1 Introduction

In 2002, the retail sector in the U.S. accumulated $3,173 billion in sales and rivaled the manufacturing sector with a total employment of approximately 15 million workers. Currently, a dramatic transformation is reshaping the retail industry as stores differentiate across formats, pricing, and location. At the forefront of this change is the expansion of Wal-Mart. Over the past decade, Wal-Mart has grown from a modest, family-run business to the leading U.S. retailer with approximately $250 billion in revenues in 2002. Dubbed the "Beast from Bentonville", Wal-Mart's phenomenal growth has revolutionized retailing by offering a wide assortment of products at discount prices; every week, Wal-Mart's 4,750 stores attract nearly 138 million consumers, and an estimated 82% of U.S. consumers purchased at least one item from Wal-Mart in 2002.[1] Wal-Mart represents 9% of U.S. retail spending[2]. Its reach extends into almost every major U.S. consumer-products company; Wal-Mart is also "Hollywood's biggest outlet, accounting for 15% to 20% of all sales of CDs, videos, and DVDs."[3]

What attracts consumers to Wal-Mart? Does Wal-Mart maintain an advantage in theto almost 3 0 specialty, and online. For instance, the rise of e-commerce has added a new dimension to retail competition by reducing search and travel costs; Amazon.com has emerged as the leading online retailer by attracting $1.39 billion in sales during 2004.

Retail competition with Wal-Mart is an important public policy issue because of the information on the location and distance to nearby stores from the top 15 chains, using a chain's online store locator form and Yahoo! Yellow Pages. I also identify the local sales tax rate charged by each store based on its zip code and data from Tax Data Systems.

I estimate a consumer's choice of store among the top 15 chains, conditional on purchasing a DVD, through a discrete choice model that allows for unobserved heterogeneity in preferences for store types and disutility of travel. I find that stores of the same type compete more intensely and are closer substitutes than stores of differing types. A striking result is that conditional on price and distance, the average consumer still prefers Wal-Mart over most other stores; any advantage that Wal-Mart maintains over its competitors cannot be solely due to lower prices or increased proximity. This advantage cannot be wholly attributed to one-stop shopping, since the model controls for preferences over other mass merchants – such as Target and Kmart. The price and distance to the nearest Wal-Mart exerts the greatest influence on the market shares of Target and Kmart.

My simulation results indicate that the entry of 15 proposed Wal-Mart stores in California during 2004 increases the predicted probability of choosing Wal-Mart for the affected households within my sample by 27%. These proposed sites are often located in urban regions with several existing Wal-Mart stores in adjacent cities; the average decrease in distance to the nearest Wal-Mart store falls by 2.6 miles.

This paper is directly related to the literature on cross-channel competition and consumer choice over stores. Empirical work in this area has been limited due to the lack of rich data on consumer choices across retailers. Goolsbee (2001) examines competition between online and offline stores and finds that the cross-price elasticity is in excess of one; he concludes that online and offline stores are not separate markets. In addition, Goolsbee (2000) looks at whether taxes affect a consumer's decision to purchase a computer online versus offline. Forman, et al. (2006) examine how local competition, availability and selection of books, and prices affect a consumer's decision to purchase online versus offline, and Brynjolfsson et al. (2008) investigate how the number of local stores affect online and catalog demand. Ellison and Ellison (2006) also examine factors that drive a consumer's decision to purchase goods in-state instead of online. In contrast, I consider competition across a wide format of stores (not just offline versus online): mass merchants, video specialty, music, and online stores.

The next section contains a brief background of the video retail industry, followed by a description of my data. Then I proceed with a description of my demand model and estimation results. Finally, I describe the simulation exercise with Wal-Mart entry into Southern California.

2 The Home Video Industry

The home video industry consists of two segments: rentals and sales (also called sell-through). My paper focuses on the sell-through market for DVDs which generates the most revenues within home video retail. The leading trade group, the Video Software Dealer's Association, reported that sales revenues for VHS and DVD format totaled $12.1 billion in 2002, outweighing the $8.38 billion accumulated from rental revenues. Video Business Research estimated that DVD sales accounted for 72% of all sell-through revenues in 2002 and totaled $8.7 billion. In recent years, the increasing penetration of the DVD format into households has continued to fuel growth in the market for DVDs.

