The Diffusion of Wal-Mart and Economies of Density

by

Thomas J. Holmes[1]

University of Minnesota

Federal Reserve Bank of Minneapolis

and National Bureau of Economic Research

May, 2006

Preliminary and Incomplete Draft for Workshop Presentation

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

1. Introduction

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

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

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

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

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

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

The latter is what happened. Wal-Mart started with its first store near Bentonville, Arkansas, in 1962. The diffusion of store openings radiating out from this point was very gradual. Wal-Mart did not first grab the "low hanging fruit" in the most desirable location throughout the county and then come back for the "high hanging fruit," with fill-in stores. Desirable locations far from Bentonville had to wait to get their Wal-Marts.

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

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

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

In addition to contributing to the literature on economies of density, the paper also contributes to a new and growing literature about Wal-Mart itself (e.g., Basker (forthcoming), Stone (1995), Hausman and Leibtag (2005), Ghemawat, Mark, and Bradley (2004)), Neumark et al (2005), Jia (2005). Wal-Mart has had a huge impact on the economy. It has been argued that this one company contributed a non-negligible portion of the aggregate productivity growth in recent years. Wal-Mart is responsible for major changes in the structure of industry, of production, and in of labor markets. One good question is: what exactly is a Wal-Mart, why is it different from a K-Mart or a Sears? One thing that distinguishes Wal-Mart is its emphasis on logistics and distribution. (See, for example, Holmes (2001)). It is plausible that Wal-Mart's recognition of economies of density and its knowledge of how to exploit these economies distinguished it from K-Mart and Sears and is part of the secret of Wal-Mart's success.

2. Model

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

2.1 Demand

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

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

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

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

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

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

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

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

The first term is the utility decrease from travelling to the Wal-Mart j that is a distance $y_{\ell j}$ from the consumer's home. The weight $\tau\left(m_\ell\right)$ the consumer places on distance depends upon population density. This is another reduced form relationship; because of differences in the availability of substitutes induced by differences in population density, consumers in areas with high population density may respond differently with distance than consumers in low density areas. The second terms allows utility to depend upon other characteristics x_j of Wal-Mart store j. In the empirical analysis, the store-specific characteristic that I will focus on is store age. In this way, it will be possible in the demand model for a new store to have less sales, everything else the same. This captures in a crude way that it takes a while for new store to ramp up sales. The final two terms are random utility components specific to store j.

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

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

for

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

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

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

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

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

Total revenue of store j is

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

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

2.2 Cost Structure and Openings of New Stores

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

2.2.1 Operating Profit

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

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

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

for parameters ν_{Land} and ν_0.

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

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

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

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

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

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

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

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

$$C_{misc} = \nu_{Misc} Rev.$$

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

The operating profit equals gross margin less labor costs, land rentals, and miscellaneous costs,

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

2.2.2 Urbanization Costs

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

2.2.3 The Density Economies

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

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

$$s_j = -\frac{1}{\sum_k \exp(-\alpha y_{jk})}, \tag{6}$$

where y_{jk} is the distance in miles from store j to store k. Spillover takes values in the range $s_j \in [-1, 0)$. It is somewhat awkward to have the spillover always take negative values, but it simplifies notation in other respects. The spillover attains its minimum value of -1 for a store in isolation (In this case the denominator is one because $y_{jj} = 0$). As we increase the number of stores nearby, the denominator increases. As the number of neighbors within a fixed distance gets arbitrarily large, the denominator gets arbitrarily large and the spillover goes to its upper bound of zero.

The parameter α is a decay parameter. If $\alpha = 0$, then distance would be irrelevant, all that would matter would be the total number of stores. So there would be no economies of having stores close together. At the other extreme if $\alpha = \infty$, then $s_j = -1$ for all stores, so again economies of density are irrelevant. So the interesting case is $\alpha \in (0, \infty)$.

The profit generated by store s_j depends upon its spillover level s_j. Let ϕs_j denote the density economies enjoyed by store j. This is an additive term to profits. The level of density economies will be governed by α and ϕ.

2.2.4 Random profit term and fixed cost

In addition to all of the profit terms listed above, there is a random profit term ε_j for operating a store at location j. This is distributed i.i.d. across locations and has mean zero.

