The Costs of Growth:

Estimating Entry Costs During Rollouts

Mauricio J Varela*

January 2011

Abstract

Retailers typically expand gradually, rather than all at once, into new markets. Policy that aims to accelerate growth requires distinguishing among the different factors limiting growth. I propose and estimate a model that quantifies alternative explanations for gradual growth, where one such explanation is diseconomies of scale in the opening of new outlets. Such diseconomies of scale may arise from scarce managerial talent and/or lack of capital. Alternative explanations allowed in the model include demand changes, decreasing cost of capital, intramarket dynamics (e.g. learning), intramarket chain economies, and changes in competition. An entry game is modeled, in which players choose how many stores to open and close and in which markets to do so. The diseconomies of scale in store opening cause entry decisions to be interdependent across markets. A variant of Maximum Simulated Likelihood is used to estimate the game with these interdependencies. I estimate the model using entry patterns in the Mexican supermarket industry during the early 2000s. I find that WalMart's per-store entry costs increased by 2% for each additional store opened. I also find that a firm whose financial performance improves by 10% enjoys a 9% reduction in per-store entry costs. Business-stealing effects and cannibalization effects are similar to those found in previous work.

1 Introduction

When expanding its roster of outlets across geographic markets, retailers typically do so gradually rather than all at once. WalMart's and Starbucks' expansions are two textbook examples of

*Kellogg School of Management, Northwestern University (email: m-varela@kellogg.northwestern.edu). I am very grateful to Mike Mazzeo, Kate Ho, Aviv Nevo, Paul Greico, and Guy Arie for their useful comments. All errors are my own.

this gradual growth processes, of "rollout of stores". Although much theoretical work exists[1] on plausible causes for such behavior, there are few empirical studies of this phenomenon. In this paper I propose an empirical model that evaluates alternative explanations for the gradual growth process. In particular, the model allows for diseconomies of scale in the expansion rate to be one explanation for gradual growth. Diseconomies of scale in the expansion rate may originate from limited resources, either capital or labor. For example, in retail, firms need capital to fund new stores and managers to open the new stores efficiently. Both of these resources may be limited, albeit renewable: positive cash flows replenish capital and managers are free to work on new stores upon finishing the opening of a previous store. As such, the rate at which a firm can grow may be determined, in part, by the market for these renewable resources.

The diseconomies of scale in the expansion rate are modeled as a per-store sunk cost of entry that increases in the number of stores being opened in a given year. The increase in per-store cost may originate from rising cost of capital, as larger volumes of capital are required, or from managers working beyond their efficient productivity levels. Beyond diseconomies of scale, other explanations for gradual growth embedded in the model are fluctuations in demand, changes in competition, changes in the firms' financials, and intra-market dynamics (e.g. learning by doing within a market).

Distinguishing among the drivers of growth may be important for both firm strategy and public policy. A firm whose growth is constrained by limited resources can target policies to increase these resources, for example, by instituting management training programs or by developing a franchise system. If, on the other hand, lack of demand constrains growth the firm would do better by investing in advertising and product development. In addition, a social planner concerned with long-run concentration may wish to accelerate the growth of particular firms. Distinguishing among the drivers of growth is crucial for this goal. For example, making capital available will be effective only if it is lack of capital what limits growth. If the market for managers is constraining growth, a policy that simplifies store entry, for example, by easing zoning laws and property transfers, may have an impact in accelerating growth. The Mexican FTC encountered this problem in the early

[1] Penrose (1995) provides a summary of the classic models, and Sutton (2001), of more recent work.

2000s. WalMart was expanding at a much faster pace than the other national retailers. In an attempt to equalize the playing field the FTC approved a purchasing alliance among three large Mexican retailers.[2] Had the FTC known the retailers limited expansion was due to demand and cost of capital[3] it might have taken a different course of action.

The model I propose builds on the insights of Bresnahan and Reiss (1990, 1991) in which entry costs are infered from observing entry and exit decisions. I model an entry game in which firms decide on how many stores to open or close over a set of markets. Competing firms maximize the difference between the added variable profits from each market minus the costs of opening stores in all the markets. Variable profits in each market depend on the number of stores the firm has in that market and the number of stores the firm's competitors have in that market. The costs of opening stores is a non-linear function of the total number of stores being opened. This non-linear cost causes markets to be interdependent, as the entry decision in one market increases the cost of entry in other markets. I make the estimation tractable for this high dimensional choice problem[4] by using a variant of Maximum Simulated Likelihood: I use Bayes' rule to separate the likelihood function into two parts, only one of which needs to be simulated and requires relatively few simulation draws to obtain adequate accuracy.

I estimate the model using data on the Mexican supermarket industry during the early 2000s. It is a time of expansion for both WalMart and other discount retail chains.[5] I find that WalMart's sunk cost per store increases between 1% and 2% for each additional store they open in a given year. I also find that firms enjoy a 9% reduction in sunk entry costs when the firm's financial performance, as measured by the firm's debt to assets ratio, improves by 3 percentage points

[2]Sinergia, approved in 2004, was a JV for joint purchasing and distribution created by Gigante, Comercial Mexicana and Soriana. cfr. Recurso de Reconsideracion, RA-022-2004, Comision Federal de Competencia

[3]Gigante exited the market in 2008 after years of poor performance. It sold its stores to Soriana, who repositioned the brand and restructured the store network. Comercial Mexicana filed for bankruptcy protection in 2010 after huge losses in the US derivative market. Had they financed their operation in pesos they would have avoided bankruptcy.

[4]All entry games have complex interdependencies arising between players. In my model the interdependencies are exacerbated, as they are present across players and across markets. Previous work that has dealt with interdependencies across markets and players is Jia (2008). I do not follow her technique as it requires restricting the game to two players and the Mexican supermarket industry has more than two players.

[5]Durand (2007) and Iacovone et al. (2009) attribute this expansion of discount retailers in Mexico to the fall of trade barriers implemented by NAFTA.

(10%). Also, demand fluctuations alone cannot explain the growth rate as markets must grow 90% to justify the entry of a third store after opening a second store. The estimated competition effects are very similar to those found in Jia (2008). I find that a market must grow 45% to recover the loss in profitability caused by the presence of a competitor's store. I also find that the effect of cannibalism (the profit-loss on the inframarginal stores upon opening an additional store) is twice as strong as that of competition, which could imply that brand differentiation is significant in this industry. Counterfactuals in which the diseconomies of scale are completely removed for a subset of firms show how WalMart's competitors, Comercial Mexicana and Soriana, would have expanded faster had these diseconomies not existed.

Finally, it is important to recognize that the model does not allow for falling costs that are not due to financial constraints, nor does it allow for inter-market dynamics, as explanations for gradual growth. As such, I conclude that the observed patterns of entry can be rationalized by limited resources (increasing marginal costs) but I cannot reject the possibility that the observed patterns could also be explained by falling costs or other inter-market dynamics (e.g. learning *across* markets as opposed to learning *within* markets).

The next section provides background on the Mexican supermarket industry and discusses how particular features of the industry will support various assumptions of the model. Section 3 describes the data. Section 4 sets forth the model and the identification strategy. Section 5 demonstrates how Bayes' Rule can be applied to the likelihood function so as to make the estimation tractable. Section 6 provides the results and is followed by conclusions in section 7.

2 Industry Background

Suppliers

The industry of interest is the supermarket industry. It is comprised of self-service stores that sell mostly groceries but may also offer general merchandise, drugs, optics, and services. In Mexico the industry comprises 7 national chains, a few regional chains, and local stores. All national chains, with the exception of Waldo's Dólar Mart, are publicly traded and have been in the industry for decades. Waldo's entered in 2002 and is not publicly traded. During the late 90s and early 2000s the national chains undertook a rapid expansion, opening stores in new markets and closing unprofitable stores. This expansion is said to have been an effect of the trade liberalization that followed NAFTA[6]. Figure 1 shows the growth of both national chains and regional/local firms in Mexico during this period. It is clear that national chains were expanding quite rapidly during this time period.

Absent in the model are the non-formal grocery retailers: tent markets[7] and behind-the-counter stores. I do not consider these retailers as competitors to the chain stores given the significant quality differences. Non-formal retailers manage cash-only transactions, carry very few products, and do not advertise. Their main advantage lies in their strategic locations and their ability to operate outside a regulatory environment (tax evasion[8] and FDA regulation).

Regional chains and local stores are included in the model but their entry decisions are taken as exogenous. These firms have stores that are smaller (with a few exceptions) than those of national chains but offer similar products and services. They open and close stores at a much slower rate than national chains: between 1999 and 2006, 41% of the firms did not open or close a single store, 31% opened or closed one or two stores, and 28% opened or closed more than two stores. The largest number of stores opened by a regional chain during this time period is 16. Thus, I model the regional chains and local stores as firms that affect national chains' profitability but whose entry and exit decisions are independent of those of national chains.

National chains operate stores under four different formats: supermarkets, hypermarkets, clubs, and bodegas. Stores that are run by different firms but have the same format are very similar, both in floor size[9] and product assortment. Stores that have different format differ significantly on sales

[6]Iacovone et al. (2009) discusses WalMart's entry and expansion in Mexico.

[7]Ambulant tent markets represent a significant portion of the general merchandise industry in Mexico. According to Instituto Nacional de Estadística (2004b), ambulant retailers accounted for 1.6M jobs in 2003, approximately the same amount as Mom & Pop stores. Although some of these retailers are located in rural areas 65% of them are located, as of 2003, in cities of more than 15k people and 47% in cities of more than 100k people. As a reference, the formal sector accounted for approximately 7.8M jobs (cfr. sector 46 in Instituto Nacional de Estadística (2004a)).

[8]Fuentes et al. (2008) estimate ambulant-retailer tax evasion at 700 million dollars per year.

[9]The average floor size of bodegas, in thousands of square feet, is 54 for Soriana, 55 for Comercial Mexicana, and 49 for Gigante. The respective numbers for hypermarkets are 90, 83 and 69. The numbers are more dispersed for supermarkets, ranging from 8k square feet for Waldo's to 38k square feet for Chedrahui. Source: ANTAD 2006 Directory. WalMart data are not available.

floor size and the number of products offered. They do, however, use the same distribution system and suppliers. Table 1 shows the distribution of store formats for the national chains. Waldo's is the only chain whose only format is supermarkets. They are also the only young chain in the industry. All other chains have been operating in Mexico for decades. WalMart, who entered the Mexican supermarket industry in 1993[10], did so through acquisition of a long standing national chain. The differences in format distribution will be important when interpreting the differences in entry costs across firms. The recent entry of Waldo's will be important in interpreting the diseconomies of scale in entry since the experience curve may bias the estimates of the diseconomies of scale.

Distribution centers (DCs) are important for store location since most merchandise is channeled through the DC. The paper does not model the opening decisions of DCs but rather takes the locations of DCs as exogenous. Such assumption does not seem restrictive given firms do not add DCs at the same pace they open stores. Only two firms, Soriana and Chedrahui, add DCs during the period under study. Soriana opens a major DC in the Mexico City area and opens two smaller DCs, focused on perishable goods, in the northwest and southeast. Chedrahui opens two DCs, one in Mexico City and one in Monterrey. Gigante inherits a DC in the south of Mexico with the acquisition of a regional chain in that area.

Markets

Urban areas in Mexico are well defined and clearly marked: people reside in a concentrated area (figure 2 shows an aerial snapshot of the median market) beyond which is the countryside. The average distance between two neighborhing towns is 22 miles (flying distance) and the median town is 3.2 miles wide.[11] Defining a town as the relevant market for supermarkets is natural given the lack of suburban sprawl. Unfortunately, Mexico's geopolitical structure is based on municipalities and not towns. A municipality is smaller than a county but larger than a town. Most municipalities encompass a single town. Since Census and ANTAD data are reported at the municipality level

[10] WalMart opens its first Sam's Club in Mexico in 1993 in a JV with Cifra. In 1997 WalMart adquires Cifra and converts Cifra's stores into Supercenters.

[11] The median market, according to its 2006 expenditure on groceries, is Sahuayo, MICH; population of 63k; expenditure of $104M.

and not the town level, I use municipalities to define markets.

Some larger cities are comprised of more than one municipality. For these large cities I merge the municipality data to form a single market. The exception to this rule is Mexico City, which is excluded from the analysis altogether,[12] Monterrey, and Guadalajara. These last two cities are so large that defining them as a single economic market would be unrealistic. Instead, I define each municipality within Monterrey and Guadalajara as a market and try to control for the possible demand spillovers between neighborhing municipalities.

I include all municipalities with a population larger than 30,000 (as of 2005) or that appear in the ANTAD directory (adding 33 more municipalities). This is a very inclusive rule: most municipalities with fewer than 30,000 people are rural municipalities, which have less than 30% of the population concentrated in or around a town.[13] The final sample represents 80% of the national population and includes all major urban areas.

Tourist and border towns require special attention. The demand for grocery stores in tourist towns (e.g., Cancun) can be larger than that in otherwise similar-sized towns due to the affluence of tourists. Border towns have neighboring American counterparts (e.g., El Paso - Cd. Juarez). These neighboring towns can potentially have demand and/or supply spillovers that affect store profitability. The model partially controls for these factors with demand shifters.

Behavior

I now discuss some findings on firms' behavior obtained in a set of interviews with employees from ANTAD and from two chains in Mexico.

Firms' entry and exit decisions are made by a centralized planner who may be the board of directors or a vice president. These decisions are revised once a year or once every year and a half. Between revisions, a real estate department within the firm searches for optimal locations. Incorporated in the search is information obtained from regional managers, store managers, and

[12] Entry decisions in Mexico City are very different than in other markets due to the scarce store locations, the small correlation between demand and localized demographics (due to the commuting patterns), and the abundance of non-formal retailers.

