Job Creation or Destruction?

Labor-Market Effects of Wal-Mart Expansion

Emek Basker *,†

University of Missouri - Columbia

November 2002

Abstract

The phenomenal expansion of Wal-Mart provides a clean case for studying the labor-market effects of increased efficiency. I estimate the effect of Wal-Mart entry on retail employment at the county level. Using an instrumental-variables approach to correct for both measurement error in entry dates and possible endogeneity of the timing of entry, I find that Wal-Mart entry increases retail employment by 100 jobs in the year of entry. Half of this gain disappears over the next five years, leaving a statistically significant net gain of 50 jobs at the five-year horizon. The decline in retail employment in the years immediately following entry is associated with the closing of both small and large retail establishments. At the same time, retail employment in neighboring counties declines by approximately 30 jobs, and wholesale employment in the entered county declines by a similar number.

*emek@missouri.edu. I thank Daron Acemoglu, Olivier Blanchard and Sendhil Mullainathan for their support and guidance throughout this project. I have also benefited from conversations with Josh Angrist, Saku Aura, David Autor, Glenn Ellison, Bengte Evenson, Amy Finkelstein, Guido Kuersteiner, Jeffrey Miron, Whitney Newey, Marko Terviö, Ken Troske and seminar participants at the Board of Governors of the Federal Reserve System, Case Western Reserve University, the Federal Reserve Banks of Atlanta and St. Louis, Hebrew Univeristy, MIT, the Univesity of Chicago GSB, the University of Missouri, Tel Aviv University, the US Census Bureau, Washington University in St. Louis and Wellesley College. I also thank Amy Mok, Steven Sadoway and David Von Stroh for their help with typing, and Dorothy Carpenter, Kathy Cosgrove, Rich Lindrooth, Erich Muehlegger, Mike Noel, Jon Zinman, and especially Maurice Drew, Catherine Friedman and Steven Sadoway for helping me gather data for this project.

†This paper was completed while visiting the Federal Reserve Bank of St. Louis, which I thank for its hospitality. Needless to say, the views expressed are not necessarily those of the Federal Reserve Bank of St. Louis or the Federal Reserve System.

"By contributing overwhelmingly to the productivity growth jump in general merchandise retail, Wal-Mart demonstrates the impact that managerial innovation and effective use of IT can have on market structure, conduct, and performance."
– McKinsey Global Institute, 2001

1 Introduction

A recent study by McKinsey Global Institute (2001, henceforth MGI) attributes the increase in the productivity growth rate in the 1990s to only six industries: retail, wholesale, securities, telecommunications, semiconductors, and computer manufacturing. Within the retail industry, Wal-Mart has emerged as a clear industry leader. The MGI report states unequivocally that Wal-Mart has "directly and indirectly caused the bulk of the productivity acceleration through ongoing managerial innovation that increased competitive intensity and drove the diffusion of best practice (both managerial and technological)" in the general-merchandise subsector. This sentiment is shared by industry analysts and, grudgingly, by Wal-Mart competitors.[1] How has the expansion of Wal-Mart affected local labor markets?

An extensive body of literature in empirical macroeconomics analyzes the effect of technological change on aggregate employment (see, for example, Galí 1999). There is also a growing body of literature in labor economics concerning the effect of firm-level technology adoption on wages and employment within the firm (a recent example is Bresnahan, Brynjolfsson and Hitt 2002).

In light of these thriving areas of research, there is a surprising dearth of studies of the effect of entry (or expansion) of a more efficient firm on an industry. There is convincing evidence that much of the growth in aggregate productivity in the U.S. over the last decade can be traced to entry of more efficient firms and the concurrent exit of less-efficient firms, rather than productivity improvements in existing firms. Roughly 30% of productivity growth in

[1] A Wal-Mart competitor is quoted by the Washington Post saying, "The real problem [with Wal-Mart] is that they're so good at what they do" (1990). For the perspective of industry analysts see Muller 1999 and Feiner 2001.

US manufacturing, and nearly all productivity growth in the retail and service sectors in the last decade can be accounted for by reallocation due to net entry of firms (Foster, Haltiwanger and Krizan 2001 and 2002); these numbers appear roughly in line with figures for other counties (see, for example, Griliches and Regev 1995). Entry of more efficient firms can affect industry-level employment and the distribution of employment across firms within an industry, as well as industry-level output and both input and output prices.

One reason for the scarcity of research in this area may be that entry of new firms, like expansion of existing ones, is rarely exogenous. Because firms respond to local conditions when they decide to open or relocate plants, it is difficult, if not impossible, to disentangle the direct effect of expansion from the indirect effects of the conditions that lead to it. One study that finds a way around this problem is Bertrand and Kramarz (forthcoming 2002), which uses an exogenous source of variation in permits given to large retailers to analyze the effect of large-retailer entry on French labor markets, and finds that regulation limiting entry of large retailers has slowed employment growth in the French retail industry.

This paper contributes to this "missing link" in the literature and concerns just this sort of intermediate analysis with exogenous variation in the timing of store entry. The focus of the paper is on the employment consequences of entry of a more efficient firm – Wal-Mart – into local retail markets.[2] I use a unique data set containing the locations and opening dates of all US Wal-Mart stores and employ a case-study approach to track the effect of Wal-Mart entry on retail employment in the county, as well as on employment in other industries and in surrounding counties. To address endogeneity concerns, I use an instrumental-variables specification that exploits the variable lag between

[2]To see why the sign of the effect of Wal-Mart entry on sectoral emploment is ambiguous, note that firms with higher productivity need to hire fewer workers to produce a constant level of goods or services, but as they lower prices, quantity demanded increases as well. The latter effect is due to a combination of demand substitution between firms in the industry and demand substitution across industries. Which effect dominates depends on the elasticity of demand for the industry's product. This intuition is formalized in a simple model in Appendix A.

store-planning dates and store-opening dates. Store numbers, assigned by Wal-Mart during the planning process, are used to proxy for planning dates.

By examining the dynamics of county-level retail employment in the ten-year period surrounding Wal-Mart entry, I am able to disentangle the immediate effect of Wal-Mart entry from its long-run effect. In the first year after entry, retail employment in the county increases by approximately 100 jobs; this figure declines by half over the next five years as small and medium-sized retail establishments close. I present a similar analysis for other sectors (wholesale employment, which declines by approximately 20 jobs over the five years after entry, and restaurant employment, which increases slightly) and for retail employment in neighboring counties, where I find a decline of 30 jobs over 5 years. This same methodology can also be used to estimate general-equilibrium effects of Wal-Mart entry on total employment in the county, although in this case the general-equilibrium effects are too small to be estimated with precision.

The remainder of the paper is organized as follows. Section 2 provides background information on the retail industry in general and Wal-Mart in particular. Section 3 describes the data. My empirical strategy is explained in Section 4, and evidence of Wal-Mart's productivity advantage over other retailer is provided in Section 5. Results are presented and discussed in Section 6. Section 7 concludes.

2 Background

2.1 The Retail Industry

In a recent study of churning in the retail industry over the period 1987-1997, Foster, Haltiwanger and Krizan (2001) find a very large dispersion in the distribution of productivity (measured as the difference between log real output and log labor input), relative to the dispersion found in the manufacturing industries. They also find extremely high rates of job reallocation, due mostly to entry and exit: 70% of gross job and output creation (destruction) in the retail industries is accounted for by firm entry (exit). While new firms enter throughout the productivity distribution, exit is concentrated at the lower tail of the productivity distribution; this fact drives productivity growth in the retail industry. Unlike the manufacturing industries, without entry and exit, the retail industries would have experienced no productivity growth over the period studied.

Mass merchandisers, such as Wal-Mart and its main competitors K-Mart and Target, differ from traditional retailers in that they sell a large variety of products at low prices. Data from the 1997 Census of Retail Trade show that about 7.5% of retail workers were employed by "discount or mass merchandising department stores" the week of March 12, 1997. A recent study of international productivity differences across industries found that traditional stores in the United States are only 60% as productive as U.S. mass merchandisers (Baily and Solow 2001).

2.2 Wal-Mart

> "The purchasing power of a chain unquestionably gives it certain advantages. But I believe that it owes much, if not most, of its success to the intelligence with which it is operated."
> – W.D. Darby, *The Story of the Chain Store*, 1928.

The first Wal-Mart store opened in Benton County, Arkansas in 1962. By the time the company went public in 1969, it had 18 stores throughout Arkansas, Missouri, and Oklahoma. The company slowly expanded its geographical reach, building new stores and accompanying distribution centers further and further away from its original location, and continued, at the same time, to build new stores in areas already serviced. Figure 1 shows maps of the 48 contiguous states with approximate locations of Wal-Mart stores over time to illustrate this point. By 1998 Wal-Mart had approximately 2400 stores in all 50 states and about 800,000 employees in the United States. At the end of 2001 Wal-Mart had 1.2 million employees worldwide, of which about 962,000 (77%) were employed in the United States.

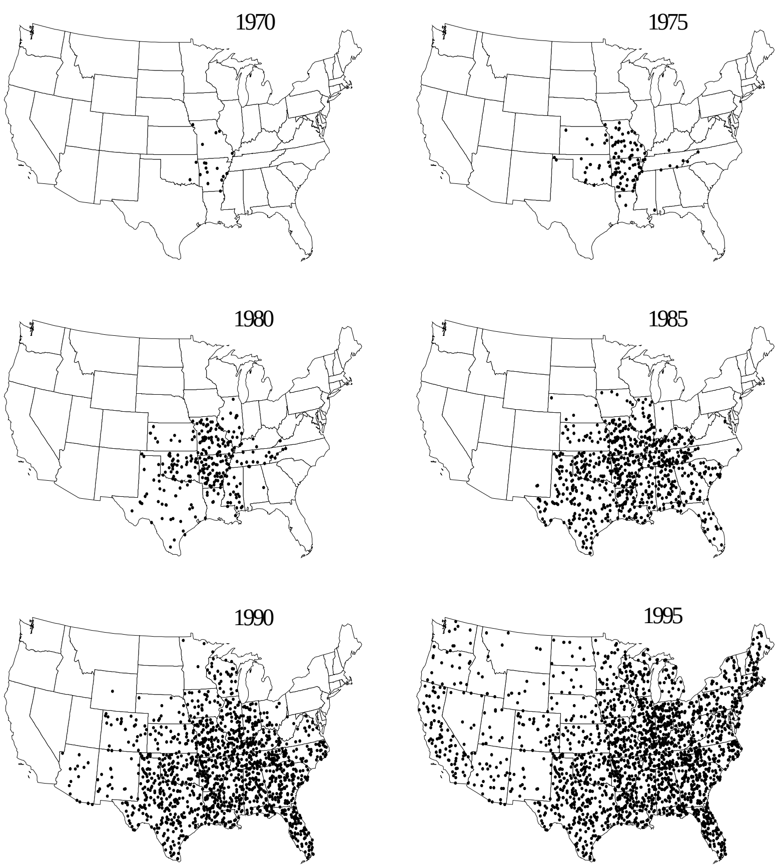

Figure 1: Location of Wal-Mart Stores, Various Years

Wal-Mart is the largest retailer in both the United States and the world. The company operates Wal-Mart discount stores as well as Wal-Mart "supercenters" which include grocery departments and constitute approximately one third of all current Wal-Mart stores. The typical Wal-Mart store spans 100,000-150,000 square feet and employs 150-350 people, many of them in part-time jobs. By 1998, one quarter of the 1614 counties entered by Wal-Mart had more than one store; of these, 234 counties had two stores, and 151 counties had 3 or more stores (among them Harris County, Texas, with 19 stores in 1998).