There is a fixed rental cost of operating a store f_t that depends upon time t in a continuous fashion. Assume it strictly decreases in time t. This captures in a crude way that everything else the same, Wal-Mart has a greater incentive to open more stores over time. This could be made endogenous and depend upon the number of stores opened to date. The f_t can be thought of as a Lagrange multiplier or shadow price of operating one more store at time t. For what I do, I won't have to go into the details, I can leave it as a black box. One strong restriction is that, endogenous or not, f_t is assumed to be continuous in t.

2.2.5 Dynamics

Everything that has been discussed so far considers quantities for a particular time period, i.e., revenues or fixed cost. I now explain the dynamic aspects of the model. It is convenient to work in continuous time. Let t index a particular instant of time.

Let A, a finite set, denote the set of all possible locations where Wal-Mart could conceivably open a store. Assume no exit so that once a store opens at a particular location j, the store remains open forever after. This is a good assumption for Wal-Mart since it rarely exits a location once it opens a store. Let $B_t \subseteq A$ be the set of locations with a store open at instant t. Let B_t^{new} be the set of store locations newly opened precisely at instant t. Let B_t°, be the set of stores open in the instant before t. Then $B_t = B_t^{\circ} + B_t^{new}$. Since the set of locations is finite, the set of dates on which stores are opened is finite. Let $\bar{t} = \{t^1, t^2, ...t^{last}\}$ be a vector of opening states. It may be on some of these opening dates more than one store is opened. Let $\bar{B} = \{B_{t^1}, B_{t^2}, ..., B_{t^{last}}\}$ be a particular sequencing of store openings. Let $r = \{\bar{B}, \bar{t}\}$ be a particular "rollout" plan. It provides a complete description of the sequencing and timing of store openings.

Suppose Wal-Mart discounts future payments with discount rate ρ (henceforth assumed to equal $\rho = .05$ at an annual rate). Suppose next there is an exogenous growth factor term g_t that varies continuously with time and is strictly increasing. This is meant to capture in a crude way the fact that Wal-Mart stores have gotten bigger as Wal-Mart has scaled up and added more things to its product line. I assume that all costs are scaled up in exactly this same way.

Take as given a roll-out plan r. The present value of the rollout plan is

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

To recap, in the outside we have the discount term $e^{-\rho t}$ as well as the growth factor g_t. We sum up across all B_t stores open at an instant t. The operating profit $\pi_j(B_t)$ of store j depends upon the entire configuration of all stores open at t, because stores may compete for the same sales. The next term ϕs_j are the density economies. The spillover s_j depends upon the stores opened. Next we subtract the urban cost at location j. We subtract the time-varying fixed cost. Last we add in the random profitability term. This is fixed. The growth factor g_t applies to this as well, but otherwise ε_j is constant over time.

2.3 Properties of a Solution

Wal-Mart's problem is to pick a rollout policy r to maximize $v(r)$. Let r^* be a solution. Perturbation methods can be used to characterize properties of a solution. Suppose in a solution r^* there store j is opened at time t and it is the unique store opening at that time. All variables in the objective function are continuous in time. Then because of continuity, it must be that the incremental profit of opening store j at instant t must be exactly zero, i.e.,

$$\Delta\pi(B_t, B_t^\circ) + \phi\Delta s(B_t, B_t^\circ) - c_{jt}^{urban} - f_t + \varepsilon_j = 0 \tag{7}$$

where

$$\Delta\pi(B_t, B_t^\circ) = \sum_{j\in B_t} \pi_j(B_t) - \sum_{j\in B_t^\circ} \pi_j(B_t^\circ)$$

is the change in operating profit and

$$\Delta s(B_t, B_t^\circ) = \sum_{j\in B_t} s_j(B_t) - \sum_{j\in B_t^\circ} s_j(B_t^\circ)$$

is the change in total spillovers added up over all stores. Note that the change in operating profit is not necessarily the same as the operating profit of store j, because the addition of this store may take sales from other stores. Equation (7) must hold with equality because if it were positive, the rollout r^* could not be optimal. Present value of profits could be increased by opening this store a little sooner. If negative, it would be better to wait.