The .l-through an

video release date, genre, and theatrical box office revenues through the Titles Database from Adams Media Research. Since each househol sample consists of 3136 transactions that correspond to 2221 households with a complete set of demographic and purchase variables.

Tables 2, 3, and 4 provide some summary statistics. The demographics of the surveyed individuals resemble the overall U.S. population with the exception that they are slightly more well-educated. The purchased DVDs encompass a wide variety of films with varying box-office success in the theatrical market. Variation in prices exists across stores and videos; the average price paid for a DVD was $17.56 with a standard deviation of 4.12. The typical consumer did not have to travel far to purchase a DVD; the average distance to the closest and second closest stores were 2.5 miles and 4.4 miles.

The dataset provides a rich set of variables on household choices and location of neighboring stores. The one dimension for which it lacks information is the set of prices across all stores that a consumer may potentially visit. The dataset contains prices for each transaction, so I observe the price of the DVD at the actual store of purchase but not at other stores. For instance, I can observe that a consumer buys "Shrek" at Wal-Mart for $15, but I do not observe the price of "Shrek" at Best Buy, Kmart, or other stores that the consumers could have visited instead. I therefore construct estimates of prices that a consumer would face at each possible store. Taking the sample of all videos with observed prices, I regress the log of the price paid for each DVD on characteristics of the video, store, and location of purchase. Table 5 presents the results from this hedonic log-price regression. For each store in the consumer's choice set, I calculate the predicted log of price using the estimated coefficients.[5] Figure 1 graphs the ratio of

[5] Suppose the hedonic regression is given by $\ln p = X\beta + \varepsilon$ where N = number of observations and k = number of independent variables. Then the equation for price is given by $p = \exp(X\beta + \varepsilon) = \exp(X\beta)\exp(\varepsilon)$, and $E[p \mid X] = \exp(X\beta)E[\exp(\varepsilon)]$. If ε

the predicted price to the actual price for all transactions within my sample. The ratio lies between 0.8 to 1.2 for 80% of the transactions, and it has a mean of 1.02 and a standard deviation of 0.27. Some of the differences between the actual and predicted prices may be attributed to misreporting by certain individuals or “focal” responses whereby surveyed individuals give round figures.

4 Model of Demand for Store Choice

Estimating demand is the first step towards investigating consumers’ preferences over Wal-Mart and other retailers. Using the data described in the previous section, I estimate a discrete choice model where consumers choose among retailers, conditional on buying a DVD title. This mixed nested logit model is equivalent to a standard mixed logit model with random coefficients on the attributes of alternatives and dummies for each nest (Train, 2003). The utility of purchasing a DVD at store *j* will depend on the price of the DVD at store *j* and the distance to store *j* as well as other store and consumer characteristics. I specify a random coefficient on the distance variable to allow heterogeneity across the population in the disutility of traveling. In addition, I group the stores into five nests and allow a consumer’s unobservable taste for stores to be correlated for stores within the same nest; the five nests coincide with the five store types: online, mass merchant, video specialty, electronics, and music store. McFadden and Train (2000) mixed logit model with the appropriate choice of variables" and distribution of the random coefficient.

While alternative models exist for estimating demand, I chose the nested logit model because of its discrete choice framework and its flexibility as well as tractability in capturing consumers' substitution patterns. The alternative Almost Ideal Demand System assumes that consumers spends a fraction of their income at every store (Hausman and Leonard, 1997); moreover, since cross-price elasticities must be separately and directly estimated among all stores, the number of parameters to estimate can be quite large. Under the logit model, consumers choose exactly one of the 15 stores to make their purchase, and consumer preferences are mapped to the characteristics of each alternative (store) in their choice set. The tastes over these characteristics are used to derive the own- and cross-price elasticities. The nesting allows for flexibility in consumers' substitution patterns as alternatives within the same nest may be closer substitutes, and the nested logit model has the additional advantage of yielding a closed form expression for the purchase probabilities.[6] The nested logit model suits the data and question at hand by capturing richness in substitution patterns in a parsimonious way.

In the following sub-sections, I first describe the specific functional forms and distributional assumptions used to estimate the model, and then I briefly interpret the estimated parameters from the demand model.

4.1 Empirical Specification

Since I am interested in estimating substitution patterns across different stores, I define the relevant market as a geographic area and DVD title pair. Conditional on a purchasing a given

[6] In contrast, an analytic formula does not exist for the equivalent logit model with random coefficients on nest dummies.

