JOB CREATION OR DESTRUCTION?
LABOR MARKET EFFECTS OF WAL-MART EXPANSION
Emek Basker*

Abstract—This paper estimates the effect of Wal-Mart expansion on retail employment at the county level. Using an instrumental variables approach to correct for both measurement error in entry dates and 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 as other retail establishments exit and contract, leaving a long-run statistically significant net gain of 50 jobs. Wholesale employment declines by approximately 20 jobs due to Wal-Mart’s vertical integration. No spillover effect is detected in retail sectors in which Wal-Mart does not compete directly, suggesting Wal-Mart does not create agglomeration economies in retail trade at the county level.

I. Introduction

Wal-Mart Corporation employs nearly one million workers in the United States—more than any other private company—and over 300,000 additional workers worldwide. It is rumored to have plans to hire as many as 800,000 additional workers in the next five years. USA Today quotes a retail-industry consultant as saying that Wal-Mart “created more jobs in the 1990s than any other company” (Hopkins, 2003). Has Wal-Mart created more jobs than it destroyed?

Given the level of public interest in Wal-Mart and other “big box” retailers, there has been surprisingly little independent research on their impact on local labor markets.1 Research into this question is hampered by paucity of data on Wal-Mart and the other large retail chains and by concerns about endogeneity of the entry decision. Firms respond to local conditions when they expand or relocate establishments—more so in the nontradable retail sector than in tradable sectors (like manufacturing)—so it is difficult to disentangle the direct effect of expansion from the indirect effects of the conditions that lead to it.2

This paper attempts to quantify the impact of Wal-Mart entry on county-level retail employment by exploiting exogenous variation in the timing of store entry. I use a unique data set containing the locations and opening dates of all U.S. Wal-Mart stores to estimate 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, my instrumental variables (IV) specification exploits the variable lag between store-planning dates and store-opening dates. Store numbers, assigned by Wal-Mart during the planning process, are used to proxy for planning dates. Because my data cover a long time period (1977–1998) and approximately 1,750 counties, I am able to examine the dynamics of county-level retail employment over a 10-year period surrounding Wal-Mart entry, separately estimating short- and long-run effects.

One way Wal-Mart entry could affect labor markets is by increasing average efficiency in the retail sector, so fewer workers are needed per sale. Foster, Haltiwanger, and Kri- zan (2002) find that nearly all productivity growth in the retail sector in the last decade can be accounted for by reallocation of workers due to net entry of establishments; they do not name individual companies in their analysis, but Wal-Mart expansion likely represents an important force in this reallocation. Diffusion of Wal-Mart’s efficient practices—perhaps due to learning/imitation by competitors—may lead other retailers to decrease employment more than proportionately to the decrease in their market share.3

Another possible effect could arise from externalities Wal-Mart creates for other retailers in the county. If Wal-Mart increases customer traffic in the store’s vicinity—like an anchor store in a traditional mall (see Pashigian and Gould, 1998)—the number and size of other retailers could increase, leading to an increase in retail employment. At the county level, this effect is likely to be small because of the nontraded nature of retail services, which operates against concentration in the industry (Holmes and Stevens, 2004). To test this hypothesis against the alternative that Wal-Mart merely captures some of the business of existing retailers, I estimate the effect of Wal-Mart entry on the number of retail establishments in different size categories and its effect on

---

1 Exceptions include Stone’s (1997 and elsewhere) Iowa case studies and other local studies modeled on his approach. These studies do not employ any exogenous variation to predict Wal-Mart entry, so their results are difficult to interpret. Findings range from significant job loss to mild job creation.

2 Bertrand and Kramarz (2002) cleverly avoid this problem, using variation in permits given to large retailers due to exogenous variation in the composition of zoning boards, to analyze the effect of entry on French labor markets. They find that regulation limiting entry of large retailers has slowed employment growth in the French retail industry.

3 Wal-Mart’s lower prices—diffused throughout the local market—may partially offset this effect by increasing demand for retail services (see Basker, 2004).

Received for publication December 5, 2002. Revision accepted for publication February 19, 2004.

* University of Missouri.

