

The Evolving Food Chain: Competitive Effects of Wal-Mart's Entry into the Supermarket Industry*

Emek Basker
University of Missouri

Michael Noel
University of California–San Diego

July 2007

Abstract

We analyze the effect of Wal-Mart's entry into the grocery market using a unique store-level price panel data set. We use OLS and two IV specifications to estimate the effect of Wal-Mart's entry on competitors' prices of 24 grocery items across several categories. Wal-Mart's price advantage over competitors for these products averages approximately 10%. On average, competitors' response to entry by a Wal-Mart Supercenter is a price reduction of 1–1.2%, mostly due to smaller-scale competitors; the response of the “Big Three” supermarket chains (Albertson's, Safeway, and Kroger) is less than half that size. We confirm our results using a falsification exercise, in which we test for Wal-Mart's effect on prices of services that it does not provide, such as movie tickets and dry cleaning services.

JEL Codes: L11, L13, L81

Keywords: Wal-Mart, Retail Prices, Supermarkets, Price Competition

*Contact: emek@missouri.edu or mdnoel@ucsd.edu. We thank Saku Aura, Roger Betancourt, Paul Dobson, Luke Froeb, Jerry Hausman, Günter Hitsch, Ephraim Leibtag, Saul Lach, Daniel Levy, and David Parsley for helpful comments and conversations, seminar participants at the Federal Reserve Bank of San Francisco, UC Santa Cruz, Bar-Ilan University, the 2006 Comparative Analysis of Enterprise (Micro) Data Conference (Chicago) and the 2007 IIOC (Savannah) for comments, Wal-Mart for providing the administrative data used in this paper with no strings attached, and Tom Holmes for the Distribution Center data. This research was supported by University of Missouri Research Board Grant #1344, and was started while Basker was visiting the Haas School of Business at the University of California–Berkeley, which she thanks for its hospitality.

1 Introduction

Productivity growth in the retail sector is driven almost entirely by entry (and expansion) and exit (and contraction): more productive firms open additional establishments, and less productive ones contract and exit. Yet little is known about the consequences of this productivity-enhancing churn on competition and prices. In this paper, we aim to fill that gap by analyzing the competitive consequences of Wal-Mart's expansion into the supermarket industry. Wal-Mart's entry into the supermarket industry has shaken up the previously-stagnant sector and has had a profound effect on its organization.

Since opening the first Wal-Mart Supercenter — which carries a full line of grocery items — in Washington, Missouri, in 1988, Wal-Mart has averaged more than 100 new Supercenter openings per year. More than half of all U.S. Wal-Mart stores now sell a full line of grocery items; in dollar terms, it is the leading supermarket chain in the U.S.¹ Wal-Mart's aggressive pricing has been credited with reducing operating margins of competing supermarkets and lowering consumer prices for many food items (Hausman and Leibtag, forthcoming). There has also been much speculation about the competitive effects of Wal-Mart's expansion; for example, the February 2005 bankruptcy filing of Winn-Dixie, a large supermarket chain based in Florida, was widely blamed on Wal-Mart's rise (see, e.g., Mccarthy, 2005; Mnyandu, 2005). In fact, the vast majority of supermarket bankruptcy cases in the last decade have cited Wal-Mart as a catalyst (Callahan and Zimmerman, 2003).

This paper analyzes the short- and medium-run price effects of Wal-Mart's fast rise to the top of the retail food chain. To do this, we combine a unique store-level panel data set consisting of both Wal-Mart's and competitors' prices with a separate data set containing the opening dates of all U.S. Wal-Mart Supercenters. The price data come from the American Chamber of Commerce Research Association (ACCRA) and cover 24 specific grocery items

¹Wal-Mart's other stores are known as "discount stores" and specialize in general merchandise including clothing, housewares, toys, and drugstore items.

from several categories including dairy products, meats, produce, canned and frozen goods, and miscellaneous items; and six services, including dry cleaning and movie tickets, which Wal-Mart does not provide. We use micro data from the July 2001, 2002, 2003 and 2004 surveys in 175 local markets; in all, the data contain nearly 100,000 individual prices. Data on the exact timing of Wal-Mart Supercenter entry into these markets were obtained from Wal-Mart Stores, Inc. and supplemented with public data. Over this time period, the number of Wal-Mart Supercenters grew by 60%, from 972 in July of 2001 to 1562 by July of 2004.

We find that both the direct and indirect price effects of Wal-Mart's expansion are significant. The direct effect is the raw difference between the price Wal-Mart charges and the price competitors charge. For most items in our sample, this effect is substantial, averaging 10%. The indirect price effect of Wal-Mart's expansion is due to price reductions at competing supermarkets; we find this effect to be approximately 1–1.2%.

We use several specifications to analyze Supercenters' effect on competitors prices. For comparison with many existing studies, we start with a cross-sectional analysis which is subject to omitted variable bias due to the endogeneity of Wal-Mart's location decisions. We also estimate long-difference and panel OLS specifications, which are substantially more reliable. Finally, we examine two instrumental-variables solutions to the endogeneity problem in the cross-sectional analysis. The number of Wal-Mart discount stores (which do not sell most groceries) in the late 1990s provides the best correction for endogeneity in the cross-section. To check our specifications we use a falsification exercise and test for Wal-Mart's effect on prices of six services that it does not provide — appliance repair, movie ticket, bowling, man's haircut, woman's beauty salon appointment, and dry cleaning. In our preferred specification, which best controls for unobserved factors influences Wal-Mart's entry decision, we find no effect of Wal-Mart Supercenter entry on the prices of these unrelated services.

The remainder of the paper is organized as follows. Section 2 provides a brief background on Wal-Mart's expanding presence in the supermarket industry. Section 3 describes the data.

Section 4 reports our estimation results. Section 5 concludes.

2 Wal-Mart and the Supermarket Industry

The supermarket industry has undergone several transformations in the past century. Tedlow (1990) reports that the first upheaval occurred between 1900 and 1930 when the rise of A&P (the Great Atlantic and Pacific Tea Company) and several other, smaller chains (including Kroger and Safeway) displaced many small, expensive and inefficient local grocers. Tedlow (1990) speculates that the reasons for A&P's spectacular early success were vertical integration and efficiency in production. By 1930, A&P operated more than 15,000 stores in much of the U.S. When A&P declined — its number of stores had dropped by 50% by 1940 — other large chains took over. With cars becoming commonplace in the 1950s, A&P was replaced by suburban supermarket chains operating large self-service stores stocking ever-more varieties.

The most recent shakeup involves the growth of “superstores”: general merchandise stores which have added a full line of groceries. Between 1997 and 2002, sales of grocery products in traditional grocery stores fell by approximately 2% in real terms, while sales of grocery products in “general merchandise” stores, which include Wal-Mart, grew by 48% in real terms (and by 57% in the narrower category of warehouse clubs and superstores). The share of grocery sales accounted for by general merchandise firms rose from 12.1% to 18.2%.² Pressures on traditional grocery stores are exacerbated by upscale chains (such as Whole Foods and Trader Joe's) specializing in fresh, organic, and ready-made foods.

Wal-Mart, currently the top grocer in the U.S. by dollar sales, is the main new player in this market. Kmart and Target, which also operate superstores, have not grown nearly as fast as Wal-Mart (Graff, 2006). Wal-Mart opened its first Supercenter as an experimental

²In the 1992 Census of Retail Trade, only 6.9% of food sales are reported by the general merchandise sector. But due to the change from SIC to NAICS, this figure is not directly comparable to the later figures.

format in 1988; today, more than 2,200 of Wal-Mart's stores are Supercenters. Figure 1 shows the number of Wal-Mart stores and Supercenters up to 2004. Wal-Mart currently accounts for 14% of Kraft and Kellogg's sales, 16% of General Mills' sales, and 11% of Pepsi's sales (Warner, 2006). Wal-Mart's success is often attributed to its expert logistics systems (Ellickson, 2006; Westerman, 2001) and its cost-conscious "corporate culture." In the retail sector as a whole, the productive gap between large national chains and single-unit retailers is quite large (Foster, Haltiwanger, and Krizan, 2006) and it is conceivable that food retailing exhibits even more asymmetries than the retail sector average.

Several recent studies have found that Wal-Mart charges lower prices than competitors. For example, an April 2002 UBS Warburg survey of 100 grocery and non-grocery items in 4-5 grocery stores in three markets with both Wal-Mart and non-Wal-Mart grocery stores found that Wal-Mart's prices were 17-39% lower than competitors' prices (Currie and Jain, 2002). Hausman and Leibtag (forthcoming) find a 30% premium at traditional supermarkets over superstores, mass merchandisers and club stores (SMCs) for a similar array of products.