DVD title, consumers choose which store to shop at. I condition on the particular DVD title and the decision to buy, so I may focus on substitution across different store types.[7]

Consumer *i*'s utility from traveling to store *j* location of the brick-and-mortar store. Online st decomposed into a component that is common among stores of the same nest and an independent term :

$$\varepsilon_{ijvmt} = \sum_{h=1}^{5} \zeta_{ih} TYPE_{jh} + \lambda \eta_{ijvmt} .$$

For instance, consumer *i* will have a common valuation for Amazon.com and Bestbuy.com given by $_{i,online}$, but in addition, she also has independent valuations $_{i,amazon}$ and $_{i,bestbuycom}$ that may differ for each store. The common valuation $_{i,online}$ induces a correlation between her unobserved tastes for each online store, $_{i,amazon}$ and $_{i,bestbuycom}$.

More specifically, I assume that the unobservable tastes for store types $_{ih}$ are independent and follow the unique distribution as described by Cardell (1997).[8] The distribution $_{ih}$ depends on a parameter zero. For each individual, I predict the probability of making her observed choice, and I estimate the model using Simulated Maximum Likelihood Estimation. Please refer to Appendix A for the details of the model and estimation.

4.2 Results

I now interpret the estimated coefficients of the benchmark demand model: the nesting parameter, consumers' tastes by demographics, disutility of distance, and travel costs. Then in the following Section 5, I will apply these results to directly investigate why people shop at Wal-Mart, and I will use the estimated demand parameters to perform a counterfactual simulation of Wal-Mart entry into California.

Table 6 reports the estimated utility parameters of my benchmark demand model. Tables 7 and 8 present the estimated price and distance elasticities. The estimated nesting parameter indicates that competition occurs more intensely among stores of the same type. The log-sum coefficient is 0.74 and statistically significant, indicating that a consumer's unobserved tastes for stores are correlated by store types; in other words, nesting by store types matters. Wal-Mart competes more intensely with mass merchants than stores of other types.

t the 88n otherettes for

Shopping patterns vary significantly by demographics - gender and the presence of children in the household. A consumer's education plays an important part in explaining her decision to shop at Wal-Mart and mass merchants in general. The omitted store type is the online dummy, so all coefficients on the interactions between household demographics and store types dummies must be interpreted relative to the online store option. The estimated coefficient of -0.67 on the interaction between the dummy for female and electronic store indicates that men have a higher marginal valuation of electronics stores (over online stores) relative to women. In addition, the presence of children is associated with a higher marginal utility for music stores relative to online stores.

A consumer's willingness to travel to Wal-Mart will depend upon her disutility of traveling. The demand model allows the marginal disutility of distance to vary by unobservable consumer characteristics (as captured by the random coefficient $_i$ on the distance variable) and observable consumer characteristics (as captured by interactions of the distance variable with dummies for income bracket and residence in an urban region). As shown in Table 6, the estimated mean (b) and standard deviation (s) of the log of the random coefficient on distance are -2.415 and 0.127. Very little unobserved variation exists in consumers' attitudes toward a8(TD0in consum)2i traveling, since I cannot reject the hypothesis that s = 0. Also, the random coefficient on distance is given by $\gamma_i = \exp(b + su_i)$, where u_i is distributed as a standard normal. Using the estimated parameters b and s, I calculate the mean of the coefficient on distance according to the formula: , where and disutility of distance. The magnitudes of the coefficients of relative sensitivity of consumers to distance and price. Table 9 reports the marginal cost of travel for high- and low-income consumers in rural and urban areas. The average low-income consumer in a rural area faces a marginal cost of 41 cents per mile while her counterpart in an urban area has a marginal cost of 76 cents per mile. Similarly, a high-income consumer experiences a higher marginal cost of travel of 63 cents and $1.37 in rural and urban areas. The marginal costs capture a consumer's implied value of time as well as any costs of transport, which the U.S. General Services Administration estimates as 31 cents per mile in a privately owned vehicle.[10]

5 Wal-Mart

Using the demand estimates from my benchmark model in the previous section, I now examine the nature of consumer demand for Wal-Mart. First, I consider whether Wal-Mart's advantages in the retail sector are due solely to lower prices or increased proximity. Next, to illustrate the degree of business-stealing among stores and the magnitude of consumer substitution, I use the estimated demand parameters to simulate a counterfactual experiment of Wal-Mart entry into California.

