

# The effects of Wal-Mart on local labor markets <sup>☆</sup>

David Neumark <sup>a,\*</sup>, Junfu Zhang <sup>b</sup>, Stephen Ciccarella <sup>c</sup>

<sup>a</sup> Department of Economics, UCI, 3151 Social Science Plaza, Irvine, CA 92697, USA

<sup>b</sup> Department of Economics, Clark University, 950 Main St., Worcester, MA 01610, USA

<sup>c</sup> Department of Economics, 404 Uris Hall, Cornell University, Ithaca, NY 14853, USA

Received 25 December 2006; revised 8 July 2007

Available online 14 August 2007

## Abstract

We estimate the effects of Wal-Mart stores on county-level retail employment and earnings, accounting for endogeneity of the location and timing of Wal-Mart openings that most likely biases the evidence against finding adverse effects of Wal-Mart stores. We address the endogeneity problem using a natural instrumental variables approach that arises from the geographic and time pattern of the opening of Wal-Mart stores, which slowly spread out from the first stores in Arkansas. The employment results indicate that a Wal-Mart store opening reduces county-level retail employment by about 150 workers, implying that each Wal-Mart worker replaces approximately 1.4 retail workers. This represents a 2.7 percent reduction in average retail employment. The payroll results indicate that Wal-Mart store openings lead to declines in county-level retail earnings of about \$1.4 million, or 1.5 percent. Of course, these effects occurred against a backdrop of rising retail employment, and only imply lower retail employment growth than would have occurred absent the effects of Wal-Mart.

© 2007 Elsevier Inc. All rights reserved.

*JEL classification:* J2; J3; R1

*Keywords:* Wal-Mart; Employment; Earnings

## 1. Introduction

Wal-Mart is more than just another large company. It is the largest corporation in the world, with total revenues of \$285 billion in 2005. It employs over 1.2 mil-

lion workers in the United States, at about 3600 stores.<sup>1</sup> To put this in perspective, the Wal-Mart workforce represents just under 1 percent of total employment and just under 10 percent of retail employment in the United States. It exceeds the number of high school teachers or middle school teachers, and is just under the size of the elementary school teacher workforce. Wal-Mart is reported to be the nation's largest grocer, with a 19 percent market share, and its third-largest pharmacy, with a 16 percent market share (Bianco and Zellner, 2003).

<sup>☆</sup> This is a revised version of a preliminary draft presented at the Wal-Mart Economic Impact Research Conference, Washington, DC, November 2005.

\* Corresponding author.

*E-mail addresses:* [dneumark@uci.edu](mailto:dneumark@uci.edu) (D. Neumark), [juzhang@clarku.edu](mailto:juzhang@clarku.edu) (J. Zhang), [smc227@cornell.edu](mailto:smc227@cornell.edu) (S. Ciccarella).

<sup>1</sup> See <http://www.walmartfacts.com/newsdesk/wal-mart-fact-sheets.aspx#a125> (as of September 8, 2005).

During the past two decades, as Wal-Mart sharply expanded its number of stores in the United States, it increasingly encountered resistance from local communities. Opponents of Wal-Mart have tried to block its entry on many grounds, including the prevention of urban sprawl, preservation of historical culture, protection of the environment and “main-street” merchants, and avoidance of road congestion.<sup>2</sup> Yet two of the most commonly-heard criticisms are that Wal-Mart eliminates more retail jobs than it creates for a community, and that it results in lower wages, especially in retail.<sup>3</sup> Wal-Mart executives dispute these claims, especially with regard to employment. For example: its Vice President Bob McAdam has argued that there are many locations where Wal-Mart creates jobs in other businesses in addition to what Wal-Mart itself offers (PBS, 2004); the Wal-Mart web-site *Walmartfacts.com* trumpets the positive effects of Wal-Mart stores on retail jobs in the communities where stores open<sup>4</sup>; and an advertisement run in the *USA Today*, *The Wall Street Journal*, and *The New York Times* on January 14, 2005, displayed an open letter from Lee Scott, Wal-Mart President and CEO, stating “This year, we plan to create more than 100,000 new jobs in the United States.”<sup>5</sup> Of course Wal-Mart offers other potential benefits in the form of lower prices for consumers (Basker, 2005a; Hausman and Leibtag, 2004).

In this paper, we seek to provide a definitive answer regarding whether Wal-Mart creates or eliminates jobs in the retail sector, relative to what would have happened absent Wal-Mart’s entry. Also, because of concern over the effects of Wal-Mart on wages, and because policymakers may be interested in the impact of Wal-Mart on taxable payrolls, we also estimate the effects of Wal-Mart on earnings in the retail sector, reflecting the combination of influences on employment, wages, and hours. We believe that our evidence improves substantially on existing studies of these and related questions, most importantly by implementing an identification strategy that accounts for the endogeneity of store location and timing and how these may be correlated with future changes in earnings or employment. Indeed, it has been suggested that Wal-Mart’s explicit strat-

egy was to locate in small towns where the population growth was increasing (Slater, 2003, p. 92) and it is reasonable to expect that Wal-Mart entered markets where projected retail growth was strong. If Wal-Mart tends to enter fast-growing areas in booming periods, then we might expect to observe employment and earnings rising in apparent response to Wal-Mart’s entry, even if the stores actually have negative effects.

Our identification strategy is driven by a systematic pattern in the openings of Wal-Mart stores. Sam Walton, the founder of Wal-Mart, opened the first Wal-Mart store in 1962 in Rogers, Arkansas, in Benton County. Five years later, Wal-Mart had 18 stores with \$9 million of annual sales. Wal-Mart first grew into a local chain store in the northwest part of Arkansas. It then spread to adjacent states such as Oklahoma, Missouri, and Louisiana. From there, it kept expanding to the rest of the country after closer markets were largely saturated (Slater, 2003, pp. 28–29). The relationship between Wal-Mart stores’ opening dates and their distance to the headquarters is primarily a result of Wal-Mart’s “saturation” strategy for growth, which was based on control of and distribution of stores, as well as word-of-mouth advertising. In his autobiography, Sam Walton describes the control and distribution motive as follows:

“[Our growth strategy] was to saturate a market area by spreading out, then filling in. In the early growth years of discounting, a lot of national companies with distribution systems already in place—K-Mart, for example—were growing by sticking stores all over the country. Obviously, we couldn’t support anything like that. . . . We figured we had to build our stores so that our distribution centers, or warehouses, could take care of them, but also so those stores could be controlled. We wanted them within reach of our district managers, and of ourselves here in Bentonville, so we could get out there and look after them. Each store had to be within a day’s drive of a distribution center. So we would go as far as we could from a warehouse and put in a store. Then we would fill in the map of that territory, state by state, county seat by county seat, until we had saturated that market area. . . . So for the most part, we just started repeating what worked, stamping out stores cookie-cutter style” (Walton, 1992, pp. 110–111).

One might wonder whether this need to be near a distribution center requires a steady spreading out from Arkansas. Why not, for example, open distribution centers further away, and build stores near them? The explanation seems to lie in the word-of-mouth advertising

<sup>2</sup> See, for example, Bowermaster (1989), Ingold (2004), Jacobs (2004), Kaufman (1999), Nieves (1995), and Rimer (1993).

<sup>3</sup> See, for example, Norman (2004), Quinn (2000), and Wal-Mart Watch (2005).

<sup>4</sup> See <http://www.walmartfacts.com/newsdesk/wal-mart-fact-sheets.aspx> (as of December 15, 2005).

<sup>5</sup> If this refers to gross rather than net job creation, it could be consistent with Wal-Mart destroying more jobs than it creates.

advantage perceived to result from the growth strategy Wal-Mart pursued:

“This saturation strategy had all sorts of benefits beyond control and distribution. From the very beginning, we never believed in spending much money on advertising, and saturation helped us to save a fortune in that department. When you move like we did from town to town in these mostly rural areas, word of mouth gets your message out to customers pretty quickly without much advertising. When we had seventy-five stores in Arkansas, seventy-five in Missouri, eighty in Oklahoma, whatever, people knew who we were, and everybody except the merchants who weren’t discounting looked forward to our coming to their town. By doing it this way, we usually could get by with distributing just one advertising circular a month instead of running a whole lot of newspaper advertising” (Walton, 1992, p. 111).<sup>6</sup>

Wal-Mart’s practice of growing by “spreading out” geographically means that distance from Benton County, Arkansas, and time—and more specifically their interaction—help predict when and where stores opened.<sup>7</sup> Thus, the key innovation in this paper is to instrument for the opening of Wal-Mart stores with interactions between time and the distance between Wal-Mart host counties and Benton County, Arkansas, where Wal-Mart headquarters are located and the first Wal-Mart store opened.

## 2. Literature review

There are a number of studies that address claims about Wal-Mart’s impacts on local labor markets, emphasizing the retail sector. However, we regard much of this literature as uninformative about the causal impact of Wal-Mart on retail employment and earnings. Some of the existing work is by advocates for one side or the other in local political disputes regarding

Wal-Mart’s entry into a particular market. These studies are often hastily prepared, plagued by flawed methods and arbitrary assumptions, and sponsored by interested parties such as Wal-Mart itself, its competitors, or union groups (e.g., Bianchi and Swinney, 2004; Freeman, 2004; and Rodino Associates, 2003), and can hardly be expected to provide impartial evidence on Wal-Mart’s effects. Hence, they are not summarized here.

There is also an academic literature on the impact of Wal-Mart stores, focusing on the effects of Wal-Mart openings on local employment, retail prices and sales, poverty rates, and the concentration of the retailing industry, as well as the impact on existing businesses. This research is limited by three main factors: the restriction of much of it to small regions (often a single small state); its lack of focus on employment and earnings effects; and its failure to account for the endogeneity of Wal-Mart locations, either at all or, in our view, adequately.

Many of these studies, especially the early ones, focus on the effects of Wal-Mart at the regional level, spurred by the expansion of Wal-Mart into a particular region. Most of these studies focus on the effects of Wal-Mart on retail businesses and sales, rather than on employment and earnings. The earliest study, which is typical of much of the research that has followed, is by Stone (1988). He defines the “pull factor” for a specific merchandise category as the ratio of per capita sales in a town to the per capita sales at the state level, and examines the changes in the pull factor for different merchandise categories in host and surrounding towns in Iowa after the opening of Wal-Mart stores. Stone finds that, in host towns, pull factors for total sales and general merchandise (to which all Wal-Mart sales belong) rise after the arrival of Wal-Mart. Pull factors for eating and drinking and home furnishing also go up because Wal-Mart brings in more customers. However, pull factors for grocery, building materials, apparel, and specialty stores decline, presumably due to direct competition from Wal-Mart. He also finds that small towns surrounding Wal-Mart towns suffer a larger loss in total sales compared to towns that are further away.<sup>8</sup> Related results for other regions—which gener-

<sup>6</sup> See Holmes (2005) for an explanation of Wal-Mart’s expansion pattern based on the cost savings of locating stores close to one another. He calls such savings the “economies of density” that are derived from not just the word-of-mouth advertising and shared distribution system. For example, opening new stores near existing ones also makes it cost-effective for Wal-Mart to hire and train employees at existing stores and then transfer them to nearby new stores once they open.

<sup>7</sup> For example, Wal-Mart’s expansion did not reach California until 1990. It first entered New England in 1991. In 1995 Wal-Mart opened its first store in Vermont and finally had a presence in all 48 contiguous states.

<sup>8</sup> Stone’s study was updated regularly (see, for example, Stone, 1995, 1997), but its central message remained the same: Wal-Mart pulls more customers to the host town, hurts its local competitors, but benefits some other local businesses that do not directly compete with it. Using the same methods, Stone et al. (2002) show similar results regarding the effects of Wal-Mart Supercenters on existing businesses in Mississippi.

ally, although not always, point to similar conclusions—are reported in Artz and McConnon (2001), Barnes et al. (1996), Davidson and Rummel (2000), and Keon et al. (1989). All of these studies use administrative data, and employ research designs based on before-and-after comparisons in locations in which Wal-Mart stores did and did not open.<sup>9</sup>

The studies reviewed thus far do not address the potential endogeneity of the location and timing of Wal-Mart's entry into a particular market. In addition, these studies do not focus on the key questions with which this paper is concerned—the effects of Wal-Mart on retail employment and earnings. A few studies come closer to the mark. Ketchum and Hughes (1997), studying counties in Maine, recognize the problem of the endogenous location of Wal-Mart stores in faster-growing regions. They attempt to estimate the effects of Wal-Mart on employment and earnings using a difference-in-difference-in-differences (DDD) estimator that compares changes in retail employment and earnings over time in counties in which Wal-Mart stores did and did not locate, compared to changes for manufacturing and services. However, virtually none of their estimated changes are statistically significant and the data appear very noisy, so the results are generally uninformative. More important, their approach does not address the key endogeneity questions of whether Wal-Mart location decisions were based on anticipated changes after stores opened, or instead only prior trends that were already different (despite the authors posing these questions).<sup>10</sup> Hicks and Wilburn (2001), studying the impact of Wal-Mart openings in West Virginia, estimate positive impacts of Wal-Mart stores on retail employment and the number of retail firms. They do not explicitly

account for endogeneity, although they do address the issue. In particular, they report evidence suggesting that Wal-Mart location decisions are independent of long-term economic growth rates of individual counties in their sample, and that current and lagged growth have no significant effect on Wal-Mart's decision to enter. However, these results do not explicitly address endogeneity with respect to future growth. The latter, in particular, could generate apparent positive impacts of Wal-Mart stores.

