Preliminary: Do Not Cite Without Permission

Mom-and-Pop Meet Big-Box: Complements or Substitutes?

John Haltiwanger[1], Ron Jarmin[2], and C.J. Krizan[34]

April 2009

[1] University of Maryland, U.S. Census Bureau and NBER, haltiwan@econ.bsos.umd.edu
[2] U.S. Census Bureau, ron.s.jarmin@census.gov
[3] U.S. Census Bureau, cornell.j.krizan@census.gov
[4] We thank Paul Bailey for excellent research assistance as well as Randy Becker, Dan Driscoll, Lucia Foster, Kristin McCue and Javier Miranda for helpful comments. We thank the Kauffman Foundation for financial support. Any opinions and conclusions expressed herein are those of the author(s) and do not necessarily represent the views of the U.S. Census Bureau. All results have been reviewed to ensure that no confidential information is disclosed.

I. Introduction

The rise of Big-Box retailing has been the source of considerable debate. Big-Boxes, or large retail outlets operated by large national or multinational chains, have been criticized for their labor market practices, their contribution to the trade deficit and many other things. On the other hand, Big-Boxes are very popular shopping venues and have been a boon to consumers because they offer expansive product lines at low prices.

Perhaps the most often heard and most intuitively appealing criticism of the Big-Box retail format is that it displaces smaller, often family owned (a.k.a. Mom-and-Pop) retail establishments and contributes to the decline of traditional retail districts such as the main streets of small towns and the downtown shopping districts of large cities.[5] Both community leaders and policymakers are interested in knowing whether or not Big-Box retailers displace more retail employment than they create and a number of studies by economists have attempted to help shed light on these issues.

Several authors have examined the empirical evidence on Big-Boxes' impact on retail employment but they often reached different conclusions. For example, recent papers focusing on the impact of Big-Box (particularly Wal-Mart) entry on retail employment at the county level find both a positive and negative impact. Basker (2005) finds that, while it's likely that employment at other retailers shrinks, on average Wal-Mart entry leads to an overall increase in county level retail employment of about 50 jobs. Neumark, Zhang and Ciccarella (2007) highlight potential endogeneity problems with Basker's empirical methodology and use an alternative instrumental variable estimation approach that yields results showing that Wal-Mart entry reduces county retail employment.

Recent work also notes the important role of large retail chains in retail market dynamics and productivity growth. Doms, Jarmin and Klimek (2004) showed that large chains accounted for nearly all retail employment growth in the early 1990's and that they also accounted for a disproportionate share of retail IT investment. Furthermore, they were the only type of retail firm that exhibited IT-related productivity growth. Foster, Haltiwanger and Krizan (2006a) showed the importance of retail establishment net entry in generating productivity growth in the sector and noted that new establishments of large retail chains were an especially important factor.

More recently, Jarmin, Klimek and Miranda (2009) document the changes in the structure of county-level retail markets. They note that the trend away from single unit stores to those operated by chains has been underway for many decades. Importantly, however, they document that substantial entry (as well as exit) of single unit retailers persists. That is, despite seeing their share of overall retail activity decline steadily over decades (not just since the advent of the Big-Box, discount format), single unit retailers still perceive a

[5] We often refer to the single unit stores as Mom-and-Pop stores in this analysis for labeling purposes. This is not meant to indicate that all single unit establishments are family-owned businesses but rather simply that they are single unit stores. Single unit stores are typically small (less than 10 employees) and often are sole proprietorships.

niche and enter retail markets at high rates. This suggests that chain stores are not perfect substitutes for single unit retailers.

Still, the question remains: how frequently do Big-Box stores displace or become a substitute for single unit or small chain stores and when (if ever) is the presence of a Big-Box complementary to smaller retail stores? We argue that the county level data used in prior analyses is too coarse to address these questions and use retail establishment data from the Census Bureau's Longitudinal Business Database (LBD) to more directly measure the impact of changes in Big-Box activity on the activity of nearby single unit and smaller chain stores.

The establishment records on the LBD provide information on firm structure (e.g., whether a retail establishment is part of a chain) and location. We use this rich information to measure the average effects of Big-Box activities within a few miles of smaller retailers, controlling for local retail market conditions. We focus on one metro area – the Washington, D.C. metro area. We use only one metro area given the substantial time and computer resources required in generating our analytic database that has distance measures between all retail stores in the D.C. area along with information on local retail conditions for narrowly defined geographic areas. The resulting analytical database is very rich but it would clearly be of interest in future work to extend the data and analysis to more metro areas.

Our primary objective is descriptive. We attempt to quantify the impact of Big-Box store growth and entry on single unit and smaller chain stores, operating in both the same and other retail sectors, by detailed location controlling for local retail market conditions. Our empirical approach uses a rich set of controls of local retail market conditions such as the population growth within a small area around the stores as well as demographic and income variables characterizing the population in the same area.

Our main finding is that there is a substantial negative impact of Big-Box entry and growth on employment growth at single unit and especially smaller chain stores only if the Big-Box activity is in the immediate area *and* in the same detailed industry. A general pattern is that this impact declines with distance. That is, the impact tends to be the largest if the Big-Box activity is within 1 mile or 1 to 5 miles as opposed to 5 to 10 miles of the store in question. The impact of increased big-box activity manifests itself through a substantial reduction in net employment growth and smaller retailers, which is mostly accounted for by an increase in job destruction from store exit. We find no systematic relationship between the growth and entry of Big-Box activity in *other* sectors on single unit and smaller retail chain stores.

The paper proceeds as follows. Section II provides a discussion of the recent literature studying the changing structure of retail trade that helps provide background and motivation for our analysis. Section III discusses the data sources and measurement methodology used in this paper. In section IV, we discuss our empirical strategy for estimating the impact of Big-Box entry and growth on single unit and smaller chain stores. Concluding remarks are in section V.

II. Background

In this section, we review the recent literature on the dynamics of the retail trade industry with a focus on the key findings and facts that both motivate and provide perspective for our analysis. A number of studies have found that the U.S. retail trade sector has been undergoing a large transformation for many decades. This transformation has seen the small, single location retailer replaced by large retail chains as the dominant retail form. For example, Jarmin, Klimek and Miranda (2009 – hereafter JKM) show that the share of retail trade employment in the U.S. accounted for by single units declined from 53 percent in 1976 to 39 percent in 2000. In this work and the work of Foster, Haltiwanger and Krizan (2006a, 2006b – hereafter FHK), it is clear that this transformation is closely linked to the entry and exit of establishments. That is, adjustment on the extensive margin (i.e., entry and exit of stores) is a primary means by which adjustment occurs in the retail trade industry.

FHK show that over a 5-year horizon about 78 percent of job creation for single unit establishments is accounted for by establishment entry and about 79 percent of job destruction for single unit establishments is accounted for by establishment exit. The shares accounted for by entry and exit for chains are smaller but still quite large. For large chains (using a definition similar to what we use below), FHK find that about 77 percent of job creation is accounted for entry and about 65 percent of job destruction is accounted for by exit. In interpreting these statistics, it is also important to emphasize that job creation and job destruction rates are much higher for single units relative to large chains. The main message that emerges is that there is extensive churning of jobs and establishments in the retail trade industry especially for single unit establishments.

JKM document other important trends in retail trade. For example, they document the increasing size (as measured by employment) of retail establishments, even Mom-and-Pops, over time and across local retail markets of different sizes. They also find that the rate of churning for both single unit and chain retail stores is higher in large metropolitan area like Washington D.C. than it is in smaller urban or rural areas. They find large metro areas have fewer retail stores and firms per capita than do smaller areas and that rural areas have fewer retail employees per capita than larger areas. Thus, retail stores in metro areas are larger than those in smaller areas.

There are two primary strands in the literature that seek to explain the increased scale and scope of retail outlets. First, studies such as Bresnahan and Reiss (1991) and Dinerlsoz (2004) focus on market size to explain the scale and scope of retail trade stores. Schiff (2009) pushes in a related direction for restaurants showing that larger and denser cities not only have more cuisines but there is a clear hierarchy in how less common cuisines appear across cities of increasing size. The other main strand of research is comprised of studies that focus on changes in technology to explain the evolution of the scale and scope of retail trade businesses. For example, Holmes (2001) shows how barcodes lead to more frequent deliveries and larger store sizes. In related work, Doms, Jarmin and Klimek (2004) show that IT investments are related to productivity improvement only for large retail chains. Basker, Klimek and Hoang Van (2009) show retail technology

changes are complementary to consumer preferences for one-stop shopping which leads to more products in larger stores operated by larger chains.

Other authors focus on informing the public debate surrounding the entry of Wal-Mart and other "big-box" stores. In particular, Basker (2005) looks at retail employment at the county level after the entry of a Wal-Mart. She finds that, while it's likely that employment at other retailers shrinks, on average Wal-Mart entry leads to an overall increase in county level retail employment of about 50 jobs. Neumark, Zhang and Ciccarella (2007) criticize Basker's IV strategy and use an alternative that show Wal-Mart entry has a negative impact on county employment.

This brief review helps both motivate and distinguish the approach we take in this paper. We focus on detailed location effects within a metro area rather than between county variation across the nation that is used in the recent literature. Our starting point is to view different classes of retailers operating within a metro area as inputs to the provision of retail services. Motivated by the recent literature, we view the changes that are ongoing within a metro area to be driven by factors such as changing technology, consumer preferences, transportation networks and regulation that induce changes in the way that retail services are provided. That is, we view that the underlying conceptual framework is one where changes in the structure of local retail markets reflect retailers' attempts to find the most efficient way to provide retail services given technology and preferences. With this conceptual framework in the background, we are interested in investigating what happens to single unit and small chain stores when a Big-Box store enters or grows in their immediate local area.

III. Data and Measurement Issues

While we obtain data from several sources, our core data that allow us to track retail trade establishments and firms come from the Longitudinal Business Database (LBD) developed by Jarmin and Miranda (2002). The LBD covers all establishments and firms in the non-farm business sector from 1976 to 2005 and includes information about payroll, employment, detailed location, detailed industry and ownership structure. The data appendix explains the specifics of how we use the information from the LBD to construct our measures.