The causal impact Wal-Mart has on competitors' prices has been less firmly established. The UBS study, for example, finds that prices in Las Vegas, Houston and Tampa, which had Wal-Mart Supercenters, were on average 13% lower than prices in Sacramento, which did not have any Supercenters (Currie and Jain, 2002), but does not establish Wal-Mart's causal role in this difference. Basker (2005) uses a data set of all Wal-Mart's store locations and opening dates to estimate its effect on average prices of drugstore items such as shampoo and toothpaste, but the prices she uses are average prices across a large number of surveyed establishments, and may include Wal-Mart's own prices after entry. As a result, Basker's (2005) estimates may confound the direct and indirect effects of Wal-Mart's presence. Zhu, Singh, and Dukes (2005) analyze the effect of entry of a Wal-Mart discount store, which does not sell groceries, on nearby Dominick's Finer Foods supermarkets using a case-study approach. But they focus on store traffic and revenue, and do not study the effect on prices charged by the incumbents. In this paper, we estimate Wal-Mart's *causal effect* on

competitors' grocery prices.

In related work, Hausman and Leibtag (forthcoming) use data from AC Nielsen Home-scan data which contain both prices paid and quantities purchased for a large panel of consumers to estimate the competitive effects of increased spending at SMCs on competitors prices.³ They find that from 1998 to 2001, a period during which the market share of SMCs in their sample increased from 10.9% to 16.9% of grocery sales, average prices paid by consumers declined, on average, by 3%. The effect varies from product to product. For a subset of items that overlap our sample — bananas, bread, butter and margarine, cereal, chicken breast, coffee, eggs, ground beef, lettuce, milk, potatoes and soda — their estimates imply that consumer prices fell by 2.6%. While Hausman and Leibtag's (forthcoming) analysis focuses on the causal effects of shifting *expenditure shares*, due to a combination of factors including entry, exit of competitors, and changes in consumers' shopping venues or movements of relative prices, the focus of our paper is on the well-defined causal impact of entry alone.⁴

Not all studies show such a large effect, however. Singh, Hansen, and Blattberg (2006) use a case study to analyze the impact of a new Wal-Mart Supercenter on an incumbent chain grocery store 2 miles away. Using high-frequency data from the incumbent, covering a period from one year before until one year after Wal-Mart's entry, they find an increase in promotions ("sales") and a decline in quantities purchased, but no consistent pattern for prices at the competing supermarket.

³Hausman and Leibtag do not distinguish between Wal-Mart stores and other SMCs in their data, although Wal-Mart is likely responsible for the bulk of SMC expenditure.

⁴Hausman and Leibtag regress the average quantity-weighted price charged by traditional supermarkets for a given grocery category on the share of consumer expenditures at SMCs for that item. Because expenditure share at SMCs is a function of the prices that traditional grocery stores charge, they instrument for product-specific expenditure share using the expenditure share at SMCs aggregated across all products, arguing that each good contributes a negligible share to overall expenditure. The instrument is valid as long as the price of each good in one venue has no direct effect on the price of any other good in another venue, and no omitted variable affects all prices.

3 Data

Price data for 24 grocery items come from the American Chamber of Commerce Research Association (ACCRA). ACCRA, through local Chambers of Commerce, surveys up to 10 grocery stores in the first week of each quarter in participating cities. Participating cities vary from quarter to quarter, with some cities moving in and out of the sample frequently, while others are included more regularly; 250-300 cities are surveyed each quarter during the sample period. We obtained store-level prices from ACCRA for four points in time: July 2001, July 2002, July 2003 and July 2004.⁵ We use a sample of 175 cities in the continental United States that appear in the data in all four quarters; Figure 2 shows a map of their locations.

The stores surveyed include supermarket chains, “superstores” such as Wal-Mart, and smaller grocers, but exclude membership clubs such as Sam’s Club or Costco. The number of stores surveyed in any single market in a given survey period ranges from one (in 31 instances) to ten (in 905 instances); the mean, median and modal number of stores surveyed in each market is five.⁶ Approximately 14% of prices were collected at Wal-Mart stores, and Wal-Mart is included in 82% of pricing surveys in cities with at least one Supercenter.⁷

The products are listed in Table 1. They include dairy products (milk, eggs), meats (chicken, sausage), produce (bananas, potatoes), canned goods (peaches, tuna), and miscellaneous items (sugar, dishwashing powder).

In each store, the price recorded is the lowest prices available to all shoppers arriving at the store. So, for example, discounts that require obtaining coupons in advance are never

⁵ACCRA did not keep records of individual prices collected until mid-2001, when the company switched from paper to electronic data collection. Aggregate data from ACCRA have been used in numerous academic studies, including Parsley and Wei (1996), Aaronson (2001), and Frankel and Gould (2001).

⁶This seems appropriate given Ellickson’s (2004) finding that 4-6 stores per market capture 60-70% of food sales.

⁷ACCRA’s instruction to price collectors is to “[s]elect only grocery stores [...] where professional and managerial households normally shop. Even if discount stores are a majority of your overall market, they shouldn’t be in your sample at all unless upper-income professionals and executives really shop there” (American Chamber of Commerce Research Association, 2000, p. 1.3).

factored in to price. Promotions that do not require coupons or for which coupons are available in the store, and are redeemable on the spot, are factored in. In addition, when supermarkets offer lower prices to customers with “frequent shopper” cards, ACCRA’s policy is to report these lower prices as long as the cards are free and there is no waiting period for their use, conditions that are met by all supermarket frequent-shopper programs we are aware of.

Wal-Mart provided the opening dates of all its U.S. stores, as well as dates of major renovations, including conversions of “discount stores” to “Supercenters.” The initial opening dates in the Wal-Mart administrative data are extremely accurate, but approximately a quarter of the Supercenters that were converted from discount stores are missing conversion dates in the file. We used press releases from Wal-Mart’s web site to obtain the conversion dates for stores that were converted after March 2001.⁸ Maps showing the locations of Wal-Mart Supercenters as of July 1 of each year are shown as Figure 3; new Supercenters since the previous July (for 2002-2004) are shown in red.

Of the 175 sample cities, 23 had their first Supercenter open between July 2001 and July 2004. The number of Wal-Mart Supercenters increased in a total of 59 cities in the sample during the study period; of these, 19 had more than one Supercenter added.⁹

The locations and opening dates of Wal-Mart’s food Distribution Centers were obtained from Tom Holmes, who compiled the data from various sources including ammonia permit requests and subsidy information.

Finally, we use the 2000 Census of Population to obtain population size and 1999 median household income by city.

⁸The press releases from 2001-2003 have since been removed from the site.

⁹In contrast, only four Costco stores opened in the sample cities during the sample period.

4 OLS and IV Estimates

4.1 OLS

4.1.1 Cross-Sectional Estimates

We start by estimating the difference between Wal-Mart's price and competitors prices in markets that already include a Wal-Mart Supercenter. To do this, we calculate the average non-Wal-Mart price as the unweighted average of log prices across all other surveyed retail establishments, excluding Wal-Mart, for each product and city at each point in time, and the average Wal-Mart price as the unweighted average of log prices across all Wal-Mart Supercenters in the market, if there is more than one. Consistent with both popular perception and previous studies, on average, Wal-Mart's prices are 10.5% lower than competitors' prices for these items. Figure 4 presents a graphical depiction of the average log difference between Wal-Mart's price and the non-Wal-Mart price in cities with at least one Wal-Mart Supercenter by product. The differences range from 2.4% (for milk) to 19.8% (for frying chicken).

While these price differences are substantial and statistically different from zero for all products, they are smaller than ones described by Hausman and Leibtag (forthcoming), who found an average price difference of 27% between supercenters, mass merchandisers, and clubs (SMCs) on the one hand and traditional supermarkets on the other. The most likely reason for the gap between our estimates is data differences. The ACCRA data we use provide comparable prices for specific items across stores, both within and across markets. The AC Nielsen Homescan data used by Hausman and Leibtag attempt to capture the prices consumers pay for broader product categories, allowing for substitution across products and outlet types. For example, ACCRA data specifically exclude membership clubs, which use two-part tariff pricing (an annual membership fee, and lower unit prices) and tend to sell larger packages which are not directly comparable with other stores. Since clubs charge lower prices than other retailers, the Homescan SMC prices are typically lower than Wal-

Mart’s prices alone. Another difference is that ACCRA products are very narrowly defined, specifying the package size and often the brand of the item priced, whereas Homescan product categories are fairly wide and include products that vary in brand, quality, and package sizes. Finally, ACCRA prices are actual prices collected in a single week whereas the Homescan data used by Hausman and Leibtag are aggregated to the monthly level, using quantity weights that reflect consumers’ product and intertemporal substitution patterns.

For the remainder of the paper, we use \mathbf{p}_{jkt} to denote the average log price of product k across non-Wal-Mart stores in market j at time t .

Because many studies have employed simple comparisons of prices in “Wal-Mart cities” and “non-Wal-Mart cities” (e.g., Currie and Jain, 2002), our first regression is a simple cross-section using price data for each year separately and treating the number of Wal-Mart stores in the market as exogenous. For 2002, we estimate:

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

where $\mathbf{p}_{jk,2002}$ is the average non-Wal-Mart log price of product k in city j in July 2002, $\text{WMSC}_{j,2002}$ is the number of Wal-Mart Supercenters in city j at the beginning of July 2002, and product_k is a product indicator. The control variables in \mathbf{x}_j are log population (from the 2000 Census) and log 1999 median household income (we also report estimates from regressions omitting these controls). To account for correlation in the error term across products within a city, standard errors are clustered at the city level. There are no time effects in the model because it contains a single cross-section, and no city effects because they are perfectly correlated with the Wal-Mart variable.

