 168.9 miles from the closest regular distribution center when it opens and a new supercenter is 137.0 miles from the closest food distribution center.

Incremental sales and operating profit can be compared to what sales and operating profit would be if a new store were a stand-alone operation. That is, what would sales and operating profits be at a store if it were isolated so that none of its sales are diverted to or from other Wal-Mart stores in the vicinity? Table 8 shows for the average new Wal-Mart, there is a big difference between stand-alone and incremental values, implying a substantial degree of market overlap with existing stores. Average stand-alone sales is \$41.4 million compared to an incremental value of \$36.3 million, approximately a 10 percent difference. Two considerations account for why the big cannibalization numbers found here are not inconsistent with the one percent cannibalization rates reported earlier in Table 5. First, the denominator of the cannibalization rate from Table 5 includes all pre-existing stores, including those areas of the country were Wal-Mart is not adding any new stores. Taking an average over the country as a whole understates the degree of cannibalization taking place where Wal-Mart is adding new stores. Second, stand-alone sales include sales that a new store would never get because the sales would remain in some existing store (but would diverted to the new store if existing stores shut down).

Define the *Wal-Mart Age* of a state to be the number of years that Wal-Mart has been in the state.[15] The remaining rows in Table 8 classify stores by the Wal-Mart age of their state at the store's opening. Those stores in the row labeled "'1-2" are the first stores in their respective states. Those stores in the row labeled "21 and above" are opened when Wal-Mart has been in their states for over 20 years.

Table 8 shows that incremental operating profit in a state falls over time as Wal-Mart adds stores to a state and the store market areas increasingly begin to overlap. Things are actually flat the first five years at 3.5 million in incremental operating profit for general merchandise. But it falls to 3.3 million in the second five years and then to 2.9 million and lower beyond that. An analogous pattern holds for food. This pattern is a kind of diminishing returns. Wal-Mart is getting less incremental operating profit from the later stores it opens in a state.

The table also reveals a benefit from opening stores in a state where Wal-Mart has been there for many years. The incremental distribution center distance is relatively low in such states. It decreases substantially as we move down the table and to states with higher Wal-Mart ages. The very first stores in a state average about 300 miles from the closest distribution center. This falls to less than 100 miles when the Wal-Mart age of the state is over 20 years. There is a tradeoff here: the later stores deliver lower operating profit but are closer to a distribution center. The magnitude of the tradeoff is on the order of 200 miles for one million in operating profit. This tradeoff is examined in a more formal fashion in the next section and the result is roughly of this order of magnitude.

As argued in Section 2, the fixed cost of operating a store may be higher in high density areas. The comparisons in Table 8 control for variable costs across locations but not fixed costs that depend upon population density. Table 9 runs the regression analog of Table 8 with a control for population density that is quadratic in logs, following the specification of the fixed cost (1). In addition, fixed state effects are included in the regression. The

[15] 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.

idea is to hold fixed population density and determine how incremental profit varies within a state depending on whether a store is an early opener or a late opener in the state. Adding controls for population density and state fixed effects makes little difference. For example, the difference between the 11-15 group and the 1-2 group is .63 in the regression and .6=3.5 - 2.9 in the raw data. These differences between early openers and late openers are highly statistically significant.

6 Bounding Density Economies

There remain three parameters to pin down, $\theta = (\tau, \omega_1, \omega_2)$, all relating to density of one form or another. The τ parameter is the coefficient on distance between a store and its distribution center. It captures the benefit of *store density*. The parameters ω_1 and ω_2 relate to how fixed cost varies with *population density*. They are the coefficients on log population density and its square in the fixed cost specification (1).

As discussed in the introduction, the cost of distance τ includes the physical costs of moving goods. It is also intended to capture the idea when a store and its distribution center are far apart, the Wal-Mart model of quick response to demand shocks does not work so well. It would be difficult to directly measure the indirect ways that distance impedes the Wal-Mart way of doing business. Analogously, while it is certainly possible to take account the higher land prices and higher wages in big cities (and I do this), it is difficult to directly measure some of the disadvantages alluded to earlier of implementing the Wal-Mart model in big cities. So rather then try to estimate these parameters through direct measurement, the approach taken here is to infer the parameters from the way Wal-Mart behaves.

6.1 The General Method

The revealed preference approach taken here generates a set of inequalities. A bounds estimation strategy (see Manski (2003)) is a natural way to extract information from the set of inequalities implied by choice behavior. In my implementation of this strategy, I follow Pakes, Porter, Ho, and Ishii (2006) (Hereafter PPHI).

Recall that action a denotes a particular choice of Wal-Mart, a particular feasible solution to problem (2). Let a_0 denote the choice Wal-Mart actually made. For each policy a, let $\Pi(a)$ be the present value at date $t = 1$ of operating profits from general merchandise and food over all stores and all time periods given policy a,

$$\Pi(a) \equiv \sum_{t=1}^{T} (\rho_t \beta)^{t-1} \left(\sum_{j \in B_t^{Wal}(a)} \pi_{jt}^g(a) + \sum_{j \in B_t^{Super}(a)} \pi_{jt}^f(a) \right). \tag{9}$$

Analogously, let $D(a)$ be the present value of all distribution miles. This is the same as the formula for $\Pi(a)$ except the distribution distance $d_{jt}^k(a)$ of store j in year t in segment k replaces the operating profit $\pi_{jt}^k(a)$. Similarly, let $F_1(a)$ be the present value of the (log) population density for each store and year and let $F_2(a)$ be the present value of the square of (log) population density. (Recall specification (1) for how fixed cost varies with population density.) Then rewriting (2), the value to Wal-Mart of choosing policy a, given a vector of density parameters $\theta = (\tau, \omega_1, \omega_2)$, is

$$v(a, \theta) = \Pi(a) - \tau D(a) - \omega_1 F_1(a) - \omega_2 F_2(a). \tag{10}$$

Let θ_0 be the true parameter. The policy a_0 chosen by Wal-Mart solves problem (2). Hence at θ_0,

$$v(a_0, \theta_0) \geq v(a, \theta_0), \text{ for all } a \neq a_0.$$

Or

$$\Delta v(a, \theta_0) \geq 0,$$

for

$$\Delta v(a, \theta) \equiv v(a_0, \theta) - v(a, \theta). \tag{11}$$

We can decompose this as

$$\Delta v(a, \theta) = \Delta \Pi(a) - \tau \Delta D(a) - \omega_1 \Delta F_1(a) - \omega_2 \Delta F_2(a).$$

In the econometrics to follow, the error term will arise on account of measurement error. Recall that operating profit in market segment k at a particular store j in period t given some policy a can be written in the form

$$\pi_{jt}^k = (\mu - \nu_{other}) R_{jt}^k - w_{jt} \nu_{labor} R_{jt}^k - r_{jt} \nu_{land} R_{jt}^k.$$