In the case where multiple stores are opened at the same time things are slightly more complicated. Suppose at time t, the set of stores in B_t^{new} is greater than one. Consider a store $j \in B_t^{new}$. Consider a deviation where the is store is opened just a little earlier than at time t. The profit from this deviation cannot be positive, i.e.,

$$\Delta\pi(B_t^\circ + j, B_t^\circ) + \phi\Delta s(B_t^\circ + j, B_t^\circ) - c_{jt}^{urban} - f_t + \varepsilon_j \leq 0 \tag{8}$$

To understand this, note that with rollout r^*, just before time t the stores that are open is the set B_t°. With a deviation to open j early, the set becomes $B_t^\circ + j$. Next consider a deviation of opening store j a little after t. The condition that this deviation not increase profit is

$$-\Delta\pi(B_t, B_t - j) - \phi\Delta s(B_t, B_t - j) + c_{jt}^{urban} + f_t - \varepsilon_j \leq 0 \tag{9}$$

We can rewrite these as

$$\begin{aligned} \varepsilon_j &\leq \varepsilon_j^U \equiv -\Delta\pi(B_t^\circ + j, B_t^\circ) - \phi\Delta s(B_t^\circ + j, B_t^\circ) + c_{jt}^{urban} + f_t \\ \varepsilon_j &\geq \varepsilon_j^L \equiv -\Delta\pi(B_t, B_t - j) - \phi\Delta s(B_t, B_t - j) + c_{jt}^{urban} + f_t \end{aligned}$$

Next I consider perturbations that resequence the openings of stores, leaving the number of stores opened at a given time as fixed. Let r' be such a perturbation. The difference in value from doing r^* rather than r' must be weakly positive,

$$\begin{aligned} 0 &\leq v(r^*) - v(r') \\ &= \int_0^\infty e^{-\rho t} g_t \left(\Delta\pi(B_t^*, B_t') + \Delta D(B_t^*, B_t')\right) dt \\ &\quad + \int_0^\infty e^{-\rho t} g_t \left(\sum_{j \in B_t^*} \left[-c_j^{urban} + \varepsilon_j\right] - \sum_{j \in B_t'} \left[-c_j^{urban} + \varepsilon_j\right] \right) dt \end{aligned} \tag{10}$$

Note that with this perturbation, the fixed cost term drops out.

3. The Data and Some Facts

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

3.1 Data

There are five main data elements used in the analysis. The first element is store-level data on sales and other store characteristics that I have obtained from a commercial source. The second element is the exact date of store openings posted by Wal-Mart on its corporate web site. The third element is demographic information from the Census. The fourth is land price data for Wal-Mart stores obtained from tax records. The fifth element is data on how retail wages vary with population density from the Census.

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

Table 1 presents summary statistics of the TradeDimensions data for the 2,936 Wal-Marts in existence in the contiguous part of the United States as the end of 2003.[4] [5] (Alaska and Hawaii are excluded in all of the analysis.) As of the end of 2003, slightly over half of Wal-Mart's stores are supercenters. The average Wal-Mart racks up annual sales of $60 million. The breakdown is $42 million per regular store and $76 million per supercenter. The average employment is 223 and the average square feet is approximately 150,000.

The second date element is store opening dates provided by Wal-Mart itself. Wal-Mart has a relatively precise numbering system. Store #1 is the first Wal-Mart, store # 2 the second, and so forth. Wal-Mart frequently updates it stores, sometimes by remodeling and sometimes by just moving to a new facility a block or two away. When it relocates a facility, Wal-Mart keeps the same store number and does not consider this a new store. So the date is the opening of the original store, not the construction date of the remodeling or relocation. While it is commonplace for Wal-Mart to tear down a facility (or abandon it) and build a new one down the street, it is extremely rare for Wal-Mart to give up on a location completely and exit. This has happened on the order of 30 times over a 42 year period in which Wal-Mart has opened 3,000 stores (See Neumark et al.). Since it is negligible, I am going to ignore exit in the analysis and focus only on openings.

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

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

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

years using the CPI as the deflator.[6]

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

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

3.2 Facts about the Diffusion of Wal-Mart

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

Along its expansion path, Wal-Mart made choices along the way about priority locations. Some evidence of Wal-Mart's priorities can be obtained by looking at where they are at now. Table 4 presents information on the median distance to the nearest Wal-Mart across block groups. The median person in the United States is 4.2 miles from a Wal-Mart. People who live in rural areas tend to have to go the furthest. Population density in this paper is

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

measured as the number of people within a five-mile radius. In the most rural category in the table, less than 5,000 people within five miles, the median distance to a Wal-Mart is 14.3 miles. Wal-Mart is famous for bringing discount stores to small towns and there are many in small towns. But they are not very dense in small towns. So the typical rural person has to go 14.3 miles to a Wal-Mart which in many cases itself will be in a small town.