[13] For example, the largest municipalities that did not make the sample are Otumba, MEX and Bocoyna, CHIH, each with an urban population of 8k, according to the Department of State (INAFED-SEGOB).

suppliers. Upon finding a location, the real estate department may purchase the land. Negotiating with both sellers and government agencies to settle disputes on price, zoning regulation, encumbrances, etc. involves a significant amount of managerial time. Other hurdles that need to be solved before opening a store include gathering the capital required for building and operating the store, the construction of the store, and adjusting the supply chain to supply the store. In summary, the time lapsed between the decision to open a store and the day the store is opened can be as much as 18 months. The process requires much managerial talent, acquired mostly through experience in the industry.

Most firms are family owned, and as such, rely on internal cash flows and bank loans for financing. Thus, one of the important limitations on growth is the cost of capital.

Firms learn about each others entry plans through media reports, suppliers, and other firms' real estate purchases. Firms use this information when defining their own entry plans. Thus, optimal store location is based on estimates of current demand, potential demand growth, and competitors' current entry decisions. This suggests that a simultaneous move static game could be appropriate in modeling how firms compete.

3 The Data

The principal data source is the ANTAD annual directory. The directory contains, at the municipality level, the number of stores each firm has by format type. The WalMart data that was missing from the ANTAD directories for certain years was acquired from Walmex and complemented by the maps in Iacovone et al. (2009) and annual reports. The seven national chains are WalMart, Comercial Mexicana (CCM), Gigante (GG), Soriana, Waldo's Dólar Mart, Casa Ley, and Tiendas Chedrahui. All other grocery retailers in the ANTAD database are considered regional or local firms. A firm is classified as a national chain according to the number of stores it has and its geographic scope. The distinction between a national chain and one that is not is clear: the smallest of the national chains, Chedrahui, operated 96 stores in more than 20 states in 2006. The largest regional/local firm in the ANTAD directory is HEB, which, as of 2006, had only 25 stores in 6 states.[14]

Table 2 provides a snapshot of the industry in 2006. The average market had three national-chain stores and one local or regional chain store. Tijuana had the maximum number of stores, with 57 national-chain stores and 32 local supermarkets. Non-grocery retailers (apparel stores, sporting goods, and pharmacies)[15] averaged about seven stores per market, twice as many as grocery retailers.

Entry and exit decisions are inferred by differencing the number of stores from year to year. For years in which information is not disclosed at the store-format level entry and exit decisions are inferred from the changes in aggregate floor size (also reported in the directories) and annual reports.

Entry and exit decisions are not directly observed. They are inferred by differencing. This can lead to three types of mismeasurements: (1) the inability to observe a simultaneous entry and exit within a market, as the number of stores from one year to the next remains constant; (2) the inability to observe store refurbishing (without a change store format), which may involve significant investments; and (3) the double counting that may arise from store refurbishing that changes the store format (this would appear in the data as simultaneous entry of one format and exit of another format). The first two mismeasurements will cause the increase in entry costs to be overestimated, while the third will cause the increase in entry costs to be underestimated. Information from interviews and annual reports suggest the first type of mismeasurement is not common in the industry, but that the second and third do arise from time to time. Of all the entry decisions in the panel, only 3%[16] occur with a simultaneous closure of a store of different format within the same market. I do not count these entries but rather assume the change is solely one of format. I do not count changes in store format as entry decisions.

Entry and exit patterns are corrected for mergers and acquisitions. I do not consider a store acquired via merger as an entry. I expect the strategic decisions behind mergers to be different than

[14]Carrefour, also excluded from the list, operated up to 29 stores before exiting the market in 2004. It sold off its operation to Chedrahui.

[15]This data was also obtained from the ANTAD yearly directories.

[16]Twenty six out of 740.

those behind opening or closing individual business units. There were only three mergers during this time period: GG acquired Super Maz's 13 stores in 2001, CCM acquired Auchan's 4 stores in 2003, and Chedraui acquired Carrefour's 29 stores in 2005. In the year following the mergers, the acquiring firms did not close any of the acquired firms' stores; thus, I am not concerned that the new stores may bias variable profits, identified off exit decisions.

After correcting for mergers and excluding entry and exit decisions in Mexico City, the total number of entries over the seven-year span is 714. Table 4 shows the distribution of entry decisions according to different variables. Firms open between 0 and 69 stores a year, with a mean of 17. Most entry decisions regarded opening a single store in the market (65%) or opening two stores in the market (21%). The outliers involve opening 6 and 17 stores, respectively, in a single market. Of all entries, 47% occurred in new markets, with a firm starting operations in that market for the first time. Fourteen percent occurred in markets where the firm already had one store in operating and 11% where the firm already had two stores in operation. Most entries (85%) occurred in markets where a competitor was already present. Overall, there is a lot of variation in where entries occurred, which will be convenient for estimating competitive effects.

Although exits are not as common, there is still a lot of variation in exit patterns. Of a total of 91 exits, 70% were closing of a single store, and 15% were of closing of two stores, simultaneously, in the same market. There is one observation of a firm closing 7 stores simultaneously in the same market. In 17 of these exits, the firm closed its last store in the market; and in 20 more, the firm downsized from two stores to one. Of all exits, 68% occurred when a competitor simultaneously entered the market. This richness in exit patterns will be helpful in separating fixed costs from sunk entry costs.

Demand information comes from various sources. Population and state-employee[17] information is obtained from the 1995, 2000 and 2005 censuses. CONAPO[18] provides population projections for the remaining years. Projections are scaled to fit the census data. Income is obtained from

[17]State employees not only proxy for the size of government but also have access to state-run, and subsidized, grocery stores. These stores do not appear in my data. Thus, the presence of large concentrations of state employees may proxy for unobserved competition for the national chains.

[18]Consejo Nacional de Poblacion: the Mexican government agency that oversees population dynamics.

INEGI[19] at the state level. Population density is calculated by dividing population by municipality land area, obtained from INEGI.

The principal demand covariate is the total expenditures on groceries in the market. Per-person grocery expenditures are obtained from ENIGH,[20] a household expenditure survey. The included categories are (1) in-house food consumption, (2) personal care items, and (3) household care and cleaning items. ENIGH contains survey data for the years 2000, 2002, 2004, 2005, 2006, and 2008. I divide the ENIGH surveys into two databases according to household income. Low-income households are defined at the poverty line as determined by CONAPO: two minimum wages per adult.[21] Thus I obtain, for each municipality, the per-low-income-person yearly grocery expenditure, and the per-high-income-person yearly grocery expenditure. Since it is a survey, it does not always have data on all municipalities. To fill in the gaps, a non-parametric regression is fit to the data and used to predict per-person expenditures. The predicted per-person expenditure is used in the model even if the specific market-year had actual data. The non-parametric regression fits per-person expenditures on a 4th order polynomial of population, population density, income, ratio of state employees to population, fraction of low-income population, and state and year fixed effects. Outliers are censored at the 5 and 95 percentiles.

The adjusted R-squared of the high-income and low-income regressions is 0.37 and 0.24, respectively. Despite the low R-squared, the fit is relatively good. For the high-income regression, the standard deviation of the ratio of residuals to predicted values is 0.26 (the mean is zero by construction) and the 5th and 95th percentiles of this ratio are -0.37 and 0.43, respectively. This means that for 90 percent of the observations for which data exists, the difference between the observed and the predicted values is less than 43% of the predicted value.

Measuring the goodness of fit for out-of-sample predicted values is more difficult since no good reference points exist. Nevertheless, I can compare the out-of-sample predicted values with the in-sample predicted values for the municipalities that will be used in the model (ENIGH provides

[19] Instituto Nacional de Geografia, Estadistica e Informatica : the Mexican Census Bureau.

[20] Encuesta Nacional de Ingresos y Gastos del Hogar: a household expenditure survey administered by INEGI.

[21] For a household to be considered low-income, the average monthly income per adult in that household must be less than two Zone B minimum wages in a 26-working-day month.

data on more municipalities than the number used in the model). That is, for those markets that will be used in the model, I can form two sets of predicted values: (a) those market-years for which I do observe expenditure (in sample predictions), and (b) those market-years for which I do not observe expenditure (out of sample predictions). The mean value of the out-of-sample predictions is \$2,239 dollars per person, per year. The 5th and 95th percentiles (before censoring) are \$1,735 and \$2,847, while the extreme outliers are \$777 and \$6,158. For the in-sample predictions, the mean value is \$2,417, the 5th and 95th percentiles are \$1,839 and \$3,229, and the extrema are \$935 and \$6,184. Thus, the distribution of predicted values for those markets for which data exists is very similar to the distribution of predicted values for those markets where data does not exist. After censoring, the average yearly expenditure per person for the sample that will be used in the model is, for the high-income group, \$2,270 dollars. For the the low-income group it is \$1,100 dollars. The median values are just slightly lower, indicative of a non-skewed distribution.

The total market expenditure is calculated by multiplying predicted per-person expenditure by population in that income group (obtained from CONAPO's Indice de Marginacion publication). The average market expenditure is \$277M dollars[22] per year (corresponding to Cd. Cuauhtemoc, Chih., population 134k). Market expenditure is highly skewed, with a median of \$95M corresponding to Zacatlan, Pue (population 70k). The skewness in market expenditure is the result of both population skewness and a positive correlation between the household expenditure of the high-income tier and population: perhaps not surprisingly, the wealthy live in the large cities. The largest market (\$3.7B) is Guadalajara (the municipality within the greater Guadalajara area), whose population is 1.6M. The smallest market (\$0.9M) is a small municipality in the state of Oaxaca. Table 3 provides a summary of the demand covariates and some supply covariates.

The location of distribution centers is obtained from firms' websites and annual reports. The distance to a DC is calculated as the flying distance from the geographical center of the market to the geographical center of the city in which the DC is located. Although some DCs are specialized for perishable products, the model assumes all DCs provide similar support and, thus, takes the distance from the market to the nearest DC. The distance to the nearest and farthest markets in

[22] All economic figures are in 2008 US dollars.

wich a firm has a store is also calculated as flying distance. Table 3 shows these distances for 2006. The average distance to the nearest DC is 225 miles, a large distance considering it is calculated as a flying distance and not a travel distance. Firms are very spread out in Mexico, with the largest distance between any two active markets averaging 1,230 miles. As a reference the flying distance from Mexico City to Tijuana is 1,400 miles.

Financial data on the firms is obtained for five of the seven firms. Only Chedrahui and Waldo's lack financial data since neither firm was publicly traded during this time period.[23] Two measures are used to proxy for the firms' cost of capital: the Debt to Assets ratio and the Operating Cash Flow to Debt ratio. The values used are those of the previous year, since I want to proxy for the cost of new capital. Table 5 shows a summary of these statistics.

4 The Model

In this section I extend the model in Bresnahan and Reiss (1990) to account for multi-business-unit firms. The objective is to estimate a cost curve that may be non-linear in the number of stores a firm opens per year and to account for a firm's financial situation as it relates to entry costs. In the first subsection I set up the general model and discuss the normalizations required to identify parameters even when the econometrician has access to a very rich data set. In the second subsection I impose additional assumptions to the model so as to recover the cost curve in the context of the Mexican supermarket industry data, whose richness is not sufficient to identify all types of complementarities and heterogeneities.

4.1 A Revealed Preference Approach

The primitives of the model are formed by a set of I firms with access to a set of M markets. A firm obtains a benefit from an action ($\mathbf{d}_i$)[24] and incurs a cost from such action. The benefits may depend on other firms' actions. The action is the number of stores (business units) a firm opens in

[23] Chedrahui entered the Mexican exchange in 2009. Casa Ley is not in the Mexican exchange. It is 50% owned by Safeway. I use Safeway's financial numbers, as published on the NYSE.

[24] Bold face denotes vectors. i will index a firm, m a market, and t a year.

each market ($\mathbf{d}_i^e$) and the number of stores it closes ($\mathbf{d}_i^{ex}$). A firm is allowed to open any number of stores and to close as many stores as it had active at the start of the period. This extends BR's work in that firms in their model had a dichotomous action: entry or no entry ($d_i \in \{0,1\}$), while in my framework the action is a vector ($\{\mathbf{d}_i^e, \mathbf{d}_i^{ex}\} \in \mathcal{A}_i \subset \mathbb{Z}^{2M}$).[25]

Firms' profits are a function of the number of stores each firm has per market, denoted by $\mathbf{s} \in \mathbb{Z}^{IM}$. The number of stores in a market is given by the firms' decisions and the previous store allocation. In vector notation the number of stores at time t is: $\mathbf{s}_t = \mathbf{s}_{t-1} + \mathbf{d}_t^e - \mathbf{d}_t^{ex}$.

The benefit from opening a store is the increase in variable profits achieved by doing so. The benefit from closing a store is the reduction in losses of variable profits achieved by doing so. The costs of the actions ($\mathbf{d}_i$) are sunk and thus properly called sunk entry costs. The number of active stores ($\mathbf{s}$) may also carry costs but are not sunk and thus are referred to as fixed costs. One may think of sunk entry costs as the set-up costs (permits, furbishing, contracting, etc) and of the fixed costs as operating costs (wages, utilities, etc). The distinction is subtle but important in determining market structure. If sunk entry costs are large relative to fixed costs, first movers will dominate the industry even if they are inefficient. If sunk costs are small relative to fixed costs, efficient firms will eventually enter and displace inefficient firms. In a BR-like model, fixed costs are captured in the intercept of the variable profits function, although this intercept may also measure other variables. Identification of both sunk entry costs and fixed costs requires information on both entry and exit decisions.

