# Competing Retailers and Inventory: An Empirical Investigation of U.S. Automobile Dealerships*

Marcelo Olivares · Gérard P. Cachon

The Wharton School, University of Pennsylvania, Philadelphia PA, 19104
maolivar@wharton.upenn.edu · cachon@wharton.upenn.edu
opim.wharton.upenn.edu/~maolivar · opim.wharton.upenn.edu/~cachon

June 4, 2007

## Abstract

In this study we estimate empirically the effect of local market conditions on inventory holdings of U.S. automobile dealerships. We show that the influence of competition on a retailer's inventory holdings can be separated into two mechanisms: (1) the entry or exit of a competitor can change a retailer's demand (a sales effect); (2) the entry or exit of a competitor can change the amount of buffer stock a retailer chooses to hold, which influences the probability a consumer finds a desired product in stock (a service level effect). The sales effect can influence inventory through the presence of economies of scale. Theoretical arguments of inventory competition are ambiguous on the expected sign of the service level effect. We obtained data (via a web crawler) on inventory and sales of auto dealerships of a large manufacturer. Using cross-sectional variation of dealers in isolated markets, we estimated the effect of market structure (number and type of competitors) and sales on inventory levels. We used market population as an instrumental variable to control for the endogeneity of market structure. Our results suggest a strong positive non-linear effect of the number of rivals on service levels, an effect that is comparable to the sales effect. Counterfactual experiments indicate that reducing the dealership network of this manufacturer (thereby reducing competition) could reduce the remaining dealers' days-of-supply (inventory divided by average sale rate) from 14% to 27%.

*The authors thank Joel Waldfogel for his guidance and suggestions. We are grateful to Alan Abrahams, Michelle Gallert and Gabe Silvasi for their help in the implementation of the web-crawler. We thank the suggestions of seminar participants at Washington University, University of Maryland, University of North Carolina, New York University, Georgia Tech, Columbia, Harvard, University of Virginia, University of Texas at Austin, Stanford, INSEAD, University of Chicago, Northwestern University and UCLA. We also thank the Fishman-Davidson Center, The Wharton School, University of Pennsylvania, for financial support.

**Keywords**: Inventory competition, empirical, entry, supply chain management, automobile industry.

# 1. Introduction

There exists a substantial literature on how firms should manage inventory (e.g. Zipkin (2000)), but less is known about how firms actually manage inventory. In particular, why does there exist considerable heterogeneity in inventory holdings, even across firms within a single industry? For example, Table 1 displays data on inventory holdings in the U.S. automobile industry supply chain. During the years 1999-2004, Chevrolet held on-average a 73 day supply of inventory in its supply chain whereas Toyota held only about a 42 day supply. Cachon and Olivares (2006) demonstrate that differences in manufacturing flexibility and product assortments across auto makes (i.e., brands) explain nearly half of this gap, but, due to the nature of their data, they are unable to assess whether differences in dealership structure is also a factor. According to Table 1, there are indeed significant differences in dealership structure: e.g., Toyota has fewer than a third of the number of dealerships as Chevrolet (1200 vs. 4227), but each of their dealership sells on average twice as many vehicles per year (1251 vs. 627).

Inventory theory suggests that there generally are economies of scale in managing inventory. Hence, Toyota may carry less inventory in part because their dealerships operate at a higher sales rate. Some Chevrolet dealers sell considerably more vehicles per year than other Chevrolet dealerships, so economies of scale (with respect to sales) could potentially explain heterogeneity in inventory holdings even within a make. However, the U.S. auto makes exhibit other significant differences across their dealership structures. For example, we expect that General Motors (GM) dealerships are geographically located closer to each other than Toyota dealerships because there are simply more of them in the U.S. Furthermore, dealerships are not uniformly distributed across the country. Figure 1 plots the relationship between the ratio of the number of GM dealerships to Japanese brand dealerships (Toyota, Honda and Nissan) by state relative to each state's population growth. In states with large population growth from 1950 - 2004, such as California and Arizona, there is approximately the same number of GM and Japanese brand dealerships, whereas in slow growth states, such as Iowa and South Dakota, GM dealerships are much more numerous in a relative sense. As a result there exists significant variation in both the number and type of dealerships in local markets. This variation regulates the degree to which dealerships compete and theoretical models predict that inventory is influenced by the intensity of competition. (Interestingly, as we elaborate on in section 2, there is disagreement in the theoretical literature as to how competition influences inventory, i.e., does competition lead firms to hold more or less inventory.) Finally, in addition to scale and competition effects, we expect that inventory holdings could vary due to differences in consumer characteristics across markets. For example, all else being equal, a firm's optimal inventory decreases as its customers become more patient, i.e., more willing to wait for their preferred product rather than choose a competitor's product. Indeed, European consumers expect to wait to receive their vehicles, so dealerships in Europe carry a very limited supply of on-hand vehicles. Although the U.S. market operates in a very different fashion (most vehicles in the U.S. are purchased from dealer inventory), there is significant heterogeneity in consumer demographics across the country, which could lead to differences in buying behavior.[1]

The objective of this paper is to identify how inventory holdings of U.S. automobile dealerships are influenced by economies of scale with respect to sales and market structure (the term we use to describe local competition). We developed a web-crawler to collect data on GM dealerships located in more than 200 isolated markets. Because there is little heterogeneity in dealership structures across time, our empirical strategy is to exploit the cross-sectional variation in these markets to identify the effects of interest. We use instrumental variables to control for the endogeneity of market structure with respect to unobserved market characteristics. We focus on the auto industry because it is economically significant and detailed data on local inventory holdings are available (via our web-crawler). Although our results are specific to this industry, our econometric methods could be applied to study inventory in other retail industries. Furthermore, some of our findings may apply broadly to other forms of retailing.

Our study is related to the growing empirical literature on inventory. Wu et al. (2005) study the relationship between firm inventory holdings and financial performance, while Hendricks and Singhal (2005) study the impact of supply chain disruptions (including problems with inventory) on short term financial and accounting measures. Gaur et al. (2005a) find that as a retailer's margins decreases and capital intensity increases, it tends to carry less inventory (as measured by inventory turns). Roumiantsev and Netessine (2006) use aggregate inventory data to measure the relationship between demand uncertainty, lead times, gross margins and firm size on inventory levels. Rajagopalan (2005) estimates the effect of product variety on inventory levels of publicly listed US retailers. These studies use data on publicly traded firms and they do not measure the effect of market structure on inventory. Amihud and Mendelson (1989) use public data on manufacturing firms to estimate the effect of market power (proxied by the firms' margins and

[1] Holweg and Pil (2004) report that 89% of the sales in the U.S. are filled from dealership stock, versus 38% in Europe.

market shares) on inventory levels and variability. They find that firms lower their inventory as market power decreases, i.e., as competition intensifies. Our study is different because we track individual units of inventory and we are able to measure differences in local market structures.

Other empirical studies analyze the effect of market structure on product variety. For example, Berry and Waldfogel (2001) find that higher market concentration leads to higher product variety in the radio broadcasting industry, showing evidence that firms use product line expansion to preempt entry. Alexander (1997) finds a non-monotonic effect of market concentration on product variety in the music recording industry, where the highest level of variety is achieved at a moderately concentrated structure. Inventory is not a primary concern with either of those industries. Watson (2004) also finds a non-monotonic effect of competition on product variety among eyeglass retailers.[2]

In the next section, we provide a general econometric framework to measure the effect of sales and competition on inventory. Section 3 describes the data and the specification of the model. Section 4 shows our main results and section 5 provides a sensitivity analysis and further empirical evidence. Section 6 analyzes a counterfactual experiment of changing the dealership network. We conclude and discuss our findings in section 7.

## 2. An empirical model of retail inventory

We use a basic single-item periodic review base-stock model to motivate our empirical framework. Orders are received at the beginning of each period with zero lead-time. Let $D$ be i.i.d. normal demand in each period with mean $\mu$ and standard deviation $\sigma$. Some fraction of the demand that is not fulfilled from in-stock inventory is backordered; the remaining demand is lost. Let $Q$ be the order-upto level and $z = (Q - \mu)/\sigma$. In this model the service level is the probability that all demand within a period is satisfied from inventory. The service level is increasing in $z$, so for convenience we refer to $z$ as the service level with the understanding that it is really a proxy for the service level. The expected inventory at the end of each period, $I$, is then

$$I = \sigma \left( z + L\left(z\right)\right) \tag{1}$$

where $L\left(z\right)$ is the standard normal loss function.

[2] He finds that stores with two or three nearby rivals offer the highest level of variety, but product variety falls as competition increases beyond that threshold. In auto dealerships, inventory levels can be viewed as a measure of variety because it is rare to have two identical vehicles in stock. Watson (2004) does not observe sales, so unlike in our model, he is unable to distinguish whether competition influences his variable of interest (variety/inventory) via a sales effect or a service level effect as discussed in Section 2.

It is empirically inconvenient to work with (1) directly because demand is not directly observable. However, it can be shown (see the online appendix for details) that (1) can be written as

$$I = \sigma_s K(z) \tag{2}$$

where $\sigma_s$ is the standard deviation of *sales* ($\min\{Q, D\}$) and $K(z)$ is an increasing function. As in van Ryzin and Mahajan (1999), we use

$$\sigma_s = A \cdot S^{\beta_s} \tag{3}$$

to approximate the standard deviation of sales, where $S$ is observed sales over a sample period and $A$ and $\beta_s$ are coefficients. The $\beta_s$ coefficient reflects the degree to which there are economies of scale in inventory management with respect to sales. If $\beta_s = 1$, then days-of-supply (inventory divided by daily demand rate) is independent of expected sales whereas if $\beta_s < 1$, then higher sales retailers carry a lower days-of-supply for the same service level.[3] Combining (2) and (3) and taking logarithms yields.

$$\log I = \text{constant} + \beta_s \log S + \log K(z) \tag{4}$$

The above equation suggests that a firm's inventory level can be decomposed into two separate components: a sales component, $\beta_s \log S$, and a service level component, $\log K(z)$.

According to (4), market structure can influence a firm's inventory either through its sales or through its service level. Suppose the number of firms competing in a market is taken as the proxy for market structure. If a market's potential sales is reasonably fixed, then it is intuitive that entry could reduce each firm's sales (the fixed market potential is allocated among more firms). However, entry could increase a retailer's sales either because price competition is sufficiently severe to increase total sales (i.e., total potential demand increases) or via a retail agglomeration effect - consumers may be more likely to search a retailer located near other retailers rather than an isolated retailer because the consumer wishes to economize on search costs.[4] We are not directly concerned with the specific mechanism by which market structure influences sales because we conjecture that these mechanisms influence inventory only through their effect on sales.

We conjecture that there are three mechanisms by which market structure influences the service level component of (4). Two of these are related to the cost of holding too much inventory (the

[3] Gaur et al. (2005b) measures $\beta$ using public data from the U.S. retail sector, obtaining estimates from .55 to .73.

[4] See Dudey (1990), Eaton and Lipsey (1982), Stahl (1982) and Wolinsky (1983) for models of consumer search in which firm location decisions are endogenous.

overage cost) and the cost of holding too little (the underage cost). Taking demand as exogenous, a retailer sets a service level, $z$, to balance these costs optimally. The overage cost is primarily composed of the opportunity cost of capital, storage costs and depreciation. The underage cost depends on the behavior of consumers when they do not find their preferred product either because it is not carried by the retailer or because it is temporarily out of stock. In such a situation a consumer could purchase some other product at the retailer (substitute), defer purchase of the most preferred product to a later time (backorder) or leave the retailer without making a purchase (the no-purchase option). A retailer's underage cost is increasing in the retailer's margin - the larger the margin on each sale, the more costly it is to lose a sale. Furthermore, the underage cost is decreasing in consumers' propensity to substitute or backorder but increasing in the consumers' propensity to choose the no-purchase option. These behaviors probably depend on numerous consumer characteristics (i.e., local market demographics), such as the intensity of their preference for the products in the retailer's assortment, their perception of the cost to search/shop, and their ability and willingness to defer their purchase. Furthermore, we assume these demographics, as well as the overage costs, are not affected by market structure. In contrast, we conjecture that market structure influences underage costs through a *margin mechanism* and/or a *demand-retention mechanism*. The margin mechanism is simply that additional competitors increases the intensity of price competition, which lower margins, thereby decreasing the underage cost. The demand-retention mechanism influences underage costs via consumer behavior. As more competitors enter a market, consumers are more likely to choose the "no-purchase" option relative to the "substitute" or "backorder" option, thereby leading to higher underage costs. Therefore, the margin and demand-retention mechanisms counteract each other. Finally, dropping the assumption of exogenous demand, the *demand-attraction mechanism* is the third mechanism by which market structure influences the service level: a higher service level may attract more demand to a retailer (because, all else being equal, a consumer prefers to shop at a retailer with a higher service level) and more competition causes firms to increase their service level in an effort to attract more demand. (See Dana and Petruzzi (2001) and Gerchak and Wang (1994) for single-firm models in which service level is used to attract demand.)