Wal-Mart is extremely efficient even compared with other "big-box" retailers. It has been cited for its technological advantage by many industry analysts. Lehman Brothers analysts have noted Wal-Mart's "leading logistics and information competencies" and compared it favorably to the world's second-largest retailer, Carrefour, saying that Wal-Mart's "operational and technological superiority has allowed Wal-Mart to gain a comparative advantage over every competitor it has faced, including Carrefour" (Feiner 2001). The *Financial Times* is more expressive, calling Wal-Mart "an operation whose efficiency is the envy of the world's storekeepers" (Edgecliffe-Johnson 1999).

Wal-Mart's competitive edge is driven by a combination of conventional cost-cutting and sensitivity to demand conditions, and by superior technology: the company uses software-based logistics and distribution systems, and its divisions are well-integrated. Wal-Mart's most-cited advantage over small retailers is probably economies of scale and access to capital markets, whereas against other large retailers, such as K-Mart and Target, commonly-cited factors include:[3]

- Superior logistics, distribution, and inventory control: Wal-Mart's proprietary software, Retail Link, links stores directly to Wal-Mart's distribution centers, and links those directly with suppliers (like GE and Proctor & Gamble) who get daily sales data and are able to plan their own invento-

[3] Many of the details cited here on Wal-Mart's operations are from Harvard Business School's three Case Studies about Wal-Mart (Ghemawat 1989, Foley and Mahmood 1996, and Ghemawat and Friedman 1999). Similar points are also made by McKinsey Global Institute (2001).

ries accordingly. This system has reduced Wal-Mart's inventory costs to levels substantially below its competitors' (Stalk and Hout 1990).

- Size: This distribution network is made even more efficient by the geographic proximity of its many stores; Wal-Mart's size also gives it market power in some goods as well as input markets.
- Cost-conscious "corporate culture".
- Demand-sensitivity: Inventories and prices differ from store to store based on climate and consumer demographics. Reorders are made based on actual store needs (communicated to the nearest distribution center) rather than centralized forecasting, and pricing is competitive given market conditions.

There is no single best measure of productivity in the retail industry. One commonly used measure is sales per square foot. Figure 2 shows sales per square foot at K-Mart and Wal-Mart stores (in nominal dollars) for selected years. By way of comparison, a series of studies by the Urban Land Institute put average sales per square foot for mall stores slightly below K-Mart's sales over the period 1978-1997 (Urban Land Institute, various years). Wal-Mart does well also by other measures of productivity. Figures for sales for employee, cited in Johnson (2002), show Wal-Mart consistently ahead of other firms by a large margin; the effect of Wal-Mart on county-wide sales per employee are investigated in Section 5.

3 Data

3.1 Wal-Mart Stores

I use data on the locations and opening dates of 2,382 Wal-Mart stores in the United States, collected primarily from Wal-Mart annual reports, Wal-Mart editions of Rand McNally Road Atlases and annual editions of the Directory

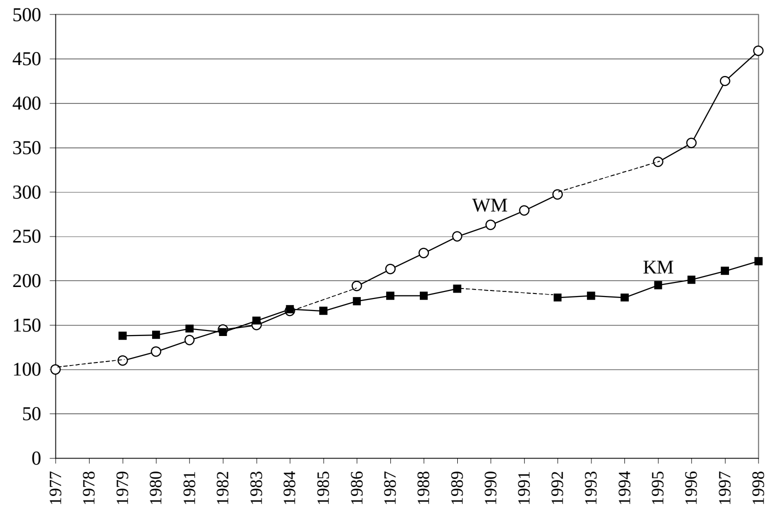

Figure 2: Annual Sales per Square Foot, Wal-Mart vs. K-Mart

of Discount Department Stores. The available data include store location (by town) and store number.

The following data sources, which I refer to collectively as "directories", provide one measure of opening dates: Vance and Scott (1994) list store entries to 1969, the year the company become publicly traded. Annual reports between 1970 and 1978 include lists of current stores. After 1978 annual reports became largely uninformative, listing only the current number of stores per state. The annual *Directory of Discount Department Stores* provides store lists between 1979 and 1993. The directory is published in the beginning of each calendar year, and contains the store list for the end of the previous calendar year. Finally, for recent years I use a special edition of the popular Rand McNally road atlas, sold only at Wal-Mart stores, which contains a list of store locations, and includes each store's company-assigned number. The variable $\mathbf{WMopen_{jt}}$ gives the number of new stores to open in county j in year t based on these directories and store lists.

I also construct an alternate set of Wal-Mart entry identifiers using a combination of company-assigned store numbers (from the Rand McNally atlases) and the net change in the number of stores each year (from company annual reports). This alternate set of entry dates is then used in an instrumental-variables specification to correct for measurement error in, and potential endogeneity of, the timing of entry. Wal-Mart assigns store numbers roughly in sequential order, with store #1 opening first, followed by store #2, and so on. I therefore assign entry dates to stores sequentially, based on their store numbers. This assignment method provides a very good approximation to the true distribution of entry dates and represents the likely order in which the stores were planned. Aggregating these store-level entry dates to the county-year level, I construct $\mathbf{WMplan_{jt}}$: the number of Wal-Mart stores in county j whose store numbers correspond to those opened in year t.[4]

For more details on the construction of the Wal-Mart variables, see Appendix B.1 .

3.2 Labor Market Data

My unit of observation is a county-year. Although there are currently 3111 counties in the contiguous 48 states, some counties have been created (usually by splitting one county in two) and others have merged over the period studied; in those cases, I have merged the observations into one observation for the entire study period.[5] I limit the data set to the 1777 counties with 1964 employment above 1500, positive employment growth between 1964 and 1977, and no Wal-Mart entry prior to 1977. The counties included in the analysis are shown in Figure 3.

Annual county-level employment by SIC (or NAICS) for 1977-1999 comes from the Census Bureau's County Business Patterns (CBP) serial. The panel

[4] Alternatively, WMplan_{jt} gives the number of stores that would have opened in county j in year t had the stores opened in the order in which they were planned.

[5] For details on these newly-created and merged counties, see Appendix B.2.

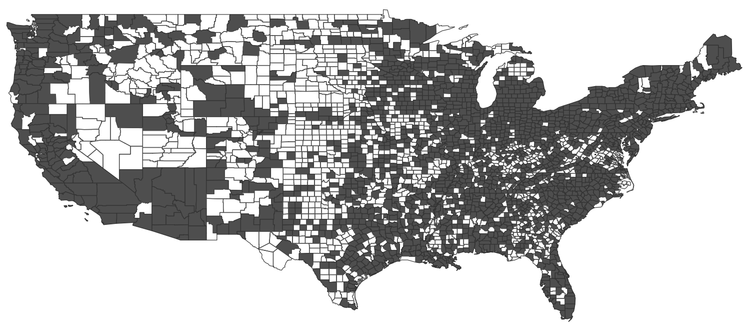

Figure 3: Counties Included in Analysis (Shaded Region)

contains 40,871 observations (1777 counties * 23 years).[6] Unfortunately, no wage data are available from CBP.

Table 1 lists some summary statistics for labor-market data. More details are available in Appendix B.3.

4 Methodology

4.1 OLS Regressions

Because the data do not appear to contain unit roots, the analysis is done using employment levels (see Appendix B.4 for details on unit root tests). The

[6] The relevant SIC (NAICS) codes are:
Retail: SIC 52-- except 5800, NAICS 44; Wholesale: SIC 50--, NAICS 42; Restaurants: SIC 5800, NAICS 721.

Table 1: Summary Statistics

	Sample Counties	Excluded Counties
Total Employment (Mean)	42,000	6,000
Total Employment (Median)	11,000	1,500
Fraction Retail Employment	15.4%	16.8%
Fraction Wholesale Employment	5.0%	5.8%
Fraction with Wal-Mart	75%	13%
Median Number of Small Establishments[a]	172	37
Median Number of Medium Establishments[a]	13	1
Median Number of Large Establishments[a]	1	0

[a] Small establishments: 1-19 employees; medium: 20-99; large: 100+

regressions are:

$$\frac{\mathrm{retail}_{jt}}{\mathrm{pop}_{jt}} = \alpha + \sum_k \sum_t \delta_{tk}\, \mathrm{urban}_{jk}\, \mathrm{year}_t + \sum_j \psi_j \mathrm{county}_j + \theta(L)\frac{\mathrm{WalMart}_{jt}}{\mathrm{pop}_{jt}} + \mathrm{u}_{jt} \quad (1)$$

where retail_{jt} is retail employment in county j in year t; pop_{jt} is the population of county j in year t; year_t is a year dummy; $\mathrm{urban}_{jk} \in \{\text{urban, suburban, rural}\}$ is an urbanization dummy allowing for different year fixed effects for urban, suburban, and rural counties;[7] $\mathrm{WalMart}_{jt}$ is the number of new Wal-Mart stores built that year in county j; county_j is a county dummy; and $\theta(L)$ is a lag polynomial with six lags and five leads (the sixth lag represents the collective period "six or more years after year t"; the omitted category (reference period) is six or more years before a given store was opened.[8] Note that employment 6 or more years before entry is normalized to be zero for all counties. The error term u_{jt} is clustered at the county level.[9]

[7] Urban_{jk} = urban if county j was inside an MSA (metropolitan statistical area) in 1960; suburban if it was $\leq$25 miles from the nearest MSA in 1960; and rural otherwise.

[8] In other words,
$\theta(L) = \theta_1 F^5 + \theta_2 F^4 + \theta_3 F^3 + \theta_4 F^2 + \theta_5 F + \theta_6 + \theta_7 L + \theta_8 L^2 + \theta_9 L^3 + \theta_{10} L^4 + \theta_{11} L^5 + \theta_{12} \sum_{\tau \geq 6} L^\tau$ where L is the lag operator and F is the lead operator.

