

Wal-Mart versus the Little Guy:
Measuring Wal-Mart's Effect on Supermarket Prices

Matt Sveum

Masters Paper

Final Draft

May 11, 2009

1 Introduction

The 2008 Christmas shopping season was a rough one for retailers around the country. Just about all retailers - from the mom-and-pop shop on Main Street to the large nation-wide chains - felt the pinch of a slowing economy. But one brick-and-mortar store was seemingly immune from the downturn. According to Wal-Mart, the chain had sales almost \$31 billion during the month of December, up 4.3% from December 2007. For 2008 as a whole, Wal-Mart's sales increased 6.3%. Target, one of Wal-Mart's closest competitors, saw only a slight increase for December of 0.2%, and an increase of 2.4% for the year.

There is no question whether or not Wal-Mart is a force to be reckoned with in the discount retail industry. City councils around the country have debated letting Wal-Mart build a Supercenter within their borders. Local establishments often argue that Wal-Mart will have a detrimental impact on their business. They say that Wal-Mart may not only drive them out of business, but leave entire downtown districts void of customers. Is that true? What is the true extent of Wal-Mart's power over the marketplace? This paper seeks to replicate *The Evolving Food Chain: Competitive Effects of Wal-Mart's Entry into the Supermarket Industry* by Emek Basker and Michael Noel. Using data collected from the American Chamber of Commerce Research Association and Wal-Mart, I test Wal-Mart's ability to lower competitors' prices on the city level. Using three different estimation techniques, I find that Wal-Mart has a statistically significant impact on the prices charged by competitors. Extending the work done by Basker and Noel, I find that the geographic size of the city acts as a counter-effect on Wal-Mart's power but does not make it disappear until a city gets far larger than any city in the sample. I also test to see if Wal-Mart Supercenters that were converted from Discount Stores have the same effect as Supercenters that were built from scratch. I find that converted stores have a statistically significant effect while new stores do not. The two effects, however, are not statistically different.

The paper takes the following form: first, I review the history of Wal-Mart and then some previous, relevant literature on Wal-Mart and prices. I present the data used in the analysis in section 2. I replicate Basker and Noel's three basic models in section 3. I expand their analysis in section 4 and conclude in section 5.

1.1 Wal-Mart History

Sam Walton opened the first Wal-Mart store in Rogers, Arkansas in 1962. The next store opened two years later, and the third a year after that. From their Arkansas home base, Wal-Mart gradually expanded out. The first out-of-state location opened in Sikeston, Missouri in 1968, but only after six stores were opened in Arkansas. Following a clear pattern, Wal-Mart opened more and more stores by never opening a new store too far from an existing store (Holmes, 2008).¹ Wal-Mart became a publicly traded company in 1972 and was in 11 states with 276 stores by the time 1980 rolled around (walmartstores.com).

Until 1988, Wal-Mart had not ventured into the grocery business. While its Discount Stores may sell some basic grocery items, none of them have a full line of food. In 1988, Wal-Mart combined their Discount Store model with a grocery store when they opened their first Supercenter in Washington, Missouri. Wal-Mart followed the same outward growth pattern with their Supercenters as they did with their Discount Stores (Holmes, 2008). By the end of the 1980s, Wal-Mart had 1,402 stores and \$26 billion in sales (walmartstores.com).

Today, Wal-Mart is the largest retailer in the world and one of the world's largest employers. Wal-Mart now operates more than 3,000 international stores, including buying the United Kingdom chain ASDA in 1999. There are currently more than 4,100 Discount Stores, Supercenters, Neighborhood Markets, and Sam's Clubs² in the United

¹A video of year-by-year store openings can be found on Dr. Holmes' website.

²A Discount Store carries non-grocery items and a few select foodstuffs, such as cereal, milk, and snacks. A Neighborhood Market carries only grocery items. A Supercenter carries both grocery and non-grocery items. A Sam's Club is a membership club.

States (walmartstores.com). According to Nielsen, 200 million people shop at Wal-Mart each year (walmartstores.com).

1.2 Literature Review

Much of the Wal-Mart focused literature has focused on Wal-Mart's impact on prices. Researchers have looked at Wal-Mart's impact on competitors' prices in a number of different ways. Most have estimated simple OLS regressions, including various variables to account for price changes not attributable to Wal-Mart. Other differences hinge on data sets used.

Emek Basker published *Selling a Cheaper Mousetrap: Wal-Mart's Effect on Retail Prices* in 2005. She gathers data from the American Chamber of Commerce Research Association to investigate the impact that Wal-Mart's presence in a community has on prices. The ACCRA data in the paper cover 165 cities from across the United States and span 20 year (1982-2002). The data from ACCRA present a major limitation in this paper. Because prices data are averaged over a city, they may or may not include Wal-Mart in the sample. The goal is to estimate Wal-Mart's effect on other stores' prices, and having Wal-Mart in the sample might cause the estimates to be off. Basker chooses ten products that a typical Wal-Mart discount store is likely to carry, primarily drug store-type goods. She formulates a panel model that says that the average log price of product k in city j during time t is a function of the price in the previous period, city and quarter fixed effects, a linear trend, and a dummy variable indicating whether or not Wal-Mart is located in the city during that quarter.

Estimating the model for each of the ten products, Basker finds downward pressure from Wal-Mart for nine of the ten products. Detergent, shampoo and toothpaste are the only products with a significant coefficient, and cigarettes and underwear have an insignificant positive coefficients.