I thank Daron Acemoglu, Olivier Blanchard, and Sendhil Mullainathan for their support and guidance. I have also benefited from the comments of Josh Angrist, Saku Aura, David Autor, Glenn Ellison, Bengte Eivenson, Amy Finkelstein, Guido Kuersteiner, Jeffrey Miron, Peter Mueser, Whitney Newey, Marko Terviö, Ken Troske, and Howard Wall, two anonymous referees, and seminar participants at Arizona State University, the Board of Governors of the Federal Reserve System, Case Western Reserve University, the Federal Reserve Banks of Atlanta and St. Louis, Hebrew University, MIT, the University of Chicago GSB, the University of Missouri, Tel Aviv University, the U.S. Census Bureau, Washington University in St. Louis, and Wellesley College. I 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 Zinnman, and especially Maurice Drew, Catherine Friedman, and Steven Sadoway for helping me gather data for this project.

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

job creation or destruction? labor market effects of Wal-Mart expansion

II. Wal-Mart Background

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 served. 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 2,400 stores in all 50 states and approximately 800,000 employees in the United States. The company grows—as measured by the number of employees and the number of stores it operates—by the week. The largest retailer in both the United States and the world, Wal-Mart currently operates in 10 countries.

Wal-Mart operated discount stores as well as 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. (Supercenters employ 400-500 workers each.) By 1998, one-quarter of the 1614 counties entered by Wal-Mart had more than one store.

Wal-Mart is extremely efficient even compared with other “big-box” retailers. Lehman Brothers analysts have noted Wal-Mart’s “leading logistics and information competencies” (Feiner et al., 2001). The Financial Times has called 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 logistics and distribution systems. The chain’s most-cited advantages over small retailers are economies of scale and access to capital markets, whereas against other large retail chains the most commonly cited

These sectors are the only two retail two-digit SIC sectors in which it can be argued Wal-Mart is not a direct competitor. Some Wal-Mart stores do include fast-food restaurants and/or gas stations, but the prevalence of these departments during the sample period is very small.
factor is superior logistics, distribution, and inventory control.5

Wal-Mart’s cost savings extend to its employment practices; it has been accused of requiring employees to work off the clock and using illegal-immigrant labor (through contractors) (see, e.g., Greenhouse, 2002; Buckley and Daniel, 2003). Such practices, if true, could reduce Wal-Mart’s measured employment without reducing its actual labor inputs. Wal-Mart’s low wages are also said to contribute to its measured productivity. Wal-Mart wage data are not publicly available, but several sources estimate the current typical hourly wage of a Wal-Mart “associate” to be $7–$8/hour (Hopkins, 2003). These wages are on a par with wages paid by other large discount chains (like K-Mart and Target), but are typically below union rates.6

III. Data

A. Wal-Mart Stores


The following data sources provide one measure of opening dates: Vance and Scott (1994) list store entries to 1969, the year the company became publicly traded. Annual reports between 1970 and 1978 include lists of current stores; after 1978 annual reports list only the number of stores per state. The annual Directory of Discount Department Stores provides store lists between 1979 and 1993. For recent years I use a special edition of the popular Rand McNally road atlas which contains a list of Wal-Mart store locations, and includes each store’s company-assigned number. The variable \( WMopen_{jt} \) 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 (counterfactual) 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). Wal-Mart assigns store numbers roughly in sequential order, with store 1 opening first, followed by store 2, and so on. Following this practice, I assign entry dates to stores sequentially, based on their store numbers. This assignment method provides a very good approximation to the likely order in which the stores were planned. Aggregating these store-level entry dates to the county-year level, I construct \( WMplan_{jt} \), which 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.

More details on variable construction are in section 1 of the Appendix.

B. Employment Data

The unit of observation is a county-year. Of the 3,111 counties in the contiguous 48 states, I limit the data set to the 1,749 counties with total employment in 1964 above 1,500, 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 2. Annual county-level employment data by two-digit SIC (three-digit NAICS) for 1977–1999 come from the Census Bureau’s County Business Patterns (CBP) serial (U.S. Bureau of the Census, 1964–1999). The panel contains 40,227 observations (1,749 counties over 23 years). Unfortunately, CBP does not contain wage data.

5 Details on Wal-Mart’s operations can be found in Harvard Business School’s three case studies about Wal-Mart (Ghemawat, 1989; Foley and Mahmood, 1996; Ghemawat and Friedman, 1999).