[9] There has been some confusion in the literature about the use of clustered standard errors with fixed-effect models. Kézdi (2001) and Bertrand, Duflo, and Mullainathan (2001)

Both employment and the number of Wal-Mart stores are divided by the current county population, so the coefficients $\theta(L)$ can be interpreted as the effect of one additional Wal-Mart store per-capita on retail employment per capita.[10] Plots of the coefficients $\theta(L)$ are therefore used to show the evolution of employment over a 10-year period, starting five years before and ending five years after Wal-Mart entry into a county. The coefficient θ_{12}, intended to capture the permanent effect of Wal-Mart entry on employment six or more years after entry, is omitted from the graphs because it is identified using a relatively small number of observations.

The OLS estimates are valid if Wal-Mart entry is correctly measured and exogenous to employment changes. Unfortunately, both measures of Wal-Mart entry – WMopen_{jt}, which uses directory data, and WMplan_{jt}, which uses store numbers to impute planned opening dates – are measured with error. Concerns about endogeneity in the timing of entry offer further complications. An instrumental-variables specification is therefore used to correct these problems.

4.2 Measurement Error

Measurement error in the Wal-Mart entry variables – WMopen_{jt} and WMplan_{jt} – takes a particular form: while the entered counties are correctly identified, the *timing* of entry may be incorrectly measured for a variety of reasons.

The timing error in WMopen_{jt} is due to errors in the directories. A particular egregious example of such errors is the lack of updating of the *Directory of Discount Department Stores* between 1990 and 1993. In addition, stores may appear in directories a year or two late, planned stores may appear before they open, and typos can cause a single store to appear multiple times in one year.

show that clustering with fixed-effect models is not a problem, and is, in fact, generally recommended when the autocorrelated process is not well understood.

[10] The use of per-capita terms on both the left- and right-hand sides of Equation (1) could in principle cause a spurious correlation between the variables that would bias the estimated coefficients. In practice, however, the year-to-year variation in county population is small enough that it is not driving the results presented here; similar results arise when retail employment and the Wal-Mart variables are normalized by a constant such as the 1990 population of the county.

The error in WMplan_{jt}, constructed using store numbers, is due to the variable lag between store planning and store entry, as well as to random assignment of store numbers to approximately 40 stores whose numbers are not known.

An instrumental-variables approach, in which one variable is used to instrument for the other, can be used to correct for this measurement error if the measurement errors in the two variables is classical and uncorrelated. That the measurement error across the two variables is uncorrelated seems plausible.[11] But because WMopen_{jt} and WMplan_{jt} are discrete variables, their measurement error is not classical, as noted by Kane, Rouse, and Staiger (1999). This induces bias in the instrumental-variables results reported here.[12]

The 12 first-stage regressions are

$$\frac{\text{WMopen}_{j,t-s}}{\text{pop}_{jt}} = \widetilde{\alpha} + \sum_k \sum_t \widetilde{\delta}_{tk}\, \text{urban}_{jk}\, \text{year}_t + \sum_j \widetilde{\psi}_j \text{county}_j + \widetilde{\theta}(L) \frac{\text{WMplan}_{jt}}{\text{pop}_{jt}} + \widetilde{\text{u}}_{jt} \quad (2)$$

where $s = -6^+, -5, ..., 4, 5$. Predicted values are denoted by $\frac{\widehat{\text{WMopen}}_{jt}}{\text{pop}_{jt}}$, and the second stage is

$$\frac{\text{retail}_{jt}}{\text{pop}_{jt}} = \alpha + \sum_k \sum_t \delta_{tk}\, \text{urban}_{jk}\, \text{year}_t + \sum_j \psi_j \text{county}_j + \theta(L) \frac{\widehat{\text{WMopen}}_{jt}}{\text{pop}_{jt}} + \text{u}_{jt}. \quad (3)$$

[11] This assumption would be violated if some stores, for example in metropolitan areas, experience shorter planning phases – for example due to quicker zoning changes – and were also more likely to appear in the directories sooner, due to better directory coverage. This does not appear to be the case.

[12] Kane, Rouse and Staiger (1999) suggest a GMM estimator to address this problem. Unfortunately, due to the size of the panel and the hundreds of covariates, their solution is not computationally feasible in this setting.

4.3 Endogeneity

Another difficulty in assessing the impact of Wal-Mart entry on the level and composition of county employment is the possible endogeneity of Wal-Mart's entry decision. This endogeneity has two dimensions: Wal-Mart chooses both the locations to enter (location dimension) and the timing of entry into those counties (timing dimension).

If Wal-Mart selects counties whose growth rates exceed those of non-entered counties, a spurious positive effect will be registered by the estimated coefficients $\widehat{\theta}(L)$. To address this concern, I limit the analysis to counties that constitute a good control group for entered counties: counties with a 1964 population above 1500 and a positive average growth rate of total employment between 1964 and 1977. Finally, I remove counties entered by Wal-Mart before 1977 to eliminate concerns about the endogeneity of employment growth. Wal-Mart entered 75% of the remaining 1777 counties between 1977 and 1998, compared with only 13% of the excluded counties.[13]

Moreover, the timing of entry may be endogenous to employment outcomes if Wal-Mart enters counties during growth spurts (or, what is less likely, during temporarily slowdowns). If entry is timed to coincide with growth spurts, estimated coefficients would reveal a spurious positive relationship between Wal-Mart entry and employment growth.

The instrumental-variables strategy described also corrects for this endogeneity concern. The correction is valid if store numbers represent store planning dates, plans are made well in advance of entry, and, as above, the lag between the two measures of entry is independent of employment outcomes. The first condition appears to hold. Consider, for example, stores 762, 763, and 764. All three are located in Jefferson County, Alabama, and their sequential numbering suggests they were probably planned together. Two of them (763 and 764) opened in 1984, while the third opened in 1990. If this difference in

[13] Indistinguishable results are obtained if the sample is limited instead to entered counties.

entry dates is random, the IV strategy should be valid.[14]

The second condition is that planning must be done sufficiently in advance of entry that spurts in employment growth cannot be reasonably forecasted for the year of entry. Ideally, the order of store entry would have been planned before any stores opened (circa 1960). A more likely scenario is that Wal-Mart determines entry into blocks of counties at regular intervals, which would reduce the endogeneity of the timing of entry.

Finally, lags in entry must be uncorrelated with employment outcomes. This assumption seems reasonable in general, but it would not hold if, for example, towns that resisted and delayed Wal-Mart entry had a disproportionate number of inefficient incumbents that closed after Wal-Mart's entry, or if the residents of such towns were more (or less) avid shoppers than residents of other towns.

If the instrumental-variables strategy outlined above does not correct for endogeneity, the lead coefficients would most likely betray this fact. In other words, if Wal-Mart times entry to take advantage of temporary retail growth spurts, then unless it times its entry perfectly, we would expect to see some increase in retail employment in the years before Wal-Mart entry. As the results below show, for the most part this is not the case; in other words, the IV strategy appears to correct for endogeneity as well as measurement error.

5 Sales per Worker

In this section I attempt to quantify the productivity differences between Wal-Mart and other retailers, subject to some important caveats. I use a common proxy for productivity in the retail sector, sales per worker, available for selected years at the county level from the Census of Retail Trade (CRT).

The CRT is conducted every five years, in years ending in 2 or 7 (1972, 1977, etc.), and provides county-level sales volume and total retail employment. To

[14] One way to check the validity of this assumption is to regress the difference in assigned entry dates based on the two sources on county characteristics. Such regressions consistently yield no relationship between county characteristics (such as size and urbanization) and the difference between the two entry dates.

compute sales per worker, I use data on total sales revenues for all establishments (1972-1992), total sales revenues for establishments with paid employees (available only 1977-1992), and total number of paid employees (1972-1992).[15] I use these data to compute two measures of sales per worker. The first is the ratio of sales in establishments with paid employees to the number of paid employees; this measure is available for 1977, 1982, 1987, and 1992. The second, noisier, measure is the ratio of total sales (in establishments with *and without* paid employees) to the number of paid employees. The second measure is available back to 1972, but has the disadvantage that sales in establishments with no employees (e.g., only a proprietor) are attributed to employees in other establishments. In practice, sales in establishments without employees are a very small fraction of total sales, so the figures are extremely similar.

I estimate the simple regression equation

$$\frac{\text{sales}_{jt}}{\text{retail}_{jt}} = \alpha + \sum_t \delta_t \text{year}_t + \sum_j \psi_j \text{county}_j + \theta \sum_{s \leq t} \text{WalMart}_{js} + \text{u}_{jt} \qquad (4)$$

where sales_{jt} is sales revenue (in real 1982-1984 dollars) of all retail establishments in county j in year t, retail_{jt} is retail employment in county j in year t, $\sum_{s \leq t} \text{WalMart}_{js}$ is the number of Wal-Mart stores in existence in county j in year t, and the other variables are as defined above. t runs from 1972 to 1992. The error term is clustered at the county level to allow for arbitrary autocorrelation at the county level.

Table 2 shows OLS and IV results from these regressions. The OLS results are shown separately for $\sum_{s \leq t} \text{WMopen}_{js}$ and $\sum_{s \leq t} \text{WMplan}_{js}$. The IV results use $\sum_{s \leq t} \text{WMplan}_{js}$ to instrument for $\sum_{s \leq t} \text{WMopen}_{js}$. "All sales" indicates that sales from all establishments are used in the computation of the LHS variable (allowing inclusion of the 1972 data); "Sales by employees" indicates that sales data used refer only to sales in establishments with employees. The IV estimates suggest that every Wal-Mart store increases sales per worker in the county by

[15] County-level data from the 1997 CRT have not yet been released.

Table 2: Estimated Effect of WalMart on Sales per Worker

	OLS		
	WMopen	WMplan	IV
All sales[a]	\$602.9	\$720.8	\$765.2
	(160.8)	(161.6)	(156.8)
Sales by employees[b]	442.3	579.0	578.0
	(159.1)	(159.9)	(143.9)

[a] Sales in all establishments per paid employee

[b] Sales in establishments with paid employees per paid employee

\$550-\$750 per year (in constant 1982-1984 dollars). Mean and median sales per worker for the period studied are approximately \$75,000, so this increase, while highly significant, is smaller than 1%.[16]

6 Results

6.1 Retail Employment

To begin, I present OLS results using the two alternative measures of Wal-Mart entry dates. Figure 4 shows the OLS estimates using the RHS variable WMopen_{jt}; Figure 5 shows OLS estimates for the same regressions, using WMplan_{jt} instead.[17] In both cases, retail employment increases by an estimated 40 jobs in the year of entry, up to half of which are eliminated within five years. In both cases as well, 20 jobs are estimated to have been created in the year *before* Wal-Mart entry. While this number is small in absolute magnitude, it is disconcertingly large relative to the estimated post-entry effect.

The IV results are shown in Figure 6. The effect of entry is estimated much

[16] These figures should be interpreted with caution. While they show that Wal-Mart entry coincides with an increase in county-level sales per employee, this coincidence is neither necessary nor sufficient to prove that Wal-Mart is more productive than its competitors. To see why it is not sufficient, note that Wal-Mart entry may increase sales per employee by substituting capital for labor; measures of capital are not available for the retail sector. At the same time, Wal-Mart prices tend to be lower than its competitors', so sales figures may decline even as productivity increases, implying that the coincidence is not even necessary.

[17] Like all regression results presented here, unless otherwise noted, the 95% confidence intervals shown use asymptotic standard errors and allow for any intertemporal correlation of errors for a given county.