As explained in Section 4, there is measurement error in the wage and land-rent measures. So the observed store operating profit is

$$\tilde{\pi}_{jt}^{k} = (\mu - \nu_{other}) R_{jt}^{k} - \left(w_{jt} + \varepsilon_{jt}^{wage}\right) \nu_{labor} R_{jt}^{k} - \left(r_{jt} + \varepsilon_{jt}^{rent}\right) \nu_{land} R_{jt}^{k},$$

for measurement error ε_{jt}^{wage} and ε_{jt}^{rent}. Assume ε_{jt}^{wage} is mean zero and independent of R_{jt}^{g}, d_{jt}^{g}, R_{jt}^{f}, d_{jt}^{f}, $F_{1,jt}$ and $F_{2,jt}$. Make the analogous assumption on ε_{jt}^{rent}. Aggregate across stores and time like in (9) to get the present value of observed operating profits $\tilde{\Pi}(a)$ for each action. Substitute this into the analog of (10) to get the observed value $\tilde{v}(a,\theta)$ under policy a given θ. Let $\Delta\tilde{v}(a,\theta)$ be the observed difference in value between the chosen policy a_0 and some other policy a. It can be written as

$$\Delta\tilde{v}(a,\theta) = \Delta\Pi(a) - \tau\Delta D(a) - \omega_1\Delta F_1(a) - \omega_2\Delta F_2(a) + \eta_a, \tag{12}$$

for $\eta_a = \varepsilon_{a_0} - \varepsilon_a$, and

$$\varepsilon_a \equiv \sum_{t=1}^{T} (\rho_t\beta)^{t-1} \left(\sum_{j\in B_t^{Wal}(a)} \left(\varepsilon_{jt}^{wage}\nu_{labor} + \varepsilon_{jt}^{rent}\nu_{land}\right) R_{jt}(a) \right). \tag{13}$$

To ease the notational burden, let $y_a = \Delta\Pi(a)$, $x_a = (\Delta D(a), \Delta F_1(a), \Delta F_2(a))'$.and $\theta = (\tau, \omega_1, \omega_2)$. Then

$$\Delta\tilde{v}(a,\theta) = y_a - x_a\theta + \eta_a. \tag{14}$$

Given the assumptions made on ε_{jt}^{wage} and ε_{jt}^{rent}, the composite measurement error η_a is mean zero and independent of x_a. Note that at the true parameter θ_0,

$$y_a - x_a\theta \geq 0, \text{ for all feasible } a.$$

Consider a set of feasible deviations defined in a manner that is unrelated to the measurement error η_a. Let there be N such sets indexed by i and let A_i denote the i-th set. Let $w_{a,i} \geq 0$ be weighting variables. Define the *basic moment* for each i by

$$m_i(\theta) = \sum_{a\in A_i} w_{a,i}\Delta\tilde{v}(a,\theta). \tag{15}$$

Next consider more complicated moments that are derived from interactions. Let x_a^k be the k-th element of x_a and suppose a lower bound $\underline{x}^k$ exists so that $x_a^k \geq \underline{x}^k$ for all a. Define

$$\underline{z}_a^k \equiv x_a^k - \underline{x}^k \geq 0.$$

Analogously, suppose an upper bound $\bar{x}^k$ exists so that $x_a^k \leq \bar{x}^k$ for all a. Define

$$\bar{z}_a^k \equiv \bar{x}^k - x_a^k \geq 0.$$

For each i define the *interaction moments* by

$$\begin{aligned} \underline{m}_i^k(\theta) &= \sum_{a \in A_i} w_{a,i} \Delta \tilde{v}(a,\theta) \underline{z}_a^k, \; k \in \{1,2,3\} \\ \bar{m}_i^k(\theta) &= \sum_{a \in A_i} w_{a,i} \Delta \tilde{v}(a,\theta) \bar{z}_a^k, \; k \in \{1,2,3\} \end{aligned} \tag{16}$$

The $\underline{z}_a^k$ and $\bar{z}_a^k$ are valid instruments because they are uncorrelated with the measurement error. If we think about $\Delta\tilde{v}(a,\theta)$ as a kind of residual, the use of the interaction moments here is analogous the familiar moment conditions for OLS, $0 = (Y - Xb)'X$.

There are N basic moments (one for each set A_i) and 6 additional interaction moments (see (16)) for each basic moment for a total of $7 \times N$ moments. Stack the moments in a vector $M(\theta)$ and index the moments by h. At the true parameter θ_0,

$$E[M_h(\theta)] \geq 0, \text{ for all } h. \tag{17}$$

The set of θ satisfying these moment inequalities is the identified set. To estimate this set, find the values that satisfy the sample analog of the moment inequalities. If no such values exist, then, following PPHI, take the value of θ closest to satisfying the inequalities. In particular, let the estimate of the identified set solve

$$\hat{\theta} = \arg\min \sum_{h=1}^{7N} M_h^-(\theta)^2$$

for

$$M_h^-(\theta) = \min\{M_h(\theta), 0\}.$$

In the estimation below, there exists a parameter region in which all $7N$ moment inequalities are satisfied. It is still necessary to specify what happens when there is no solution satisfying all the inequalities, because the issue comes up while simulating the standard errors.

The focus of the estimation is bounding the distribution cost parameter τ. A lower bound $\hat{\tau}_{lower}$ is obtained by minimizing τ over the set of feasible $\theta = (\tau, \omega_1, \omega_2)$. The feasible θ satisfy the sample analogs of (17) as well as the restrictions $\omega_1 \geq 0$ and $\omega_2 \leq 0$ (so the fixed cost (1) is weakly concave in log population density). An upper bound $\hat{\tau}^{upper}$ is obtained similarly. Linear programming techniques are employed.

6.2 Specifics of Method

Like Bajari and Fox (2005) and Fox (2005), I restrict attention to pairwise resequencing, i.e., deviations a in which the opening dates of *pairs* of stores are reordered. For example, store #1 actually opened 1962 and #2 opened 1964. A pairwise resequencing would be to open store #2 in 1962, store #1 in 1964 and to leave everything else the same.

Three classes of deviation sets are constructed.[16]

1. *Store-Density Decreasing.* Begin with the set of stores that open when the Wal-Mart age of their states is ten or more (i.e., there is at least one store in the state that is at least ten years old). For each such store j, find the set of stores, indexed by j', that (i) are opened three or more years after store j and (ii) open in a different state when the Wal-Mart age of that state is less than or equal to four years. Flipping the opening of j and j', this is set A_1.

2. *Store-Density Increasing.* Begin with the set of stores that open when the Wal-Mart age of their states is five or less. For each such store j, find the set of stores, indexed by j', that (i) are opened three or more years after store j and (ii) open in a different state when the Wal-Mart age of that state is more than ten years and (iii) if the timing of the two is flipped then the present value of distribution miles is higher under the actual policy by 150 ($\Delta D > 150$). Flipping the opening of j and j', this is set A_2.