6 In markets where Wal-Mart competes directly with unionized retailers, it is said to match the union wage (Saporito, 2003).

7 The relevant SIC (NAICS) codes are as follows. Retail: SIC 52–59 except 55 and 58, NAICS 44 except 441 and 447; wholesale: SIC 50–51, NAICS 42; restaurants: SIC 58, NAICS 721; automobile: SIC 55, NAICS 441 and 447; manufacturing: SIC 20–39, NAICS 31.
Table 1 lists some summary statistics for labor-market data. More details are in Appendix A.2.

IV. Methodology

A. OLS Regressions

Because the data do not appear to contain unit roots, the analysis is done using employment levels. (See section 3 of the Appendix for details on unit root tests. Results using first differences are extremely similar to the ones reported here.) For county \( j \) in year \( t \) I estimate

\[
\frac{retail_{jt}}{pop_{jt}} = \alpha + \sum_{k} \delta_{k} \text{urban}_{jk} \text{year}_{t} + \sum_{j} \psi_{j} \text{county}_{j} \nonumber
\]

\[+ \theta(L) \frac{WMopen_{jt}}{pop_{jt}} + \epsilon_{jt}, \tag{1} \]

where \( retail_{jt} \) is retail employment; \( pop_{jt} \) is population size; \( year_{t} \) is a year dummy; \( urban_{jk} \in \{ \text{urban, suburban, rural} \} \) is a set of three urbanization dummies allowing for different year fixed effects for urban, suburban, and rural counties; \( \text{county}_{j} \) is a county dummy; \( WMopen_{jt} \) is the number of Wal-Mart store openings in the county; and \( \theta(L) \) is a lag polynomial with five leads, a contemporaneous value, and six lags (the sixth lag represents the cumulative period 6 or more years after year \( t \); the reference period is 6 or more years before year \( t \)):

\[\theta(L) = \theta_{1}F^{5} + \theta_{2}F^{4} + \theta_{3}F^{3} + \theta_{4}F^{2} + \theta_{5}F + \theta_{6} + \theta_{7}L^{10} + \theta_{8}L^{5} + \theta_{9}L^{3} + \theta_{10}L^{2} + \theta_{11}L^{1} + \theta_{12} \sum_{\tau \geq 6} L^{\tau}, \nonumber\]

where \( L \) is the lag operator and \( F \) is the lead operator. The error term \( \epsilon_{jt} \) is clustered at the county level to allow for arbitrary autocorrelation.

Both employment and the Wal-Mart variable are divided by the current county population, so the coefficients \( \theta(L) \) can be interpreted as the effect of one additional Wal-Mart store on the level of retail employment.\(^9\) Plots of the coefficients \( \theta(L) \) show the evolution of employment over a 10-year period, starting 5 years before and ending 5 years after Wal-Mart entry into a county. The coefficient \( \theta_{12} \), intended to capture the permanent effect of Wal-Mart entry on employment 6 or more years after entry, is omitted from the graphs because it is identified using relatively few observations.

The OLS estimates are valid if Wal-Mart entry is correctly measured and exogenous to employment changes. Unfortunately, \( WMopen_{jt} \) is measured with error, and may be endogenous to retail employment outcomes. An IV specification is used to correct these problems.

B. Measurement Error

Measurement error in the Wal-Mart entry variable \( WMopen_{jt} \) takes a particular form: though the entered counties are correctly identified, the \textit{timing} of entry may be incorrectly measured due to errors in the directories. (An extreme example of such errors is the lack of updating of the \textit{Directory of Discount Department Stores} between 1990 and 1993, but other errors exist as well.) The counter factual variable \( WMplan_{jt} \) is also measured with error, by construction: it represents the number of stores that would have opened had stores always opened in the order in which they were planned.