The results for the 2002 cross-section are shown in the first two columns of Table 2. (We chose 2002 as an arbitrary mid-sample year. Results for other years, not shown, are very similar.) In the first column we omit the control variables, which we add in the second column. The estimates show that each Supercenter is correlated with competitors’ average

prices that are 2.5–3% lower, significantly different from zero at the 1% level.

Using the number of Wal-Mart Supercenters as an explanatory variable allows us to efficiently exploit the maximum amount of variation in the cross-sectional data. We have experimented with two alternative specifications (not shown). First, we use an indicator function, which equals 1 if there is at least one Supercenter in the city, to replace the number of Wal-Mart Supercenters. In July 2002, 108 of our 175 cities had at least one Supercenter. (That number increased from 98 to 121 between July 2001 and July 2004.) The indicator function misses, however, the variation in the number of Supercenters among these cities; on average, a city with one or more Supercenters in 2002 had 1.6 such stores, and the standard deviation was 0.9. Estimates using the indicator function imply that price reductions by competing grocery stores are larger than those that use the number of stores, averaging 5.5% per Supercenter (standard errors are also larger).¹⁰ When we allow the effect of the first Supercenter to differ from the effect of additional Supercenters, we find that the first Supercenter reduces prices by 8–9% and additional Supercenters have marginal effects of 2–3%, all statistically significant at the 1% level.

Unfortunately, the cross-sectional estimates cannot be interpreted causally due to omitted variable bias. Wal-Mart stores are not assigned randomly to sample cities, and factors that affect the number of Wal-Mart stores in a market are likely to be correlated with the error term in these price regressions. Because Wal-Mart caters to low-income consumers, and because it has traditionally operated in low-cost areas in the South and Midwest, we expect this bias to drive the estimated coefficients downward (away from zero). Including the control variables may mitigate this problem, but is unlikely to solve it. We next turn to some OLS specifications that are better designed to address this issue. We return to the cross-sectional specification in Section 4.2 where we consider two instrumental variables methods that attempt to address the omitted-variable bias.

¹⁰We compute the effect per Supercenter by dividing the coefficient of -8.8% by the average number of Supercenters in a market that had at least one Supercenter (1.6).

4.1.2 Long Difference Estimates

As a better alternative to the cross sectional model, we also estimate a long-difference regression,

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

which has the change in log average non-Wal-Mart prices of product k in city j between July 2001 and July 2004 on the LHS, and the change in the number of Wal-Mart Supercenters in city j over this time period, along with product fixed effects on the RHS. This specification has the advantage that time-invariant characteristics of the market are removed by the differencing, removing many possible omitted variables. Estimates obtained from this specification are valid if the number of *new* Wal-Mart Supercenters over this period is uncorrelated with the error term. An implicit assumption in these regressions is that only the entry of Wal-Mart, and not the exact timing of this entry — as long as it occurs within the 3-year window from July 2001 to July 2004 — matters.

Results are shown in the third and fourth columns of Table 2. Estimates of price reductions are around 1.2% and are significantly different from zero at the 1% level. Combining this indirect effect of Wal-Mart Supercenter entry on the prices of competing supermarkets with the post-entry price difference of 10.5% calculated above, we estimate that competing grocery stores do cut prices when Wal-Mart enters their market, but only by 1/10th of their initial (pre-entry) price premium over Wal-Mart.

This estimate of the indirect effect is smaller by 50% or more than the cross-sectional estimates, consistent with our concern that the cross-sectional estimates attribute too much of the differences across cities to Wal-Mart. As a specification check, we also estimated this model with the initial (2001) log price on the right-hand side, along with the other control variables; coefficient estimates are very similar to those shown.

4.1.3 Panel Estimates

Our third and preferred OLS specification is a panel with market fixed effects. As with the long-difference regression, the panel regression controls for all unobserved time-invariant factors that influence market demand. In addition, while the long-difference specification does not use information on the exact timing of Wal-Mart’s entry or the timing of price responses, the panel specification is more efficient because it uses all available information.

We estimate

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

where \mathbf{p}_{jkt} is the average non-Wal-Mart log price of product k in city j at time t , \mathbf{WMSC}_{jt} is the number of Wal-Mart Supercenters in city j at time t , \mathbf{city}_j is a city indicator to capture cost and price differences across cities for any reason other than Wal-Mart’s entry, \mathbf{year}_t is a year indicator and $\mathbf{product}_k$ is a product indicator; their interaction is intended to capture overall cost differences and changes at the product level.¹¹ To account for correlation in the error term across products and over time within a city, standard errors are clustered at the city level.

OLS estimates of Equation (3) are interpretable as causal effects of Wal-Mart’s entry and expansion if and only if, conditional on Wal-Mart’s entry or expansion into market j over the period July 2001–July 2004, the exact timing (year) of entry is uncorrelated with the ε_{jkt} . This is a relatively weak condition: time-invariant city characteristics are captured with the city fixed effects and the panel is short enough that there are unlikely to be many large changes in city characteristics. This is effectively a difference-in-difference estimator.

¹¹We do not include covariates in the regression because there are very few market-level time-varying factors for which data are available at annual frequency for this sample period. The exception is population — July population estimates are available from the Census for nearly all cities in the data set. We included log population in all regressions as a robustness check, but it was never significant and did not change the coefficients of interest in any meaningful way.

Results are shown in the last column of Table 2. The effect of a Wal-Mart Supercenter is now estimated to be a price reduction of approximately 1.1% among other grocery stores. This effect is identified using 77 Supercenter openings in the sample cities over the period studied, distributed roughly equally across time (23 between July 2001 and July 2002, 28 the following year, and 26 in the final year). This estimate is very slightly smaller — and slightly more precise — than the long-difference estimates reported earlier. The 1.1% price reduction by competing grocery stores is modest compared to Wal-Mart’s direct effect — prices that are 10% lower — suggesting the benefits from Wal-Mart Supercenters accrue mostly to consumers who shop there, while consumers who do not modify their shopping habits benefit little.

These results are best interpreted within the context of a monopolistically-competitive industry. Let the absolute value of the elasticity of residual demand for a given store be ε_0 before Wal-Mart’s entry, and ε_1 after entry. Conceivably, ε_1 could be higher *or lower* than ε_0 , depending on the composition of the residual consumers, but our estimates of price reductions indicate that the elasticity of residual demand increases when Wal-Mart enters the market.¹² Assuming constant marginal cost, competitors’ price reductions can be used to infer the relationship between the post-entry and pre-entry elasticities. A 1% price reduction is consistent with post-entry elasticity of residual demand of

$$\varepsilon_1 = \frac{\varepsilon_0}{1.01 - 0.01 \cdot \varepsilon_0}.$$

This could represent a very small increase in the elasticity of demand — if ε_0 is around 3, a 1% price reduction implies an increase of only 2% in the elasticity of residual demand, to 3.06 — or a larger increase if ε_0 is large. For ε_0 of 10, for example, a 1% reduction in price indicates the elasticity of residual demand increased by 10%, to 11. Importantly, we cannot

¹²For a given marginal cost c and elasticity of residual demand ε , the static profit-maximizing price is $p = \frac{\varepsilon}{\varepsilon-1} \cdot c$.

infer from these modest price changes the degree to which quantities purchased changed.¹³

We prefer using the number of Supercenters on the LHS to other specifications because it uses all available information, and efficiently exploits the most variation in the data. If we replace this variable with an indicator variable which equals 1 if there is a Supercenter in the market, we identify the Supercenter effect using only the 23 *initial* entries into a market (19 of them in the first two years), compared with the 77 openings used above. In this specification, the estimated effect of each Supercenter entry (not shown) increases to 1.65%, but standard errors are 2-3 times larger than in our main specification.¹⁴ We also experimented with a specification in which we allow for the first Supercenter to have a different effect from subsequent ones. The results (not shown) suggest the effects of the first and later Supercenters are almost identical, and statistically indistinguishable, at -1.2% and -1.1%, respectively. The first of these is not statistically different from zero but the second is significant at the 1% level.

These estimates of 1–1.2% price reductions per Supercenter mask larger average price reductions over our sample period. The 59 cities in our sample in which the number of Wal-Mart Supercenters increased over the three-year period July 2001–July 2004 averaged an increase of 1.5 Supercenters over this time. Our estimates therefore imply that, in the average city with Supercenter entry or expansion grocery prices at competing supermarkets fell by 1.5–1.8% over this three-year period.

We also estimate Equation (3) separately for each product. Rather than presenting individual estimates for these product-by-product regressions, Figure 5 depicts the point

¹³These price reductions are in line with other estimates of the effect of increased elasticity of demand on retail prices. Lach (forthcoming) finds that the immigration wave from the former Soviet Union to Israel in 1990 reduced prices of groceries and household goods in Israeli cities that received a disproportionate number of immigrants, due to the immigrants' higher elasticity of demand. He estimates that a one percentage point increase in the ratio of immigrants to natives in a city decreased prices, on average, by half a percentage point.