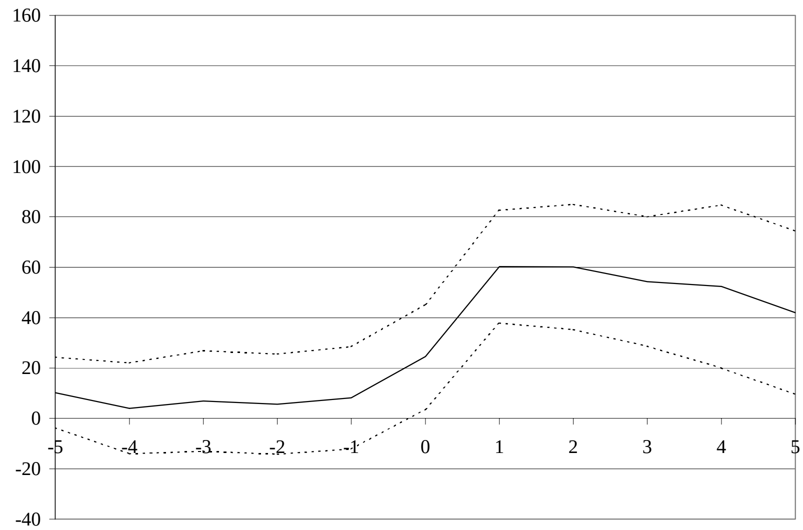

Figure 4: OLS Retail Employment Results (WMopen)

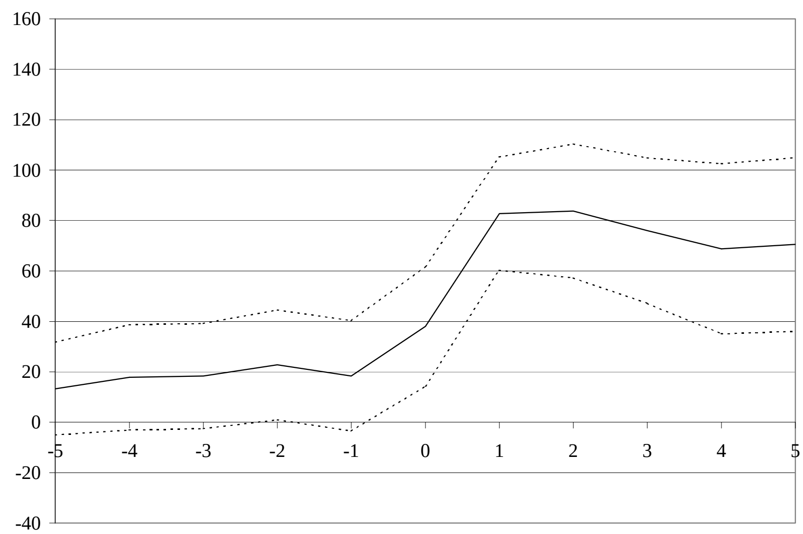

Figure 5: OLS Retail Employment Results (WMplan)

more cleanly at approximately 100 jobs. In the years immediately following entry, there is a loss of 40-70 additional jobs. The net effect at the five-year horizon, however, is positive and significant (p-value 0.0003).[18]

Recall that the typical Wal-Mart store employs 150-350 workers. These results suggest that employment increases by less than the full amount of Wal-Mart's hiring, even before allowing other firms time to fully adjust to Wal-Mart's entry. Part of this discrepancy can be explained by buyouts of existing chain stores by Wal-Mart Corporation, and prompt exit and cutbacks by other retailers.[19] Another (albeit unlikely) possibility is that Wal-Mart replaces existing part-time jobs with full-time jobs. CBP employment figures do not control for hours worked, so full-time and part-time employees are weighted equally.

It is difficult to compare hours of work for the typical Wal-Mart employee with hours of work at other retailers, because very little is known about employment conditions at Wal-Mart. A reasonable prior is that Wal-Mart employees work *fewer*, not more, hours than other retail workers, based on the finding by Bertrand and Kramarz (forthcoming) that as more entry is allowed into the French retail industry, part-time employment increases relative to all retail employment. Wal-Mart claims that 70% of its employees work 28 hours a week or more (Wal-Mart 2001a). This figure appears to be within the norm for workers in the discount retail industry.[20] It is also in keeping with the rest of the retail industry: the 30th percentile of hours worked by retail employees, obtained from the March Current Population Survey (CPS) for 1978-1999, is 28 hours across employer size, state, and year.

As a specification check, Figure 7 shows IV results from a regression that allows year fixed effects to vary not by urbanization but by Census region. This

[18] The long-run effect, six or more years after entry (not shown in Figure 6) is a net increase of 15 jobs over employment in year 0; this long-run increase is not statistically significant (p-value 0.3749).

[19] This possibility is explored in more detail in Section 6.2.

[20] See `http://www.pbs.org/storewars/stores3.html`.

second-stage regression is

$$\frac{\text{retail}_{jt}}{\text{pop}_{jt}} = \alpha + \sum_k \sum_t \lambda_{tk}\,\text{region}_{jk}\,\text{year}_t + \sum_j \psi_j \text{county}_j + \theta(L)\,\frac{\widehat{\text{WMopen}}_{jt}}{\text{pop}_{jt}} + \text{u}_{jt}. \quad (5)$$

(with appropriate modification to the first stage regression). The results are extremely similar to those results presented in Figure 6, where year fixed effects are allowed to differ by 1960 urbanization status. The instantaneous effect of entry is estimated at 100 jobs, with a decline of 20-50 jobs in the years immediately following entry.[21,22]

As noted in Section 4.3, if the timing of entry were endogenous, we would expect to see deviations from the county's long-run level of per-capita retail employment, relative to other counties, prior to entry. No such effect is evident in the leading coefficients. Given that construction alone can take several months (Murzell 1993), it is unlikely that Wal-Mart could introduce a store – with its requisite rezoning, planning, construction, and set-up – in less than a year.

All regression results reported for the remainder of the paper will be IV results with WMplan_{jt} instrumenting for WMopen_{jt}, and year fixed effects allowed to differ across urbanization categories.

6.2 Distribution of Retailer Size

Because Wal-Mart competes with retailers across categories – not only with general-merchandise stores, but with apparel stores, drug stores, etc. – it is interesting to look at changes in the overall distribution of retailer size in the years following Wal-Mart entry. We expect a decline in the number of competing retailers following Wal-Mart entry, with less-efficient (and disproportionately

[21] Ideally, we would like to control for state*year fixed effects or state*year*urban status fixed effects. Unfortunately, this is computationally infeasible given the simultaneous inclusion of county fixed effects in all models.

[22] The number of retail workers in Benton County, Arkansas, where Wal-Mart headquarters are located, has increased by 5-6 workers for each new Wal-Mart store over the last 20 years.

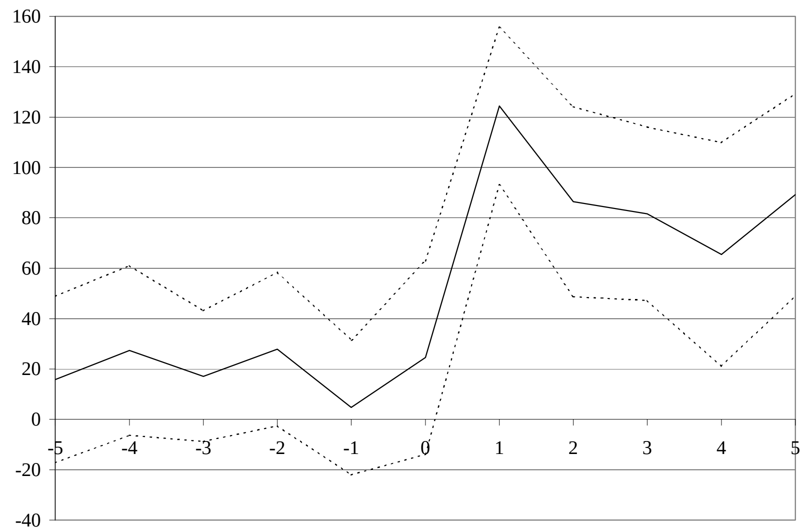

Figure 6: IV Retail Employment Results

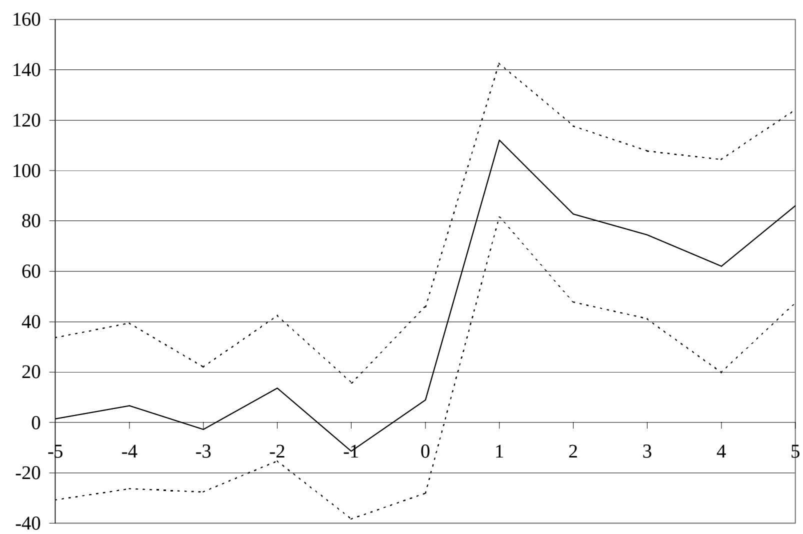

Figure 7: IV Retail Employment Results, Region*Year Fixed Effects

small) retailers most likely to exit. As noted by Stone (1987 and elsewhere), some retailers selling complementary products may be positively affected by Wal-Mart entry. In addition, retailers located near Wal-Mart may benefit from the externality of increased customer traffic, if Wal-Mart behaves like an anchor store in a traditional mall (see Pashigian and Gould 1998, Gould, Pashigian, and Prendergast 2002).

To capture the effect of Wal-Mart on the number of retail firms in each size category, I estimate instrumental-variables regressions with second stage

$$\frac{\mathrm{estab}_{jt}}{\mathrm{pop}_{jt}} = \alpha + \sum_k \sum_t \delta_{tk} \mathrm{urban}_k \mathrm{year}_t + \sum_j \psi_j \mathrm{county}_j + \theta(L) \frac{\widehat{\mathrm{WMopen}}_{jt}}{\mathrm{pop}_{jt}} + \mathrm{u}_{jt} \quad (6)$$

where estab_{jt} are, respectively, the number of small retail establishments (under 20 employees) in county j at year t; the number of medium-sized establishments (20-99 employees); and the number of large establishments (100+ employees). IV results are presented below.

Figure 8 shows the effect of Wal-Mart on the number of small establishments, defined as having fewer than 20 employees. There is a significant decline of 3 small retail establishments in the years after Wal-Mart entry (p-value 0.0009).