In more recent work, Basker (2005b) studies the effects of Wal-Mart on retail employment using nationwide data. Basker attempts to account explicitly for endogeneity by instrumenting for the actual number of stores opening in a county in a given year with the planned number. The latter is based on numbers that Wal-Mart assigns to stores when they are planned; according to Basker, these store numbers indicate the order in which the openings were planned to occur. She then combines these numbers with information from Wal-Mart *Annual Reports* to measure planned and actual openings in each county and year; the data reflect some measurement error in store opening dates.<sup>11</sup> Her results indicate that county-level retail employment grows by about 100 in the year of Wal-Mart entry, but declines to a gain of about 50 jobs in five years as other retail establishments contract or close. In the meantime, possibly because Wal-Mart streamlines its supply chain, wholesale employment declines by 20 jobs in the longer term.<sup>12</sup>

The principal problem with this identification strategy, however, is that the instrument is unconvincing. For the instrument to be valid, two conditions must hold. The first is that planned store openings should be correlated with (predictive of) actual openings; this condition is not problematic. The second condition is that the variation in planned openings generates exogenous variation in actual openings that is uncorrelated with the unobserved determinants of employment that endogenously affect location decisions. This second condition holds if we assume, to quote Basker, that “the number of

<sup>9</sup> A couple of studies rely on surveys of local businesses rather than administrative data. McGee (1996) reports results from a small-scale survey of small retailers in five Nebraska communities conducted soon after Wal-Mart stores entered. He finds that 53 percent of the responding retailers reported negative effects of Wal-Mart's arrival on their revenues while 19 percent indicated positive effects. In a survey of Nebraska and Kansas retailers, Peterson and McGee (2000) find that less than a third of the businesses with at least \$1 million in annual sales reported a negative effect after Wal-Mart's arrival, while close to one half of the businesses with less than \$1 million in annual sales indicated a negative effect, with negative effects most commonly reported by small retailers in central business districts. The research design in these surveys fails to include a control group capturing changes that might have occurred independently of Wal-Mart openings. In addition, reported assessments by retailers may not reflect actual effects of these openings.

<sup>10</sup> The second question can only be addressed via an instrumental variables approach, and the first requires looking at changes in growth rates, not changes in levels; their study only does the latter.

<sup>11</sup> This is detailed in an appendix available from the authors.

<sup>12</sup> Using the same instrumental variables strategy, Basker (2005a) estimates the effects of Wal-Mart entry on prices of consumer goods at the city level, finding long-run declines of 8–13 percent in prices of several products including aspirin, detergent, Kleenex, and toothpaste, although it is less clear to us why location decisions would be endogenous with respect to price. Ordinary least squares (OLS) estimation also finds long-run negative effects of Wal-Mart on prices for nine out of ten products, although only three are significant and all are smaller than the IV results. Hausman and Leibtag (2004) study the effects of Wal-Mart on food prices.

*planned* Wal-Mart stores . . . for county  $j$  and year  $t$  is independent of the error term . . . and *planned* Wal-Mart stores affect retail employment per capita only insofar as they are correlated with the *actual* construction of Wal-Mart stores” (Basker, 2005b, p. 178). The second part of the assumption is potentially problematic; actual stores should, of course, be the driving influence, although planned stores—even if they do not materialize or do so only with delay—may still affect decisions of other businesses.<sup>13</sup> The first part of the assumption is a more serious concern, though, as it seems most likely that planned openings will reflect the same unobserved determinants that drive endogenous location as are reflected in actual openings, and we cannot think of an argument to the contrary (nor does Basker offer one). In this case if the OLS estimate of the effect of Wal-Mart stores on retail employment is biased upward because of endogeneity, the instrumental variable (IV) estimate may also be upward biased, and more so than the OLS estimate.<sup>14</sup>

Using various data sources, Goetz and Swaminathan (2004) study the relationship between Wal-Mart openings between 1987 and 1998 and county poverty rates in 1999, conditional on 1989 poverty rates (as measured in the 1990 and 2000 Censuses of Population). They also use an IV procedure to address the endogeneity of Wal-Mart entry, instrumenting for Wal-Mart openings during 1987–1998 in an equation for county poverty rates in 1999. Their IVs include an unspecified pull factor, access to interstate highways, earnings per worker, per capita property tax, population density, percentage of households with more than three vehicles, and number of female-headed households. The results suggest that county poverty rates increase when Wal-Mart stores open, perhaps because Wal-Mart lowers earnings (although the authors offer other explanations as well). However, why the IVs should affect Wal-Mart openings only, and not changes in poverty directly (conditional on Wal-Mart openings), is not clear, and it is not hard to construct stories in which invalid exclusion restrictions

would create biases towards the finding that Wal-Mart openings increase poverty.<sup>15</sup>

Our research addresses the four principal shortcomings of the existing research on the effects of Wal-Mart on local labor markets. First, we estimate the effects of Wal-Mart openings on retail employment (as well as earnings). Second, we have—we believe—a far more convincing strategy to account for the potential endogeneity of Wal-Mart openings. Third, we are able to use administrative data on Wal-Mart openings that eliminate the measurement error in recent work. And finally, we use a data set that is national in scope.

### 3. Data

The analysis is done at the county level. Aside from following Basker’s strategy, this strikes us as a reasonable geographic level of disaggregation at which to detect the effects of Wal-Mart stores—not so small that many of the effects may occur outside of the geographic unit, and not so large that the effects may be undetectable.

Employment and payroll data are drawn from the US Census Bureau’s County Business Patterns (CBP).<sup>16</sup> CBP is an annual series that provides economic data by industry and county. The series includes most economic activity, but excludes data on self-employed individuals, employees of private households, railroad employees, agricultural production workers, and most government employees. Payroll in the CBP includes salaries, wages, reported tips, commissions, bonuses, vacation allowances, sick-leave pay, employee contributions to qualified pension plans, and taxable fringe benefits, and is reported before deductions for Social Security, in-

<sup>13</sup> A telling example occurred in 2003, after Wal-Mart announced plans to build 40 Supercenters in California. Supermarket chains in southern California including Albertsons, Ralphs, and Vons immediately sought to lower costs and avoid being undercut by Wal-Mart, leading to four months of labor conflict, all before Wal-Mart opened its first Supercenter in California (in March 2004). See, for example, Goldman and Cleeland (2003) and Hiltzik (2004).

<sup>14</sup> See the discussion in the longer working paper version of this study (Neumark et al., 2005). Basker also motivates the IV as correcting for bias from measurement error in the actual opening dates of Wal-Mart stores in her data. But the same argument against the validity of planned openings as an instrument applies.

<sup>15</sup> For example, consider the use of the number of female-headed households as an IV, and suppose that this variable is positively correlated with changes in poverty rates (because of rising inequality over this period), and also positively correlated with Wal-Mart openings (because they locate in lower-income areas). In this case the IV estimate of the effect of Wal-Mart openings on changes in poverty rates is biased upward because of the positive correlation between the instrument (female-headed households) and the error term in the equation for the change in the poverty rate. In addition, although the authors do not specify how they construct their pull factor, we assume it is similar to the measure described above—a ratio of county to statewide retail sales. Given that this is a dependent variable in other studies of the effects of Wal-Mart, it is hard to justify using it as an IV for Wal-Mart openings.

<sup>16</sup> A good description of the CBP data is available at <http://www.census.gov/epcd/cbp/view/cbpview.html> (as of August 15, 2005).

come tax, insurance, etc.<sup>17</sup> It does not include profit or other compensation earned by proprietors or business partners. Payroll is reported on an annual basis. The employment measure in the CBP data is a count of jobs, rather than the number of people employed (in one or more jobs). Employment covers all full- and part-time employees, including officers and executives, as of the pay period including March 12 of each year. Workers on leave are included, while proprietors and partners are not.

The most significant limitation of the CBP data for studying the effects of Wal-Mart is that a wage cannot be computed. The CBP data do not provide a breakdown of employment into full-time and part-time workers, or any information on skill composition. Thus, we cannot tell whether changes in payrolls reflect changes in pay rates for comparable workers or shifts in skill composition or hours. As a consequence, these data cannot be used to address questions of the effects of Wal-Mart on wages. We can, though, estimate Wal-Mart's effect on total retail payrolls.

We use CBP data from 1977 through 2002. We began with 1977 because CBP data are not continuously machine-readable for the years 1964–1976, and ended with 2002 because that was the last year available. As explained below, however, most our analysis goes only through 1995—a period for which our identification strategy is most compelling.

We study the retail sector as a whole, as well as the general merchandising subsector—which includes Wal-Mart and other general department-style stores.<sup>18</sup> Looking at results for these different retail sectors is useful for assessing and understanding the results. For example, if Wal-Mart reduces retail employment, we would expect the employment reduction to show up for the aggregate retail sector, but if it competes mainly with small specialized retail businesses there should also be a sizable increase in general merchandising. Some complications arise in working with the CBP data because by federal law no data can be published that might dis-

close the operations of an individual employer. As we look at more disaggregated subsets of industries, it is more likely that data are suppressed and so our sample becomes smaller. We therefore constructed two samples with which we can consistently compare at least some retail industry sectors for the same set of observations:

- *A sample*: all county-year observations with complete (non-suppressed) employment and payroll data for aggregate retail, and in total.
- *B sample*: all observations in the A sample that also have complete data for the general merchandising retail subsector to which Wal-Mart belongs.

Because the rules for whether or not data are disclosed depend on the size of the retail sector and the size distribution of establishments within it, sample selection is endogenous. We therefore emphasize results for the A sample, which includes nearly all counties and years. We use the B sample only to compare estimates for aggregate retail and general merchandising; as long as any biases from selection into the B sample are similar across retail subsectors, the estimates for aggregate retail and general merchandising can still be meaningfully compared.<sup>19</sup>

Wal-Mart provided us with administrative data on 3066 Wal-Mart Discount Stores and Supercenters. The data set contains every Discount Store and Supercenter still in operation in the United States at the end of fiscal year 2005 (January 31, 2005), as well as the opening date.<sup>20</sup> Because employment in the CBP data is measured as of early March in each year, we code store openings as occurring in the first full calendar year for which they are open. If we instead coded the store opening as occurring in the previous calendar year (some-time during which the store did in fact open), in most cases (unless the store opened before early March) we

<sup>17</sup> Given that it excludes non-taxable fringe benefits, of which the most important is health insurance, we refer to the CBP measure as earnings, rather than compensation.

<sup>18</sup> General merchandising includes retail stores which sell a number of lines of merchandise, such as dry goods, apparel and accessories, furniture and home furnishings, small wares, hardware, and food. The stores included in this group are known as department stores, variety stores, general merchandise stores, catalog showrooms, warehouse clubs, and general stores. (See <http://www.census.gov/epcd/ec97sic/def/G53.TXT>, as of June 13, 2007). Based on 1997 Census data, general merchandising makes up about 2.2 percent of all retail establishments, and 10.7 percent of all retail employment.

<sup>19</sup> Basker (2005b) also focuses on the retail subsector excluding eating and drinking places and automotive dealers and gasoline service centers—subsectors that are least likely to compete directly with Wal-Mart. However, constructing data for this subsector results in nearly two-thirds of county-year pairs being discarded because of non-disclosure, in which case endogenous sample selection can be severe.

<sup>20</sup> Unfortunately, we are often missing information on when Discount Stores converted to Supercenters, which is frequent in the latter part of the sample. All we know is the current store type. We also received data identifying a small number of stores (54, as of 2005) that closed. We return to this issue in some of our robustness analyses. This data set also indicated some store relocations within counties, which are treated as continuing stores because Wal-Mart replaced smaller, older stores with larger ones in nearby locations.

would be using an employment level prior to the store opening as a measure of employment post-opening.<sup>21</sup>

After dropping stores in Alaska and Hawaii, we used 2211 stores in our main analysis through 1995, and 2795 stores when we use the full sample period through 2002. By 2005, Wal-Mart also had 551 Sam's Club stores in the United States (the first opened in 1983), on which we also obtained data, although the data were less complete (for example, lacking information on square footage). We do most of our analysis considering the Wal-Mart stores other than Sam's Clubs, but also some analysis incorporating information on the latter.<sup>22</sup>

We constructed a county-year file by collecting county names and FIPS codes for the 3141 US counties from the US Census Bureau.<sup>23</sup> We created time-consistent geographical areas accounting for merges or splits in counties during the sample period. For counties that split during the sample period we maintained the definition of the original county, and for counties that merged during the sample period we created a single corresponding county throughout.<sup>24</sup> This led to a file of 3094 counties over 19 years (26 years when we use the full sample), to which we merge the CBP and Wal-Mart data. We assigned to the counties population data for each year from the US Census Population Estimates Archives.<sup>25</sup> Finally, we compiled latitude and longitude data for each county centroid from the US Census Bureau's Census 2000 Gazetteer Files.<sup>26</sup> We constructed distance measures from each county to Wal-Mart headquarters in Benton County, Arkansas, for reasons explained below,<sup>27</sup> creating dummy variables for county centroids in rings within a radius of 100 miles from

<sup>21</sup> The measurement error corresponding to stores opening in January through early March of the previous year is less important in the IV estimation than in OLS estimation, because the IV estimation identifies the effect from the predicted probability of store openings rather than actual openings, and these predictions typically will not vary much between adjacent years.

<sup>22</sup> Sam's Clubs are different because customers have to become members, like at Costco. The small number of Wal-Mart Neighborhood Markets is not included in any data we have, but the first one did not open until 1998, beyond the sample period used for most of our analysis.

<sup>23</sup> Downloaded from <http://www.census.gov/datamap/fipslist/AllSt.txt> (as of April 5, 2005).

<sup>24</sup> The code for creating consistent counties over time through 2000 was provided by Emek Basker, and supplemented by us.

<sup>25</sup> Downloaded from <http://www.census.gov/popest/archives/> (as of April 5, 2005).

<sup>26</sup> Downloaded from <http://www.census.gov/tiger/tms/gazetteer/county2k.txt> (as of April 5, 2005).