5.1 The Wal-Mart Advantage

To understand Wal-Mart's dominance in the retail sector, I first consider substitution patterns between Wal-Mart and other retailers. Tables 7 and 8 present the price and distance elasticities across all 15 stores in my sample. Wal-Mart competes most intensely in price with Kmart and Target and to a lesser extent with Sam's Club. If Wal-Mart decreases its price by 1%,

10

then the market shares of Kmart and Target fall by 1.69% and 1.57%. The distance elasticities exhibit a similar pattern to the price elasticities. If the distance to the nearest Wal-Mart decreases by 1% for all households, then the market shares of Kmart and Target decrease by 0.26% and 0.24%.

To quantify how consumer's value a shopping trip to Wal-Mart, I use my demand estimates to calculate a consumer's willingness to pay to shop at Wal-Mart. The estimated utility coefficients on store type and store dummies from the benchmark model imply that the average consumer prefers Wal-Mart to most other stores even conditional on price and distance. For instance, if all retailers charged the same price and were located in the same proximity, a consumer with "average" characteristics would still prefer to shop at Wal-Mart.

Several possible explanations exist for this finding. The average consumer's preference for Wal-Mart may reflect the convenience of one-stop shopping, the expectation of lower prices in other items in the consumer's shopping bundle, or an unobserved Wal-Mart "quality" effect. I investigate each possibility below.

First, differences in product assortment may account for why a consumer would prefer to shop at Wal-Mart (where they can purchase a variety of other goods in addition to DVDs) as opposed to Blockbuster (a video specialty store that mainly sells DVDs.) Recall that the benchmark model of demand contains store-type dummies, which can capture systematic differences in consumers' market baskets across different types of stores. However, product assortment cannot entirely account for the preference for Wal-Mart. Under the estimates from the benchmark model of model, consumers still prefer to shop at Wal-Mart even relative to other mass merchants that offer one-stop shopping (e.g., Target, K-Mart).

under the age of 18, a college education, lives in an urban area, and income of $40,000) favors Wal-Mart over all other mass merchants; he is willing to pay $7.09, $4.70, $2.61, and $2.37 to shop at Wal-Mart instead of Kmart, Sam's Club, Costco, and Target for a $15 DVD, assuming both stores are located 5 miles away. His female counterpart also values Wal-Mart over other mass merchants; she would be willing to pay $6.56, $4.17, $2.09, and $1.85 to shop at Wal-Mart instead of Kmart, Sam's Club, Costco, and Target. In contrast, individuals with above average age or education levels experience a lower utility of shopping at Wal-Mart; a 55-year old male with kids under the age of 18, a graduate school education, lives in an urban area, and income above $75,000 would actually prefer to shop at Target instead of Wal-Mart, and he is willing to pay $1.74 to do so.

Finally, this striking result suggests that Wal-Mart's advantage might not solely be due to lower prices and increased proximity. A Wal-Mart "quality" effect still persists even when we allow for a Wal-Mart specific effect on prices, distance, and all other explanatory variables.[11]

5.2 Simulation of Wal-Mart Entry into California

As previously discussed, the estimated demand coefficients indicate a strong preference for Wal-Mart by the average consumer, even conditional on price and distance. To quantify the magnitude of this preference, I examine a particular public policy issue of Wal-Mart's entry into California.

In 2004, Wal-Mart announced its intention to open 40 more store sites as part of its aggressive expansion plans into California, particularly in the Southern California region. Previous attempts to construct new store sites have met with "intensifying grassroots

[11] Wal-Mart may not raise its price due to advantages in cost. Wal-Mart's cost advantages stem from low labor costs and the retail chain's logistics and distribution innovations (Emek, 2005b).

opposition", and many agree that Wal-Mart's "biggest barrier to growth is ... opposition at the local level".[12] In 2003, a fierce struggle ensued in Contra Costa County near San Francisco, as Wal-Mart collected signatures to compel a referendum over its entry. Wal-Mart has also met staunch local resistance at other California cities such as West Covina, Oakland, Bakersfield, and Inglewood by local merchants and labor unions. The United Food and Commercial Workers union has been a long-time opponent of the chain, and in 2003, it organized campaigns against Wal-Mart in 45 locations across the U.S.

