Does Adding Inventory Increase Sales? Evidence of a Scarcity Effect in U.S. Automobile Dealerships

Gérard P. Cachon  
*The Wharton School, University of Pennsylvania

Santiago Gallino  
*Tuck School of Business, Dartmouth

Marcelo Olivares  
*Columbia Business School and Universidad de Chile

June 28, 2013

Abstract

What is the relationship between inventory and sales? Clearly, inventory could increase sales: expanding inventory creates more choice (options, colors, etc.) and might signal a popular/desirable product. Or, inventory might encourage a consumer to continue her search (e.g., on the theory that she can return if nothing better is found), thereby decreasing sales (a scarcity effect). We seek to identify these effects in U.S. automobile sales. Our primary research challenge is the endogenous relationship between inventory and demand - e.g., dealers influence their inventory in anticipation of demand. Hence, our estimation strategy relies on weather shocks at upstream production facilities to create exogenous variation in downstream dealership inventory.

We find that the impact of adding a vehicle of a particular model to a dealer’s lot depends on which cars the dealer already has. If the added vehicle expands the available set of sub-models (e.g., adding a four-door among a set that is exclusively two-door), then sales increase. But if the added vehicle is of the same sub-model as an existing vehicle, then sales actually decrease. Hence, expanding variety should be the first priority when adding inventory - adding inventory without expanding variety is actually detrimental. Based on this insight, given a fixed set of cars, vehicles should be allocated among a group of dealers so as to maximize each dealer’s variety. Our data indicate that the implementation of this strategy could increase expected sales by about 2.5% without changing the total number of vehicles in the market, which vehicles are

*The authors thank seminar participants at Columbia University, Cornell University, Dartmouth College, Duke University, Harvard University, Johns Hopkins University, Northwestern University, University of Michigan, University of North Carolina at Chapel Hill, University of Pennsylvania, University of Texas at Austin, University of Texas at Dallas, University of Wisconsin-Madison and the Empirical Workshop in Operations Management at the Wharton School.
produced or the number of vehicles at each dealership. If the firm is willing to reduce aggregate inventory by a modest 2.9% (and no more than 10% at any one dealer), then the sales impact of the “maximize variety, minimize duplication strategy” doubles, to 5.0%.
1 Introduction

In early 2008, before the financial crisis, car dealerships in the United States (U.S.) held enough vehicles to cover sales for 75 days (WardsAuto market data). However, immediately following the financial crisis automakers began drastic reductions in their inventories. By January 2010, days-of-supply for the industry had dropped to less than 49, leading many dealers to complain that their low inventories were negatively affecting sales (AutomotiveNews (2010)). Were those complaints justified?

Adding inventory can increase sales for several reasons. With more inventory, a dealer can expand the variety of options (e.g., trim, colors, options) immediately available to a customer. This increases the chance a customer finds a vehicle that sufficiently matches her preferences, thereby increasing the likelihood of a sale. While this variety effect does not directly change the customer’s fundamental demand, it is also possible that in some product categories, like automobiles, more inventory actually modifies preferences. For example, seeing many cars on a dealers lot might cause a customer to infer that the car is popular (a dealer carries many cars only if the model is popular), thereby making the car more desirable to the customer and increasing the chance the customer purchases the vehicle. In contrast, ample inventory could create the opposite inference: if there are many cars, then the car must not be popular, and it must not be popular for a reason, so the customer becomes less likely to purchase. Or, seeing that many units are available, a customer may become more likely to continue her shopping/search process at other dealers because she believes that if she decides to return to the dealership, the car she desires will still be available. Once the customer leaves, she might not return (because there is a chance she will find a better alternative elsewhere), so this more active search lowers the dealership’s sales.

In general, for simplicity, we use the label “variety effect” for any mechanism that assigns a positive relationship between inventory and demand, with the understanding that some mechanisms may not be directly related to product variety (as in when multiple units of the same product signals popularity/desirability). And we use “scarcity effect” for any mechanism with a negative relationship. Our objective is to empirically evaluate the strength of these effects in the U.S. auto industry.