An IV approach, in which leads and lags of \( WMplan_{jt} \) are used to instrument for leads and lags of \( WMopen_{jt} \), can be used to correct for this measurement error if the measurement errors in the two variables are classical and uncorrelated. That the measurement error across the two variables is uncorrelated seems plausible.\(^10\) But because \( WMopen_{jt} \) and \( WMplan_{jt} \) are both discrete, their measurement error is not classical: the actual number of Wal-Mart stores in city \( j \) in year \( t \) differs from the measured number by an integer whose expected mean is different from 0. This induces a slight bias in the IV results reported here.\(^11\)

C. Endogeneity

Another difficulty in assessing the impact of Wal-Mart entry on the level and composition of county employment is

\(^9\)The use of per capita terms on both the left- and right-hand sides of equation (1) could cause a spurious correlation between the variables that would bias the estimated coefficients. In practice, the year-to-year variation in county population is small enough that it is not driving the results; the results are robust to normalization by a constant such as the 1990 county population. See also footnote 18.

\(^10\)This assumption would be violated if some stores, for example in metropolitan areas, experienced shorter planning phases and were also more likely to appear in the directories sooner, due to better directory coverage. This does not appear to be the case.

\(^11\)Because store closings are exceedingly rare, when the directories report zero new Wal-Marts in town, the expected number of openings is some (small) positive number. Similarly, when the reported number of new Wal-Marts is 1, the expected number is less than 1. 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.

\(^8\)A county is defined as urban if it 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.
the possible endogeneity of Wal-Mart’s entry decision with respect to retail employment. This potential endogeneity has two dimensions: Wal-Mart may select the counties it enters nonrandomly, and it may choose the timing of entry nonrandomly.

If Wal-Mart selects counties whose growth rates exceed those of nonentered counties, a spurious positive effect will be registered by the estimated coefficients $\hat{\theta}$ (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 1,500 and a positive average growth rate of total employment between 1964 and 1977. I also remove counties entered by Wal-Mart before 1977, to eliminate concerns about the endogeneity of employment growth. Wal-Mart entered 75% of the remaining 1,749 counties between 1977 and 1998, compared with only 13% of the excluded counties.\textsuperscript{12}

To address endogeneity of the timing of entry (conditional on the counties selected for entry), I rely again on store planning dates described above. This identification strategy assumes that Wal-Mart plans its store entries well in advance of entry and cannot accurately forecast exact market conditions at the time for which entry is planned. Because the company may fine-tune entry dates based on current market conditions, the actual entry date may be endogenous, but can be instrumented for using the date for which entry was planned. For the purposes of exposition, suppose Wal-Mart has a once-and-for-all effect on retail employment (so we can estimate a simple differences-in-differences model). We would like to estimate the coefficient $\hat{\theta}$ from the equation

$$ retail_{jt}^{WM} = retail_{jt}^0 + \theta, $$

where $retail_{jt}^{WM}$ is retail employment county $j$ in year $t$ in the presence of a Wal-Mart store, and $retail_{jt}^0$ is retail employment in the absence of the Wal-Mart store. Because we cannot observe both $retail_{jt}^{WM}$ and $retail_{jt}^0$ for a given county-year pair, OLS estimates implicitly assume that

$$ retail_{jt}^0 = \alpha + \sum_i \delta_i \text{year}_i + \sum_j \psi_j \text{county}_j + u_{jt}, $$

$$ \mathbb{E}(u_{jt}|WMopen_{jt}) = 0. $$

That is, the presence of a Wal-Mart store in county $j$ in year $t$ is uncorrelated with the error term in the retail employment equation: controlling for some basic county characteristics (in this case, simply county fixed effects), Wal-Mart entry is exogenous. This assumption is a very strong one, and unlikely to be true.

The IV strategy described above corrects for this endogeneity concern under two identifying assumptions: the number of planned Wal-Mart stores (per capita) for county $j$ and year $t$ is independent of the error term in equation (2); and planned Wal-Mart stores affect retail employment per capita only insofar as they are correlated with the actual construction of Wal-Mart stores. That is,

$$ \mathbb{E}(u_{jt}|WMplan_{jt}) = 0. $$

As this discussion suggests, the IV estimator $\hat{\theta}$ will be biased if plans to build a Wal-Mart store spur the building of a strip mall—or the closing of an existing store—in anticipation of Wal-Mart’s entry, even if Wal-Mart does not actually open a store in that county the year it planned to.\textsuperscript{13} The estimator will also be biased if Wal-Mart’s planners can foresee employment fluctuations at the time of the store’s planning, or if planning dates anticipate that a growth spurt will occur over the next few years and the timing of entry is then adjusted to coincide (on average) with such a spurt.

