Job Creation or Destruction?

Labor-Market Effects of Wal-Mart Expansion

Emek Basker *,†
University of Missouri
January 2004

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.

JEL Numbers: J21, L11, L81

*emek@missouri.edu. 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 Evenson, 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 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 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 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.

1 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 surprising 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 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 ten-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 Krizan (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 (forthcoming)). 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 retail employment in establishments not directly competing with Wal-Mart. For the second exercise, I use employment in two sectors, restaurants and automobile dealerships and service stations.[4] If Wal-Mart merely substitutes for other retailers, the number of small and medium-sized retail establishments should decline, and Wal-

Mart will not have any impact on retailers in sectors where Wal-Mart does not compete directly. If Wal-Mart entry creates positive externalities for other retailers, the number of small and medium-sized retail establishments could increase, as could employment in restaurants and automobile dealerships. I also test for a negative indirect effect of Wal-Mart entry on retail employment in neighboring counties to see whether Wal-Mart entry imposes negative externalities on nearby communities.

To see how much of Wal-Mart's measured effect on retail employment is due to a change in the classification of workers in the county from wholesale to retail, I also estimate the effect of Wal-Mart entry on wholesale employment. Finally, a falsification test estimates Wal-Mart's effect on manufacturing employment.

I find that immediately after entry, retail employment in the county increases by approximately 100 jobs; this figure declines by half over the next five years as some small and medium-sized retail establishments close. Wholesale employment declines by approximately 20 jobs over five years. Restaurant employment increases slightly; there is no change in employment in manufacturing or in automobile dealerships and service stations. No effect can be detected on retail employment in neighboring counties, due to very large confidence bands.

The remainder of the paper is organized as follows. Section 2 provides background information on Wal-Mart. Section 3 describes the data. My empirical strategy is explained in Section 4, and results are presented and discussed in Section 5. Section 6 concludes.

2 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 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. 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 U.S. 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 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 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. While Wal-Mart wage data are not publicly available, several sources estimate the current typical hourly wage of a Wal-Mart "associate" to be \$7-\$8/hour (Hopkins 2003). These wages are on par with wages paid by other large discount chains (like K-Mart and Target), but are typically below union rates.[6]

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

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 $\mathbf{WMopen_{jt}}$ 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 (counter-factual) 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 $\mathbf{WMplan_{jt}}$. $\mathbf{WMplan_{jt}}$ 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.

More details on variable construction are in Appendix A.1 .

3.2 Employment Data

The unit of observation is a county-year. Of the 3111 counties in the contiguous 48 states, I limit the data set to the 1749 counties with total employment in 1964 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 2. Annual county-level employment data by 2-digit SIC (3-digit NAICS) for 1977-1999 come from the Census Bureau's County Business Patterns (CBP) serial. The panel contains 40,227 observations (1749 counties over 23 years).[7] Unfortunately, CBP does not contain wage

data.

Table 1 lists some summary statistics for labor-market data. More details are in Appendix A.2.

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 A.3 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

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

where $\mathbf{retail_{jt}}$ is retail employment; $\mathbf{pop_{jt}}$ is population size; $\mathbf{year_t}$ is a year dummy; $\mathbf{urban_{jk}} \in \{\text{urban, suburban, rural}\}$ is a set of three urbanization dummies allowing for different year fixed effects for urban, suburban, and rural counties;[8] $\mathbf{county_j}$ is a county dummy; $\mathbf{WalMart_{jt}}$ is the number of new Wal-Mart stores in the county; and $\theta(L)$ is a lag polynomial with six lags and five leads (the sixth lag represents the cumulative period six or more years after year t; the reference period is six or more years before

year t):

$$\theta(L) = \theta_1 F^5 + \theta_2 F^4 + \theta_3 F^3 + \theta_4 F^2 + \theta_5 F + \theta_6 + \theta_7 L + \theta_8 L^2 + \theta_9 L^3 + \theta_{10} L^4 + \theta_{11} L^5 + \theta_{12} \sum_{\tau \geq 6} L^\tau$$

where L is the lag operator and F is the lead operator. The error term $\mathbf{u_{jt}}$ 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 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 relatively few observations.