In *The Evolving Food Chain: Competitive Effects of Wal-Mart's Entry into the Supermarket Industry*, Basker and Michael Noel examine Wal-Mart's effect on prices using

three different estimation techniques. While the data still come from ACCRA, Basker and Noel obtained store-level data instead of the city-level data Basker used in her previous paper. This is a major change because it allows the authors to remove Wal-Mart's prices from the data set and only focus on non-Wal-Mart prices. This presents a much clearer picture of Wal-Mart's effect than the city-level data from ACCRA. Basker and Noel find that prices at Wal-Mart are about 10 percent less than prices at other stores. One other significant change to the data is that a store count is now used instead of a binary variable indicating Wal-Mart's presence. They start out using a simple cross-sectional estimation but reject the model because they say it places too much weight on Wal-Mart's presence in a market and not enough weight on other factors. This is partly because of an endogeneity problem surrounding store locations. They suggest a few instruments to help fix this endogeneity problem: distance from a distribution center to a store and the number of Wal-Mart discount stores in a city in a given base year. The first IV actually increases the bias due to its direct effect on prices. The second IV reduces bias, although Basker and Noel conclude that it is still not quite as good as their next two estimation techniques.

Basker and Noel next use a long-difference estimation, using the difference between the first year (2001) and the last year (2004) of their sample. They argue that this eliminates the endogeneity problem because all factors that do not change over time are eliminated. Their results show other stores' prices fall by about 1.2 percent, which is statistically different from zero. The problem that the authors have with this approach is that it does not use all of the available data.

The third and final approach that Basker and Noel use is a panel data model, and is the closest model to the one used in *Mousetrap*. They prefer this model because it uses all available data. This allows the timing of Wal-Mart's entry to be taken into account, which is not possible in their previous two models. Overall, Basker and Noel find a 1.1 percent decrease in competitors' prices. When they estimate each product separately, they find that 22 out of 24 products had negative coefficients, seven of which were

significant.

Jerry Hausman and Ephraim Leibtag look at the effect on prices of supercenters, mass merchandisers and club stores, not just Wal-Mart in a 2005 paper. Hausman and Leibtag use ACNielsen Homescan data as opposed to the ACCRA data Basker and Noel use. This data set comes from randomly selected homes, and records all food purchases made. It also records where the purchase was made. This allows the authors to account for different types of stores and measure how prices determine where people shop. Of the 20 goods Hausman and Leibtag look at, they find that the non-traditional stores have a lower average price than the traditional supermarkets for 19 of the goods (soda was the only good with a higher price). The model says that price in market i during time t is a function of market and month fixed effects, and the percentage of expenditures from supermarkets. They define price quite differently than Basker and Noel do, however. Price, in this case, is really more of a measure of expenditure on a certain category of goods. Whereas ACCRA carefully defines products down to size and brand, the Homescan data used here contain more widely defined categories. The difference in the definition of price also means that Hausman and Leibtag are measuring how consumers' expenditures change as their shopping patterns change while Basker and Noel measure how competitors' prices change when Wal-Mart expands in a city. These are related, yet quite different, estimations.

They find all twenty products have negative coefficients, meaning that as the percentage of people shopping at non-traditional stores increases, the average price paid falls. All coefficients are significant, although their magnitude varies widely. The authors state that they are not able to determine a theme among the coefficients. Branded and unbranded goods both see some large and some small coefficients. Overall, Hausman and Leibtag find that supercenters' prices are 5-48 percent lower than traditional grocery stores. This is significantly higher than the ten percent that Basker and Noel found. Basker and Noel attribute this difference to the data set used by the two teams of researchers.

2 Data

This paper uses three sets of data: price data, Wal-Mart store location and opening date data, and city data. The Wal-Mart and price data come from Basker and Noel, and city data come from the Census.

The price data used are from the American Chamber of Commerce Research Association (ACCRA) and were obtained from Emek Basker. ACCRA uses local Chambers of Commerce to calculate the average price of a number of goods each quarter. The purpose of this data is to calculate “a useful and reasonably accurate measure of living cost differences among urban areas” (ACCRA, 2008). ACCRA instructs the Chambers of Commerce that are collecting the data to collect prices only from stores where professional households typically shop. This means that, depending on the city, Wal-Mart may or may not be included in the sample. However, the data that I am using are store-level data, not the published average price data. I am only interested in the non-Wal-Mart observations, making it a moot point that ACCRA may not sample from Wal-Mart.

While ACCRA’s data includes prices for a wide variety of goods and services, ranging from dry cleaning to bananas, I am only concerned with products that are likely carried by Wal-Mart Supercenters. Table 1 lists all 24 products included in the sample along with ACCRA’s description and average price. I exclude products that Wal-Mart does not sell because I make the assumption that Wal-Mart should have no measurable effect on those products.

The second component I include is the Wal-Mart store data. The data come primarily from Wal-Mart, and are described in depth in Basker and Noel’s paper. Included in the data are all Wal-Mart Supercenter locations and the date the store opened as either new construction or as a remodeled Discount Store. This allows me to determine how many Supercenters were open in each city in each of the four years in the sample. I make no distinction between stores that were converted from existing Discount Stores

and stores that were new construction until section 4.3. Not every city in the sample has a Wal-Mart Supercenter during the sample period. Of the 175 cities in the sample, 63 cities have zero Wal-Mart stores, or about 36 percent. One store is the most common number of stores with 65 cities. Table 2 lists statistics for the cities in the sample based on the number of stores.