3. *Population-Density Changing.* Define store locations by population density groupings in 1990. Let grouping 1 be locations with "less than 15" (thousand people within

[16] The opening of a general merchandise store and a food store are considered two different opening events. In cases where a supercenter opens from scratch rather that a conversion of an existing Wal-Mart, there are then two opening events. In all the pairs considered, a general merchanside opening is paired with another general merchandise opening, and a food opening with another food opening.

five miles), grouping 2 be "15 to 40", grouping 3 be "40 to 100", and grouping 4 be "over 100." Take pairs of stores opening more than two years apart in the same state in different population density groupings. Flip the order. The various combinations result in 6 different sets of deviations $A_3 - A_8$.

With *Store-Density Decreasing* deviations, instead of adding yet another store j in a state where Wal-Mart has already been in for over ten years, an alternative store j' that would have been one of the early stores in a another state is opened in its place. And store j doesn't open until store j' would have opened. With this kind of deviation, there is less cannibalization at the cost of more distribution miles, since the j' store will tend to be far from the distribution network at this early date. That the deviation was not optimal to pursue helps provide information on a lower bound $\hat{\tau}_{lower}$ on the importance of distribution miles. Analogously, the second set of deviations goes in the other direction reducing store density, providing information helpful for pinning down an upper bound $\hat{\tau}^{upper}$. The third category defines sets of 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.

6.3 Estimates

The set of *Store-Density Decreasing* deviation has 240,000 elements that involve reordering pairs of store opening dates that satisfy the given criteria. Calculating the terms of (12) for each pair is a lengthy process because the sales model is calculated at the block group level and there are so many block groups and often many Wal-Marts to be considered for each block group. To reduce computation costs, a random one-third sample was taken of the A_1 and A_2 deviation sets. There are not as many deviations involving reordering within the same state, so for the A_3 through A_8 deviations (the population-density changing deviations), all elements of the sets were included in the analysis.

Table 10 reports summary statistics for each deviation set. The third column reports the mean value of $\Delta\tilde{\Pi}(a)$ for each set. As explained earlier, this is the mean measured present value operating profit difference from following the actual policy rather than the deviation. Note that the units of $\Delta\tilde{\Pi}(a)$ are such that if a store opening in 1992 is interchanged with store opening in 1997, then it is the present value as of 1992 of (inflation adjusted) operating profits. Analogously, ΔD is the present value of the difference in year-distribution miles from doing the actual rather than a deviation (as of 1992 if it is a 1992/1997 switch).

Looking at Store-Density Decreasing deviations (type 1), we see that the mean measured present value difference from doing the actual rather than the deviation is -\$3.8 million. The negative value means that on average Wal-Mart was losing operating profit by doing what it did rather than a deviation. So random deviations to disperse stores raises the present value of profits by \$3.8 million. But note these deviations come at a cost, since mean $\Delta D = -1740$, i.e. present value year-distribution miles are higher by 1,740 on average.

Next consider Store-Density Increasing deviations (type 2). By following the actual policy a_0 rather than deviate this way, Wal-Mart enjoyed $\Delta\Pi = 4.8$ more in operating profits then it would on average obtain from these deviations. But if it deviated in this way, store-year distribution miles would have declined on average by 1,201, so again there is a tradeoff, but in a different direction.

For now, ignore the fixed cost from population density by assuming $\omega_1 = 0$ and $\omega_2 = 0$. We can then use the first two moment inequalities to bound τ. Since Wal-Mart chose a_0, the expected value of the first moment inequality must be nonnegative at the true τ,

$$\begin{aligned} E[M_1] &= E[\Delta\Pi_1] - \tau E[\Delta D_1] \\ &= -3.8 + \tau 1740 \\ &\geq 0 \end{aligned}$$

This implies that

$$\tau \geq \hat{\tau}_{lower} = \frac{3.8}{1740} = .0218(\ \$\text{million}) = \$2,180. \tag{18}$$

This says that the cost of one store distribution mile (per year) is at least \$2,180. Analogously, we can use type 2 deviations to derive an upper bound of

$$\tau \leq \hat{\tau}^{upper} = \frac{4.8}{1201} = .0400(\$\text{million}) = \$4,000.$$

These findings are intriguing but incomplete as they don't take into account the population density components of fixed costs (ΔF_1 and ΔF_2). Note that by doing the actual policy a_0 rather than the type 1 or type 2 deviations, Wal-Mart tended to stay out of higher population density locations longer (ΔF_1, the difference in present value log population density from doing a_0, is negative). If Wal-Mart is putting weight $\omega_1 > 0$ on this, then this is another reason why Wal-Mart was willing to give up operating profits by doing a_0 rather then deviate in the type 1 way. Inequality (18) gives all of the credit to τ in getting moment equality one to hold. If we increase ω_1, we can lower τ and the moment inequality still holds. Analogously, increasing ω_1 puts slack into $E[M_2]$ permitting an increase in $\hat{\tau}^{upper}$. Now if we are free to pick any ω_1 and ω_2, then moments 1 and 2 would put no restriction on τ. This is where moments 3 through 8 come in. With these deviations, the stores being flipped are in the same state so the impact on distribution miles is minimal, as can be seen in the ΔD column in the table. These deviations do change population density and so put discipline on the choice of parameters ω_1 and ω_2.

Table 11 presents the results of moment inequality estimation when the full set of constraints is imposed. (There are eight basic moments plus 6 interaction moments for each for a total of $56 = 8 \times (1+6)$ moments.) The linear programming problem of minimizing or maximizing τ subject to the constraints is solved. The exercise is conducted with all store openings and then a breakdown by whether the early opening store in the deviation pair was opened after 1990 or before. The period after 1990 is particularly interesting because that is when supercenters began to open. The estimated range is narrowest for this later period, with a lower bound of \$1,780 and an upper bound of \$5,190.

For each estimation, four sets of 95 percent confidence intervals are reported. The first two sets follow PPHI's procedure for the linear case and the reader should refer to PPHI for specifics. Set 1 corresponds to what PPHI call the *inner* approximation. It is constructed by first simulating draws from the empirical distribution of the four components of each of the 56 moments. (Note each moment (17) is a linear combination of four components with coefficients $(1, -\tau, -\omega_1, -\omega_2)$). Next the full linear programming estimation over all 56 constraints is implemented with the simulated data, for both the lower bound and the upper bound. Table 11 reports the 2.5 and 97.5 percentiles of the estimated bounds across the simulations. Set 2, the *outer* approximation, is simulated for the lower bound as follows. Simulate moment components as in set 1. Then take the estimated lower bound $\hat{\theta}_{lower} = (\hat{\tau}_{lower}, \hat{\omega}_{1,lower}, \hat{\omega}_{2,lower})$ and the simulated components to create "fitted values." Add a simulated error to the fitted values and estimate a lower bound on θ, using only the moments that were binding in the calculation of the original $\hat{\theta}_{lower}$.[17] As explained in Ho (2007), the outer confidence interval (Set 2) is asymptotically conservative. In all three cases, the 2.5 percentile of the lower bound estimate is lower for Set 2 than Set 1, while the 97.5 percentile of the upper bound is higher for Set 2 than Set 1. The notable point to be made about the confidence intervals is that the 2.5 percentiles of the lower bound confidence region are of similar magnitude as the point estimates of the lower bounds. Likewise, the

[17] Luttmer (1999) discusses a related approach that focuses on the binding moments.