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

4.2 Measurement Error

Measurement error in the Wal-Mart entry variable $\mathbf{WMopen_{jt}}$ takes a particular form: while the entered counties are correctly identified, the *timing* of entry may be incorrectly measured 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, but there are other errors as well.) The counter-factual variable $\mathbf{WMplan_{jt}}$ 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 instrumental-variables (IV) approach, in which leads and lags of $\mathbf{WMplan_{jt}}$ is used to instrument for leads and lags of $\mathbf{WMopen_{jt}}$, 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.[10] But because $\mathbf{WMopen_{jt}}$ and $\mathbf{WMplan_{jt}}$ 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 zero. This induces a slight bias in the instrumental-variables results reported here.[11]

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 with respect to retail employment. There are two dimensions to this potential endogeneity: Wal-Mart may select the counties it enters non-randomly, and it may choose the timing of entry non-randomly.

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. 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 1749 counties between 1977 and 1998, compared with only 13% of the excluded counties.[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 θ from the equation

$$\text{retail}_{jt}^{WM} = \text{retail}_{jt}^{0} + \theta$$

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

$$\frac{\text{retail}_{jt}^{0}}{\text{pop}_{jt}} = \alpha + \sum_{t} \delta_t \text{year}_t + \sum_{j} \psi_j \text{county}_j + \text{u}_{jt} \tag{2}$$

$$\mathbb{E}\left(\text{u}_{jt} \mid \text{WMopen}_{jt}\right) = 0.$$

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 instrumental-variables 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 are 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,

$$\mathbb{E}\left(\mathrm{u}_{jt} \mid \mathrm{WMplan}_{jt}\right) = 0.$$

As this discussion suggests, the IV estimator $\widehat{\theta}$ 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 anticipate of Wal-Mart's entry, even if Wal-Mart does not actually open a store in that county the year it planned to.[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 (actual) 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 $\mathbf{WMplan_{jt}}$ instrument for 12 leads and lags of $\mathbf{WMopen_{jt}}$), 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 lead coefficients $\theta_1 - \theta_5$. This increase would be apparent in OLS estimates, but will be absent in IV estimates if the instrumental-variables strategy corrects for this endogeneity.[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.

5 Results

5.1 Retail Employment

To begin, I present OLS results using $\mathbf{WMopen}_{jt}$ entry dates in Figure 3.[15] Retail employment is shown to increase by about 40 jobs in the year of entry, up to half of which are eliminated within five years. In addition, 20 jobs are estimated to have been created in the year *before* Wal-Mart entry. While this increase is small in absolute magnitude, it is statistically significant and disconcertingly large relative to the estimated post-entry effect.[16]

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 five-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 4.3, 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, footnote 14 makes clear, this is not conclusive evidence in support of the identifying assumption.

5.2 Retail Establishments

To capture the effect of Wal-Mart on the number of retail establishments, I estimate instrumental-variables regressions replacing the LHS variable $\frac{\mathbf{retail_{jt}}}{\mathbf{pop_{jt}}}$ by $\frac{\mathbf{estab_{jt}}}{\mathbf{pop_{jt}}}$, where $\mathbf{estab_{jt}}$ 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 (100 or more employees). IV results are shown in Figure 5; 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 may have chosen to exit preemptively.[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 (fewer than 20 employees). It shows a decline of 4 retail establishments within 5 years of Wal-Mart entry, 3 of them within two years of entry. The number of medium-sized establishments (20-99 employees), shown in Figure 7, decreases by about 0.7 in the second year following entry, then remains flat.

5.3 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, while Discount Stores do not; all Wal-Mart stores compete with apparel stores, hardware stores, bookstores, music stores, etc. Moreover, since 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 manu-

facturing 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 due to 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. While restaurant employment per capita grows throughout the period surrounding Wal-Mart entry, there is no discontinuity in that trend associated with entry of Wal-Mart. 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 there are 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. Since 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 – it is unlikely to have a significant effect on total county employment, and indeed, the estimated effect on total employment (not shown) is statistically zero.[18]

5.4 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 (since 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 the distributive trade in neighboring counties.