There is theoretical support for these three mechanisms that link market structure to service level. Deneckere and Peck (1995) consider a model in which consumers can observe both the prices and quantities of $n$ firms before choosing from which firm to purchase. As a result, both the margin and demand-attraction mechanisms are active. If an equilibrium exists, they find that equilibrium prices are decreasing in $n$, but service levels are nevertheless independent of $n$, which suggests the two mechanisms offset each other. Dana (2001) modifies the Deneckere and Peck (1995) model and indeed finds that entry can reduce service levels - the effect of price competition on the firms' underage costs can dominate the demand-attraction effect. The analogous conclusion can be inferred from Bernstein and Federgruen (2005).[5] Other models obtain the same result, but the causality is reversed: competition induces firms to reduce their service level, even if a 100% service level is costless, because lower service levels dampen price competition (see Balachander and Farquhar (1994) and Daugherty and Reinganum (1991)). Consistent with the hypothesis that competition leads to lower service levels, Gaur et al. (2005a) find that retailers with lower margins carry lower inventory and Amihud and Mendelson (1989) provide evidence of a direct link between market power and inventory levels.[6] However, Cachon (2003) develops the opposite hypothesis. He considers the special case of the Deneckere and Peck (1995) model with fixed prices. As a result, entry has no impact on margins. However, the demand-attraction mechanism remains active and he finds that service levels are *increasing* in $n$: firms use service level more aggressively to attract demand when they face more competition.

There is no demand-retention effect in the Deneckere and Peck (1995) model (and its derivatives), because, in part, demand is assumed to be lost when a retailer stocks out rather than "spilling over" to another retailer. Numerous models do study retail competition with spillover demand (e.g., Lippman and McCardle (1997), Mahajan and van Ryzin (2001), Netessine and Rudi (2003)) but those models neither have a demand-attractive effect (the demand allocated to a retailer does not depend on his inventory) nor a margin-effect (price is assumed to be fixed), nor a demand-retention effect (firms do not influence whether consumers choose to purchase or continue shopping). As a result, market structure and service level are independent of each other in those models. Cachon et al. (2006) and Watson (2006) do develop formal models with a demand-retention effect. If service level is interpreted as the probability a firm carries a consumer's most preferred product, then they show that firms increase their service level as they face more competition because a higher

[5] If prices decrease, Bernstein and Federgruen (2005) find that service levels decrease, but they do not explicitly study the impact of market structure, and their model is ill suited to do so. To explain, their model of retailer $i$'s demand is $D_i(p) = d_i(p)\varepsilon_i$, where $p$ is the vector of retail prices, $d_i(p)$ is a deterministic demand function and $\varepsilon_i$ is a stochastic shock. It is not clear how to modify $\varepsilon_i$ to account for firm entry. It is possible to make assumptions regarding the impact of entry on $\varepsilon_i$ but they do not do so.

[6] Gaur et al. (2005a) do not directly link retail competition to inventory level - they only observe a correlation between margins and inventory turnover. Amihud and Mendelson (1989) also explore the variability of inventory holdings.

service level reduces the chance a consumer continues searching/shopping.

To summarize, theoretical models of inventory competition predict service levels are either decreasing, independent or increasing with respect to entry. Additional competition reduces service levels if the impact of price competition on margins is severe, whereas additional competition increases service levels if higher service levels either attract additional demand or help to retain demand.

Given this discussion, we now further elaborate on (4). For each retailer $r$ and product category $b$, based on the periodic review model, inventory, $I_{rb,}$ is determined by a combination of sales, $S_{rb}$, and the service level, $z_{rb}$. (We now distinguish products by category because it is plausible that inventory levels across categories at the same retailer have different motivations to hold inventory.) We use the index $i$ to denote each $(r, b)$ combination and $m(i)$ to denote the relevant market for observation $i$. Market structure can influence sales and service level, but there are other factors describing a market that could influence service level (e.g., consumer characteristics). Let $W_{m(i)}$ be a (column) vector of observable covariates capturing the characteristics of the local market that affect the service level of observation $i$.[7] However, different observations from the same local market can have different service levels; that is, there may be factors specific to a retailer or product category that affect its service level. The vector $V_i$ captures observable factors of this kind. For example, $V_i$ may include factors describing the supply process of a retailer or a vector of brand dummies.

We assume the following reduced form for the effect of service level:

$$\log K(z_i) = \beta_v V_i + \gamma W_{m(i)} + \xi_{m(i)} + \nu_i \tag{5}$$

The error term $\xi_{m(i)}$ captures unobserved factors relevant to local market $m(i)$; $\nu_i$ denotes other unobserved factors specific to observation $i$. The term $\gamma W_{m(i)} + \xi_{m(i)}$ is the effect of local market conditions on service level. A subset of the covariates in $W$, referred to as $C_{m(i)}$, capture the intensity of competition in market $m(i)$, what we refer to as market structure, such as the number of rival stores in the local market. The term $\gamma_c C_{m(i)}$ measures the *overall* impact of competition on service level, including price competition and inventory competition effects (e.g. demand attraction/retention effects); therefore, its sign is ambiguous. Other covariates in $W$ include an intercept and demographic characteristics of the markets that capture differences in consumer characteristics

[7] Throughout the paper, we use column vectors for covariates and row vectors for parameters.

which influence a retailer's optimal service level. For example, if a certain demographic of consumers (e.g., income) affects their propensity to purchase new vehicles (so that they are unwilling to purchase used vehicles), then new car dealerships in markets with a high concentration of these types of consumers may have a lower optimal service level.

Replacing (5) in (4) gives the following model, which we seek to estimate using data from a cross section of retailers:

$$y_i = \beta X_i + \gamma W_{m(i)} + \xi_{m(i)} + \nu_i. \tag{6}$$

where $y_i = \log I_i$, $X_i = (\log S_i, V_i)$ and $\beta = (\beta_s, \beta_v)$. The parameters to be estimated are $\theta = (\beta, \gamma)$. We are interested in the magnitude of the coefficient of sales $\beta_s$ ($\beta_s = 1$ means there are no economies of scale with respect to sales) and the sign and magnitude of the competition effect ($\gamma_c C_{m(i)} < 0$ suggests that the price effect of competition dominates whereas $\gamma_c C_{m(i)} > 0$ suggests that the demand attraction/retention effects dominate).

**Estimation method**

There are several challenges associated with the identification of $\theta$. It is important to define each retailer's market appropriately, otherwise $W$ may be a poor measure for local market characteristics. We attempt to alleviate this concern by identifying geographically isolated markets (a similar approach was used by Bresnahan and Reiss (1991)). To estimate $\gamma_c$ precisely, it is important that the selected markets have sufficient variation in market structure.

The endogeneity of some of the variables in $X$ and $W$ is of particular concern with respect to the identification of $\theta$. Sales is affected by product popularity, which may also affect customer purchase behavior (e.g., the propensity to backorder) and therefore the service level chosen by retailers. While the demographic variables in $W$ capture part of the heterogeneity in consumer characteristics across markets, some customer characteristics are unobservable and will enter in $\xi$. Following on our previous example, if a market has consumers with a high affinity to purchase new vehicles and these consumer tastes are not fully captured by the covariates in $W$, then we would expect $\xi$ to be correlated with sales. Hence, estimating (6) with Ordinary Least Squares (OLS) leads to biased estimates of $\theta$.

Measures of market structure are subject to a similar endogeneity bias. Retailers choose which markets to enter and they may do so based on market characteristics that they observe but are unobserved by the econometrician. Inventory costs affect dealership profits, therefore entry decisions are affected by local market characteristics that influence inventory, including $\xi$. If such is the case,

$C$ and $\xi$ may be correlated. Intuition suggests this correlation is negative: high service levels (high $\xi$) raise total inventory costs, leading to lower profits and fewer entrants (low $C$). This suggests a downward bias in estimating $\gamma_c C$ through OLS.

We use a two step method to estimate $\theta$. In the first step, we use a within-market estimator of $\beta$ which accounts for the endogeneity of sales. In the second step, we replace $\beta$ in (6) with this estimate and estimate the modified (6) using Instrumental Variables to account for the endogeneity of market structure. We describe in detail this two step method in what follows.

In the first step, we seek to estimate $\beta$ by comparing dealers located in the same local market. Define the set $M_m = \{i : m(i) = m\}$ which contains all observations from market $m$. Also, let $\bar{X}_m = \frac{1}{|M_m|}\sum_{i\in M_m} X_i$ and $\bar{y}_m = \frac{1}{|M_m|}\sum_{i\in M_m} y_i$. We use a transformation of the dependent variable $\dot{y}_i = y_i - \bar{y}_{m(i)}$ and the covariates $\dot{X}_i = X_i - \bar{X}_{m(i)}$. to re-write (6) as

$$\dot{y}_i = \beta\dot{X}_i + \nu_i \tag{7}$$

Assuming $E\left(\dot{X}_i\nu_i\right) = 0$, estimating (7) using OLS gives a consistent estimate of $\beta$. The main advantage of this model with respect to (6) is that it allows consistent estimation of $\beta$ even when some of the covariates in $X$ (e.g. sales) are correlated with $\xi$. Its main disadvantage is that the effect of local market conditions, $\gamma W_{m(i)} + \xi_{m(i)}$, are not estimated.

The second step estimates $\gamma$ using the estimated coefficient $\hat{\beta}$. Replacing $\beta$ in (6) with $\hat{\beta}$ and rearranging gives

$$y_i - \hat{\beta}X_i = \gamma W_{m(i)} + \varepsilon_i, \tag{8}$$

where $\varepsilon_i = \xi_{m(i)} + \nu_i$. We estimate (8) using Instrumental Variables (IV) to instrument for the endogeneity of market structure. We seek factors excluded from $W_{m(i)}$ that are correlated with market structure but uncorrelated with unobservable consumer characteristics that enter in $\xi_{m(i)}$. We use measures of market population as our main instruments on the assumption that population is correlated with entry (more firms enter as a market's population increases) and population is uncorrelated with unobserved consumer characteristics that influence service level conditional on the observed controls in $W_{m(i)}$.[8] The exogenous instruments, denoted by $Z$, include several measures of population and the demographic in $W$. $Z$ *does not* include covariates in $X$ or the measures of market structure $C$. Assuming $E(Z_i\varepsilon_i) = 0$, estimating (8) using Two Stage Least Squares (2SLS) gives a consistent estimator of $\gamma$.

[8] The assumption that larger markets lead to more entry can be verified empirically when markets are well defined. See Bresnahan and Reiss (1990) for empirical evidence of the effect of population on entry in auto dealership markets.

Our two-step method estimates $\theta$ consistently based on two moment conditions: $E(\dot{X}_i\nu_i) = 0$ and $E(Z_i\varepsilon_i) = 0$. Instead of using a two-step method, we also estimate these moment conditions jointly using Generalized Method of Moments (GMM). (See the online appendix for details on this estimation procedure.) There are two main advantages of the GMM approach. First, it is more efficient (the estimation is more precise). (See Wooldridge (2002), Section 8.3 for details on the statistical properties of GMM.) Second, the standard errors provided by the 2SLS in the second step of our two-step method are not correct because the regression includes variables which are estimated ($\hat{\beta}X_i$). The standard errors from the joint estimation using GMM are correct. The main drawback from using GMM is that $\beta$ is biased when the second moment condition $E(Z_i\varepsilon_i) = 0$ is misspecified (i.e., when some of the covariates in $Z$ are not exogenous). In addition, common statistics used to evaluate the goodness of fit in regressions (e.g. $R^2$) are not available for GMM. We found that the point estimates from GMM were similar to those obtained through the two-step method, and the statistical significance was also similar. However, the standard errors of GMM are correct and we use them to validate our hypothesis testing.