<sup>27</sup> We use the Haversine distance formula for computing distances on a sphere (Sinnott, 1984).

Benton County, Arkansas, 101–200 miles, etc., out to the maximum radius of 1800 miles (with the vector of these dummy variables denoted *DIST* below).

#### 4. Empirical approach and identification

We estimate models for changes in retail employment and payrolls. We generally capture increased exposure to Wal-Mart stores via a measure of store openings in a county-year cell—i.e., the change in the number of stores. We define changes in employment, payrolls, and number of stores on a per person basis, to eliminate the undue influence of a small number of large employment changes in extraordinarily large counties. As long as we divide all of these changes by the number of persons in the county, the estimated coefficient on the Wal-Mart variable still measures the effect of a Wal-Mart store opening on the change in the level of retail employment or earnings. To control for overall income growth that may affect the level of demand for retail, we include changes in total payrolls per person as a control variable, capturing economic shocks specific to counties and years that could coincidentally be associated with the distance-time interactions that make up the instrumental variable, since these shocks likely affect specific regions in specific periods. In addition, all models include fixed year effects to account for aggregate influences on changes in retail employment or earnings that might be correlated with Wal-Mart openings, which occur with greater frequency later in the sample.

We denote the county-level measures of retail employment and payrolls (per person) as *Y*, the number of Wal-Mart stores (per person) as *WM*, total payrolls per person as *TP*, and year fixed effects (in year *s*) as *YR<sub>s</sub>*. Indexing by county *j* ( $j = 1, \dots, J$ ) and year *t* ( $t = 1, \dots, T$ ), and defining  $\alpha$ ,  $\beta$ ,  $\gamma$ , and  $\delta_s$  as scalar parameters, our baseline model for the change in the dependent variable for each observation *jt* is:

$$\Delta Y_{jt} = \alpha + \beta \Delta WM_{jt} + \gamma \Delta TP_{jt} + \sum_{s=1}^T \delta_s YR_s + \varepsilon_{jt}. \quad (1)$$

Fixed county differences in the levels of the dependent variables drop out of the first-differenced model. However, there may be systematic variation in these first differences across different regions, corresponding to faster or slower growth. We allow for this in a highly flexible manner by also estimating the first difference models including county fixed effects (*CO*), which al-

lows for a different linear trend for each county, so the model becomes

$$\Delta Y_{jt} = \alpha + \beta \Delta WM_{jt} + \gamma \Delta TP_{jt} + \sum_{s=1}^T \delta_s YR_s + \sum_{i=1}^J \phi_i CO_i + \varepsilon_{jt}, \quad (1')$$

where the  $\phi_i$  are also scalar parameters.

#### 4.1. Endogeneity of Wal-Mart location decisions and identification

Consistent estimation of Eqs. (1) or (1') requires that  $\varepsilon_{jt}$  is uncorrelated with the right-hand-side variables. If Wal-Mart location decisions are based in part on contemporaneous and future changes in employment or payrolls, then this condition could be violated. This endogeneity is natural, since Wal-Mart would be expected to make location decisions (including the location and timing of store openings) based on current conditions and future prospects, which might be related to both employment and payroll. As but one example, Wal-Mart may open stores where real estate development and zoning have recently become favorable to retail growth.

Our identification strategy in light of this potential endogeneity is based on the geographic pattern of Wal-Mart store openings over time. Fig. 1 illustrates how Wal-Mart stores spread out geographically throughout the United States, beginning in Arkansas as of 1965, expanding to Oklahoma, Missouri, and Louisiana by 1970, Tennessee, Kansas, Texas, and Mississippi by 1975, much of the South and the lower Midwest by 1985, more of the Southeastern seaboard, the plains, and the upper Midwest by 1990, and then, in turn, the Northeast, West Coast, and Pacific Northwest by 1995. After 1995, when the far corners of the country had been entered, there was only filling in of stores in areas that already had them. This pattern is consistent with Wal-Mart's growth strategy, discussed in the Introduction.

This pattern of growth generates an exogenous source of variation in the location and timing of Wal-Mart store openings that provides natural instrumental variables for Wal-Mart store openings.<sup>28</sup> In particular, time and distance from Benton County, Arkansas

predict where and when Wal-Mart stores will open. However, this does not necessarily imply that time and distance can serve as instrumental variables for exposure to Wal-Mart stores. If we posit an equation for Wal-Mart openings of the form

$$\Delta WM_{jt} = \kappa + \pi \Delta TP_{jt} + \sum_{i=1}^J \lambda_i DIST_i + \sum_{s=1}^T \mu_s YR_s + \eta_{jt}, \quad (2)$$

then because Eq. (1') already includes year and county fixed effects, and the latter capture all the variation in distance, the time and distance variables in Eq. (2) give us no identifying information.

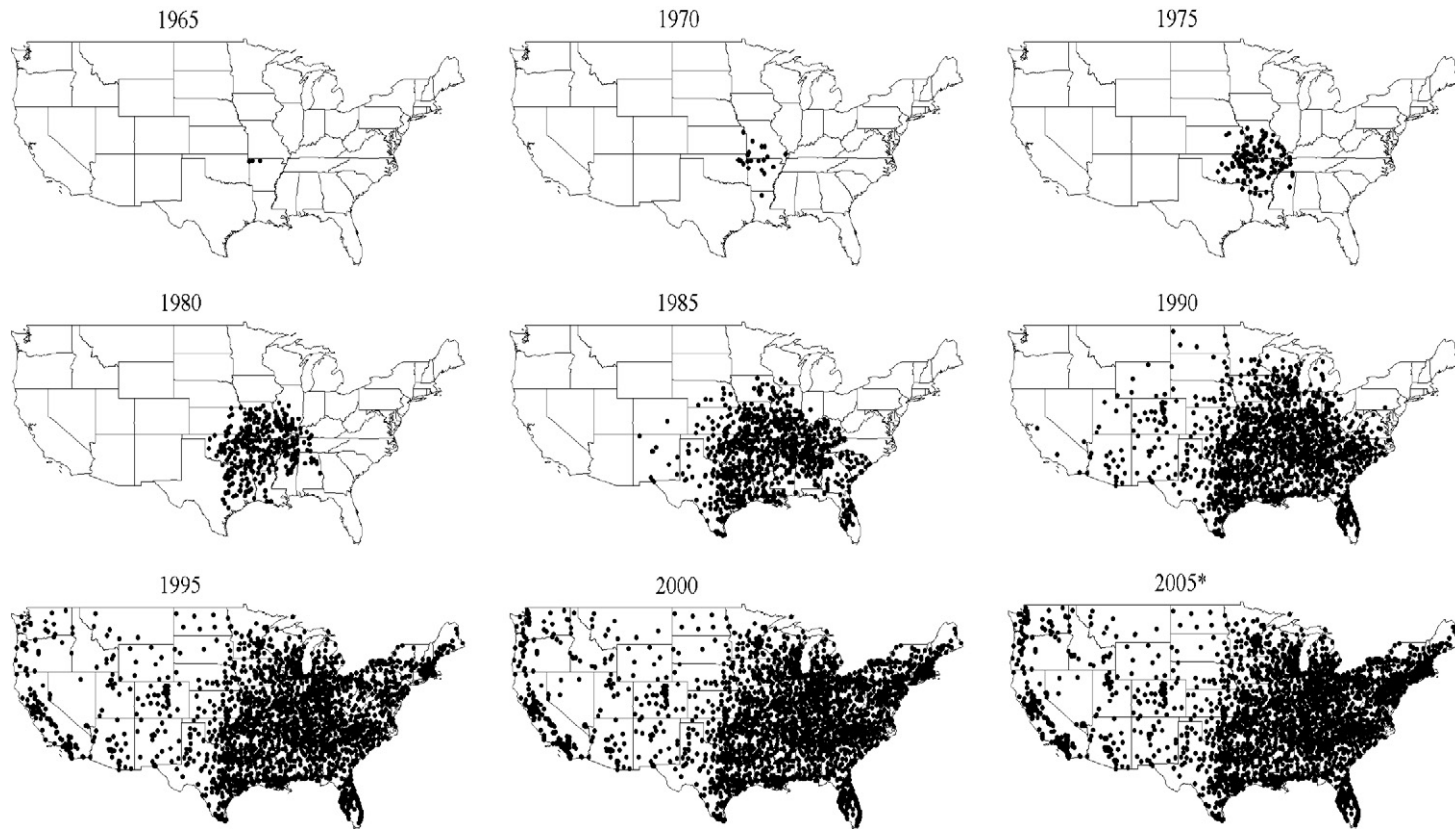
However, Eq. (2) does not accurately capture the geographic and time-related pattern of Wal-Mart openings. In particular, the additivity of the distance and year effects in Eq. (2) implies that differences across years in the probability of Wal-Mart openings are independent of distance from Benton County. But as shown in Fig. 2, which depicts openings only, in the area near Arkansas openings are concentrated in the earliest years, whereas in the 1981–1985 period, for example, openings are more concentrated further away from Arkansas. This pattern becomes more obvious in the 1986–1990 and 1991–1995 maps, where openings thin considerably in the area of Wal-Mart's original growth, and are more common first in Florida, then in the Southeast and the lower Midwest, and finally in California, the upper Midwest, and the Northeast.

The fact that the rate of openings slows considerably in the Southeast, for example, in the later years, and increases in areas further away—in a rough sense spreading out from Benton County like a wave (albeit irregular)—contradicts the additivity of distance and time effects in Eq. (2), and instead implies that the model for Wal-Mart openings should have distance-time interactions, with the probability of openings higher early in the sample period in locations near Benton County, but higher later in the sample period further away from Benton County. Because this relationship holds through 1995, when Wal-Mart had begun to saturate border areas, we restrict most of our analysis to

<sup>28</sup> Jia (2006) addresses the endogeneity of Wal-Mart store locations in a more structural fashion. In particular, she studies the entry of chain stores such as Wal-Mart and K-Mart into local markets and their effects on the numbers of small retailers (rather than retail employment) as a two-stage game in which Wal-Mart and K-Mart first enter

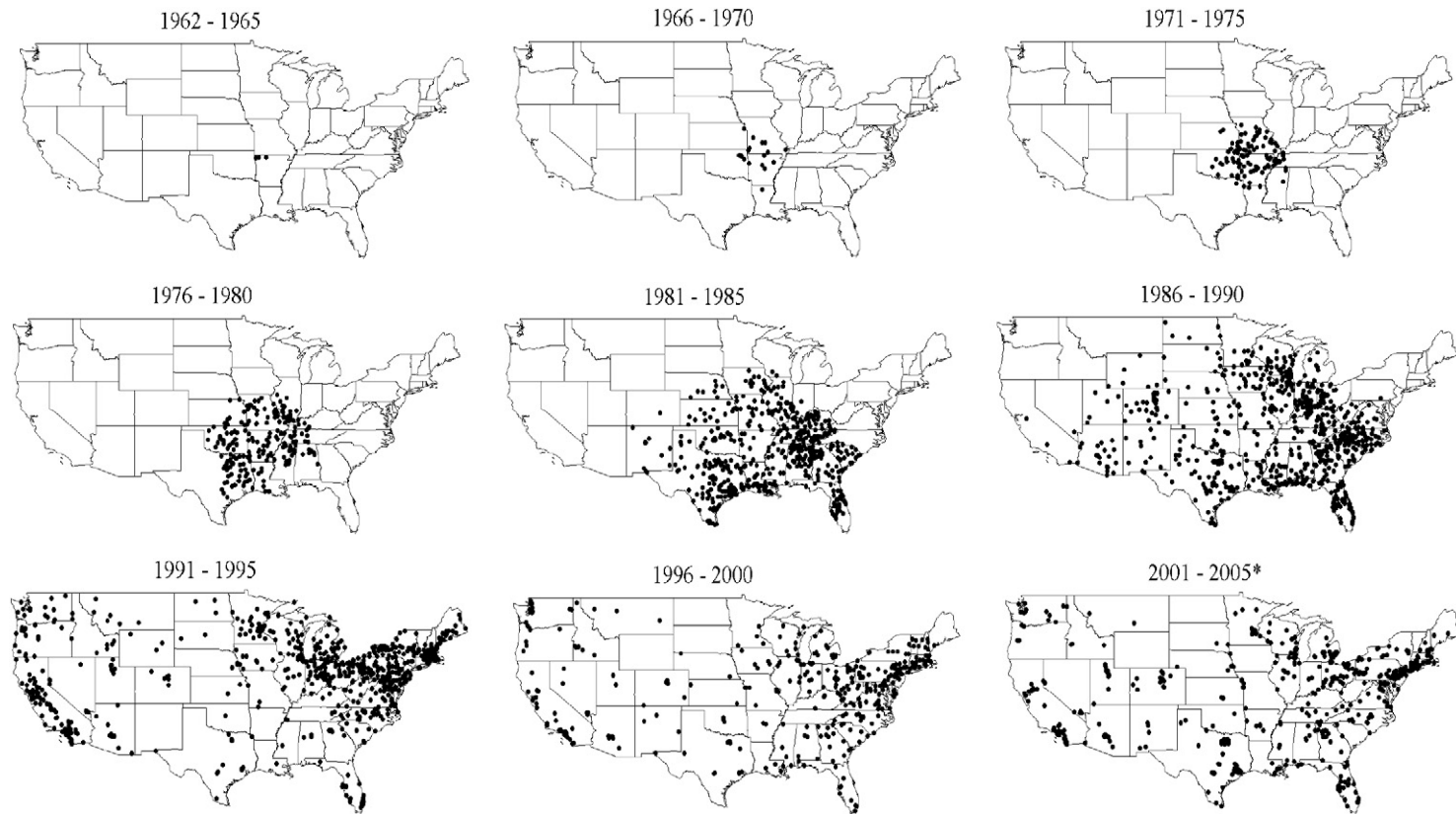
local markets, and in the second stage small retailers decide whether to enter. Her estimates suggest that during the 1988–1997 sample period Wal-Mart's expansion is responsible for half to three quarters of the decline in the number of small retailers. In addition, she shows that estimating a reduced-form model and ignoring the endogeneity of entry decisions could underestimate the adverse effects of Wal-Mart on small retailers by 50–60 percent.