While the data appendix provides many details about our measurement methodology, our basic approach is to use the LBD to assemble a set of the retail trade establishments operating in our target counties during the 1976 to 2005 period. We use the longitudinal nature of the data to minimize spurious entry and exit due to missing geographic or industry information. The LBD's firm identifiers allowed us to label establishments according to whether they were part of a single or multiunit firm and classify the multiunits according to how many states they operate in. For the remainder of the paper

we refer to "small" chains that have establishments in 1 to 14 states and "large" chains operating in 15 or more states.[6]

The LBD does not contain a classification for Big-Box retail stores so we use information from a variety of outside data sources to identify them for our analysis. Due to the confidential nature of the Census Bureau micro data used in our study, we can't list the names of the well-known retailers that make up our list of Big-Boxes but in the appendix we describe the sources and criteria for classifying stores as Big-Boxes. Virtually all of our Big-Boxes satisfy the criteria of large chains (i.e., they are part of multi-unit firms that operate in 15 or more states). However, in the analysis in this paper, the large chain category is defined as all stores that satisfy the large chain criteria and are not Big-Boxes. To sum up, we have four mutually exclusive categories: single units (a.k.a. Mom-and-Pops), small retail chains, large retail chains and Big-Boxes.

Another critical part of our analysis is to assign a detailed location to all our establishments. As described in the data appendix, we assign a latitude and longitude measure to the retail trade establishments in the D.C. metro area (where again this is defined in a consistent manner as described in the appendix). We use the latitude and longitude information to construct distance measures between all the retail trade establishments in the D.C. region. As described below, by using the distance measure we can construct activity rings of various sorts (Big-Box, demographic, etc.) with 1 mile, 1 to 5 mile and 5 to 10 mile radii.[7]

Our demographic data come from the 1980, 1990, and 2000 Decennial Censuses. We use these microdata files to construct measures of population characteristics, income, and population growth in local areas. These methodologies are also described in detail in the data appendix.

Much of our analysis is based on analyzing employment growth at the establishment level and the decomposition of employment growth into components like job destruction from exit. In the remainder of this section, we provide details about our measures of employment and growth. Let E_{it} be employment in year *t* for establishment *i*. In practice, this is a point-in-time measure reflecting the number of workers on the payroll for the payroll period that includes 12 March. We measure establishment-level employment growth as follows:

$$g_{it} = (E_{it} - E_{it-1}) / X_{it},$$

where

[6] In the recent literature some have used the number of states and some the number of stores to classify establishments into chain types (see, e.g., FHK (2006b) and Basker, Klimek, Hoang Van (2009)). Our method follows FHK (2006b) and a large chain in our case by construction has at least 15 stores operating in at least 15 states. Results on the changing composition of retail trade towards larger chains are not sensitive to the precise cutoff for large chains.

[7] We recognize that this simple measure of 'nearness' ignores natural barriers such as rivers, lakes, and other obstacles.

$X_{it} = .5*(E_{it} + E_{it-1})$.

This growth rate measure has become standard in analysis of establishment and firm dynamics, because it shares some useful properties of log differences but also accommodates entry and exit. (See Davis et al 1996, and Tornqvist, Vartia, and Vartia 1985).[8] In what follows, we refer to this as the DHS growth rate measure. Note that the DHS growth rate measure can be defined at any level of aggregation (establishment, local area, industry, etc.)

Measures of job creation and destruction at the establishment level can be given by:

$$JC_{it} = \max(g_{it}, 0)$$
$$JD_{it} = \max(-g_{it}, 0)$$

Job creation from entry at the establishment level is given by:

$$JC_{it} = \max(g_{it}, 0) * I\{g_{it} = 2\}$$

where I is an indicator variable equal to one if expression in brackets hold, zero otherwise, and $g_{it} = 2$ denotes an entrant.

Similarly job destruction from exit at the establishment level is given by:

$$JD_{it} = \max(-g_{it}, 0) * I\{-g_{it} = 2\}$$

where $g_{it} = -2$ denotes an exit.

Using these measures it is straightforward to generate aggregate measures of job creation and destruction as well as job creation and destruction from entry and exit, respectively (at any level of aggregation) given by:

$$JC_t = \sum_i (X_{it} / X_t) \max\{g_{it}, 0\}$$
$$JD_t = \sum_i (X_{it} / X_t) \max\{-g_{it}, 0\}$$
$$JC\_Entry_t = \sum_i (X_{it} / X_t) I\{g_{it} = 2\}.$$
$$JD\_Exit_t = \sum_i (X_{it} / X_t) I\{g_{it} = -2\}$$

Given these definitions, the following simple relationships hold:

[8] The DHS growth rate like the log first difference is a symmetric growth rate measure but has the added advantage that it accommodates entry and exit. It is a second order approximation of the log difference for growth rates around zero.

$g_t = JC_t - JD_t$, $JC_t = JC\_Cont_t + JC\_Entry_t$ and $JD_t = JD\_Cont_t + JD\_Exit_t$

where JC_Cont and JD_Cont are job creation and job destruction for continuing establishments respectively.

Given that the aggregates at any level are employment weighted, we consider both employment weighted and unweighted regression results. The employment-weighted regressions by construction yield that the mean of the dependent variable is equal to the appropriate employment weighted mean.

IV. Results

A. Basic facts about the evolution of retail sector in the D.C. area.

We find that the retail trade establishments in our set of D.C. metro counties largely mirrors national trends in retail trade. Figure 1 shows the employment growth for the four types of retail establishments over the 1976 to 2005 period. The rapid employment growth for the large chains and Big-Box beginning in the early nineties is striking. Interestingly, judging solely from Figure 1, it does not appear that single unit Mom-and-Pop stores bore the brunt of the displacement effects of the larger chains' growth. Instead, the smaller chains retreated in the face of increased competition from the Big-Boxes and other large chains.

An important part of the patterns in Figure 1 is that they reflect the overall growth patterns for the D.C. metro area. That is, since D.C. was growing over this period, all types of establishments in retail trade were likely to grow. Figure 2 presents the shares of employment by establishment type by year. It is evident in Figure 2 that there has been substitution away from smaller chain store employment towards Big-Box and large chain employment. There is a modest overall downward trend in the share of single unit employment especially through 1996 but interestingly single unit establishments recovered some of their share in the 1996-2005 period.

In the empirical analysis that follows, we exploit establishment-level employment growth rates and the components of establishment-level employment growth such as the job creation from entry and job destruction from exit. Figures 3-5 provide perspective on the aggregate trends in these measures by establishment type. Figure 3 depicts the net employment growth rates for the four retail establishment types. The rapid growth of Big-Box stores in the 1990's is clearly evident from the figure. Until 1990, the time series patterns for the four types moved in sync. Starting in 1990, however, the pattern for Big-Box stores begins to deviate. To a smaller degree the cycle for large chain stores runs counter to smaller chain chains and single location stores. Also of note is that during the 1999-2000 trough in large chain employment growth, single location stores show their strongest growth since the mid-80s when all store types exhibited robust growth. The figure shows that employment growth at single unit retailers remained on par with that at larger chain establishment, at least for the D.C. metro area.

Underlying the net employment growth rate patterns are large rates of gross job creation and destruction. For example, the average annual net growth rate of single unit establishments is 3.5 percent per year but this is accomplished by an average gross job creation rate of 22 percent and an average gross job destruction rate of 19.5 percent. Much, although not all, of these patterns are associated with establishment entry and exit. That is, as emphasized by FHK, in retail trade the extensive margin is a very important margin of adjustment. For single units, for example, establishment entry accounts for about 52 percent of job creation and 46 percent of job destruction on an annual basis.[9] These patterns make sense as once a store is created it becomes relatively more difficult to change its size and scope.

Given the importance of the extensive margin in retail trade, Figures 4 and 5 provide detail on the contribution of store openings and closings to job creation and destruction (computed as described above) respectively. In each case we see that for most years both job creation from new store openings and job destruction is highest for single location stores and lowest for Big-Box. This hierarchy has been noted by JKM and FHK. The obvious exception to this is the large spike in the job creation from the entry of new Big-Box stores in the 90's.

In what follows, we exploit both the spatial and the industry variation in the growth of Big-Box store employment. The variation that we are exploiting is within the D.C. area. As such, we always control for year effects so we are not using the variation at the aggregate D.C. area level depicted in Figures 1-5 to identify the effects of interest. In the next section, we discuss how we measure distance from Big-Box store activity and provide information about the variation across locations. Table 1 provides information about which major retail sub-sectors in the D.C. area have Big-Box activity[10] for two sub-periods. We report only through 2000 in Table 1 because we are using an SIC-based definition. In the analysis that follows, however, we use much more detailed industry (e.g., 4-digit SIC and 6-digit NAICS). However, Table 1 is useful to provide a broad overview. It is clear that in terms of major industry groups, Big-Box activity is concentrated in a relatively small number of major industry groups – building materials, hardware and garden supply; general merchandise; and home furnishings, furniture and equipment. There is also substantial growth of Big-Box activity in each of these broad categories (quantitative information about growth suppressed for disclosure reasons).

B. Empirical Specifications: Focusing on the Role of Distance Within the D.C. Metro Area.

As discussed above, existing evidence on the impact of Big-Box stores on retail employment comes from analyses using county level data to examine changes in payroll

[9] Appropriate caution is required in comparing the shares of job creation from entry and job destruction from exit on an annual basis and over a five-year horizon as in FHK.

[10] To avoid disclosure problems for some of the cells, we simply report information about existence of activity in the cell and whether employment in the cell grew over time.

and/or employment in the wake of entry of a particular chain such as Wal-Mart. Prominent papers in this literature use Wal-Mart's geographically based expansion strategy to construct instruments (planned opening dates in the case of Basker (2006) and distance from Bentonville, AR in the case of Neumark, Zhang and Ciccarella (2007) and Dube, Eidlin and Lester (2005)).

Rather than focusing on Wal-Mart per se or on between county variation across the nation, our approach focuses on the role of distance within the D.C. metro area. Our approach relies on detailed store level micro data geocoded to permit distance to play an explicit role and to permit more precise control for local retail market conditions. Given transportation costs faced by shoppers, the impact – positive or negative – of entry by a Big-Box store within in a large metro area should be localized. For example, in the D.C. area the entry of a Big-Box store in Prince George's County, MD shouldn't be expected to impact retail employment in Loudon County, VA (on the opposite side of the D.C. area) as much as retail trade activity in the immediate area in Prince George's County.[11]

We focus our attention on the impact of Big-Box stores on the dynamics of retail establishments operated by single location Mom-and-Pops and smaller chain stores. Our regression approach is to estimate the impact of changes in Big-Box employment along both the intensive and extensive margins on single unit and smaller chain store activity in the immediate area, controlling for common factors that impact retail trade activity in the immediate area. In the remainder of this section, we provide an overview of our estimation approach. In particular, we discuss the regression strategy for single units but note that we use the same approach for smaller chain store activity. The main regression specification is given by:

$$Y_{it} = \alpha_1 BB\_1\_mile\_same_{it-1} + \alpha_{15} BB\_1\_to\_5mile\_same_{it-1} + \alpha_{5to10} BB\_5\_to\_10mile\_same_{it-1}$$
$$+ \beta_1 BB\_1\_mile\_other_{it-1} + \alpha_{15} BB\_1\_to\_5mile\_other_{it-1} + \alpha_{5to10} BB\_5\_to\_10mile\_other_{it-1}$$
$$+ X_{it}^{'}\delta + \lambda Year_t + \varepsilon_{it}$$

where Y_{it} is the outcome variable of interest (either the employment growth rate of the establishment or a component of the single unit establishment growth rate such as the job destruction from exit), X_{it} is a vector of controls and $Year_t$ represents year effects. We describe the main variables of interest (e.g., $BB\_1\_mile\_same_{it-1}$) as well as the controls below.