To formalize the preceding paragraphs, firms' profits are given by (a) variable profits, which are a function of the number of a firm's stores, and (b) entry and exit costs, which are a function of the firms' entry and exit decisions. Both variable profits and entry costs are affected by L exogenous observed shifters, $\mathbf{x}_t \in \Re^{LIM}$. Variable profits are also affected by unobserved shifters, $\varepsilon_{it} \in \Re^M$. Both functions are parametrized by a finite unknown vector, θ. Thus, the profits of a firm may be written as (to simplify on notation I omit the time script and group $\chi_i \equiv (\mathbf{x}_i, \varepsilon, \theta)$)

$$\pi(\mathbf{d}_i^e, \mathbf{d}_i^{ex} | \mathbf{d}_{-i}, \mathbf{s}, \chi_i) = V(\mathbf{s}_i + \mathbf{d}_i^e - \mathbf{d}_i^{ex}, \mathbf{s}_{-i} + \mathbf{d}_{-i}^e - \mathbf{d}_{-i}^{ex}, |\chi_i) + C(\mathbf{d}_i^e, \mathbf{d}_i^{ex} | \chi_i) \tag{1}$$

[25]Previous extensions of BR involved firms choosing a single location on an exogenous choice set of vertically (Mazzeo (2002)) or horizontally (Seim (2006)) differentiated locations; firms choosing multiple entry in a single location (Ishii (2008)); and firms choosing multiple entry in multiple locations (Ellickson et al. (2010); Holmes (2011); Jia (2008); Nishida (2009)).

where $V(\cdot)$ are the variable profits and $C(\cdot)$ are the costs. It is this cost function that is of main interest in the paper. The rest of the paper consists of trying to recover the parameters that characterize both $V(\cdot)$ and $C(\cdot)$. Observed by the econometrician are the realized actions, $\mathbf{d}^\star$, and the state variables $(\mathbf{s}, \mathbf{x})$. Assumed by the econometrician are the functional forms $V(\cdot)$ and $C(\cdot)$, although these may be very flexible. Also assumed by the econometrician are the rules by which $\mathbf{d}^\star$ arises from the unobserved profits, $\pi(\mathbf{d})$. I will assume firms play a complete information simultaneous move game in pure strategies, choosing actions $\mathbf{d}^\star$ that maximize profits.

The profit function is allowed to be very flexible. With enough variation in the data, one could estimate very complicated functional forms, including dynamic value functions. Nevertheless, there are two critical assumptions required to identify the profit function when observing solely entry and exit decisions along with state variables. The first regards the fact that the profit function is identified up to level and scale. The second regards the fact that variable profits, entry costs, and scrap values are not separately identified.

Assumption L&S The profits from having 0 stores in the market is 0. The variance of the unobservable ε is 1.

This normalization is common in almost[26] all entry papers and is required since the observed action $\mathbf{d}^\star$ is optimal under positive, strictly monotonic, transformations of the profit function.[27] In the model, a firm with no stores accrues a profit of 0. In reality, the profits from not operating in any market are not 0 but the value of investing in some other venture. The level normalization affects how the estimate on the constant is to be interpreted and the scale normalization implies that estimates are relative to some base measure. Nevertheless, these normalizations do not bias other estimates and, as such, can be called a trivial normalization. In contrast, the next normalization on scrap values is not a trivial one.

[26] The alternative most common normalization (Manski (1975); Fox (2010)) is to normalize the parameter on one of the covariates that has full support to 1. This is done when the econometrician does not want to make a distributional assumption on the unobservable, such that the higher moments of the unobservable are unknown.

[27] The profit function is normalized up to an affine transformation, and not a strictly monotonic transformation, since affine transformations are the only strictly monotonic transformations that preserve additive separability. That is, if $f(x) = f_1(x_1) + f_2(x_2)$, then $g(f(x)) = g(f_1(x_1) + f_2(x_2)) = h_1(x_1) + h_2(x_2)$ only if $g(\cdot)$ is an affine function.

Assumption SV. Scrap Values are normalized to 0. Any complementarities (either positive or negative) in multiple exits are also normalized to 0.

This is required since Entry Costs, Variable Profits and Scrap Values (the profits accrued/lost following exit) are not separately identified. The reason why this is so is that the econometrician observes only two decisions (entry and exit) but would like to infer three values. A simple example can make this clear: a monopolistic firm is deciding to enter a market it has not yet entered. Variable profits are $\pi + \epsilon$, where ϵ varies from market to market and its scale is normalized to 1, profits of not entering are normalized to 0, entry costs are E, and scrap values of exiting are S. From observing many entry decisions, the econometrician can infer the entry threshold, $\epsilon^{\star}$, where $E - \pi = \epsilon^{\star}$. From many exit decisions, the econometrician can infer the exit threshold, $\epsilon^{\star\star}$, where $S - \pi = \epsilon^{\star\star}$. Now the econometrician has two equations but three unknowns: E, S, π. To identify E and π some assumption needs to be made on S. This assumption can be setting S to 0 or to some other value.

If a firm can operate only one store the normalization on Scrap Values is just one of level, affecting the constant of the variable profits but otherwise trivial. When the firm operates more than one store the normalization is no longer trivial. For example, if the firm operates two stores, the econometrician would want to infer the scrap values accrued when exiting both markets simultaneously, when exiting each market individually, and when exiting one market and entering another. Assumption SV states that the scrap values in all these cases would be the same: zero.

In the current context, normalizing complementarities on multiple exits to zero can be restrictive if multiple closures affect the cost of entry (by absorbing the managers' time) or the firm's overall profitability (e.g., the loss of brand value or increased labor costs due to union contracts). Multiple exits are not common in the Mexican supermarket industry, which mitigates these concerns. Only four firm-year observations close eight or more stores in a given year, and the next largest closure is four stores. Nevertheless, I add a robustness check to the estimation. In an alternative specification I assume each exit increases the per-store sunk entry costs by the same amount that an entry would.

With these normalizations in place the cost function varies solely with entry decisions: $C(\mathbf{d}_i^e, \mathbf{d}_i^x|\chi_i) = C(\mathbf{d}_i^e|\chi_i)$. Variable profits are identified separately from entry costs in that entry costs are not affected by exit decisions and variable profits are.

One final comment regarding $V(\cdot)$, the variable profits function: no structure has been imposed on it. $V(\cdot)$ may be the value function of a dynamic program where the state space is defined by $(\mathbf{s}, \chi)$, as in Bajari et al. (2007). Furthermore, any arbitrary complementarity between the *vectors* $\mathbf{d}_i$, $\mathbf{d}_{-i}$, and the state space $(\mathbf{x}, \mathbf{s})$ may be recovered given the econometrician has enough data. This data requirement is not trivial. With many markets the vectors are of high dimensionality and independent observations are given by players acting in an *industry*, where the industry is composed of a set of markets. The Data Generating Process must be able to generate such rich data. In other words, the industry must have gone through significant variance to produce different combinations of entry and exit for each state, where the state is given by the number of stores *in each market*. In most settings a single rise and decline of an industry is observed. Additional restrictions need to be added to recover some parts of the variable profits and cost curves.

The Mexican supermarket industry data is not as rich as one would like. As such, I do not estimate a dynamic value function and I restrict significantly the complementarities that may arise across markets. The following section makes these restrictions explicit. The exact structural forms of variable profits and entry costs are introduced at the end of the section.

4.2 Adapting the Model to the Mexican Supermarket Industry

The main interest of the paper is to estimate a cost curve where the cost per store may increase with the number of stores opened per year. I assume per-store entry costs can be separated into two parts, one at the national level that is independent of location and depends solely on the total number of stores being opened; and a second part at the market level that is dependent solely on observed market covariates but not on how many stores are being opened in that specific market.

This restriction limits the market interdependencies that may arise due to entry costs. Opening a store in Tijuana will raise the cost of opening a store in Cancun, but it will do so by the same amount in which it raises the cost of opening a store in Oaxaca. This is the only intermarket dependency that will be allowed.[28] The two main complementarities that are assumed away are (1)

[28] The number of arbitrary complementarities that the general framework allows for is much larger. For example,

economies of scale in entry within the same market, for example, if WalMart opened a Supercenter and a Sam's Club next to each other, with a shared parking lot; and, (2) economies of scale across a subset of markets, for example, opening a store in Cancun may decrease the cost of opening a store in Playa del Carmen but not in Oaxaca. This latter complementarity is unlikely to exist in the industry given the isolation of markets. Complementarities of the former form may exist and would bias estimates of variables profits, especially those regarding competition between a firm's own stores. It is not likely to bias the nationwide diseconomies of scale in entry if the firms open stores in many markets, as is the case. Nationwide economies of scale (e.g. national advertising) may exist but cannot be separately identified from nationwide diseconomies of scale and will bias downward the diseconomies of scale.

The above restriction does not reduce the relevant state space, as entry decisions across markets remain interdependent.[29] As such, I must resign myself to estimating a model with a huge state space. Estimating a dynamic model is infeasible since a reduced form of either Value Functions or Policy Functions that can be related to the primitives of the model is unknown and the usual techniques for solving for either Value Functions or Policy Functions are infeasible with such large state space. Thus, the following assumption is made:

Assumption SG. Firms play a static game. They do not consider they, or any other competing firm, will revisit their decisions next year when making decisions this year.

The assumption is not as stringent as it appears at first glance. By posing a static game, the estimates of the variable profits should be interpreted as NPV and not as per-period profits. Covariates will capture both a direct effect of the covariate on profits and an indirect effect of that covariate on the expectations of future profits. For example, market growth does not contribute directly to profits, but a positive estimate on market growth indicates that firms value growing markets for their future rents.

with 400 markets and a simple dichotomous entry/exit decision, the set of different possible entry costs is 2^{400}, which is a number with 120 zeros.

[29] A reduction in the state space would require removing the market label from the state. This would make the model unrealistic: it would imply that having two stores in different markets is the same (at least stochastically so) as having two stores in the same market.

The two important dynamic concerns are learning and preemption. I will try to control for these two in the static model. I control for learning within a market by allowing for a level shift in entry costs for the first time a firm opens stores in a market. If "learning by doing" within a market is prevalent in the industry this estimate will be positive, as opening subsequent stores is easier than opening the first. If the "option value" of learning is prevalent, this estimate will be negative as firms open the first time in a market to learn about the market and, following a bad outcome, do not open subsequent stores. Although I could allow for similar shifts in entry costs for the second and third time a firm opens in a market, I do not think it is necessary to do so as most the the intra-market learning happens with the first store: the store manager will have a good idea of the local demand, local suppliers and local real estate.

I control for preemption within a market by allowing for a level shift in entry costs for the first time a firm opens stores in a market where no other competitor is present. If preemption is prevalent the estimate will be negative.

What is lost in the static model is learning and preemption across markets. It is these inter-market dependencies that make the dynamic model infeasible. The isolation of markets mitigares the concern of intermarket learning but not of intermarket preemption. It may be the case that a firm may open in its opponents' favorable markets before opening in its own favorable markets. In interviews with industry players this does not seem to be the case, as managers focus on opening the most profitable stores given competitors current expansion plans but not competitors' potential future expansion plans.

The optimal strategy in the static game is very complex given the many different store opening configurations that are possible. I reduce the complexity of the problem by assuming away any market interdependencies that are not related to diseconomies of scale in entry costs. Thus, the gains in variable profits from the entry decision in many markets is solely the added value of the gains in variable profits of each market where entry occurs.

Assumption IVP. Variable profits are assumed to be independent and additive across markets.

This assumption eliminates any spillover across markets due to variable profits. There are two reasons for this assumption. The first is that the data are not rich enough to estimate these spillovers. Estimating arbitrary complementarities across markets would require observing multiple entry and exit *vectors*. The mexican data does not have such richness. The second reason is that it simplifies the optimal entry/exit strategy. By assuming no spillovers in variable profits, no spillovers in exit costs (assumption 4.1), and equal spillovers in entry costs, exit decisions are independent across markets and entry decisions in any given market do not alter the optimal ordering of entry decisions in other markets. The combinatorial search problem is simplified to one of finding the optimal ordering of entry possibilties.

Independence is a strong assumption. It rules out the possibility that consumers may benefit from stores in nearby markets. It also rules out the possibility of sharing advertising, procurement, overhead, and/or distribution costs across markets. For example, a firm cannot increase its overall purchasing power by increasing the number of stores it operates.

Assuming away spillovers on the demand side is not much of a concern in this industry since markets are isolated (cfr. section 2). On the supply side, stores acquire most of their merchandise either from a DC (operated by the parent company or by a large supplier, e.g., CocaCola) or from local suppliers. One of the main spillovers to worry about is firms clustering stores in anticipation of a DC (as in Holmes (2011)). During the observed time span, only two firms added DCs to their networks (and one firm inherited a DC from an acquired firm). Thus, with the exception of these two firms, it is reasonable to consider the network of DCs as exogenous and the clustering of stores as a response to current DCs and not future DCs.

The other supply-side complementarity of concern is increasing purchasing power. A firm that opens stores in a given market may increase its purchasing power, increasing the variable profits of stores in all markets. If such behavior is prevalent, entry costs may be biased downwards. The firm has an incentive to open many stores since that increases the profitability of existing stores. The bias on the *increase* in the entry cost should be negligible given the large base of stores the firms already operate.

The last assumption is on heterogeneity. Ciliberto and Tamer (2009) and Toivanen and Waterson (2005) are examples of entry models that allow for rich competition patterns that depend on firms' identity. Estimating such entry models requires many independent observations of the same firms. Limitations on the length of the panel and the fact that I allow for heterogeneity in entry costs across firms requires reducing the degrees of freedom in other parts of the model.

Assumption HVP. Variable profits are homogeneous across firms and across markets conditional on covariates.

Variable profits are identified off exit decisions. Allowing for variable profits to differ across markets/firms requires observing multiple exit decisions in all markets and/or for all firms. As stated earlier, although exit does happen, it is not overly common. Complementarities would be estimated off very few observations, making them dubious and biasing entry costs in unknown ways. By assuming homogeneity, exit decisions of different firms and/or in different markets may be pooled.

To the extent that firms are different, the estimated coefficients would capture the average effect. Other coefficients may be biased to accommodate the heterogeneity. For example, if firms differ in variable profits, estimated differences in entry costs will capture differences in variable profits and not only true differences in entry costs. Another example of this type of bias occurs when firms with more stores are also more efficient. Since variable profits are not allowed to vary across firms, competition effects between a firm's own stores will be biased downwards, as firms with lower costs will also be the firms that have more stores in the same market.

As described in Section 2, stores owned by different firms but of the same format (supermarkets, hypermarkets, bodegas, and clubs) are very similar. They tend to be of the same size, operate with similar personnel and carry similar product lines. Assuming homogeneous variable profits should not be too much of a concern once the covariates include the store-format mix and different intercepts (purchasing power) for the different firms. I add both of these in an extension to test the homogeneity assumption.