¹⁴The 23 cities that got their first Supercenters over this period averaged 1.3 new Supercenters between July 2001 and July 2004. We compute the effect per Supercenter by dividing the coefficient of -2.1% by this number.

estimates graphically. Effects range from price reductions of 3.65% for margarine to price increases of 0.65% for soda, with twenty two of the 24 coefficients negative. Point estimates are statistically different from zero at the 5% level (and negative) for seven products: bananas, lettuce, canned tomatoes, margarine, frying chicken, dish washing powder, and tissues (estimates for sausage are significant at the 10% level). Neither of the positive coefficients is significantly different from zero.

The panel estimates provide the cleanest identification and the smallest coefficient estimates (in absolute terms) among the OLS specifications. Estimates of the marginal effect of a new Wal-Mart Supercenter fall from approximately 3% in the cross-sectional regression to 1.2% in the long-difference regression and to 1.1% in the panel regression. These differences are driven by the assumptions required for causal interpretation. The cross-sectional regressions assume that the locations of Wal-Mart stores are not correlated with anything that influences price. Long-difference estimates assume that this is true only of the *new* locations over the period studied and that the timing of entry does not matter for the price impact. Panel estimates assume that the choice to enter a particular market during this 3-year period may be endogenous, but conditional on entry, the exact timing is uncorrelated with other determinants of prices. The fact that the panel estimates, which impose the weakest conditions on causal interpretation, provide the smallest point estimates is consistent with the common perception that Wal-Mart tends to open stores in low-price locations.

There is another difference between the cross-sectional, long-difference, and panel OLS estimates, however, that has to do with the time horizon over which Wal-Mart's Supercenters can affect the competition. The panel regression has the shortest horizon, in the sense that we only allow contemporaneous — that is, short-run — effects of the change in the number of Supercenters on competitors' prices. The long-difference regression allows the effects to accumulate over the study period. Because entry of Supercenters in our sample cities is spread approximately uniformly over the 3-period, the long-difference estimates represent an average of the 1-year, 2-year, and 3-year price changes due to competition

with a Supercenter. That point estimates in the long-difference specification are slightly larger than those in the panel specification suggests that the medium-run (2–3 year) effects of Supercenter competition may be slightly larger than the short-run effects. The cross-sectional estimates take this a step further. Several of the Supercenters in our sample, such as the ones in Augusta, GA, and Jefferson City, MO, are among the very first Supercenters to open in the late 1980s and 1990. Others opened just before, or during, the sample period. As a result, in addition to the omitted-variable bias, the cross-sectional confound short-, medium-, and long-run effects.

Overall, these results are similar to estimates in Basker (2005) for Wal-Mart’s effect on drugstore prices. Table 5 in Basker (2005) shows short-run price declines for 7 drugstore and 3 clothing items averaging 1.1%, with the effect on drugstore items alone averaging 1.3% (clothing prices were essentially unaffected), and long-run price declines of 5.2% and 7.3%, respectively.

4.1.4 Differential Responses by Chain Size

Before turning to alternative identification methods, we test to see whether the “Big Three” supermarket chains — Albertson’s, Kroger, and Safeway — respond differently to Wal-Mart’s competition than do other, smaller, supermarkets. There has been some speculation in the press that these chains might suffer most from Wal-Mart’s expansion (see, e.g., Callahan and Zimmerman, 2003). Combined, these three firms had approximately \$140 billion in sales in 2005, accounting for more than a quarter of total food- and grocery-store sales in the U.S. that year (U.S. Census Bureau, 2006; Supermarket News, 2007).¹⁵ Approximately 23% of the non-Wal-Mart price observations in the data are from the Big Three supermarkets.

¹⁵Since the end of the sample period, Albertson’s was acquired by Supervalu.

The panel specification is given by

$$\begin{aligned}
 p_{ijkt} = & \alpha + \rho \mathbf{big3}_{ijkt} + \theta \text{WMSC}_{jt} + \psi \mathbf{big3}_{ijkt} \text{WMSC}_{jt} \\
 & + \sum_j \gamma_j \text{city}_j + \sum_{kt} \delta_{kt} \text{product}_{k,\text{year}_t} + \varepsilon_{ijkt} \quad (4)
 \end{aligned}$$

where \mathbf{p}_{ijkt} is the average non-Wal-Mart log price of product k in city j at time t in store type i (i is either the Big Three, or non-Big Three stores). The variable $\mathbf{big3}_{ijkt}$ is an indicator that equals 1 when the price is from the Big Three. The cross-sectional and long differences formulations are a similar adaptation of Equations (1) and (2). Given our earlier finding that the cross-sectional and long-difference results were slightly closer to the panel estimates when control variables are omitted, no additional control variables are included in these regressions. (Results are insensitive to this omission.)

We report the results in Table 3. The Big Three’s prices are, on average, 2-3% higher than other stores’ (depending on whether we use the cross-sectional or the panel estimates), and their position relative to other stores’ prices is virtually unchanged over this period (from the long-difference estimates). The main result is that the Big Three’s response to Wal-Mart’s entry is muted relative to other stores’ response. In the panel specification, which we trust most, we estimate that while other stores lower prices, on average, by 1.26% when a new Supercenter opens, the Big Three’s price response is a price reduction of a statistically insignificant 0.48%.

These results give support to the notion that smaller retailers are the most affected by Wal-Mart’s competition. They also help explain why Singh, Hansen, and Blattberg (2006) do not find strong effects of a Wal-Mart Supercenter entry on prices charged by a competing traditional supermarket. One possible explanation is that many supermarket chains have a “uniform pricing” policy whereby prices are set centrally for a broad geographic area.¹⁶ It is

¹⁶Uniform pricing is used, for example, at Dominick’s Finer Foods in Chicago (see Hoch, Kim, Montgomery, and Rossi, 1995; Chevalier, Kashyap, and Rossi, 2003). Conversations with managers at other supermarket

also possible that the large supermarket chains attempt to differentiate themselves from Wal-Mart by offering higher service levels, wider selection, or other amenities for which (some) consumers are willing to pay a premium.¹⁷ Further evidence of differentiation between Wal-Mart and the Big Three is provided by Ellickson and Misra (forthcoming), who cite evidence that Wal-Mart sets relatively stable prices from week to week (“every day low pricing”) in three quarters of its stores, while the Big Three use this pricing scheme only in a fifth of their stores, more often opting for frequent price promotions or a hybrid scheme.

Another possibility is that, when a new Wal-Mart Supercenter opens, it attracts many consumers who are sensitive to price differences; consumers who continue to shop at Big Three stores may have less price-elastic demand, inducing those stores to increase their relative prices. This market-segmentation hypothesis is consistent with evidence from Singh, Hansen, and Blattberg (2006) that households that tend to purchase name-brand items (rather than the cheapest, store-brand, varieties) and those buy more speciality meats and home meal replacements are less likely to “defect” to Wal-Mart.¹⁸

4.2 Instrumental Variables

While we are comfortable with the weak exogeneity assumptions in the long-differences and panel regressions, omitted-variable bias in the cross-sectional analysis prevents us from making a causal interpretation of $\hat{\theta}$, the coefficient estimate from Equation (1). Our goal in this section is to examine two instrumental-variables methods that might alleviate this

chains reveal that this is a common practice. This strategy could well be profit maximizing if there are substantial managerial “menu costs” involved in price setting (see, e.g., Levy, Bergen, Dutta, and Venable, 1997), or if used as a strategic device to “soften” competition in contested markets (see Dobson and Waterson, 2005, for a model of this phenomenon).

¹⁷This view is implicit in the Bureau of Labor Statistics’ method for addressing price differences across establishments. In contrast, Hausman and Leibtag (2004) argue that Wal-Mart and other grocery stores offer homogeneous products.

¹⁸In a similar vein, Forman, Ghose, and Goldfarb (2007) find that the mix of products sold to any given market by Amazon.com changes to more-obscure titles when a Wal-Mart or Target store opens in that market. To borrow the terminology of Betancourt (2005), Wal-Mart, small grocers, and the Big Three supermarket chains all offer different combinations of distribution services, including depth of assortment.

problem and move the point estimates in the cross-sectional analysis towards the more-plausible results obtained in the long-difference and panel specifications.

That said, we do not expect the IV results to match the panel and long-difference OLS estimates exactly. As noted above, absent endogeneity concerns, cross-sectional estimates of Wal-Mart's price impact represent a weighted average of short- and long-run effects, whereas panel estimates shed light on Supercenters' short-run impact only.

The first instrument we consider is the distance from a market to Wal-Mart's nearest food Distribution Center (DC). Because of Wal-Mart's historical expansion pattern in circles around Bentonville, Arkansas (see Holmes, 2006), geographically-based instruments have been popular in recent papers studying its local effects. The argument for this instrument is that operating a store near a food DC is cheaper than operating a store further away due to lower trucking costs, so distance to a food DC may predict Wal-Mart's Supercenter locations. Two confounding factors for this instrument are, first, that unobserved market characteristics are spatially correlated, and second, that the distance to the nearest DC has a direct effect on marginal cost and therefore on Wal-Mart's — and competitors' — prices. As a result, the instrument is likely to be correlated with the error term in the second-stage regression.