97.5 percentiles of upper bound confidence regions are of similar magnitude as the point estimates of the upper bound.

The PPHI confidence interval theory assumes the error terms are independent across deviations. This assumption is not valid here. The error term arises here from store-level measurement error; two alternative deviations may involve the same store and therefore share a common error component. Put in another way, the estimates are based on the 160,000 deviations listed in Table 10. But these are constructed from pairs of 5,000 store openings (general merchandise and food) so there are at best 5,000 draws underlying the analysis, not 160,000. A formal asymptotic theory for this case has not been developed. To get a sense of the importance of the issue, the following exploratory simulation exercise was conducted. (The description here is brief; see Appendix B for details.) For each simulation, a bootstrap sample of deviations was drawn. A lower bound was estimated using the simulated data with only the binding moments as in Set 2. Analogously, an upper bound was estimated with the binding moments. Each deviation was given an error term drawn in two ways. In the first—used to create Set 3—the errors terms were drawn independently for each deviation. In the second—used to create Set 4—the error terms were drawn at the level of each store, so two deviations with a common deviating store do not have independent error terms. Sets 3 and 4 are otherwise created the same way. Of course, the confidence intervals when dependence is taken into account (Set 4) are wider than when it is not (Set 3). In particular, the 2.5 percentile for the lower bound with the 1990s data goes from \$1,640 to \$940, while the 97.5 percentile for the upper bound goes from \$5,250 to \$5,520. Nevertheless, it is encouraging that the estimates remain the same order of magnitude, despite the vast differences in the underlying number of independent draws.

7 Discussion of Estimates

In the previous section, the tightest interval for the estimate of τ is obtained for the sample of store openings in the 1990s where the lower bound on τ is \$1,780 and the upper bound is \$5,190. The unit of τ is a distribution mile year. If a store were right next to a distribution center rather than 100 miles away, over the course of a year, the savings would be at least \$178,000 and no more than \$519,000. If all 5,000 Wal-Mart stores (here supercenters are counted as two stores) were each 100 miles further from their distribution centers, costs would go up almost a billion dollars at the lower bound estimate.

To get a sense of the direct cost of trucking, I have talked to executives in the discount industry and have been quoted a marginal cost estimate of \$1.20 per truck mile for "in house" provision. If a store is 100 miles from the distribution center (200 miles round trip) and if there is a delivery every day in a year, then the trucking cost is \$1.20*200*365=\$85,400. This is less than half of the lower bound and sixteen percent of the upper bound. I conclude the value of proximity extends well beyond savings in trucking. The difference between the lower bound \$178,000 and the direct trucking figure of \$85,400 leaves much room for Wal-Mart to place a high valuation on the ability to quickly respond to demand shocks. My industry source on trucking costs emphasized the value of quick turnarounds as an important plus factor beyond savings in trucking costs.

A second perspective on the τ parameter can be obtained by looking at Wal-Mart's choice of when to open a distribution center (DC). An in-depth analysis of this issue is beyond the scope of this paper but some exploratory calculations are useful. Recall that problem (2) held fixed DC opening dates and considered deviations in store opening dates. Now hold fixed store opening dates and consider deviations in DC openings. Denote t_k^{open} to be the year DC k opens. Define $D_{k,t}^{inc}$ to be DC k's *incremental contribution* in year t to reduction in store-distribution miles. This is how much higher total store-distribution miles would be in year t if distribution center k were not open in that year. Assume there is a fixed cost ϕ_k of operating distribution center k in each year. Optimizing behavior implies that the following inequalities must hold for the opening year $t = t_k^{open}$,

$$\begin{aligned} \phi_k &\leq \tau D_{k,t}^{inc} \\ \phi_k &\geq \tau D_{k,t-1}^{inc} \end{aligned} \tag{19}$$

The first inequality says that the fixed cost of operating the distribution center in year t_k^{open} must be less than the distribution cost savings from it being open.[18] Otherwise, Wal-Mart can increase profit by delaying the opening by a year. The second inequality states that the fixed cost exceeds the savings of opening it the year before (otherwise open it a year earlier). Now if $D_{k,t}^{inc}$ changes gradually over time then (19) implies $\phi_k \approx \tau D_{k,t}^{inc}$ holds approximately at the date of opening. (Think of this as a first-order condition.)

Table 12 reports the mean values of the $D_{k,t}^{inc}$ statistic across distribution centers. The statistic is reported for the year the distribution center opens, as well as the year before

[18] There is also a marginal cost involved with distribution. But assume this is the same across distribution centers. So shifting volume across distribution centers doesn't affect marginal cost.

opening and the two years after opening. For example, to calculate this statistic in the year before opening, the given DC is opened one year early, everything else the same, and the incremental reduction in store miles is determined. I also report how the mean number of stores served varies when the DC opening date is moved up or pushed back. At opening, the mean incremental reduction in store miles of a distribution center is 5,809 miles and the DC serves 52 stores. The later the DC opens, the higher the incremental reduction in store miles and the more stores served. This happens because more stores are being built around it.

If we knew something about the fixed cost, then the condition $\phi_k \approx \tau D^{inc}_{k,t}$ provides an alternative means of inferring τ. A very rough calculation suggests a "ballpark" fixed cost of \$18 million a year.[19] Since the mean value of $D^{inc}_{k,t}$ when DCs open equals 5,809 miles, we back out an estimate of τ equal to

$$\hat{\tau} = \frac{\phi}{meanD^{inc}_{k,t}} = \frac{\$18 \text{ million}}{5,809} = \$3,098.$$

This is roughly midway between the lower bound estimate \$1,780 and the upper bound estimate of \$5,190. It is encouraging that these two approaches—coming from two very different angles—give consistent results.[20]

8 Concluding Remarks

This paper examined the dynamic store-location problem of Wal-Mart. Using the moment inequality approach outlined in PPHI, the paper was able to bound a preference parameter relating to the benefits Wal-Mart obtains when stores are close to distribution centers. The paper illustrates the power of this type of approach in getting a sensible analysis out of what would otherwise be complex and likely intractable.

While the analysis is rich in many dimensions—notably in its fine level of geographic detail and in the way it incorporates numerous data objects—it has limitations. One

[19] Distribution centers are on the order of one million square feet. Annual rental rates including maintenance and taxes are on the order of \$6 per square foot, so \$6 million a year is a rough approximatation for the rent of such a facility. A typical Wal-Mart DC has a payroll of \$36 million. If a third of labor is fixed cost, then we have a total fixed cost of \$18=\$6+\$12 million.