4.3 Structural Forms

In this section I set forward the specific functional forms of the variable profit and cost functions. To summarize the previous section, variable profits are additive across markets and are the same for all firms and all years conditional on covariates. Variable profits in a given market depend solely on the number of stores, both of the firm and its competitors, in that market. Entry costs consist of two parts: a per-market entry cost and a national cost function, which depends solely on the total number of entries. Thus, firms profits are given by:

$$\pi(\mathbf{d}_i^e, \mathbf{d}_i^{ex} | \mathbf{d}_{-i}, \mathbf{s}, \chi_i) = \sum_{m \in M} \left[v(s_{imt}, s_{-imt} | \chi_{imt}) - c_1(d_{imt}^e | \chi_{it}) \right] - c_2(\omega_{it} | \chi_{it})$$

$$\mathbf{s}_{it} = \mathbf{s}_{it-1} + \mathbf{d}_{it}^e + \mathbf{d}_{it}^x \qquad s_{-imt} = \sum_{k \in I \setminus \{i\}} s_{kmt} \qquad \omega_{it} = \sum_{m \in M} d_{imt}^e \tag{2}$$

Sunk Entry Costs

The cost function is given in two parts: the per-store cost in a given market and the overall national cost that is increasing in the number of stores:

$$c_1(d_{imt}^e | \chi_{it}) = d_{imt}^e \left(\theta_i^L + \theta^F x_{it}^{Fs} + \theta^{\text{Learning}} 1_{\{s_{imt-1}=0\}} + \theta^{\text{Preemption}} 1_{\{s_{imt-1}=0, s_{-imt-1}=0\}} \right) \tag{3}$$

$$c_2(\omega_{it} | \chi_{it}) = \left(\theta_i^C + \theta^{FC} x_{it}^{Fs} \right) \left(1 + \theta^{\text{Trend}} t \right) \left(\omega_{it} \ln [\omega_{it}] \right) \tag{4}$$

Equation 3 is the linear part that is common in previous literature. It contains a constant linear cost, θ_i^L, which is allowed to vary by firm. This parameter is what is recovered in most of the previous literature when referring to sunk costs. It can capture the sunk costs of a specific store: building the store, issuing permits, hiring personnel, contracting with suppliers, etc. I expect θ_i^L to be lower for the more efficient firms (e.g. WalMart) or for those firms whose store-format mix is dominated by small stores (e.g. Waldo's, which only operates supermarkets). I also expect θ_i^L to be lower for firms with unobservably larger variable profits, since variable profits are not allowed to differ arbitrarily across firms.

The term x^{Fs} is the firm's debt-to-assets ratio. It proxies for the firm's cost of capital, so I expect θ^F to be positive: the larger is the debt, the more expensive the capital becomes for the firm. As a robustness check, I use the ratio of operating cash flow to debt as an alternative measure of the cost of capital.

Also affecting the per-store entry costs are θ^{Learning} and $\theta^{\text{Preemption}}$. The former is a level increase in per-store entry costs when the stores being opened are the first stores in that specific market. The term $\theta^{\text{Preemption}}$ is a level shift in entry costs when a firm is entering a market for the first time and no competitor is present in that market.

Equation 4 is the main interest of the paper. It is a non-linear function of the total number of stores being opened in a given year (ω_{it}). A positive estimate on θ_i^C implies the cost per store is increasing in the total number of stores opened. That is, the cost per store of five new stores is less than the cost per store of ten new stores. I use $\omega \ln[\omega]$ instead of ω^2 since I expect the increases to be less than proportional. The term θ_i^C is allowed to vary by firm since I expect that resources across firms for store opening to differ significantly.

The only caveat is that I do not instrument for unobserved heterogeneity in entry costs. A firm that has an unobservably low entry cost will open more stores because its entry cost is lower, and this will bias downward the estimate on θ_i^C. Allowing for the per-store linear cost (θ_i^L) to differ across firms mitigates this concern to some extent. Nevertheless, a firm's per-store cost may be dropping with time. This would be the case if the firm were undergoing a learning process (experience curve). I would not be surprised if this were the case with Waldo's, which is new to the market. If this were the case, I would expect θ_i^C to be biased downwards, and the bias could be large enough to make θ_i^C negative. It would be surprising if the other firms, which have been in the market for decades, experienced a similar learning curve.

The other terms affecting the non-linear costs are $\theta^{FC} x^{Fs}$ and the time trend. The former captures how a firm's cost of capital affects increases in marginal costs. A negative estimate on θ^{FC} implies increases in marginal costs are smaller when the firm has a high cost of capital. If increases in marginal costs captured purely managerial constraints, the estimate would be indicative of capital and managerial resources being complements in store opening. That is, when capital abounds, the *increases* in the cost of entry are larger (the shadow price to the constrained managerial input is larger), probably because managers are more spread out. Nevertheless, the overall entry cost may be lower, as captured by θ^F. If, on the other hand, the increases in entry costs are due to both managerial and financial limitations, then θ^{FC} may be capturing the non-linearities that the functional form restricts.

The time trend, $(1+\theta^{\text{Trend}}t)$, allows for the increases in cost to relax over time. As firms build capabilities, each year they can open more stores than in the previous year. A negative value on θ^{Trend} would capture this. The specific structure implies a linear discounting. The time trend is important for rationalizing firms opening more stores each year when demand is not growing.

Variable Profits

Variable profits are given similarly to Ellickson et al. (2010), where the variable profits per market are just the added value of a reduced form of profits per store:

$$v(s_{imt}, s_{-imt}|\chi) = s_{imt}\left[\mathbf{x}_{imt}\beta + \theta^{\text{Own}}\ln\left[s_{imt}+1\right] + \theta^{\text{Comp}}\ln\left[s_{-imt}+1\right] + \varepsilon_{imt}\right] \tag{5}$$

The term θ^{Own} captures the complementarities/substitution stores in the same market have on each other. A positive θ^{Own} would capture the Chain Effect of Jia (2008), while a negative value would be more in accordance with Holmes (2011), who finds that stores cannibalize each other. The term θ^{Comp} captures the effect of competitor stores and I expect this value to be negative. Supply and demand shifters in $\mathbf{x}$ include the log of market potential of the high- and low-income populations; the log-number of local stores (non-national grocery retailers); the log-number of state employees; the log of the average income per capita in the market; the market's average growth rate for the previous three years; the log of population density in the market; the log-distance to the nearest DC, to the nearest market in which the firm has stores, and to the farthest market in which the firm has stores; the dummy variables for tourist, border and metro markets; and a constant.

The market potential is given by the total dollar expenditure, as given by the ENIGH survey (cfr section 3). I posit that the market potential of the high-income population will be important for variable profits while the market potential of the low-income population will be less so since the lower-income tier tends to buy more in the smaller mom and pop shops. I expect the number of state employees to have a negative impact on variable profits since they have access to state-run and -subsidized grocery stores and, as such, do not buy at the national chains even though their expenditure has been accounted for in the market-potential variables. I also expect the distance to the DC and to the nearest active market to be important for variable profits but the distance to the farthest market to be insignificant.

The term ε_{imt} is a unobservable that is known to the firms. I assume ε is iid across market-firm-years and is distributed Normal with mean 0 and variance 1. The infinite support of ε_{imt} is important for rationalizing certain market outcomes. If θ^{Own} is negative, such that stores are partial substitutes, the optimal number of stores to have opened is increasing with market size $(\mathbf{x}\beta + \varepsilon\ )$ and there is always a market size that justifies a given number of stores. Figure 3 shows how different values of market size map into the optimal number of stores when θ^{own} is negative. On the other hand, if θ^{Own} is positive, the optimal number of stores is always infinite independently of market size and, thus, closure of a single store cannot be justified under the current model. Thus, I assume θ^{Own} to be negative hereafter. I am not explicitly assuming θ^{Own} to be negative (which rules out the "Chain Effect" within a market). I am affirming that under the current model the existence of a single exit rules out positive θ^{Own}. I *implicitly* assumed θ^{Own} to be negative when I assumed free exit. If exit were costly, single exit could be justified even if θ^{Own} were positive. In Jia (2008), a Chain Effect can arise because the author does not observe exits and she restricts the number of stores per location to one and because the effect dies out with distance.

As is normal in these models, I am concerned that ε_{imt} may be unobservably large in markets where entry occurs, biasing both θ^{Own} and θ^{Comp}. Instead of using instruments for the endogenous variables, I will proxy for the unobserved profitability by using the number of non-grocery retailers in the market. Non-grocery retailers include the number of department stores, pharmacies, and specialty stores (sporting goods, electronics, autoparts, etc). These stores draw employees from the same labor pool as grocery stores and share similar real estate costs. As such, they should capture the unobservable supply and demand conditions in the market that affect both grocery and non-grocery retailers.

5 Estimation

5.1 Best Response and Equilibrium

Under a simultaneous-move complete-information game, a player optimizes over his own strategy space, taking the other players' strategy as given. These best-response functions along with an equilibrium selection rule are sufficient to provide a mapping between the observed market outcomes ($\mathbf{d}^\star =(\mathbf{d}^\star_1, ..., \mathbf{d}^\star_I)$) and the unobserved profit function. I now characterize the best-response function of the players.

Let $\mathbf{d}^\star_i$ be the observed action of firm i. The optimality of (2) implies

$$\pi(\mathbf{d}^\star_i|\mathbf{d}^\star{}_{-i}, \mathbf{s}, \chi_i) \geq \pi(\mathbf{d}|\mathbf{d}^\star_{-i}, \mathbf{s}, \chi_i) \; \forall \mathbf{d} \in \mathcal{A}_i \subset \mathbb{Z}^{2M} \tag{6}$$

This set of inequalities is very large, but most of them are redundant. In the appendix, I show how the non-redundant inequalities can be separated almost to a per-market basis. They cannot be fully separated into a per-market basis due to the market interdependencies caused by the increasing entry costs.

The following are some relevant notations. $\mathcal{D}^e$ is the set of markets where the firm entered (opened at least one store). $\mathcal{D}^{ex}$ is the set of markets where the firm exited. $\mathcal{D}^n$ is the remaining set of markets. Define:

$$\begin{array}{rcll}
ms_{imt} & \equiv & \mathbf{x}_{imt}\beta + \theta^{\text{Comp}} \ln\left[s_{-imt}+1\right] & \\
\varphi_{imt} & \equiv & \theta^L_i + \theta^F x^{Fs}_{it} + \theta^{\text{Learning}} 1_{\{s_{imt-1}=0\}} + \theta^{\text{Preemption}} 1_{\{s_{imt-1}=0, s_{-imt-1}=0\}} & \text{if } d_{imt} \geq 0 \text{ \& } 0 \text{ otherwise} \\
Y_m & \equiv & ms_{imt} + \theta^{\text{Own}} \left(s_{imt} \ln\left[s_{imt}+1\right] - \left(s_{imt}-1\right) \ln\left[s_{imt}\right]\right) - \varphi_{imt} & \text{if } s_{imt} \geq 1 \text{ \& } \infty \text{ otherwise} \\
Z_m & \equiv & ms_{imt} + \theta^{\text{Own}} \left(\left(s_{imt}+1\right) \ln\left[s_{imt}+2\right] - s_{imt} \ln\left[s_{imt}+1\right]\right) - \varphi_{imt} & \\
E_+ & \equiv & \left(\left(\omega_{it}+1\right) \ln\left[\omega_{it}+1\right] - \omega_{it} \ln\left[\omega_{it}\right]\right) \left(\theta^C_i + \theta^{FC} x^{Fs}_{it}\right) \left(1+\theta^{\text{Trend}} t\right) & \\
E_- & \equiv & \left(\omega_{it} \ln\left[\omega_{it}\right] - \left(\omega_{it}-1\right) \ln\left[\omega_{it}-1\right]\right) \left(\theta^C_i + \theta^{FC} x^{Fs}_{it}\right) \left(1+\theta^{\text{Trend}} t\right) & \text{if } \omega_{it} \geq 1 \text{ \& } \infty \text{ otherwise}
\end{array}$$

Y_m is the non-random change in variable profits due to reducing the number of stores in market m from the observed number of stores (including the decision d^e_{imt}, d^{ex}_{imt}) to one less store. This reduction can be accomplished by opening one less store or by closing a store, depending on the observed decision. When it is the former, the market specific entry costs (φ_{imt}) are substracted from the variable profits. Z_m is the non-random change in variable profits from reducing the number of stores in market m from the observed number of stores plus one to the observed number of stores. E_+ is the non-market specific increase in costs of entry when opening one more store when $\omega^\star$ stores have already been opened. E_- is the same as E_+ but with $\omega^\star - 1$ stores.

The non-redundant inequalities of the optimality condition (6) are then

$$\begin{array}{lcc} \forall m \in \mathcal{D}^e & \varepsilon_{imt} \geq -Y_m + E_- & \varepsilon_{imt} < -Z_m + E_+ \\ \forall m \in \mathcal{D}^{ex} & \varepsilon_{imt} \geq -Y_m & \varepsilon_{imt} < -Z_m \\ \forall m \in \mathcal{D}^n & \varepsilon_{imt} \geq -Y_m & \varepsilon_{imt} < -Z_m + E_+ \\ \multicolumn{3}{c}{\min_{m \in \mathcal{D}^e} \{Y_m + \varepsilon_{imt}\} \geq \max_{k \in \mathcal{D}^n \cup \mathcal{D}^e} \{Z_k + \varepsilon_{ikt}\}} \end{array} \tag{7}$$

The last inequality is not irrelevant since convex costs imply $E_+ > E_-$. The last inequality establishes that the firm could not have done better by opening one more store in some alternative market in exchange for not opening a store in one of the markets where it did open one.