The business-stealing effects of Wal-Mart are a hotly debated topic as Wal-Mart looks to expand its presence in California. Target and Kmart have already situated 184 and 163 stores within California, and as Wal-Mart's closest competitors, they stand to suffer from the entry of Wal-Mart. I simulate the effects of entry of Wal-Mart at 15 store sites in California, which include 10 new stores constructed in 2004, 3 proposed store sites that were rejected by city votes (Inglewood, West Covina, and Oakland), and 2 proposed store sites that were approved by the city (Palm Springs and Rosemead). Table 13 lists each city and the corresponding zip code used for the simulation. As seen in Figure 2, the majority of these sites are located in Southern California.

A total of 37 households, that comprise slightly over 1% of my sample, are affected by the entry of these 15 new stores, and the average change in distance to the nearest Wal-Mart was 2.6 miles. I simulate the predicted probability of choosing each store before and after the entry of the 15 Wal-Mart stores. Table 14 reports the estimates and standard errors for the average predicted probability of choosing each store for the 37 households before and after the entry of Wal-Mart, and the table also shows the average change and percentage change in the predicted probabilities. The average change in probability of choosing Wal-Mart increased by 5.92

[12] Business Week Online, "Is Wal-Mart Too Powerful?", October 6, 2003.

percentage points, which accounted for 27% increased probability, and the average change in

I find that substitution occurs proportionately more among stores of the same type. For instance, a change in the price or distance to a Wal-Mart store has the largest impact on the market shares of Target and Kmart. A striking result is that even conditional on the price of a DVD and distance, the average consumer still prefers to shop at Wal-Mart over most other stores. This result remains even after allowing for a Wal-Mart specific effect in my demand model, and it suggests that Wal-Mart's dominant market share may not be due solely to low prices and location. This preference cannot wholly be attributed to Wal-Mart's one-stop shopping convenience, since the average consumer prefers Wal-Mart even relative to other mass merchants, such as Kmart and Target.

To capture the magnitude of consumers' prefer

Research Council (Schultze and Mackie, 2002) supports the underlying assumption by the Bureau of Labor Statistics that stores such as Wal-Mart may not have a lower “service-adjusted” price. However, my results suggest the contrary: even conditional on store and consumer characteristics, Wal-Mart appears to be a desirable place to shop relative to most other stores for the average consumer. In fact, if Blockbuster Video can be thought of as the “traditional” place to purchase a video while Wal-Mart is the “new” discount retailer, then my results imply that an “average” 35-year old female who lives in an urban area and has a college education and children under the age of 18 is willing to pay $6.06 to shop at Wal-Mart instead of Blockbuster Video (for a $15 DVD if both stores are 5 miles away.)

Appendix A: Details of Demand Model and Estimation

A.1 Model

Following Berry, Levinsohn, and Pakes (1995), I model a consumer's choice of store as a function of store and consumer characteristics while allowing for unobserved heterogeneity in preferences over store characteristics and correlation in tastes among store of the same type. Consumer *i*'s utility from traveling to store *j* is given by:

$$U_{ij} = U(z_j, h_i, d_{ij}, p_j, \xi_j, \omega_i, \varepsilon_{ij}, \theta)$$

where z_j is a vector of observable store characteristics, h_i is a vector of observable consumer characteristics, d_{ij} is the distance to store *j* for consumer i, p_j is the price at store *j*, $_j$ captures any unobserved characteristics of store *j*, $_i$ is a vector of unobserved characteristics of consumer *i*, $_{ij}$ is individual *i*'s idiosyncratic taste for store *j*, and is a vector of parameters to be estimated.

The terms and capture the two sources of unobserved heterogeneity in consumer preferences over store types. Interactions of the unobservable consumer characteristics and observable store characteristics *z* allow tastes for store characteristics to differ among the population in unobservable ways. Furthermore, specifying an error structure that allows for correlations in the idiosyncratic taste over particular stores generates more flexible substitution patterns.