We use the DHS growth rate measure described in the previous section for both LHS and RHS variables that accommodates including entering, exiting and continuing establishments. The primary explanatory variables of interest are measures of the (one

[11] The role of distance likely varies by type of product and sector. For example, retail customers might be more likely to travel extensive distances for certain types of durable goods (e.g., autos). As we discuss in the conclusion, an area for future research is to explore variation in the patterns by product and sector. We leave this for future research since this approach would be facilitated by constructing a database for multiple metropolitan areas.

year) lagged Big-Box activity in various concentric rings around each single unit store. Specifically, we consider the lagged growth rate of employment of Big-Boxes in the same sector and in other sectors in concentric 1 mile, 1 to 5 mile and 5 to 10 mile rings (so in the above expression we distinguish lagged Big-Box activity by distance and same/other sector). In like fashion, we also consider specifications where the Big-Box activity is a dummy variable indicating the first entry of Big-Box activity in the same sector and other sectors for the same concentric rings. For purposes, of defining sector, we use 4-digit SIC and 6-digit NAICS as appropriate. Our measure of the growth rate of Big-Box activity is the overall growth of Big-Box activity in the respective ring on an employment-weighted basis. In practice, it is the DHS growth rate of Big-Box activity in the respective rings. We use the one year lagged Big-Box growth for two related reasons. We are interested in the response of Single Unit and Smaller chain establishments to Big-Box growth and such a response likely takes some time. In addition, the time precedence helps in the identification of the response.

An example will aid interpretation of these measures. Consider a single unit hardware store in operation in both years t and t-1 that is located within 1 mile of general merchandise Big-Box with employment in years t, t-1 and t-2 of 100, 90 and 80, respectively, and is also located within 1 mile of a Big-Box home improvement store that entered in year t-1 with 100 employees. In year t, our single unit hardware store will have a lagged "same sector Big-Box within 1 mile" DHS growth rate of (100-0)/((100+0)/2)=2 (the value for entrants). Similarly, it will have a lagged "other sector Big-Box with 1 mile" DHS growth rate of (100-90)/((110+90)/2)=.10.

We estimate the regression specifications below using OLS. As such, the results should be interpreted appropriately as descriptive regressions providing quantitative information about how single unit (and in turn smaller chain) establishment growth and survival responds to changes in Big-Box activity in their local area. However, we include a rich set of controls in our descriptive regressions to control for local retail trade conditions as well as characteristics of establishments under investigation. These include year effects, local demographic, population and income characteristics and measures of the proximity of retail establishments to transportation infrastructure (the X matrix in the above equation includes all of these controls).

These controls are intended to, as fully as possible, soak up retail market conditions at the local level so that our Big-Box distance measures capture the impact of proximity to Big-Box stores. In addition, one can think about the other sector Big-Box variables as additional controls for the retail conditions at the local level. In other words, we think our identification of the impact of same sector effects of Big-Box activity is on stronger grounds than the other sector effects. For the latter, the effect we may be identifying is the combination of the effect of other sector Big-Box effects and effects of local retail conditions not captured by our other controls.

The detailed construction of our controls is described in the data appendix. We include the following to capture local demographic and income characteristics: quartiles of income in the 10-mile concentric ring, shares of the local population (10 mile ring) by

education class, age class, and gender.[12] We also include the growth rate of the population within a 5-mile ring. The population growth is intended to capture fast growing areas and the demographic effects to capture the characteristics and resources of local consumers. To measure proximity to transportation infrastructure, we compute the number interstate exits within 1, 5 and 10 miles for each single location and smaller chain store. This is included not only as a proxy for ease of access of shoppers, but to capture the location preferences of Big-Box stores that require many truck deliveries. We also compute the number of Metro (subway) stops within 1 mile. We assume shoppers traveling on Metro will not patronize establishment located far from a station.

For both single unit and smaller chain stores, we analyze two sets of stores in our descriptive regressions. Recall the dependent variable is the DHS growth rate, g_{it}, which is computed as the growth between period t-1 and t or a component of the DHS growth rate. This can be computed for stores active in period t-1, which we refer to here as incumbents, and births, stores not active in period t-1 but active in period t. Since much of the debate about the impact of Big-Box stores focuses on those retail businesses that existed in an area before a Big-Box enters, we run regressions that focus only on the incumbent single unit and smaller chain stores in addition to more general regressions that allow period t single unit and smaller chain entry. The incumbent only samples are also the relevant sample when we explore exit since it is only the incumbents that are at risk for exit.[13]

We also include establishment-specific controls for establishment age. While many factors impact growth, we have found that amongst the most robust patterns are for establishment age. We have found, for example, that young establishments exhibit a very high exit rate (and thus a very high job destruction rate from exit -- see, Haltiwanger, Jarmin and Miranda (2008)). As such, we include a dummy variable indicating whether the establishment is less than five years old. We also include a dummy variable indicating whether the establishment has a left censored age (was in existence in 1976).

Before proceeding to the empirical analysis, it is useful to emphasize that there is considerable variation across single unit (and smaller chain) stores in terms of their exposure to Big-Box stores. Table 2 presents the standard deviations of the key Big-Box employment growth measures for the within 1, 1 to 5 mile and 5 to 10 mile rings for both the incumbent only sample and the sample with incumbents and entrants. The unit of observation in Table 2 is single unit establishments in a given year in the respective

[12] For the demographic and income variables we use a 10 mile ring both because we think this is a reasonable area definition for the local population characteristics but also as described in the appendix our 10 mile concentric measures are more reliable for these measures. We use a 5 mile ring for population growth since we have more reliable and robust measures of overall population in smaller concentric rings.

[13] We don't consider analyses of entry of single unit or smaller chain stores separately since analysis of entry per se requires a different approach. That is, the at-risk group for entrants is potential entrants, which is difficult to measure (although see Dunne, Roberts and Klimek (2007) for an interesting way to measure potential entrants). Even though estimating the probability of entry by itself is not so straightforward; we do include analyses of the patterns of employment growth that include entering establishments. We regard these as asking the reasonable question: which establishments have more rapid growth (including the contribution of entry) given the presence of Big-Boxes?

samples. Consider, for example, the 1 to 5 mile ring. The standard deviation in same sector Big-Box employment growth rate in this concentric ring is 0.131 for the incumbent sample and the 0.153 in the incumbent plus birth sample. The standard deviation in other sector Big-Box employment growth rates in the respective rings is even larger. We also note that we observe substantial variation across single unit establishments in terms of the local population growth rate. The latter is obviously one important factor to control for in the analysis.

C. *Results for Single Unit "Mom-and-Pop" Stores*

Tables 3 through 7 report the results of the descriptive regressions for single units. Table 3 shows the results of regressions measuring the relationship between the growth in employment in Big-Box stores and employment growth at single unit (Mom-and-Pop) retailers. We report both unweighted and employment regressions for both the incumbent only and incumbents plus births samples. We report the main coefficients of interest as well as the impact of the controls that are especially important. For the latter, we find in virtually all of our specifications that local population growth is an important determinant of growth and survival of single units. We also find that establishment age is important as we suggested above. That is, we tend to find that incumbents have a high job destruction rate from exit that contributes to young incumbents on average having negative net employment growth. We find that when births are added to the sample, the coefficient on establishment age becomes positive reflecting the contribution of births (who are by construction less than five years old) to the average growth of young establishments. For the other non-reported controls, we find sensible patterns with, for example, higher income areas experiencing greater growth and survival.

Our main coefficients of interest are those measuring the growth of same and other sector Big-Box employment broken out by distance. For the "other sector" Big-Box growth we find mixed results on single unit employment growth and survival – in some cases a positive and significant effect (see, for example, the 1 to 5 mile ring effects in the incumbent sample in Table 3 and the impact on exit in the same ring in Table 4) but more often a small and insignificant effect in others. In that respect, we find no evidence of an impact of other sector Big-Box employment growth with some modest evidence of a positive effect. One interpretation is that these patterns suggest our other controls are doing a reasonable job of sweeping out local retail market conditions. The reason is that, at least in the absence of other controls, one could interpret the other sector Big-Box effects as essentially an indicator of local retail conditions. Under the latter interpretation, the effect should be positive throughout. We also note that these patterns suggest there is not much evidence that, holding local retail conditions constant, Big-Box employment growth in other sectors is a complement for single unit growth and survival. Nevertheless, caution is required in interpreting the other sector effects because the modest estimated effects may reflect potentially offsetting effects. For example, suppose that there is a negative direct effect of other sector Big-Box activity but positive indirect effect reflecting the extent to which other sector Big-Box activity is capturing local retail market conditions.

The estimated coefficients for "same sector" Big-Box growth generally indicate Big-Boxes have a negative impact on the growth of single unit retailers in the surrounding area.[14] The results are stronger for the weighted regressions suggesting that larger single units are more adversely affected by growth in Big-Box employment. Since the policy debate is typically about the jobs created and or destroyed, we have a preference for the employment-weighted results. However, the unweighted results are also of interest since they treat the smallest Mom-and-Pop stores equally with other single unit establishments. In other words, they provide insights into what happens at the Mom-and-Pop establishments regardless of size.

For the results with incumbents only (columns 1 and 2), the larger negative effects from same sector Big-Box effects tend to be closer in. The concentric ring with the most consistent negative effects is the 1-5 mile ring. For the 1-mile ring, the impact of same sector Big-Boxes is insignificant in the unweighted case but significant and quite negative and large in the weighted case. For the incumbents only specifications, the same sector Big-Box effects in the 5 to 10 mile ring yield an inconsistent pattern.

The magnitude of the negative effect of same sector Big-Box effects increases with the sample that includes both single unit incumbents and births. These patterns suggest that growth in Big-Box employment acts as a deterrent to entry for single unit retailers in the same sector. In this specification, the point estimates for same sector Big-Box growth effects are negative in all of the concentric rings with the largest effects in the 5-10 mile ring. Moreover, in almost all cases the effects are statistically significant.