We then consider a second instrument: the number of pre-existing discount stores in a market in a base year. Because conversion of a discount store into a Supercenter is likely cheaper than *de novo* entry into a market, the number of discount stores is likely to predict Supercenter entry without having a direct effect on grocery prices. Unlike the distance instrument, the number of pre-existing discount stores affects only the *fixed* cost of entry, and not the marginal cost of goods, so it is less likely to be correlated with the error term. As expected, we find that the first instrument *increases* the bias in the cross-sectional estimates whereas the second reduces the omitted-variable bias and pushes the cross-sectional estimates towards the panel and long-difference estimates we believe.

4.2.1 Distance to Distribution Center

Wal-Mart supplies its stores primarily through several dozen Distribution Centers (DCs) around the country. As of July 2001, only 20 DCs dealt with groceries; between July 2001 and July 2004, twelve more food DCs were opened. According to data from Wal-Mart Watch, an anti-Wal-Mart organization, the median DC serves approximately 90 stores within a 250 mile radius. If it is cheaper for Wal-Mart to supply markets the closer they are to a food DC, we should observe more Supercenters in closer markets. This logic leads us to use each market's distance to the nearest food DC to instrument for the number, or change in the number, of Supercenters. However, this argument implies that, conditional on opening a Supercenter in a market, distance from the nearest food DC could also affect the store's costs and therefore prices. Thus, the instrument may still be correlated with the error term in Equation (1).

A second problem is that Wal-Mart determines the current (and future) locations of its Supercenters and DCs simultaneously (see Holmes, 2006, for a model of this dynamic problem). Even though each DC serves many stores, so that any one market has a minimal effect on the location of a DC, market characteristics tend to be spatially correlated, which can exacerbate the endogenous element in the DC location decision.

Table 4 reports results using linear distance (as the crow flies) to the nearest DC as the instrument, as well as using the exponent of distance.^{19,20} First-stage F statistics are in the range of 10–20 (except for one specification which has an F statistic of 42). Second-stage estimates are 3–5 times larger in absolute value than the cross-sectional OLS estimates from Table 2, which were themselves larger than the more-plausible panel and long-difference

¹⁹Estimates using log distance had a weak first stage.

²⁰Because the number of food DCs increased over the time period studied, with some markets getting “closer” to a DC while other were unaffected, in principle this instrument can be used in a panel regression with city fixed effects, as well as the long-difference regression. In practice, both the panel and long-difference specifications have extremely weak first-stage results (F statistics of 1 or below), rendering the IV results meaningless, and we do not report them here. Results for the other cross-sections are indistinguishable from the ones reported here.

estimates. The IV estimates suggest a price impact of 12-17% per Supercenter, more than 10 times our other estimates. These numbers are implausible in light of the fact that grocery stores' gross margins — the fraction of revenues not accounted for by wholesale costs — average below 30% (U.S. Census Bureau, 2006).

We conclude from this exercise that the instrument is not valid for our purpose. One must be cautious about distance-based instruments in this context, because of spatial correlation and the potential for a direct effect of distance on price. In this case these problems exacerbate the selection problem instead of correcting for it.²¹

4.2.2 Discount Store Conversions

Many of Wal-Mart's Supercenters are conversions from pre-existing discount stores. (In keeping with Wal-Mart's terminology, we refer to a store as a discount store if it does not sell a full line of groceries. Discount stores may sell soda, snacks, and the like but do not carry most of the grocery items in our price sample.) One reason for this is that such entry or expansion of the chain is cheaper than *de novo* entry into a market. We exploit this fact by using the number of pre-existing discount stores in a market to instrument for Supercenter entry. The instrument is valid if the number of pre-existing discount stores is uncorrelated with unobservable characteristics of the market that also affect competitors' grocery prices, before and after Wal-Mart's (potential) entry into the grocery market. To minimize such potential correlations we use the number of Wal-Mart discount stores in a relatively early time period, before Wal-Mart's introduction of the Supercenter format became widespread. We report estimates using base years ranging from 1988 to 1997.

In Table 5 we present IV estimates for Equation (1).²² For each base year, we show the

²¹A similar problem arises when distance from Wal-Mart's corporate headquarters in Bentonville, Arkansas, is used to instrument for store openings; see Basker (2006) for a discussion.

²²Because the instrument takes on a single value for each city — the number of discount stores in year t_0 — we cannot employ it in the panel setting. In principle, it can be used in the long-difference specification, but first-stage F statistics are always below 10, indicating that the instrument is too weak for the second-stage results to be interpretable (Staiger and Stock, 1997).

first-stage F statistic and the estimated second-stage (IV) coefficient. In all these regressions, the first-stage results are strong and significant, with F statistics ranging from 30–50. In every case, the first-stage coefficient (not reported) is positive and ranges from 0.5 to 0.6. IV estimates of competitors' response range from 2–3% price reductions, falling as the base year increases. Interestingly, F statistics are highest for the early base years: the number of discount stores in 1988 is a better predictor of Supercenters in 2002 than is the number of discount stores in 1997. This is probably because many discount stores were converted to Supercenters between 1988 and 1997, so a location that still has many discount stores in 1997 is likely to be a relatively bad location for a Supercenter. At the same time, the early base years also yield coefficient estimates that are larger in absolute value than the OLS estimates, suggesting that the same selection concerns that apply to Supercenters in 2002 may also apply to the choice of locations of discount stores in earlier years. This implies that the number of discount stores in later base years, 1997 in particular, may be the most reliable instrument.

In the next section, we conduct a falsification exercise to test all our specifications and identification strategies, and confirm that the panel and long-difference OLS specifications, and the discount-store IV specification, provide robust and trustworthy results.

4.3 Falsification Exercise

So far, we have taken the position that the panel and long-difference results are most plausible because they impose the weakest exogeneity conditions and we have used the panel results as a benchmark against which to evaluate the IV estimates. In this section, we provide a falsification exercise that further bolsters this argument.

The ACCRA data include prices of several services Wal-Mart does not provide and for which we expect to find no economically meaningful price impact due to Wal-Mart's entry. If we were to find an impact following Wal-Mart's entry, it would suggest that unobservable aggregate demand or cost differences that are correlated with Wal-Mart's entry are driving,

or contributing to, our estimated price effects. We use prices of six services to test all the specifications above. The services are: appliance repair (home service call for clothes washing machine repair; price includes minimum labor charges if applicable but excludes parts); movie ticket (adult admission for showings of first-run films on a Saturday evening at indoor movie theaters); bowling (one bowling lane at non-league rates on a Saturday evening, excluding equipment rental); man’s haircut (“standard” haircut at regular barbershops, excluding styling, razor cuts, etc.); woman’s beauty salon appointment (shampoo, trim, and blow-dry for shoulder-length or shorter hair, no style change, at beauty shops that accept appointments and allow the client to select her beautician); and dry cleaning (of a man’s two-piece suit).

We report OLS estimates for the cross-sectional, long-difference, and panel equations in Table 6. Except for the cross-sectional specification, the estimates are very close to zero and statistically insignificant despite tight confidence intervals. The cross-sectional estimates are statistically different from zero, consistent with our concern that differences in the number of Wal-Mart stores across cities cannot be treated as exogenous, and unobserved demand and cost differences across markets are in part driving the estimates. These estimates confirm our concern that endogeneity *increases* the estimated effect of Wal-Mart’s Supercenters in the cross-sectional regressions, because Wal-Mart’s traditional markets have been in low-income, low-cost cities. The cross-sectional estimates are only about 1 percentage point smaller (in absolute terms) than the corresponding estimates for grocery products shown in Table 2. One way to interpret this 1% figure is as a difference-in-difference estimate, where the two “difference” dimensions are the number of Wal-Mart stores in a market and whether the product is sold by Wal-Mart, and unobservable city-specific effects are differenced out. The identifying assumption is that unobserved shocks to cost and demand affect grocery and service prices equally. The difference of 1% is, in fact, almost exactly the estimate we get in the difference-in-difference panel regression in the last column of Table 2, in which the difference dimensions are the number of Wal-Mart stores in a market and time.

The long-difference and panel point estimates are very small and statistically insignifi-

cant, adding to our confidence in these specifications. The panel estimates are particularly small, at less than one fifth of one percent. We conclude from these falsification exercises that the indirect effect of Wal-Mart entry on competitors' prices, estimated in the long differences and panel specifications, are true competitive effects. They do not appear to be driven by correlation with unobserved aggregate or city-specific demand shocks. Effectively, our panel estimate is also the difference-in-difference-in-difference estimate, which exploits differences in three dimensions: the number of Wal-Mart stores in a market, grocery products sold at Wal-Mart versus services not provided at Wal-Mart, and time.

Table 7 shows the same falsification exercise for our cross-sectional IV specifications. We report six coefficient estimates. The first two show estimated effects from the distance-to-DC instrument using, respectively, linear and exponential distance measures. Both estimates are implausible at about -9% and significantly different from zero at the 5% level, confirming our concerns that this instrument is invalid. The next four columns use different base years for the discount-store IV specifications: 1988, 1991, 1994, and 1997. As before, all four have strong first stages, demonstrated by F statistics in the 30–50 range. All point estimates are insignificantly different from zero, with the exception of the first (which is significantly different from zero only at the 10% level). Consistent with our finding that later base years produced smaller second-stage estimates of Supercenters' impact on competitors' grocery prices, we find that estimates of Wal-Mart's price effect on these six services using the 1997 base year instrument are very close to zero. This increases our confidence in this IV for the cross-sectional regression, although the long-difference and panel regressions are still preferred.