[20] An intriguing pattern in Table 12 is that the mean miles saved at opening for food operations is approximately half that for general merchandise. If the fixed cost of these kinds of operations are similar, this suggests that the mile cost is twice as high for food then general merchandise. Given issues of perishability, this is plausible, and warrants further study.

limitation is that all economies of density are channeled through the benefits of stores being close to distribution centers. Economies of density can be potentially enjoyed through other channels, including: *management* (it is easier for upper-level management to oversee a given number of stores when the stores are closer together); and *marketing* (satisfied Wal-Mart customers might tell their friends and relatives on the other side of town about Wal-Mart—this benefits Wal-Mart only if it has a store on the other side of town). A caveat then is that my estimate of τ may be picking up some economies of density from management and marketing. I have chosen to focus on *distribution* both because (1) I can measure it (i.e. the locations of distribution centers) but can't measure management and marketing activities and (2) my priors tell me distribution is very important for Wal-Mart. My findings in the previous section with DC opening dates that were consistent with my baseline findings are particularly helpful here. DC openings should be unrelated to management and marketing sources of economies of density.

A second limitation of the paper is that it leaves out oligopolistic interaction. Motivated by the seeming constancy of Wal-Mart's location behavior over several decades and across market segments (general merchandise and food), I have taken a decision-theoretic approach as a first step. As a second future step, I am very interested in expanding the analysis to incorporate oligopolistic interaction with K-Mart and Target. The moment inequality approach appears very promising for this purpose.

Wal-Mart has attracted much attention and various interest groups, particularly labor unions, have attempted to slow its growth, e.g. by trying to get local governments to use zoning restrictions to block entry of stores. These kinds of policies limit store density. The analysis here is not at the stage where it is possible to run a policy experiment to evaluate the welfare effects of limiting Wal-Mart's growth. Among other things, that would require uncovering how such limits would impact Wal-Mart's DC network and (except briefly at the end of the paper) this has been held fixed in the analysis. Nevertheless, the estimates of this paper suggest any policy that would substantially constrain store density—that in turn had the effect of reducing DC density—would result in significant cost increases.

Appendix A: Data

The selected data and programs used in the paper are posted at www.econ.umn.edu/~holmes/research.html. In particular, data on facility locations and opening dates are posted there. The store-level sales and employment data used in this paper can be obtained from TradeDimensions.

Facility Locations and Opening Dates

The data on facility locations and opening dates and other data for the project is posted at www.econ.umn.edu/~holmes/research.html.

The data for Wal-Mart stores were constructed as follows. In November 2005, Wal-Mart a file which listed for each Wal-Mart store the address, store number, store type (supercenter or regular store) and opening date. This data was combined with additional information posted at Wal-Mart's web site about openings through January 31, 2006 (the end of fiscal year 2006). The opening date mentioned above is the date of the original store opening, not the date of any later conversion to a supercenter. To get the date of supercenter conversions, I used two pieces of information. Wal-Mart's website posts information about store openings 2001 and after and conversions are announced as store openings. To get the dates of conversions taking place before 2001, I used data collected by Emek Basker (see Basker (2005)) based on published directories.

The data on Wal-Mart's food distribution centers (FDC) are based on reports that Wal-Mart is required to file with the EPA, as part of a *Risk Management Plan.*[21] (The freezers at FDCs use chemicals that are potentially hazardous.) Through these reports, all FDCs are identified. The opening dates of most of the FDCs were obtained from the reports. Remaining opening dates were obtained through a search of news sources.

Data on Wal-Mart general distribution centers (GDC) that handle general merchandise were cobbled together from various sources including Wal-Mart's annual reports and other direct Wal-Mart sources, Mattera and Purinton (2004) and various web and news sources. Great care was taken to distinguish GDCs from other kinds of Wal-Mart facilities such as import centers and speciality distribution centers such as facilities handling internet purchases.

The longitude and latitude of each facility was obtained from commercial sources and manual methods.

In the analysis, I aggregate time to the year level, where the year begins February 1

[21] The EPA data is distributed by the "Right-to-Know Network" at http://www.rtknet.org/.

and ends January 31 to follow the Wal-Mart fiscal year. January is a big month for new store openings (it is the modal month) and February and March are the main months for distribution center openings. New January-opened stores soon obtain distribution services from new February-opened DCs. To be conservative in not overstating the number of distribution store miles, I assume that the flow of services at a DC begins the year prior to the opening year. Put differently, in my analysis I shift down the opening year of each DC by one year.

Wage Data

The wages are average retail wage in the county containing the store. (Payroll divided by March employment.) The source is County Business Patterns (1977, 1980, 1985, 1990, 1995, 1997, 2000, 2002, 2004) with interpolation in intervening years and years with missing values. For 2000 and beyond the NAICS definition of the retail sector *does not* include eating and drinking establishments. For 1997 and earlier, the SIC code definition of retail *does* include eating and drinking and these are subtracted out for these years.

Property Value Data

For each store location and each Census year, I created an index of residential property value as follows. I identified the block groups within a two-mile radius of the store and called this the neighborhood. Total property value in the neighborhood was calculated as the aggregate value of owner-occupied property plus one hundred times monthly gross rents of renter-occupied property. This was divided by the number of acres in a circle with radius two miles and the CPI was used to convert it into 2005 dollars.

County property tax records were obtained from the web for 46 Wal-Mart locations in Minnesota and Iowa. (All stores in these states were searched, but only for these locations could the records be obtained.) Define the *land-value/sales ratio* to be land value from the tax records as a percent of the (fitted) value of 2005 sales for each store. For these 46 stores, the correlation of this value with the 2000 property value index is .71. I regressed the land-value/sales ratio on the 2000 property value index without a constant term and obtained a slope of .036 (standard error of .003). The regression line was used to obtain fitted values of the land-value/sales ratio for all Wal-Mart stores.

Appendix B

Simulation Exercise for Set 3 and Set 4 Confidence Intervals

This appendix discusses the procedure used to construct the Set 3 and Set 4 confidence intervals in Table 11. In the PPHI procedure, random draws take place at the level of a moment which aggregates deviations. In what I do here, I need to make draws at the level of a deviation, because I need to use the store information to connect the error structure of deviations that share stores. So the first step in a given simulation is to take a bootstrap sample of deviations. For example, there are 163,061 deviations used for the whole sample, so a bootstrap sample of 163,061 deviations is drawn from this. An error term η_a is drawn for each of the bootstrap deviations a as will be explained. Define $\tilde{y}_{lower,a} = x_a\hat{\theta}_{lower} + \eta_a$, where $\hat{\theta}_{lower} = (\hat{\tau}_{lower}, \hat{\omega}_{1,lower}, \hat{\omega}_{2,lower})$, for $\hat{\tau}_{lower}$ in Table 11 and $\hat{\omega}_{j,lower}$ the value of ω_j where the minimum $\hat{\tau}_{lower}$ is obtained. The procedure of estimating the lower bound is applied to the simulated data points $(x_a, \tilde{y}_{lower,a})$, imposing only the moment inequalities that are binding in the estimate of θ_{lower}, analogous to the PPPI outer approximation. The analogous procedure is applied to the upper bound.