In terms of the quantitative implications, the results in Table 3 suggest non-trivial effects. For example, a one standard deviation increase in lagged Big-Box employment in the same sector in the 1-5 mile ring yields about .7 percentage point decrease in single unit (weighted) employment growth in the incumbent sample and a 1.4 percentage point decrease in single unit (weighted) employment growth in the incumbent plus birth sample.

While Table 3 yields insights into the relationship between Big-Box and single unit employment growth, the regressions in Table 3 are quite general in that we're measuring the impact of changes in Big-Box employment on both the intensive and extensive margins on changes in single unit employment on both the intensive and extensive margins. Indeed, much of the popular debate about the impact of Big-Boxes focuses on the extensive margins of adjustment for Big-Boxes and incumbent single units. That is, does the entry of a Big-Box lead to the exit of single unit retail establishments? We now turn to this question but note that the analysis of this question focuses on the incumbents. The reason is that births are by construction not at risk of exit in the current period.

[14] As we have argued, we think the demographic, income and population variables control for key differences in local retail market conditions and in addition the same sector effects hold constant the impact of other sector Big-box activity. If there remains omitted factors in the residual that jointly influence single unit (or smaller chain store) growth and Big-Box growth in the local area, such factors should bias the estimated effect upwards. Thus, if anything, our negative estimated effects are an underestimate of the true negative effect.

In Tables 4 and 5, we report results of regressions focusing on changes along the extensive margin for single units and Big-Boxes, respectively. The dependent variable for Table 4 is job destruction from establishment exit (as described above) for incumbent single units.[15] In interpreting Table 4, it is important to note that the LHS moves in the opposite direction as the LHS in Table 3. Note, for example, that the mean of the LHS in Table 4 is 0.26 in the unweighted column and 0.09 in the weighted column. The former implies that the unweighted average job destruction rate from exit is 26 percent while the weighted average job destruction from exit is 9 percent. The finding that the job destruction from exit averages are almost exactly equal to but the opposite sign of the net employment growth rate for incumbents reflects the finding that, on average, most of the net adjustment among incumbents is on the extensive margin.

Results for Table 4 are broadly similar to those in Table 3; weighted results are stronger and "other sector" Big-Box growth is not a factor. We find that the growth of own-sector Big-Box employment is associated with an increase in job destruction from exit for single units. Again the most consistent findings are for the 1 to 5 mile ring. For our preferred employment weighted results, we now see a pattern of Big-Box impacts declining with distance from the single unit. That is, the rate of job destruction from establishment exit is greater for establishments within a mile of "same sector" Big-Box stores than it is for those more isolated from one or more Big-Boxes. For example, in the weighted results we find that same sector Big-Box growth in the 5-10 mile ring is associated with a decline in job destruction from exit.

In Table 5, we utilize the same dependent variable as in Table 3 but instead of the Big-Box employment growth rate we include a dummy variable indicating whether the prior year was the first year that a Big-Box (in the same sector or other sectors, respectively) entered the concentric ring in question. That is, we focus on the impact on single unit employment growth associated with changes in Big-Boxes on the extensive margin. As in Table 4, we see that focusing on change along the extensive margin yields a pattern of declining impact with distance. The entry of a Big-Box, even in the same sector, has no impact on t-1 to t employment growth for single unit retail establishment located in the 5 to 10 mile ring.

The specification in Table 5 lends itself to easily interpretable quantitative implications. The entry of a Big-Box store in the same sector yields a 16 percentage point decline in employment growth for single units in the weighted incumbent sample and almost a 25 percentage point employment growth decline for single units in the weighted, incumbent plus birth sample.

Table 6 shows results where we focus on extensive margin effects of Big-Boxes on the extensive margin (exit) effects of single units. Here we obtain that Big-Box new entry

[15] This LHS variable and specification is very similar to a standard exit equation with the LHS variable being a discrete variable that takes on a value of 1 if exit and 0 otherwise. We have estimated the latter using a linear probability model and obtain essentially the same patterns as in Table 3. As in the standard exit equation, the relevant establishments at risk for exit are the incumbents.

yields an increase in job destruction from exit for single units for close in rings. For the unweighted results, the results are significant in the 1 to 5 mile ring. For the weighted results, we obtain the pattern of Tables 4 and 5 with a declining impact the furthest the distance from the Big-Box entry. The quantitative implications of Table 6 are also readily interpretable. The first entry of a Big-Box in the same sector within a 1 mile ring of the incumbent single unit store yields an increase of about 21 percentage points in job destruction from exit in the preferred weighted employment results.

To further explore the patterns we have shown in this section, in Table 7 we report results where we use the specification from Table 4 but permit the effects of Big-Box employment growth on job destruction to vary over time. In particular, as illustrated in Figures 1 and 2 Big-Box employment growth accelerated in the 1990s. Table 7 shows the results when we permit the impact of lagged Big-Box employment growth to impact single units differently pre- and post-1992 (where post-1992 includes 1992).[16] The findings in Table 7 show that much of the impact of Big-Box employment growth on single unit job destruction from exit is a post-1992 phenomenon.

D. *Results for Smaller Chain Stores*

We now turn our attention to establishments that are part of smaller chain firms. We are interested in the impact of Big-Box employment on these types of establishments to compare and contrast with the single unit establishment results. In addition, recall that in Figures 1-5 we observe that the aggregate patterns suggest establishments from smaller chain stores have exhibited substantial employment loss and associated declining share. In some ways, the aggregate patterns suggest Big-Box activity is a more direct substitute for the type of retail trade activity offered by the establishments from smaller chain chains as opposed to Mom-and-Pop single unit establishments. Here we use the micro variation in the local area to investigate these relationships.

For this analysis, we focus on the analogues of Tables 3, 4 and 5.[17] That is, in Table 8 we present results showing the impact of same and other sector Big-Box employment growth on employment growth for establishments from smaller chain stores. Again, we present results for both an incumbent sample and a sample with incumbents plus births. Table 9 presents the results for job destruction from exit for the incumbent sample. In turn, Table 10 presents results on the impact of entry of Big-Box employment (the extensive margin) on employment growth of smaller chain establishments.

Before discussing the main results, it is useful to note that for the most part the other controls have similar qualitative effects in this setting but the estimates for some key controls tend to be smaller in magnitude and less significant. This can be seen by

[16] Note that the 1992 growth rate is the growth rate from 1991 to 1992 and the lagged growth rate for Big-Boxes in 1992 is the 1990 to 1991 growth rate so we are essentially breaking the sample into the pre and post 1990 growth patterns for Big-Boxes.

[17] One complicating but interesting issue that we do not consider for smaller chain stores is the extent to which the entry or growth of Big-Boxes (or the effects of other controls) has an impact on the propensity to yield adjustments of stores that are part of the same smaller chain. That is, to consider the unit of observation as the local firm rather than the establishment.

observing that local population growth is often not significant in Tables 8, 9, and 10. This may be because establishments for smaller chain stores are more likely serving a larger market than single unit establishments that have found a niche in the local market.

In Tables 8, 9 and 10, we find that growth in other sector Big-Box employment does not yield a consistent pattern and is mostly insignificant. Relative to the results for single unit establishments, we find even less of a relationship between other sector Big-Box employment growth and the employment growth of establishments from smaller chain stores.

In terms of same sector effects for establishments of smaller chain stores, we find a large, negative impact on growth and survival of lagged Big-Box employment growth in the 1-mile and 5 to 10 mile rings. We think the absence of much of an effect in the 1 to 5 mile ring likely reflects the spatial pattern of activity of such establishments, although verifying this would involve additional analysis. The largest impact, interestingly, is in the 1-mile ring. For example, Table 10 shows that for the weighted incumbent sample that the entry of a Big-Box in the same sector within a 1-mile ring yields a 34-percentage point decline in smaller chain establishment employment growth. In the incumbent plus birth sample, the effect is even larger with a 40-percentage point decline.

Compared to the results for single units, the findings suggest that small chains facing increasing Big-Box employment nearby take a big hit. Compare for example the first row of Table 10 to the first row of Table 5. The impact within 1 mile for smaller chain employment growth is about twice as large as it is for single units for the preferred weighted results.

V. Concluding Remarks

Our main findings are summarized as follows:

1. Within the D.C. metro area, the share of employment accounted for by Big-Box stores and larger chain stores has risen substantially at the expense of both single unit and especially smaller chain stores.
2. Much of the margin of adjustment of retail trade at the establishment level is via establishment entry or exit rather than changes in the scale of operations at the establishment level. This pattern is especially true for single unit and small chain stores.
3. The entry and growth of Big-Box stores has a substantial negative impact on employment growth and survival of single unit and smaller chain stores that operate in the same detailed industry as the Big-Box. This negative impact tends to decline with distance from the Big-Box. That is, the impact is largest if the single unit or smaller chain store is within 1 mile or 1 to 5 miles of the Big-Box store relative to being 5 to 10 miles from the Big-Box. These patterns are observed in regressions controlling for local retail conditions in the immediate area.

4. We find no systematic relationship between the entry and growth of Big-Box stores and the growth and survival of single unit and smaller chain stores that operate in a different detailed industry from the Big-Box. In that respect, we find no evidence of single unit and smaller chain stores operating in a different sector than the local Big-Box being complements or substitutes.

We think our findings are novel relative to the existing recent literature in that we are exploring the impact of Big-Box entry and growth within a metropolitan area and in particular exploit spatial variation within very narrowly defined geographic areas. A core message of our findings is that distance and sector both matter. Single unit and smaller chain stores in the same sector and close by location as recent Big-Box entry and growth take the biggest hit in terms of growth and survival.

While we think the findings are novel and interesting, the analysis here is very much a first step. For one, we look at only one metro area. Exploring additional metro areas is of interest not only to consider how robust our findings are to other areas but would permit richer investigation into the nature and mechanisms underlying the results. Additional areas would permit us, for example, to explore whether the same sector or other sector effects are being driven by specific sectors (and or put differently, how do the findings vary by sector?). One could easily imagine that Big-Boxes in the General Merchandise industry might have different same and other sector effects than Big-Boxes in the Building Supplies industry. We note however that constructing the data infrastructure that permits the type of analysis using detailed location information for many metro areas requires substantial work. While the LBD has the source information that permits detailed geocoding (i.e., latitude and longitude), the LBD has not yet been geocoded on a national basis. In addition, constructing all of the distance and controls in detailed geographic areas requires considerable time and computing resources as well. Hopefully, this paper is a step towards showing the payoff of exploiting spatial variation on such a detailed location basis.

Another limitation is that our findings are primarily descriptive. Developing and then ultimately estimating the underlying model that helps us understand and provide structure for characterizing how retail services in local communities is evolving is obviously needed.[18] Our novel descriptive findings should help provide the basic facts to guide the development of such models.