As density increases, median distance to the closest Wal-Mart decreases, until we get to heavily urbanized areas, those with over 200,000 people in a five mile radius. Wal-Mart is famous for avoiding very large cities; it still doesn't have a store in New York City. One issue with interpreting the increase in distance for the most urbanized areas is that the most urbanized areas tend to be on the coasts. Even if Wal-Mart wasn't purposely avoiding the largest cities, Wal-Mart might be less represented in big cities because they tend to be far from Arkansas. To examine this issue, I examine how the relationship between density and distance varies for different parts of the country. I use Census Divisions. For every Census Division (except the Mountain Division), the areas with the lowest median distance to Wal-Mart has density less than 200,000 people. This is strong evidence that Wal-Mart has at least some aversion to highly dense areas. This can be contrasted with the location behavior of Target and K-Mart illustrated at the bottom of Table 4. Distance to the closest Target and K-Mart decreases with population density. People who live in the most urbanized areas are closest to Targets and K-marts.

Wal-Mart's aversion to very dense areas can be used to learn something about the weight they place on density economies. If density economies did not matter while urbanization costs were huge, the following location strategy would be optimal. Build a dense network of stores in small and medium size towns. Initially skip the very large cities. Come back and fill in the very large cities later only after all the good locations in small and medium sized town have been taken. Note one can skip big cities and still have a very dense network of stores since big cities comprise a negligible portion of the landscape.

Table 4B shows that Wal-Mart did not use this strategy. It shows the date of the first Wal-Mart in a given state and for each of the large cities in the state the year that the median distance to Wal-Mart in the city was 10 or 5 miles to the closest Wal-Mart. (Cities here include the county containing the central city). Wal-Mart went to Little Rock and

Oklahoma right away, the first states in the expansion process. OK, these are not big cities. But St. Louis and Kansas City are and Wal-Mart got into these cities in 1989 and 1987, before it opened any stores in Michigan or Pennsylvania.

4. First Stage Parameter Estimation

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

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

4.1 Demand Estimation

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

$$\eta_j = \ln(Rev_j) - \ln(\hat{R}ev_j(\theta)).$$

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

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

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

for

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

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

The second issue is what to do about supercenters. As can be seen in Table 1, supercenter sales are almost twice as large as regular store sales. What is going on here is clear:

supercenters have a broader product line, so everything else the same we would expect supercenters to have larger sales. But this is not something that fits easily into the model just outlined. Even if I were able to put supercenters cleanly in the demand model, in my later analysis I would have the problem that I don't know the dates when a given supercenter was converted from a regular store, I only know store openings. (A large percentage of supercenter were once regular stores.) My project would be a lot simpler if Wal-Mart had never got into the supercenter format.

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

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

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

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

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

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

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

4.2 Labor Costs

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

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

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

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

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

4.3 Land Costs

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

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

4.4 Other Costs

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

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

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

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

4.5 Extrapolation to Other Years

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

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

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

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

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

5. Stage Two: Density Economies

It remains to pin down the density cost parameters as well as the urbanization cost parameters. I use perturbation approaches.

I consider three strategies. These strategies differ in complexity and in assumptions about error terms. In the first, I consider selected resequencings to increase operating profit. These strategies result in a more dispersed rollouts. I examine the terms of the tradeoffs involved to draw inferences about the magnitude of density economies. This approach is less complicated than the next two approaches.

The third strategy examines resequencings in a more comprehensive way than strategy one and allows for measurement error..

The third strategy uses the first-order conditions derived earlier. These conditions result on bounds on the random profit terms. I use these as the basis of orthogonality conditions.

5.1 Deviate to Increase Operating Profit

The idea of this approach is derive a lower bound on the importance of spillovers by considering alternative rollouts that deliver higher operating profits in present value terms. This approach requires that there is no structural error term, $\varepsilon_j = 0$, and that there is no measurement error for my estimates of operating profit derived in the first stage. These assumptions are relaxed in the alternative procedures discussed next.

One issue is urbanization costs. I skirt this issue here as follows. I assume that there is a critical population density $\overline{m}$, such that if $m < \overline{m}$, then $c^{urban}(m) = 0$. This is plausible. In very large cities, highly urbanized areas, Wal-Mart has to do very different things if it wants to open a store, a multi-story structure with escalators, limited parking, etc., all different than then standard Wal-Mart model. But in a medium-sized town, Wal-Mart can go to the outskirts of town and build a big box structure that would not look any different than what it would be able to do in a small town. For this section I assume that $\overline{m} > 20$ thousand per square mile.