6 Conclusion

Wal-Mart entry has raised concerns in many communities about the changes it may cause to the size and structure of the retail industry. Wal-Mart's reputed efficiency, combined with its market power, could cause a decline in the number of retail jobs in the community. In this paper, I do not take a position on whether such a decline is

favorable or harmful, but merely attempt to quantify the effect of Wal-Mart entry on retail employment and on the number of retail establishments. Using an instrumental-variables approach to correct for endogeneity of Wal-Mart entry, and measurement error in the data, I conclude that Wal-Mart entry has a small positive effect on retail employment at the county level while reducing the number of small retail establishments in the county. I also find a small negative effect of Wal-Mart entry on wholesale employment. No effect can be seen on retail sectors in which Wal-Mart does not compete directly: restaurants and automobile sales and service. The latter fact suggests that Wal-Mart does not create large agglomeration externalities at the county level. Wal-Mart's effect on neighboring counties cannot be precisely estimated, so I cannot determine whether Wal-Mart entry reduces retail employment in neighboring counties.

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 to adjust to the shock over a period of several years. Finally, because I use a large panel of approximately 1750 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 1500, while the average Wal-Mart store had approximately 300 employees – the effect can be estimated with relative 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 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.

A Data and Empirical Issues

A.1 Wal-Mart Data

Table 2 shows the sources from which store opening dates, used in the construction of the variable $\mathbf{WMopen_{jt}}$, were drawn. Chain Store Guides' *Directories 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., exist 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. Since there are very few store closures, I use the net increase to proxy for 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. Each store that opened in Arizona during this period is assigned the opening date 1990 with probability $\frac{5}{14}$; 1991 with probability $\frac{1}{2}$; and 1992 and 1993 with probability $\frac{1}{14}$ 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.[19]

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.

A.2 Employment Data

In counties with a small number of employers, 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 $\frac{2}{3}$ on the lower bound and $\frac{1}{3}$ on the upper bound). For example, a firm with 1-19 employees is assigned a value of 7.[20]

A.3 Unit Roots

To test whether county-level employment per capita contain 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. Because DF tests sometimes fail to reject unit roots even when none are present, I also use the Maddala and Wu (1999) and Levin and Lin (1993) more powerful panel unit root tests. Validity of panel unit root tests depend on the series being independent realizations of a single common process.

Table 3 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. A common unit root process is rejected by both tests for each series.

Notes

[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 signficant 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 (Basker 2004).

[4]These sectors are the only two retail 2-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.

[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, and 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.

[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.

[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, experience 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]Since 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 one, the expected number is less than one. 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.

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

[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).

[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 would be no pre-entry growth in the IV estimates even if Wal-Mart entry is not causally associated with any increase in employment.

[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.

[16]Rduced-form estimates using $\mathbf{WMplan_{jt}}$ are extremely similar.

[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.

[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.

[19]Entry dates assigned in this way are measured with error, but they are unbiased.

[20]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.

References

[1] Basker, Emek (2004). "Selling a Cheaper Mousetrap: Entry and Competition in the Retail Sector." University of Missouri Working Paper 04-01.

[2] Bertrand, Marianne, and Francis Kramarz (2002). "Does Entry Regulation Hinder Job Creation? Evidence from the French Retail Industry." *Quarterly Journal of Economics* 117, 1369-1414.

[3] Buckley, Neil and Caroline Daniel (2003). "Wal-Mart vs. the Workers." *Financial Times*, November 20, 2003.

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

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

[6] Edgecliffe-Johnson, Andrew (1999). "A Friendly Store from Arkansas." *Financial Times*, June 19, 1999.

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

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

[9] 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.

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

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

[12] Greenhouse, Steven (2002). "Suits Say Wal-Mart Forces Workers to Toil Off the Clock." *New York Times*, June 25, 2002.

[13] Holmes, Thomas J. and John J. Stevens (forthcoming). "Geographic Concentration and Establishment Size: Analysis in an Alternative Economic Geography Model." *Journal of Economic Geography.*

[14] Hopkins, Jim (2003). "Wal-Mart's Influence Grows." *USA Today*, January 29, 2003.

[15] 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.

[16] Levin, Andrew, Chien-Fu Lin, and Chia-Shang James Chu (2002). "Unit Root Tests in Panel Data: Asymptotic and Finite-sample Properties." *Journal of Econometrics* 108:1, 1-24.

[17] 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.

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

[19] Peled, Micha (2001). *Store Wars: When Wal-Mart Comes to Town*. Independent Television Service film.

[20] Saporito, Bill (2003). "Can Wal-Mart Get Any Bigger? Yes, a Lot Bigger... Here's How." *Time*, January 13, 2003.