Equations 1-4 in (7) are necessary conditions that each firm must satisfy in any equilibrium. They fully characterize the firms' best-response function. As is well known, they are not sufficient to derive a mapping between the observed outcomes, $\mathbf{d}^\star$, and the underlying parameters. It is because of multiple equilibria that such mapping cannot be established. I assume an equilibrium selection rule and test how sensitive are the estimated parameters to the assumed equilibrium selection rule.[30] I choose the equilibrium selection rule by applying the iterative best response heuristic and choosing an order of movement in the heuristic.[31] Given the multitude of markets, many different equilibrium selection rules can be used. For example, I could allow firm A to decide where to open a single store, then to allow firm B to decide where to open a single store, then allow firm A to open a second store, and so on. Reversing the order of the firms or of the markets

[30] The other three approaches I could have taken are (a) to estimate the equilibrium selection rule along with the parameters, (b) to be agnostic about the equilibrum selection rule and set identify the parameters that could arise under any equilibrium selection rule (Tamer (2003); Ciliberto and Tamer (2009)), and (c) identify parameters off necesary conditions through moments of profit inequalities (Pakes (ming); Pakes et al. (2006)). The first two require knowledge of the equilibria set, which in my application is so large that it would be difficult to find. The third option requires solving for a selection issue on the unobservables that is infeasible in the current setting.

[31] The iteratuve best response heuristic consists in assigning a strategy to all players (usually the no-action strategy) and then allowing each player in turn to best response to the strategy of all other players. The process is iterated until no player alters his strategy. When the process converges it does so at an equilibrium of the simultanous move game. When players' strategies are strategic substitutes, as in entry, the process has been show to converge.

results in a different equilibrium selection rule.

When using the equilibrium selection rule described above, there is no closed form solution for the ε space that maps into the observed market outcomes. Monte Carlo integration is required to obtain such probability: many ε draws are simulated. For each draw, the game is solved. Counting how many times the outcome of the game matches the outcome observed in the data gives the probability over the observed outcome. The key complication is that a single observation is a set of firms competing over a set of markets. As such, the necessary conditions for all firms must hold in all markets simultaneously, in addition to the given equilibrium selection rule. For any given random draw, the probability of all these conditions holding simultaneously is very small. Thus, a very large number of simulations is required to achieve positive probabilities.[32]

To circumvent this problem I apply Bayes' Rule to the likelihood function. Let $\Pr[\Gamma_{\mathbf{d}}(\theta)]$ be the probability that market outcome $\mathbf{d}$ arises when firms play according to a specific equilibrium selection rule. Let $\Pr[\Omega_{\mathbf{d}}(\theta)]$ be the probability defined by the intersection of the half spaces of equation (7) for all firms and $\Pr[\Omega_{\mathbf{d}_i}(\theta)]$ be that for firm i. Using Bayes' Rule and the independence of unobservables across firms, the following must hold:

$$\Pr[\Gamma_{\mathbf{d}}(\theta)] = \Pr[\Gamma_{\mathbf{d}}(\theta) \cap \Omega_{\mathbf{d}}(\theta)] = \Pr[\Gamma_{\mathbf{d}}(\theta)|\Omega_{\mathbf{d}}(\theta)]\Pr[\Omega_{\mathbf{d}}(\theta)] = \Pr[\Gamma_{\mathbf{d}}(\theta)|\Omega_{\mathbf{d}}(\theta)]\prod_{i\in I}\Pr[\Omega_{\mathbf{d}_i}(\theta)]$$

The first equality indicates that the probability of the specific equilibrium arising is the probability that it arises *and* that it is *an* equilibrium. The second equality separates this probability into two parts: the probability that the specific equilibrium arises given it is an equilibrium times the probability that it is an equilibrium.

Note that when one wrongly ignores multiple equilibria and uses the best-response functions directly in the likelihood, she is implicitly assuming the conditional probability is one. The econometrician who wrongly specifies such a likelihood will obtain inconsistent estimates only if the conditional probability varies with the parameters. As I will show later, the Hessian of the condi-

[32]An alternative way to estimate the model would be to use Moment Inequality techniques (Pakes et al. (2006); Pakes (ming)) on the best-responses alone. Doing so could come at the cost of point identification and standard-errors computation being quite challenging (Chernozhukov et al. (2007)). Furthermore, the requirements on the unobservables in Pakes (ming) may be difficult to meet in the current setting. The technique of bounding the probabilities, developed in Tamer (2003) and Ciliberto and Tamer (2009), are also infeasible since they require knowledge of the equilibria set.

tional probability with respect to the parameters is, in the Mexican data, very flat compared to the Hessian of the second probability (the probability of the best responses). As such, the estimated parameters are insensitive to the specified equilibrium selection rule at the estimated parameter vector. It would not be surprising if they were also insensitive to other equilibrium selection rules.

I apply Bayes' rule once more so as to simplify $\Pr\left[\Omega_{\mathbf{d}_i}(\theta)\right]$. Let $\Pr\left[\Lambda^4_{\mathbf{d}_i}(\theta)\right]$ be the probability space given by line 4 in equation (7) and $\Pr\left[\Lambda^{1-3}_{\mathbf{d}_i}(\theta)\right]$ be the probability space given by lines 1 through 3 of equation (7). Then

$$\Pr\left[\Omega_{\mathbf{d}_i}(\theta)\right] = \Pr\left[\Lambda^4_{\mathbf{d}_i}(\theta)|\Lambda^{1-3}_{\mathbf{d}_i}(\theta)\right]\Pr\left[\Lambda^{1-3}_{\mathbf{d}_i}(\theta)\right]$$

The last probability does have a closed-form solution. This makes the simulations very convenient as relatively few draws are necessary to obtain adequate coverage.

The mixed maximum simulated log-likelihood for estimation is then

$$MMSL(\theta) = \sum_{t\in T}\log\left[\tilde{\Pr}\left[\Gamma_{\mathbf{d}_t}(\theta)|\Omega_{\mathbf{d}_t}(\theta)\right]\right] + \sum_{t\in T}\sum_{i\in I_t}\log\left[\tilde{\Pr}\left[\Lambda^4_{\mathbf{d}_{it}}(\theta)|\Lambda^{1-3}_{\mathbf{d}_{it}}(\theta)\right]\right] +$$

$$\sum_{t\in T}\sum_{i\in I_t}\left\{\sum_{m\in\mathcal{D}^e_{it}}\log\left[G(-Z_m+E_+)-G(-Y_m+E_-)\right] + \sum_{m\in\mathcal{D}^{ex}_{it}}\log\left[G(-Z_m)-G(-Y_m)\right] + \sum_{m\in\mathcal{D}^n_{it}}\log\left[G(-Z_m+E_+)-G(-Y_m)\right]\right\}$$

where

$$\tilde{\Pr}\left[\Gamma_{\mathbf{d}_t}(\theta)|\Omega_{\mathbf{d}_t}(\theta)\right] = \frac{1}{R_2}\sum_{r=1}^{R_2}\mathbf{1}_{\{\Gamma_{\mathbf{d}_t}(\theta,\varepsilon_r)\}} \quad \varepsilon_r \sim N(0,1)\ \cap\ \varepsilon_r \in \Lambda^{1-4}_{\mathbf{d}_{it}}(\theta)$$

$$\tilde{\Pr}\left[\Lambda^4_{\mathbf{d}_{it}}(\theta)|\Lambda^{1-3}_{\mathbf{d}_{it}}(\theta)\right] = \frac{1}{R_1}\sum_{r=1}^{R_1}\mathbf{1}_{\{\Lambda^4_{\mathbf{d}_{it}}(\theta,\varepsilon_r)\}} \quad \varepsilon_r \sim N(0,1)\ \cap\ \varepsilon_r \in \Lambda^{1-3}_{\mathbf{d}_{it}}(\theta)$$

The first conditional probability, $\tilde{\Pr}\left[\Gamma_{\mathbf{d}_t}(\theta)|\Omega_{\mathbf{d}_t}(\theta)\right]$, is a count on how many times the observed equilibrium ($\mathbf{d}^\star$) arises when modeling firms as playing under the assumed equilibrium selection rule. The ε draws used for this censored probability are such that an equilibrium is guaranteed to exist, but it need not be the equilibrium observed in the data. The second probability, $\tilde{\Pr}\left[\Lambda^4_{\mathbf{d}_{it}}(\theta)|\Lambda^{1-3}_{\mathbf{d}_{it}}(\theta)\right]$, is a count on how many times line 4 in equation 7 is satisfied when ε is drawn from a censored Normal distribution. The censoring is given according to inequalities $\Lambda^{1-3}_{\mathbf{d}_{it}}(\theta)$ being satisfied.[33] Part B in the Appendix provides more details on consistency and efficiency of

[33] Note that the censoring depends on the parameters. There is a closed form transformation between the unit-uniform distribution and the censored multivariate-normal distribution used in $\tilde{\Pr}\left[\Lambda^4_{\mathbf{d}_{it}}(\theta)|\Lambda^{1-3}_{\mathbf{d}_{it}}(\theta)\right]$, so draws are

the likelihood function when using conditional probabilities. I set R_1 to 3000 and R_2 to 2000.

Two different equilibrium selection rules are used to form the conditional probability, $\Pr\left[\Gamma_{\mathbf{d}}(\theta)|\Omega_{\mathbf{d}}(\theta)\right]$. Both involve finding the equilibrium through iterating best responses. The first entails firms moving according to the number of stores they each owned at the end of the sample period: WalMart moves first, followed by Waldo's, Soriana, GG, CCM, Ley, and Chedrahui, respectively. The second involves firms moving in alphabetical order.[34] Note that the game is still a simulatenous-moves game. The order of movement refers to the heuristic by which I find an equilibrium through iterating best responses.

6 Results

Tables 6 and 7 contain the results of the main specification. Table 6 shows the results for the entry-cost parameters under all specifications, while table 7 displays the results of the variable-profits estimates only for the main specification and the specifcation without constraints to entry. The variable profits of the other specifications are all very similar to those of the base model and are not shown.

6.1 Entry Costs

The parameters governing the convex part of entry costs, θ_i^C in equation (3), are positive and statistically different from zero for two firms: WalMart and CCM. These two firms along with Soriana show a steady expansion over the time span, although WalMart grows much faster than the other two. The estimate for Soriana is also positive, but not significant. Casa Ley and

taken form the uniform distribution at the beginning of the optimization and held constant until the end. There is no closed form transformation for the conditional distribution used in $\tilde{\Pr}\left[\Gamma_{\mathbf{d}_t}(\theta)|\Omega_{\mathbf{d}_t}(\theta)\right]$, and as such may introduce "chattering" into the optimization. Part B in the appendix provides more details.

[34] In a single-market entry game, the equilibrium selected by iterating over best responses would be the equilibrium that is most beneficial to the first mover. This is not the case with multiple markets. For example, assume there are two markets and two players. Duopoly is never profitable for A but is profitable for B in market 2. Player B can only open one store and prefers market 2 over 1 even when Player A is in 2. Player A prefers to be a monopolist in 2 than in 1. Player A moves first and opens in both markets. Player B then opens in market 2. Player A revisits his strategy and opens only in market 1. This is not the most beneficial equilibrium for player A. The most beneficial equilibrium for player A is for him to be a monopolist in market 2. He would have achieved that equilibrium if player B had moved first.

Chedrahui, who have an erratic expansion over the 7 year panel, show negative estimates, albeit noisy. Gigante and Waldo's have negative and significant estimates. Waldo's estimate is probably due to learning since they entered the industry in 2001. Gigante's negative estimate are puzzling. It could reflect failure of the firm to profit maximize, which would be consistent with them exiting the industry in 2008. The estimates are consistent across all specifications, increasing slightly in other specifications. Furthermore, the constraints are being relaxed over time at a rate of 7% a year. Thus firms can continue to expand at a greater pace each year.

The economic significance of the estimates cannot be addressed in isolation of the other estimates that affect costs. To address the economic significance, Table 8 displays the percentage increase in the per-store entry cost of the marginal store opened and the marginal store not-opened in a given year for each firm. That is, if WalMart opened 24 stores in 2001, the number under WalMart-2001 is the difference in cost between the 25th store and the 24th store, over the cost of the 24th, for WalMart in 2001. Notice that it is precisely this difference between the marginal store opened and the marginal store not opened that identifies the estimates. Entry costs of inframarginal stores is not identified, and as such it would be wrong to comment on the entry costs of the inframarginal stores. A Wald test is specified for each element in the table. The tested hypothesis is that the percentage increase is 0 at a 90% and 95% confidence level.

For WalMart and Comercial, each additional store increases the per-store cost of entry by 1% to 2%, and this increase is statistically significant. As such it appears both firms have many entry opportunities but can only capitalize on a few of them each year. I asses the size of this 1% in two different ways. From the model's perspective, a 1% increase in entry cost is offset by a 6% increase in market size (measured by high-income-expenditure) in the market where a firm is just indifferent between opening a store or not. To asses the 1% increase in economic terms, note that a WalMart store generates approximately \$6 M in gross profit a year.[35] This implies WalMart would incur an additional \$60,000 dollars if it were to open an additional store this year instead of next. Although significant, it is not much different than the labor cost of an additional skilled

[35] The average sales per store is ~\$ 30M and the gross profit margin is of 20%, as calculated from there 2006 Annual Report

manager.

Soriana has economically insignificant estimates. Chedrahui and Ley have noisy negative estimates that fluctuate across time. This fluctuation is driven by fluctuations in the firms' financials. As such, it appears these firms are not constrained in the rate of growth, but sometimes present economies of scale in the rate of growth. When they are not constrained it must be because they lack the entry opportunities WalMart and Comercial have. This lack of opportunities may be driven by high entry costs (financing costs) or high operating costs.[36] Either way, since their growth is limited by other costs they do not present diseconomies of scale in the rate of entry: they cannot afford entry and thus, for those stores they do open their management suffices. The negative estimate on Waldo's is consistent with an experience curve that declines over time. Since Waldo's is a new entrant in the industry it is not surprising that their costs drop in the number of stores they open.