## 3. Data

This section provides a brief description of the U.S. auto industry and details the data in our study. Six companies account for about 90% of sales in the U.S. auto market: Daimler-Chrysler (DC), Ford, GM, Honda, Nissan and Toyota. We refer to DC, Ford and GM as domestic manufacturers. Each company offers vehicles under several brands or auto makes. For example, GM makes include Chevrolet, GMC, Pontiac, Buick, Saturn, Cadillac and Hummer. Each auto make produces several models. Examples include the Chevrolet Malibu, the Toyota Camry and the Ford Explorer. Models can be classified into vehicle classes, including cars, sports cars, Sport Utility Vehicles (SUV) and pickups, among others. Each model is offered with multiple options, which include different body styles, engines, transmission types and breaking systems, among other features.

In the U.S., auto distribution is regulated by franchise laws, which require that all new vehicles must be sold through a network of dedicated franchised dealers. (See Smith (1982) for details on dealership franchise laws.) As of 2006, there are approximately 22,000 dealerships in the U.S. The number of dealerships has been declining in the U.S. since it peak in 1930 when there were about 50,000 dealerships (Marx (1985)).

### 3.1 Definition of Markets

Based on (6), we seek to define isolated markets so that we can accurately proxy for the level of competition within the market. We begin with Urban Areas (UA) defined in the 2000 Census and with population below 150,000.[9]. We designate an UA as isolated if it meets the criteria summarized in Table 2. These criteria impose minimum distance requirements to markets of equal or larger size with the rationale that consumers who do not find their desired product inside their market will try to find that product in the closest more populous market. Dranove et al. (1992) and Bresnahan and Reiss (1991) use similar criteria to define isolated markets. From this set of markets, we selected for our study the 235 markets that have at least one GM dealership. (As we describe later, our data are from GM dealerships.) We obtained demographic data and geo-coded information (latitude and longitude) for these markets from the 2000 decennial census. 37%, 5%, 26% and 31% of the markets are located in the Mid-West, North-East, South and West census regions, respectively. (See the online appendix for a map indicating their locations.) Data on new vehicle dealerships located in each market were obtained from edmunds.com.[10]

Table 3 describes the selected markets, grouped according to the total number of Ford, DC, GM, Honda, Toyota and Nissan dealerships. The second column shows the number of markets with the observed number of dealers. For example, there are 17 monopoly markets with one GM dealership. In more than 90% of the markets there are 10 or fewer dealerships. The number of dealerships increases with market size, measured by population (third column). The last three columns show the percent of markets with at least one dealership of a non-GM domestic manufacturer (Ford or DC), Japanese manufacturer (Toyota, Honda or Nissan) and a second GM dealership, respectively. The first competitor faced by a GM dealer is usually a non-GM domestic dealership. Japanese dealerships usually enter markets with three or more dealerships. All of the observed markets with six or more dealerships have at least one GM, one non-GM domestic and one Japanese dealership. Several markets have more than one GM dealership. In all but 5 markets, the GM dealers carry franchises of different brands. The table shows that the selected markets have sufficiently rich variation in market structure, both in the number and type of dealerships.

[9] These include: (i) urbanized areas "consisting of territory with a general population density of at least 1,000 people per square mile of land area that together have a minimum residential population of at least 50,000 people"; and (ii) urban clusters of "densely settled territory with at least 2,500 people but fewer than 50,000". (quoted from Census glossary, www.census.gov)

[10] We matched dealers to UA based on 5 digit zipcodes. Matching tables were obtained from the Missouri Census Data Center (http://mcdc2.missouri.edu).

We obtained the following demographic data for each market: percent of population above 60 years old *(ELDER)*, that is African-American (*BLACK*), with a college degree (*COLLEGE*), active in the army (*ARMY*), involved in a farming occupation (*FARMING*) and that commutes to work with public transportation (*PUBTRANS*). We also obtained median household income for each UA (*INCOME*). Summary statistics of these variable are shown in Table 4.

We included *BLACK*, *INCOME*, *COLLEGE, ELDER,* and *FARMING* in $W$ because these variables have substantial partial correlation with the number of dealerships in a market (see online appendix Table 10). In addition, we included *PUBTRANS* and *ARMY* (to capture potential differences in consumer characteristics and their affinity for domestic brands) and indicators of the census region where the UA is located.[11]

### 3.2 Model specification

We obtained inventory and sales data from a website offered by GM (http://www.gmbuypower.com) that enables customers to search new vehicle inventory at local dealerships. We developed a web-crawler that each day monitored inventory in all the GM dealerships located in our selected markets (and only GM dealerships[12]) from August 15, 2006 to February 15, 2007 (six months of data). The web-crawler recorded the number and type of vehicles available at each dealership (e.g., the number of GMC Yukon 2007 4WD available at each dealer) along with specific information on each vehicle, such as color, options, list price and, most importantly, the vehicle identification number (VIN). VINs uniquely identify all new vehicles in the U.S. Therefore, by keeping track of the VINs available at each dealership, we are able to identify replenishments (a vehicle is added to a dealer's inventory) and sales (a vehicle is removed from a dealer's inventory). We also can identify dealer transfers (a vehicle removed from one dealer's inventory and added to another dealer's inventory) among the dealers in our sample. However, to identify all dealer transfers would require monitoring all dealers in the U.S., which was not feasible. Instead, we monitored all dealerships in seven states, which we believe allows us to identify most of the transfers occurring in our sample markets in

[11] We also estimated specifications which included other demographics, including voter turnout, the percent of Republican votes, the percent Latino in the population, and the average number of vehicles per household, among others. The results in these specifications were similar to those reported in Section 4. Some of these additional variables were not available for all markets, so we decided to exclude them from our main results.

[12] Developing web-crawlers for each manufacturer would require substantial additional effort. Some websites offer inventory search for dealerships of several brands (e.g. nada.org) but many of these are not suitable for the large-scale data collection that we require. We monitored our web-crawler frequently in case changes were made to the website. In fact, during our study period GM did change its website. Substantial effort was required to repair the crawler.

those states.[13]

To validate our data, we visited three dealerships in the Philadelphia area[14]. Most of the vehicles found at these dealers on June 2, 2006 were posted on the website during that day[15]. The dealerships visited declined to provide data on the specific vehicles sold.

To estimate model (6), we defined the dependent variable as the average vehicle inventory of each brand at a GM dealership ($INV$). (HUMMER is excluded from our analysis because it is present in only one of our study markets.) We imputed total sales ($SALES$) of each make during the study period to measure expected sales. (Sales includes vehicles transferred to other dealerships.)

We estimated several specifications for the competition effect, $\gamma_c C$. The simplest measure is the number of dealerships in the market ($NC$). We restricted the dealership counts to the following manufacturers: GM, DC, Ford, Toyota, Honda and Nissan. We included the square of this variable ($NCSQ$) to capture non-linearities in the effect of competition. We also estimated the effect of the number of rivals using a flexible non-parametric specification, with indicator variables of the form $\delta_x = 1\{NC = x\}$, with $x \in \{1..N_{\max}\}$. We restricted our sample to markets with 8 or fewer dealerships to measure this effect more precisely ($N_{\max} = 8$).[16] In some specifications, we also include the number of GM dealerships in the market ($NGM$) to test whether the effect of competition varies across different types of dealerships.

To measure potential competition from outside the market, we included the driving time (from http://www.randmcnally.com) to the closest GM dealership outside the UA ($OUTSIDE$) as a covariate in $W$. Driving time was used to capture the effect of nearby highways on transportation costs. We also estimated models with "bird-fly" distance (using latitude and longitude data) and to GM dealerships carrying the same brand. Our results were similar with these alternative measures.

GM dealerships can own multiple franchises of GM brands. If customers substitute between different GM brands, a stock-out in one brand is less likely to become a lost sale for a multi-brand dealership, because customers may buy a vehicle from another brand on the lot. If inter-brand substitution within GM is substantial, we expect the number of franchises carried by a dealership ($NFRANCH$) to have a negative effect on the service level. This dealer specific measure is included

[13] The selected states are Colorado, Nebraska, Florida, Wisconsin, Maine, California and Texas. These states are geographically relatively isolated (they border Mexico or Canada, they have a substantial coastline and/or their border areas are sparsely populated) and exhibit variation in population growth (see Figure 1).

[14] We selected this dealerships by convenience. None of the selected markets are in the Philadelphia area.

[15] The dealership lots include many vehicles (sometimes more than 100) and the authors could not verify all of them.

[16] We also expanded our sample including markets with 9 and 10 dealerships and our results were similar.

as a covariate in $X$. Make dummies were also included in $X$ to control for differences in customer loyalty and preferences that can influence service level.

The effective service level may also be affected by a dealership's supply process. For example, transfers between dealerships enable dealerships to share inventory, which helps to reduce inventory.[17] Therefore, we include a measure of transfers as a control variable in $X$. Let $T_{rb}$ be the total amount of transfers received of product category $b$ by dealership $r$ and let $Q_{rb}$ be the total incoming orders (without transfers from other dealerships) received. For observation $i = (r, b)$, we measure the percent of transfers received as:

$$TRANSF_i = \frac{T_i}{T_i + Q_i}$$

We expect $TRANSF$ to have negative effect on average inventory levels. Recall, we are unlikely to observe all of the transfers for all dealerships. We include a dummy, $ALLSTATE$, to indicate whether the dealership is located in one of the states where we monitored all dealerships.[18]

The structural model underlying equation (6) suggest that coefficient $\beta_s$ captures statistical economies of scale associated with sales volume. In auto dealerships, there are additional economies of scale in sales volume. For example, there can be economies of scale arising from fixed ordering costs (such as order processing and transportation). These other sources of economies of scale will be captured in the estimated $\beta_s$ coefficient. However, visual inspection of time-series inventory level data does not reveal a strong "saw-tooth" pattern, suggesting that batching is not a main factor determining inventory levels.[19]

Table 5 shows summary statistics and the correlation matrix of the main variables in the econometric model.

[17] Anupindi and Bassok (1992) show that centralization of inventory stocks of multiple retailers usually decreases total inventory relative to the descentralized case where each retailer chooses their inventory level independently. Rudi et al. (2001) analyze a model of two newsvendors with transshipments of left-over inventory. It can be shown that their model implies a negative association between the average number of transfers received by a retailer and its service level. Narus and Anderson (1996) report inventory reductions from inventory sharing initiatives in several industries operating with descentralized distribution networks.

[18] If the coefficient on $TRANSF$ is negative, we expect $ALLSTATE$ to be positive because for the observations with $ALLSTATE=0$ a fraction of the transfers are unobserved. In all the specifications analyzed, the coefficient on $ALLSTATE$ was positive and significant. The average percent of transfers for dealerships with $ALLSTATE=0$ and $ALLSTATE=1$ is 4.5% and 10%, respectively. $ALLSTATE$ is market specific and is therefore included in $W$.

[19] We also included measures that capture heterogeneity in batch sizes across dealerships, such as the coefficient of variation of weekly incoming orders. The results including these measures were similar. We noted that the measures of batching are sensitive to the unit of time aggregation (e.g. weekly, bi-weekly) and therefore decided to exclude these measures from our main results.

### 3.3 Instrumental variables

We use total population in the UA *(UAPOP)* and fringe population (*FRINGEPOP*) to instrument for market structure. The fringe population of a UA is defined as the population of all zipcodes outside the UA within a 100 miles radius for which the UA is the closest UA with dealerships[20]. We also used measures of past population as instruments: county population in 1950 and 1970 (*POP50, POP70*). Franchising laws impose costs on the manufacturer to close existing dealerships. Markets with current low population which had higher population in the past are likely to have more dealerships than those which never had a large population. Due to this "stickiness" in dealership exit, past population has positive partial correlation (conditional on current population) with the number of dealerships. All population measures were included with natural log transformation because it provided better fit in the first stage estimates of the 2SLS regressions. Motivated by Figure 1, we defined two additional instruments that depend on county population growth between 1950 and 2000 (denoted $g$): $PGWTH{=}\max(0, g)$ and $NGWTH{=}\max(0, -g)$.[21] UA population was obtained from the 2000 decennial census. Historical county population was obtained from the Inter-University Consortium for Political and Social Sciences (ICPSR).