Now I explain the draws of η_a. For each deviation a, a store i opening earlier than some j switches with j. It is convenient here to call this deviation ij instead of a and to let η_{ij} be the error term. Suppose we focus on a binding moment k. Then for this moment

$$E_k \left[z_{k,ij}(y_{ij} - x_{ij}\underline{\theta} + \eta_{ij}) \right] = 0$$

for instrument z_{ij}. Now since $E\left[\eta_{ij}|y_{ij},x_{ij}\right] = 0$,

$$E\left[z_{k,ij}\eta_{ij}\right]^2 \leq E\left[z_{k,ij}\left(y_{ij} - x_{ij}\underline{\theta} + \eta_{ij}\right)\right]^2$$

Or

$$E[\eta_{ij}]^2 \leq \zeta^2 \equiv \min_{\{k,\text{ binding moments}\}} \frac{E\left[z_{k,ij}\left(y_{ij} - x_{ij}\underline{\theta} + \eta_{ij}\right)\right]^2}{E[z_{k,ji}^2]}$$

Let $\hat{\zeta}^2$ be the sample analog of ζ^2. To be conservative, let $\hat{\zeta}^2$ be an estimate of $\sigma_\eta^2 = E[\eta_{ij}]^2$. For Set 3, the η_{ij} are drawn i.i.d. from $N(0, \hat{\zeta}^2)$. For Set 4, for each store i, draw ε_i i.i.d. from $N(0, \frac{1}{2}\hat{\zeta}^2)$. Then define $\eta_{ij} \equiv \varepsilon_i - \varepsilon_j$.

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, Patrick, Benkard, C. Lanier., and Jonathin. Levin, "Estimating Dynamic Models of Imperfect Competition" *Econometrica*, Vol. 75, No. 5, 1331–1370 (September, 2007).

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" *Review of Economics and Statistics* 87:1 (February 2005) 174-183

Basker, Emek, "The Causes and Consequences of Wal-Mart's Growth, *Journal of Economic Perspectives*, 21:3 (Summer 2007) 177-198.

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

Caves, Douglas W. ; Laurits R. Christensen; Michael W. Tretheway, "Economies of Density versus Economies of Scale: Why Trunk and Local Service Airline Costs Differ," *The RAND Journal of Economics*, Vol. 15, No. 4. (Winter, 1984), pp. 471-489.

Ellickson, Paul E., "Does Sutton Apply to Supermarkets?" *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.

Ho, Katherine (2007), "Insurer-Provider Networks in the Medical Care Market" NBER Working Paper #11822.

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

Jia, Panle, "What Happens When Wal-Mart Comes to Town: An Empirical Analysis of the Discount Retailing Industry" MIT Manuscript, 2007

Luttmer, Erzo G.J., "What Level of Fixed Costs Can Reconcile Consumption and Stock Returns?," Journal of Political Economy, 1999, Vol. 107, No. 5, pp. 969-997

Mattera, Philip and Anna Purinton, (2004) "Shopping for Subsidies: How Wal-Mart Uses Taxpayer Money to Finance Its Never-Ending Growth," manuscript, *Good Jobs First*

Manski, C (2003), *Partial Identification of Probability Distributions*, Springer: New York

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, November, 2006.

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,

U.S. Bureau of the Census, *County Business Patterns*, (1977, 1980, 1985, 1990, 1995, 1997, 1998, 2000, 2002, 2004).

U.S. Bureau of the Census, *Summary File 3, Census of Population and Housing*, (1980, 1990, 2000).

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

Wal-Mart Stores, Inc. (1971-2006). *Annual Report*.

Walton, Sam (1992) *Made in America: My Story* New York: Doubleday

Table 1
Summary Statistics of Store-Level Data
(End of 2005, Excludes Alaska and Hawaii)

Store Type	Variable	N	Mean	Std. Dev	Min	Max
All	Sales ($millions/year)	3,176	70.5	30.0	9.1	166.4
Regular Wal-Mart	Sales ($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 Wal-Mart	Employment	1,196	123.5	40.1	57.0	410.0
SuperCenter	Employment	1,980	333.8	91.5	31.0	812.0

Source: TradeDimensions Retail Database.

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

	General Merchandise (Including Supercenters)		Food Store (Part of Supercenter)		General Distribution Centers		Food Distribution Centers	
Decade Open	Opened this decade	Cumulative	Opened this decade	Cumulative	Opened this decade	Cumulative	Opened this decade	Cumulative
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	15	43	26	35

Source: See Appendix.

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 Young (21 and below)	0.35	0.31	0.31
Share Black	0.10	0.13	0.13

Source: Summary File 3, Census of Population and Housing (1980, 1990, 2000)

Table 4
Parameter Estimates for Demand Model

Parameter	Definition	Unconstrained	Constrained (Fits Reported Cannibalization)
λ^g	General Merchandise Spending per person (annual in $1,000)	1.686	1.938
		(.056)	(.043)
λ^f	Food spending per person (annual in $1,000)	1.649	1.912
		(.061)	(.050)
ξ_0	Distance disutility (constant term)	.642	.703
		(.036)	(.039)
ξ_1	Distance disutility (coefficient on population density)	-.046	-.056
		(.007)	(.008)
α	Outside Alternative valuation parameters		
	constant	-8.271	-7.834
		(.508)	(.530)
	ln(mbar)	1.968	1.861
		(.138)	(.144)
	$\ln(\text{mbar})^2$	-.070	-.059
		(.012)	(.013)
	Per Capita Income	.015	.013
		(.003)	(.003)
	Share of block group black	0.341	.297
		(.082)	(.076)
	Share of block group young	1.105	1.132
		(.464)	(.440)
	Share of block group old	.563	.465
		(.380)	(.359)
γ	Store-specific parameters		
	store age 2+ dummy	.183	.207
		(.024)	(.023)
σ^2	measurement error	.065	.065
		(.002)	(.002)
N		3176	3176
SSE		205.117	206.845
R^2		.755	0.753
ln (L)		-155.749	-169.072

Table 5
Cannibalization Rates
From Annual Reports and in Model

Year	From Annual Reports	Demand Model (Unconstrained)	Demand Model (Constrained)
1998	n.a.	.62	.48
1999	n.a.	.87	.67
2000	n.a.	.55	.40
2001	1	.67	.53
2002	1	1.28	1.02
2003	1	1.38	1.10
2004	1	1.43	1.14
2005	1	1.27	1.00*

*Cannibalization Rate is imposed to equal 1.00 in 2005.
Source: Estimates from the model and Wal-Mart Stores, Inc (2004, 2006)

Table 6
Comparative Statics with Demand Model
(Uses Constrained Model)