Each consumer will choose the store that maxi where *k* indexes all possible stores in consumer *i*'s choice set. If   has distribution $f_1($ $)$ and   has distribution $f_2($ $)$, then the probability of consumer *i* choosing store *j* is:

$$P_j(h_i) = \int_{\varepsilon \in A_{ij}} f_1(\varepsilon) \; f_2(\omega) \; d\omega \, d\varepsilon \; .$$

To obtain the market shares of the stores, I need to integrate

$$ {}_{ij}(\ {}_{i}) = \frac{\exp(\ {}_{ij}\ {}_{i}/\ )\left(\sum_{k\in TYPE}\exp(\ {}_{ik}\ {}_{i}/\ )\right)^{-1}}{\sum_{h=1}^{5}\left(\sum_{k\in TYPE}\exp(\ {}_{ik}\ {}_{i}/\ )\right)} $$

In general, a mixed nested logit model relaxes the Independence of Irrelevant Alternatives (IIA) assumption among alternatives in of the parameters with respect to the number of draws, since models may appear identified at lower numbers of draws when they are in fact not. The parameter estimates and standard errors were stable with respect to different start values and to 200, 1000, and 4000 Halton or random draws.[15]

A.3 Unobserved Consumer and Store Characteristics

In the benchmark demand model, consumer tastes are correlated among stores of the same type in unobserved ways. The model also allows consumers to have an unobserved taste over distance and traveling. The store fixed effects capture a store's unobserved quality that is fixed over time.

One concern is that additional unobserved characteristics (not captured by the store dummies) may still exist and be correlated with price. I conduct a series of checks to implicitly test for the magnitude of any endogeneity bias.[16] First, I examine whether the estimates from my benchmark model of demand suffer from the classic symptoms of endogeneity bias. Then, I consider a direct extension to my structural model to check for the extent of any endogeneity bias.

[15] I tried more general specifications of the mixed logit model, e.g., a full correlation matrix for idiosyncratic tastes across store types, but the estimates were not stable with respect to the number of draws. For the creation of my

First, the results of my benchmark model do not appear to exhibit the classic symptoms of endogeneity bias. Although the model contains store dummies which control for aspects of (unobserved) store quality that are constant over time, any time-varying unobserved quality of a store could be correlated with price. A classic symptom of not accounting for this correlation and endogeneity is an upward-sloping demand curve. With this endogeneity bias, demand estimates and elasticities may mistakenly indicate that consumers prefer higher pric where *X* contains demographics and store characteristics, *p* is the price at store *j* at time *t*, *d* is the distance of consumer *i* to store *j*, ξ is an unobserved store quality that may vary over time, and is an idiosyncratic error term. If store quality does not vary over time, then $_{jt}$ = $_{j}$ for all *t*. Including store dummies in the utility specification will deal with the endogeneity problem (Nevo, 2000). The benchmark model of demand includes store fixed effects.

However, if store quality varies over time, then we can decompose the unobserved store quality into two components:

$$\xi_{jt} = \xi_j + \Delta\xi_{jt} .$$

where $_{j}$ is the component of quality that does not vary over time, and Δ $_{jt}$ is the component that varies over time (Nevo, 2000). Endogeneity bias can arise through correlation changes in store quality over time Δ $_{jt}$ and variables such as price *p* and distance *d*. While including interactions of store and weekly dummies would control for the endogeneity, this requires a large number of parameters to be estimated, which

Since I use an estimate of the price variable in the utility specification, I need to adjust the standard errors of the demand coefficients to account for noise in the price estimates obtained in the first step. I employ the following procedure: I bootstrap the price regression 100 times. If N denotes the number of observations in the price dataset, then each bootstrapped sample consists of N observations drawn with replacement from the price data. For each bootstrapped sample, I re-estimate the price regression, use the results to calculate the estimates of price for each store in the consumer's choice set, and re-estimate the mixed nested logit model with the new price estimates. I add the variance in parameter estimates over the bootstrapped price samples to the variance in estimates from the original dataset. The standard errors were calculated using the BHHH approximation to the Hessian with a numeric gradient. The bootstrap procedure produces a valid correction for the standard errors if the moment conditions from the price regression and the demand estimation are orthogonal (Newey, 1984). This is a plausible assumption, since my sample consists of individuals from several different markets dispersed across the U.S. A sampled individual's demand comprises a very small portion of the aggregate demand in each market and very little influence on market price.

References

Adler, T. and M. Ben-Akiva (1976): "Joint-Choice Model for Frequency, Destination and Travel Mode for Shopping Trips," *Transportation Research Record*, No. 569, pp. 136-150.