5 Conclusion

Wal-Mart rose to the top of the retail food chain very fast, becoming the largest grocer in the United States in 2002, only fourteen years after opening its first Supercenter. Wal-

Mart’s entry into the grocery business has not been uncontroversial, and while consumers shop at Wal-Mart in record numbers, many people have been critical of Wal-Mart and concerned about its impact on competing businesses. This paper contributes to the debate by quantifying an important component of the overall impact of Wal-Mart on the grocery industry — namely, its effect on the prices consumers pay.

Based on the evidence presented here, we conclude that the competitive effect of Wal-Mart Supercenter entry on the prices charged by other supermarkets and grocers is 1–1.2%. The conclusion is consistent across our long-difference and panel specifications and is precisely estimated. Our falsification exercises for these specifications, which use products whose prices should be unaffected by Wal-Mart’s entry, confirm that our estimates are not influenced by unobserved demand or cost shocks correlated with Wal-Mart’s entry. In contrast, we caution against causal interpretation of our cross-sectional OLS estimates, which are likely to suffer from omitted variable bias due to unobserved city-level differences that affect Wal-Mart’s entry decision. However, when we difference our cross-sectional estimates (which include the effects of Wal-Mart’s entry and of the omitted variables) and our corresponding falsification exercise estimates (which include the effect of omitted variables only) we recover the same 1% price effect that matches our preferred specifications. We also find that the price effects are concentrated in smaller supermarket and grocery chains. The largest supermarket chains — Kroger, Albertson’s, and Safeway — reduce their prices in response to Wal-Mart’s entry by less than half as much as its smaller competitors. Wal-Mart’s own prices on the basket of food items in our study are about 10% lower than its competitors, inducing this modest competitive response.

Our analysis is based on a unique data set of store-level prices of twenty-four specific grocery items and six non-food items in 175 markets. Because of the specific nature of the products in the data (brand, weight, etc.), we are able identify price effects more cleanly than one could with more aggregated data. In addition, because of the broad coverage of markets, we are able to report price effects more representative than one could with a case-study style

of analysis. Our indirect price effect of 1–1.2% is smaller than one obtained by Hausman and Leibtag (forthcoming) using more aggregated data, but larger than the case-study findings of Singh, Hansen, and Blattberg (2006).

There are many potential dimensions along which competitors may respond to Wal-Mart's entry. In addition to short-run price adjustments, Wal-Mart's entry may also change the quantities or mix of products sold at competing supermarkets, which could have implications for long run prices. For example, in the long run, competing supermarkets — especially inefficient ones — may contract or exit, causing prices to rise (due to less competition) or fall (due to a shift in the average efficiency in firms). Potential entry by Wal-Mart, rather than actual entry, may also have an effect if markets are contestable and incumbents suppress their prices to deter entry. This paper addresses the causal effect of one of the most important and immediate effects, that of short- and medium-run price responses. The large *direct* price effect we find demonstrates that consumers who shop at Wal-Mart receive a sizeable price benefit, while the indirect price effects are quite small, especially for shoppers at the Big Three supermarkets.

References

- Aaronson, D. (2001) “Price Pass-through and the Minimum Wage,” *Review of Economics and Statistics*, 83(1), 158–169.
- American Chamber of Commerce Research Association (2000) “ACCRA Cost of Living Index Manual,” .
- Basker, E. (2005) “Selling a Cheaper Mousetrap: Wal-Mart’s Effect on Retail Prices,” *Journal of Urban Economics*, 58(2), 203–229.
- (2006) “When Good Instruments Go Bad,” unpublished paper, University of Missouri.
- Betancourt, R. (2005) *The Economics of Retailing and Distribution*. Edward Elgar Publishing, Cheltenham, UK, and Northampton, MA.
- Callahan, P., and A. Zimmerman (2003) “Price War in Aisle 3,” *Wall Street Journal*, May 27, 2003.
- Chevalier, J. A., A. K. Kashyap, and P. E. Rossi (2003) “Why Don’t Prices Rise During Periods of Peak Demand? Evidence from Scanner Data,” *American Economic Review*, 93(1), 15–37.
- Currie, N., and A. Jain (2002) “Supermarket Pricing Survey,” UBS Warburg Global Equity Research.
- Dobson, P. W., and M. Waterson (2005) “Chain-Store Pricing Across Local Markets,” *Journal of Economics and Management Strategy*, 14(1), 93–119.
- Ellickson, P. B. (2004) “Supermarkets as a Natural Oligopoly,” unpublished paper, Duke University.
- (2006) “Quality Competition in Retailing: A Structural Analysis,” *International Journal of Industrial Organization*, 24(3), 521–540.
- Ellickson, P. B., and S. Misra (forthcoming) “Supermarket Pricing Strategies,” *Marketing Science*.
- Forman, C., A. Ghose, and A. Goldfarb (2007) “Geography and Electronic Commerce: Measuring Convenience, Selection, and Price,” unpublished paper, University of Toronto.
- Foster, L., J. Haltiwanger, and C. J. Krizan (2006) “Market Selection, Reallocation and Restructuring in the U.S. Retail Trade Sector in the 1990s,” *Review of Economics and Statistics*, 88(4), 748–758.
- Frankel, D. M., and E. D. Gould (2001) “The Retail Price of Inequality,” *Journal of Urban Economics*, 49(2), 219–239.

- Graff, T. O. (2006) "Unequal Competition among Chains of Supercenters: Kmart, Target, and Wal-Mart," *Professional Geographer*, 58(1), 54–64.
- Hausman, J., and E. Leibtag (2004) "CPI Bias from Supercenters: Does the BLS Know that Wal-Mart Exists?," National Bureau of Economic Research Working Paper 10712.
- (forthcoming) "Consumer Benefits from Increased Competition in Shopping Outlets: Measuring the Effect of Wal-Mart," *Journal of Applied Econometrics*.
- Hoch, S. J., B.-D. Kim, A. L. Montgomery, and P. E. Rossi (1995) "Determinants of Store-Level Price Elasticity," *Journal of Marketing Research*, 32(1), 17–29.
- Holmes, T. (2006) "The Diffusion of Wal-Mart and Economies of Density," unpublished paper, University of Minnesota.
- Lach, S. (forthcoming) "Immigration and Prices," *Journal of Political Economy*.
- Levy, D., M. Bergen, S. Dutta, and R. Venable (1997) "The Magnitude of Menu Costs: Direct Evidence from Large U.S. Supermarket Chains," *Quarterly Journal of Economics*, 12(3), 791–825.
- Mccarthy, M. J. (2005) "Winn-Dixie Files for Chapter 11," *Wall Street Journal*, February 23, 2005.
- Mnyandu, E. (2005) "Winn-Dixie Files Bankruptcy - Loses out to Wal-Mart," *Reuters*, February 22, 2005.
- Parsley, D. C., and S.-J. Wei (1996) "Convergence to the Law of One Price without Trade Barriers or Currency Fluctuations," *Quarterly Journal of Economics*, 111(4), 1211–1236.
- Singh, V. P., K. T. Hansen, and R. C. Blattberg (2006) "Market Entry and Consumer Behavior: An Investigation of a Wal-Mart Supercenter," *Marketing Science*, 25(5), 457–476.
- Staiger, D., and J. H. Stock (1997) "Instrumental Variables Regression with Weak Instrument," *Econometrica*, 65(3), 557–586.
- Supermarket News (2007) "SN Top 75 2006," *Supermarket News*, May 1, 2007, <http://supermarketnews.com/top75/top-75-2006/>.
- Tedlow, R. S. (1990) *New and Improved: The Story of Mass Marketing in America*. Basic Books, Inc., New York.
- U.S. Census Bureau (2006) *Current Business Reports, Series BR/05-A: Annual Revision of Monthly Retail and Food Services: Sales and Inventories—January 1992 Through February 2006*. Washington, DC.
- Warner, M. (2006) "Wal-Mart Extending Dominance of the Grocery Business," *New York Times*, March 3, 2006.

Westerman, P. (2001) *Data Warehousing: Using the Wal-Mart Model*. Morgan Kaufmann Publishers, San Francisco.

Zhu, T., V. Singh, and A. Dukes (2005) “Local Competition and Impact of Entry by a Dominant Retailer,” unpublished paper, Carnegie Mellon University.