## 4. Results

Table 6 displays the estimation results. Columns 1 shows the estimates of the first step of our two step method. Columns (2)-(5) show different specifications for the second step of the method. Columns (6) and (7) show the joint GMM estimates. The coefficients for the demographics and the dummies for make, region and *ALLSTATE* are omitted for ease of visualization. The complete results for some of the specifications are displayed in the online appendix, Table 9. In this section, we discuss the results reported in Table 6.

Column (1) shows that the point estimates of the coefficient of *logSALES* ($\beta_s$) is measured with precision and is below one with statistical significance. The magnitude of the $\beta$ coefficient suggest substantial economies of scale: a 10% increase in sales translates into a 3.6% decrease in days-of-supply ( $(1-\beta_s)\cdot 10\%$ ). The use of transfers from other dealerships, measured by *TRANSF*, has a large economic (and statistically significant) effect in reducing inventory levels. Increasing *TRANSF*

[20] A similar measure was used by Dranove et al. (1992). We calculated distances using latitude and longitude. The census proxy of zipcodes (Zip Code Tabulation Area, ZCTA) were used.

[21] Bresnahan and Reiss (1990) uses similar functions of population growth to capture entry in auto dealership markets.

by 0.1 (a 10% increase in the fraction of supply received from transfers) reduces inventory by 7.4%. The coefficient on *NFRANCH* is small and not significant. The coefficient of determination ($R^2$) is high, suggesting that a substantial fraction of the within-market variation on inventory can be explained by the covariates included in $X$.

Column (2) shows the estimates of the service level effect of competition ($\gamma_c$) using OLS. The specification includes a linear and a quadratic term of the number of dealerships in the market (*NC* and *NCSQ*). The estimates suggest that the effect of competition is positive and marginally decreasing. Figure 2 illustrates the estimated impact of the number of dealerships on inventory, measured by the percent change relative to a monopolist. The figure shows the effect of competition through *service level only* (sales is kept constant). Upper and lower bounds of the 95% confidence interval are illustrated with + and - symbols, respectively (standard errors are calculated using the delta method, see Hayashi (2000)). The squares in the figure plot the estimates from a flexible non-parametric specification. Interestingly, the more parsimonious quadratic polynomial model approximates very well the non-parametric model. In all the specifications analyzed, the coefficient of *OUTSIDE* is small and not significant, suggesting that our market definition capture well the local competition faced by a dealership.

Column (3) estimates equation (8) using IVs to instrument for the endogenous variables *NC* and *NCSQ*. IVs include *UAPOP, FRINGEPOP, PGWTH, NGWTH, POP50* and *POP70*.[22] Even though the estimates are less precise than in (2), they suggest a similar pattern for the competition effect. In fact, the competition effect suggested by the IV estimates is larger: the elasticity at the mean of *NC* is 0.54 versus 0.35 for specification (2).[23] The $R^2$ of the first stage of 2SLS is 0.69. (The first stage estimates are displayed in online appendix, Table 10).

Columns (4) and (5) includes the number of GM dealerships (*NGM*) as an additional measure of competition. Column (4) reports the OLS estimates and column (5) the IV estimates. Both specifications suggest that the effect of entry of a rival GM dealership has a larger positive effect compared to the effect of an average dealer. As before, the implied elasticity at the mean is larger for the IV estimates.

Columns (6) and (7) report the joint GMM estimates. The instruments used in these estimations include exogenous variables in $W$ and *UAPOP, FRINGEPOP, PGWTH, NGWTH, POP50* and

[22] The number of observations of the specifications using IVs is smaller because we could not obtain past population for all markets.

[23] A Hausman test rejects the estimates of columns (2) and (3) are equal (p-value <0.01).

*POP70*. Hence, the estimates of columns (6) and (7) are comparable to those of columns (3) and (5). The point estimates and statistical significance of the estimated coefficients obtained through GMM and the two step method are similar. Because the asymptotic standard errors of the GMM estimates are correct, this validates the statistical significance of our results.[24]

## 5. Sensitivity analysis and further empirical evidence

In this section, we report on a sensitivity analysis and provide additional empirical evidence to test the robustness of our results.

Model (6) suggests a linear relationship between the logarithms of inventory and sales, and requires a constant $\beta_s$ across markets with different market structures. A scatter plot (available in the online appendix) of *logINV* versus *logSALES* reveals a strong linear relationship between the two variables in three types of markets: GM monopoly markets, markets with GM and non-GM domestic dealers, and markets with all kinds of dealerships (GM, non-GM domestic and Japanese). A regression of *logINV* on *logSALES* allowing for different slopes and intercepts across the three groups yields $R^2 = 0.95$ and fails to reject the hypothesis of equal slopes across the three series ($p = 0.36$)[25]. This analysis suggests there are no interaction effects between *logSALES* and market structure, i.e., the effect of competition on service level is separable from the effect of sales.

Regressions over the sub-sample of dealerships with *ALLSTATE=1* yield estimates that are similar in magnitude, sign and statistical significance to those reported in Table 6.

Model 6 can be subject to measurement error bias if average inventory and sales are estimated from a short time series. To explain, suppose only one week of daily observations are available to evaluate *INV* (*average* inventory level) and *SALES* (a dealer's *expected* sales). If sales during that week were below average, then *INV* overestimates average inventory and *SALES* underestimates expected sales. The measurement errors of *INV* and *SALES* are then negatively correlated, and so the coefficient on sales, $\beta_s$, is likely to be downward biased. To assess the magnitude of the bias, we replicated our analysis using three months of data. The results were basically identical to our main results (data from a six month period), which suggests that this potential measurement error bias is small in our analysis.

We use market population as an IV to identify a causal effect of competition on service level (section 4). The main concern with OLS is that the positive correlation between competition and

[24] We also estimated specifications (2) and (4) through GMM and the estimates were also similar.

[25] In pairwise tests of the coefficients the smallest $p$ value was 0.18.

service level could be driven by unobserved factors that affect both variables rather than a causal effect of competition. For example, non-GM dealerships may have a stronger incentive to enter markets where customer loyalty to GM brands is lower. The number of dealerships in a market becomes a proxy of consumer's lack of loyalty for GM, which could have a positive association with the service chosen by GM dealers.[26] Given the demographic controls included in $W$, we believe it is unlikely that unobserved consumer characteristics that affect service level are correlated with market population. Hence, the IV estimates should be consistent. Nevertheless, we provide additional results following a different identification strategy which corroborate our findings.

In equation (6) we use the number of dealerships in a market as a measure of competition. We argue, due to the demand attraction and retention effects, that dealerships raise their service level when they face more intense competition to prevent losing customers to rival stores. If so, then the effect of entry on service level should depend not only on the number of dealerships in a market but also on the number and type of *products* they offer. An entrant that offers more models which are close substitutes to the products offered by the incumbents should trigger a larger increase in the service level. In fact, Cachon and Olivares (2006) show that the aggregate days of supply of a model tends to increase with the number of models offered in the same segment.

To validate our conjecture, we estimate equation (6) using the number of *models* offered by rival dealerships as a measure of competition. Following the literature of spatial competition (e.g. Seim (2006)), we define different *bands* where the products offered by rival dealerships can be located. These bands define a measure of "distance" between product $b$ and products offered by rivals. The definition of the bands is based on a market segmentation commonly used in the auto industry. While these definitions can be subjective, we feel they work reasonably well to capture the degree of similarity across products in this industry. We conjecture that the number of products in closer bands should have a higher impact on the service level than products located in the outer bands. On the contrary, if the association between service level and market structure is driven by unobserved customer loyalty for GM, then products in all bands should have a similar positive association.[27]

[26] For example, suppose there exists some consumer characteristic describing loyalty for GM brands which is observed by firms and unobserved by the econometrician. This characteristic must be particular to a subset of markets because we control (via brand dummies) for the overall preference for GM brands. Based on this characteristic, Ford dealers are attracted to markets where loyalty for GM is low. GM dealerships raise their service level in these markets because of the presence of this characteristic, not *per se* because of the presence of the Ford dealership.

[27] Online appendix Table 11 shows that GM's assortment is similar to the variety offered by the industry. Hence, preferences for GM brands and vehicles segments are likely to be independent, (i.e., a preference for the GM brand is not merely a proxy for a preference for a particular vehicle segment) and the number of models offered by rivals in any band should be a good proxy for the lack of customer loyalty to GM. If we do not see the same effect on different

Let $\Omega$ be the set of all models offered in model-year 2007. For a given product $b$, we define a partition $\left\{\Omega_b^1 ... \Omega_b^K\right\}$ of the set of products $\Omega$ and refer to $\Omega_b^k$ as the $k^{th}$ band of product $b$. Bands are defined so that their distance to product $b$ is increasing in $k$. Let $C_{rb}^k$ be the number of models in band $\Omega_b^k$ offered by the rivals of dealership $r$. The number of models offered is calculated based on the brands carried by each rival dealership and the list of models offered by each brand[28]. We included dealerships of all manufacturers (not just the six included in the previous estimation). Define the column vector $C_{rb} = \left(C_{rb}^1 ... C_{rb}^K\right)^t$ and the row vector of parameters $\psi = \left(\psi^1 ... \psi^K\right)$. The parameter $\psi^k$ measures the average effect of adding a model in the $k^{th}$ band to the assortment of a rival dealership in the market.

We estimate the following linear model (6):

$$y_{rb} = \beta X_{rb} + \psi C_{rb} + \gamma_d W_{m(r)}^d + \epsilon_{rb} \tag{9}$$

were $y_{rb}$, $X_{rb}$ are defined as before and $W_{m(r)}^d$ includes demographics (it does not include measures of market structure). This model is different from (6) because the effect of competition depends not only on the number of dealerships in the market but also the number and type of models they offer. Two GM dealers located in the same local market carrying different assortments therefore face different levels of competition. We define product bands based on Ward's model segmentation, which classifies models into 26 segments based on three dimensions (see online appendix, Table 12): vehicle class (standard car, luxury car, sport utility vehicle, cross utility vehicle, van and pickup), sizes (small, medium and large) and price (lower, middle, upper, etc.).

For our analysis, we focus on groups of products for which at least three product bands can be reasonably defined. We chose small and medium sized standard cars (hereon SM cars, which exclude luxury and large cars) and light-trucks (hereon Trucks, which include SUV, CUV and mini-van)[29]. We defined bands for the segments in each of these two groups and ran two separate regressions. The dependent variable is the logarithm of the average inventory level of models in a specific model segment offered by each dealership. For example, for the SM car regression, inventory of "Lower Small Car" and "Upper Middle Car" of a specific dealership is counted as two different observations. For SM cars, four bands were defined. The first band includes standard cars which have similar

bands, then it is unlikely that the relationship between service level and market structure is driven by unobserved customer loyalty to GM.

[28] We do not know the actual number of models offered since we do not observe inventory of dealerships other than GM.

[29] Full-sized vans and pickups are excluded because we could not obtain inventory data on them. We excluded large and luxury cars because bands for these types of vehicles could not be reasonably defined.

size or price (*B(PRICE,SIZE)*). The second band includes all other standard cars (*B(STDCAR)*). The third and fourth band includes luxury cars and light-trucks, respectively (*B(ANYCAR)* and *B(OTHER)*)[30]. For Trucks, we defined three bands. The first band includes vehicles within the same class with similar size or price (*B(PRICE,SIZE)*). For example, if $b =$"Middle SUV", band 1 includes vehicles in the segments "Middle Luxury SUV" and "Large SUV" but not "Large Luxury SUV" nor "Middle CUV". Band 2 includes all other trucks (*B(TRUCK)*), and band 3 all cars (*B(OTHER)*).