[18] See, e.g., the recent paper by Jia (2007) modeling the impact of discount stores on Mom-and-Pop stores.

References

Basker, Emek (2005), "Job Creation or Destruction? Labor Market Effects of Wal-Mart Expansion," *The Review of Economics and Statistics*, 87(1), 174-83.

Basker, Emek, Shawn Klimek, and Pham Hoang Van (2008), "Supersize It: The Growth of Retail Chains and the Rise of the "Big-Box" Retail Format", working paper.

Bresnahan, Timothy F & Reiss, Peter C, 1991. "Entry and Competition in Concentrated Markets," *Journal of Political Economy*, University of Chicago Press, vol. 99(5), pages 977-1009, October.

Chamberlain, Robert G., "What Is The Best Way To Calculate The Great Circle Distance (Which Deliberately Ignores Elevation Differences) Between Two Points?" http://www.census.gov/cgi-bin/geo/gisfaq?Q5.1. and http://www.cs.nyu.edu/visual/home/proj/tiger/gisfaq.html

Davidson, Sharon M., and Amy Rummel. 2000. "Retail Changes Associated with Wal-Mart's Entry into Maine." *International Journal of Retail & Distribution Management* 28(4/5), 162-9.

Davis, Steven, John Haltiwanger, and Scott Schuh (1996) *Job Creation and Destruction*. Cambridge MA: MIT Press.

Dinlersoz, Emin M. "Firm Organization and the Structure of Retail Markets" *Journal of Economics and Management Strategy*, Summer 2004, vol. *13*(2), pp. 207-240.

Doms, Mark E.; Jarmin, Ron S. and Klimek, Shawn D. "Information technology investment and firm performance in US retail trade." *Economics of Innovation and New Technology*, October 2004, vol. *13*(7), pp. 595-613.

Dube, Arindrajit, and Ken Jacobs. 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, Arindrajit, Barry Eidlin, and Bill Lester. 2005. "Impact of Wal-Mart Growth on Earnings throughout the Retail Sector in Urban and Rural Counties." Unpublished paper, University of California, Berkeley.

Dunne, Timothy, Shawn Klimek, Mark Roberts, and Xu. 2007. "The Dynamics of Market Structure and Market Size in Two Health Services Industries," Center of Economic Studies Working Paper CES 07-26.

Foster, Lucia, John Haltiwanger, and C.J. Krizan (2006a), "Market Selection, Reallocation and Restructuring in the U.S. Retail Trade Sector in the 1990s," *The Review of Economics and Statistics*.

Foster, Lucia, John Haltiwanger, and C.J. Krizan (2006b), "The Evolution of National Retail Chains: How We Got Here", working paper.

Haskel, Jonathan and Raffaella Sadun (2005), Entry, Exit and Labour Productivity in U.K.Retailing: Evidence from Micro Data," March.

Haltiwanger, John, Ron S. Jarmin, and Javier Miranda (2008), "Business Formation and Dynamics by Business Age: Results from the New Business Demography Statistics", working paper.

Holmes, Thomas J. "Bar Codes Lead to Frequent Deliveries and Superstores." *RAND Journal of Economics*, Winter 2001, vol. *34*(4), pp. 708-725.

Holmes, Thomas J. (2006), "The Diffusion of Wal-Mart and Economies of Density," May.

Jarmin, Ron S. and Javier Miranda (2002), "The Longitudinal Business Database," Center for Economic Studies Working Paper No. 02-17.

Jarmin, Ron S., Shawn Klimek, and Javier Miranda (2009), "The Role of Retail Chains: National, Regional, and Industry Results," in *Producer Dynamics* (eds. Dunne, Jensen and Roberts), Chicago, NBER/University of Chicago Press, pp. 237-270.

Jarmin, Ron S., Shawn Klimek, and Javier Miranda (2004), "Firm Entry and Exit in the U.S. Retail Sector, 1977-1997," Center for Economic Studies Working Paper No. 04-17.

Jia, Panle (2007), "What Happens When Wal-Mart Comes to Town: An Empirical Analysis of the Discount Retail Retailing Industry," working paper.

Lorch, B. (2006) "The evolution of a big-box landscape: a case study of the Winnipeg market" *Canadian Journal of Regional Science* 29 (3), 375 - 392

Lorch, B. (2005) "Auto-dependent Induced Shopping: Exploring the Relationship between Power Centre Morphology and Consumer Spatial Behaviour" *Canadian Journal of Urban Research* 14 (2), 364-383

Neumark, D., J. Zhang, and S. Ciccarella (2007), "The Effects of Wal-Mart on Local Labor Markets" IZA Discussion Paper Series No. 2545.

Schiff, Nathan (2009), "Cities and Product Variety in the Restaurant Industry," working paper.

W. Sinnott, "Virtues of the Haversine", Sky and Telescope, vol. 68, no. 2, 1984, p. 159

Tornqvist, L., P. Vartia, and Y. Vartia (1985), "How Should Relative Change Be Measured?", *American Statistician* 39 (February), 43-46.

Data Appendix

A. Big-Box Definitions:

Our first task in assembling our data was to determine which stores are Big-Boxes. There is not a standard definition but there are a variety of lists and sources that we used to develop our list. A good starting point is the Colombia University Graduate School of Architecture, Preservation and Planning's Web Site.[19] It identifies several key elements of Big-Boxes including size (50-200 thousand square feet), often rectangular (hence Big-Box label), with ample parking to facilitate access by shoppers who travel to the site by car.

We located several institutional studies of big-box stores effects on local communities that proved very useful in developing our working definition. For example, the Public Law Research Institute of the University of California's Hastings College of Law surveys studies performed by the states of California, Hawaii, and Maryland as well as the city of Los Angeles and the differences in their definitions prove to be as useful as the similarities.[20] For example, the state of California defines a Big-Box as any store measuring over 75,000 square feet in area. By contrast, the Maryland Department of Planning and the city of Los Angeles use 20-200,000 and of 60-130,000 square feet respectively. A Hawaii Legislative Reference Bureau study points out that using a measure of size alone is misleading because what constitutes a "big" store depends partly on what is being sold in it. That is, a 25,000 square foot bookstore seems small compared to a 120,000 sq ft Costco but it is very large compared to other bookstores. For this reason the University of California study recommends using a combination of factors to identify Big-Boxes. They are: size, industry, design, and stock diversity/depth.

Our work follows the spirit of this definition and uses information from Wikipedia and the National Retailer's Federation list of the top 100 U.S. retail firms as well. Wikipedia describe a Big-Box as a large, freestanding, rectangular, generally single-floor structure built on a concrete slab with floor space several times greater than traditional retailers in the sector. They also note that store sizes vary across geography and industry[21] and provide a link to a list of "Superstores" – another common term for Big-Boxes.[22] This list is our starting point. To it we add firms from the National Retail Federation's list of top 100 retailers that meet our working definition of a Big-Box.[23] Finally, we made a couple of additions to the list based on our knowledge in the D.C. retail market. While the sources we use for our list are in the public domain, we have not included our actual list of Big-Boxes to avoid any potential complementary disclosure issues. But suffice it to say that the Big-Box list we use are the well-known set of stores commonly thought of as Big-Boxes.

19 http://www.columbia.edu/itc/architecture/bass/newrochelle/extra/big_box.html
20 http://www.uchastings.edu/centers/state-local-gov/docs/plri_big_box_paper_04.pdf
21 http://en.wikipedia.org/wiki/Big-box_store
22 http://en.wikipedia.org/wiki/List_of_superstores
23 http://www.stores.org/pdf/08TOP100.pdf

B. MSA Definition:

A second fundamental concept we had to define is what counties we should include in our definition of the Washington D.C. metro area. Since our data span three decades and the official definition of the Washington MSA (and even what an MSA is) changed significantly over that time, we had to settle on a reasonable approximation of the region. We began by looking at how the list of counties in the MSA has changed during our timeframe in Table A.1

Table A.1

Component Name	When Present in Definition				
	2003	**1993**	**1983**	**1973**	**Working Definition**
District of Columbia	X	X	X	X	X
Calvert County, MD	X	X	X		(removed)
Charles County, MD	X	X	X	X	X
Frederick County, MD	X	X	X		X
Montgomery County, MD	X	X	X	X	X
Prince George's County, MD	X	X	X	X	X
Arlington County, VA	X	X	X	X	X
Clarke County, VA	X	X			
Culpeper County, VA		X			
Fairfax County, VA	X	X	X	X	X
Fauquier County, VA	X	X			
King George County, VA		X			
Loudoun County, VA	X	X	X	X	X
Prince William County, VA	X	X	X	X	X
Spotsylvania County, VA	X	X			
Stafford County, VA	X	X	X		X
Warren County, VA	X	X			
Alexandria city, VA	X	X	X	X	X
Fairfax city, VA	X	X	X	X	X
Falls Church city, VA	X	X	X	X	X
Fredericksburg city, VA	X	X			
Manassas city, VA	X	X	X		X
Manassas Park city, VA	X	X	X		X
Berkeley County, WV		X			
Jefferson County, WV	X	X			

Table A.1 shows us that there have been additions and subtractions to the list of counties in the D.C. MSA over time (for example Culpeper County VA and Berkeley County WV). Because of this we decided to choose the counties most consistently present but that also were not too far away from the District of Colombia. Our final set of counties is

noted in the column titled "working definition". As will be discussed later, we dropped Calvert County MD because it was not covered by Census Tracts in 1980.

C. Establishment and Firm Data:

Having identified Big-Box firms and narrowed our geographic focus to a few counties, we selected our establishment data from the Longitudinal Business Database (LBD). The LBD is a research dataset constructed at the Center for Economic Studies containing the universe of all U.S. Business establishments with paid employees from 1976 to 2005 (Jarmin and Miranda (2002)). It is based on the administrative data in the Census Bureau's Business Register (BR) but improves those data in several key ways. For example, it contains a time-invariant establishment numeric identifier that allows us to link stores in the D.C. area over time. Also, it provides information on the establishments' most consistent industry code. Finally, it re-times establishment births to avoid clustering caused by administrative processing. From the LBD we selected all establishments with a "best" industry code within retail trade.[24] We also selected all those establishments in our target counties. We used the longitudinal nature of the LBD to fill-in missing industry and geographic information as necessary to avoid spurious entry and exit from our sample. The resulting sample of retail trade establishments yields about the same counts of employment and establishments as for County Business Patterns.

Although the LBD has been used extensively in other research projects and is relatively free of outliers, since our study focuses on a narrow industry in a small geography, did some additional checking for outliers. We edited the employment patterns of a very small number of observations that had unusually erratic employment growth and loss. For those few cases where employment changed by several orders of magnitude for a single year and then returned to earlier levels (or disappeared), we replaced the large (small) number in the series with the nearest adjacent value that had the same magnitude as the rest of the series.