The final important component of the data set is the city data. Obtained from the Census Bureau, this includes information on the population, average income, and square mileage for each city. The average population in the 175 sample cities is 175,123 and the average income is just over \$35,000. The smallest city (Gunnison, CO) has 5318 residents while the largest city's (Houston, TX) population is 1,954,848. Cities also vary greatly in geographic size. The average size is 72.3 square miles, with the smallest city (Indiana, PA) being 1.7 square miles and the largest (Jacksonville, FL) being 757 square miles. While there is certainly a relationship between population and square mileage, it is not a perfect one. The correlation coefficient between the two variables is 0.775.

3 Model and Results

This paper sets out to determine the power that Wal-Mart has to lower prices at competing stores. To achieve this goal, I present three different models and a few extensions of these models. All three models exhibit the same basic structure, but use different aspects of the data set. The first model presented is the cross-sectional model, which produces a biased estimate. Next is the long-difference model, which is better than the cross-section, but not as good as the panel data model that I finish with. All three models in this section are replications of Basker and Noel's work. The models presented in section 4 are extensions of the models here.

3.1 Cross-Section

The simplest approach is to treat each year with a separate equation. This means that Wal-Mart determines store locations exogenously, an assumption that is most likely not true. Using 2002 as the year of choice, the model takes the form:

$$p_{jk,2002} = \alpha + \theta \text{WMSC}_{j,2002} + \sum_k \phi_k \text{product}_k + \mathbf{x}_j \beta + \epsilon_{jk} \quad (1)$$

where $p_{jk,2002}$ is the average price of product k in city j in 2002 at Wal-Mart's competitors, $\text{WMSC}_{j,2002}$ is the number of Wal-Mart Supercenters in city j in 2002, product_k is a product indicator variable to capture product fixed-effects, and \mathbf{x}_j is a vector of city-specific variables income and population. All non-dummy variables, with the exception of the store count, are used in their log forms.

I estimate the regression both with and without the \mathbf{x}_j city variable vector. I present my findings in the first two columns of table 3. Following Basker and Noel, the standard errors that are reported are clustered by city. While it is safe to assume that the errors are not correlated across cities, it is not safe to assume that there is no correlation between products within the same city. Many of the things that impact price but are unobservable affect a wide range of products, but are probably not present across cities.

These results show that prices are two to three percent lower for each Wal-Mart in a city. For both cases, the Wal-Mart coefficient is statistically significant. At first glance, this is significant because it implies that *each* Wal-Mart causes prices to fall at competing stores. These results, however, should be taken with a grain of salt. The nature of this model is bound to cause some bias in the results with this type of data. Wal-Mart does not choose locations in a vacuum, and the factors that influence their location decisions are also very likely to influence the price competitors are charging in a particular city (Basker and Noel, 2007). Because these factors are not measurable, and thus not included in the model, much of their effect is attributed to Wal-Mart. It also should not be surprising that these factors would inflate Wal-Mart's effect instead of mitigate it.

The results that I find closely resemble the results found by Basker and Noel. They find that Wal-Mart forces competitors' prices down by 2.53 percent without \mathbf{x}_j included and 3.17 percent with \mathbf{x}_j in the model. The differences are in the thousandths place in the estimated coefficient, and are not statistically significant. The difference most likely comes from differences in the programming used to sort and process the price data. Basker and Noel find the log of each price and then take the average while I find the average price and then take the log.

3.2 Long-Difference

The next model that I look at is a long-difference model. This model is better than the cross-section because it uses more of the data. More importantly, it removes all of the city fixed effects that are correlated with Wal-Mart's decision to locate to that particular city. The model follows the same logical structure as the cross section, but includes the change in price and the change in store count. The model takes the form:

$$\Delta p_{jk} = \alpha + \theta \Delta \text{WMSC}_j + \sum_k \phi_k \text{product}_k + \mathbf{x}_j \beta + \epsilon_{jk} \quad (2)$$

where I now have the change in price at non Wal-Mart stores and the change in the number of stores between the beginning and ending years of the sample (2001 and 2004 respectively). Once again, this model is estimated with and without the city variable vector and the results are shown in the second two columns of table 3. As with the cross section results, the WMSC coefficient are lower when the city variables are not included. Even when they are included their coefficients are statistically insignificant. The two estimates here are closer to each other than the two cross section estimates, and both show that each Wal-Mart store forces competitors' prices down by about 1.6 percent.

In contrast to the cross section, the long-difference results seem much more reasonable. This hints that most of the bias found in the cross sectional model is removed

in the long-difference model. This also makes sense intuitively. The bias in the cross sectional model comes from certain city characteristics that affect both the price and Wal-Mart’s decision to open a store. Because those characteristics are not measurable, the cross sectional model attributes their affect on prices to Wal-Mart. The long-difference model removes these stationary characteristics, leaving a better picture of Wal-Mart’s true power.

As with the cross section, my results follow Basker and Noel’s very closely. My estimation of Wal-Mart’s effect³ of 1.56 percent and 1.64 percent with and without income and population in the model is very close to Basker and Noel’s estimation of 1.23 percent and 1.27 percent. Once again, the differences are not statistically significant.