\*Includes all locations as of January 31, 2005.

Fig. 1. Location of Wal-Mart stores, 1965–2005.



\*Includes all openings as of January 31, 2005.

Fig. 2. Location of Wal-Mart openings, 1962–2005.

this period, although we also report results using the full sample through 2002. The most flexible form of this interaction, expanding on Eq. (2), includes interactions between dummy variables for year and for the different distance ranges, as in

$$\Delta WM_{jt} = \kappa + \pi \Delta TP_{jt} + \sum_{i=1}^J \lambda_i DIST_i + \sum_{s=1}^T \mu_s YR_s + \sum_{s=1}^T \sum_{i=1}^J \vartheta_{is} (DIST_i \times YR_s) + \eta_{jt}. \quad (3)$$

Given this specification, the endogenous effect of Wal-Mart openings is identified in Eq. (1') by using the distance-time interactions as instruments for exposure to Wal-Mart stores.<sup>29,30</sup> Throughout, we report standard errors that cluster on state and year, which are robust to heteroskedasticity of the error across state-year cells, and to spatial autocorrelations across counties within states.<sup>31</sup>

<sup>29</sup> Soon after our research was completed, we discovered recent research done concurrently by Dube et al. (2005), which also exploits the geographic pattern of Wal-Mart openings to identify their effect on retail earnings growth. They focus only on earnings, and restricts the analysis to the 1992–2000 period. The discussion of the maps in Figs. 1 and 2, which suggest that Wal-Mart's "fanning out" growth strategy had run its course after 1995 or so, raises concerns about identification of the effects of Wal-Mart stores using data only from 1992–2000. In addition, Dube et al. estimate the models in levels—i.e., retail earnings growth on the number of stores (both on a per person basis)—instrumenting for the number of stores with distance-time interactions. However, as Figs. 1 and 2 make clear, it is openings that are predicted by the distance-time interaction, not the total number of stores; for example, the number of stores remains high near Benton County late in the sample period. As a consequence, this model is misspecified. If the distance-time interactions appear in the equation for openings (or the first difference for the number of stores), the equation in levels would have to have these interactions multiplied by a linear time trend, which would capture the fact that the number of stores *grew faster*, for example, at distances far from Benton County late in the sample period. An even more recent paper is by Drewianka and Johnson (n.d.), who echo our concern about the IV identification strategy in the more recent period they study (n.d.). They instead simply control for county-specific trends. However, the fact that the IV strategy we use probably makes little sense in part of the period they study does not negate the endogeneity problem, which they do not address.

<sup>30</sup> See a longer working paper version of this paper (Neumark et al., 2005) for a simple theoretical model that gives more rigorous justification for our empirical specifications.

<sup>31</sup> Below, we also report estimates clustering on state only, which also allows temporal autocorrelation. The standard errors are only slightly larger and the qualitative conclusions are unchanged. One problem with this level of clustering is that given the large set of instrumental variables we have, the implicit number of degrees of freedom falls below the number of exclusion restrictions, so we cannot calculate an *F*-statistic for the exclusion restrictions in the first stage.

A few comments regarding identification of the model are in order. First, although there are, technically speaking, more instruments than are needed to identify the model, overidentifying tests are in our context uninformative about the validity of the instruments. Conceptually, there is one instrument—the interaction of distance and time—and all we do here is to use a flexible form of this variable. Were there a good reason—based, for example, on location strategies gleaned from Sam Walton's writings—to think that there were particularly exogenous location decisions in a narrower time period or at a specific distance from Benton County, then arguably one might want to use this information to identify the model, and test for the validity of distance-time interactions from other periods or further distances using overidentification tests. Walton's writings, however, bolstered by the maps in Fig. 2, suggest that distance and time acted in a similar fashion throughout the entire period during which Wal-Mart stores spread to the borders of the United States. Thus, the entire set of instruments is either jointly valid, based on a priori arguments, or it is not.<sup>32</sup>

Second, in addition to the distance-time interactions predicting store openings, we also must be able to exclude distance-time interactions from the models for changes in retail employment and payrolls. Given that points at a given distance from Benton County are located on a circle with Benton County at its center, there is no obvious reason why this condition would be expected to be violated. A particular area—say, 500 miles straight east from Benton County—may have economic conditions or structure that differ from those in Benton County and hence also have systematically different trends (or changes) in retail employment or earnings. But this does not imply that we should expect similar influences at the same time at points that are a common distance from Benton County, aside from the effects of Wal-Mart in these areas. Indeed this seems likely to be *particularly* true for retail, which should exhibit little or no regional concentration, but instead be roughly the same across regions with similar levels of income. In contrast, this is less likely to hold for other industries, which—because they need not be located near their customers—can be regionally concentrated; manufacturing is the prime example, but the same is true of many other industries. Nonetheless,

<sup>32</sup> More generally, suppose there is a continuous instrumental variable *Z* that is argued to be valid on a priori grounds. One could not create a set of dummy variables corresponding to each value of *Z* and do an overidentification test to establish the validity of *Z* as an instrument.

the model with fixed county effects allows for different trends by county over the sample period, and thus should largely capture differential changes in economic conditions across regions. Hence, even when looking at retail, we have more confidence in the exclusion restrictions for the validity of the instruments when the county fixed effects are included. Moreover, for this specification there is less reason to be skeptical of using this framework to study employment changes in other industries, which can serve as a falsification test; in particular, we estimate the model for manufacturing employment, which should not be affected, locally, by the opening of Wal-Mart stores.<sup>33</sup>

Third, one might view the pattern of roll-out of Wal-Mart stores as indicating that store openings are exogenous and therefore it is not necessary to correct for endogeneity. However, while the overall pattern and timing appear to be exogenous, the exact counties in which stores locate are not. That is, Wal-Mart's decision to enter an approximate "ring" around Benton County, Arkansas, in a particular period may be exogenous, but where in the ring they enter is not, and it is this endogeneity with which we are concerned, and for which our instrument corrects.<sup>34</sup>

## 5. Empirical results

### 5.1. Descriptive statistics

Descriptive statistics for population, employment, and payroll are reported in Table 1. Most of the statistics are for the A sample, with the exception of those for general merchandising. Statistics for the B sample were

<sup>33</sup> On the other hand, we cannot completely rule out some correlation of retail activity with county- and year-specific shocks. Of course the most flexible model is the fully saturated one that includes county-specific fixed effects as well as their interactions with the year fixed effects in the equations we estimate, allowing for arbitrary differences in the changes in the dependent variables by county and year. In this case, of course, the distance-time interactions would provide no identifying information. But even in the absence of attempts to account for endogeneity, a model with unrestricted time effects for each county would not permit the identification of the effects of Wal-Mart stores. Thus, a more restricted version of how distance-time interactions enter the employment and earnings equations would have to be imposed whether or not we were addressing the endogeneity of Wal-Mart openings.

<sup>34</sup> Khanna and Tice (2000, 2001) have also noted the geographic pattern of the spread of Wal-Mart stores that we exploit. Based on this pattern, they study the effect of Wal-Mart entry on incumbent firms, treating Wal-Mart entry as exogenous. They use somewhat larger areas than we do (based on the first three digits of zip codes), but even so, decisions as to which of these to enter may be endogenous.

Table 1

County-level summary statistics, population, employment, and payroll, 1977–1995

	Means (1)	Medians (2)
Population	78,223 (256,595)	22,644
Number of counties/share of total	3032 0.98	...
Aggregate retail employment ( $N = 57,964$ )	5659 (19,758)	1105
Retail payrolls (\$1000s) ( $N = 57,964$ )	92,208 (364,022)	14,990
General merchandising employment (B sample, $N = 37,999$ )	966 (2691)	195
General merchandising payrolls (\$1000s) (B sample, $N = 37,999$ )	14,170 (43,099)	2477

Figures are for full (A) sample except for general merchandising. In column (1) standard deviations are reported in parentheses. Payrolls are in thousands of 1999 constant dollars.

similar, although counties in the B sample are larger since data are less likely to be suppressed for larger counties and counties with larger retail sectors. On average, counties have approximately 78,000 residents. The data cover 3032 counties, 98 percent of the total. The average level of aggregate retail employment is 5659, or about 5.5 percent of total employment. Employment in general merchandising is about one-sixth of the retail total, or 966 workers on average. Average aggregate retail payrolls are around \$92 million, and general merchandising payrolls are about \$14 million. Although not reported in the table, average retail payrolls per worker across counties is around \$13,700, and the figure for general merchandising is around \$13,100.

Table 2 provides some descriptive statistics on Wal-Mart stores, as of 1995. As indicated in the first row, for counties with one or more stores open the average number of stores is 1.43. The second row indicates that just over half of counties have at least one Wal-Mart store. The remaining rows of the table report the distribution of number of stores per county (for counties with one or more stores). Around 78 percent of the counties with Wal-Mart stores have only one store, about 13 percent have two stores, and around 4 percent have three stores. There is then a smattering of observations with more stores (with a maximum of 17 stores at the end of the sample period in Harris County, Texas, which includes Houston).

### 5.2. Preliminary evidence on endogeneity bias

It is useful to ask whether we can detect evidence of endogeneity, and infer something about the direction of endogeneity bias, without relying explicitly on the

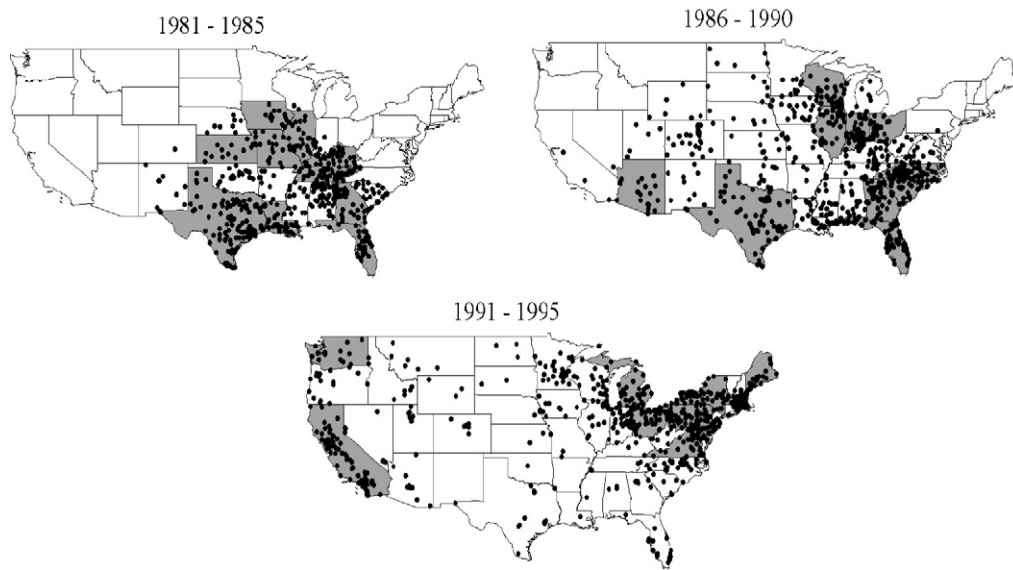


Fig. 3. Geographic concentration of Wal-Mart openings by period and geographic area.

Table 2

County-level summary statistics, Wal-Mart stores, 1995

Average number of stores, for counties with stores	1.43
Share with one or more stores open, all counties	0.505
Share with given number of stores, for counties with stores	
1 store	0.783
2 stores	0.128
3 stores	0.042
4 stores	0.017
5 stores	0.013
6 stores	0.006
7 stores	0.004
8 stores	0.003
9 stores	0.002
10 stores	0.001
11 stores	0.001
12 stores	< 0.001
13 stores	< 0.001
14 stores	< 0.001
15 stores	< 0.001
16 stores	< 0.001
17 stores	< 0.001
Number of counties	3064

Numbers are for full (A) sample of observations with aggregate retail employment data in 1995.

identifying assumptions underlying the IV estimation. One way to get some indirect evidence on endogeneity is to look at the relationship between past growth in retail employment and decisions to open Wal-Mart stores. As Figs. 1 and 2 indicate, the correct comparison for any period is not between all counties where stores did and did not open. Rather, for a particular period we should identify the region in which many stores were

opening, and ask whether—within that region—stores were opening in fast-growing counties.<sup>35</sup>

We did this comparison for different selected periods and regions to give a sense of what this calculation yields throughout the sample. We first focused on store openings in the early 1980s, in states close to Arkansas, restricting attention to openings in the 1983–1985 period. We computed openings per state, and chose the ten states with the highest number of openings. These ten states are shaded in the first map in Fig. 3, which clearly shows that these are states in relatively close proximity to Arkansas (as well as Florida). The map also shows the same openings, in the 1981–1985 period, which were displayed in Fig. 2. The overlaying of these openings and the shaded states makes clear that these were the states in which openings were concentrated in this period. For each county in these ten states, we computed the annualized rate of growth of retail employment over the immediately preceding five-year period 1977–1982.<sup>36</sup> We then estimated linear probability models for whether Wal-Mart entered a county (opening its first store, which, as Table 2 shows, would be the case for most openings), as a function of the prior growth rate of retail employment. We add controls for county population as well as state dummy variables, to account for the effects of population density as well

<sup>35</sup> A recent study commissioned by Wal-Mart (Global Insight, 2005) instead does the first comparison, and erroneously concludes that there is no evidence of endogeneity bias.