With this assumption on urbanization costs in hand, I consider perturbations that re-sequence stores in areas with $m_j < 20$. Since none of these stores are in urban areas, I don't have to know the parameters of the urbanization cost to determine the effect of the perturbation on profit. Also, since the total number of stores is fixed at any point in time, I don't need to know the parameters of the fixed cost function either.

I fix a time interval $[\underline{t}, \bar{t}]$ over which I resequence stores. Ideally, I would select a resequencing of stores $m_j < 20$ in this time interval to maximize the present value of operating profits. This is a complicated given all the combinations and the fact that there is interactions in demand between stores. If there were no demand interactions and if demand stayed constant over time, the rollout that would maximize present value of operating profits would be easy to calculate. Just rank the store locations in operating profits and open them in the order of highest to lowest. In Wal-Mart's early years the amount of demand overlap was fairly small. So I use this sequential algorithm as an approximation.

Throughout this section I impose that $\alpha = .02$. This allows me to calculate the spillover. The question of interest is how big is ϕ.

I consider two time intervals, 1971-1980, and 1982-1989. Table 9 reports the results for the two cases. Resequencing over the period 1970-1980 reallocates 200 stores. The actual policy compared to the deviation has 106 million dollars less in operating profit, $\Delta\pi = -106$. But there is a trade-off as spillover is higher in the actual policy, $\Delta s = 15$. Using (10), optimality of the chosen path implies that the difference between the actual and the deviation is positive,,

$$\begin{aligned}
\Delta\pi + \phi\Delta s &\geq 0 \\
-106 + \phi 15 &\geq 0 \\
7.07 &= \phi^L_{1970-1980} \leq \phi
\end{aligned}$$

Analogously, using the openings over the period 1982-1989, we obtain a lower bound of

$$\phi^L_{1982-1989} = 22.07.$$

To put this number in perspective, a stand alone store with no neighbors has a spillover of -1 while a store with an infinity of close neighbors has a spillover of 0. So ϕ is the annual change in profit per year from such an extreme change in density. Of course, such extreme changes in density are not in the relevant range of data.

These calculations are interesting as a first cut. But the approach has problems if there is any measurement error for operating profits in the first states. Suppose $\pi_j(B_t)$ is the operating profit that the firm uses in its calculations while $\tilde{\pi}_j(B_t) = \pi_j(B_t) + \tilde{\varepsilon}^{measure}_j$ is what we see, where $\tilde{\varepsilon}^{measure}_j$ is classical measurement error. If measurement error is important, by resequencing on the basis $\tilde{\pi}_j(B_t)$ I tend to overstate the change in operating profit $\Delta\pi$ and hence overstate the bound ϕ^L.

5.2 Pairwise Resequencings

In this approach, I continue to assume the structural error term $\varepsilon_j = 0$, i.e. the profit component that the firm sees and makes decisions upon, but I allow for classical measurement error $\tilde{\varepsilon}^{measure}_j$. As in Pakes, Porter, Ho, Ishii, PPHI, I construct inequality moments in which the measurement error is averaged out.

Like Bajari and Fox (2005) and Fox (2005), I consider pairwise resequencing. That is, I take two stores opening on different dates and interchange the opening dates, leaving everything else fixed. Here I define *different dates* as opening at least two years apart. For this draft, I only consider stores opening over the period 1970-1980. Approximately 300 stores opened during this period. Examining openings that are two or more years apart, leaves me with 28,741 different perturbations. Table 1 reports that the mean value of $\Delta\pi$ and Δs across these perturbations is $\overline{\Delta\pi} = -.025$ and $\overline{\Delta s} = .183$.