Basker and Noel make the point that the coefficient generated in the long-difference model only takes the change in stores into account. Timing of price and store changes are unknown to the model, so it must be assumed that only the change between 2001 and 2004 matters. This assumption may be fairly reliable, but the data set contains more data than is being used here. Next, I present the panel data model that Basker and Noel use to fix this problem.

3.3 Panel Data

The panel data model, like the long-difference model, eliminates bias from unobserved city effects, but it improves on the long-difference by using all four years worth of data. The model takes the form:

$$p_{jkt} = \alpha + \theta \text{WMSC}_{jt} + \sum_j \gamma_j \text{city}_j + \sum_{kt} \delta_{kt} \text{year}_t \text{product}_k + \epsilon_{jkt} \quad (3)$$

where p and WMSC are the same as above but are now specific to time t , city_j is a city indicator variable, and year_t is a year indicator variable. With the panel model, there is

³The estimations of Wal-Mart’s effect are negative in all models presented in this paper. To avoid confusion with terms such as “increases” or “larger,” I report the coefficients in absolute value. The tables at the end retain the negative sign on the coefficients.

no need to worry about bias because unobservable city characteristics are internalized. The WMSC coefficient should be a very close representation of Wal-Mart's true power over competitors' prices.

The WMSC estimate from the panel model is the lowest of the three different models. Whereas the cross sectional model said that Wal-Mart forces prices down by two percent, the panel model shows that Wal-Mart's actual price-lowering power is 1.2 percent. Basker and Noel's paper contains a detailed explanation of the interpretation of the panel estimate. For my purposes here, it will suffice to say that this is the best estimate to contrast with in section 4 where I present my extensions on their work.

My results, once again, closely resemble the results from Basker and Noel's paper. My WMSC panel estimation is a statistically significant 1.24 percent while Basker and Noel found a statistically significant 1.09 percent. These two estimates are very similar, and are in fact, not statistically different. All three different specifications of the model - the cross section, the long difference, and the panel - are successfully replicated in the results found by Basker and Noel. With the replication successful, I now have a firm ground to build on in the next section where I present two different extensions.

4 Extensions

Up to this point, the results presented have been a replication of Basker and Noel's results. In this section, I present two extensions of their work. First, I add geographic size to the model to see how that impacts Wal-Mart's effect. Second, I test to see if Supercenters that were originally Discount Stores carry the same effect as Supercenters that were built from scratch.

4.1 Wal-Mart's Effect and Area

One question has not been touched by previous literature: Does Wal-Mart carry the same effect in a large city as in a small city? The cross section and the long-difference

models treat two cities as identical in all aspects except for store count. Thus, they assume that Wal-Mart’s effect is identical in a small city as in a large city. This does not seem very reasonable. If a city is small, say 10 square miles, all competing stores are within a fairly short distance to Wal-Mart. This means that just about all consumers will have a choice to drive to Wal-Mart to do their shopping. As the area of the city grows, the expected distance between two stores also grows. The further the distance between Wal-Mart and a competitor, the less pressure the competing store is likely to feel to lower prices. If a consumer must drive 20 miles across town to get to the closest Wal-Mart, he or she may decide to stay closer to home and spend a few dollars more. In this section, I test to see how area impacts Wal-Mart’s estimated effect on prices. I do this, not because I believe area has a direct impact on prices, but because it is a good indicator of how far apart different stores are likely to be. Thus, Wal-Mart should have more influence over competitors’ prices in cities that have a smaller square mileage simply because stores are likely to be closer together in these cities. There could be other reasons that Wal-Mart’s influence erodes as square-mileage grows, but they are not detectible by this model.

To test this hypothesis, I add an interaction term involving the Wal-Mart store count and the square mileage to the model. My new panel model takes the form:

$$p_{jkt} = \alpha + \theta \text{WMSC}_{jt} + \sigma \text{WMSC}_{jt} \times \text{sqmi}_j + \sum_j \gamma_j \text{city}_j + \sum_{kt} \delta_{kt} \text{year}_t \text{product}_k + \epsilon_{jkt} \quad (4)$$

where all variables are the same as in equation (3) and sqmi_j is the log of city j ’s square mileage. I also estimate similar adaptations of equations (1) and (2). Both of these models also include sqmi_j in addition to the interaction term. This term is excluded from the panel estimate because it does not vary over time.

The coefficient for WMSC should still be negative in this model for the same reason it was negative in all three models presented in section 3. The hypothesis of this research is that Wal-Mart exercises downward pressure over competitors’ prices, and that was confirmed in the previous section. There is no reason to expect that, when area is

added to the mix, Wal-Mart would now force competitors to raise their prices. The other coefficient of interest here is the Wal-Mart and square mileage interaction term. I expect that this new interaction term will have a positive coefficient. If it is true that Wal-Mart's effect is mitigated by a larger metropolitan area, then the Wal-Mart effect coefficient ($WMSC_{jt}$) should have the opposite sign as the interaction term between WMSC and sqmi. If this is correct, the marginal effect of a new Wal-Mart Supercenter is:

$$\frac{\partial p_{jkt}}{\partial WMSC_{jt}} = \theta + \sigma \cdot (sqmi_j) \quad (5)$$

This is, in fact, what the estimates show. When area is added to the panel model, the WMSC coefficient increases dramatically to 0.0240 (from 0.0124) for the panel model. This indicates that prices fall by 2.4 percent if a city has an area of zero square miles. The results for all three model specifications are in table 4. For the panel model, the interaction term has a coefficient of 0.0022, which is the expected sign. The other two models also have positive interaction terms. The cross-sectional model is the only model that generated coefficients that are significant at the 1 percent level. The long difference specification produces a store count coefficient that is significant, but only at the 10 percent level. As has been discussed already, though, the cross sectional model produces biases which could cause the estimated coefficient to also have an inflated test statistic giving a false positive.