The parameters regarding the financial constraints are also of the expected sign and significant (table 6). When a firm is more financially distressed, proxied by a high debt to assets ratio, entry costs are larger. An estimate of 0.09 is equivalent to a 3% increase in entry costs[37] for each additional point in the debt to assets ratio. Given a standard deviation of 3.9 points in the firm-demeaned distribution of Debt/Assets ratio, a firm's entry costs can vary by up to 12% due to financial costs over the time span under study.

When diseconomies of scale in entry are ignored (column 7) the coefficient on financial resources is economically and statistically insignificant. This reinforces the argument that, to be able to estimate the effect of financials on entry costs, one needs to account for managerial constraints or else the endogeneity will work to eliminate the effect: a firm with financial resources (low per-store entry costs) can open stores, but as it does, it raises costs due to limited renewable resources. If the model does not allow for this second effect, entry costs will be high when financial resources are low (due to unobserved managerial costs) as well as when financial resources are high (due to

[36]The model assumes homogenous demand across firms, thus, the lack of opportunities must be driven by covariates that are firm specific: financial constraints, location of distribution centers, or extension of network of stores (nearest and farthest active markets).

[37]For all comparisons to entry costs I will take the average cost of entry of the marginal store opened, which is 3.0 units. 3% = 0.09 x 1 / 3

observed financial costs), and they will estimate a zero effect on financial resources.

The interaction between managerial constraints and financial constraints is also statistically significant, albeit small. A negative estimate implies that a firm that lacks capital has lower *increases* in entry costs. If the increases in entry costs capture the degree of managerial constraints, the estimate is indicative of managerial and financial resources being complements. As the cost of capital increases, firms open less stores, requires less managerial resources, and, thus, the constraints due to managerial resources also decrease.

The learning estimate is positive, significant, and stable across all models. An estimate value of 1.16 represents a 39% premium in entry costs for those markets where the firm does not operate a single store. This implies learning by doing (adquiring knowledge of local taste, local suppliers and real estate locations) is much stronger than the learning option value of opening a store (opening a store as a way to resolve uncertainty). The effect of preepmtion is, surprisingly, small and insignificant. This could be justified if demand has not been fully controlled for: markets that are unobservably more profitable are also those in which competitors have already entered. As such the firm will open in such markets, biasing downwards the estimate on preemption.

One final comment on entry costs. As more controls are added to the specification, specifically column 3 in table 6, the effects discussed above are stronger. The percentage increase in entry costs between the marginal store opened and the marginal store not-opened are approximately 3% for WalMart and CCM and 1% for Soriana[38]. For the other competitors it remains insignificant.

6.2 Variable Profits

The estimates on variable profits are shown in Table 7. The effect of cannibalism among a firm's own stores is significant, as is the effect of competitor stores on profits. Surprisingly, the effect of the same-firm-stores is stronger than that of the competitors, indicating that brand differentiation may be significant. With an estimate of -0.29 for the competitor stores and an estimate of 0.51 for high-income expenditure, the market size must grow 48% to recover the losses on the firm's existing stores inflicted by entry of a competitor. This number drops to 26% for the second

[38] Tables available on request.

store a competitor opens and to 18% for the third store. As for the own-store competition effect (cannibalization effect), the estimate is -1.14. Thus, a market must grow 147% for the profitability per store to be the same as it was before the opening of the second store and 90% when opening the third store. Considering entry costs, the market must grow 160% from the size at which it was profitable to open the second store in order to justify the entry of the third store.[39]

Other market covariates are statistically significant and as expected. The effect of high-income expenditure is large, while that of low-income expenditure is negligible. Using high-income expenditure as the base for interpreting the estimates, a 1% increase in the number of state employees is offset by a 0.2% increase in high-income expenditure. Similarly, a 1% increase in the distance to the nearest DC is offset by a 0.29% increase in high-income expenditure. Noticeably, the effect of the nearest market is stronger than that of the DC; a 1% increase in the distance to the nearest market is offset by a 0.45% increase in high-income expenditure. This supports the "Chain Effect" in Jia (2008), where stores in nearby markets reinforce each other, possibly through the sharing of overhead and advertising and through improved logistics.

Border towns appear to be particularly attractive, possibly due to unobserved access to American DCs or to greater unobserved demand from American counterpart towns or from travelling consumers stalled at customs. Similarly, metro markets appear to be more attractive than the average market, possibly proxing for unobserved higher income. Surprisingly, however, tourist towns do not appear to have a larger unobserved demand. Finally, the effect of non-grocery retailers is positive and significant. This is expected given they control for unobserved demand. With an estimate of 0.094, a median of 2 non-grocery stores in a market, and a median market size of 100M, an additional non-grocery retailer is equivalent to a $4 M increase in market size.

In related work, Jia (2008) finds that population must increase 40% to offset WalMart's lost in profitability when a Kmart enters a market where WalMart has a store. This is very similar to the 48% discussed earlier. Regarding her estimates on entry costs, she finds that the profit loss

[39]Comparison is done from second to third store and not from first to second store given the Learning effect. If the learning effect is considered, the market size at which the first store is profitable is larger than that at which it is optimal to open the second store. As such I should observe firms open a store in a given market and open a second store in the following year.

a Kmart store inflicts on a WalMart store is equivalent to 4% of entry costs (her entry costs also include fixed costs since she does not distinguish between the two). This number is 6.7% with my estimates, indicative of slightly lower entry costs in the Mexican industry than in the US industry. It is surprising that, despite two very different data sets and estimation techniques, the two sets of estimates are similar, at least regarding the effect of competition relative to entry costs and market potential. Unfortunately, the estimates regarding cannibalization of a firm's own stores are not comparable for reasons described earlier.

Ellickson et al. (2010) find that the competition effect is stronger than the cannibalization effect. In Table 3 of their paper, they show that population must increase 180% for a WalMart store to offset the loss incurred by the entry of a second competitor store. This contrasts with the 48% value discussed earlier. Furthermore, they find that the population increase required to offset the profit loss due to cannibalization when opening a second store amounts to 27,000 people, or 54%. This contrasts with the 147% required increase in high-income market expenditure implied by the base-model estimates.

6.3 Multiple Equilibria and Likelihood

The estimates do not change when using the alternative equilibrium selection rule. That does not guarantee that they would change under some third equilibrium selection rule. Nevertheless, breaking down the likelihood into three parts may shed additional light on how the estimates are insensitive to the equilibrium selection rule. Table 9 shows the values of the likelihood for each of the three parts the conform the likelihood. It also shows a norm[40] for the second derivative of each of those parts of the likelihood with respect to the parameter vector (for the Hessian of the objetive function). A small norm implies the objective function is flat, and as such insensitive to the parameter vector. Since the log-likelihood is the sum of three diferent sets of probabilities, the total hessian can also be expressed as the sum of three different hessians. The table shows how the hessian for the probability that varies with the different equilibrium selection rules is

[40] I use the Euclidean operator norm: the square root of the largest eigenvalue of the square positive definite matrix X'X.

very flat. That is, the specific assumed equilibrium selection rule has no power in identifiying the parameters of interest. The parameters are being identified off the markte-specific necessary condtiions ($\lambda^1 - \lambda^3$). As noted before, these hessians are evaluated at the optimal paramter vector given the equilibrium selection rule and could be very different at an alternative parameter vector or under an alternative equilibrium selection rule. Nevertheless, the fact that under these two equilibrium selection rules provides sugestive evidence that alternative equilbrium selection rules may also provide little information towards identifying the parameters.

7 Conclusions

In this paper I attempt to quantify the determinants of gradual growth observed in the Mexican retail industry. One of such determinants is the firm constraints to the rate of expansion; and, more specifically, increasing marginal entry costs. Other determinants include firms' financial situation, increasing demand, and intermarket dynamics. By appropiately accounting for firm's financials, the increases in marginal entry costs may be attributed mostly to managerial constraints. Such constraints arise because of the market for specialized managers, whose labor input is renewable over time: after opening a store the manager is free to open a second store. With increasing costs for the managers' time, a firm expands gradually over time, in the form of a rollout process.

In the context of the Mexican supermarket industry, I do find evidence that managerial constraints can explain the rollout patterns observed in the data. After controlling for the firms' financial situation, I find that marginal sunk entry costs rise by 1 to 2%. This effect is economically significant: for a store whose expected profits per year are $6M, a 1% increase in cost implies the firm would have to pay $60k to open the store a year earlier. With such high entry costs, it could be in the firms' interest to invest in managers early on. If there is a tension between investing in managerial capacity and investing in increasing demand, the former may be better than the latter when managerial constraints are so large. I also find that financial differences across firms explain a large portion of the differences in entry patterns. A firm whose financial performance is one standard deviation above the mean enjoys a 12% reduction in the cost of opening a store.

Competition effects are similar to those found in other literature, in which markets must grow 48% to offset the losses a competitor induces on entry.

To estimate the constraints on entry a structural model is proposed. Given the limited data that is available, strong simplifying assumptions are imposed on the model so as to reduce the level of dependency of outcomes in one market on those in other markets. However, these dependencies cannot be completely removed. To be able to estimate the model, a static game between firms is solved and an equilibrium selection mechanism is assumed. A likelihood function is formed based on the predictions of the game and a GHK type Montecarlo integration is used to obtain probabilities. Some robustness tests are used to show that the equilbrium selection rule specified has little impact on the parameters. These tests can also be used in other work where multiple equilbria is present to find out, on a first approach, if the equilbirum selection rule affects or not the parameters of interest.

The model is highly simplified to retain tractability. With constrained growth, a dynamic model would have been more appropriate but would have required estimating a dynamic model with a huge state space. This leaves open the door to future research: estimating dynamic models with large state spaces. Other future lines of research include estimating entry complementarities that may exist among multi-business unit firms that do depend on the location of entry, such as opening multiple stores in the same market compared to opening multiple stores in multiple markets.

References

Bajari, P., L. Benkard, and J. Levin: 2007, 'Estimating Dynamic Models of Imperfect Competition'. *Econometrica* **75**(5), 1331–1370.

Bresnahan, T. and P. Reiss: 1990, 'Entry in Monopoly Markets'. *The Review of Economic Studies* **57**(4), 531–553.

Bresnahan, T. F. and P. C. Reiss: 1991, 'Entry and Competition in Concentrated Markets'. *The Journal of Political Economy* **99**(5), 977–1009.

Cameron, A. C. and P. Trivedi: 2005, *Microeconometrics Methods and Applications*. Cambridge University Press.

Chernozhukov, V., H. Hong, and E. Tamer: 2007, 'Estimation and Confidence Regions for Parameter Sets in Econometric Models¹'. *Econometrica* **75**(5), 1243–1284.

Ciliberto, F. and E. Tamer: 2009, 'Market Structure and Multiple Equilibria in the Airline Industry'. *Econometrica*.

Durand, C.: 2007, 'Externalities from foreign direct investment in the Mexican retailing sector'. *Camb. J. Econ.* p. bel046.

Ellickson, P. B., S. Houghton, and C. Timmins: 2010, 'Estimating Network Economies in Retail Chains: A Revealed Preference Approach'. *NBER Working Paper 15832* p. 44.

Fox, J.: 2010, 'Estimating Matching Games with Transfers'. *NBER Working Paper No. 14382*.

Fuentes, H. J., A. Zamudio, C. Ortega, J. Mendoza, and J. Soto: 2008, 'Servicio de Elaboracion del Estudio de: "Evasion Fiscal Generada Por el Ambulantaje"'. *ITESM, Campus Ciudad de Mexico*.

Gourieroux, C. and A. Monfort: 1991, 'Simulation Based Inference in Models with Heterogeneity'. *Annals d'Economie et de Statistique* pp. 69–107.

Gourieroux, C. and A. Monfort: 1996, *Simulation Based Econometric Models*. CORE Lectures, OXFORD University Press.

Holmes, T. J.: 2011, 'The Diffusion of Wal-Mart and Economies of Density'. *Econometrica* **79**(1), 253–302.

Iacovone, L., B. Javorcik, W. Keller, and J. Tybout: 2009, 'Supplier Responses to Wal-Martís Invasion of Mexico'. *working paper*.

Instituto Nacional de Estadística, G. e. I.: 2004a, 'Censos Economicos 2004'.

Instituto Nacional de Estadística, G. e. I.: 2004b, *La Ocupación en el Sector no Estructurado en México, 1995-2003*.

Ishii, J.: 2008, 'Compatibility, Competition, and investment in Netowrk Industries: ATM Networks in the Banking Industry'. *working paper*.

Jia, P.: 2008, 'What Happens When Wal-Mart Comes to Town: An Empirical Analysis of the Discount Retailing Industry'. *Econometrica* **76**(6), 1263–1316. 10.3982/ECTA6649.

Manski, C. F.: 1975, 'Maximum score estimation of the stochastic utility model of choice'. *Journal of Econometrics* **3**(3), 205–228.

Mazzeo, M. J.: 2002, 'Product Choice and Oligopoly Market Structure'. *The RAND Journal of Economics* **33**(2), 221–242.

Nishida, M.: 2009, 'Estimating a Model of Strategic Netwrok Choice: the Convenience-Store Industry in Okinawa'. *NET Institute Working Paper No. 08-27*.

Pakes, A.: forthcoming, 'Alternative Models for Moment Inequalities'. *Econometrica*.

Pakes, A., J. Porter, K. Ho, and J. Ishii: 2006, 'Moment Inequalities and Their Application'. *working paper*.

Penrose, E.: 1995, *The Theory of the Growth of the Firm*. OXFORD University Press, 3rd edition edition.

Seim, K.: 2006, 'An Empirical Model of Firm Entry with Endogenous Product-Type Choices'. *The RAND Journal of Economics* **37**(3), 619–640.

Sutton, J.: 2001, *Technology and Market Structure*. MIT Press, 1st edition.

Tamer, E.: 2003, 'Incomplete Simultaneous Discrete Response Model with Multiple Equilibria'. *The Review of Economic Studies* **70**(1), 147–165.

Toivanen, O. and M. Waterson: 2005, 'Market Structure and Entry: Where's the Beef?'. *The RAND Journal of Economics* **36**(3), 680–699.