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

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

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

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

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

Table 1: Summary Statistics

	Mean	Median
Population	120,500	42,000
Retail Employment	5,000	1,300
Automotive Retail Employment	1,000	350
Restaurant Employment	3,000	800
Wholesale Employment	3,000	550
Manufacturing Employment	7,500	2,300
Number of Small Retail Establishments[a]	360	130
Number of Medium Retail Establishments[a]	35	11
Number of Large Retail Establishments[a]	8	1

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

Table 2: 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 3: Unit Root Tests

Employment	Dickey-Fuller % Rejected	Maddala-Wu p-Value	Levin-Lin p-Value
Retail	9.43	0.0000	0.0000
Wholesale	13.72	0.0000	0.0000
Restaurant	13.72	0.0000	0.0000
Automotive	12.12	0.0000	0.0000
Manufacturing	6.12	0.0000	0.0000

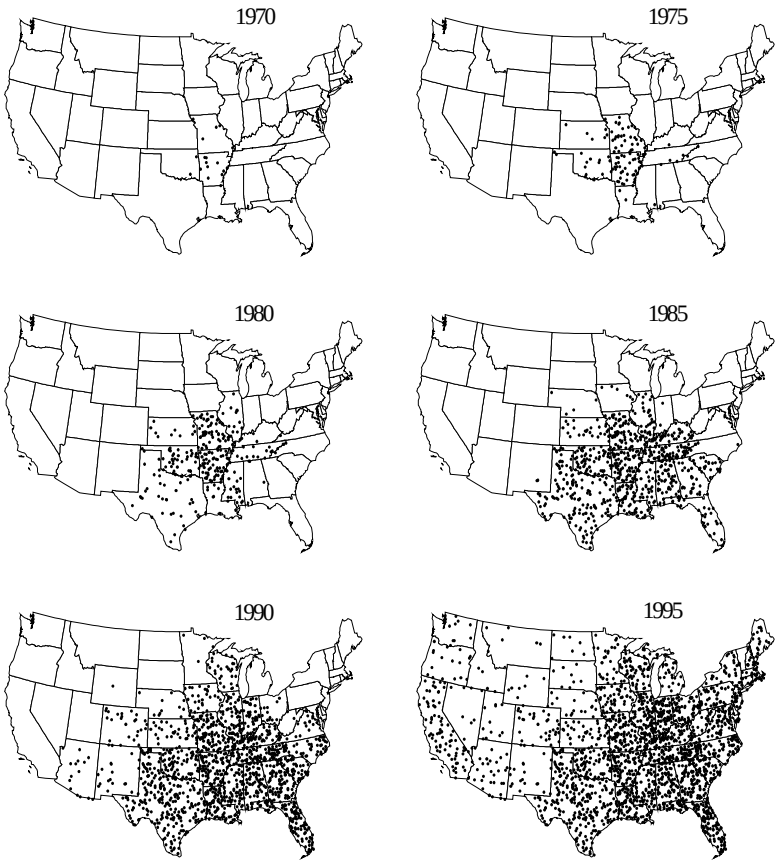

Figure 1: Location of Wal-Mart Stores, Various Years

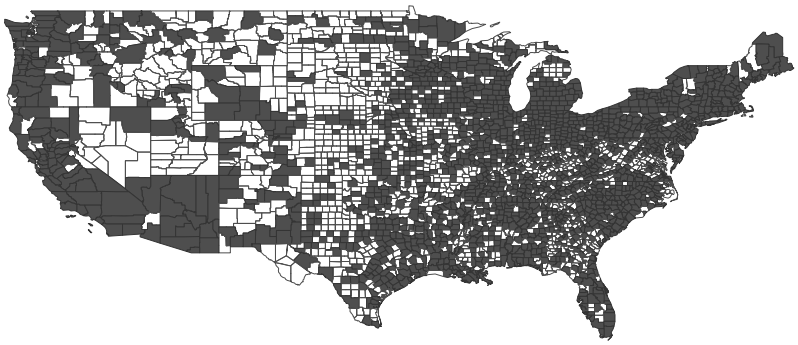

Figure 2: Counties Included in Analysis (Shaded Regions)