While the coefficients for WMSC are larger, they are now counteracted by city size. From table 2, a city with one Wal-Mart has an average size of 40.98 square miles. That means that the marginal effect of a single Wal-Mart store is just under 1.6 percent with the panel estimate using equation (5). For an average size city (72.31 square miles), the marginal effect an additional Wal-Mart from equation (5) has an F-statistic of 6.19, which is significant at the five percent level. This means that the marginal effect of Wal-Mart on competitors' prices is statistically different from zero even if the two individual components of that marginal effect are not significant. It is possible that area

and Wal-Mart’s store count are correlated enough that the individual effects cannot be independently determined, giving the insignificant coefficients while the combined effect is significant.

The largest city in the sample is 757 square miles, giving Wal-Mart a marginal effect of 0.94 percent. A single Wal-Mart store has some downward pressure on prices as long as the city is smaller than 55,000 square miles, well larger than any city in the sample. For cities larger than 195 square miles, Wal-Mart’s effect is smaller than was estimated in the model from equation (3). There are 10 cities in the sample larger than 195 square miles, meaning that the estimated impact of Wal-Mart is larger for the overwhelming majority of cities than is indicated in section 3.3.

Store-Level Effect

One restriction of the data is that the precise location of included stores is not known. This means that I can estimate an average effect for Wal-Mart across an entire city, but different stores may react differently based on their proximity to Wal-Mart’s store. As was mentioned earlier, the expected distance between two points within a circle grows as the radius of the circle grows. In fact, the expected distance between two points in a circle with radius R is $\frac{128R}{45\pi} \cong 0.9054R$. The proof of this can be found in Dennis Blumenfeld’s *Operations Research Calculations Handbook* on page 124. Using this fact, I can manipulate my results to generate a store-level effect. The store-level effect for a store x miles away from a Wal-Mart store will be $\theta + \tau x$ where

$$\tau = \frac{\hat{\sigma} \cdot \sqrt{\pi}}{0.9054} \tag{6}$$

and $\hat{\sigma}$ in equation (6) is estimated from equation (4) using the square root of square mileage instead of the log of square mileage. The square root of area is used here because it allows for a more straight-forward adaptation of the result found in Blumenfeld. The city-level estimates using this form of area can be found in the last column of table 4. The estimated results indicate that Wal-Mart has a statistically significant 1.71 percent

effect on competitors prices for cities with a square mileage of zero. For the average size city, Wal-Mart's effect is 1.45 percent. Applying equation (6), the marginal effect of an additional Wal-Mart on a store x miles away from Wal-Mart is $-1.71 + 0.001x$. Since I do not have the data to state the mean or median store-level effect, the equation form will have to suffice. Because different competitors are likely to respond to Wal-Mart differently based on their proximity to Wal-Mart, this estimation is an important extension as it pinpoints exactly how distance influences the impact of Wal-Mart.

4.2 Wal-Mart's Effect on the Big Three with Area

In their paper, Basker and Noel expand their analysis by distinguishing between prices that come from one of the three largest supermarket chains (Albertson's, Kroger, and Safeway), which they call the Big Three, and all other stores. They find that a Wal-Mart Supercenter forces prices at all other stores by 1.26 percent while the Big Three reduce their prices by a statistically insignificant 0.48 percent. They attribute this, in part, to large national, retailers having set pricing policies. This would reduce the impact Wal-Mart has on their pricing due to pricing policy being set from afar instead of at the store level. Their full explanation can be found in their paper.

I combine the model presented in equation (4) with Basker and Noel's Big Three model to assess how Wal-Mart's effect changes with area and competition type. It is possible that Big Three stores are more likely to locate in geographically larger cities. If this is the case, then maybe some of the counter-effect to Wal-Mart attributed to the Big Three is really attributable to area. Therefore, the key here will be to see what happens to the estimate for the interaction term between Wal-Mart store count and the Big Three dummy (as described in Basker and Noel's paper and below). If the interaction term's coefficient stays the same after area is added, it can be concluded that the original estimate of the Big Three's resilience is accurate.

When I separate Big Three prices from all other stores, my data set grows to 25,665 observations, of which 9,121 are from the Big Three. As was mentioned above, the

average size of the 175 cities in the sample is 72.31 square miles. There are 94 cities in the sample that have one of the Big Three stores, and those cities have an average size of 89.05 square miles. I expand the Big Three model that Basker and Noel use to include area. The expanded model takes the form:

$$p_{ijkt} = \alpha + \theta \text{WMSC}_{jt} + \rho \text{Big3}_{ijkt} + \sigma \text{WMSC}_{jt} \times \text{sqmi}_j + \psi \text{WMSC}_{jt} \times \text{Big3}_{ijkt} + \sum_j \gamma_j \text{city}_j + \sum_{kt} \delta_{kt} \text{year}_t \text{product}_k + \epsilon_{jkt} \quad (7)$$

where Big3_{ijkt} is an indicator variable that takes the value “1” if the price of product k comes from one of the Big Three stores (type i) in city j during time t and “0” otherwise.