Distance	Population Density (thousands of people within a 5 mile radius)						
(miles)	1	5	10	20	50	100	250
0	.999	.989	.966	.906	.717	.496	.236
1	.999	.979	.941	.849	.610	.387	.172
2	.997	.962	.899	.767	.490	.288	.123
3	.995	.933	.834	.659	.372	.206	.086
4	.989	.883	.739	.531	.268	.142	.060
5	.978	.803	.615	.398	.184	.096	.041
10	.570	.160	.083	.044	.020	.011	.006

Table 7
Distribution of Variable Input Costs
(Percentiles of Distribution are Weighted by Sales Revenue)

Estimated 2005 Labor Costs

Quartile	Store Location	Annual Payroll per Worker ($)	Wages as Percent of Sales
Minimum	Pineville, MO	12,400	4.5
25	Litchfield, IL	19,300	7.0
50	Belleville, IL	21,000	7.6
75	Miami, FL	23,000	8.3
Maximum	San Jose, CA	37,900	13.7

Estimated Land-Value/Sales Ratios (Expressed as a Percent)

Quartile	Store Location	Index of Residential Property Value Per Acre ($)	Land-Value as Percent of Sales
Minimum	Lincoln, ME	1,100	.0
25	Campbellsville, KY	32,100	1.2
50	Cleburne, TX	67,100	2.4
75	Albany, NY	137,300	5.0
Maximum	Mountain View, CA	1,800,000	65.0

Table 8
Incremental and Stand-Alone Values of New Store Openings
(All evaluated at 2005 Demand Equivalents)

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

	N	Incremental Sales	Incremental Operating Profit	Incremental Distribution Center Distance (miles)	Stand-alone Sales	Stand-alone Operating Profit
All	3,176	36.3	3.1	168.9	41.4	3.6
By State's Wal-Mart Age at Opening						
1-2	288	38.0	3.5	343.3	38.7	3.6
3-5	614	39.5	3.5	202.0	41.5	3.7
6-10	939	37.6	3.3	160.7	40.9	3.6
11-15	642	36.1	2.9	142.1	42.2	3.4
16-20	383	32.9	2.8	113.7	41.2	3.5
21 and above	310	29.5	2.4	90.2	44.4	3.6

Part B: Food (New supercenters)

	N	Incremental Sales	Incremental Operating Profit	Incremental Distribution Center Distance	Stand-alone Sales	Stand-alone Operating Profit
All	1,980	40.2	3.6	137.0	44.8	4.0
By State's Wal-Mart Age at Opening						
1-2	202	42.4	3.9	252.9	3.9	4.0
3-5	484	42.7	4.0	171.2	4.1	4.2
6-10	775	41.0	3.6	113.5	4.0	4.0
11-15	452	36.7	3.2	95.3	3.9	3.9
16-20	67	30.1	2.8	94.0	3.5	3.5

Table 9
Incremental Operating Profit Regression
2005 Demand Equivalents
Includes State Fixed Effects

	General Merchandise	Food
State's Wal-Mart Age at Opening		
3-5	-0.04	-0.10
	(.05)	(.07)
6-10	-0.30	-0.60
	(.05)	(.07)
11-15	-0.63	-1.10
	(.05)	(.08)
16-20	-0.76	-1.35
	(.06)	(.12)
21 plus	-1.32	
	(.06)	
log population density	5.78	6.39
	(.18)	(.31)
(log population density)^2	-0.26	-0.28
	(.01)	(.01)
R^2	.51	.50
N	3176	1980

Table 10
Summary Statistics Deviations by Type

Deviation Type	Description	Sample Number	Mean Values			
			$\Delta\tilde{\Pi}$ (in millions of 2005 dollars)	ΔD (miles)	ΔF_1 (log popden)	ΔF_2 (log popden^2)
1	Store-Density Decreasing	83,625	-3.8	-1,740.3	-2.1	-13.1
2	Store-Density Increasing	17,999	4.8	1,201.5	-4.8	-37.7
	Population-Density Changing					
3	Group 4 to Group 3	5,579	1.6	-9.9	3.6	34.8
4	Group 3 to Group 2	8,226	4.4	25.0	3.5	26.6
5	Group 2 to Group 1	12,176	5.0	-64.3	3.2	17.4
6	Group 1 to Group 2	10,182	-2.2	-53.3	-3.5	-19.8
7	Group 2 to Group 3	12,110	1.0	-89.7	-4.0	-29.9
8	Group 3 to Group 4	13,164	2.7	-28.1	-4.9	-47.3

Table 11

Estimated Bounds on Distribution Cost τ
Units are Dollars per Mile Year
(95 Percent Confidence Intervals in Parentheses Constructed Four Ways)

Sample of Perturbations Used for Estimate	Lower Bound	Upper Bound
All Openings (N=163,061)	724 #1(700,4810) #2(530,910) #3(700,740) #4(550,910)	6712 #1(2940,6990) #2(5560,8880) #3(6620,6800) #4(6120,7290)
Openings 1990 and beyond (N=48,558)	1780 #1(1820,4180) #2(1730,1830) #3(1640,1910) #4(940,2630)	5190 #1(1950,5310) #2(4370,7930) #3(5140,5250) #4(4850,5520)
Openings before 1990 (N=114,503)	590 #1(570,4660) #2(420,760) #3(560,620) #4(330,850)	8330 #1(6800,9180) #2(7000,13660) #3(8180,8510) #4(7340,9350)

Table 12
Mean Incremental Miles Saved and Stores Served for Distribution Centers
Across Alterative Opening Dates Including Actual

	One Year Prior to Actual	Actual Year Opened	One Year after Actual	Two Years After Actual
All Distribution Centers				
(N = 78)				
Mean Incremental Miles Saved	4439	5809	6719	7145
Mean Stores Served	24	52	58	63
By Type of DC				
Regional Distribution Centers				
(N = 43)				
Mean Incremental Miles Saved	6144	7709	8698	8866
Mean Stores Served	37	69	76	79
Food Distribution Centers				
(N = 35)				
Mean Incremental Miles Saved	2332	3462	4275	5020
Mean Stores Served	7	32	36	43