We also used the BR for our analysis. The BR contains information on establishment name and address that we needed to geocode the establishments and to flag our Big-boxes. We assigned a latitude and longitude to each establishment in our sample by first selecting it's address from the BR and then using an algorithm in ARCGIS that first attempts to geocode the stores exact address. If the address cannot be found, we instructed ARCGIS to assign the latitude and longitude of the centroid of the establishment's zip code.

Both the BR and the LBD contain a numeric firm identifier but we used the business name information in the BR to flag Big-Box firms. We identified the Big-Box firms by choosing the identification number associated with the largest firm with the Big-Box name (or key parts of the name) in the Big-Boxes' specific industry. Once we had the

[24] Prior to 2002, we selected establishments with SIC codes from 52 to 59. From 2002 forward we chose those with NAICS codes between 44 and 45 as well as 722 (restaurants).

Big-box firm ids and the establishment geocodes from the BR, we linked the information back into the sample from the LBD.

Because previous studies on the retail trade industry have shown that the extent of the firms' geographic coverage is correlated with important aspects of firm behavior (see Foster, Haltiwanger and Krizan (2006b), Jarmin, Klimek and Miranda (2009)) we created similar categories of firm types. Single Unit firms (SUs) are those firms that have only one establishment. Small chain retail firms are those firms operating multiple establishments in 1 to 14 states and "Large" chain firms are those businesses operating in 15 or more states. All of the categories are mutually exclusive with the large chains, for example, those stores part of large chains that are not classified as Big-Boxes.

D. Demographic Data:

We use a wide range of demographic controls in our analysis. The data underlying these variables come from the 1980, 1990 and 2000 Decennial Census microdata (sample) files. Our basic data on population, education, age, and gender, etc. were constructed by tabulating the weighted person-level observations in the files to the Census Tract level and assigning the tract's latitude and longitude as the cells' location. Income was measured by calculating the quartiles of the region's income distribution each year and then totaling the number of households in the tract that fell in each quartile.

Once we had the tract-level data for 1980, 1990, and 2000 we calculated the distance between each tract and each establishment (see below for details) and summed the data into 1, 5, and 10 mile rings around each store. Next we imputed the levels for each variable in the non-censual years by calculating an annual inter-censual growth rate, counting the number of years since the last census and multiplying the number of years by the annual rate. For example, 1-mile population in 1984 was imputed as follows:

$$POP^1_{1984} = POP^1_{1980} + \left(\left(\frac{POP^1_{1990} - POP^1_{1980}}{\frac{POP^1_{1980} + POP^1_{1990}}{2}} \right) / 10 \right) * 4$$

Although conceptually simple, this was a somewhat difficult algorithm to apply because of a couple of the characteristics of census tracts. For example, census tracts not only change over time, they did not cover the entire country until 1990. Until then only a few metropolitan areas had tracts assigned to them and while coverage in the D.C. are was fairly complete by 1980, we dropped Calvert County MD from our analysis because it was not yet tracted. Another complication is that there are many tracts with a radius greater than 1 – or even 5 – miles. The result is that a non-trivial number of our establishments had missing 1 or 5 mile values for many of the demographic characteristics, particularly in 1980. Observations with missing values for the 1-mile 1980 population also had missing 1-mile population growth rates for the 1980-1990 period. We addressed these problems by imputing backwards from the 1990 value using

the 10-mile growth rate of the variable being imputed. This problem with missing 1 and 5 mile distance characteristics files made these characteristics files less robust than the 10 mile distance files. We use the latter in the analysis for demographic and income characteristics of the local area.

E. Highway Exits and Metro Stations:

The highway exit data was created by using the search feature on latlon.com which allows a user to point and click a particular point on a map to find the latitude and longitude of the point on the map. This allowed us to create a database of the location of current exits for the limited access highways in our target counties. The more difficult part of the exercise was to search through a series of highway maps from the 1970s and 1980s to assign opening years to the exits.

We obtained a dataset of Metro station latitude and longitudes from Matthew Graham (affiliation). As in the highway exits, we also researched their opening dates and include only those stations currently open in our ring totals.[25]

D. Measuring Distances:

Having geocoded our establishment, demographic, highway and Metro Station data, we compute the distances between elements using a variant of the Haversine formula (Sinnott (1984)).[26] Starting with the following terms:

dlon = longitude2 - longitude1
dlat = lattitude2 - lattude1
$a = (\sin(dlat/2))^2 + \cos(lat1) * \cos(lat2) * (\sin(dlon/2))^2$
$c = 2 * \arcsin(\min(1,\sqrt{a}))$

the distance between any two points on the earth is given by:

$d = R * c$

where R is the radius of the earth (3963 miles according to Chamberlin (1996)).[27]

[25] See http://www.wmata.com/about_metro/docs/history.pdf for and excellent concise history of the history of the D.C. region's Metro transit system.
[26] See http://www.census.gov/cgi-bin/geo/gisfaq?Q5.1 and http://www.cs.nyu.edu/visual/home/proj/tiger/gisfaq.html
[27] We checked the results of this formula by assuming a flat earth, using the Pythagorean Theorem and a correction for the distance between a degree of latitude or longitude at the D.C. region's latitude provided by the National Geospatial-Intelligence Agency's Web Site: http://www.nga.mil/MSISiteContent/StaticFiles/Calculators/degree.html . We found a difference of only 10 feet on average.