Figure 9 shows a decline also in the number of medium-sized establishments (with 20-99 employees) following Wal-Mart entry. Note that the scaling is not as in Figure 8, because there is less fluctuation in the number of medium-sized establishments. There is a slight, marginally significant, decline in the number of medium-sized establishments (p-value 0.0337).

Finally, Figure 10 shows the effect of Wal-Mart on the number of large establishments (with 100 or more employees). Note that the estimated coefficients mirror those on retail employment shown in Figure 6. The increase in the number of large retail establishments, of approximately 0.7, reinforces the interpretation that Wal-Mart's entry coincides with exit or contraction of other

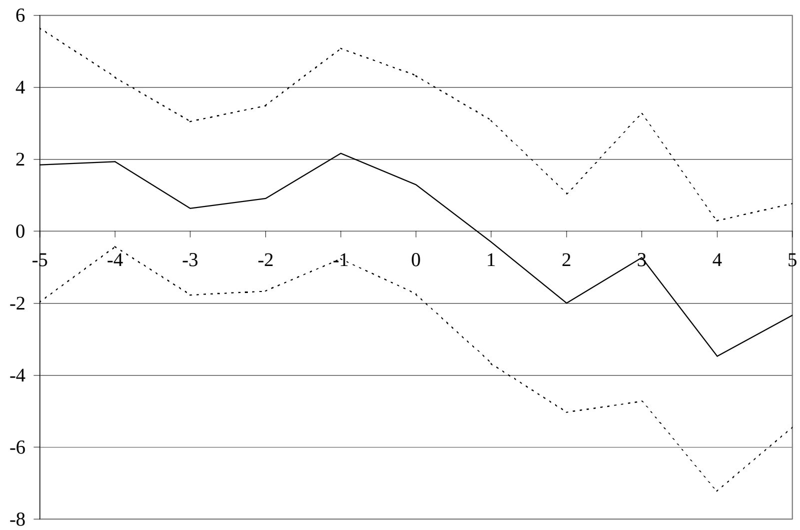

Figure 8: Small Retail Establishments (1-19 Employees)

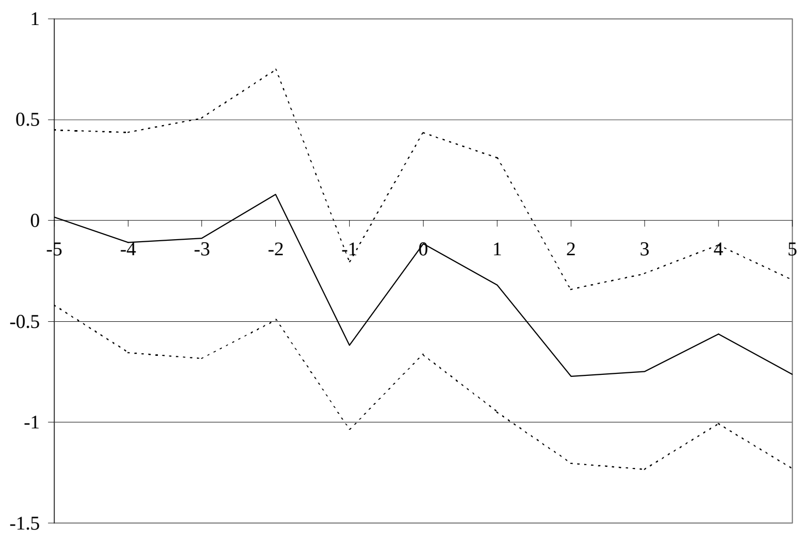

Figure 9: Medium-Sized Retail Establishments (20-99 Employees)

large retailers.[23] Additional firms exit, or shrink in size below 100 employees, in the years following Wal-Mart entry.

6.3 Retail Employment in Neighboring Counties

Shopping in neighboring counties is an imperfect substitute for shopping in one's county of residence, since additional travel is involved. Nevertheless, due to their low prices and large selection, Wal-Mart stores in rural areas typically draw customers from a wide radius which may include several neighboring counties. In a series of studies, Kenneth Stone argues that much of the negative effect of Wal-Mart on retail employment occurs not in the communities in which Wal-Mart located, but in nearby communities (see, e.g., Stone 1997).

At the same time, competition in the labor market may drive workers from retail establishments in their own counties to neighboring counties. (County Business Patterns data attribute jobs to the county in which the employer is located, rather than the workers' counties of residence.) Both effects are expected to work in the same direction, so the expected effect on neighboring counties' retail employment is unambiguously negative.

I define counties as "neighbors" if the distance between their geographic centers is under 10 miles.[24] Formally, let $J = \{\text{neighbor}\,(j)\}$ be the set of county j's neighbors, and define for any variable X,

$$\mathrm{X}_{Jt} \equiv \sum_{k \in J} \mathrm{X}_{kt}. \tag{7}$$

To estimate the effect of Wal-Mart entry in county j on retail employment in

[23] In some cases, Wal-Mart acquired a large number of stores from a competitor; in those cases Wal-Mart entry was not associated with a net increase in the number of large retail establishments. Examples include the 1977 purchase by Wal-Mart of 16 Mohr Value Discount Department Stores in Missouri and Illinois, and the 1981 purchase of 106 stores in nine states from Kuhn's-Big K Stores Corp.

[24] Other definitions of "neighbors" were also tried as robustness checks; the results were not sensitive to the exact definition.

the surrounding area J, I estimate an IV regression with second stage

$$\frac{\text{retail}_{Jt}}{\text{pop}_{Jt}} = \alpha + \sum_t \delta_t \text{year}_t + \sum_j \psi_j \text{county}_j + \phi \frac{\sum_{s \le t} \widehat{\text{WMopen}}_{Js}}{\text{pop}_{Jt}} + \theta(L) \frac{\widehat{\text{WMopen}}_{jt}}{\text{pop}_{Jt}} + \text{u}_{jt} \quad (8)$$

(to economize on estimated parameters, year fixed effects are assumed constant across urbanization categories). $\frac{\sum_{s \le t} \widehat{\text{WMopen}}_{Js}}{\text{pop}_{Jt}}$ and $\frac{\widehat{\text{WMopen}}_{jt}}{\text{pop}_{Jt}}$ are predicted values from the appropriate first-stage regressions, with instruments $\frac{\sum_{s \le t} \widehat{\text{WMplan}}_{Js}}{\text{pop}_{Jt}}$ and appropriate leads and lags of $\frac{\widehat{\text{WMplan}}_{jt}}{\text{pop}_{Jt}}$. The variable $\frac{\sum_{s \le t} \widehat{\text{WMplan}}_{Js}}{\text{pop}_{Jt}}$ is the number of existing stores, per-capita, in counties J at year t (in other words, the cumulative number of new stores opened no later than year t). It is included in the regression to avoid confounding the effect of Wal-Mart entry in county j with the effect of entry in county j's neighbors, and is assumed to have a once-and-for-all effect on employment in those counties. Note that as the estimated effect is on retail employment per capita in neighboring counties, the appropriate normalization of the number of Wal-Mart stores also uses neighboring counties' population.

IV results are shown in Figure 11. No significant effect of Wal-Mart entry on retail employment in neighboring counties can be detected, although the mean of the pre-entry coefficients is roughly 50 jobs above the post-entry mean. To interpret the coefficients, note that the regression includes retail employment in *all* neighboring counties on the LHS. The average county in the sample has about 5 neighbors, so the annual fluctuations shown are on the order of 10-20 jobs per neighboring county.[25]

[25] The estimated coefficient ϕ is approximately 47 with 95% confidence-interval $[7, 86]$. This number is in line with the estimated five-year effects in the own-county IV regressions. When own-county effects are constained to be once-and-for-all, with second-stage regression

$$\frac{\text{emp}_{jt}}{\text{pop}_{jt}} = \alpha + \sum_t \delta_t \text{year}_t + \sum_j \psi_j \text{county}_j + \phi \frac{\widehat{\text{WMopen}}_{jt}}{\text{pop}_{jt}} + \text{u}_{jt}$$

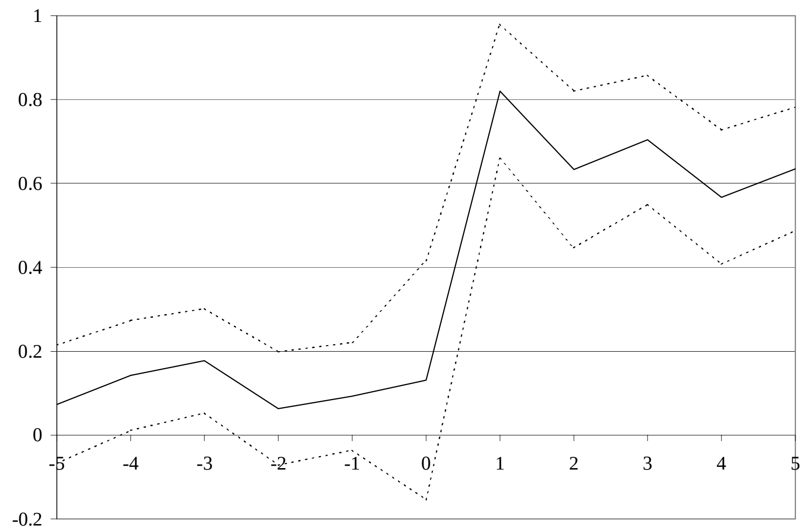

Figure 10: Large Retail Establishments (100+ Employees)

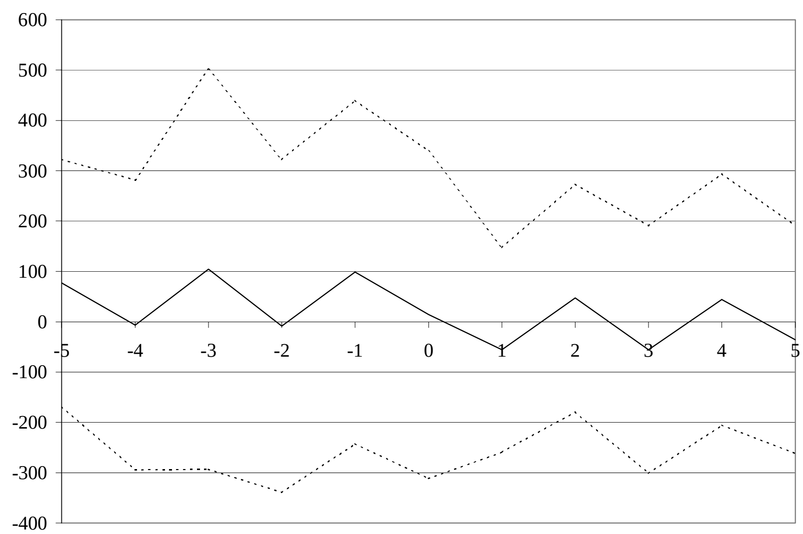

Figure 11: Retail Employment in Neighboring Counties

As an alternative specification with more power, I estimate a simple once-and-for-all IV regression with second-stage regression

$$\begin{aligned}\frac{\text{retail}_{Jt}}{\text{pop}_{Jt}} =& \alpha + \sum_t \delta_t \text{year}_t + \sum_j \psi_j \text{county}_j \\ &+ \phi \frac{\sum_{s\leq t} \widehat{\text{WMopen}}_{Js}}{\text{pop}_{Jt}} + \theta \frac{\sum_{s\leq t} \widehat{\text{WMopen}}_{js}}{\text{pop}_{Jt}} + \text{u}_{jt} \end{aligned} \tag{9}$$

where $\frac{\sum_{s\leq t} \widehat{\text{WMopen}}_{js}}{\text{pop}_{Jt}}$ is the (predicted) number of existing Wal-Mart stores county j per J capita. The estimated coefficient θ is -29.119 and is significant at the 1% level. In other words, while Wal-Mart entry into county j permanently increases retail employment in county j by approximately 50 jobs, retail employment in neighboring counties decreases by approximately 30 jobs.