This model shows Wal-Mart’s effect on different types of competitors, and it also shows how that impact is adjusted by the size of the city. It makes sense that larger cities are more apt to have other large chain stores in addition to Wal-Mart. As with equation (4), there is no reason to expect that Wal-Mart’s effect should become positive in this new model. Also, the interaction term with square mileage should be positive here as well because Wal-Mart’s influence decreases as city size grows. Finally, the interaction term with Big3 should be positive if it is true that Wal-Mart impacts large competitors less than small or local competitors.

The results are shown in the first column of table 5. The coefficients have the expected sign and magnitude, even if they are statistically insignificant. Each additional Supercenter has a 1.79 percent effect on competitors’ prices, which is higher than the 1.24 percent found from the model in equation (3). This effect is counteracted by the area, which shows that Wal-Mart’s effect falls (gets closer to zero) by .08 percent for every one percent increase in area. For the average city in the sample, Wal-Mart’s effect is 1.44 percent for non-Big Three stores and 0.76 percent for the Big Three. For a city with one Wal-Mart, the average size is 40.98 square miles, giving Wal-Mart an effect of 1.49 percent for non-Big Three stores and 0.81 percent for the Big Three.

4.3 Differing Effects of Converted and Newly Constructed Stores

Up until this point, I have been making the assumption that all Supercenters have the same effect whether they are a newly constructed store or a converted Discount Store. This, however, may not be the case. It is not hard to imagine that Supercenters of differing origins may have different effects on competitors' prices. Wal-Mart began opening Supercenters in 1988, converting many Discount Stores in the process. The model that I present in this section allows different effects from stores that were brand new locations and stores that were converted. The model does not, however, distinguish between stores converted in the 1980s and stores converted during the sample period.

To model different effects for different store origins, I replace the Supercenter store count variable with a "new store" count and a "converted store" count. A store is considered "converted" if, for a constant store number, the year the store opened is different from the year it opened as a Supercenter. For example, the store in Rogers, Arkansas was store numbered store 1 when it was opened as a Discount Store in 1962. It retained its store number when it became a Supercenter in 1997. Because 1997 and 1962 are not equal, this store would go into converted store count for Rogers. In general, Wal-Mart does not change store numbers during the conversion process. The altered panel model takes the form:

$$p_{jkt} = \alpha + \pi \text{NWMSC}_{jt} + \lambda \text{CW MSC}_{jt} + \sum_j \gamma_j \text{city}_j + \sum_{kt} \delta_{kt} \text{year}_t \text{product}_k + \epsilon_{jkt} \quad (8)$$

where all variables keep the same meaning as in equation (3) and NW MSC_{jt} is the number of Supercenters that began their life as new construction and CW MSC_{jt} is the number of Supercenters in city j during time t that began their life as Discount Stores and were later converted to a Supercenter. These two variables sum to the WMSC value used in previous models, so it is not included in this model to avoid perfect multicollinearity.

The results of the panel model suggest that stores with different beginnings do, indeed, have different effects on competitors' prices. The estimated coefficients are

presented in the second column of table 5. The results show that Supercenters that began as a Discount Store have a statistically significant downward effect on competitors' prices of 1.69 percent. Stores that were built as Supercenters have a statistically insignificant downward effect on prices of 0.74 percent. The two coefficients, however, are not statistically different from one another. The F-stat for the null of $\pi = \lambda$ is 0.93, which means that I cannot reject that new stores have the same power as converted stores. The coefficients are different enough, however, to still be interesting.

This indicates that new stores have less of an impact on competitors' prices than converted stores, a fact that seems somewhat counterintuitive on face value. But there is a very plausible explanation for this result. In their paper, Basker and Noel discuss an endogeneity problem in the cross section due to Wal-Mart picking locations using criteria that may also impact prices at other stores. In other words, Wal-Mart does not pick locations in a vacuum. Even if the endogeneity problem goes away with our panel specification, it may still show up when comparing types of stores.

As mentioned in section 1.1, Wal-Mart has had two very strategic expansions. Wal-Mart first expanded its Discount Store concept in the 1970s by never moving too far away from a previous store. It did not cherry pick markets around the country first, but the company instead slowly moved out from its base. In the 1990s, Wal-Mart used the same pattern to expand Supercenters around the country. This time, the company started from southern Missouri instead of northern Arkansas, but it covered roughly the same area in the same order as the first expansion. This means that a store going into the new store count in equation (8) is not a converted store for one of two reasons: that market was either passed up the first time around (i.e. it never got a Discount Store because it was deemed to be a bad market for one), or the market is getting more Supercenters than it had Discount Stores. If the former is the case, then it can be expected that cities being filled in with new Supercenters are most likely cities that are less vulnerable to Wal-Mart's power. If the latter is the case, then the new Supercenters are probably following converted Supercenters in most cases. If this is true, Wal-Mart

may have “used up” its effect with the converted stores and competing stores see little reason to lower their prices further when a new, non-converted, store opens.