<sup>36</sup> This window is arbitrary, but follows the calculation in the Global Insight study (Global Insight, 2005).

Table 3  
Estimated relationships between store openings and prior retail employment growth

Openings in:	1983– 1985	1988– 1990	1993– 1995	1993–1995, Excluding CA and WA
	(1)	(2)	(3)	(4)
Annualized rate of growth of retail employment in previous five-year period	0.439*** (0.143)	0.191 (0.177)	0.122 (0.515)	0.407 (0.613)
<i>N</i>	854	615	363	299

Estimates are from linear probability models, with standard errors robust to heteroskedasticity. The states covered in each sample period are shown in Fig. 3. The previous five-year periods corresponding to the three columns are: 1977–1982, 1982–1987, and 1987–1992. County population in the ending year of these five-year periods, and state dummy variables, are included as controls. Growth rates are computed relative to the average at the beginning and end of each period, because there are a handful of observations with retail employment of zero. Growth rates are measured on an annualized basis. Thus, for example, the estimate in column (1) implies that a 1 percent higher annual growth rate boosts the probability of a store opening by 0.0044; this is a 2.7 percent increase in the probability of a store opening.

\*\*\* Significant at the 1% level.

as other unspecified features of states on store openings. We then did the same for the 1988–1990 period, using retail employment growth from 1982–1987, and the 1993–1995 period, using retail employment growth from 1987–1992. The corresponding maps are shown in the second and third panels of Fig. 3.

The results, reported in Table 3, indicate that Wal-Mart stores entered counties—among an appropriate comparison group—that had previously had faster retail employment growth. For the 1993–1995 estimation, the set of ten states with the most openings includes California and Washington, which are quite far geographically from the other eight states and therefore the comparison to counties in those other states may not be appropriate; we therefore also show results excluding these two states. As the table shows, in every case the relationship between prior growth and whether a store opened is positive (and sometimes significant, although that is not of foremost interest here). This evidence does not directly address the issue of endogeneity bias—which concerns the link between store openings and future retail employment absent store openings. But it does suggest that endogeneity is a concern.

### 5.3. First-stage estimates

Turning to our main analysis, Fig. 4 displays information on the estimated coefficients of the distance-time

interactions from the first stage, although for ease of interpretability we present estimates for the change in the number of stores rather than the number of stores per person. The omitted categories are 900–1000 miles and 1988. The first-stage regression includes 272 distance-time interactions (between 17 year dummy variables and 16 distance dummy variables). In Fig. 4, we break these up into the 16 distance groups, for each distance range around Benton County.

Starting with distances near Benton County, Arkansas, we see that store openings were more likely early in the sample period, and less likely later on. Conversely, at distances far from Benton County store openings were much more common later in the sample period. Alternatively, store openings were much more likely at nearby distances early in the sample, and at farther distances later in the sample. Thus, these regressions confirm the impression from the maps in Fig. 2 that store openings spread out over the sample period, concentrated near Benton County early in the sample period and farther later.

### 5.4. Effects on retail sector employment

We now turn to the OLS and IV estimates. The OLS estimates identify the effect of Wal-Mart simply from differences in outcomes in counties and years where stores opened compared to those where they did not. In contrast, the IV estimates identify the effect from differences in outcomes in counties and years with a high versus a low predicted probability of store openings, as depicted in Fig. 4. Thus, for example, a negative effect of Wal-Mart openings on retail employment will be inferred when retail employment fell (or grew more slowly) in county-year pairs that were in geographic regions (defined by distance from Benton County, Arkansas) and years in which store openings were more likely, without reference to the actual counties within these regions in which stores opened.

Table 4 reports estimates for retail employment. The dependent variable is the change in retail employment per person at the county level, and the Wal-Mart measure is the change in the number of stores per person. We report estimates both excluding and including county-specific time trends (Eqs. (1) and (1')); the latter is our preferred specification. The estimates for retail employment (and earnings) turn out to be insensitive to the inclusion of the county fixed effects, which bolsters the validity of the assumption that the instrumental variables are not correlated with the error terms.

The OLS estimates point to increases of 40–55 workers in both aggregate retail and general merchandising

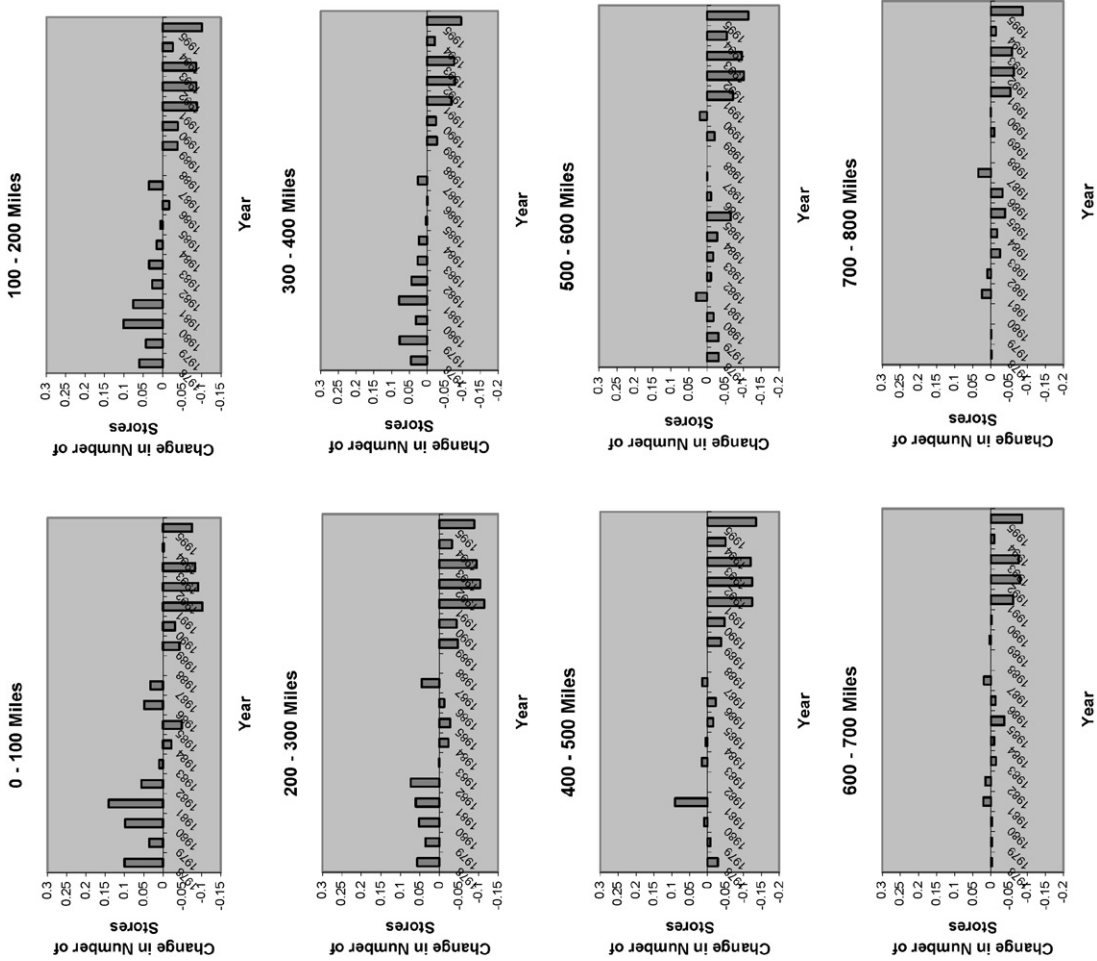
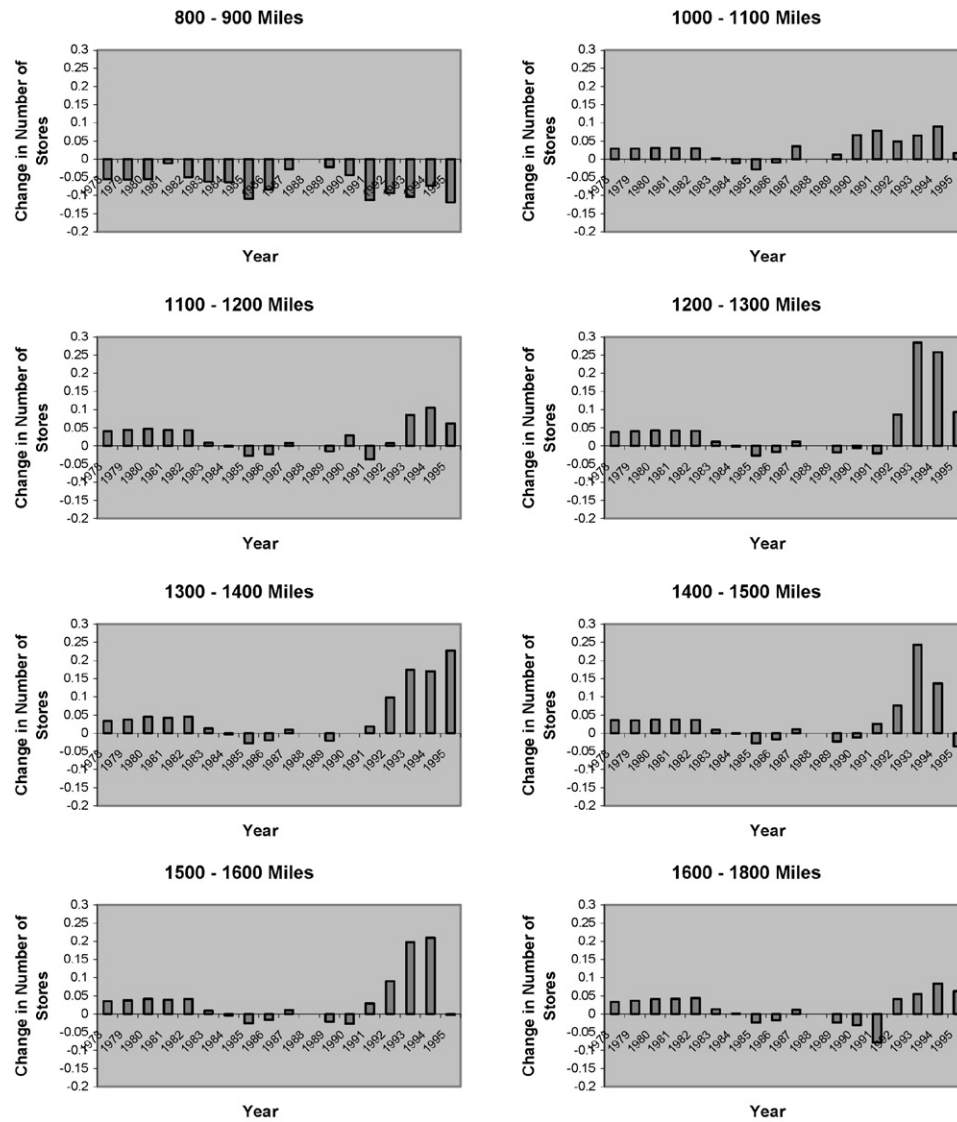


Fig. 4. First-stage results for change in number of stores.



Panels show estimated coefficients of distance-year interactions from estimates of Eq. (3), although estimated for the change in the number of stores rather than the change in the number of stores per person. The omitted reference categories are 1988 and 900–1000 miles.

Fig. 4. (continued)



Table 4  
Estimated effects of Wal-Mart stores on retail employment

	County-specific time trends excluded (Eq. (1))			County-specific time trends included (Eq. (1'))		
	Agg. Retail (1)	Agg. Retail (2)	Gen. Merch. (3)	Agg. Retail (4)	Agg. Retail (5)	Gen. Merch. (6)
	A Sample	B Sample	B Sample	A Sample	B Sample	B Sample
<i>OLS</i>						
Wal-Mart stores	44.52*** (6.29)	42.33*** (9.11)	54.74*** (3.94)	40.01*** (6.02)	42.75*** (9.61)	53.81*** (3.99)
Total payrolls per person	0.001*** (0.0002)	0.001*** (0.0003)	0.0001*** (0.00003)	0.001*** (0.0003)	0.001*** (0.0003)	0.0001*** (0.00003)
Adjusted $R^2$	0.084	0.109	0.081	0.068	0.089	0.038
<i>IV for Wal-Mart stores</i>						
Wal-Mart stores	-166.98*** (48.83)	-134.48** (54.26)	29.36** (11.80)	-146.35*** (47.77)	-105.80* (56.69)	17.92 (14.20)
Total payrolls per person	0.001*** (0.0002)	0.001*** (0.0004)	0.0001** (0.00003)	0.001*** (0.0003)	0.001*** (0.0003)	0.0001*** (0.00003)
<i>F</i> -statistics	12.8	6.8	6.8	4.2	3.9	3.9
Reject null of no endogeneity bias at 5% level	Yes	Yes	Yes	Yes	Yes	Yes
<i>N</i>	54,554	32,668	32,668	54,554	32,668	32,668

The Wal-Mart variable is the change in the number of stores per person, and the dependent variable is the change in employment per person. The coefficients measure the effect of one Wal-Mart store opening on county-level retail employment. The instrumental variables are interactions between dummy variables for years and dummy variables for counties with centroids with a radius of 100 miles from Benton County, Arkansas, 101–200 miles, 201–300 miles, etc., out to 1800 miles. All of the radii except the last cover a 100 mile range; the last covers 1601–1800 miles, because there is only one county beyond the 1700 mile radius. All specifications include year fixed effects, and the specifications in columns (4)–(6) also include county dummy variables (in both the first and second stages). Standard errors (shown in parentheses below the estimates) and *F*-statistics are calculated clustering on state and year, and hence allow arbitrary contemporaneous correlations across counties in a state. The sample period covers 1977–1995. See notes to Table 1 for additional details. The results for the test of the null hypothesis of no endogeneity bias are based on bias-corrected bootstrapped empirical distributions for the difference between the OLS and IV estimates, with the bootstrapping based on states and years rather than individual observations (the level of clustering used in the standard error calculations); this is preferable to the conventional Hausman test (Hausman, 1987) because OLS is inefficient if there is either heteroskedasticity across counties or autocorrelation within counties.