Following PPHI, I consider subsets of the deviations, and average them, where the weights can depend on the decisions taken. Like the previous subsection, I avoid the issue of figuring out the urbanization costs by restricting attention to markets in which $m_j < \overline{m}$ so I don't need to know the urbanization cost to calculate the change in present values. Looking at cases where both cities in the pair are in markets with $m_j < 20$, there are 14,458 deviations. We expect the measurement error to average out to something close to zero, so that

$$\overline{\Delta\pi} + \phi\overline{\Delta s} \geq 0.$$

From this we obtain a bound of $\phi^L = .61$. This is not a very tight bound. I am basically considering random deviations here. I consider subsets of deviations which depend upon the decisions taken. Consider cases where the older store of the pair (the one opened first in the actual rollout) is closer to Wal-Mart's original store # 1. As we would expect, conditioning in this way, the mean $\overline{\Delta s}$ for such pairs is greater than the unconditioned level. Conditioning further that the older is in a smaller town (has a lower population density), $\overline{\Delta s}$ increases more and $\overline{\Delta\pi}$ gets more negative. So we are looking at cases where there is a trade-off. The older store is in a smaller market with lower operating profit, but has the advantage of being close to Wal-Mart's network. Conditioning upon this set, the upper bound increases to $\phi^L = 2.64$.

We can obtain an upper bound on ϕ by conditioning on cases where the older store has lower spillovers than the younger store, $\Delta s < 0$. As predicted by the theory, in these cases, the older store has higher operating profit. We obtain an upper bound for $\phi^U = 9.71$.

If we assume the threshold for urbanization costs is above 40, then we can expand the range of cities sizes over which deviations can be considered. The results are similar.

5.3 Opening Date FONC Approach

The third and final approach allows for the structural error term ε_j observed by Wal-Mart but not by us, as well as classical measurement error. I make assumptions about orthogonality conditions that the ε_j satisfy. I then use the inequality conditions (8) and (9) to derive moment inequalities.

To proceed, I will have to look at all stores, not truncate by city size as before. I need information about the urbanization cost and as well as the time-varying fixed cost. I begin with functional form assumptions. Assume

$$c_j^{urban} = \gamma_0 + \gamma_1 \ln \underline{m}_j,$$

where as before

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

Assume

$$f(t) = \tau_0 - \tau_1 t,$$

where $\tau_1 > 0$. With these function form assumptions, the bounds on the ε_j can be written

$$\begin{aligned} \varepsilon_j &\leq \varepsilon_j^U \equiv -\Delta\pi(B_t^\circ + j, B_t^\circ) - \phi\Delta s(B_t^\circ + j, B_t^\circ) + \gamma_0 + \gamma_1 \ln \underline{m}_j + \tau_0 - \tau_1 t \\ \varepsilon_j &\geq \varepsilon_j^L \equiv -\Delta\pi(B_t, B_t - j) - \phi\Delta s(B_t, B_t - j) + \gamma_0 + \gamma_1 \ln \underline{m}_j + \tau_0 - \tau_1 t \end{aligned}$$

As above, the decay factor $\alpha = .02$. Then the remaining parameters are $\theta = (\phi, \gamma_0, \gamma_1, \tau_0, \tau_1)$. The differences $\Delta\pi(B_t^\circ + j, B_t^\circ)$ are estimated in the first stage and the spillovers are pinned down by α and store distances and so can be directly calculated.

Assume that over a time interval $[\underline{t}, \bar{t}]$, for the stores that open in this period the ε_j are i.i.d. Now this assumption holds if we start out with a finite number of locations and set $[\underline{t}, \bar{t}] = (-\infty, \infty)$, if we assume that $\lim_{t\rightarrow\infty} f_t = -\infty$, and if we assume that ε_j is i.i.d. over this finite set of locations. Let $\underline{\overline{J}}$ be the set of stores opened over this time period. With this structure, all store locations are eventually opened. In my application I won't take this extreme approach as I will be looking a bounded interval of time $[\underline{t}, \bar{t}]$, so there is potentially selection issues about the stores opened before $\underline{t}$ and the stores opened after $\underline{t}$ My approach will ignore this issue. But I will be looking at a very long time period and *will* be taking into account the selection going on throughout the time period. Everything else the same, locations with higher ε_j will be opened earlier than locations with lower ε_j. My procedure takes this explicitly into account.

Let z_k be an instrument so that for stores $j \in \underline{\overline{J}}$,

$$E[\varepsilon_j z_k] = 0.$$

I use as my instruments (1) a constant term, (2) log neighboring population $ln(\underline{m}_j)$, (3) distance d_j^1 of store j from Wal-Mart's store #1 in Rogers, Arkansas. All take nonnegative values. Now the ε_j^L and the ε_j^U are random variables depending on the draw of the ε_j for $j \in \underline{\overline{J}}$. Since $\varepsilon_j^U \geq \varepsilon_j$ must hold and since z_k is nonnegative, we have

$$E[\varepsilon_j^U z_k] \geq 0. \tag{11}$$

Analogously,

$$E[\varepsilon_j^L z_k] \leq 0. \tag{12}$$

So with the three instruments, we have 6 moment inequalities. Following the literature, I look for parameters in which these inequalities are satisfied in the sample averages.