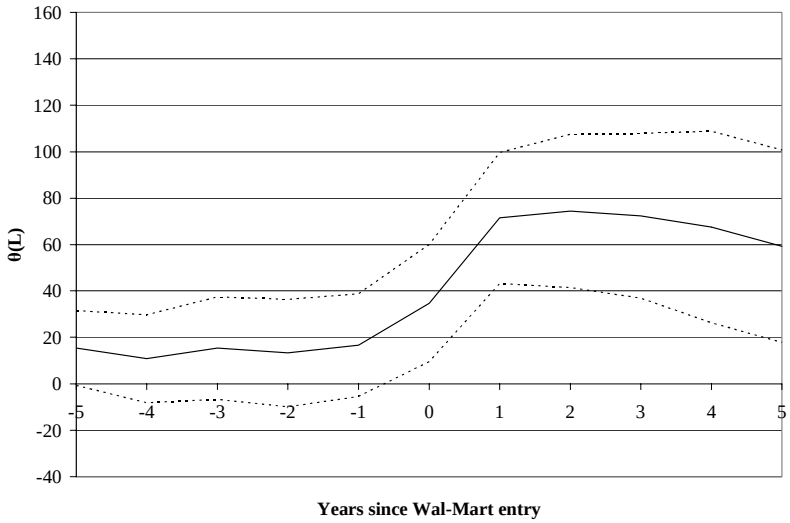

Figure 3: Evolution of Retail Employment (OLS)

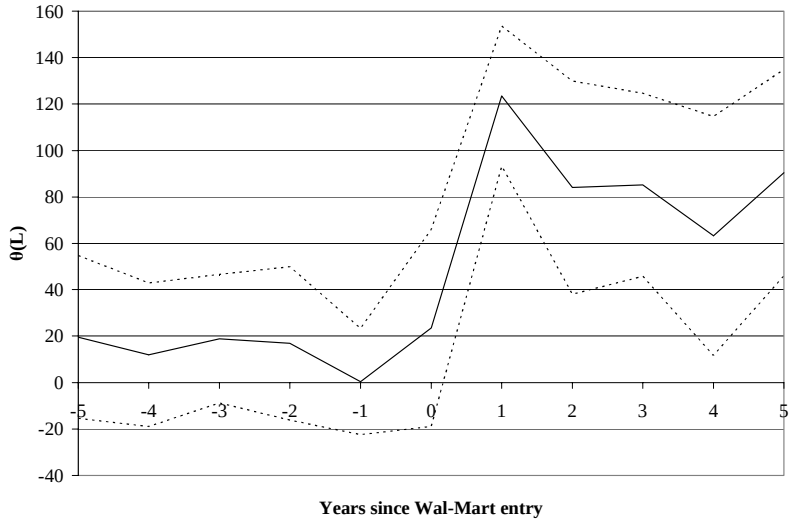

Figure 4: Evolution of Retail Employment (IV)

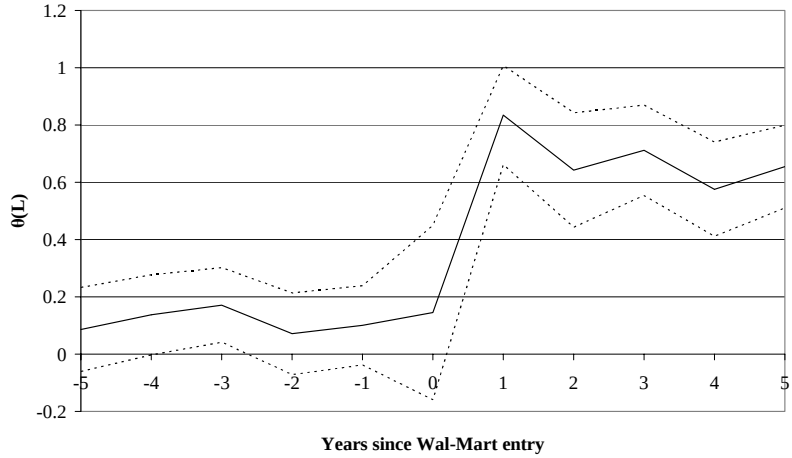

Figure 5: Evolution of Number of Large Retail Establishments (IV)