Figure 1: Retail Employment in D.C. Metro Counties

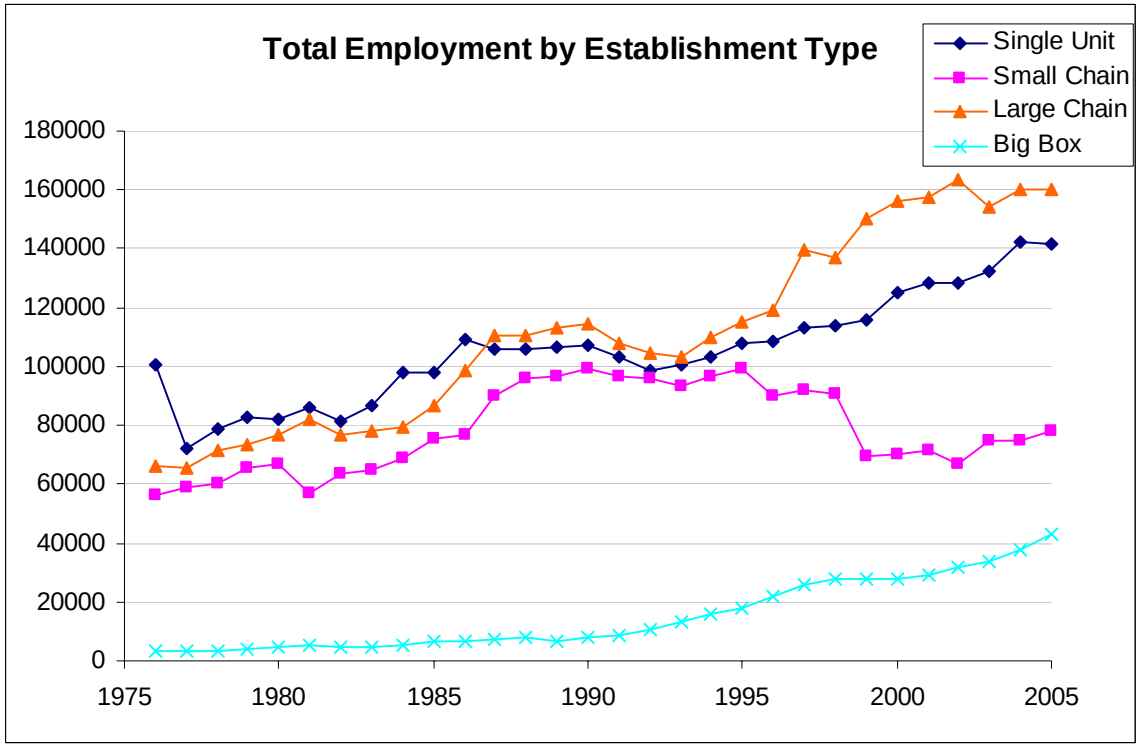

Figure 2: Shares of Retail Employment in D.C. Metro Counties

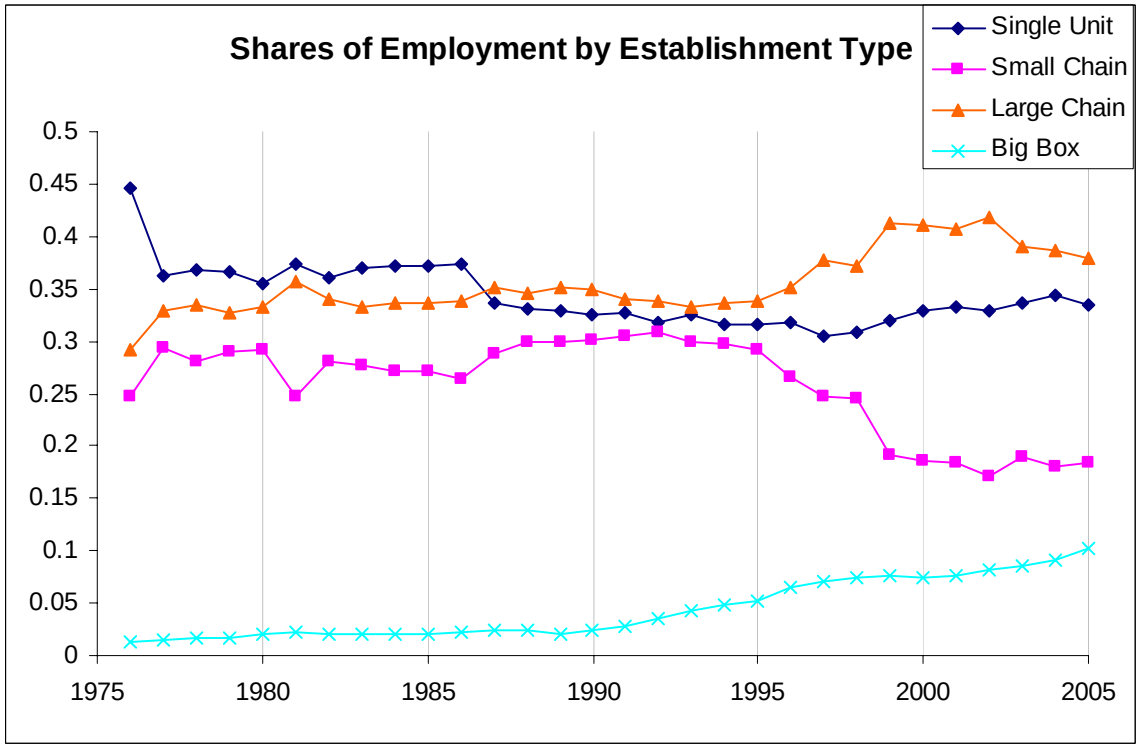

Figure 3: Net Employment Growth by Type

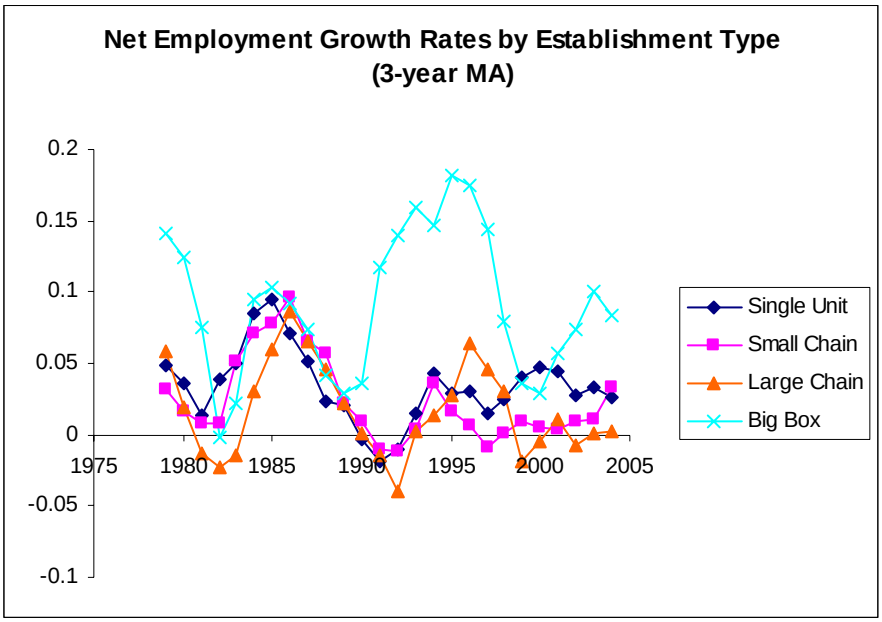

Figure 4: Job Creation from New Retail Stores by Type

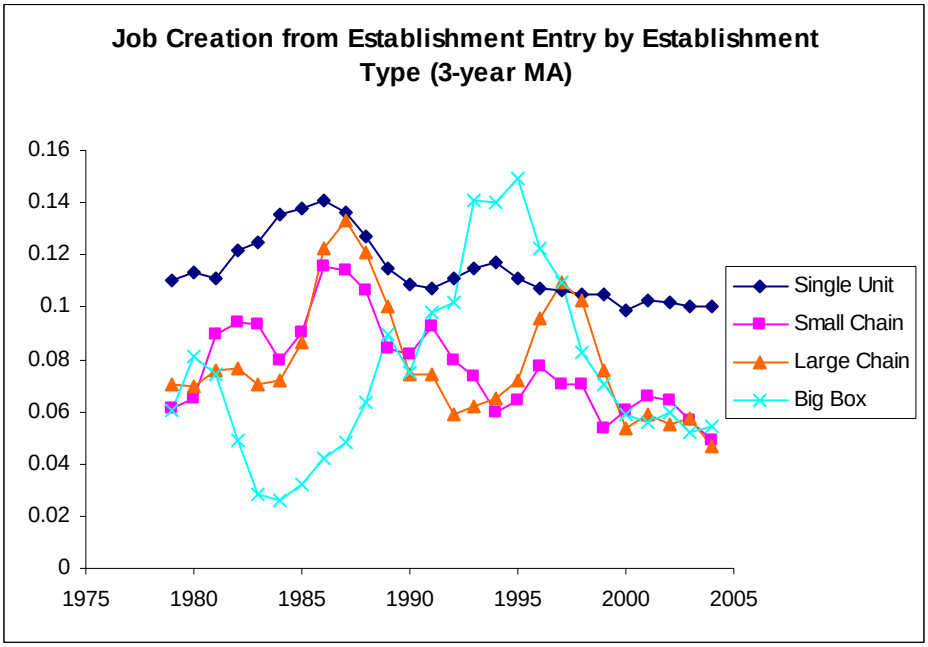

Figure 5: Job Destruction from Store Closings by Type

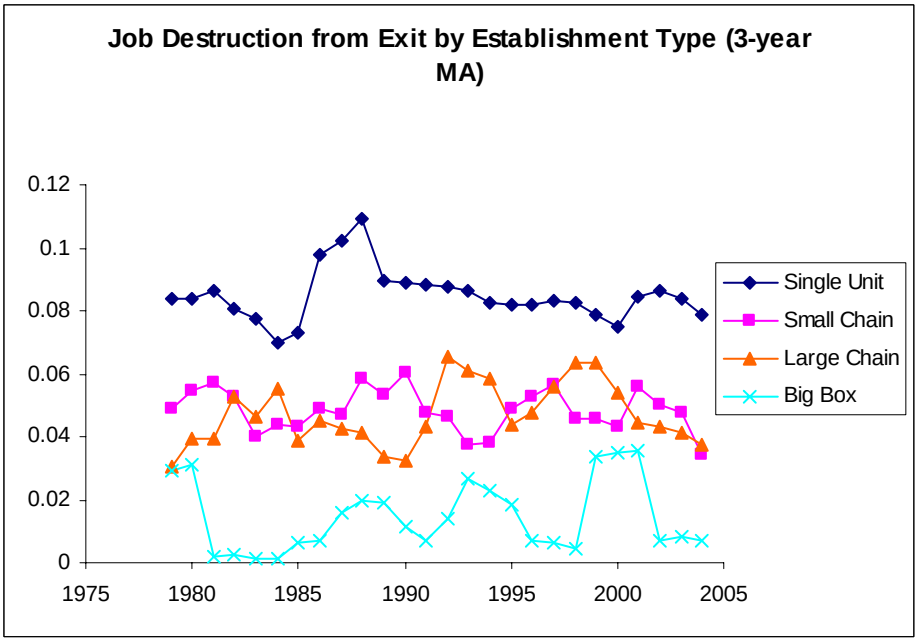

Table 1: Presence and Growth of Big-Box Activity by Broad Industry (2-digit SIC) in D.C. Metro Area

Broad SIC	1978-1991	1992-2000
Building Materials, Hardware, Garden Supply	Yes	Yes (+)
General Merchandise	Yes	Yes (+)
Food Stores	No	No
Automotive Dealers, Service Stations	Yes	Yes (+)
Apparel and Accessory	No	No
Home Furnishings, Furniture and Equipment	Yes	Yes (+)
Eating and Drinking	No	No
Miscellaneous	Yes	Yes (+)

Notes: A "Yes" in a column indicates the presence of Big-Box activity. A "+" or "-" in the second column indicates that Big-Box activity in that sector has expanded (contracted).

Table 2: Standard deviations of Selected Explanatory Variables

	Single Unit Incumbents in t-1	Single Unit Estabs with Emp>0 in t-1 or t
Lagged Growth in:		
Same Sector Big-Box Emp (1 mile)	0.057	0.062
Other Sector Big-Box Emp (1 mile)	0.482	0.482
Same Sector Big-Box Employment (1 to 5 mile)	0.131	0.153
Other Sector Big-Box Employment (1 to 5 mile)	0.253	0.254
Same Sector Big-Box Employment (5 to 10 mile)	0.159	0.193
Other Sector Big-Box Employment (5 to 10 mile)	0.210	0.212
Population Growth in 5 mile ring	0.032	0.033

Table 3: Relationship Between Growth in Big-Box Employment and SU Employment Growth

	SU Incumbents in t-1		SU Estabs with Emp>0 in t-1 or t	
Dependent Variable: Establishment net employment growth				
Explanatory Variables:				
Lagged Growth in				
Same Sector Big-Box Emp (1 mile)	0.008	-0.062	-0.025	-0.110
	0.027	*0.018*	*0.031*	*0.020*
Same Sector Big-Box Emp (1 to 5 mile)	-0.045	-0.050	-0.157	-0.108
	0.012	*0.010*	*0.013*	*0.011*
Same Sector Big-Box Emp (5 to 10 mile)	-0.005	0.023	-0.233	-0.161
	0.010	*0.009*	*0.010*	*0.009*
Other Sector Big-Box Emp (1 mile)	-0.004	-0.002	-0.002	-0.002
	0.003	*0.002*	*0.004*	*0.003*
Other Sector Big-Box Emp (1 to 5 mile)	0.014	0.018	-0.002	0.009
	0.007	*0.005*	*0.008*	*0.005*
Other Sector Big-Box Emp (5 to 10 mile)	-0.009	0.001	-0.001	0.003
	0.008	*0.006*	*0.010*	*0.007*
Population Growth in 5 mile ring	0.151	0.073	0.265	0.139
	0.054	*0.043*	*0.004*	*0.052*
Establishment less than 5 years old	-0.110	-0.046	0.529	0.291
	0.003	*0.002*	*0.004*	*0.003*
R-squared	0.01	0.01	0.07	0.04
Number of Observations	251,949	251,949	290,930	290,930
Mean of Dependent Variable	-0.26	-0.09	0.04	0.03
Weighted by Employment	No	Yes	No	Yes

Notes: All Specifications include year effects, local demographic and income controls, and controls for the number of highway exits within 1, 5 and 10 miles and the number of metro stops within 1 mile. Standard errors in italics.