6.4 Other Sectors

6.4.1 Wholesale

Because Wal-Mart is vertically integrated, Wal-Mart entry is unlikely to complement wholesale employment even though it is associated with an increase in measured retail employment. Moreover, anecdotal evidence suggests that some competing retailers find that Wal-Mart's prices are better than their wholesalers. *Inc.* magazine, in an article about the effect of Wal-Mart on small businesses, interviewed a retailer who noted that his local Sam's Club, a membership warehouse club owned by Wal-Mart, carried some items at a lower price than his distributor (Welles 1993). Thus, though a retail store, Wal-Mart may be a substitute for wholesalers. In addition to this direct competition from Wal-Mart, wholesalers may also be affected indirectly as small retailers, who traditionally buy from regional wholesalers, shut down (Shills 1997).

the estimated coefficient $\widehat{\phi}$ is approximately 62, with confidence interval $[11, 113]$.

The estimated effect of Wal-Mart entry on county-level wholesale employment is shown in Figure 12. The observed decline of 25 wholesale jobs following Wal-Mart entry is statistically significant (p-value 0.0334).[26]

6.4.2 Restaurants

Restaurant employment is used as a control for retail employment, because the two are likely to be highly correlated, but restaurant employment is not expected to be substantially affected by Wal-Mart entry. There are several caveats to this claim: retail and restaurant employees may be drawn from the same labor market (though restaurant employees are on average younger and less-educated than retail workers); and there is some anecdotal evidence to suggest that restaurants, at least fast-food restaurants (which cannot be separated from other eating establishments in County Business Patterns data), may complement shopping at Wal-Mart.

There is no perceptible impact of Wal-Mart on restaurant employment, as Figure 13 shows: neither a discontinuity as with retail employment (see Figure 6 above), nor a change in the pattern of growth as with wholesale employment (Figure 12). The observed trend in restaurant employment is most likely due to other factors not captured by the regression, and suggests that the instrumental-variables specification addresses some, but not all, concerns about endogeneity in the timing of entry.[27]

6.4.3 Total Employment

As noted above, the typical Wal-Mart store has 150-350 employees, less than 2% of total employment in the average county at the time of the Wal-Mart entry. The chances of finding a statistically-significant effect on total county employment are therefore slim, and in fact, Figure 14 shows the estimated effect is

[26] In long-run – six or more years after entry – a further 10 jobs are lost in the wholesale sector. The long-run effect is therefore a loss of 35 wholesale jobs (p-value 0.0029).

[27] The long-run coefficient (not depicted) is not significantly different from the coefficient for year 0.

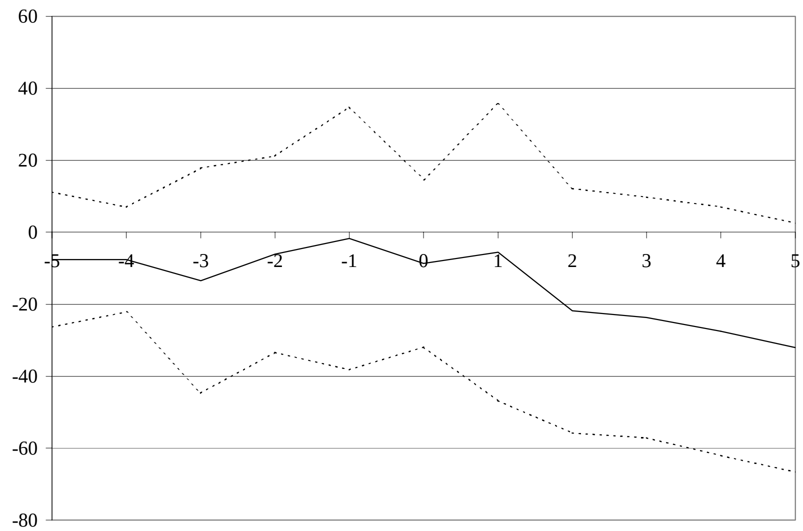

Figure 12: Wholesale Employment

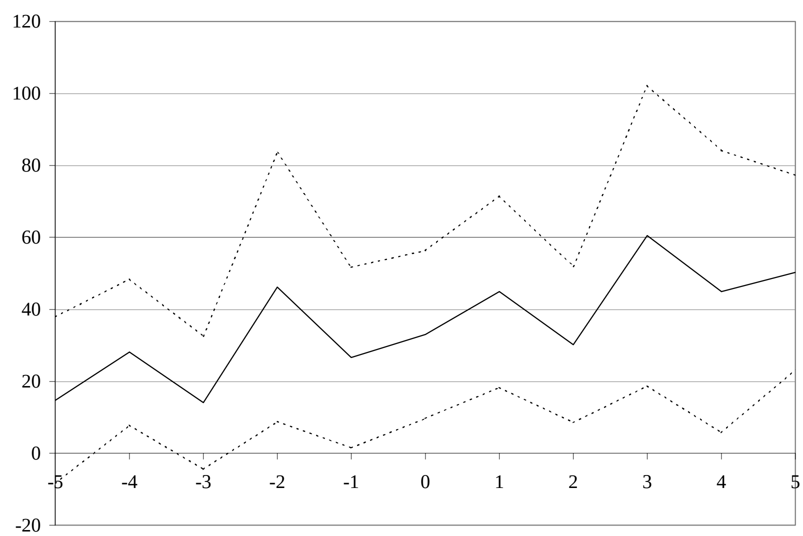

Figure 13: Restaurant Employment

statistically zero. Neither the five-year effect nor the long-run effect is statistically different from the coefficient in the entry year. The increasing trend in the years before Wal-Mart entry, however, again suggests that some endogeneity in the timing of entry remains.

7 Conclusion

The effect of productivity increases on employment has been analyzed at both the macroeconomic level and the plant level, but the intermediate level has been neglected by researchers. This level of analysis is important, because much of the productivity growth we have observed in the past decade was associated with entry, rather than technology adoption by existing firms and establishments. This paper takes a first stab at estimating the effect of entry of a more productive firm on sector-level employment and the reallocation of jobs across firms.

The experiment is a clean one, because I am able to identify the date of entry precisely, using an instrumental-variables specification. The effect I estimate is a flexible reduced-form effect, allowing both Wal-Mart and other firms in the county of entry as well as in surrounding counties to adjust to the shock over a period of several years. Finally, because I use a large panel of nearly 1800 counties over 23 years, and because Wal-Mart entry is a "large" shock relative to the size of the local retail market in most counties – median retail employment in 1990 was only 850, while the average Wal-Mart store had approximately 200 employees – the effect can be estimated with relative precision.

I find an increase of 100 retail jobs in the county at entry; half of that increase remains five years after entry. This effect is substantially mitigated when neighboring counties are also considered, where there is a decline of approximately 30 retail jobs. There is also a negative effect on county-level wholesale employment. Combined, these negative effects are large enough to fully offset the gains to retail employment in the entered county.

In closing, it should be emphasized that this paper does not attempt to answer the question whether entry of Wal-Mart has a positive or negative net impact on a local economy. The answer to that question depends on many other factors, which are beyond the scope of this paper; these include concerns about market concentration, income effects, distribution of rents, and more.

References

[1] Baily, Martin Neil and Robert M. Solow (2001). "International Productivity Comparisons Built from the Firm Level." *Journal of Economic Perspectives* 15.

[2] Banerjee, Anindya (1999). "Panel Data Unit Roots and Cointegration: An Overview." *Oxford Bulletin of Economics and Statistics* 61.

[3] Barnes, Norma Ganim, Allison Connell, Lisa Hermenegildo, and Lucinda Mattson. "Regional Differences in the Economic Impact of Wal-Mart." *Business Horizons* 39.

[4] Bertrand, Marianne, and Francis Kramarz (2002, orthcoming). "Does Entry Regulation Hinder Job Creation? Evidence from the French Retail Industry." *Quarterly Journal of Economics* 117.

[5] Bertrand, Marianne, Esther Duflo and Sendhil Mullainathan (2001). "How Much Should We Trust Difference-in-Difference Estimates?" MIT mimeograph.

[6] Bresnahan, Timothy F., Erik Brynjolfsson and Loin M. Hitt (2002). "Information Technology, Workplace Organization, adn the Demand for Skilled Labor: Firm-Level Evidence." *Quarterly Journal of Economics* 117.

[7] Chain Store Guide (1979-1982, 1987-1993). *Directory of Discount Department Stores*. New York: Business Guides, Inc.

[8] Chain Store Guide (1983-1986). *Directory of Discount Stores*. New York: Business Guides, Inc.

[9] Darby, W. D. (1928). *The Story of the Chain Store: A Study of Chain Store Policies and Methods, Particularly as They Affect the Independent Merchant in the Dry Goods Field, Together With a General Survey of Chain Store Developments*. New York: Dry Goods Economist.

[10] Dobbs, I. M., M. B. Hill, and M. Waterson (1987). "Industrial Structure and the Employment Consequences of Technical Change." *Oxford Economic Papers* 39.

[11] Edgecliffe-Johnson, Andrew (June 19, 1999). "A Friendly Store from Arkansas: Wal-Mart's Devotion to Customer Service and Cost-Cutting Technology is Starting to Reshape Retailing Around the World." *Financial Times*.

[12] Entorf, Horst, Michel Gollac, and Francis Kramarz (1999). "New Technologies, Wages, and Worker Selection." *Journal of Labor Economics* 17.

[13] Feiner, Jeffrey M., et al. (1993). *The Wal-Mart Encyclopedia: A Portfolio Manager's Guide to Growth*. New York: Salomon Brothers.

[14] Feiner, Jeffrey M., et al. (2001). *Wal-Mart Encyclopedia X: Building a Global Brand*. New York: Lehman Brothers.

[15] Foley, Sharon, and Takia Mahmood (1996). "Wal*Mart Stores, Inc." Harvard Business School Case Study 9-794-024.

[16] Foster, Lucia, John Haltiwanger, and C. J. Krizan (2001). "Aggregate Productivity Growth: Lessons from Microeconomic Evidence," in *New Developments in Productivity Analysis*. Edward Dean, Michael Harper and Charles Hulten, eds. Chicago: University of Chicago Press.

[17] Foster, Lucia, John Haltiwanger, and C. J. Krizan (2002). "The Link between Aggregate and Micro Productivity Growth: Evidence from Retail Trade." NBER Working Paper 9120.

[18] Franz, Lori, and Edward Robb (1990). "Effect of Wal-Mart Stores on Economic Environment of Rural Communities." Business and Public Administration Research Center mimeograph.