\* Significant at the 10% level.

\*\* Idem., 5%.

\*\*\* Idem., 1%.

employment from a Wal-Mart store opening.<sup>37</sup> For the A sample, the estimates imply an increase in county-level aggregate retail employment of about 0.8 percent at the mean. The estimate for general merchandising is slightly greater than the estimate for aggregate retail, for the B sample for which these estimates are comparable, suggesting that there is at most a small reduction in employment in the rest of the retail sector; however, the difference is not significant. Finally, as we would expect, the estimated coefficient of total payrolls per person is positive.

In contrast to the OLS estimates, the IV estimates—which conditional on the identification strategy being valid are interpretable as causal effects of Wal-Mart openings on retail employment—point to employ-

ment declines in the aggregate retail sector.<sup>38</sup> Without county-specific trends, the estimates for the A sample indicate that a Wal-Mart store opening reduces employment at the county level by about 167 workers. With county-specific trends the estimate falls to 146. In the B sample the estimates are smaller by about 30. Since the average number of workers in a Wal-Mart store is about 360 (Basker, 2005b), the estimated employment decline (using a figure of 150, close to the Eq. (1') estimate) implies that each Wal-Mart worker takes the place of 1.4 retail workers. On a county basis, this estimate implies

<sup>37</sup> The retail estimates are similar to the long-run OLS estimates reported by Basker (2005b).

<sup>38</sup> The *F*-statistics for the joint significance in the first-stage regression of the distance-time interactions that serve as instruments range from 3.9 to 12.8 across the different samples and specifications in Table 4. These *F*-statistics are sometimes small enough that asymptotic results may not apply well (Staiger and Stock, 1997). However, in such cases bias in the IV estimates is in the direction of the OLS bias (Chao and Swanson, 2003), and therefore such bias cannot account for the differences between our IV and OLS results.

a 2.7 percent reduction in retail employment attributable to a Wal-Mart store opening. This could reflect a combination of closings of other retail establishments and employment reductions at those establishments. And of course these estimates do not imply absolute declines in retail employment, but only that retail employment was lower than it would have been had Wal-Mart stores not opened.

The estimated magnitude strikes us as plausible. Government data indicate that sales per employee are much higher at large retailers than at small retailers,<sup>39</sup> and a McKinsey study (Johnson, 2002) suggests much higher sales per employee at Wal-Mart than at other large retailers (in 1995, 70 percent higher than Sears, and 36 percent higher than K-Mart). In addition, a negative employment effect would be expected if Wal-Mart is indeed more efficient and can be profitable while charging significantly lower prices.<sup>40</sup> Finally, Wal-Mart is currently about 1 percent of the workforce, but it was smaller during most of sample period, so this displacement figure does not imply large-scale employment declines; as noted above, it implies that a store opening results in a 2.7 percent decline in retail employment.

For general merchandising, the IV estimates continue to point to increases in employment (sometimes significant), qualitatively similar to the OLS estimates although smaller. We might expect increases in general merchandising employment since this is the sector in which Wal-Mart is classified. However, the evidence that the estimated increase in general merchandising employment is much less than the size of an average Wal-Mart store suggests that Wal-Mart reduces employment at other employers in the general merchandising sector as well as in the remainder of the retail sector.<sup>41</sup> Thus, the estimates are consistent with Wal-Mart

reducing overall retail employment, although shifting its composition toward general merchandising; recall, though, that the evidence for general merchandising (and the B sample overall) may be tainted by endogenous sample selection because of suppressed data.

The finding that the OLS estimates of employment effects for the aggregate retail sector are generally positive, and the IV estimates negative, is consistent with Wal-Mart endogenously locating stores in places where retail growth is increasing. As shown in the table, there is always statistically significant evidence of endogeneity bias in the aggregate retail sector. We also reject exogeneity in the estimates for general merchandising, and the bias still appears to be upward, which is what we would expect if Wal-Mart enters markets where prospects for general merchandising are good. However, we should not expect much endogeneity bias in the estimated effect of Wal-Mart on employment in this subsector, and that is reflected in the much smaller difference between the OLS and IV estimates of the effects of Wal-Mart on general merchandising employment. Wal-Mart may well have chosen to locate in areas where retail growth was likely to be strong, but the composition of the retail sector is more likely to be a direct consequence of where Wal-Mart chose to locate than a determinant of its location decisions.

### 5.5. Effects on retail sector earnings

Table 5 turns to effects on retail sector earnings. The dependent variable is now the change in retail payrolls per person. The first three columns report results without county-specific trends, and the last three columns including these trends. The results are similar in both cases.

The OLS estimates indicate that a Wal-Mart store opening is associated with increases in retail payrolls of about \$0.2 to \$0.3 million in aggregate retail and about \$0.44 million in the general merchandising sector. In contrast, the IV estimates indicate that aggregate retail payrolls fall when a Wal-Mart store opens, by approximately \$1.1 to \$1.7 million for the A sample; the range is slightly smaller for the B sample. In addition, the evidence is weaker statistically for the B sample and

<sup>39</sup> See, for example, data from the 2002 Economic Census, at [http://factfinder.census.gov/servlet/IBQTable?\\_bm=y&-geo\\_id=&-ds\\_name=EC0244SSSZ1&-\\_lang=en](http://factfinder.census.gov/servlet/IBQTable?_bm=y&-geo_id=&-ds_name=EC0244SSSZ1&-_lang=en) (as of September 21, 2006).

<sup>40</sup> There is a potential countervailing positive effect from Wal-Mart's greater efficiency if lower prices in retail boost consumer demand enough through either income or substitution effects that overall retail sector employment increases. However, this "retail scale" effect may be small if the types of products sold at Wal-Marts and the retail establishments with which they compete are more likely to be inferior goods, so that much of the increase in real income stemming from lower prices is spent in other sectors of the economy (including higher-end retail).

<sup>41</sup> There is a potentially important implication of this finding for interpreting our estimates. In particular, the estimated effects of Wal-Mart openings could be exaggerated if Wal-Mart store openings tend to spur the opening of other large retailers in the same area, in which case we are effectively estimating the impact of more than just Wal-

Mart stores. (It is unlikely that the issue is one of Wal-Mart opening where other large retailers already exist, given the geographic pattern of Wal-Mart openings. Also, even if this sometimes happens, our IV estimates will not pick up the effects of these other stores because their openings will not be predicted by the first stage.) But given that general merchandising employment scarcely rises when Wal-Mart stores open, this seems unlikely.

Table 5  
Estimated effects of Wal-Mart stores on retail payrolls (\$1000s)

	County-specific time trends excluded (Eq. (1))			County-specific time trends included (Eq. (1'))		
	Agg. Retail (1)	Agg. Retail (2)	Gen. Merch. (3)	Agg. Retail (4)	Agg. Retail (5)	Gen. Merch. (6)
	A Sample	B Sample	B Sample	A Sample	B Sample	B Sample
<i>OLS</i>						
Wal-Mart stores	270.2*** (90.0)	166.4 (145.1)	458.9*** (38.9)	233.1*** (89.7)	172.5 (154.1)	427.6*** (39.3)
Total payrolls per person	0.024*** (0.005)	0.020*** (0.005)	0.001*** (0.0003)	0.023*** (0.005)	0.020*** (0.006)	0.002*** (0.0003)
Adjusted $R^2$	0.110	0.151	0.067	0.094	0.123	0.039
<i>IV for Wal-Mart stores</i>						
Wal-Mart stores	-1,711.4** (778.1)	-1,669.8* (912.6)	684.3*** (132.2)	-1,149.8 (735.8)	-1,547.6* (894.4)	525.4*** (149.2)
Total payrolls per person	0.023*** (0.005)	0.020*** (0.005)	0.001*** (0.0003)	0.023*** (0.005)	0.020*** (0.006)	0.002*** (0.0003)
Reject null of no endogeneity bias at 5% level	Yes	Yes	Yes	Yes	Yes	No
<i>N</i>	54,554	32,668	32,668	54,554	32,668	32,668

The notes from Table 4 apply. The only difference is that the dependent variable is the change in retail payrolls per person, so the coefficients now measure the effect of one Wal-Mart store opening on county-level retail payrolls (measured in units of thousands of 1999 constant dollars). The *F*-statistics for the first stage are as reported in Table 4.

\* Significant at the 10% level.

\*\* Idem., 5%.

\*\*\* Idem., 1%.

when county-specific trends are included (not significant for the A sample). However, the positive effect of Wal-Mart openings on payrolls in general merchandising remains, and the estimates are a bit larger than the OLS estimates.

The payroll results largely parallel those for employment. Of course, much of the public debate over Wal-Mart concerns the levels of wages Wal-Mart pays and whether it drives down wages of competitors. As pointed out earlier, because the CBP data do not distinguish part-time from full-time workers, the data cannot be used to compute hourly wages. In addition, the data do not allow any controls for individual worker characteristics that would be required to determine whether wages were changing for comparable workers. Nonetheless, we estimated models for changes in retail earnings per worker, with the Wal-Mart variable in this case simply measuring the change in the number of stores. The point estimates were positive for the A sample (although small—indicating increases of \$7 to \$47—and insignificant); they were negative and insignificant for the B sample. Thus, there is certainly no evidence that Wal-Mart openings reduce retail earnings per worker. Taken literally, the implication of no impact of store openings on wages would imply that the labor supply curve to the retail industry is close to perfectly elastic. A highly elastic labor supply curve may

not be unreasonable, given that retail is a small share of the overall labor market, and is a very low-skill sector in which labor supply can respond very quickly and strongly to wage changes (in contrast to a highly-skilled sector). However, we remind the reader that the CBP data could be quite misleading about wage effects. For example, if Wal-Mart induces a shift toward older or full-time workers, increased retail earnings per worker could be consistent with lower wage rates for comparable workers.<sup>42</sup> Regardless, the combined estimates indicate that overall earnings in the retail sector decline after Wal-Mart stores open.

### 5.6. Identification

A natural principal concern with our approach is identification. We therefore address a few issues related to identification in Table 6, focusing on the aggregate retail employment estimates for the full (A) sample. First, if we are actually detecting effects of Wal-Mart

<sup>42</sup> In principle, one could study these issues using data from the Current Population Survey (CPS) or the Decennial Census of Population. However, in the CPS most county identifiers are suppressed for reasons of confidentiality. Census data are less attractive because they are not available for each year but only once a decade. Furthermore, in downloadable Census estimates by county from all long-form respondents, neither hours, education, nor income are available by industry.

Table 6  
Assessments of identification

	County-specific time trends excluded (Eq. (1))	County-specific time trends included (Eq. (1'))
	Employment (1)	Employment (2)
1. Table 4 estimates, retail employment	−166.98*** (48.83)	−146.35*** (47.77)
<i>Falsification test, manufacturing employment</i>		
2. IV [first-stage <i>F</i> -statistic]	160.74** (80.67) [12.8]	−60.45 (73.05) [4.2]
<i>Restricted IVs isolating Wal-Mart entry</i>		
3. Instrument with year-distance interactions for pairs satisfying (i) 1982 or earlier and within 500 miles of Benton County, or (ii) 1992 or later and 1000 miles or more from Benton County [first-stage <i>F</i> -statistic], retail employment	−190.47*** (59.74) [7.7]	−97.15* (54.45) [7.8]
<i>Exclude counties near Benton County</i>		
4. Exclude counties within 100 miles of Benton County ( <i>N</i> = 53,716), retail employment	−179.27*** (52.59)	−159.34*** (50.86)
5. Exclude counties within 200 miles of Benton County ( <i>N</i> = 51,539), retail employment	−197.59*** (58.22)	−177.03*** (56.80)
6. Exclude counties within 300 miles of Benton County ( <i>N</i> = 47,321), retail employment	−207.38*** (64.84)	−185.66*** (64.31)

Table reports IV estimates of coefficient on change in number of Wal-Mart stores per person. See notes to Tables 1 and 4. All estimates are for A Sample. Sample sizes are the same as in Table 4. Estimated coefficients of total payroll per person variable are not shown.

\* Significant at the 10% level.

\*\* Idem., 5%.

\*\*\* Idem., 1%.

stores, then paralleling the evidence of adverse effects on retail, we should *fail* to find evidence of effects of store openings on industries unlikely to be affected. Employment in service industries and retail sectors other than department stores do not necessarily provide useful falsification exercises of this type, since the location of retail activity may affect these other industries and sectors. However, it is useful to look at manufacturing. Although this industry might be affected by cost pressures exerted by Wal-Mart, there is no reason why the effects of Wal-Mart openings should be felt on *local* manufacturers. As noted earlier, we are skeptical of using our instrumental variables approach to study an industry like manufacturing when the county fixed effects are excluded, since the other industry may have strong region-specific trends; but we are more confident with the county fixed effects included. The results for manufacturing are reported in row 2 of Table 6. We find that in the specification with the county dummies, there is no effect of Wal-Mart openings. On the other hand, the specification without the county dummies show a positive effect. The fact that we only get an effect when we exclude the county dummies (which means we do not allow for county-specific trends) helps confirm our concerns about the validity of the instrument for man-

ufacturing, as the differences in results imply that the instrument(s) is (or are) correlated with these trends. But the results with the county dummies clearly provide a better falsification test, and show no effect of Wal-Mart precisely where we should see none.