My main interest is in the parameter ϕ. I use the 6 moment inequalities to bound ϕ. To obtain a lower bound for ϕ, I minimize ϕ subject to (11) and (12) and parameter restrictions

$$\begin{aligned} \phi &\geq 0 \\ \gamma_0 + \tau_0 &\geq 0 \\ \gamma_1 &\geq 0 \\ \tau_1 &\geq 0. \end{aligned}$$

This is a standard linear programming problem. To obtain an upper bound for ϕ, I maximize ϕ subject to the same constraints. The results are presented in Table 11. I estimate the bounds using different time periods of the data. For example, using data for the period 1970 through 1989, the lower bound for ϕ is 1.45 and the upper bound is 2.14. For this case the bounds are relatively tight. The estimates here are roughly the same order of magnitude as the estimates with the alternative procedure.

REFERENCES

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

Table 1
Summary Statistics: TradeDimensions Data

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

Table 2
Distribution of Wal-Mart Stores by Year Open

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

Table 3
Summary Statistics for Census Block Groups

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

Table 4
Median Distance To Nearest Store across Census Blockgroups
By Population Density

	All	Maximum Population Density (1,000 within 5 miles) for Density Category						
		5	25	50	100	200	300	300 and above
Median Distance to Wal-Mart	4.2	14.3	7.2	3.7	3.2	3.0	3.3	4.7
By Division								
NEW ENGLAND DIVISION	4.6	14.9	7.5	4.7	3.8	2.9	4.2	7.0
MIDDLE ATLANTIC DIVISION	5.8	14.5	8.4	4.9	4.6	3.7	4.2	8.0
EAST NORTH CENTRAL DIVISION	4.3	12.7	7.9	3.7	3.5	3.2	3.4	4.5
WEST NORTH CENTRAL DIVISION	5.4	17.4	6.1	3.2	2.8	3.0	4.0	6.5
SOUTH ATLANTIC DIVISION	4.1	13.5	7.4	3.9	3.1	2.9	3.3	4.8
EAST SOUTH CENTRAL DIVISION	4.5	12.3	5.5	2.9	2.6	3.0	5.1	4.9
WEST SOUTH CENTRAL DIVISION	3.2	13.9	4.6	2.7	2.5	2.4	2.6	3.5
MOUNTAIN DIVISION	3.1	25.1	7.6	3.1	2.8	2.7	2.3	2.4
PACIFIC DIVISION	4.5	19.5	10.9	6.4	3.6	3.7	3.8	4.1
Median Distance to Target	5.0	34.8	21.3	10.2	4.5	3.3	2.8	2.9
Median Distance to K-Mart	5.8	25.5	15.0	7.0	4.8	4.0	3.8	3.5

Table 4B

State	City	Year of First Wal-Mart in State	Year avg distance in city to closest Wal-Mart is: <10 miles	<5 miles
Arkansas	Little Rock	1962	1967	1975
Oklahoma	Oklahoma City	1968	1981	1983
Missouri	St. Louis	1968	1984	1989
Missouri	Kansas City	1968	1981	1987
Tennessee	Memphis	1973	1985	1991
Tennessee	Nashville	1973	1981	1981
Kentucky	Louisville	1974	1984	1988
Texas	Harris	1975	1984	1990
Texas	Dallas	1975	1985	1988
Illinois	Chicago	1977	1992	2000
Alabama	Birmingham	1979	1984	1988
Georgia	Atlanta	1981	1988	
Florida	Miami	1982	1992	
Indiana	Indianapolis	1983	1990	2000
Minnesota	Minneapolis	1988	1992	
Ohio	Columbus	1988	1994	1999
Ohio	Cleveland	1988	1994	1998
Michigan	Detroit	1989	1998	
Pennsylvania	Philadelphia	1990	1994	2002
Pennsylvania	Pittsburgh	1990	1994	2003
California	Los Angeles	1990	1998	2002
New York	New York City	1991		
Massachusetts	Boston	1992	1995	
Washington	Seattle	1993	1996	