Table 7 summarizes the OLS estimation results of model (9) for SM cars and Trucks. All of the specifications include dummies for region, price (based on the model segmentation), *ALLSTATE* and demographic characteristics. Columns 1 and 3 include the number of vehicles in each band as the measure of competition. While none of the measures are statistically significant, the first band *B(PRICE,SIZE)* has the largest positive point estimate of all the bands. In columns 2 and 4 we add a quadratic term on the number of vehicles on the first band (*NSQ* variables) to capture non-linearities. The results show that the number of vehicles in the first band has a positive effect on the service level, and the marginal effect is decreasing in the number of vehicles. The number of vehicles in the outer bands have no significant effect on the service level (conditional on the number of models in the first band). The results are similar in sign and magnitude across the SM cars and Truck regressions, but the statistical significance of the Truck results are smaller.

To compare the magnitude of the competition effect between this product competition model and our initial “number-of-dealerships” competition model, we calculated the elasticities at the mean implied by each model. For the dealership competition model estimated in Table 6, column 2, the implied elasticity is 35%. For the SM car and Truck product competition models, the implied elasticities are 34% and 19%, respectively. The average elasticities across the models are similar in order of magnitude, suggesting that they are capturing a similar effect: the impact of competition on service level.

Model (9) is estimated with OLS, which can produce biased estimates because $C_{rb}$ is endogenous. The concern is that idiosyncratic consumer tastes for specific type of vehicles will affect product line decisions of dealers and their service levels at the same time, confounding the causal effect of $C_{rb}$. But these specific idiosyncratic consumer tastes are unlikely to be correlated with market population, and therefore should not bias the IV estimates reported in Table 6. On the other hand,

[30] A regression that merges bands 3 and 4, obtains similar results.

the IV regressions can give biased estimates if unobserved customer loyalty for GM is correlated with population. But this confounding effect is unlikely to produce the pattern observed in the product competition model (Table 7). In short, it is hard to find a confounder that biases the estimates of the competition effect in *all* the models we consider, i.e., the estimated effect of competition is robust to different specifications and identification strategies.

To summarize, our empirical results can be interpreted as follows. First, the number of vehicles offered by rivals has a positive effect on the service level of the products offered by a dealership. Second, most of the effect of competition on service level is captured by products which are close substitutes, i.e., a dealer does not respond to the entry of another dealer selling products in different segments but the incumbent dealer does increase its service level in response to the entry of another dealer who sells products in similar segments to the incumbent dealer. Third, there is a saturation effect: the first close substitutes have a large impact on service level, but the effect becomes smaller as more products enter the first band. Overall, these empirical results provide good support for our conjecture that the intensity of inventory competition depends on the number and type of products offered in a market. This pattern is unlikely to be driven by unobserved market characteristics affecting service level.

## 6. The effect of reducing the dealership network

Domestic manufacturers expanded their dealership networks in the early 1900's when a large fraction of the U.S. population lived in rural areas and transportation was difficult. As a result, many dealerships were established so that they could be close to population centers. Japanese manufacturers established their dealership networks in the second half of the century. Due to improved transportation and a greater concentration of the population in urban areas, there are fewer Japanese brand dealerships and they are concentrated in different regions of the country than the domestic manufacturers. (See Figure 1.) Because franchise laws impose high costs on manufacturers for forcing the closure of a dealership, domestic manufacturers still have considerably more dealerships than the Japanese brands. (GM paid more than one billion dollars to Oldsmobile dealers to close that brand, see Welch (2006)) Domestic manufacturers are concerned that the large number of dealerships in their network is causing inefficiencies. As a result, there is discussion on the value of reducing the number of dealerships, despite the costs of doing so (Rechtin and Wilson (2006)).

Closing a dealership in a market has two effects on the inventory of the remaining dealerships. First, some of the sales of the closed dealership are captured by the remaining dealerships. Due to economies of scale, the remaining dealerships will reduce their days of supply. Second, as shown in Table 6, columns 4 and 5, the presence of another GM dealership in a market increases a dealership's service level. Hence, removing a dealership from a market is likely to decrease the remaining dealer's service level, further decreasing that dealers' days of supply.

We used the estimates of Table 6, column 4, to measure the effect of closing some of GM's dealerships.[31] We selected markets with eight or fewer dealerships, and these markets have three or fewer GM dealerships. In the counterfactual, all but one of the dealerships, the one with highest sales volume, remains in the market. To obtain a lower bound, we assumed that all the of sales from the closed dealerships are lost. Thus, the lower bound provides the inventory reduction due only to the lower service level from the reduced competition. To obtain an upper bound, we assumed the remaining dealership captures all of the sales of the closed dealerships. If the remaining dealership carries different brands, the sales are allocated proportionally to the actual sales of each brand. Table 8 summarizes the results from the counterfactual experiment.

The improvement in inventory performance is substantial, between 20 and 38 days-of-supply on average. Interestingly, the service level effect (measured by the lower bound) is similar (and even larger) in magnitude to the potential gains from economies of scale (the difference between the upper and lower bounds). However, we caution that we do not conclude that a large number of dealerships (possible due to restrictive franchise laws) is harmful to consumer welfare - although they may lead to high inventory holding costs, they also provide consumers with potentially lower prices (due to price competition) and a greater selection of products. Similarly, we do not conclude that GM would benefit from reducing its dealership network - we do predict that GM would carry less inventory, but this is only beneficial if the impact on sales is sufficiently small. Whether or not GM benefits from closing dealerships requires an estimate of the sales impact of such closures, an estimation that is beyond the scope of this research.

## 7. Conclusion

We develop an econometric model to estimate the effect of market structure on inventory holdings. We identify two drivers of inventory holdings: a sales effect and a service level effect. We find that

[31] The OLS estimates are more precise than the IV estimates and yield a more conservative reduction.

the sales effect reflects strong economies of scale in managing inventory - increasing a dealer's sales reduces the dealer's inventory when measured in terms of days-of-supply. Based on our estimates, Chevrolet could reduce its days-of-supply by 20% (a 14.8 decrease in days-of-supply) by matching Toyota in sales volume per dealership.

We are particularly interested in the impact of market structure (local competition) on service levels (buffer inventory held by dealerships conditional on sales). Some theoretical models predict that increased competition has no impact on service level (Deneckere and Peck (1995), Lippman and McCardle (1997), Mahajan and van Ryzin (2001)). Others predict increased competition decreases service levels via a margin effect - entry reduces margins via price competition, thereby reducing the incentives to hold inventory (Dana (2001)). Finally, there are models that predict entry increases service levels - firms may increase their service level to attract demand (as in Cachon (2003)) or to better retain demand (i.e., to prevent customers from searching other retailers, as in Cachon et al. (2006) and Watson (2006)). Among dealerships in the automobile industry we find that competition increases service levels, i.e., any margin effect associated with entry is dominated by the demand attraction and/or retention effects. This result contrasts somewhat with the findings in Gaur et al. (2005a) and Amihud and Mendelson (1989). Gaur et al. (2005a) finds that as a retailer lowers its margin, it tends to carry less inventory. (Amihud and Mendelson (1989) has a similar finding but they study manufacturing firms.) If low margins are taken as a proxy for more intense competition, then they find that inventory decreases with competition. They do not study auto retailing, so it is possible that in other retail markets the margin effect of entry dominates the demand attraction/retention effects. Alternatively, lower margins may proxy the use of markdowns to reduce inventory at the end of the season. If retailers use markdowns more aggressively in one year relative to another, margins and inventory will have a positive correlation across years which is unrelated to the margin effect we describe. Further research is needed to reconcile these issues and findings.

Competition increases inventory in the auto industry, but we find that the marginal effect of competition is decreasing: the first entrant into a monopoly market causes a 16% increase in inventory whereas entry beyond the seventh dealership has no positive effect on inventory (conditional on sales). We provide additional empirical evidence showing that the service level of products depends on the number of close substitutes offered by rivals, but is insensitive to the number of dissimilar products. Our results are robust to different econometric specifications.

Our findings suggest that inventory may vary across automobile makes in part because auto makes vary in their dealership structures. As the dealership network becomes more dense there are two reinforcing effects on inventory. One, if sales per dealership declines as the number of dealerships increases (which is plausible), then the presence of economies of sales with respect to sales suggests that inventory, measured in days-of-supply, will increase. Second, an increases in the density of dealerships increases competition among them (i.e., there will be more dealerships per market), which also increases inventory via higher service levels. Thus, when comparing two automobile distribution networks, we expect (all else being equal) the one with the greater number of dealerships to carry more inventory. This conjecture is consistent with aggregate data for U.S. inventory holdings of the major automobile manufacturers. For example, Toyota has approximately half as many dealerships as Chevrolet, and carries 42% less inventory (42 days of supply vs. 73 days of supply). (See Cachon and Olivares (2006) for other factors that explain differences in aggregate inventory holdings.) Furthermore, in the markets that we study, reducing the number of GM dealerships reduces days-of-supply by at least 14% and by as much as 27%.

## References


Alexander, Peter. 1997. Product variety and market structure: A new measure and a simple test. *Journal of Economic Behavior and Organization* **32** 207–214.

Amihud, Yakov, Haim Mendelson. 1989. Inventory behavior and market power. *International Journal of Industrial Organization* **7** 269–280.

Anupindi, Ravi, Yehuda Bassok. 1992. Centralization of stocks: Retailers vs. manufacturer. *Management Science* **45**(2) 178–188.

Balachander, Subramanian, Peter H. Farquhar. 1994. Gaining more by stocking less: A competitive analysis of product availablility. *Marketing Science* **13**(1) 3–22.

Bernstein, Fernando, Awi Federgruen. 2005. Descentralized supply chains with competing retailers under demand uncertainty. *Management Science* **51**(1) 18–29.

Berry, Steve, Joel Waldfogel. 2001. Do mergers increase product variety? Evidence from radio broadcasting. *The Quarterly Journal of Economics* **116**(3) 1009–25.

Bresnahan, Timothy, Peter Reiss. 1990. Entry in monopoly markets. *The Review of Economic Studies* **57**(4) 531–553.

Bresnahan, Timothy, Peter Reiss. 1991. Entry and competition in concentrated markets. *The Journal of Political Economy* **99**(5) 977–1099.

Cachon, Gerard. 2003. Supply chain coordination with contracts. Handbooks in Operations Research and Management Science: Supply Chain Management, S. Graves and T. de Kok. 11:229-339.

Cachon, Gerard, Marcelo Olivares. 2006. Drivers of finished goods inventory performance in the U.S. automobile industry. Working paper, Wharton, University of Pennsylvania.

Cachon, Gerard, Christian Terwiesch, Yi Xu. 2006. The impact of consumer search, firm entry and competition on assortment and pricing. *Marketing Science* Forthcoming.

Dana, James. 2001. Competition in price and availability when availability is unobservable. *RAND Journal of Economics* **32**(3) 497–513.

Dana, James, Nicholas Petruzzi. 2001. Note: The newsvendor model with endogenous demand. *Management Science* **47**(11) 1488–97.

Daugherty, Andrew, Jennifer Reinganum. 1991. Endogenous availability in search equilibrium. *Rand Journal of Economics* **22**(2) 287–306.

Deneckere, Raymond, James Peck. 1995. Competition over price and service rate when demand is stochastic: A strategic analysis. *RAND Journal of Economics* **26**(1) 148–162.

Dranove, David, Mark Shanley, Carol Simon. 1992. Is hospital competition wasteful? *RAND Journal of Economics* **23**(2) 247–262.

Dudey, Marc. 1990. Competition by choice: The effect of consumer search on firm location decisions. *American Economic Review* **80**(5) 1092–104.

Eaton, Curtis, Richard G. Lipsey. 1982. An economic theory of central places. *The Economic Journal* 56–72.

Gaur, Vishal, Marshall L. Fisher, Ananth Raman. 2005a. An econometric analysis of inventory turnover performance in retail services. *Management Science* **51**(2) 181–194.

Gaur, Vishal, Saravanan Kesavan, Ananth Raman, Marshall Fisher. 2005b. Estimating demand uncertainty using judgemental forecasts. Working paper, http://pages.stern.nyu.edu/ vgaur/.