So far, I have been using the same set of products in this section as in previous sections. This set of 23 products includes some goods found only at Supercenters and some goods found both at Supercenters and Discount Stores. By restricting by data set to include only products sold at Supercenters or only products sold at both types of stores, I am able to estimate each type of store’s effect on prices at competing stores for both categories of goods. Table 1 includes an asterisk next to the products carried at Discount Stores. The products without asterisks are only available at Supercenters. The hypothesis here is that converted stores should have no effect on competitors prices for products the store sold when it was a Discount Store. If a Wal-Mart carried soda before it became a Supercenter, competing grocery stores shouldn’t feel nearly as much pressure to lower soda prices as hamburger prices, a product that Discount Stores do not carry.

The results are shown in the last two columns of table 5. For products sold at both types of stores, neither NWMSC nor CWMSC have statistically significant coefficients. The effect for converted stores is roughly half the size of the estimation including all products and loses its significance. This is not unexpected; the effect of converted Wal-Mart’s on prices at competing stores for products sold prior to the conversion should be small. The coefficients for NWMSC and CWMSC are not statistically different from one another. When only products that are exclusively sold at Supercenters are included in the data set, CWMSC regains significance (this time at the 1 percent level) and both coefficients are larger than either of the other two estimations. Once again, the coefficients are not statistically different from one another, but the F-stat is the largest for the model only including Supercenter products. The explanation for the disparity between the estimates is the same as was discussed above for the estimates with all products. Additionally, when Wal-Mart converts a Discount Store to a Supercenter, competing stores may feel added pressure to lower prices on grocery

items due to negative effects they felt on non-grocery items when the Discount Store opened. This could explain why converted stores have the largest estimated effect in the Supercenter-only model.

5 Conclusion

In this paper, I replicate previous results and extend the analysis of Wal-Mart's power over competitors' prices. By estimating three different models, I show that each Wal-Mart location forces prices at competing stores down by about 1.2 percent. I also show that Wal-Mart's estimated power is reduced when the area of a city is included, but its power does not go away until a city get much larger than any city in the sample. Finally, I find that Supercenters converted from Discount Stores do not have a statistically different effect than newly-constructed stores, even though they have a statistically significant coefficient and new stores do not.

These results are significant because they lend credibility to the argument that Wal-Mart can be harmful to small retailers in communities. If a local grocer has a small profit margin to begin with, he may not be able to afford to drop prices, even by one percent. Even if the city is large, Wal-Mart's lower prices are powerful enough to force competitors to follow. But what is bad for local businesses may be good for local consumers. If competing stores even in large metro areas feel a pinch from Wal-Mart, it seems logical to conclude that shoppers are willing to drive a few extra miles to get their purchases for less. If they are able to buy things they previously could not afford, their standard of living increases. From the analysis here, it is impossible to tell which effect is larger. Either way, Wal-Mart is certainly a major force in the retailing world.

Bibliography

- [1] American Chamber of Commerce Research Association. “ACCRA Cost of Living Index: Comparative Data for 318 Urban Areas.” Vol. 41, No. 2, (2008).
- [2] Basker, Emek, and Michael Noel. “The Evolving Food Chain: Competitive Effects of Wal-Mart’s Entry into the Supermarket Industry.” *Journal of Economics and Management Strategy* (Forthcoming).
- [3] Basker, Emek. “Selling a Cheaper Mousetrap: Wal-Mart’s Effect on Retail Prices.” *Journal of Urban Economics* 58 (2005): 203-29.
- [4] Blumenfeld, Dennis. *Operations Research Calculations Handbook*. CRC Press, 2001.
- [5] “Corporate Fact Sheet.” 17 July 2008. Wal-Mart. <http://www.walmartstores.com>.
- [6] Hausman, Jerry, and Ephraim Leibtag. “Consumer Benefits from Increased Competition in Shopping Outlets: Measuring the Effect of Wal-Mart.” MIT and Economic Research Service (2005).
- [7] Holmes, Thomas. “The Diffusion of Wal-Mart and Economies of Density,” unpublished paper, University of Minnesota, (2008)
- [8] Sveum, Matthew. “Rolling Back Prices: Wal-Mart’s Ability to Change the Amount Consumers Pay.” Honors Thesis. Hamline University, 2007.
- [9] “Wal-Mart Store, Inc - History Timeline.” Wal-Mart. <http://walmartstores.com/AboutUs/7603.aspx>.