Table 5
Demand Parameter Estimates

Parameter	Definition	a	b	c
λ	scaling parameter	29.742	29.057	18.702
		(.055)	(.057)	(.057)
$\rho=1/(1-\sigma)$	correlation parameter	.781	.767	.959
		(.055)	(.057)	(.057)
τ_0	constant	.616	.621	.464
		(.054)	(.056)	(.031)
τ_1	population density within 5 miles	-.046	-.049	-.001
		(.047)	(.048)	(.016)
ω	constant	-7.769	-7.586	-10.517
		(.055)	(.057)	(.057)
	lnmaxc(neig5)	1.503	1.605	2.596
		(.054)	(.056)	(.058)
	lnmaxc(neig5)^2	-.027	-.037	-.140
		(.043)	(.045)	(.010)
	pcitrun	.023	.021	.018
		(.045)	(.046)	(.004)
	blackshr	.928	.909	.841
		(.055)	(.057)	(.057)
	youngshr	1.241	.881	.633
		(.055)	(.057)	(.057)
	oldshr	1.369	1.158	1.288
		(.055)	(.057)	(.057)
γ	store age 3- dummy		.246	
			(.057)	
	store age 3-5 dummy			.240
				(.062)
	store age 6-10 dummy			.319
				(.060)
	store age 11-20 dummy			.340
				(.057)
	store age 20- dummy			.225
				(.057)
σ^2	measurement error	.092	.090	.090
		(.055)	(.057)	(.003)
N		1457	1457	1457
SSE		134.746	131.039	130.554
R^2		.674	.683	.684
ln (L)		-333.020	-312.701	-309.9963

Table 6
Comparative Statics

Distance (miles)	Neig5 (thousands)						
	1	5	10	20	40	100	250
0	.999	.984	.957	.893	.766	.499	.244
1	.997	.973	.930	.839	.678	.402	.185
2	.995	.954	.890	.765	.576	.312	.138
3	.991	.923	.829	.669	.467	.234	.102
4	.984	.875	.745	.558	.361	.171	.074
5	.971	.803	.637	.440	.267	.122	.053
10	.596	.213	.122	.069	.039	.019	.010
15	.062	.018	.011	.007	.004	.003	.002
20	.003	.001	.001	.001	.000	.000	.000
25	.000	.000	.000	.000	.000	.000	.000

Notes: These values are evaluated when black ratio, young ratio, old ratio and PCI are set to be at their

Table 7
Average Retail Wages and Population Density

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

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

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

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

Table 9
Discounted Values over Deviation Intervals
(Present Value in Millions of 2003 Dollars as of the Beginning of the Interval)

Interval 1: 1971-1980

	Revenue	Operating Profit	Spillovers ($\alpha = .02$)
Actual Policy	14,965	1,413	-48
Resequence Small Stores to Maximize Operating Profit	15,950	1,519	-63
Gain from Actual	−985	−106	15

Interval 2: 1982-1989

	Revenue	Operating Profit	Spillovers ($\alpha = .02$)
Actual Policy	133,577	12,673	-270
Resequence Small Stores to Maximize Operating Profit	136,665	13,004	-285
Gain from Actual	−3,088	−331	15

Table 10
Pairwise Resequencings 1970-1980
All Resequencings Two or More Years Apart

	Number	Mean $\Delta\pi$	Mean Δs ($\alpha = .02$)	Implied Bound
All Resequencings	28,741	-.025	.183	
Only stores in small towns ($m_j < 20$)	14,458	-.115	.189	0.61 (lower)
Older Store Closer to Store #1	10,873	-.258	.255	1.01 (lower)
And Older store in Smaller Town	5,722	-.697	.264	2.64 (lower)
Spillovers Lower in Original	3,381	.272	-.028	9.71 (upper)
Only stores in small or medium towns ($m_j < 40$)	22,433	-.057	.188	0.30 (lower)
Older Store Closer to Store #1	16,745	-.197	.258	0.76 (lower)
And Older store in Smaller Town	8,170	-.992	.266	3.73 (lower)
Spillovers Lower in Original	5,293	.390	-.045	8.67 (upper)

Table 11
Estimated Bounds on Density Parameter φ
Opening Date FONC Approach

Time Period Used	Lower Bound	Upper Bound
1970-1989	1.45	2.14
1970-1979	.65	1.62
1980-1989	1.41	4294.57
1970-2004	3.62	3.62