[19] Galí, Jordi (1999). "Technology, Employment, and the Business Cycle: Do Technology Shocks Explain Aggregate Fluctuations?" American Economic Review 89.

[20] Ghemawat, Pankaj (1989). "Wal-Mart Stores' Discount Operations." Harvard Business School Case Study 9-387-018.

[21] Ghemawat, Pankaj and Gregg Friedman (1999). "Wal-Mart in 1999." Harvard Business School Case Study N9-799-118.

[22] Griliches, Zvi and Haim Regev (1995). "Firm productivity in Israeli industry 1979-1988." *Journal of Econometrics* 65, 175-203.

[23] Pashigian, B. Peter and Eric D. Gould (1998). "Internalizing Externalities: The Pricing of Space in Shopping Malls." *Journal of Law and Economics* 41.

[24] Gould, Eric D., B. Peter Pashigian and Canice Prendergast (2002). "Contracts, Externalities, and Incentives in Shopping Malls." Mimeograph.

[25] Im, Kyung So, M. Hashem Pesaran, and Yongcheol Shin (1997). "Testing for Unit Roots in Heterogeneous Panels." Mimeograph, University of Cambridge.

[26] Johnson, Bradford C. (2002). "Retail: The Wal-Mart Effect." *The McKinsey Quarterly*.

[27] Kane, Thomas J., Cecilia Elena Rouse, and Douglas Staiger (1999). "Estimating Returns to Schooling When Schooling is Misreported." Princeton Industrial Relations Section Working Paper 419.

[28] Kézdi, Gábor (2001). "Robust Standard Error Estimation in Fixed-Effects Panel Models." University of Michigan mimeograph.

[29] KMart Corporation (1997-1999). *KMart Fact Book*.

[30] Levin, Andrew, and Chien-Fu Lin (1993). "Unit Root Tests in Panel Data: New Results." University of California San Diego Discussion Paper 93-56.

[31] Maddala, G.S., and Shaowen Wu (1999). "A Comparative Study of Unit Root Tests with Panel Data and a New Simple Test." *Oxford Bulletin of Economics and Statistics* 61.

[32] McKinsey Global Institute (2001). "US Productivity Growth 1995-2000: Understanding the contribution of Information Technology relative to other factors." Washington DC: MGI.

[33] Muller, Thomas (1999). "Impact of Anticipated Wal-Mart on the Northern Neck (including Kilmarnock and nearby communities)." Fairfax, Virginia: mimeograph.

[34] Muzeroll, Phillis A. (1993). "Upper Valley Economic Report: Ascutney, Wal-Mart Follow Growth." *Vermont Business Magazine*.

[35] Shils, Edward B. (1997). "The Shils Report: Measuring the Economic and Sociological Impact of the Mega-Retail Discount Chains on Small Enterprise in Urban, Suburban and Rural Communities." `http://www.lawmall.com/rpa/rpashils.htm`

[36] Shnuch, Mary T. (June 3, 1990). "Business Isn't the Same When Wal-Mart Comes to a Small Town." *The Washington Post.*

[37] Stalk, George, Jr., and Thomas M. Hout (1990). *Competing Against Time: How Time-based Competition Is Reshaping Global Markets*. New York: The Free Press (Macmillan).

[38] Stone, Kenneth E. (1991). "Competing with the Mass Merchandisers." *Small Business Forum*, Spring 1991.

[39] Stone, Kenneth E. (1997). "Impact of the Wal-Mart Phenomenon on Rural Communities." *Increasing Understanding of Public Problems and Policies*. Chicago: Farm Foundation.

[40] Stone, Kenneth E. (1989). "The Impact of Wal-Mart Stores on Other Businesses in Iowa." Iowa State University mimeo.

[41] Urban Land Institute (various years). *Dollars and Cents of Shopping Centers: A Study in Receipts and Expenses in Shopping Center Operations*. Washington, DC: Urban Land Institute.

[42] US Bureau of the Census (1964-1999). *County Business Patterns*. Washington, DC: US Government Printing Office.

[43] Vance, Sandra S., and Roy V. Scott (1994). *Wal-Mart: A History of Sam Walton's Retail Phenomenon*. New York: Twayne Publishers.

[44] Wal-Mart Stores, Inc. (1971-2001). *Annual Report*.

[45] Wal-Mart Stores, Inc. (2001a). *Wal-Mart Associate Handbook*.

[46] Welles, Edward O. (1993). "When Wal-Mart Comes to Town." *Inc. Magazine*, July 1, 1993.

A Theoretical Framework

This appendix presents a simple model to highlight the ambiguity of the effect of a positive technology shock on sectoral employment. I focus on the retail sector, abstracting away from other sectors in the economy. This approach is legitimate assuming the retail sector is sufficiently small and insulated, so that wage and price changes in the retail sector do not have widespread consequences for prices and production levels in other sectors (and by implication, on aggregate price and production levels).

Consider a perfectly competitive retail sector with an upward-sloping labor supply function

$$L = \eta w$$

where $\eta > 0$. The upward-sloping labor-supply function should be viewed as a reduced-form representation; it could be due, most plausibly, to workers' heterogeneous outside options in other sectors.

Two intermediate retail services are produced competitively with a single input (labor) and CRS production functions

$$\begin{aligned} R_1 &= AL_1 \\ R_2 &= L_2 \end{aligned}$$

where $A \geq 0$ parametrizes firm 1's technology.[28]

A final retail service, with price p, is produced competitively from the intermediate services with CES production function

$$R = \left[\gamma \left(R_1\right)^{\rho} + (1-\gamma)\left(R_2\right)^{\rho}\right]^{\frac{1}{\rho}} \tag{10}$$

where $\gamma \in (0,1)$ is the relative importance of R_1 in the final good produc-

[28] A more complicated model with multiple inputs would be able to capture an added dimension of uncertainty due to the elasticity of substitution between labor and capital, which may vary between the old and new technologies.

tion, $\frac{1}{1-\rho}$ is the elasticity of substitution in demand for retail services, and $\rho \in (-\infty, 1)$. The two intermediate services are gross substitutes if $\rho > 0$, and they are gross complements if $\rho < 0$. Demand for R is given by the invertible function $R(p)$.

Total retail employment is just the sum of employment in the two intermediate sectors:

$$L_1 + L_2 = L.$$

First-order conditions in the final good sector are

$$\begin{aligned} p_1 &= \gamma p R^{1-\rho} (R_1)^{\rho-1} = \gamma p R^{1-\rho} (A L_1)^{\rho-1} \\ p_2 &= (1-\gamma) p R^{1-\rho} (R_2)^{\rho-1} = (1-\gamma) p R^{1-\rho} (L_2)^{\rho-1} \end{aligned} \tag{11}$$

and in the two intermediate sectors

$$\begin{aligned} p_1 &= \frac{w}{A} \\ p_2 &= w. \end{aligned} \tag{12}$$

Equating $Ap_1 = p_2$ in the final good FOCs yields

$$L_2 = \left(\frac{\gamma A^\rho}{1-\gamma} \right)^{\frac{1}{\rho-1}} L_1.$$

From the labor supply function,

$$L_1 + L_2 = \eta w = \eta (1-\gamma) p \left[\gamma A^\rho (L_1)^\rho + (1-\gamma)(L_2)^\rho \right]^{\frac{1-\rho}{\rho}} (L_2)^{\rho-1}$$

where the second equality comes from recognizing that $w = p_2$ (from Equation (12)) and evaluating p_2, from Equation (11), using Equation (10).

Solving simultaneously yields employment equations

$$L_1 = \eta p \left(\gamma A^\rho\right)^{\frac{1}{1-\rho}} \left[\left(\gamma A^\rho\right)^{\frac{1}{1-\rho}} + (1-\gamma)^{\frac{1}{1-\rho}}\right]^{\frac{1-2\rho}{\rho}} \tag{13}$$

$$L_2 = \eta p \left(1-\gamma\right)^{\frac{1}{1-\rho}} \left[\left(\gamma A^\rho\right)^{\frac{1}{1-\rho}} + (1-\gamma)^{\frac{1}{1-\rho}}\right]^{\frac{1-2\rho}{\rho}} \tag{14}$$

$$L = \eta p \left[\left(\gamma A^\rho\right)^{\frac{1}{1-\rho}} + (1-\gamma)^{\frac{1}{1-\rho}}\right]^{\frac{1-\rho}{\rho}} \tag{15}$$

and price equations

$$w = p_2 = p \left[\left(\gamma A^\rho\right)^{\frac{1}{1-\rho}} + (1-\gamma)^{\frac{1}{1-\rho}}\right]^{\frac{1-\rho}{\rho}} \tag{16}$$

$$p_1 = p \left[\gamma^{\frac{1}{1-\rho}} + \left(\frac{1-\gamma}{A^\rho}\right)^{\frac{1}{1-\rho}}\right]^{\frac{1-\rho}{\rho}} \tag{17}$$

where p is defined implicitly by the function

$$p = R^{-1}\left(\eta p \left[\left(\gamma A^\rho\right)^{\frac{1}{1-\rho}} + (1-\gamma)^{\frac{1}{1-\rho}}\right]^{\frac{2-2\rho}{\rho}}\right). \tag{18}$$

Comparative statics cannot be analyzed without putting some structure on the demand function $R(p)$. Assume it is given by

$$R(p) = e^{1-\alpha \ln(p)}$$

with constant elasticity of demand $|\alpha| > 0$. Solving Equation (18), we get

$$p = \exp\left[\frac{1 - \ln\left(\eta \left[\left(\gamma A^\rho\right)^{\frac{1}{1-\rho}} + (1-\gamma)^{\frac{1}{1-\rho}}\right]^{\frac{2-2\rho}{\rho}}\right)}{\alpha + 1}\right].$$

The main result of this model highlights the ambiguity of the sign of the effect of increased productivity in one firm on overall sectoral employment.

Result 1 *If demand for retail services is inelastic* ($\alpha < 1$)*, total retail employment decreases with* A*. If demand for retail services is elastic* ($\alpha > 1$)*, total retail employment increases with* A*.*

Proof. Follows from differentiating Equation (15):

$$\frac{\partial L}{\partial A} = \frac{\eta}{A} \exp\left[\frac{1 - \ln\left(\eta\left[(\gamma A^{\rho})^{\frac{1}{1-\rho}} + (1-\gamma)^{\frac{1}{1-\rho}}\right]^{\frac{2-2\rho}{\rho}}\right)}{\alpha + 1}\right] \bullet$$
$$\left[(\gamma A^{\rho})^{\frac{1}{1-\rho}} + (1-\gamma)^{\frac{1}{1-\rho}}\right]^{\frac{1-2\rho}{\rho}} (\gamma A^{\rho})^{\frac{1}{1-\rho}} \left(\frac{\alpha - 1}{\alpha + 1}\right)$$

and noting that all but the last term are strictly positive. ■

To see the intuition for this result, note that the productivity increase in the retail sector has two direct effects. Given a fixed quantity demanded, fewer workers are needed to supply it. At the same time, however, lower prices will increase quantity demanded. Which effect dominates depends on the price elasticity of demand. Note that this result does not depend on the particular functional forms used here (for final-good production and demand), but is much more general.