Table 4: Relationship Between Growth in Big-Box Employment and Single Unit Job Destruction from Exit

Dependent Variable: Establishment JD from exit	SU Incumbents in t-1	
Explanatory Variables:		
Lagged Growth in		
Same Sector Big-Box Emp (1 mile)	0.015	0.086
	0.024	*0.014*
Same Sector Big-Box Employment (1 to 5 mile)	0.038	0.040
	0.011	*0.008*
Same Sector Big-Box Employment (5 to 10 mile)	-0.001	-0.027
	0.009	*0.007*
Other Sector Big-Box Emp (1 mile)	0.003	0.002
	0.003	*0.002*
Other Sector Big-Box Employment (1 to 5 mile)	-0.011	-0.009
	0.006	*0.003*
Other Sector Big-Box Employment (5 to 10 mile)	0.004	-0.011
	0.007	*0.004*
Population Growth in 5 mile ring	-0.178	-0.098
	0.047	*0.033*
Young Establishment Indicator (less than 5 years old)	0.152	0.081
	0.003	*0.002*
R-squared	0.01	0.01
Number of Observations	251,949	251,949
Mean of Dependent Variable	0.26	0.09
Weighted by Employment	No	Yes

Notes: All Specifications include year effects, local demographic and income controls, and controls for the number of highway exits within 1, 5 and 10 miles and the number of metro stops within 1 mile. Standard errors in italics

Table 5: Relationship Between Growth in Big-Box Employment and SU Employment Growth

	SU Incumbents in t-1		SU Estabs with Emp>0 in t-1 or t	
Dependent Variable: Establishment net employment growth				
Explanatory Variables:				
First entry of Big-Box in:				
Same sector within 1 mile (last year)	0.114	-0.159	0.059	-0.246
	0.073	*0.045*	*0.091*	*0.055*
Same sector in 1 to 5 mile ring (last year)	-0.069	-0.099	-0.060	-0.126
	0.034	*0.030*	*0.042*	*0.037*
Same sector in 5-10 mile ring (last year)	-0.040	-0.023	0.002	-0.007
	0.024	*0.017*	*0.029*	*0.021*
Other sector within 1 mile (last year)	0.005	-0.007	0.013	-0.001
	0.009	*0.006*	*0.011*	*0.007*
Other sector in 1 to 5 mile ring (last year)	0.037	0.059	-0.003	0.035
	0.023	*0.016*	*0.028*	*0.019*
Other sector in 5-10 mile ring (last year)	-0.036	-0.010	-0.053	-0.008
	0.031	*0.028*	*0.038*	*0.033*
Population Growth in 5 mile ring	0.151	0.072	0.182	0.136
	0.054	*0.043*	*0.004*	*0.052*
Young Establishment Indicator (<5 years old)	-0.110	-0.046	0.536	0.296
	0.003	*0.002*	*0.004*	*0.003*
R-squared	0.01	0.01	0.07	0.04
Number of Observations	251,949	251,949	290,930	290,930
Mean of Dependent Variable	-0.26	-0.09	0.04	0.04
Weighted by Employment	No	Yes	No	Yes

Notes: All Specifications include year effects, local demographic and income controls, and controls for the number of highway exits within 1, 5 and 10 miles and the number of metro stops within 1 mile. Standard errors in italics

Table 6: Relationship Between First Entry of Local Big-Box and Job Destruction from Exit

	SU Incumbents in t-1	
Dependent Variable: Establishment JD from exit		
Explanatory Variables:		
First entry of Big-Box in:		
Same sector within 1 mile (last year)	-0.075	0.209
	0.064	0.035
Same sector in 1 to 5 mile ring (last year)	0.079	0.116
	0.030	0.023
Same sector in 5-10 mile ring (last year)	0.035	-0.013
	0.021	0.013
Other sector within 1 mile (last year)	-0.007	0.000
	0.008	0.005
Other sector in 1 to 5 mile ring (last year)	-0.020	-0.008
	0.020	0.012
Other sector in 5-10 mile ring (last year)	0.046	0.046
	0.027	0.021
Population Growth in 5 mile ring	-0.178	-0.096
	0.047	0.033
Young Establishment Indicator (less than 5 years old)	0.152	0.081
	0.003	0.002
R-squared	0.01	0.01
Number of Observations	251949	251949
Mean of Dependent Variable	0.26	0.09
Weighted by Employment	No	Yes

Notes: All Specifications include year effects, local demographic and income controls, and controls for the number of highway exits within 1, 5 and 10 miles and the number of metro stops within 1 mile.

Table 7: Relationship Between Growth in Big-Box Employment and SU Job Destruction from Exit (Pre and Post 1992)

	SU Incumbents in t-1	
Dependent Variable: Establishment JD from Exit		
Explanatory variables:		
Lagged Growth in:		
Same Sector Big-Box Emp (1 mile)	-0.002	0.016
	0.049	*0.034*
Same Sector Big-Box Emp (1 mile) X POST 1992 dummy	0.028	0.093
	0.056	*0.037*
Same Sector Big-Box Employment (1 to 5 mile)	-0.010	-0.047
	0.022	*0.017*
Same Sector Big-Box Employment (1 to 5 mile) X POST 1992 dummy	0.064	0.114
	0.025	*0.019*
Same Sector Big-Box Employment (5 to 10 mile)	-0.009	-0.026
	0.016	*0.013*
Same Sector Big-Box Emp (1 mile) X POST 1992 dummy	0.028	0.093
	0.056	*0.037*
Other Sector Big-Box Emp (1 mile)	0.007	0.008
	0.005	*0.003*
Other Sector Big-Box Emp (1 mile) X POST 1992 dummy	-0.005	-0.010
	0.006	*0.004*
Other Sector Big-Box Employment (1 to 5 mile)	-0.008	-0.002
	0.008	*0.005*
Other Sector Big-Box Employment (1 to 5 mile) x Post 1992 dummy	-0.006	-0.013
	0.012	*0.007*
Other Sector Big-Box Employment (5 to 10 mile)	0.004	-0.005
	0.011	*0.007*
Other Sector Big-Box Employment (5 to 10 mile) x Post 1992 dummy	0.000	-0.009
	0.014	*0.009*
Population Growth in 5 mile ring	-0.178	-0.096
	0.047	*0.033*
Young Establishment Indicator (less than 5 years old)	0.152	0.081
	0.003	*0.002*
R-squared	0.01	0.01
Number of Observations	251949	251949
Mean of Dependent Variable	0.26	0.09
Weighted by Employment	No	Yes

Notes: All Specifications include year effects, local demographic and income controls, and controls for the number of highway exits within 1, 5 and 10 miles and the number of metro stops within 1 mile. Standard errors in italics.

Table 8: Relationship Between Growth in Big-Box Employment and Smaller Chain Store Employment Growth

	Small Chain Incumbents in t-1		Small Chain Estabs with Emp>0 in t-1 or t	
Dependent Variable: Establishment net employment growth				
Explanatory Variables:				
Lagged Growth in				
Same Sector Big-Box Emp (1 mile)	-0.115	-0.100	-0.146	-0.097
	0.025	*0.026*	*0.030*	*0.031*
Same Sector Big-Box Emp (1 to 5 mile)	-0.010	0.010	-0.038	-0.034
	0.014	*0.013*	*0.016*	*0.015*
Same Sector Big-Box Emp (5 to10 mile)	-0.053	-0.022	-0.131	-0.073
	0.011	*0.010*	*0.013*	*0.012*
Other Sector Big-Box Emp (1 mile)	-0.006	-0.004	-0.006	-0.005
	0.005	*0.003*	*0.005*	*0.004*
Other Sector Big-Box Emp (1 to 5 mile)	-0.014	0.005	0.006	-0.001
	0.009	*0.007*	*0.011*	*0.008*
Other Sector Big-Box Emp (5 to 10 mile)	0.005	-0.012	0.028	0.001
	0.011	*0.008*	*0.013*	*0.009*
Population Growth in 5 mile ring	0.027	0.125	0.004	0.110
	0.093	*0.068*	*0.006*	*0.076*
Young Establishment Indicator (less than 5 years old)	0.032	0.010	0.536	0.284
	0.005	*0.004*	*0.006*	*0.005*
R-squared	0.01	0.01	0.09	0.05
Number of Observations	83,770	83,770	96,689	96,689
Mean of Dependent Variable	-0.18	-0.06	0.01	0.02
Weighted by Employment	No	Yes	No	Yes

Notes: All Specifications include year effects, local demographic and income controls, controls for the number of highway exits within 1, 5 and 10 miles and the number of metro stops within 1 mile. Standard errors in italics.

Table 9: Relationship Between Growth in Big-Box Employment and Job Destruction from Exit of Smaller Chain Stores

Dependent Variable: Establishment JD from Exit	Small Chain Incumbents in t-1	
Explanatory Variables		
Lagged Growth in		
Same Sector Big-Box Emp (1 mile)	0.099	0.081
	0.020	*0.017*
Same Sector Big-Box Employment (1 to 5 mile)	0.013	-0.005
	0.012	*0.009*
Same Sector Big-Box Employment (5 to 10 mile)	0.045	0.011
	0.009	*0.007*
Other Sector Big-Box Emp (1 mile)	0.002	0.002
	0.004	*0.002*
Other Sector Big-Box Employment (1 to 5 mile)	0.005	0.002
	0.008	*0.004*
Other Sector Big-Box Employment (5 to 10 mile)	-0.006	-0.001
	0.009	*0.005*
Population Growth in 5 mile ring	-0.011	-0.030
	0.078	*0.046*
Young Establishment Indicator (less than 5 years old)	0.003	0.013
	0.004	*0.003*
R-squared	0.01	0.004
Number of Observations	83,770	83,770
Mean of Dependent Variable	0.17	0.05
Weighted by Employment	No	Yes

Notes: All Specifications include year effects, local demographic and income controls, and controls for the number of highway exits within 1, 5 and 10 miles and the number of metro stops within 1 mile. Standard errors in italics

Table 10: Relationship Between Growth in Big-Box Employment and Smaller Chain Employment Growth

	Small Chain Incumbents in t-1		Small Chain Estabs with Emp>0 in t-1 or t	
Dependent Variable: Establishment net employment growth				
Explanatory Variables:				
First Entry of:				
Big-Box in same sector within 1 mile (last year)	-0.270	-0.342	-0.341	-0.403
	0.063	0.068	0.078	0.083
Big-Box in same sector in 1 to 5 mile ring (last year)	-0.009	0.034	-0.043	0.024
	0.039	0.036	0.047	0.044
Big-Box in same sector in 5 to 10 mile ring (last year)	-0.012	-0.063	0.041	-0.015
	0.033	0.023	0.038	0.028
Big-Box in other sector within 1 mile (last year)	0.001	0.004	-0.007	-0.008
	0.013	0.009	0.015	0.011
Big-Box in other sector in 1 to 5 mile ring (last year)	-0.036	-0.007	0.018	-0.018
	0.033	0.023	0.038	0.027
Big-Box in other sector in 5 to10 mile ring (last year)	-0.160	-0.072	-0.108	-0.024
	0.033	0.032	0.040	0.038
Population Growth in 5 mile ring	0.025	0.124	-0.003	0.108
	0.093	0.068	0.006	0.076
Young Establishment Indicator (less than 5 years old)	0.032	0.011	0.540	0.285
	0.005	0.004	0.006	0.005
R-squared	0.01	0.01	0.09	0.05
Number of Observations	83770	83770	96689	96689
Mean of Dependent Variable	-0.18	-0.06	0.01	0.02
Weighted by Employment	No	Yes	No	Yes

Notes: All Specifications include year effects, local demographic and income controls, controls for the number of highway exits within 1, 5 and 10 miles and the number of metro stops within 1 mile. Standard errors in italics.