A second concern is that our distance-time interaction instrumental variables do not simply capture the beginning of the period when Wal-Mart stores first “reach” a region, but also periods in which they were filling in regions where numerous stores had already opened. For example, Fig. 4 shows that in the band 200–300 miles from Benton County, there are higher predicted openings through 1982, when stores were first opening in this region, as well as in 1987, when stores were filling in this region. This is not necessarily problematic for the IV strategy. The distance and time interactions predict where and when stores were opened, whereas the endogeneity we are trying to purge concerns the choice about the *particular* counties in which to open stores given that Wal-Mart was opening stores in a particular region. So the first-stage estimates will also reflect stores opening in the 200–300 mile band in 1987. But *all* counties in this band in this year have a higher predicted probability of a store opening, not just the specific counties where stores opened. Nonetheless, it might be prefer-

able to focus on the variation in store openings associated with Wal-Mart's initial wave of entry into a region, which may be more exogenous than where the company chose to fill in stores. To do this, we restricted the distance-time interactions included in the first stage to those picking up either observations early in the sample period close to Benton County (1982 or earlier and within 500 miles of Benton County), or observations late in the sample period far from Benton County (1992 or later and 1000 miles or more from Benton County). Figure 4 suggests that these periods and distances capture distinct initial waves of Wal-Mart openings. The estimates, reported in row 3 of Table 6, are relatively similar to the baseline estimates, although the specification with county dummy variables included provides weaker evidence (significant only at the 10-percent level) of negative employment effects.<sup>43</sup>

Finally, our identification strategy requires that there not be unmeasured shocks to the dependent variable that are common to points on a ring of given radius around Benton County. We argued that this condition seems likely to hold, since there is no obvious reason why all points within, say 400–500 miles should exhibit systematically different changes in particular periods relative to Benton County (aside from the effects of Wal-Mart). However, for short distances around Benton County it is possible that there are common shocks, and if it is the differences in retail employment growth between these short distances and greater distances that identifies the effects of Wal-Mart in our empirical analysis, then we could be picking up a spurious relationship. In rows 4–6 of Table 6, we present evidence assessing this possibility, dropping all observations on counties within, in turn, 100, 200, and 300 miles of Benton County. As the table shows, the results are robust to dropping these counties.

### 5.7. Robustness and other analyses

We close the empirical analysis by briefly describing some additional analyses exploring the robustness of our results on retail employment, and some other issues, in Table 7. As noted earlier, we have reported standard errors clustered on state and year. In row 2, we instead report estimates (which are the same) with standard errors clustered on state only, which allows temporal autocorrelation within and across counties in each state. The standard errors are larger in all cases, although not by that much, and the statistical conclusions are unchanged. Note also that in column (2) the

standard error changes by less. This is what we would expect if the inclusion of the county-specific trends captures sources of dependence between observations on the same county.

We have included a control for total payrolls per person in all of the specifications reported thus far. Because this variable includes retail payroll, it is potentially endogenous. In rows 3 and 4 we therefore report two additional specifications: one with retail payrolls subtracted out of total payrolls; and the other simply dropping the total payrolls control. The qualitative conclusions are very similar, although both changes—and more so dropping the total payroll control—yield slightly larger negative effects on retail employment. Next, we drop counties that never had a Wal-Mart store during the sample period. In this case, identification of the effects of Wal-Mart stores comes only from the time-series variation in store openings for the set of counties that got a store, providing a potentially “cleaner” control group. The results, reported in row 5, are qualitatively similar, although the estimated negative employment effects are smaller.

We also consider some other differences in the sample definition. Row 6 reports estimates dropping the small number of counties with stores that closed; these estimates are very close to the baseline estimates.<sup>44</sup> Following that, row 7 reports estimates extending through the entire period covered by the data—ending in 2002 rather than 1995—even though, as discussed earlier, the identifying relationship between time, location, and store openings is strongest through 1995. For the longer period the evidence suggests sharper declines in retail employment.

Thus far we have ignored information on Sam's Clubs. We do not necessarily want to treat these as equivalent to other Wal-Mart stores, so we study them in two ways. First, we omit all counties that had a Sam's Club at some point during the sample period; 97 percent of these counties also had a Wal-Mart. Second, we use the full sample but treat each Sam's Club store like other Wal-Mart stores. As the table shows in rows 8 and 9, the results dropping the counties with Sam's Clubs are qualitatively similar, although the estimated effects are smaller. The results are quite insensitive to simply treating Sam's Clubs like other Wal-Mart stores.

To this point we have simply used a count of Wal-Mart stores (on a per capita basis) to measure exposure

<sup>43</sup> This specification uses 58 instead of 272 instruments.

<sup>44</sup> An alternative is to incorporate information on store closings as reductions in the number of Wal-Mart stores. However, we are skeptical that store closings and openings have symmetric (opposite-sign) effects, and we do not have a natural instrument for store closings.

Table 7  
Robustness analyses

	County-specific time trends excluded (Eq. (1))	County-specific time trends included (Eq. (1'))
	Retail employment (1)	Retail employment (2)
1. Table 4 estimates	−166.98*** (48.83)	−146.35*** (47.77)
2. Cluster on state only	−166.98*** (55.94)	−146.35*** (49.28)
3. Subtract retail payroll out of total payroll per person control	−174.07*** (45.07)	−153.74*** (49.57)
4. Omit total payroll per person control	−181.58*** (46.12)	−162.07*** (51.23)
5. Drop counties that never have a store ( $N = 30,957$ )	−94.40*** (30.63)	−103.10*** (34.95)
6. Drop counties with closed stores ( $N = 53,746$ )	−166.68*** (43.15)	−144.08*** (47.19)
7. Through 2002 ( $N = 75,482$ )	−184.84*** (41.27)	−215.29*** (59.53)
8. Drop counties with Sam's Clubs ( $N = 49,031$ )	−150.47*** (43.63)	−130.88*** (47.31)
9. Combine Wal-Mart stores with Sam's Clubs	−161.44*** (42.72)	−175.96*** (47.81)
10. Weighted by store size	−216.12*** (59.53)	−223.48*** (62.41)
11. Counties with above median population (21,000) in 1977 ( $N = 27,290$ )	−81.13 (57.58)	−105.00** (48.90)
12. Counties with below median population (21,000) in 1977 ( $N = 27,264$ )	−103.28* (53.48)	−90.03 (59.70)
13. Two-year differences ( $N = 51,429$ )	−196.48*** (34.73)	−209.61*** (48.49)
14. Three-year differences ( $N = 48,362$ )	−206.03*** (35.81)	−224.35*** (45.24)

Table reports IV estimates of coefficient on change in number of Wal-Mart stores per person. See notes to Tables 1 and 4. All estimates are for A Sample. Sample sizes are the same as in Table 4 unless otherwise noted. Estimated coefficients of total payroll per person variable, where included, are not shown. For the estimates labeled “weighted by store size,” stores are weighted by their square footage (at the end of the sample period) divided by the average square footage for the sample, to maintain comparability with the baseline estimates. The specifications using two- and three-year differences use dummy variables for the distinct two- or three-year periods over which the differences are computed, rather than dummy variables for single years. That way, the same dummy variables appear in the first-differenced and levels regressions.

\* Significant at the 10% level.

\*\* Idem., 5%.

\*\*\* Idem., 1%.

to Wal-Mart. But store size varies—for example, counties with smaller populations may get smaller stores<sup>45</sup>—and the exposure measure should perhaps take this into account. Thus, we computed an exposure measure that weights by store size relative to the average store size. These estimates are reported in row 10, and reveal somewhat larger effects. Note, however, that in our data size is measured as of the end of the sample period only, and will not pick up things such as the conversion of a regular store to a Supercenter.

The effects of Wal-Mart stores may depend on the type of area in which they are located. In rural and less dense counties Wal-Mart may have a larger impact because there are fewer other large-scale competitors before they enter. Alternatively, in these areas the stores may be more likely to create jobs because shoppers (and workers) come from other counties. As shown in Fig. 5, Wal-Mart stores originally opened more commonly in rural areas, and over time shifted to more urban areas. (This classification is based on whether the store opened in what is now classified as an MSA.) Because we do not have a clean classification of counties located in MSAs or not, we instead focus on county population,

<sup>45</sup> This may in part reflect whether or not the Wal-Mart store is a Supercenter.

### Percentage of Wal-Mart Openings in Urban Locations, 1977-1995

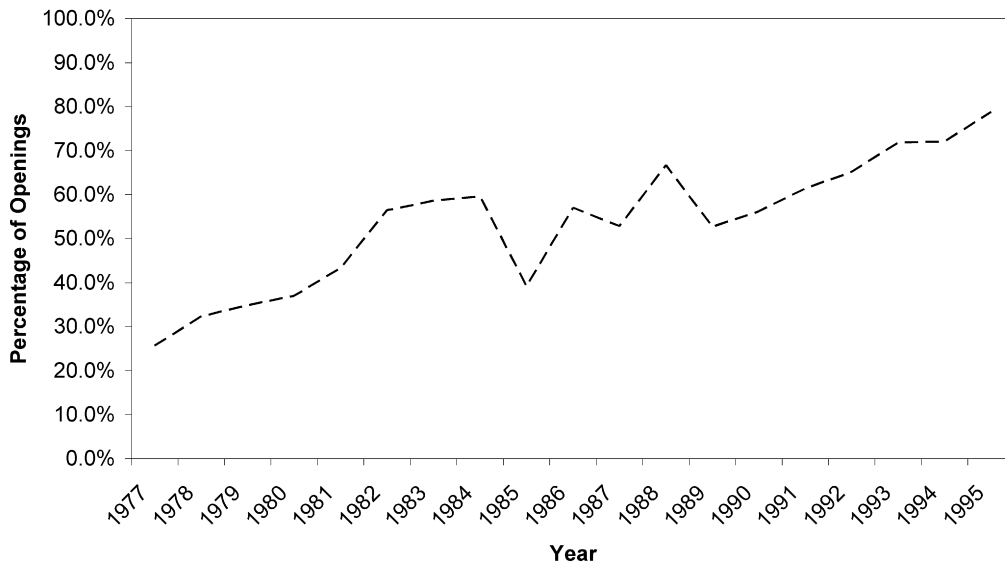


Fig. 5. Locations of Wal-Mart openings.

and estimate the models for two samples: counties with above median population at the beginning of the sample period, and counties with below median population. As shown in rows 11 and 12, there is no clear connection between density and the effects of Wal-Mart stores.

Finally, we use equations differenced over two and then three years, rather than one year. We do this for a few reasons. First, there may be some anticipatory effects of Wal-Mart store openings. This could lead to effects of store openings that are biased towards zero, if the shorter differences do not include these anticipatory effects (which seem likely to be in the same direction as those occurring after stores open). Second, the effects of store openings might not all occur within a year. Allowing for longer differences can pick up longer-term effects, and it is ambiguous whether these are likely to be larger or smaller.<sup>46</sup> Third, because of how we code store openings, in about one-sixth of cases (for stores opening between January and early March) the employment change and store opening may be misaligned, likely biasing the estimated effect towards zero. By using longer differences, we reduce this problem substantially. As the estimates in the last two rows of Table 7 show, with two-

and three-year differences the estimates are qualitatively very similar (and were for the other samples, for general merchandising, and for OLS as well as IV), but always slightly larger in absolute value the longer the differences, as we might have expected from the first and third considerations discussed above.<sup>47</sup>

## 6. Conclusions and discussion

Motivated in large part by local policy debates over Wal-Mart store openings, and the large size of Wal-Mart relative to the retail sector, we estimate the effects of Wal-Mart stores on retail employment and earnings. Critics have charged that Wal-Mart's entry into local labor markets reduces employment and wages, and the company (and others) have countered that these claims are false, and touted Wal-Mart's retail job creation effects.

Our analysis emphasizes the importance of accounting for the endogeneity of the location and timing of Wal-Mart openings that—in our view, and as borne out by the data—is most likely to bias the evidence against

<sup>46</sup> In principle, one can study contemporaneous and lagged effects more directly by including lagged store openings as well. However, experimenting with such specifications in the IV setting led to uninformative estimates, because predicted values of store openings at different lag lengths were very highly correlated, which is not surprising given the nature of our IV.

<sup>47</sup> Finally, although not reported in the table, we also estimated a simpler specification for the change in retail employment per person on the change in the number of Wal-Mart stores (with the latter not on a per capita basis, unlike the rest of the estimates reported in the paper). Both the OLS and IV estimates were qualitatively similar to those reported in Table 4, and implied similar percentage changes in retail employment.

finding adverse effects of Wal-Mart stores. Our strategy for addressing the endogeneity problem is based on a natural instrumental variables approach that arises because of the geographic and time pattern of the opening of Wal-Mart stores, which slowly spread out in a wave-like fashion from the first stores in Arkansas.

The findings in this paper rather strongly belie claims, as well as recent research findings, suggesting that Wal-Mart store openings lead to increased retail employment. On average, Wal-Mart store openings reduce retail employment by about 2.7 percent, implying that each Wal-Mart employee replaces about 1.4 employees in the rest of the retail sector. Driven in part by the employment declines, retail earnings at the county level also decline as a result of Wal-Mart entry, by about 1.5 percent. It is harder to draw any firm conclusions regarding the effects of Wal-Mart on wages, although the data do not provide any indication that retail earnings per worker are affected by Wal-Mart openings.