Gerchak, Yigal, Yunzeng Wang. 1994. Periodic-review inventory models with inventory-level dependent demand. *Naval Research Logistics* **41** 99–116.

Hayashi, Fumio. 2000. *Econometrics*. Princeton University Press, Princeton, New Jersey.

Hendricks, Kevin B., Vinod R. Singhal. 2005. Association between supply chain glitches and operating performance. *Management Science* **51**(5) 695–711.

Holweg, Matthias, Fritz Pil. 2004. *The second century*. MIT Press, 5 Cambridge Center, Cambridge, MA 02142.

Lippman, Steven, Kevin McCardle. 1997. The competitive newsboy. *Operations Research* **45** 54–65.

Mahajan, Siddarth, Garrett van Ryzin. 2001. Inventory competition under dynamic consumer choice. *Operations Research* **49**(5) 646–657.

Marx, Thomas G. 1985. The development of the franchise distribution system in the U.S. automobile industry. *The Business History Review* **59** 465–474.

Narus, James A., James C. Anderson. 1996. Rethinking distribution. *Harvard Business Review* 112–120.

Netessine, Serguei, Nils Rudi. 2003. Centralized and competitive inventory models with demand substitution. *Operations Research* **51**(2).

Rajagopalan, S. 2005. Determinants of inventory levels at U.S. retailers: An empirical study. Working paper, University of Southern California.

Rechtin, Mark, Amy Wilson. 2006. Thinning the herd. *Automotive News* **81** 82–84.

Roumiantsev, Serguei, Serguei Netessine. 2006. What can be learned from classical inventory models: A cross-industry empirical investigation. *Manufacturing & Service Operations Management* Forthcoming.

Rudi, Nils, Sandeep Kapur, David F. Pyke. 2001. A two-location inventory model with transshipment and local decision making. *Management Science* **47**(12) 1668–80.

Seim, Katja. 2006. An empirical model of firm entry with endogenous product-type choices. *RAND Journal of Economics* Forthcoming.

Smith, Richard L. 1982. Franchise regulation: An economic analysis of state restrictions on automobile distribution. *Journal of Law and Economics* **25**(1) 125–157.

Stahl, Konrad. 1982. Differentiated products, consumer search, and locational oligopoly. *Journal of Industrial Economics* **31** 97–113.

van Ryzin, Garrett, Siddarth Mahajan. 1999. On the relationship between inventory costs and variety benefits in retail assortments. *Management Science* **45**(11) 1496–1509.

Watson, Randall. 2004. Product variety and competition in the retail market for eyeglasses. Working paper, University of Texas at Austin.

Watson, Randall. 2006. Search, availability, and competition in product ranges. Working paper, Department of Economics, University of Texas at Austin.

Welch, David. 2006. The other club battering GM. *Businessweek* **3974**(38).

Wolinsky, Asher. 1983. Retail trade concentration due to consumers' imperfect information. *Bell Journal of Economics* **14**(14) 275–82.

Wooldridge, Jeffrey. 2002. *Econometric Analysis of Cross Section and Panel Data*. MIT Press, Cambridge, Massachusetts.

Wu, Owen Q., Hong Chen, Murray Z. Frank. 2005. What actually happened to the inventories of American companies between 1981 and 2000. *Management Science* **51**(7) 1015–1031.

Zipkin, Paul. 2000. *Foundations of Inventory Management*. McGraw-Hill.

| Make | Days of supply | No. of dealers | Sales per dealership |
|---|---|---|---|
| Chevrolet | 73 | 4227 | 627 |
| Ford | 74 | 3939 | 795 |
| Honda | 48 | 1001 | 1077 |
| Toyota | 42 | 1200 | 1251 |

Averages during years 1999-2004

Table 1 – Comparison of inventory performance, number of dealerships and sales per dealership across four auto makes. Days-of-supply is calculated based on all finished vehicle inventory in the supply chain, including factory lots, ports of entry, in transit to dealerships and at dealerships.

| Population of UA (thousands) | Distance (in miles) to the closest UA with the following minimum populations | | | |
|---|---|---|---|---|
| | Pop>100 | Pop>50 | Pop>25 | Pop>10 |
| [100,150] | 50 | -- | -- | -- |
| [50,100] | 50 | 30 | -- | -- |
| [25,50] | 50 | 50 | 30 | -- |
| <25 | 50 | 50 | 30 | 30 |

Table 2 – Market selection criteria. A UA with population indicated in column 1 is selected if it meets the criteria in columns 2-5.

| | | | *% of markets with dealers* | | |
|---|---|---|---|---|---|
| # of dealers | # of markets | Median population | non-GM domestic | Japanese | 2nd GM |
| 1 | 17 | 3.91 | 0% | 0% | 0% |
| 2 | 24 | 6.20 | 83% | 8% | 8% |
| 3 | 43 | 9.76 | 100% | 5% | 2% |
| 4 | 29 | 13.05 | 100% | 55% | 41% |
| 5 | 21 | 26.16 | 100% | 86% | 57% |
| 6 | 22 | 36.48 | 100% | 100% | 41% |
| 7 | 14 | 38.37 | 100% | 100% | 71% |
| 8 | 21 | 62.63 | 100% | 100% | 95% |
| 9 | 9 | 68.22 | 100% | 100% | 100% |
| 10 | 13 | 61.47 | 100% | 100% | 100% |
| 11 | 9 | 78.50 | 100% | 100% | 100% |
| 12 | 3 | 100.32 | 100% | 100% | 100% |
| 13 | 6 | 122.00 | 100% | 100% | 100% |
| 14 | 1 | 105.36 | 100% | 100% | 100% |
| 15 | 1 | 122.98 | 100% | 100% | 100% |

Table 3 – Summary statistics of the isolated markets. The last three columns show the percent of markets with at least one dealership of non-GM domestic manufacturers, Japanese manufacturers and a second GM dealership, respectively.

| Variable | Mean | Std. Dev. | Min | Max |
|---|---|---|---|---|
| INCOME | 32.6 | 5.9 | 19.7 | 63.0 |
| ELDER | 15.5 | 4.8 | 4.7 | 38.3 |
| BLACK | 5.4 | 11.9 | 0.0 | 70.1 |
| PUBTRANS | 0.8 | 1.3 | 0.0 | 11.9 |
| COLLEGE | 18.5 | 9.9 | 1.5 | 57.3 |
| FARMING | 1.3 | 1.5 | 0.0 | 12.7 |
| ARMY | 0.7 | 2.4 | 0.0 | 22.6 |

Table 4 – Summary statistics for the demographic variables in the selected markets.

| Variable | Mean | Std. Dev. | Min | Max | *INV* | *SALES* | *NC* | *NGM* | *OUTSIDE* | *NFRANCH* |
|---|---|---|---|---|---|---|---|---|---|---|
| INV | 22.53 | 25.69 | 1.14 | 157.70 | | | | | | |
| SALES | 32.07 | 40.37 | 1.00 | 314.00 | 0.93 | | | | | |
| NC | 4.59 | 2.13 | 1.00 | 8.00 | 0.41 | 0.38 | | | | |
| NGM | 1.50 | 0.63 | 1.00 | 3.00 | 0.38 | 0.35 | 0.69 | | | |
| OUTSIDE | 1.06 | 2.63 | -0.02 | 67.75 | -0.02 | -0.03 | -0.07 | -0.07 | | |
| NFRANCH | 3.40 | 1.25 | 1.00 | 5.00 | -0.37 | -0.37 | -0.34 | -0.51 | -0.04 | |
| TRANSF | 0.07 | 0.09 | 0.00 | 0.63 | -0.09 | -0.02 | -0.11 | -0.04 | -0.04 | 0.06 |

Table 5 – Summary statistics and correlation matrix.

| | Step 1 | Step 2 | | | | Joint | |
|---|---|---|---|---|---|---|---|
| | (1) | (2) | (3) | (4) | (5) | (6) | (7) |
| | OLS | OLS | IV | OLS | IV | GMM | GMM |
| *logSALES* | 0.642** | | | | | 0.674** | 0.656** |
| | (0.024) | | | | | (0.022) | (0.022) |
| *NFRANCH* | 0.027 | | | | | -0.007 | 0.015 |
| | (0.022) | | | | | (0.020) | (0.021) |
| *TRANSF* | -0.744** | | | | | -0.756** | -0.772** |
| | (0.150) | | | | | (0.148) | (0.148) |
| *NC* | | 0.178** | 0.284* | 0.179** | 0.282* | 0.316** | 0.290** |
| | | (0.035) | (0.127) | (0.035) | (0.130) | (0.094) | (0.093) |
| *NCSQ* | | -0.011** | -0.018 | -0.013** | -0.025 | -0.023* | -0.026** |
| | | (0.004) | (0.014) | (0.004) | (0.014) | (0.010) | (0.010) |
| *NGM* | | | | 0.104** | 0.342** | | 0.312** |
| | | | | (0.035) | (0.117) | | (0.090) |
| *OUTSIDE* | | -0.002 | -0.031 | -0.002 | -0.028 | -0.023 | -0.026 |
| | | (0.006) | (0.032) | (0.006) | (0.033) | (0.023) | (0.023) |
| Observations | 684 | 679 | 676 | 679 | 676 | 676 | 676 |
| R-squared | 0.84 | 0.28 | 0.25 | 0.29 | 0.22 | n/a | n/a |

Table 6 – Main estimation results. Demographic variables and dummies for make, region and ALLSTATE are not shown. Column (1) shows the results from the 1st step of the two step method; columns (2)-(5) shows the estimates from the 2nd step. Column (3) uses IVs to instrument for NC and NCSQ. Column (5) uses the same IVs to instrument for NC, NCSQ and NGM. Columns (7)-(8) show the joint estimation using GMM. Standard errors shown in parenthesis. ** and * indicate significance at the 1% and 5%.

| | (1) | (2) | (3) | (4) |
|---|---|---|---|---|
| | SM car | SM car Sq. | Truck | Truck Sq. |
| LOGSALES | 0.5711** | 0.5706** | 0.7957** | 0.7955** |
| | (0.028) | (0.028) | (0.0241) | (0.0241) |
| PERTRANSF | -0.4064* | -0.4273* | -0.3631 | -0.4030* |
| | (0.2037) | (0.2038) | (0.1886) | (0.1894) |
| N B(PRICE,SIZE) | 0.0118* | 0.0332** | 0.0069 | 0.0304 |
| | (0.0057) | (0.0118) | (0.006) | (0.019) |
| NSQ B(PRICE,SIZE) | | -0.0004* | | -0.0006 |
| | | (0.0002) | | (0.0004) |
| N B(STD_CAR) | -0.0125 | -0.0182 | | |
| | (0.0073) | (0.0154) | | |
| NSQ B(STD_CAR) | | 0.0000 | | |
| | | (0.0002) | | |
| N B(ANY_CAR) | -0.0143 | -0.0148 | | |
| | (0.0079) | (0.0129) | | |
| N B(TRUCK) | | | -0.0066 | -0.0049 |
| | | | (0.0042) | (0.0092) |
| NSQ B(TRUCK) | | | | 0.0000 |
| | | | | (0.0001) |
| N B(OTHER) | 0.0004 | -0.0012 | 0.0030 | 0.0030 |
| | (0.0036) | (0.0037) | (0.0048) | (0.0049) |
| Observations | 775 | 775 | 712 | 712 |
| R-squared | 0.6 | 0.6 | 0.69 | 0.69 |

Table 7 – Estimation results with the number of models as a measure of competition. Separate regressions were estimated for Small and Medium size cars (SM cars) and Light-trucks (Truck). N B(.) measures the number of models in each product band and NSQ B() is the square of this measure. Other controls include region and price dummies and demographics (not shown). Standard errors shown in parenthesis. ** and * indicate significance at 1% and 5%.

| # dealers closed | No. obs. | Actual days-of-supply | Reduction in days-of-supply Lower bound | Upper bound |
|---|---|---|---|---|
| 1 | 121 | 142.6 | 19.4 | 36.0 |
| 2 | 21 | 126.6 | 25.0 | 49.6 |
| Total | 142 | 140.2 | 20.2 | 38.0 |

Table 8 -- Effect of closing GM dealerships on inventory.