Because the regression equation is exactly identified (12 leads and lags of $WMplan_{jt}$ instrument for 12 leads and lags of $WMopen_{jt}$), these identifying assumptions cannot be tested directly. I employ an indirect test instead, using the lead coefficients $\theta_1-\theta_5$: if Wal-Mart times entry to take advantage of retail growth spurts, then (under most conditions) we should see some increase in retail employment in $\theta_1-\theta_5$. This increase would be apparent in OLS estimates, but will be absent in IV estimates if the IV strategy corrects for this endogeneity.\textsuperscript{14}

I also estimate the effect of Wal-Mart entry on manufacturing employment, using it as a falsification test: if Wal-Mart planned entry to coincide with general employment increases, manufacturing employment would increase with Wal-Mart entry. As the results below show, this is not the case. I conclude that the IV strategy appears to correct for endogeneity as well as measurement error.

V. Results

A. Retail Employment

To begin, I present OLS results using $WMopen_{jt}$ entry dates in figure 3.\textsuperscript{15} Retail employment is shown to increase by approximately 40 jobs in the year of entry, up to half of which are eliminated within five years. In addition, 20 jobs

\textsuperscript{12} Indistinguishable results are obtained if the sample is limited instead to entered counties.

\textsuperscript{13} Anecdotal evidence suggests that small retailers tend to continue operating as long as they can, even when this is not profit-maximizing (Peled, 2001).

\textsuperscript{14} The test is imperfect. To see this, suppose Wal-Mart forecasters can predict which counties will experience growth spurts in retail employment over the next few years, and plan to open stores in those counties. Under this scenario, Wal-Mart’s planned entry dates would coincide imperfectly with growth spurts, but actual opening dates could be adjusted (for example, by delaying construction) to fall precisely during these spurts. If employment growth of the sort Wal-Mart uses to fine-tune its entry arrives in isolated spurts (i.e., Wal-Mart does not enter counties experiencing sustained growth in retail employment that lasts for several years), there will be no pre-entry growth in the IV estimates even if Wal-Mart entry is not causally associated with any increase in employment.

\textsuperscript{15} Throughout the paper, the 95% confidence intervals shown use asymptotic standard errors clustered at the county level. The reference period is 6 or more years prior to Wal-Mart entry.
are estimated to have been created in the year before Wal-Mart entry. Though this increase is small in absolute magnitude, it is statistically significant and disconcertingly large relative to the estimated postentry effect.\footnote{Reduced-form estimates using WMplan$_j$ are extremely similar.}

The IV results are shown in figure 4. The effect of entry is estimated much more cleanly at approximately 100 jobs. In the years immediately following entry, there is a loss of 40–60 jobs. The net effect at the 5-year horizon, however, is positive and significant ($p$-value 0.0003).

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. 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.

Very little is known about employment conditions at Wal-Mart, including the prevalence of part-time work. A reasonable prior is that Wal-Mart employees work fewer hours than other retail workers [using French data, Bertrand and Kramarz (2002) find that entry of large retailers is increases part-time employment 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 is within the norm for workers in the discount retail industry (Peled, 2001), and also in keeping with the rest of the retail industry: the 30th percentile of hours worked by retail employees, computed from the March Current Population Survey (CPS) for 1978–1999, is 28 hours across employer size, state, and year.

As noted in section IV C, if the timing of entry were endogenous, we would expect to see an increase in the county’s retail employment prior to entry. No such effect is evident in the leading coefficients, although, as footnote 14 makes clear, this is not conclusive evidence in support of the identifying assumption.

**B. Retail Establishments**

To capture the effect of Wal-Mart on the number of retail establishments, I estimate IV regressions replacing the left-side variable retail$_j$/pop$_j$ by estab$_j$/pop$_j$, where estab$_j$ is the number of retail establishments in county $j$ at year $t$ in each of three size categories.