Note that the estimated adverse effects on retail employment do not imply that the growth of Wal-Mart has resulted in lower absolute levels of retail employment. Like in all studies of this type, the estimates are relative to a counterfactual of what would have happened to retail employment absent the effects of Wal-Mart. From 1961, the year before the first Wal-Mart store opened, through 2004, the last full year for which we have a count of the Wal-Mart Discount Stores and Supercenters, retail employment in the United States grew from 5.56 million to 15.06 million, or 271 percent, considerably faster than overall employment (242 percent).<sup>48</sup> If each of the 3066 stores present in January of 2005 reduced retail employment by our estimate of 146 workers (Table 4, column (4)) relative to the counterfactual, then our estimates imply that, in the absence of Wal-Mart, retail employment would have instead grown to 15.51 million as of 2004, or 3 percent higher than the observed figure. So the negative employment effects of Wal-Mart that we estimate simply imply that retail employment growth was a bit more modest than it would otherwise have been, growing by 271 percent from 1961 through 2004, rather than 279 percent. The estimates do imply, however, that retail employment is lower than it would have been in counties that Wal-Mart entered, and hence that Wal-Mart has negative rather than positive effects on net job creation in the retail sector.

The lower retail employment associated with Wal-Mart does not necessarily imply that Wal-Mart stores

worsen the economic fortunes of residents of the markets that these stores enter. Our results apply only to the retail sector, and we suspect that there are not aggregate employment effects, at least in the longer run, as labor shifts to other uses. Wage effects are more plausible, although these may operate more on the manufacturing side through the buying power that Wal-Mart exerts, as opposed to the retail side which is a low-wage sector regardless of Wal-Mart—although there are exceptions such as relatively highly-paid grocery workers who may be harmed from competition with Supercenters. If there are wage (or employment) effects that arise through cost pressures on Wal-Mart's suppliers, however, they would not necessarily be concentrated in the counties in which stores open, so that our methods would not identify them.

Moreover, Wal-Mart entry may also result in lower prices that increase purchasing power, and if prices are lowered not just at Wal-Mart but elsewhere as well, the gains to consumers may be widespread. Furthermore, the gains may be larger for lower-income families (Hausman and Leibtag, 2005), although it is also possible that labor market consequences for these families are also more adverse.

Another line of criticism of Wal-Mart is that through lowering wages it increases the burden on taxpayers by increasing eligibility for and receipt of government benefits.<sup>49</sup> However, a key implicit assumption is that in the absence of Wal-Mart, employees of the company would have higher-paying jobs, rather than, for example, no jobs. Thus, the validity of this criticism hinges on whether Wal-Mart's entry into a labor market affects overall employment and wages. It is also worth pointing out that if Wal-Mart causes both earnings and price declines for low-income families, then taxpayer burden could increase even if the price declines more than offset the earnings declines for these families, because government programs are typically tied to earnings.

Thus, aside from the question of employment effects, there are numerous remaining questions of considerable interest regarding the effects of Wal-Mart on labor markets and goods markets, on consumption, and on social program participation and expense. In addition,

<sup>48</sup> See [http://data.bls.gov/PDQ/servlet/SurveyOutputServlet?&series\\_id=CEU420000001](http://data.bls.gov/PDQ/servlet/SurveyOutputServlet?&series_id=CEU420000001) and <ftp://ftp.bls.gov/pub/suppl/empst/ceseeb1.txt> (as of December 22, 2006).

<sup>49</sup> A report by the Democratic Staff of the Committee on Education and the Workforce of the US Congress (Miller, 2004) claims that because of Wal-Mart's low wages, an average Wal-Mart employee costs federal taxpayers an extra \$2103 in the form of tax credits or deductions, or public assistance such as healthcare, housing, and energy assistance (see also Dube and Jacobs, 2004). There are many heroic assumptions needed to construct such estimates, and this is not the place to dissect them.



tion, of course, there are non-economic issues that play a role in the debate over Wal-Mart, such as preferences for downtown shopping districts versus suburban malls. The identification strategy developed in this paper may prove helpful in estimating the effects of Wal-Mart stores on some of these other outcomes as well.

## Acknowledgments

We are grateful to Ron Baiman, Emek Basker, Marianne Bitler, Jan Bruckner, Eric Brunner, Colin Cameron, Wayne Gray, Judy Hellerstein, Hilary Hoynes, Chris Jepsen, Giovanni Peri, Howard Shatz, Betsey Stevenson, Brandon Wall, Jeff Wooldridge, two anonymous referees, and seminar participants at Brookings, Clark, Cornell, HKU, IZA, PPIC, Stanford, UQAM, UC-Davis, UC-Irvine, UC-Riverside, UC-San Diego, UConn, and USC for helpful comments. We are also grateful to Wal-Mart for providing data on store locations and opening dates, and to Emek Basker for providing her Wal-Mart data set and other code; any requests for Wal-Mart data have to be directed to them. Despite Wal-Mart having supplied some of the data used in this study, the company has provided no support for this research, and had no role in editing or influencing the research as a condition of providing these data. The views expressed are those of the authors, and not those of Wal-Mart.

## References

- Artz, G.M., McConnon, J.C., Jr., 2001. The impact of Wal-Mart on host towns and surrounding communities in Maine. Working paper. Office of Social and Economic Trend Analysis, Iowa State University. Available at [http://www.seta.iastate.edu/retail/publications/artz\\_narea\\_paper.pdf](http://www.seta.iastate.edu/retail/publications/artz_narea_paper.pdf) (as of September 7, 2005).
- Barnes, N.G., Connell, A., Hermenegildo, L., Mattson, L., 1996. Regional differences in the economic impact of Wal-Mart. *Business Horizon* 39, 21–25.
- Basker, E., 2005a. Selling a cheaper mousetrap: Wal-Mart's effect on retail prices. *Journal of Urban Economics* 58, 203–229.
- Basker, E., 2005b. Job creation or destruction? Labor market effects of Wal-Mart expansion. *Review of Economics and Statistics* 87, 174–183.
- Bianchi, D., Swinney, D., 2004. Wal-Mart: A destructive force for Chicago communities and companies. Center for Labor and Community Research. Available at [http://www.clcr.org/publications/pdf/walmart\\_impact\\_analysis.pdf](http://www.clcr.org/publications/pdf/walmart_impact_analysis.pdf) (as of September 7, 2005).
- Bianco, A., Zellner, W., 2003. Is Wal-Mart too powerful? *Business Week* (October 6). Available at <http://209.157.64.200/focus/f-news/990922/posts> (as of October 31, 2005).
- Bowermaster, J., 1989. When Wal-Mart comes to town. *New York Times* (April 2). Section 6, Part 2, p. 8.
- Chao, J.C., Swanson, N.R., 2003. Alternative approximations of the bias and MSE of the IV estimator under weak identification with an application to bias correction, Discussion paper No. 1418. Cowles Foundation, Yale University.
- Davidson, S.M., Rummel, A., 2000. Retail changes associated with Wal-Mart's entry into Maine. *International Journal of Retail and Distribution Management* 28, 162–169.
- Drewianka, S., Johnson, D., n.d. Wal-Mart and local labor markets, 1990–2004. Unpublished paper. University of Wisconsin, Milwaukee.
- Dube, Arindrajit, Jacobs, K., 2004. Hidden cost of Wal-Mart jobs: Use of safety net programs by Wal-Mart workers in California. Briefing paper series. UC Berkeley Labor Center, August.
- Dube, A., Eidlin, B., Lester, B., 2005. Impact of Wal-Mart growth on earnings throughout the retail sector in urban and rural counties. Unpublished paper. University of California, Berkeley.
- Freeman, G., 2004. Wal-Mart Supercenters: What's in store for southern California? Los Angeles County Economic Development Corporation, January.
- Goetz, S.J., Swaminathan, H., 2004. Wal-Mart and county-wide poverty. Staff paper No. 371. Department of Agricultural Economics and Rural Sociology, Pennsylvania State University. Available at <http://cecd.aers.psu.edu/pubs/PovertyResearchWM.pdf> (as of September 7, 2005).
- Global Insight, 2005. The economic impact of Wal-Mart.
- Goldman, A., Cleeland, N., 2003. The Wal-Mart effect; An empire built on bargains remakes the working world. *Los Angeles Times* (November 23). Business Desk, Part A, p. 1.
- Hausman, J., 1987. Specification tests in econometrics. *Econometrica* 46, 1251–1271.
- Hausman, J., Leibtag, E., 2004. CPI bias from Supercenters: Does the BLS know that Wal-Mart exists? Working paper No. 10712. National Bureau of Economic Research, Cambridge, Massachusetts.
- Hausman, J., Leibtag, E., 2005. Consumer benefits from increased competition in shopping outlets: Measuring the effect of Wal-Mart. Unpublished paper. MIT.
- Hicks, M.J., Wilburn, K., 2001. The regional impact of Wal-Mart entrance: A panel study of the retail trade sector in West Virginia. *Review of Regional Studies* 31, 305–313.
- Hiltzik, M., 2004. Golden state; UFCW sacrifices workers while declaring victory. *Los Angeles Times* (March 4). Business Desk, Part C, p. 1.
- Holmes, T., 2005. The diffusion of Wal-Mart and economies of density. Unpublished paper. University of Minnesota.
- Ingold, J., 2004. Disdain for Wal-Mart pulls neighbors together; opposition becomes a social phenomenon. *Denver Post* (April 18). A Section, p. A-01.
- Jacobs, A., 2004. In a historic black hamlet, Wal-Mart finds resistance. *New York Times* (March 6). Section A, p. 8.
- Jia, P., 2006. What happens when Wal-Mart comes to town: An empirical analysis of the discount retailing industry. Unpublished paper. MIT.
- Johnson, B.C., 2002. Retail: The Wal-Mart Effect. McKinsey Global Institute.
- Kaufman, L., 1999. Wal-Mart casts eye northward. *New York Times* (February 16). Section C, p. 1.
- Keon, T., Robb, E., Franz, L., 1989. Effect of Wal-Mart stores on the economic environment of rural communities. Working paper. Business and Public Administration Research Center and College of Business and Public Administration, University of Missouri.
- Ketchum, B.A., Hughes, J.W., 1997. Wal-Mart and Maine: The effect on employment and wages. *Maine Business Indicators* 42, 6–8.
- Khanna, N., Tice, S., 2000. Strategic responses of incumbents to new entry: The effect of ownership structure, capital structure, and focus. *Review of Financial Studies* 13, 749–779.
- Khanna, N., Tice, S., 2001. The bright side of internal capital markets. *The Journal of Finance* 56, 1489–1528.

- McGee, J.E., 1996. When Wal-Mart comes to town: A look at how local retailers respond to the arrival of the retailing giant. *Journal of Business and Entrepreneurship* 8, 43–52.
- Miller, G., 2004. Everyday low wages: The hidden price we all pay for Wal-Mart. Democratic Staff of the Committee on Education and the Workforce, US House of Representatives. Available at <http://edworkforce.house.gov/democrats/WALMARTREPORT.pdf> (as of September 7, 2005).
- Nieves, E., 1995. Our towns; Taking on Wal-Mart, ahead of time. *New York Times* (August 1). Section B, p. 5.
- Neumark, D., Zhang, J., Ciccarella, S., 2005. The effects of Wal-Mart on local labor markets. Working paper No. 11782. National Bureau of Economic Research, Cambridge, Massachusetts (revised 2007).
- Norman, A., 2004. *The Case Against Wal-Mart*. Raphel Marketing, Atlantic City, New Jersey.
- PBS, 2004. One, two, three, four ... we don't want your Superstore. Frontline program: "Is Wal-Mart Good for America?" November 16. Available at <http://www.pbs.org/wgbh/pages/frontline/shows/walmart/transform/protest.html> (as of September 7, 2005).
- Peterson, M., McGee, J.E., 2000. Survivors of 'W-day': An assessment of the impact of Wal-Mart's invasion of small town retailing communities. *International Journal of Retail and Distribution Management* 28, 170–180.
- Quinn, B., 2000. *How Wal-Mart is Destroying America and the World and What You Can Do about It*. Ten Speed Press, Berkeley, California.
- Rimer, S., 1993. Around New England, Main Street faces peril. *New York Times* (February 28). Section 1, p. 22.
- Rodino Associates, 2003. Report on research for big box retail/superstore ordinance. Prepared for Industrial and Commercial Development Division, Community Development Department, City of Los Angeles. Available at [http://www.lacity.org/council/cd13/houscommecdev/cd13houscommecdev239629107\\_04262005.pdf](http://www.lacity.org/council/cd13/houscommecdev/cd13houscommecdev239629107_04262005.pdf) (as of September 7, 2005).
- Sinnott, R.W., 1984. Virtues of the Haversine. *Sky and Telescope* 68, 159.
- Slater, R., 2003. *The Wal-Mart Decade: How a Generation of Leaders Turned Sam Walton's Legacy into the World's #1 Company* (New York, New York: Portfolio).
- Staiger, D., Stock, J.H., 1997. Instrumental variables regression with weak instruments. *Econometrica* 65, 557–586.
- Stone, K.E., 1988. The effect of Wal-Mart Stores on businesses in host towns and surrounding towns in Iowa. Working paper. Department of Economics, Iowa State University. Available at <http://www.econ.iastate.edu/faculty/stone/Effect%20of%20Walmart%20-%201988%20paper%20scanned.pdf> (as of September 7, 2005).
- Stone, K.E., 1995. Impact of Wal-Mart stores on Iowa communities: 1983–93. *Economic Development Review* 13, 60–69.
- Stone, K.E., 1997. Impact of the Wal-Mart phenomenon on rural communities. In: Ernestes, D.P., Hicks, D.M. (Eds.), *Increasing Understanding of Public Problems and Policies*. Farm Foundation, Chicago, IL, pp. 189–200.
- Stone, K.E., Artz, G., Myles, A., 2002. The economic impact of Wal-Mart Supercenters on existing businesses in Mississippi. Mississippi State University Extension Service, Fall. Available at <http://msucares.com/pubs/misc/m1283.pdf> (as of September 7, 2005).
- Wal-Mart Watch, 2005. Low prices at what cost? Annual report. Available at <http://fivestones.sitestream.com/pdf/2005-annual-report.pdf> (as of September 7, 2005).
- Walton, S., 1992. *Made In America: My Story*. Doubleday, New York, NY.