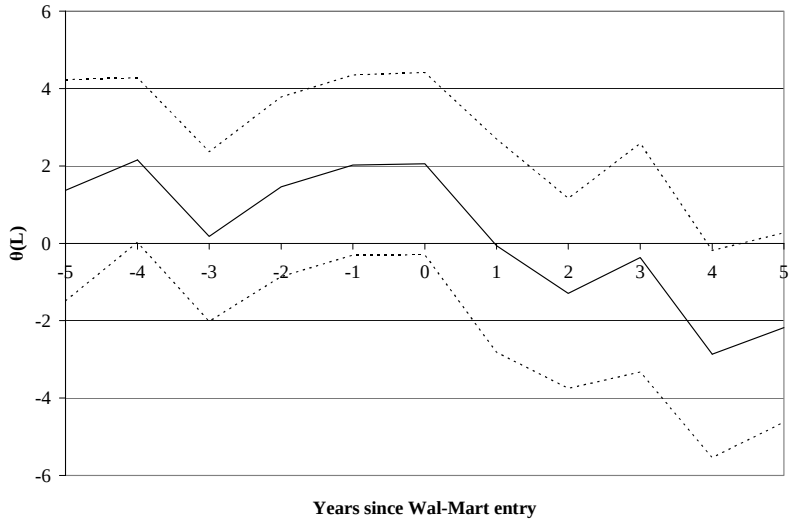

Figure 6: Evolution of Number of Small Retail Estabishments (IV)

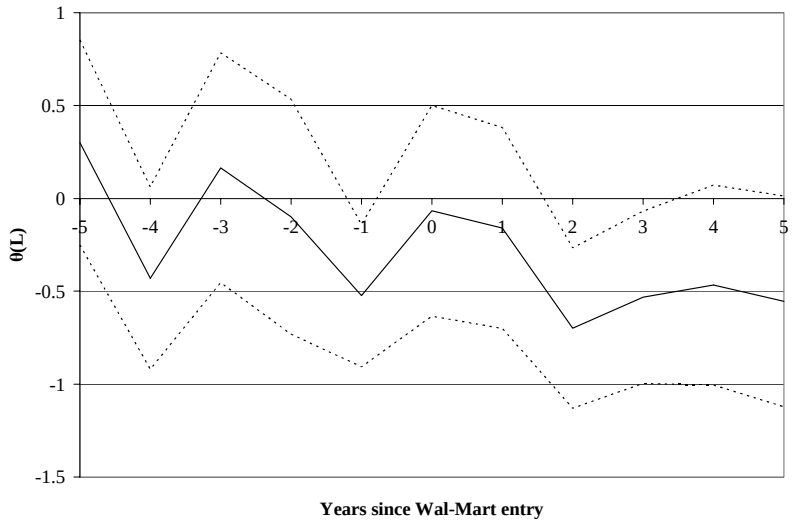

Figure 7: Evolution of Number of Medium Retail Establishments (IV)

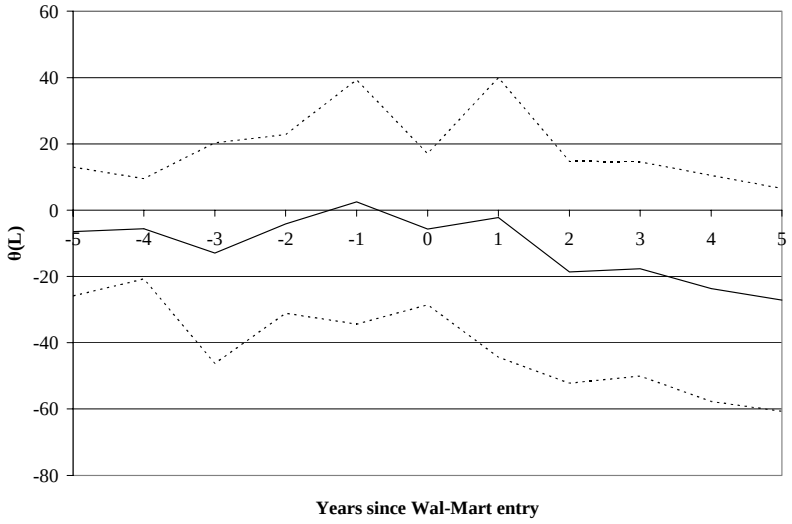

Figure 8: Evolution of Wholesale Employment (IV)