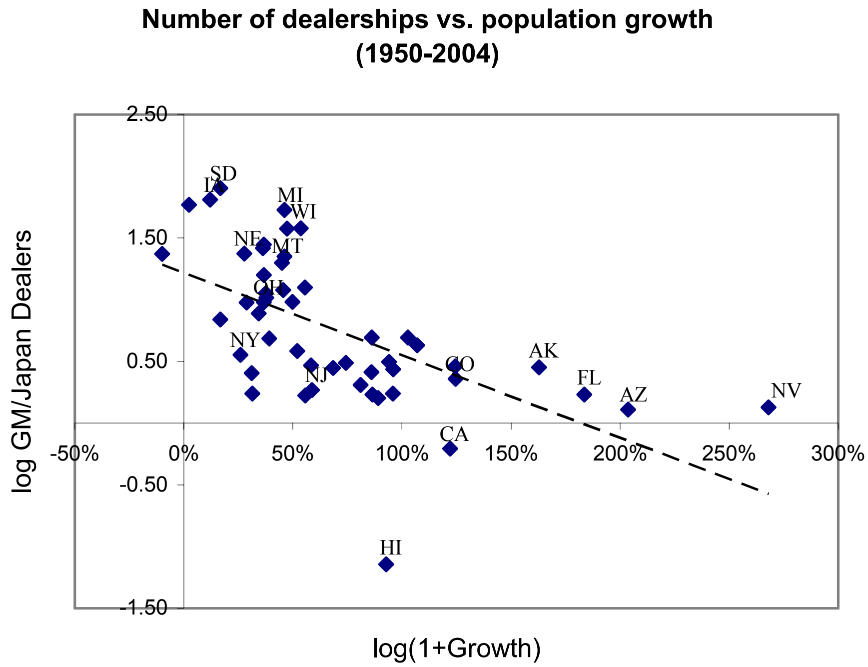


Figure 1 – Scatter plot of the ratio of GM to Japanese dealerships versus population growth between 1950-2004. Each observation is a state and both axes are with logarithm transformation. Selected labels indicate US postal service state codes.

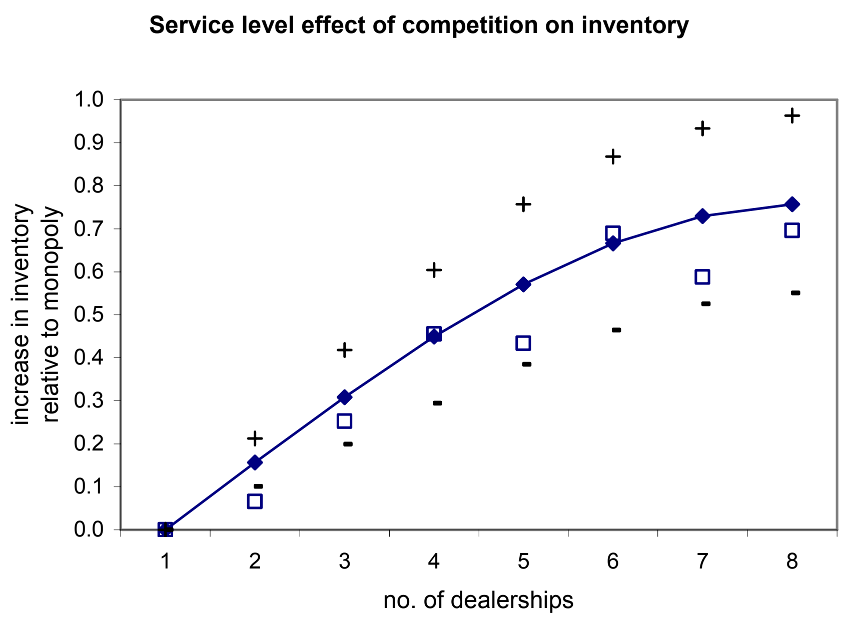


Figure 2 – Effect of competition on the targeted service level, measured as the change in inventory relative to a monopolist GM dealership. The curve shows the estimated effect using specification (2). + and – indicate upper and lower bounds on the 95% confidence interval, respectively, for these estimates. Squares show the estimated effect using a flexible non-parametric function.

## Online Appendix

### Properties of $K(z)$ function

**Lemma 1** *Let $D$ be a normally distributed random variable with mean $\mu$ and standard deviation $\sigma$. Define $Q = \mu + z\sigma$ and $S = \min\{Q, D\}$. Then, the variance of $S$, $V(S)$, is:*

$$V(S) = \sigma^2 \{\Phi(z) - \phi(z)[z + \lambda(z)]\}$$

*where $\phi(z)$ and $\Phi(z)$ are the density and distribution functions of the standard normal, respectively, and $\lambda(z) = \phi(z)/\Phi(z)$ is the hazard rate. Furthermore, $V(S)$ is increasing in $z$.*

**Proof.** Define $Y \equiv Q - D$. Using a result for the normal distribution truncated at zero (see Olsen (1980)):

$$\begin{aligned} V(D|D < Q) &= V(Y|Y > 0) \\ &= \sigma^2 \{1 + \lambda(z)[-z - \lambda(z)]\} \end{aligned} \tag{A1}$$

The variance of sales $V(S)$ can be expressed as

$$\begin{aligned} V(S) &= V(S|D \geq Q)\Pr(D \geq Q) + V(S|D < Q)\Pr(D < Q) \\ &= \sigma^2 \{\Phi(z) - \phi(z)[z + \lambda(z)]\} \end{aligned}$$

where the last line above follows from $(A1)$. It follows that $V(S)$ is increasing in $z$:

$$\frac{d}{dz}\left\{\Phi(z) - \phi(z)\left[z + \frac{\phi(z)}{\Phi(z)}\right]\right\} = \phi(z)(z + \lambda(z))^2 > 0$$

■

**Proposition 2** *In the periodic review base-stock inventory model with normally distributed demand $D$ with mean $\mu$, standard deviation $\sigma$, and order-upto level $Q$, the expected inventory can be expressed as:*

$$I = \sigma_s K(z)$$

*where $\sigma_s = V(S)^{1/2}$ and*

$$K(z) = (z\Phi(z) + \phi(z)) \cdot \{\Phi(z) - \phi(z)[z + \lambda(z)]\}^{-1/2}.$$

*Furthermore, $K(z)$ is increasing in $z$.*

**Proof.** The loss function for the standard normal is given by $L(z) = \phi(z) - z(1 - \Phi(z))$. Combining this expression with the Lemma implies $K(z) = \frac{\sigma}{V(S)^{1/2}}(z + L(z))$. Equation (1) implies $I = \sigma^s K(z)$.

Denote $f_1 = (z\Phi(z) + \phi(z))$ and $f_2(z) = \Phi(z) - \phi(z)[z + \lambda(z)]$. Note that $f_1' = \Phi(z)$ and that $f_2(z) \geq 0$ (otherwise, $V(S)$ could be negative). Taking derivatives of $K(z) = f_1(z) / [f_2(z)]^{1/2}$ we obtain:

$$\begin{aligned}
\text{sign}\left\{K'(z)\right\} &= \text{sign}\left\{\Phi(z) \cdot f_2(z) - \frac{1}{2} f_2'(z) f_1(z)\right\} \\
&= \text{sign}\left\{f_2(z) - \frac{1}{2}\phi(z)(z + \lambda(z))^3\right\} \\
&= \text{sign}\left\{\Phi(z) - \phi(z)[z + \lambda(z)] \cdot \left[1 - \frac{1}{2}(z + \lambda(z))^2\right]\right\}
\end{aligned}$$

which is positive given that $f_2(z) \geq 0$. ■

**GMM Estimation**

Our estimation can be viewed as a special case of multiple equation GMM (see Hayashi (2000), Chapter 4 for a general treatment of multiple equation GMM). We redefine our notation to fit our model into this framework. Define $\epsilon_{1i} = \nu_i$, $\epsilon_{2i} = \varepsilon_i$, $y_{1i} = \dot{y}_i$, $y_{2i} = y_i$ and the column vectors $U_{1i} = \left(\dot{X}_i, \vec{0}_{k_\gamma}\right)$, $U_{2i} = \left(X_i, W_{m(i)}\right)$, where $\vec{0}_{k_\gamma}$ is a column vector of zeros of dimension $k_\gamma$ .We estimate the following moment conditions:

$$\begin{aligned}
E(Z_{1i}\epsilon_{1i}) &= 0 \qquad (1) \\
E(Z_{2i}\epsilon_{2i}) &= 0 \qquad (2)
\end{aligned}$$

where $Z_{1i}$ and $Z_{2i}$ are vectors of exogenous instruments. $Z_{2i}$ includes measures of current and past market population (described in section 3.3) and all demographics included in $W_{m(i)}$. $Z_{1i}$ includes $\dot{X}_i$ and all the covariates included in $Z_{2i}$. Let $k_p = \dim(Z_{pi})$ for $p = 1, 2$, $k_\beta = \dim(\beta)$ and $k_\gamma = \dim(\gamma)$ be the dimensions of the exogenous instruments and the vector parameters $\beta$ and $\gamma$. The error terms are given by:

$$\begin{aligned}
\epsilon_{1i}(\theta) &= y_{1i} - \theta U_{1i} \\
\epsilon_{2i}(\theta) &= y_{2i} - \theta U_{2i}
\end{aligned}$$

Define the stacked column vector

$$g_i(\theta) = \begin{bmatrix} Z_{1i}\epsilon_{1i}(\theta) \\ Z_{2i}\epsilon_{2i}(\theta) \end{bmatrix}$$

The sample counterpart of the moment conditions (1) and (2) is given by $g(\theta) = \frac{1}{n}\sum_{i=1}^{n} g_i(\theta)$, where $n$ is the number of observations.

If $k_1 + k_2 = k_\beta + k_\gamma$ the model is said to be "just identified"; in this case, $\theta$ can be chosen to make $g(\theta) = \vec{0}$. If $k_1 + k_2 < k_\beta + k_\gamma$, the model is not identified: there are infinite values of $\theta$ that yield $g(\theta) = \vec{0}$. If $k_1 + k_2 > k_\beta + k_\gamma$, which is our case, the model is said to be over-identified and $\theta$ is chosen to solve the quadratic form

$$\hat{\theta}(H) = \arg\min_{\theta} g(\theta)' H g(\theta) \tag{3}$$

where $H$ is any square positive-definite matrix of dimension $k_1+k_2$. $H$ is referred to as the weighting matrix, and $\theta(H)$ is consistent for any choice of $H$. Because $g(\theta)$ is linear in $\theta$, (3) can be solved analytically:

$$\hat{\theta}(H) = \left(S'_{zu} H S_z\right)^{-1} S_{zu} H s_{zy} \tag{4}$$

where

$$s_{zy} = \begin{bmatrix} \sum_{i=1}^{n} Z_{1i} y_{1i} \\ \sum_{i=1}^{n} Z_{2i} y_{2i} \end{bmatrix}, \text{ and } S_{zu} = \begin{bmatrix} \sum_{i=1}^{n} Z_{1i} U'_{1i} \\ \sum_{i=1}^{n} Z_{2i} U'_{2i} \end{bmatrix}.$$

There is a choice of $H$ that makes $\theta(H)$ efficient (it minimizes its asymptotic standard error). Standard results of GMM show that efficiency is maximized by choosing $H$ as the inverse of $S = E[g_i g_i']$. The computation of this efficient weighting matrix requires approximating the expectation $E[g_i' g_i]$ whose sample counterpart depends on $\theta$. Hence, one needs to know $\theta$ before computing an estimate of $S$. GMM operates in two steps: in the first step, one can use any weighting matrix $H$ to obtain a consistent estimate of $\theta$; in the second step, we use the estimated $\hat{\theta}$ to compute a consistent estimate of $S$,

$$\hat{S} = \frac{1}{n}\sum_{i=1}^{n} g_i\left(\hat{\theta}\right)' g_i\left(\hat{\theta}\right)$$

and then re-estimate $\theta$ by solving (3) using $H = \hat{S}^{-1}$. In our case, we use the consistent estimate of $\theta$ provided by the two-step method described in section 2 to estimate $\hat{S}$. Denoting $\theta_0$ the true parameter and $\theta^*$ the efficient GMM estimator, the asymptotic variance is given by:

$$\begin{aligned} Avar(\theta^*) &= Var\left(\sqrt{n}\,(\theta^* - \theta_0)\right) \\ &= \left(S'_{zu} \hat{S}^{-1} S_{zu}\right)^{-1} \end{aligned}$$

The estimator $\theta^*$ has a multivariate normal distribution with mean $\theta_0$ and covariance matrix $n^{-1} Avar(\theta^*)$.