To confirm that Wal-Mart’s creation can be detected in the data, I estimate the regressions using the number of large retail establishments (with 100 or more employees). IV results are shown in figure 5; the estimated coefficients mirror those on retail employment shown in figure 4. The increase in the number of large retail establishments, approximately 0.7, suggests that Wal-Mart’s entry often coincides with exit or contraction of other large retailers. In some cases, Wal-Mart acquired a large number of stores from a competitor; in other cases, incumbent establishments

![Figure 3.—Evolution of Retail Employment (OLS)](image1)

![Figure 4.—Evolution of Retail Employment (IV)](image2)

![Figure 5.—Evolution of Number of Large Retail Establishments (IV)](image3)
may have chosen to exit preemptively.\textsuperscript{17} There is a small decline in the number of large establishments in subsequent years.

Figure 6 shows the effect of Wal-Mart on the number of small establishments (with fewer than 20 employees). It shows a decline of four retail establishments within 5 years of Wal-Mart entry, three of them within 2 years of entry. The number of medium-size establishments (with 20–99 employees), shown in Figure 7, decreases by approximately 0.7 in the second year following entry, then remains flat.

C. Other Sectors

Wal-Mart competes with establishments in a wide array of sectors, some more directly than others. Wal-Mart Supercenters compete directly with grocery stores, whereas discount stores do not; all Wal-Mart stores compete with apparel stores, hardware stores, bookstores, music stores, and so on. Moreover, because Wal-Mart is vertically integrated, it competes against wholesalers as well as retailers. In this section, I look for an effect of Wal-Mart on wholesale employment, which could be due to Wal-Mart’s indirect competition with wholesalers. I also look for an effect of Wal-Mart on two retail segments—restaurants and automobile dealerships and service stations—where Wal-Mart does not compete. Finally, as a falsification exercise, I estimate Wal-Mart’s effect on manufacturing employment.

The estimated effect of Wal-Mart entry on county-level wholesale employment is shown in figure 8. The observed decline of approximately 20 wholesale jobs following Wal-Mart entry is marginally significant (\( p \)-value 0.0682).

I use employment in restaurants and in automobile sales and service to test for an agglomeration effect of Wal-Mart entry. If these sectors expand following Wal-Mart entry, one interpretation could be that Wal-Mart creates positive externalities for other retailers in the area. I find no evidence that Wal-Mart entry affects either of these two sectors. The evolution of employment by restaurants is shown in figure 9. Although restaurant employment per capita grows throughout the period surrounding Wal-Mart entry, there is no discontinuity in that trend associated with the entry. Figure 10 shows employment in automobile sales and service. Here there is no trend whatsoever: Wal-Mart entry is not associated with any changes in employment in this sector. These results suggest that agglomeration economies generated by Wal-Mart (if any) must be at a level of aggregation substantially smaller than the county.

Manufacturing employment is shown in figure 11. The confidence intervals are very large and show some large fluctuations over the 10 year period shown, but no substantial increase or break in manufacturing employment can be seen before or at the time of Wal-Mart entry. It appears unlikely that Wal-Mart plans its new stores based on prior knowledge about future growth in the manufacturing sector. Because the typical Wal-Mart store has 150–350 employees—less than 2% of total employment in the average

\textsuperscript{17} Examples of establishment acquisition include Wal-Mart’s 1977 purchase of 16 Mohr Value Discount Department Stores in Missouri and Illinois, and its 1981 purchase of 106 stores in nine states from Kuhn’s-Big K Stores Corp.
county at the time of the Wal-Mart entry—it is unlikely to have a significant effect on total county employment, and indeed, the estimated effect on total employment (not shown) is statistically 0.18

D. Neighboring Counties

If Wal-Mart’s effect on the retail industry in the entered county is due to agglomeration economies, entry could produce negative effects on neighboring counties’ employment (in both the retail and wholesale sectors). Unfortunately, estimating Wal-Mart’s effect on neighboring counties with any precision is impossible, as the confidence intervals around the point estimates are very large.

I define counties as neighbors if the distance between their geographic centers is within some fixed range (e.g., 5, 10, or 25 miles). To estimate the effect of Wal-Mart entry in county \( j \) on retail employment in the surrounding area, I use the same strategy as with own-county effects, but replace retail employment per capita in the entered county with retail employment per capita in neighboring counties. I include controls for the number of Wal-Mart stores in neighboring counties to avoid estimating a spurious relationship between Wal-Mart entry in county \( j \) and employment in neighboring counties (in as much as Wal-Mart entry dates are correlated in neighboring counties).