Basker, E. (2005a): "Job Creation of Destruction? Labor-Market Effects of Wal-Mart Expansion," *Review of Economics and Statistics*, Vol. 87, pp. 174-183.

Basker, E. (2005b): "Selling a Cheaper Mousetrap: Entry and Competition in the Retail Sector," *Journal of Urban Economics*, Vol. 58, pp. 203-229.

Basker, E. and M. Noel (2007): "The Evolving Food Chain: Competitive Effects of Wal-Mart's Entry into the Supermarket Industry," University of Missouri Department of Economics Working Paper 07-12.

Ben-Akiva, M., D. Bolduc, and J. Walker (2001): "Specification, Estimation, and Identification of the Logit Kernel (or Continuous Mixed Logit) Model," mimeo, Department of Civil Engineering, MIT.

Berndt, E., B. Hall, R. Hall, and J. Hausman (1974): "Estimation and Inference in Nonlinear Structural Models," *Annals of Economics and Social Measurement*, Vol. 3/4, pp. 653-665.

Berry, S., J. Levinsohn, and A. Pakes (1995): "Automobile Prices in Market Equilibrium," *Econometrica*, Vol. 63, no. 4, pp. 841-890.

Bhat, C. (2001): "Quasi-random Maximum Simulated Likelihood Estimation of the Mixed Multinomial Logit Model," *Transportation Research B*, Vol. 35, pp. 677-693.

Brynjolfsson, E., J.Y. Hu, and M.S. Rahman (2008): "Battle of the Retail Channels: How Product Selection and Geography Drive Cross-Channel Competition", mimeo.

Cardell, N.S. (1997): "Variance Components Structures for the Extreme Value and Logistic Distributions with Application to Models of Heterogeneity," *Econometric Theory*, Vol. 13, pp. 185-213.

Chevalier, J. and A. Goolsbee (2003): "Price Competition Online: Amazon versus Barnes and Noble," *Quantitative Marketing and Economics*, Vol. 1, no. 2, pp. 203-222.

Chintagunta, P., J. Dube, and K. Goh (2005): "Beyond the Endogeneity Bias: the Effect of Unmeasured Brand Characteristics on Household-level Brand Choice Models," *Management Science*, Vol. 51, pp. 832-849.

Chiou, L. and J. Walker (2007): "Masking Identification of Discrete Choice Models under Simulation Methods," *Journal of Econometrics*, Vol. 141, pp. 683-703.

Davis, P. (2006): “Spatial Competition in Retail Markets: Movie Theaters,” *RAND Journal of Economics*, forthcoming.

Ellison, G. and S. Fisher Ellison (2005): “Search, Obfuscation, and Price Elasticities on Internet,” mimeo, MIT.

Ellison, G. and S. Fisher Ellison (2006): “Internet Retail Demand: Taxes, Geography, and Online-Offline Competition,” mimeo, MIT.

Forman, C., A. Ghose, and A. Goldfarb (2006): “Geography and Electronic Commerce: Measuring Convenience, Selection, and Price,” mimeo, Tepper School of Business Working Paper #2006-E95, Carnegie Mellon University.

Franklin, A. (2001): “The Impact of Wal-Mart Supercenters on Supermarket Concentration in U.S. Metropolitan Areas,” *Agribusiness*, Vol. 17, no. 1, pp. 105-114.

Goolsbee, A. (2000): “In a World Without Borders: The Impact of Taxes on Internet Commerce,” *The Quarterly Journal of Economics,* Vol. 115, no. 2, pp. 561-576.

Goolsbee, A. (2001): “Competition in the Computer Industry: Online versus Retail,” *The Journal of Industrial Economics*, Vol. 49, n. 4, pp. 487-499.

Massachusetts Institute of Technology Center for Real Estate, <http://web.mit.edu/cre/students/curriculum/courses/11.433 /fall04/Lecture6.ppt>, accessed 10/30/2004.

McFadden, D. (1981): “Econometric Models of Probabilistic Choice” in *Structural Analysis of*

Weisbrod, G., R. Parcells, and C. Kern (1984): "A Disaggregate Model for Predicting Shopping Area Market Attraction," *Journal of Retailing*, Vol. 60, No. 1, pp. 65-83.

	Price
Market share	

Market share	Distance

i j *i* *j* *i*

j