## References

Hayashi, Fumio. 2000. *Econometrics*. Princeton University Press, Princeton, New Jersey.

Olsen, Randall. 1980. Approximating a truncated normal regression with the method of moments. *Econometrica* **48**(5) 1099–1106.

Additional Figures and Tables

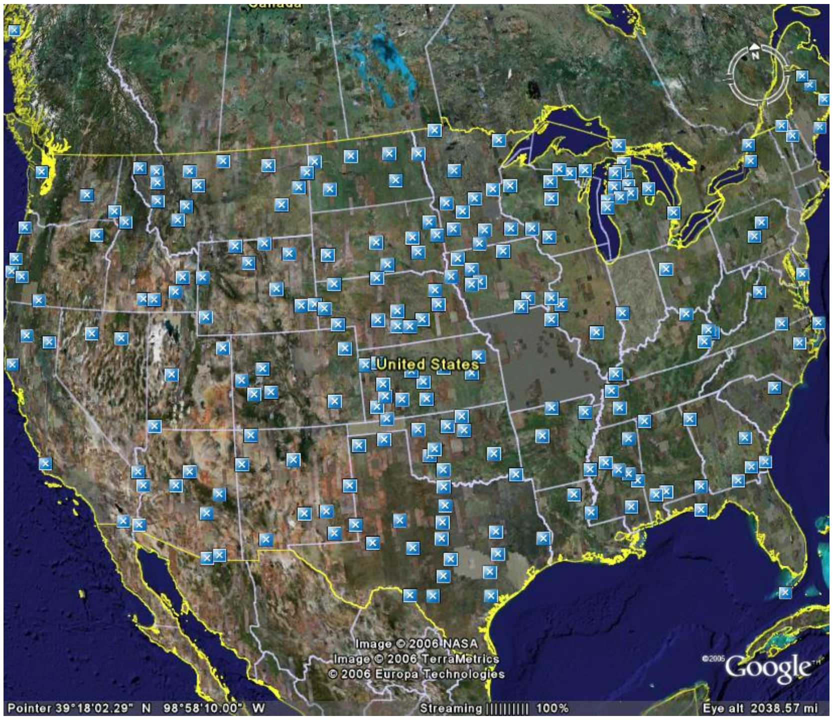


Figure 3: Location of geographically isolated markets.

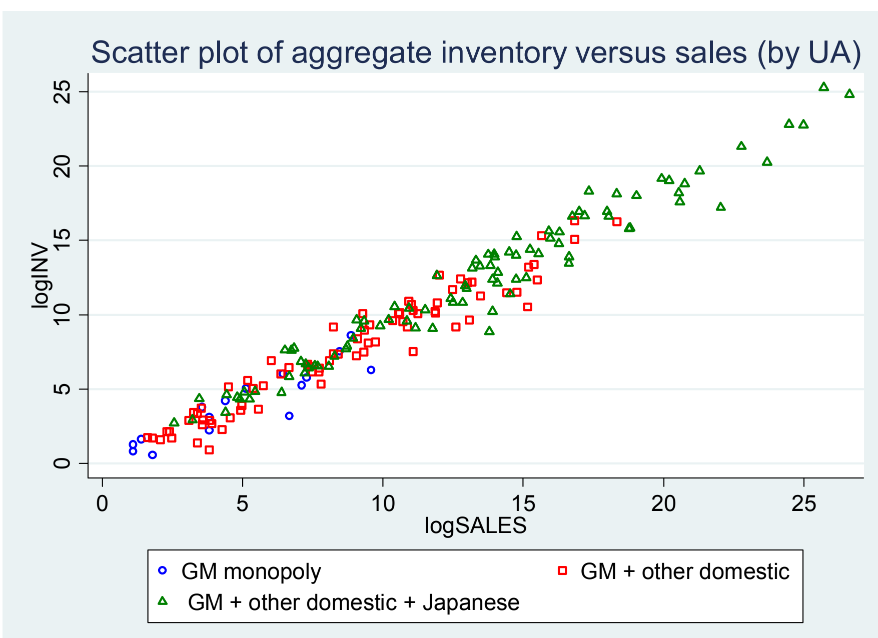


Figure 4 – Scatter plot of log inventory (logINV) and log sales (logSALES) for three types of market: monopoly markets (circles), markets with GM and non-GM domestic dealers (squares) and markets with all types of dealers (GM, non-GM domestic and Japanese). Each observation is a market, and both measures are aggregated across dealerships in each market.

| | Step 1 | | Step 2 |
|---|---|---|---|
| *logSALES* | 0.642** | *NC* | 0.178** |
| | (0.024) | | (0.035) |
| *NFRANCH* | 0.027 | *NDEALERSSQ* | -0.011** |
| | (0.022) | | (0.004) |
| *TRANSF* | -0.744** | *OUTSIDE* | -0.002 |
| | (0.150) | | (0.006) |
| *make=BUICK* | -0.448** | *ALLSTATE* | 0.076 |
| | (0.046) | | (0.040) |
| *make=CADILLAC* | -0.503** | *ELDER* | -0.013** |
| | (0.051) | | (0.004) |
| *make=GMC* | -0.503** | *BLACK* | 0.002 |
| | (0.046) | | (0.002) |
| *make=PONTIAC* | -0.220** | *PUBTRANS* | -0.015 |
| | (0.047) | | (0.012) |
| *make=SATURN* | -0.088 | *LOGINCOME* | 0.378** |
| | (0.135) | | (0.115) |
| | | *COLLEGE* | -0.001 |
| | | | (0.002) |
| | | *ARMY* | -0.002 |
| | | | (0.006) |
| | | *FARMING* | -0.028* |
| | | | (0.014) |
| | | *REGION==MW* | -0.071 |
| | | | (0.075) |
| | | *REGION==NE* | 0 |
| | | | (0.000) |
| | | *REGION==SO* | 0.051 |
| | | | (0.078) |
| | | *REGION==WE* | 0.002 |
| | | | (0.077) |
| | | *CONSTANT* | -3.196** |
| | | | (1.191) |
| Obs | 684 | Obs | 679 |
| R-squared | 0.84 | R-squared | 0.28 |

Table 9 – OLS estimates. Step 1 and Step 2 correspond to columns (1) and (2) in Table 6, respectively.

| | NC | NCSQ | NGM |
|---|---|---|---|
| *LOGUAPOP* | 1.528** | 13.860** | 0.345** |
| | (0.099) | (1.046) | (0.040) |
| *LOGFRINGE* | 0.222** | 1.142 | 0.052* |
| | (0.061) | (0.642) | (0.025) |
| *NGWTH* | 0.041* | 0.255 | 0.023** |
| | (0.018) | (0.191) | (0.007) |
| *PGWTH* | -0.001 | 0.017 | 0.003** |
| | (0.002) | (0.020) | (0.001) |
| *LOGPOP70* | 0.637* | 4.233 | -0.152 |
| | (0.279) | (2.956) | (0.114) |
| *LOGPOP50* | -0.372 | -1.753 | 0.01 |
| | (0.266) | (2.817) | (0.108) |
| *OUTSIDE* | -0.111 | -0.676 | -0.039 |
| | (0.104) | (1.097) | (0.042) |
| *ALLSTATE* | 0.182 | 2.067 | 0.116* |
| | (0.126) | (1.335) | (0.051) |
| *ELDER* | 0.040** | 0.327* | 0.016** |
| | (0.013) | (0.143) | (0.005) |
| *BLACK* | 0.015* | 0.156* | 0.002 |
| | (0.006) | (0.061) | (0.002) |
| *PUBTRANS* | -0.066 | -0.987* | -0.017 |
| | (0.038) | (0.404) | (0.016) |
| *LOGINCOME* | 0.878* | 8.011* | 0.127 |
| | (0.371) | (3.928) | (0.151) |
| *COLLEGE* | 0.011* | 0.122* | 0.004 |
| | (0.005) | (0.052) | (0.002) |
| *ARMY* | 0.02 | 0.354 | 0.013 |
| | (0.019) | (0.198) | (0.008) |
| *FARMING* | -0.183** | -2.109** | -0.118** |
| | (0.047) | (0.502) | (0.019) |
| *REGION==MW* | -0.211 | -1.328 | 0.149** |
| | (0.136) | (1.436) | (0.055) |
| *REGION==NE* | -0.281 | -3.868 | 0.621** |
| | (0.270) | (2.855) | (0.110) |
| *REGION==SO* | -0.603** | -5.468** | 0.111 |
| | (0.176) | (1.865) | (0.072) |
| *CONSTANT* | -10.667** | -111.751** | -0.787 |
| | (3.897) | (41.259) | (1.589) |
| *Observations* | 676 | 676 | 676 |
| *R-squared* | 0.69 | 0.64 | 0.42 |

Table 10 – Estimates from first step of IV regression. The dependent (endogenous) variable is shown in the header. Exogenous variables excluded from *W* are UAPOP, FRINGEPOP, PGWTH, NGWTH, POP50 and POP70.

| Manufacturer | Total no. of models | Standard car | Luxury car | SUV and CUV | Van | Pickup |
|---|---|---|---|---|---|---|
| GM | 56 | 27% | 14% | 36% | 11% | 13% |
| Ford | 38 | 24% | 26% | 37% | 5% | 8% |
| DC | 25 | 24% | 8% | 48% | 12% | 8% |
| Toyota | 26 | 35% | 19% | 35% | 4% | 8% |
| Nissan | 16 | 31% | 13% | 38% | 6% | 13% |
| Honda | 14 | 21% | 29% | 36% | 7% | 7% |
| Total | 175 | 27% | 18% | 38% | 8% | 10% |

Table 11 – Number of models offered by each manufacturer in model-year 2007. The percentages indicate the fraction of models offered on each model segment, based on Ward's Auto model segmentation.

| **Segment** | **class** | **size** | **price** | **no. of models** |
|---|---|---|---|---|
| Lower Luxury | luxury car | large | lower | 17 |
| Luxury Specialty | luxury car | middle | specialty | 5 |
| Luxury Sport | luxury car | small | specialty | 14 |
| Middle Luxury | luxury car | large | middle | 11 |
| Upper Luxury | luxury car | large | upper | 6 |
| Large Regular | standard car | large | upper | 10 |
| Lower Middle | standard car | middle | lower | 6 |
| Lower Small | standard car | small | lower | 7 |
| Middle Specialty | standard car | middle | specialty | 9 |
| Small Specialty | standard car | small | specialty | 6 |
| Upper Middle | standard car | middle | upper | 16 |
| Upper Small | standard car | small | upper | 20 |
| Large Cross/Utility Vehicle | CUV | large | std | 7 |
| Large Luxury Cross/Utility Vehicle | CUV | large | luxury | 6 |
| Middle Cross/Utility Vehicle | CUV | middle | std | 20 |
| Middle Luxury Cross/Utility Vehicle | CUV | middle | luxury | 12 |
| Small Cross/Utility Vehicle | CUV | small | std | 7 |
| Large Pickup | pickup | large | std | 9 |
| Small Pickup | pickup | small | std | 11 |
| Large Luxury Sport/Utility Vehicle | SUV | large | luxury | 8 |
| Large Sport/Utility Vehicle | SUV | large | std | 9 |
| Middle Luxury Sport/Utility Vehicle | SUV | middle | luxury | 5 |
| Middle Sport/Utility Vehicle | SUV | middle | std | 16 |
| Small Sport/Utility Vehicle | SUV | small | std | 1 |
| Large Van | van | large | std | 4 |
| Small Van | van | small | std | 12 |

Table 12 – Model segmentation (Source: Ward's Auto).