6 Tables

Product	Description	Average Price	Minimum Price	Maximum Price
Bananas	Price per lb	0.51	0.25	0.82
Bread*	24 oz loaf, white lowest price	0.04	0.01	0.12
Canned Tomatoes	14.5 oz Hunts or Del Monte	1.00	0.66	1.67
Coffee*	11.5 oz can Mazwell House Hills Bros. or Foldgers	2.56	1.48	4.16
Corn Flakes*	18 oz Kellogg's or Post	2.67	1.25	4.31
Dish Powder*	75 oz Cascade	0.05	0.03	0.09
Eggs	One dozen, grade A large	1.03	0.39	2.64
Frozen Corn	16 oz whole kernal, lowest price	1.20	0.73	2.49
Frying Chicken	Whole, uncut price per lb	1.01	0.56	1.72
Ground Beef	Price per lb, min 80% lean lowest price	1.86	0.94	3.83
Lettuce	1.25 lb head	1.09	0.52	1.95
Margarine	1 lb cubes, Blue Bonnet or Parkay	0.83	0.44	1.95
Milk*	Half-gallon whole milk	1.81	0.94	2.86
Parmesan Cheese	8 oz Kraft canister	3.61	2.49	5.28
Peaches	29 oz Hunts, Del Monte, Libby's or Lady Alberta can	1.69	1.12	3.09
Potatoes	10 lb., white or red	3.48	1.50	6.78
Sausage	Jimmy Dean or Owens, 100% pork	0.21	0.11	0.32
Shortening	3 lb can, all-vegetable, Crisco	3.21	1.87	4.27
Soda*	2 liter Coke	1.18	0.78	1.79
Steak	T-bone, price per lb	7.62	3.99	12.89
Sugar	4 lb sack, lowest price	0.39	0.24	0.78
Sweet Peas	15-15.25 oz can, Del Monte	0.81	0.47	1.50
Tissue*	200 count box, Kleenex brand	1.36	0.93	2.28
Tuna	6 oz can, Starkist or Chicken of the Sea	0.68	0.38	1.85

Table 1: Product descriptions and Price statistics. *Denotes products carried at both Supercenters and Discount Stores.

Number of Stores	Number of Cities	Average Population	Average Income	Average Square Milage
0	63	185,993	36,577	54.15
1	65	82,303	33,176	40.98
2	35	238,904	34,776	119.31
3	8	277,712	40,413	135.11
4	2	267,578	39,424	68
5	1	505,963	34,947	607
8	1	1,954,848	36,616	579.42
Total	175	175,123	35,152	72.31

Table 2: City statistics based on the number of Wal-Mart stores in 2002.

LHS Variable	Cross-Section		Long Difference		Panel
	$\log(p_{jk,2002})$		$\Delta \log(p_{jk})$		$\log(p_{jkt})$
WMSC	-0.0236*** (0.0063)	-0.0307*** (0.0059)	-0.0156*** (0.0043)	-0.0163*** (0.0051)	-0.0124*** (0.0034)
Log		0.0231*** (0.0062)		0.0016 (0.0050)	
Population					
Log Median		0.0656* (0.0366)		0.0005 (0.0323)	
Income					
Product FE	Y	Y	Y	Y	Y
Product \times Year FE	N	N	N	N	Y
City FE	N	N	N	N	Y
Observations	4,191		4,176		16,774

Table 3: Regression results. * significant at the 10% level, ** significant at the 5% level, *** significant at the 1% level

LHS Variable	Cross-Section		Long Difference		Panel	Panel
	$\log(p_{jk,2002})$		$\Delta \log(p_{jk})$		$\log(p_{jkt})$	$\log(p_{jkt})$
WMSC	-0.0845*** (0.0249)	-0.0626*** (0.0224)	-0.0556* (0.0337)	-0.0553* (0.0336)	-0.0240 (0.0200)	
WMSC (Using Square Root Area)						-0.0171* (0.0104)
Log Area	0.0041 (0.0078)	-0.0513*** (0.0160)	-0.0042 (0.0045)	-0.0177* (0.0177)		
WMSC×Log Area	0.0116*** (0.0045)	0.0084*** (0.0042)	0.0077 (0.0057)	0.0075 (0.0056)	0.0022 (0.0033)	
WMSC×Square Root Area						0.0003 (0.0005)
Log		0.0528*** (0.0139)		0.0136 (0.0104)		
Population						
Log Median		0.0707** (0.0350)		0.0071 (0.0317)		
Income						
Product FE	Y	Y	Y	Y	Y	Y
Product × Year FE	N	N	N	N	Y	Y
City FE	N	N	N	N	Y	Y
F-stat for joint effect	22.74	14.76	6.52	6.91	6.19	6.19
p-value	<0.0001	0.0006	0.0385	0.0315	0.0454	0.0454
Observations	4,191		4,176		16,774	

Table 4: Regression results with Area added. The F-stat for the joint effect measures whether $\theta + \ln(72.31)\sigma = 0$, where 72.31 square miles is the mean city size. * significant at the 10% level, ** significant at the 5% level, *** significant at the 1% level.

	Big Three with Area	Wal-Mart Store Type	Discount Store & Supercenter	Only Supercenter
LHS Variable	$\log(p_{ijkt})$	$\log(p_{jkt})$	$\log(p_{jkt})$	$\log(p_{jkt})$
WMSC	-0.0179 (0.0249)			
NWMS		-0.0074 (0.0054)	-0.0071 (0.0078)	-0.0084 (0.0058)
CWMS		-0.0169** (0.0067)	-0.0084 (0.0080)	-0.0197*** (0.0075)
Big3	0.0199* (0.0087)			
WMS \times Log Area	0.0008 (0.0041)			
WMS \times Big3	0.0068 (0.0044)			
Product FE	Y	Y	Y	Y
Product \times Year FE	Y	Y	Y	Y
City FE	Y	Y	Y	Y
F-stat	5.13	0.93	0.01	1.11
p-value	0.1625	0.3341	0.9123	0.2924
Observations:	25,665	16,774		

Table 5: Regression results for Area with Big Three and for Store Differentiation. The F-stat for Big Three with Area tests whether $\theta + \ln(72.31)\sigma + \psi = 0$, where 72.31 square miles is the mean city size. The F-stat for Wal-Mart Store Type tests whether $\pi = \lambda$.