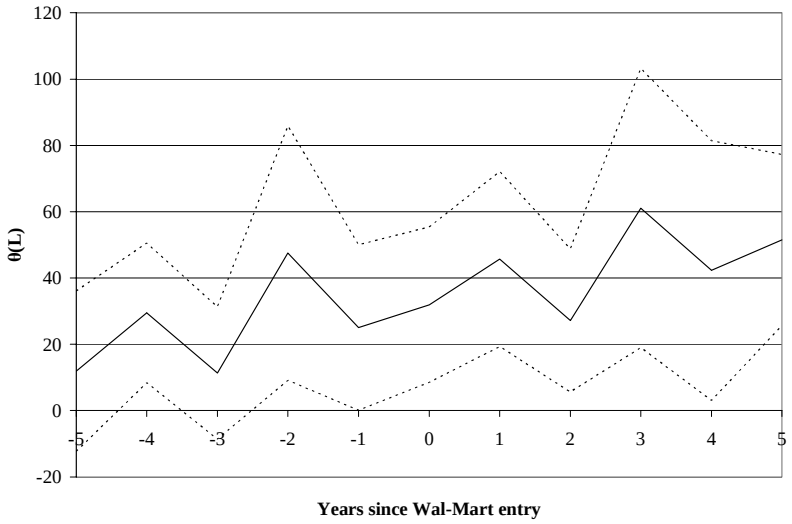

Figure 9: Evolution of Restaurant Employment (IV)

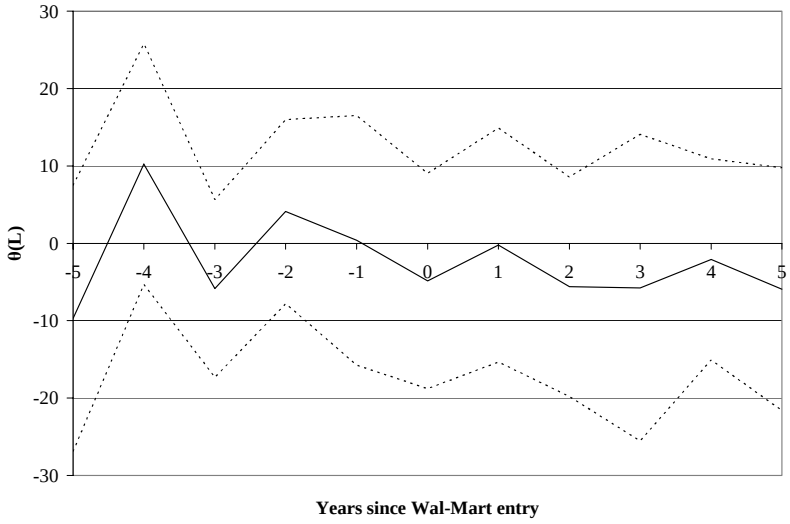

Figure 10: Evolution of Automobile Dealership & Service Station Employment (IV)

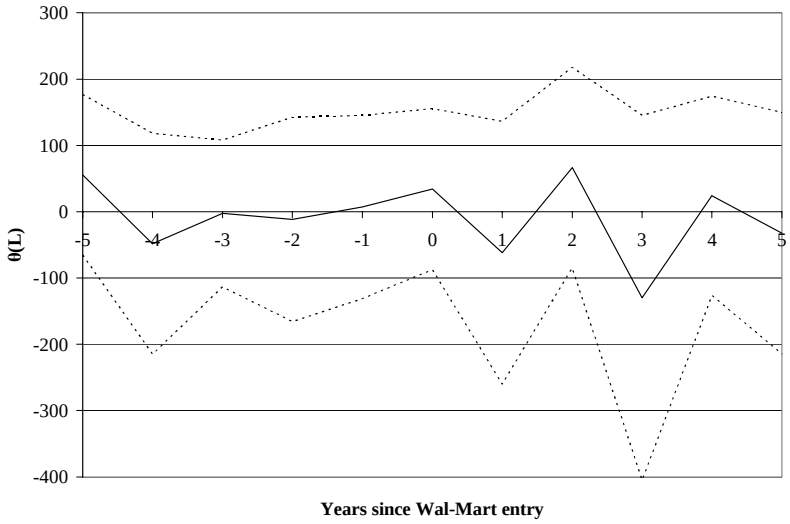

Figure 11: Evolution of Manufacturing Employment (IV)