APPENDIX

A Proof of Redundant Deviations

The non-redundant inequalities to problem 6 are: (1) in each market, holding decisions in all other markets constant, add one more store to the final count (either by opening one more store in markets where exit did not occur, or by closing one less store in markets where exit did occur); add one less store to the final count wherever possible (either by closing one more store in markets where entry did not occur and the final count has a positive number of stores, or by opening one less store in markets where entry did occur); and (2) not opening the least profitable store that was opened and open instead the most profitable store not opened. Deviation (1) is captured in the first three inequalities in 6 while deviation (2) is captured in the fourth inequality.

I prove the statement in four parts. First I show that the profit function is concave. This implies that all deviations in which the net number of stores in a market is different than the observed strategy by two or more stores are redundant once all deviations of one or less stores are accounted for. Concavity is not sufficient to reduce the set of non-redundant inequalities to those specified in (1) above because of the integer problem. It would be sufficient if the profit function were continous in the decision variables. The second part of the proof shows that it is sufficient to consider only inequalities in which the net number of stores across all markets differs from the observed decision by one or less stores. This can be achieved by the additivity of the cost function, by the absence of intermarket dependencies that are not through the cost function, and because the intermarket dependencies introduced by the cost function are the same across all markets. The third part of the proof shows that deviation (2) described above renders redundant all deviations in which the net difference in the number of stores across all markets is zero. The last part of the proof shows how (1) and (2) together render redundant all other deviations in which the net difference in the number of stores across all markets is one or minus one. This completes the proof.

Part 1. Concavity of the Profit Function in each Decision Variable ($\{d^e_{imt}, d^x_{imt}\}_{m=1}^M \in \mathbb{N}^{2M}$)

Variable profits for each market have the form $x(a - b\ln(c + x))$, which is concave in x . The variable profits of one market do not depend on the entry/exit decisions in alternative markets. The cost function is convex in entry decisions and flat in exit decisions. The cost function affects negatively profits. Thus, the profit function is concave in each entry decision variable. Regarding exit decisions, the variable profits in each market have the form $-x(a - b\ln(c - x))$, which is concave in x . Thus, the profit function is concave in each exit decision.

Since the profit function is concave in each decision variable independently of the other decision variables, all deviations from the observed strategy of two or more units in a given market are redundant.

Part 2. Consider only deviations where the total number of entries and exits are one unit different than the observed decision.

Let $\mathbf{x}$ be a valid deviation from the optimal decision $\mathbf{d}^\star$, such that $x^e_m = d^{e\star}_m$ and $x^e_n = d^{e\star}_n + 1$. Let $\mathbf{y}$ be another valid deviation such that $(y^e_l, y^x_l) = (x^e_l, x^x_l)\ \forall l \neq (m, n)$ and $y^e_m = d^{e\star}_m + 1$, $y^e_n = d^{e\star}_n$. Furthermore, let it be that $\pi(\mathbf{d}^\star_i|\mathbf{d}_{-i}, \mathbf{s}, \chi_i) \geq \pi(\mathbf{x}|\mathbf{d}_{-i}, \mathbf{s}, \chi_i)$ and $\pi(\mathbf{d}^\star_i|\mathbf{d}_{-i}, \mathbf{s}, \chi_i) \geq \pi(\mathbf{y}|\mathbf{d}_{-i}, \mathbf{s}, \chi_i)$. That is, $\mathbf{x}$ and $\mathbf{y}$ are two deviations that are known to not be better than the optimal strategy, and they are the same deviation everywhere except for in markets m and n, in which $\mathbf{x}$ follows $\mathbf{d}^\star$ in m and $\mathbf{y}$ does so in n. Let $\mathbf{z}$ be a third feasible deviation such that $(z^e_l, z^x_l) = (x^e_l, x^x_l)\ \forall l \neq (m, n)$ and $z^e_m = y^e_m$ and $z^e_n = x^e_n$. Then $\pi(\mathbf{d}^\star_i|\mathbf{d}_{-i}, \mathbf{s}, \chi_i) \geq \pi(\mathbf{z}|\mathbf{d}_{-i}, \mathbf{s}, \chi_i)$ by convexity of the cost function and additivety across markets.

A similar argument can be applied for having one less entry and similar dual arguments can be used with exit. Thus, no non-redundant deviation can have the total number of entries across all markets being more than 1 unit apart from the observed strategy as they would contradict the above result.

Part 3. Equation 4 in6 implies any deviation in which the total number of entries and exits are the same as the observed strategy (although the locations may be different) are redundant.

A deviation in which the the total number of entries/exits do not change but the location does is one in which entry/exit decisions are swapped between two or more markets. By Part 1 of the proof we need only consider deviations within a market of one unit. Thus the entry/exit decisions are swapped between pairs of markets. Because of additivity of the profit function across markets it is enough to consider each pair of swaps in isolation. That is, if it is not profitable to swap the entry decisions between markets A and B, and it is not profitable to do so between markets C and D, then it is not profitable to swap A with B and C with D simultaneously.

Equation 4 in 6 states that the most profitable entry-pair-swap among all pairs of markets is not profitable. That is, if the firm would hold back the entry of the least profitable market and enter instead the most profitable market, it would not fair better. If the best possible pair-swap is not profitable, than any other pair-swap is not profitable either.

Exit is free. Thus, if is not profitable to close a store in market A when following the firms optimal strategy in the other markets, it will not be profitable to do so under any other deviation.

Part 4. All deviations in which the net number of entries/exit are 1 unit apart fromt the observed strategy are redundant given all 4 equations in 6.

By Part 1 we need consider only deviations within a market of one unit. As such any deviation in which the net number of entries/exits is different from the observed strategy by exactly one unit will be one in which the firm opens one more (one less) store in a specific market and swaps entry/exit decisions in other markets. As stated in Part 3, it is enough to consider those deviations in which the firm swaps entry/exit decisions in a single pair of markets and additionally opens one more (one less) store in a third market. The additional opening of one more (one less) store is not profitable in isolation by the first 3 equations in 6. The pair-swap of entry is not profitable in isolation either by equation 4 in 6. Both deviations executed jointly will not be profitable either because of the additivity of the profit function.

This completes the proof.

B Simulated Maximum Likelihood and Bayes' Rule

This section shows how separating the likelihood into two parts, one of which has a closed form solution, diminishes significantly the error introduced by simulation. I follow Gourieroux and Monfort (1996) and Cameron and Trivedi (2005). Let $f_N(\theta)$ be a probability and $\theta^\star = \arg\max_{\theta\in\Theta} \mathbb{E}_N\left[\ln f_N(\theta)\right]$ be a unique maximizer. Let $\hat{\theta} = \arg\max_{\theta\in\Theta} \frac{1}{N}\sum_{i=1}^{N} \ln f_i(\theta)$ and $\hat{\theta} \xrightarrow{p} \theta^\star$ with

$$\sqrt{N}\left(\theta^\star - \hat{\theta}\right) \xrightarrow{d} N(0, \mathbf{A}^{-1}(\theta^\star)) \qquad \mathbf{A}(\theta^\star) = -\mathbb{E}_N\left[\frac{1}{N}\mathbf{H}_N(\theta^\star)\right]$$

In other words, $\hat{\theta}$ is a consistent, unbiased Maximum Likelihood estimator of $\theta^\star$, where the log-likelihood for a finite N is $\frac{1}{N}\sum_{i=1}^{N} \ln f_i(\theta)$. If a closed form of $f_i(\theta)$ is unknown but there is an $f_i(\theta|r)$ and a sequence of independent draws $\{r_j\}_{j=1}^{R}$ such that $\tilde{f}_{R,i}(\theta) = \frac{1}{R}\sum_{r=1}^{R} f_i(\theta|r)$ and $\tilde{f}_{R,i}(\theta) \xrightarrow{p} f_i(\theta)$, then an alternative estimator can be used:

$$\hat{\hat{\theta}} = \arg\max_{\theta\in\Theta} \frac{1}{N}\sum_{i=1}^{N} \ln \tilde{f}_{R,i}(\theta)$$

Cameron and Trivedi (2005) show that if $\sqrt{N}/R \to 0$, then $\sqrt{N}\left(\theta^\star - \hat{\hat{\theta}}\right) \xrightarrow{d} N(0, \mathbf{A}^{-1}(\theta^\star))$. That is, the maximum simulated likelihood estimate converges in distribution to the same distribution as does the Maximum Likelihood estimate.

Consistency requires that the number of draws, R, go to infinity much faster than N. But for a finite N, how many draws are required? Cameron and Trivedi (2005) shed some light on this question by studying the first-order bias of the log-likelihood function for finite draws (R):

$$\ln f_i(\theta) - \mathbb{E}_R\left[\ln \tilde{f}_{R,i}(\theta)\right] = \frac{1}{2}\frac{\mathbb{E}_R\left[\left(\tilde{f}_{R,i}(\theta) - f_i(\theta)\right)^2\right]}{f_i(\theta)^2} \tag{8}$$

As is evident in (8) , the bias is inversely proportional to the probability squared. If the probability is low, the bias can be significant. The bias can be reduced by increasing the number of draws, which would reduce the numerator: $\mathbb{E}_R\left[\left(\tilde{f}_{R,i}(\theta) - f_i(\theta)\right)^2\right] = \frac{1}{R}\sigma_r$. Using Bayes' Rule allows the probability being simulated to be much larger, thus achieving a small bias without having to incur a large number of draws.

A Mixed Maximum Simulated Likelihood contains a part that is simulated and a part that is not. Consistency and asymptotic properties of the MMSL are the same as those of the MSL when $\sqrt{N}/R \to 0$. The proof is a trivial extension of the proof presented in Gourieroux and Monfort (1991).[41]

When applying Bayes' Rule to the probability, the draws for the simulated probability have to be simulated from a conditional distribution. If there is a known, closed form, transformation between the uniform distribution and the conditional distribution many draws may be taken from the unit-uniform distribution and transformed, in each step of the optimization, into draws from the conditional distribution. If there is no closed form transformation, draws must be simulated from an unconditional distribution and screen accordingly to the restrictions of the conditional distribution. This may cause "chattering" in the optimization.

[41] Let $f^\star(y_i|x_i, u_{hi}, \theta) = f_1^\star(y_i|x_i, u_{hi}, \theta) + f_2^\star(y_i|x_i, \theta)$ in equation (9) of Gourieroux and Monfort (1991).

Figure 1: Total number of stores from chain retailers in Mexico over time

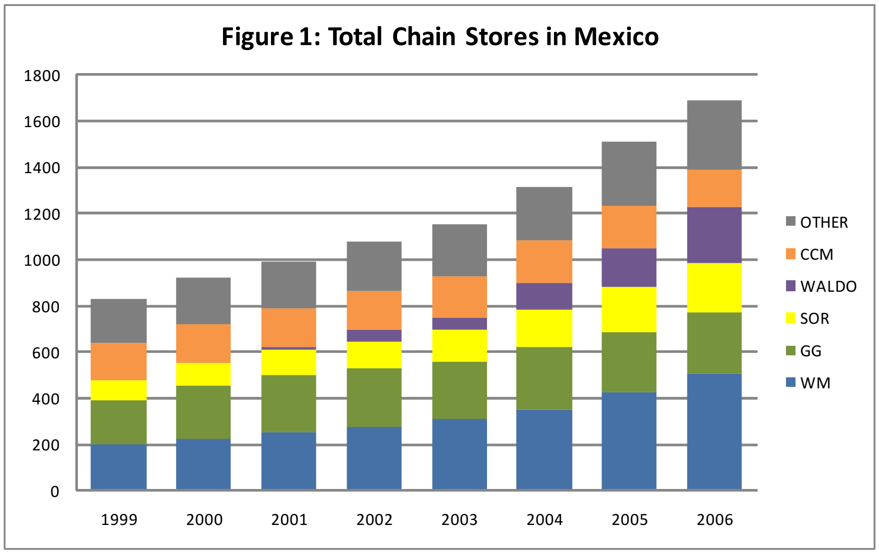

Table 1: Distribution of store formats by firm in 2006, in percentage points

	Hypermarkets (82k ft^2)	Clubs (94k ft^2)	Bodegas (52k ft^2)	Supermarkets (16k ft^2)
Walmart	23	15	50	12
Waldo's	0	0	0	100
Soriana	74	8	17	0.5
CCM	55	14	6	25
Ley	50	0	0	50
GG	31	0	11	58
Chedrahui	90	0	0	10

Table 2: 2006 Snap Shot of the Mexican Supermarket Industry

	Markets	Mean	Std. Dev	Min	Max
# National Chain stores	416	2.89	6.65	0	57
# National Chain firms	416	1.20	1.69	0	7
# Non-Nat. Chain stores	416	1.10	3.69	0	33
# Non-grocery retailers	416	6.64	16.63	0	182