B Data and Empirical Issues

B.1 Wal-Mart Data

Table 3 shows the sources from which store opening dates, used in the construction of the variable WMopen_{jt}, were drawn. Chain Store Guides' *Directories of Discount Department Stores* from 1990-1993 are available, but are largely uninformative; the directories do not appear to have been updated in those years. For stores that do not appear in the 1989 directory, but do appear in the 1995 Rand McNally road atlas (i.e., exist in 1994), opening dates are assigned according to the following algorithm. From the annual reports, I obtain the net increase (rarely, a decrease) in the number of Wal-Mart stores in each state each year. Since there are very few store closures, I use the net increase to proxy for the gross increase, i.e., the number of new stores to open each year in each state. For example, in Arizona, 5 new stores opened in 1990, 7 in 1991, and one each in 1992 and 1993. Using the list of stores that existed in 1994 but not in 1989, I assign entry dates randomly, in proportion to their probability of opening in each year. Therefore each store that opened in Arizona during this period has a probability $\frac{5}{14}$ of being assigned to 1990, probability $\frac{1}{2}$ of being assigned to 1991, and probability of $\frac{1}{14}$ each of being assigned to 1992 and 1993. In all, 680 stores' opening dates are assigned in this way, as follows: 203 in 1990, 145 in 1991, 181 in 1992, and 151 in 1993.[29]

Table 3: Directory Sources for Wal-Mart Opening Dates

Years	Source
1962-1969	Vance and Scott (1994)
1970-1978	Wal-Mart Annual Reports
1979-1982	Directory of Discount Department Stores
1983-1986	Directory of Discount Stores
1987-1989	Directory of Discount Department Stores
1990-1993	See text
1994-1997	Rand McNally Road Atlas

Table 4 shows the assignment of store opening dates by store numbers. Store openings in county j in year t implied by these assigned entry dates are aggregated to the county-year level and form the variable WMplan_{jt}. The accuracy of this method depends critically on Wal-Mart assigning store numbers in a roughly sequential order, and not reassigning numbers in the event of store closure. Only 40 stores closed over the entire period 1964-1999, so the latter condition appears to be satisfied; this also implies that reassignment of store numbers, if it takes place at all, cannot be common.[30]

Figure 15 shows the distribution of the difference between the two measures of opening dates, at the store level. Over 40% of stores are assigned the same opening year with both measures, approximately 80% of stores are assigned two opening dates within one year of one another, and 90% are within two years.

[29] Entry dates assigned in this way clearly suffer from measurement error, but they are unbiased. This method is therefore preferred to the naïve alternative of assigning all stores that opened in those years the date they first appear in the data (generally, 1993); entry dates assigned this way would be biased as well as measured with error.

[30] Relocation of stores within a community, which is much more common than store closure, does not pose a problem.

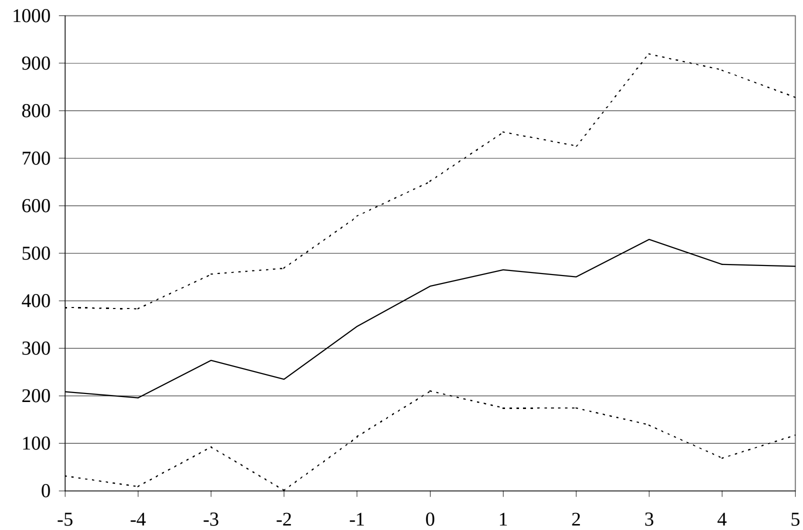

Figure 14: Total Employment

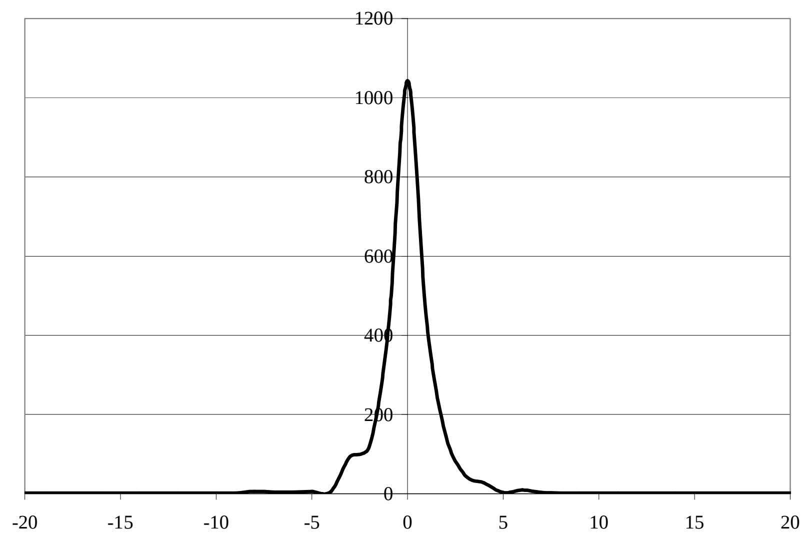

Figure 15: Differences between the Two Measures of Opening Dates (Years)

Table 4: Store Opening Year Assignment by Store Number

Store Numbers	Assigned Year	Store Numbers	Assigned Year
1-2	1964	330-489	1981
3	1965	490-549	1982
4-5	1966	550-640	1983
6-7	1967	641-743	1984
8-12	1968	744-857	1985
13-17	1969	858-977	1986
18-37	1970	978-1111	1987
38-51	1971	1112-1256	1988
52-62	1972	1257-1399	1989
63-77	1973	1400-1569	1990
78-103	1974	1570-1701	1991
104-124	1975	1702-1874	1992
125-152	1976	1875-2013	1993
153-193	1977	2014-2123	1994
194-225	1978	2124-2231	1995
226-274	1979	2232-2999	1996
275-329	1980	2300+	1997

B.2 County Merges and Splits

The following is a complete list of county merges and splits in the contiguous 48 states since 1960.[31]

Arizona: Yuma County split into two counties, La Paz and Yuma, in 1987.

Florida: Dade County changed its name to Miami-Dade County in 1997.

Georgia: Muscogee County became Columbus-Muscogee county, in 1971; reflected in Census data beginning 1974.

Nevada: Carson City and Ormsby County were consolidated into one community in 1969.

New Mexico: Valencia County split in two in 1983 to form Cibola County and Valencia County.

Virginia has seen the largest number of changes, many of them involving the creation of "independent cities" that are not part of any county. Bedford city

[31] For more information about county definitions and changes see `http://www.census.gov/population/cencounts/00-90doc.txt`.

split from Bedford County in 1968; Emporia city split from Greensville County in 1967; Lexington city split from Rockbridge County in 1966; Manassas city and Manassas Park city split from Prince William County in 1975; Nansemond County became Nansemond city in 1972 and then was annexed to Suffolk city in 1974; Poquoson city split from York County in 1975; Salem city split from Roanoke County in 1968; South Boston city merged with Halifax County in 1995.

Wisconsin: Menominee County was formed from parts of Shawano and Oconto Counties in 1961.

Wyoming: Since 1970, employment in Yellowstone National Park has been reported with Teton and Park Counties; I have combined these three counties and treat them as one throughout.

B.3 Employment Data Accuracy

Occasionally, in counties with a small number of employers, data on the total number of employees in a sector, or even in the entire county, is omitted from County Business Patterns to avoid disclosure of the number of employees in individual firms. Only the number of firms in each of eight employment-size categories (1-19, 20-99, ... 50,000-99,999, 100,000+ employees) is given in these cases. In those instances, I assume that the actual number of employees of a firm of size X is a weighted mean of the lower and upper bounds on its employment-size class (with weight $\frac{2}{3}$ on the lower bound and $\frac{1}{3}$ on the upper bound); the exception is the class-size of 100,000+, to which I assign 150% of the lower bound, or 150,000. For example, a firm with 1-19 employees is assigned a value of 7.[32]

[32] I chose to weight the lower and upper bounds of each interval by $\left(\frac{2}{3}, \frac{1}{3}\right)$, respectively, rather than $\left(\frac{1}{2}, \frac{1}{2}\right)$, because counties small enough to elicit concerns about disclosure of information on individual firms in aggregate data seem likely to have a disproportionate number of small employers. The results are robust to this specification.

B.4 Unit Roots

There is a high degree of persistence in county-level employment. To test whether the employment series used admit unit roots, I run a Dickey-Fuller (DF) test on each county series separately, after removing year fixed effects interacted with 1960 urbanization status (urban, suburban, rural). By construction, a 5% rejection rate is to be expected at the 95% confidence level if the series have unit roots. The actual rejection rates vary by series from 6%-14%, and are shown in Table 5.

Panel-data unit root tests provide a powerful alternative to county-by-county testing. I apply two such tests, one by Maddala and Wu (1999) and another by Levin and Lin (1993). Maddala and Wu use a variant of a test by Fisher which uses a combination of the p-values from the county-by-county DF tests. The Maddala-Wu test statistic is

$$-2\sum_{i=1}^{N} \ln(\pi_i) \sim \chi^2_{(2N)}$$

where π_i is the p-value from the Dickey-Fuller test for county i, and the null hypothesis is that all series share a unit root. The Levin-Lin test is a panel variant of the Dickey-Fuller test with the same null hypothesis.

The panel tests are subject to two caveats. First, the null hypothesis will be violated if even one of the series is stationary. Rejection of the null hypothesis may therefore be interpreted to imply that some series have unit roots and others do not, or, if we believe the series constitute realizations of a single process, rejection implies that the common process is stationary. The tests are therefore meaningless under the assumption that each county employment series may be the realization of a unique process. Second, both tests assume that the observations are independent; this assumption may be violated by spatial correlation.

Table 5 reports the test results. The first column shows the fraction of counties for which county-by-county Dickey-Fuller tests rejected the presence of unit roots at 95% significance. The rejection rates of 6%-14% for these series are higher than the expected 5% under the null hypothesis of unit roots. The second and third columns report p-values from Maddala-Wu and Levin-Lin tests, respectively.

Table 5: Unit Root Tests

Employment	Dickey-Fuller % Rejected	Maddala-Wu p-Value	Levin-Lin p-Value
Total	5.74	0.0016	0.000
Retail	8.50	0.0000	0.000
Wholesale	13.67	0.0000	0.000
Restaurant	13.67	0.0000	0.000

Although the tests reject unit roots for all employment series, the caveats mentioned above render these rejections less than perfectly informative. The decision to use levels or first differences is therefore somewhat arbitrary. I report the results in levels because they are somewhat easier to interpret. Results from first-difference specifications are extremely similar to the results presented.