Figure 1
Diffusion of Wal-Mart Stores and General Distribution Centers

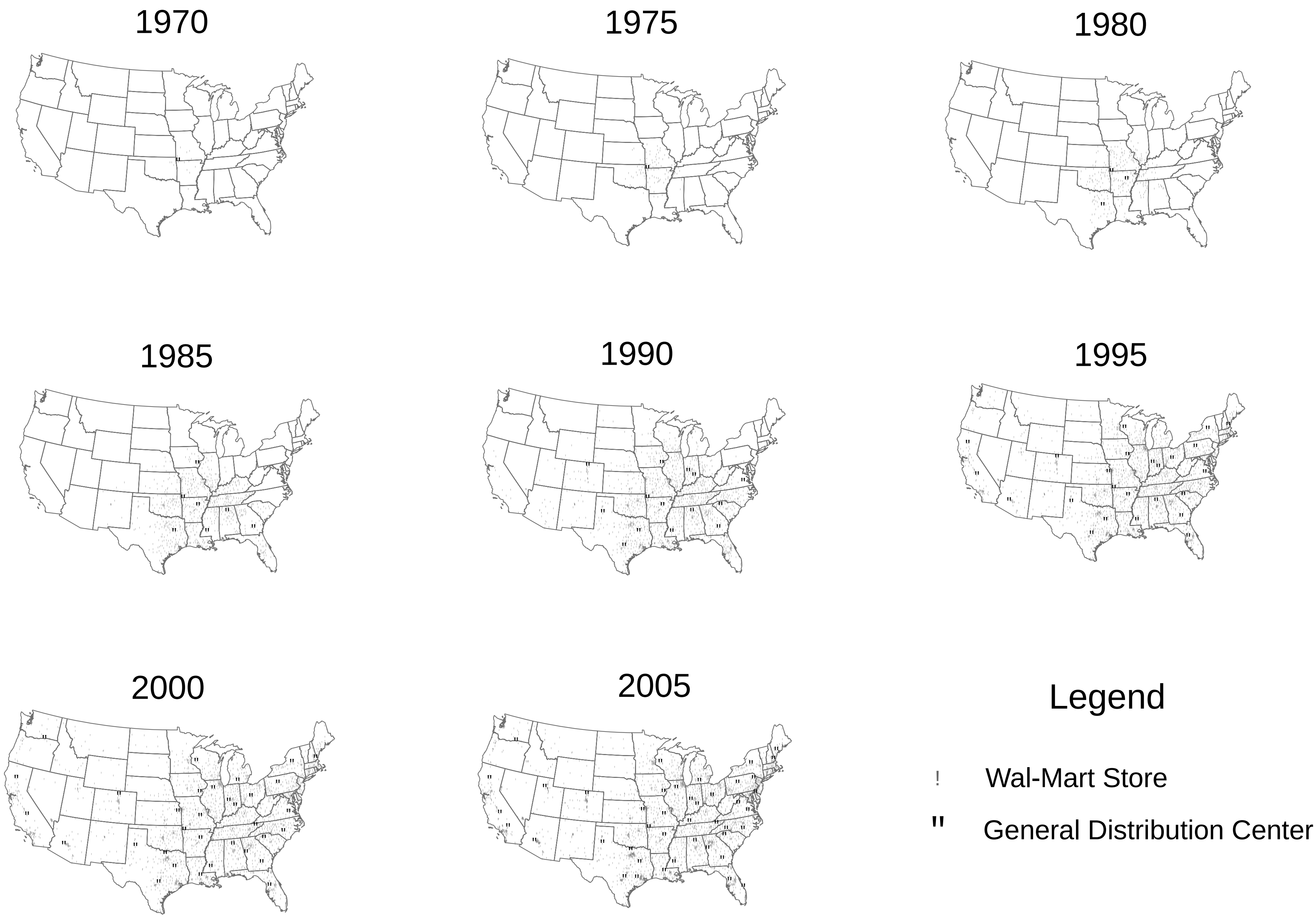

Figure 2
Diffusion of Supercenters and Food Distribution Centers

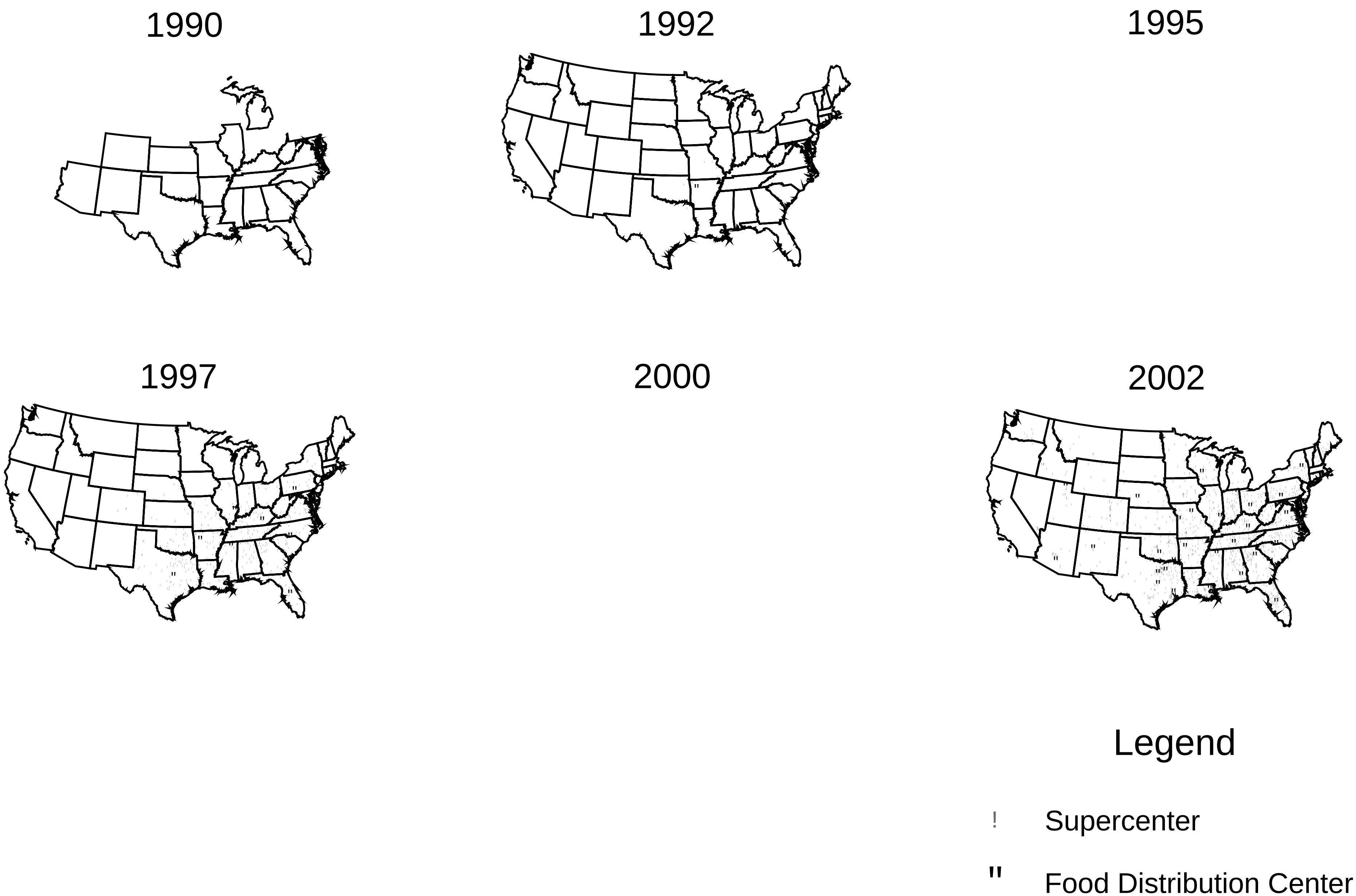

Figure 3
Sales per Employee as a Function of Square Footage of Store