Figure 2: Aerial view of the median (by market potential) market: Sahuayo, MICH

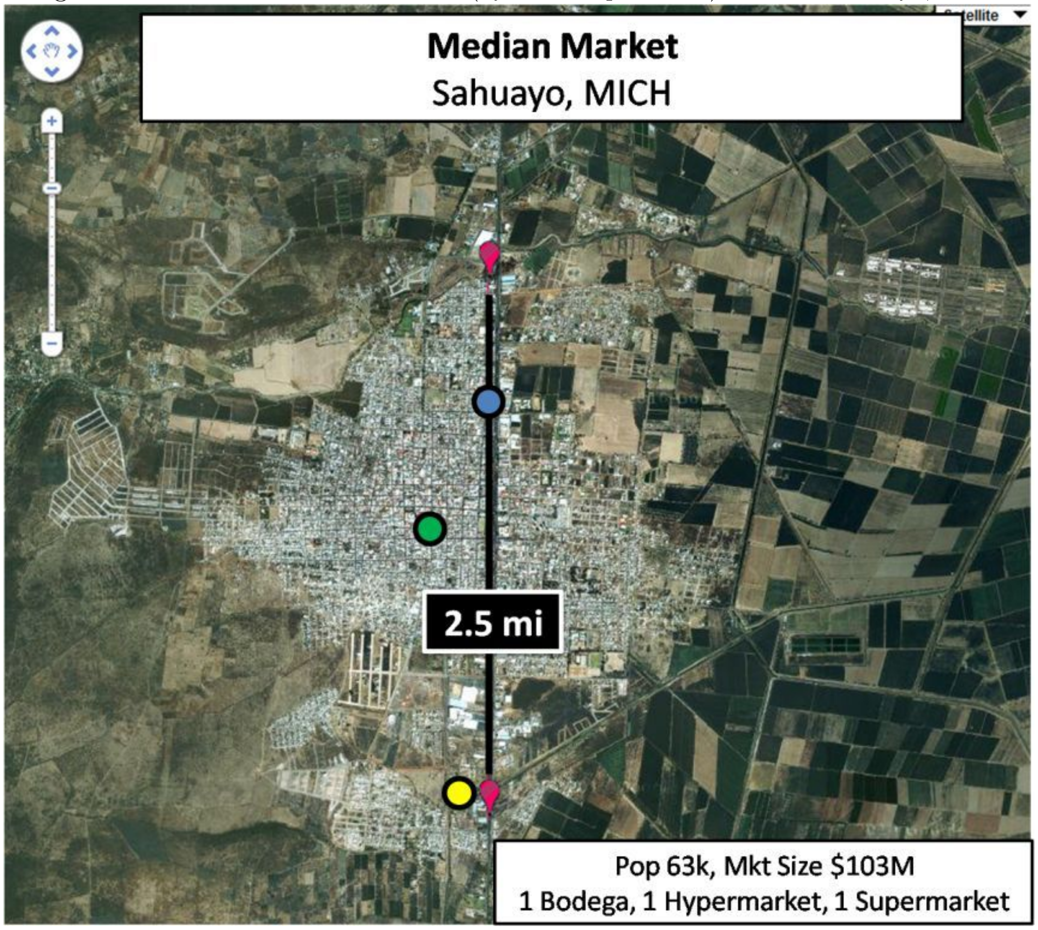

Table 3: Summary Statistics

		Obs	Mean	Std. Dev.	Min	Max
Firm-Mkt-Year	# Stores	22680	0.30	1.28	0	33
	# Stores	2369*	2.51	2.82	1	33
	Dist to DC (mi)	2369*	226	218	0	1430
	Dist to closest mkt	2369*	69	106	4	1446
	Dist to farthest mkt	2369*	1224	332	49	2025
Mkt-Year	MS high inc ($M)	2912	210	433	0.17	3,426
	MS low inc ($M)	2912	70	95	0.63	767
	Non-grocery stores	2912	4.93	13.98	0	182
	Non-chain stores	2912	1.04	3.41	0	35
	Gvt. Workers (ln)	2912	8.23	13.87	0	86
	Income (ln $k/yr)	2912	6.17	2.53	2.78	13
	Market Growth (%)	2912	1.37	0.71	-23	27
	Density (ln 1/acre)	2912	1.24	3.53	0.00	44
	MSA (GDL, MTY)	2912	0.04	0.19	0	1
	Toursit town	2912	0.09	0.29	0	1
	Border Town	2912	0.05	0.21	0	1

* Only for observations where the firm has stores
** 7 firms, 416 markets, 7 years - excludes Mexico City

Table 4: Summary Statistics on Entry and Exit patterns

Entries

Firm's Presence			Competitors' Presence			Competitors' Action		
First store	336	47%	No stores	108	15%	Leaving	55	8%
Second store	97	14%	1 Store	59	8%	Entering	259	36%
3rd+ store	281	39%	2+ Stores	547	77%	Nothing	400	56%
Total	**714**			**714**			**714**	

Exits

Last store	17	19%	No Stores	4	4%	Leaving	4	4%
2nd to last store	20	22%	1 Stores	7	8%	Entering	62	68%
Other	54	59%	2+ Stores	80	88%	Nothing	25	27%
Total	**91**			**91**			**91**	

Table 5: Summary Statistics of Firm level data

	Firm-Years	Mean	SDev	Min	Max
Entries (total per firm-year)	49	16.6	17.8	0	69
Exits (total per firm-year)	49	3.8	5.1	0	22
Entries (including Mexico City)	49	20.3	22.2	0	82
Debt to Assets ratio*	35	0.47	0.13	0.27	0.77
Firm demeaned D/A Ratio	35	0.00	3.86	-10.8	6.1
Op Cash Flow to Debt ratio*	35	0.20	0.12	-0.13	0.47
Firm demeaned CF/D Ratio	35	0.00	7.44	-21.6	17.8

* Excludes Waldo's and Tiendas Chedrahui

Figure 3: Entry Thresholds and Firms Profits

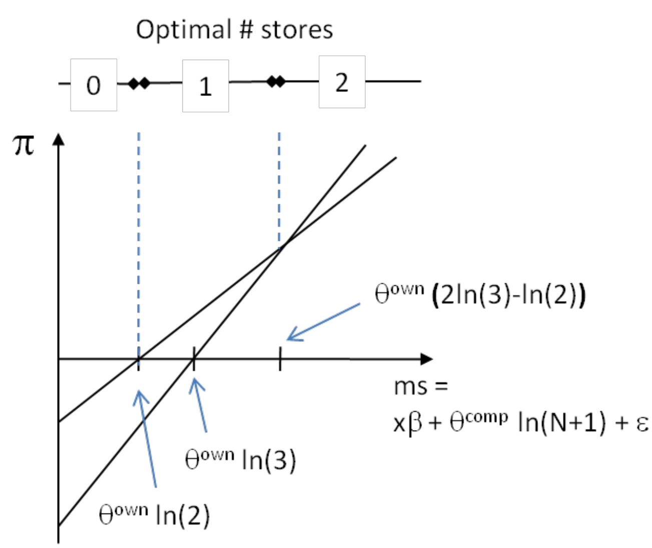

Table 6: Regression results - Entry Costs

	Entry Costs						
MODEL:	Base	Controls	Alternative Equilibrium	Alternative Financial	Excludes Dynamics	Including Exits in w	No Convex Costs
Convex Costs							
Comercial M	0.26 *	0.27	0.26 *	0.19	0.06	0.19	N/A
	(0.113)	(0.307)		(0.202)	(0.070)	(0.123)	
Chedrahui	-0.21	-0.16	-0.21	-0.22	-0.14 *	-0.20	
	(0.135)	(0.109)		(0.132)	(0.060)	(0.121)	
Gigante	-0.24 *	-0.28 *	-0.24 *	-0.25 *	-0.04	-0.29 *	
	(0.053)	(0.066)		(0.052)	(0.083)	(0.058)	
Ley	-0.18 *	-0.20 *	-0.18	-0.19 *	-0.12 *	-0.18 *	
	(0.050)	(0.103)		(0.069)	(0.030)	(0.080)	
Soriana	0.05	-0.04	0.05	-0.08	-0.16 *	0.04	
	(0.198)	(0.213)		(0.185)	(0.047)	(0.200)	
Waldo's	-0.44 *	-0.50 *	-0.44 *	-0.45 *	-0.22 *	-0.44 *	
	(0.061)	(0.078)		(0.059)	(0.066)	(0.060)	
WalMart	1.19 *	1.23 *	1.19 *	1.17 *	-0.23 *	1.17 *	
	(0.062)	(0.106)		(0.070)	(0.096)	(0.194)	
Financial Costs							
Theta F	0.09	0.09	0.09	-0.03 *	0.03 *	0.09 *	0.01
	(0.048)	(0.051)		(0.013)	(0.018)	(0.043)	(0.017)
Theta FC	-0.02	-0.02	-0.02	0.007	-0.002	-0.020	N/A
	(0.013)	(0.016)		(0.004)	(0.004)	(0.011)	
Other							
Learning	1.16 *	1.18 *	1.16 *	1.16 *	NA	1.16 *	1.16 *
	(0.075)	(0.077)		(0.075)		(0.072)	(0.069)
Preemption	-0.01	-0.04	-0.01	-0.01	NA	-0.01	0.01
	(0.115)	(0.121)		(0.116)		(0.116)	(0.109)
Time Trend	-0.07 *	-0.06 *	-0.07 *	-0.06 *	0.11	-0.07 *	N/A
	(0.005)	(0.004)		(0.005)	(0.069)	(0.006)	
Linear Costs							
Comercial M	2.58 *	2.0 *	2.58 *	2.74 *	3.54 *	2.69 *	3.18 *
	(0.290)	(0.691)		(0.469)	(0.251)	(0.320)	(0.189)
Chedrahui	3.79 *	3.76 *	3.79 *	3.80 *	4.66 *	3.80 *	3.44 *
	(0.244)	(0.276)		(0.247)	(0.212)	(0.256)	(0.209)
Gigante	3.97 *	3.74 *	3.97 *	4.04 *	3.98 *	4.24 *	3.30 *
	(0.244)	(0.374)		(0.256)	(0.407)	(0.255)	(0.181)
Ley	3.51 *	4.37 *	3.51 *	3.51 *	4.27 *	3.52 *	3.11 *
	(0.257)	(0.461)		(0.249)	(0.222)	(0.321)	(0.200)
Soriana	2.62 *	3.23 *	2.62 *	3.04 *	4.19 *	2.67 *	2.79 *
	(0.636)	(0.807)		(0.614)	(0.227)	(0.646)	(0.181)
Waldo's	3.62 *	4.14 *	3.62 *	3.70 *	4.49 *	3.65 *	2.24 *
	(0.290)	(0.283)		(0.294)	(0.286)	(0.291)	(0.167)
WalMart	-1.94 *	-1.37 *	-1.94 *	-1.94 *	4.38 *	-1.92 *	2.31 *
	(0.260)	(0.395)		(0.182)	(0.425)	(0.767)	(0.163)
Log-Likelihood	-2,620	-2,573	-2,620	-2,988	-2,772	2,621	2,700

* Sample includes 7 firms over 7 years operating in 416 markets. Hubber-White standard errors in parenthesis. Stared estimates significantly different than 0 at 5%

** Controls include, in addition to those presented in table 7, Time dummies for entry costs, Firm dummies for variable profits, and the censored number of non-retail grocery stores opened/closed.

*** Financial measure used in base model is Debt/Assets of the previous year. Alternative financial measure is Net Cash Flows / Debt of the previous year

**** Base equilibrium selection rule derived from iterated best responses with the order of movement being: Waldo's, WalMart, Soriana, Gigante, Comercial, Ley, Chedrahui. Alternative ESR has firms move in alphabetical order.

Table 7: Regression results - Variable Profits

MODEL:	Base	No Convex Costs
Variable Profits		
Competition		
Own	-1.14 *	-1.10 *
log(n+1)	(0.047)	(0.045)
Competitors	-0.29 *	-0.27 *
log(n+1)	(0.066)	(0.062)
Market Potential		
High-Inc Expenditure	0.51 *	0.49 *
log($ B)	(0.069)	(0.064)
Low-Inc Expenditure	0.039	-0.018
log($ B)	(0.065)	(0.061)
Non-grocery stores	0.09 *	0.08 *
log(n+1)	(0.032)	(0.035)
Local grocery stores	0.11 *	0.10 *
log(n+1)	(0.049)	(0.047)
State Employees	-0.095	-0.117
log('000)	(0.138)	(0.140)
GDPPC	1.20	0.90
log($ k/yr)	(0.63)	(0.59)
Mkt Growth Rate	-0.047	-0.040
	(0.027)	(0.027)
Pop Density	-0.61 *	-0.60 *
log(1/acre)	(0.092)	(0.092)
MTY/GDL metro area	0.17 *	0.13 *
Dummy	(0.062)	(0.060)
Tourist Town	0.009	0.001
Dummy	(0.101)	(0.100)
Border Town	5.94 *	5.88 *
Dummy	(0.243)	(0.805)
Constant	0.24 *	0.26 *
	(0.049)	(0.052)
Supply Covariates		
Dist to DC	-0.15 *	-0.14 *
log(mi)	(0.030)	(0.027)
Dist to nearest mkt	-0.23 *	-0.20 *
log(mi)	(0.034)	(0.048)
Dist to farthest mkt	-0.09 *	-0.13
log(mi)	(0.045)	(0.113)

* Hubber-White standard errors in parenthesis. Refer to Table "Entry Costs" for notes. Estimates of other models almost identical to Base Model.

Table 8: Estimated Percentage Increase in Marginal Cost of Entry

% Increase in Per-Store Cost of Entry evaluated at the number of stores opened that year

Firm	2000	2001	2002	2003	2004	2005	2006
Comercial	1.08**	1.78	1.59**	0.81**	0.74**	0.56**	0.50**
Chedrahui	-7.81*	-1.98	-1.82	NA	NA	-0.42	-0.73
Gigante	-0.07	-0.42**	-1.20**	NA	-0.16**	-0.43**	-0.87**
Ley	-0.76**	-1.21	-0.74	-0.91**	NA	-4.92**	-0.14**
Soriana	0.84	-0.14	-0.04	0.04	0.04	0.03	0.05
Waldo's	NA	NA	-0.52**	NA	-0.24**	-0.25**	-0.15**
WalMart	2.18**	1.32**	1.35**	1.12**	1.01**	0.50**	0.42**

NA - firm did not open a single store that year
Wald test rejects hypothesis that value is equal to 0 at a 90% confidence (*) or a 95% confidence (**).

Table 9: Breakdown of Likelihood Function and Hessian

	Likelihood	Value	Hessian L2 Norm
Base Model	log(Pr[ESR | equilibrium])	-3.92	
	log(Pr[lambda 4 | lambda 1-3])	-12.67	1.74E+03
	log(Pr[lambda 1-3])	-2603.4	4.96E+05
	Total	-2620.0	
Alterna-tive ESR	log(Pr[ESR | equilibrium])	-3.89	
	log(Pr[lambda 4 | lambda 1-3])	-7.58	1.83E+03
	log(Pr[lambda 1-3])	-2,608.5	4.95E+05
	Total	-2,620.0	