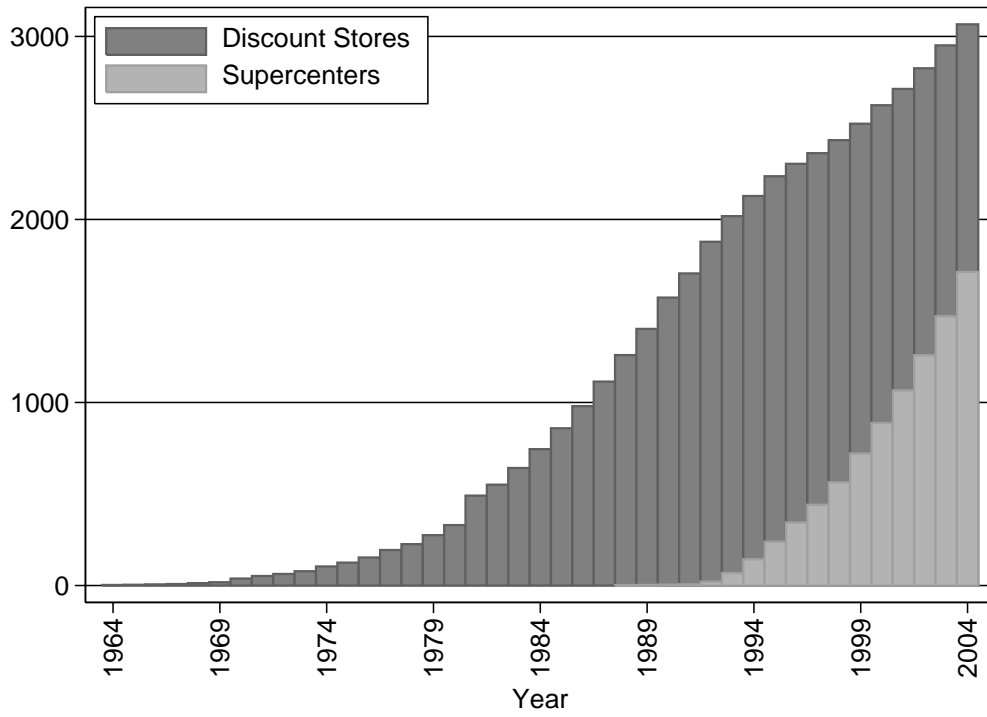


Figure 1. Evolution of Wal-Mart Stores

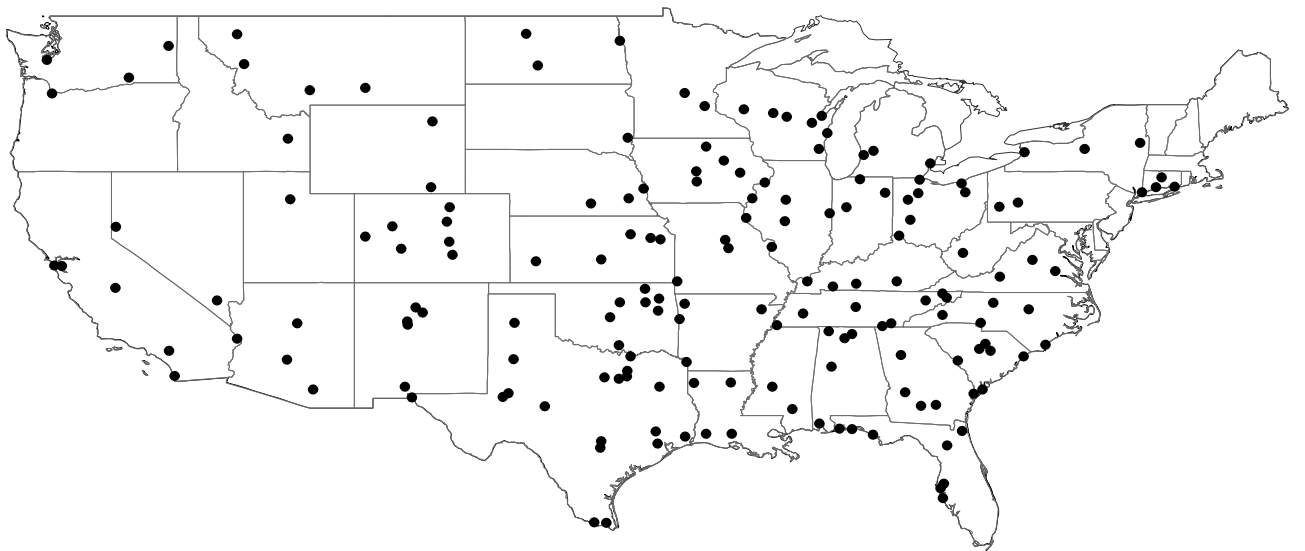
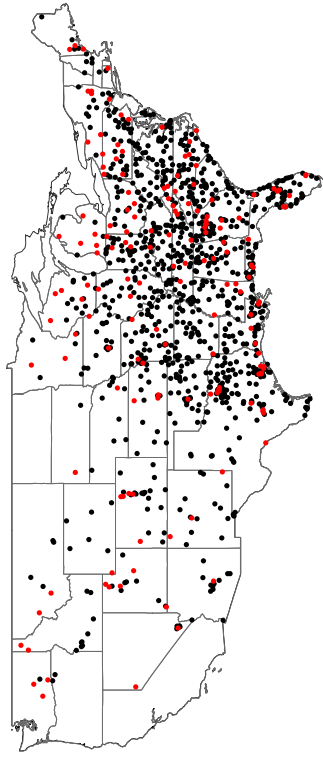
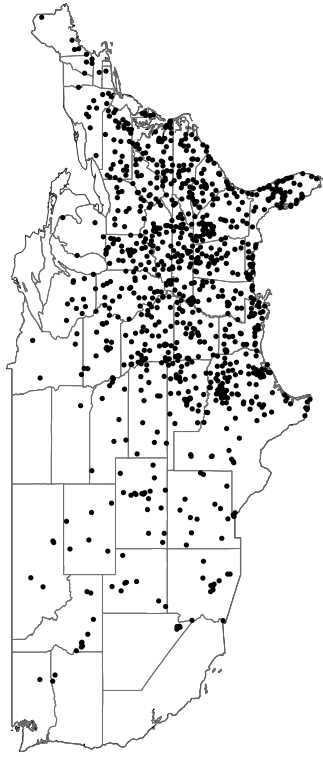


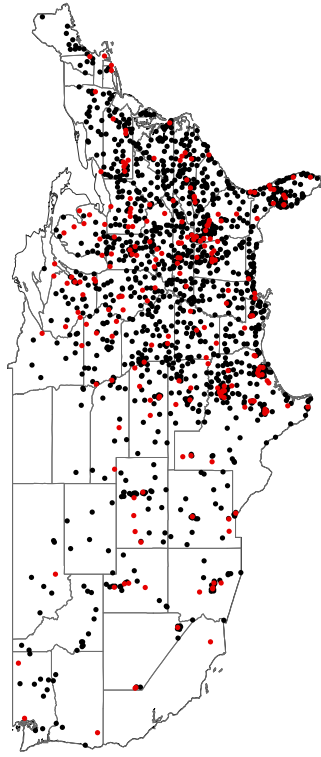
Figure 2. Sample Cities



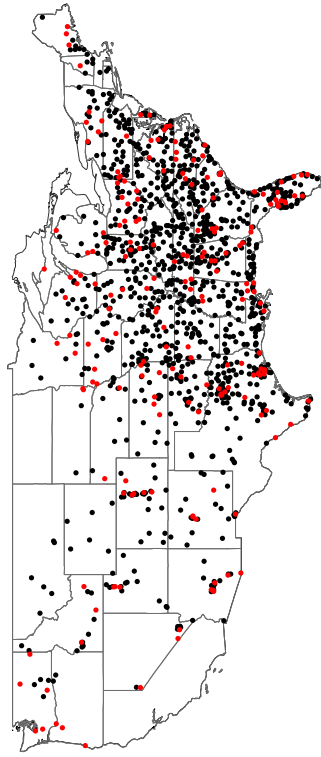
(a) July 1, 2001



(b) July 1, 2002



(c) July 1, 2003



(d) July 1, 2004

Figure 3. Wal-Mart Supercenter Locations (New Locations in Red)