In the estimated results, not shown, confidence intervals are too large to reject any effect of Wal-Mart, positive or negative, on employment in the distributive trades in neighboring counties.

VI. Conclusion

Wal-Mart entry has raised concerns in many communities about the changes it may cause in the size and structure of the retail industry. Wal-Mart’s reputed efficiency, combined

---

18 I have also tested for an effect of Wal-Mart entry on county population, using annual Census Bureau estimates of county population for the years 1977–1999. I find no effect of Wal-Mart entry on population.
Finally, because I use a large panel of approximately 1,750 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 county retail employment in 1990 was only 1,500, and the average Wal-Mart store had approximately 300 employees—the effect can be estimated with fair precision. Of course, these effects represent the average impact of Wal-Mart and may not be representative of any individual county’s experience.

The small magnitude of the estimated net effect of Wal-Mart on retail employment is striking in light of the level of public discussion on this topic. Other effects Wal-Mart entry—for example, on prices, tax revenue, or the environment—have not been ruled out by this analysis; nor has the possibility that the small net county-level effects described here mask much larger reallocations at the sub-county level. Publicly available data cannot address that concern, so it remains an open question.

REFERENCES


APPENDIX

Data and Empirical Issues

1. Wal-Mart Data

Table A1 shows the sources from which store opening dates, used in the construction of the variable $W_{t,0}$, were drawn. Chain Store Guides Directory of Discount Department Stores from 1990–1993 are available, but are largely uninformative; the directories appear not 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., existed in 1994), opening dates are assigned according to the following algorithm. From the annual reports, I obtain the net increase in the number of Wal-Mart stores in each state each year. Because very few store closures occur, I use the net increase to proxy for the number of new stores to open each year in each state. For example, in Arizona, five new stores opened in 1990, seven 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. Each store that opened in Arizona during this period is assigned the opening date 1990 with probability $2/1991$ with probability $2$; and 1992 and 1993 with probability $1$ each. In all, 680 stores’ opening dates are assigned in this way: 203 in 1990, 145 in 1991, 181 in 1992, and 151 in 1993.

The accuracy of this method depends critically on Wal-Mart 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 cannot be common.

2. Employment Data

In counties with a small number of employees, data on the total number of employees in a sector may be omitted from County Business Patterns to avoid disclosure of the number of employees in individual firms; a range (1–19, 20–99, etc.) is then given instead of an exact number of employees. In those instances, I assume that the actual number of workers is a weighted mean of the lower and upper bounds on the given employment range (with weight $w_i$ on the lower bound and $w_i$ on the upper bound). For example, a firm with 1–19 employees is assigned a value of 7.21

3. Unit Roots

To test whether county-level employment per capita contains unit roots, I run a Dickey-Fuller (DF) test on each county series separately, entry dates assigned in this way are measured with error, but they are unbiased.

I chose to weight the lower and upper bounds of each interval by $w_i = 1$ and $w_i = 1$, respectively, rather than $w_i = 1$. 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.
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. Because DF tests sometimes fail to reject unit roots even when none are present, I also use the more powerful Maddala-Wu (1999) and Levin-Lin, (Levin, Lin, & Chu, 1993) panel unit root tests. The validity of panel unit root tests depends on the series being independent realizations of a single common process.

Table A2 reports the test results. The first column shows the fraction of 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. A common unit root process is rejected by both tests for each series.

<table>
<thead>
<tr>
<th>Table A1.—Directory Sources for Wal-Mart Opening Dates</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Years</strong></td>
</tr>
<tr>
<td>1979–1982</td>
</tr>
<tr>
<td>1983–1986</td>
</tr>
<tr>
<td>1987–1989</td>
</tr>
<tr>
<td>1990–1993</td>
</tr>
<tr>
<td>1994–1997</td>
</tr>
</tbody>
</table>


<table>
<thead>
<tr>
<th>Table A2.—Unit Root Tests</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Employment</strong></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td>Retail</td>
</tr>
<tr>
<td>Wholesale</td>
</tr>
<tr>
<td>Restaurant</td>
</tr>
<tr>
<td>Automotive</td>
</tr>
<tr>
<td>Manufacturing</td>
</tr>
</tbody>
</table>