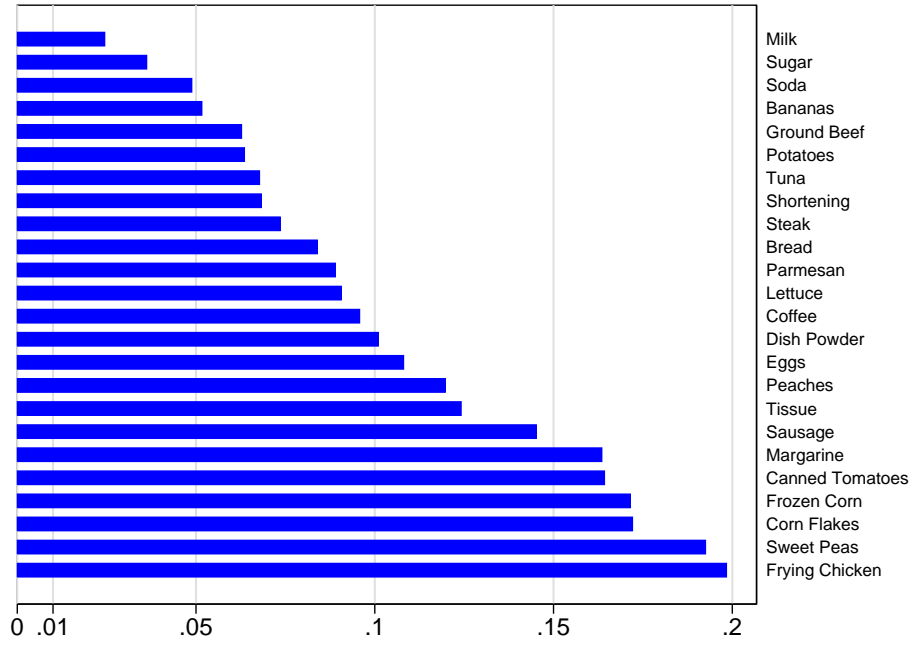


Figure 4. Raw Price Differences by Product

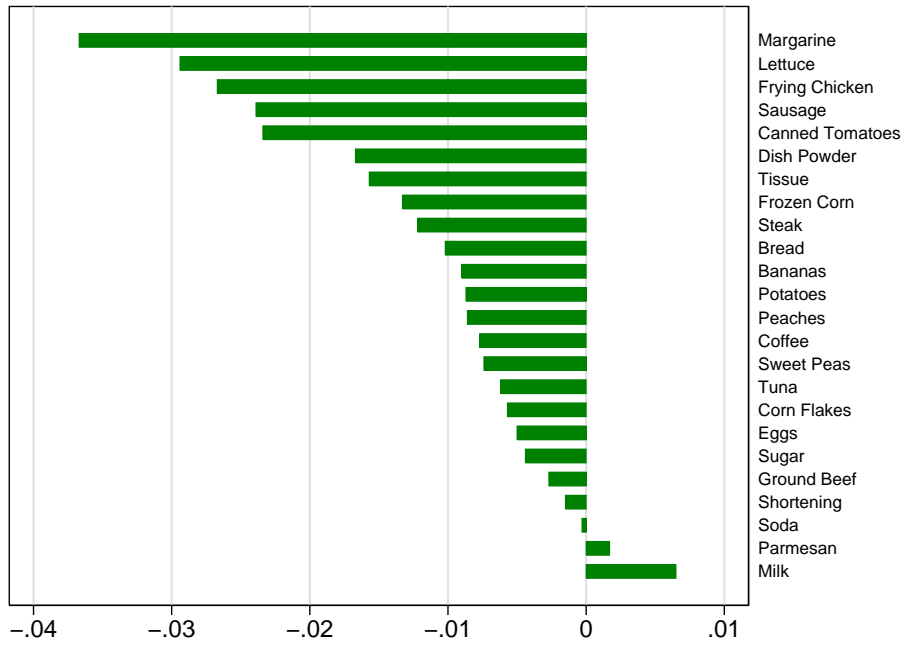


Figure 5. OLS Panel Fixed Effects Estimates by Product

Table 1. ACCRA Products

Product	Description	Average Price
Bananas	Price per lb.	0.50
Lettuce	Head of iceberg lettuce	1.09
Potatoes	10 lb. sack	3.49
Canned Tomatoes	14.5 oz. Hunts/Del Monte	0.98
Canned Peaches	29 oz. Hunts/Del Monte/Libby's/ Lady Alberta halves or slices	1.67
Canned Sweet Peas	15 oz. Del Monte/Green Giant	0.79
Frozen Corn	16 oz. whole kernel	1.18
Eggs	1 dozen grade A or AA large eggs	1.03
Milk	0.5 gal. whole milk	1.82
Margarine	1 lb. Blue Bonnet/Parkay sticks	0.81
Parmesan Cheese	8 oz. Kraft cannister grated parmesan	3.58
Frying Chicken	whole uncut, price per lb.	0.99
Ground Beef	Price per lb.	1.86
Pork Sausage	1 lb. package, Jimmy Dean/Owen	3.24
T-Bone Steak	Price per lb.	7.58
Canned Tuna	6 oz. chunk light tuna, Starkist/Chicken of the Sea	0.67
Bread	Price per oz. for loaf with lowest cost per oz.	0.04
Coffee	11.5 oz. can or brick, Maxwell House/Hills Brothers/Folgers	2.54
Cereal	18 oz. Kellogg's/Post corn flakes	2.63
Shortening	3 lb. can Crisco all-vegetable shortening	3.20
Sugar	4 lb. cane or beat	1.58
Soda	2-liter Coca Cola, excl. deposit	1.17
Dishwashing Powder	75oz Cascade dish washing powder	4.05
Tissue	175-count Kleenex tissues	1.35

All prices are average *nominal* prices for July 2001–2004.

Table 2. OLS Estimates of Wal-Mart's Price Impact

	Cross-Section		Long Difference		Panel
LHS Variable	$p_{jk,2002}$		Δp_{jk}		p_{jkt}
WMSC ^a	-0.0253*** (0.0063)	-0.0317*** (0.0062)	-0.0123*** (0.0037)	-0.0127*** (0.0045)	-0.0109*** (0.0032)
Log Population		0.0206*** (0.0065)		0.0004 (0.0052)	
Log Median Income		0.0679* (0.0382)		0.0067 (0.0341)	
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		4,186		16,773

^a WMSC is Supercenter count in cross-sectional and panel regressions, and the *change* in Supercenter count in the long-difference regression.

Robust standard errors in parentheses (clustered by city)

* significant at 10%; ** significant at 5%; *** significant at 1%

Table 3. Estimates of Average Price Effect by Chain Size

	Cross-Section	Long Difference	Panel
LHS Variable	$p_{ijk,2002}$	Δp_{ijk}	p_{ijkt}
big3	0.0323* (0.0164)	0.0060 (0.0130)	0.0262*** (0.0088)
WMSC ^a	-0.0268*** (0.0068)	-0.0138*** (0.0052)	-0.0126*** (0.0034)
big3 \times WMSC	0.0091 (0.0087)	0.0066 (0.0085)	0.0078* (0.0045)
Observations	6,407	6,371	25,735

All regressions include product FE.

Panel regression also includes city and product \times year FE.

^a WMSC is Supercenter count in cross-sectional and panel regressions, and the *change* in Supercenter count in the long-difference regression.

Robust standard errors in parentheses (clustered by city)

* significant at 10%; ** significant at 5%; *** significant at 1%

Table 4. Cross Sectional IV Estimates Using Distance to Distribution Center

Instrument	Linear Distance		Exponential Distance	
WMSC	-0.1678*** (0.0640)	-0.1240*** (0.0411)	-0.1424*** (0.0412)	-0.1006*** (0.0254)
Controls	N	Y ^a	N	Y ^a
First Stage F Statistic	10.2	12.9	42.6	19.6

LHS variable is $p_{jk,2002}$. All regressions include product FE.

^a Log 2000 population and log median income

The number of observations in each regression is 1,049.

Robust standard errors in parentheses clustered by city

* significant at 10%; ** significant at 5%; *** significant at 1%

Table 5. Cross Sectional IV Estimates Using Pre-Existing Discount Stores

IV Base Year	First Stage F Statistic	IV Coefficient	IV Base Year	First Stage F Statistic	IV Coefficient
1988	45.9	-0.0322*** (0.0111)	1993	49.7	-0.0264** (0.0105)
1989	59.8	-0.0281*** (0.0102)	1994	53.8	-0.0232** (0.0102)
1990	49.3	-0.0284*** (0.0105)	1995	40.4	-0.0208** (0.0102)
1991	52.9	-0.0274*** (0.0103)	1996	40.8	-0.0206** (0.0102)
1992	49.5	-0.0279*** (0.0106)	1997	33.1	-0.0187* (0.0102)

LHS variable is $p_{jk,2002}$. Each base year represents a different regression.

All regressions include product FE.

The number of observations in each regression is 4,191.

Robust standard errors in parentheses clustered by city

* significant at 10%; ** significant at 5%; *** significant at 1%

Table 6. Falsification Exercises (OLS)

LHS Variable	Cross-Section		Long Difference		Panel
	$p_{jk,2002}$		Δp_{jk}		p_{jkt}
WMSC ^a	-0.0153** (0.0067)	-0.0228*** (0.0059)	-0.0036 (0.0038)	-0.0043 (0.0068)	-0.0018 (0.0027)
Log Population		0.0228*** (0.0065)		0.0010 (0.0048)	
Log Median Income		0.1650* (0.0416)		0.0082 (0.0293)	
Product FE	Y	Y	Y	Y	Y
Product \times Year FE	N	N	N	N	Y
City FE	N	N	N	N	Y
Observations	1,040		1,045		4,194

^a WMSC is Supercenter count in cross-sectional and panel regressions, and the *change* in Supercenter count in the long-difference regression.

Robust standard errors in parentheses (clustered by city)

* significant at 10%; ** significant at 5%; *** significant at 1%

Table 7. Falsification Exercises (IV)

Instrument	Distance to Food DC		Pre-Existing Discount Stores			
	Linear	Exponential	1988	1991	1994	1997
WMSC	-0.0941** (0.0468)	-0.0882** (0.0368)	-0.0217* (0.0115)	-0.0158 (0.0107)	-0.0103 (0.0101)	-0.0029 (0.0108)
First Stage	10.2	42.6	45.6	52.4	53.4	33.1
F Statistic						

LHS variable is $p_{jk,2002}$. All regressions include product FE.

Robust standard errors in parentheses (clustered by city)

The number of observations in each regression is 1,049.

* significant at 10%; ** significant at 5%; *** significant at 1%