While it is possible to identify several mechanisms that connect inventory to sales, estimating the relationship between inventory and sales is complex primarily because it is reasonable to believe that inventories are chosen endogenously. For example, a simple plot reveals a positive relationship
between the amount of inventory a dealer carries and the dealer’s average weekly sales. But dealers that operate in larger markets are expected to carry more inventory and have higher sales even if inventory has no influence on demand merely because a firm rationally needs to carry more inventory when it serves more demand. To overcome this selection effect, we estimate the influence of inventory using only observed variation within dealer-model pairs rather than variation across dealerships and models. This approach is valid given the assumption that a dealer’s market conditions are reasonably constant in our six-month study period (e.g., there is little change in local factors like demographics, population, or the degree of competition the dealer faces). However, even within a dealer-model pair, there is a concern that a dealer may change her inventory level in anticipation of changes in demand. For example, the dealer may build inventory due to a planned promotion. In that situation it is incorrect to conclude that the larger inventory caused the higher sales. To overcome this issue, we exploit shocks to dealers’ inventories due to weather disruptions in upstream production. Extreme weather disrupts production via a number of mechanisms (e.g., delays in inbound or outbound shipments, worker absenteeism, etc.) and also is independent of dealer demand (as production generally occurs at a considerable distance from the dealership), thereby providing a valid instrument that allows us to estimate the causal impact of inventory on sales. Given our results, we are then able to estimate the increase in sales that could be achieved if vehicles were allocated differently across dealerships.

2 How inventory impacts sales

In this section we first describe several mechanisms by which more inventory can increase sales and then discuss mechanisms that predict the opposite.

At a basic level, it is intuitive that more inventory can increase sales because of stockouts: focusing on a single item with \( q \) units of inventory and stochastic demand, \( d \), that is independent of \( q \), as inventory increases, so does expected sales, \( E[\text{min}(q,d)] \), simply because a stockout occurs when \( d > q \) (assuming backorders are undesirable). However, the magnitude of this effect diminishes with inventory, i.e., the effect is small when \( q \) is large. Agrawal and Smith (1996) and Anupindi et al. (1998) develop methods to estimate retail demand when lost sales are unobservable.

Inventory can also increase sales through its influence on demand. For example, Balakrishnan et al. (2004) assume that inventory increases demand and study how this effect influences single-product inventory decisions in a deterministic environment (which eliminates the confound of a
stockout effect due to randomness). Inventory can also influence demand in an environment with product variety. For example, say a retailer stocks similar items that differ in several attributes (e.g., engine size, body style) and consumers have heterogeneous preferences over those attributes. Increasing inventory may increase the breadth of attributes available to consumers, thereby increasing demand (because consumers are more likely to find an item that matches their preference), and in turn this leads to higher expected sales. This is similar to a stockout effect in which each possible variant is considered separately. There is an extensive literature on consumer choice that offers a number of approaches for modeling variety (e.g., multinomial logit, nested logit, etc.). See Train (2009) for an overview. Musalem et al. (2010), Conlon and Mortimer (2008) and Kök and Fisher (2007) use some of these consumer choice models to estimate how changes in an assortment affect sales. There is also work that combines the inventory choice decision with one of these consumer choice models (see Talluri and van Ryzin (2004), Smith and Achabal (1998), Gaur and Honhon (2006)).

Continuing to hold preferences constant, inventory could increase sales by influencing a consumer’s engagement in the purchasing process. For example, if a consumer is not aware of an item, the consumer cannot even consider purchasing it - a large inventory may act like a billboard and increase awareness. Or, a consumer may infer that a large inventory implies a low price (e.g., the item must be on promotion or the dealer will be willing to negotiate a good deal), thereby motivating the consumer to include the item in her consideration set (see Zettelmeyer et al. (2006) for a study on the effect of dealership inventory on prices). Finally, if search is costly, then consumers are more likely to visit (and therefore buy from) a dealer that has a reputation for higher inventory - nobody likes to go to a store only to find out that the desired item is unavailable (e.g., Deneckere and Peck (1995), Dana Jr. and Petruzzi (2001), Bernstein and Federgruen (2004), Su and Zhang (2009), Matsa (2011)). Craig et al. (2011) studies a similar mechanism in the context of a manufacturer selling to retailers, showing that an increase in the fill-rate of a retailer’s orders increases its demand. Alternatively, inventory could influence demand by directly influencing preferences. For example, a consumer might infer from a large inventory that the item has good quality (why else would the dealer have so many), thereby making the item more desirable to the consumer - a good quality item has useful features and durability.\footnote{See Guajardo et al. (2012) for an empirical study on the effect of quality and reliability on automobile demand.}

In contrast to the various mechanism by which inventory increases sales, there are several mechanisms that lead to a scarcity effect in which more inventory actually lowers sales. This could
happen if consumers infer that an item with ample inventory is unpopular or low quality - there must be many units because nobody is buying the item (e.g., Balachander et al. (2009), Stock and Balachander (2005)). Or, a consumer might prefer an item that is perceived to be exclusive or rare, as in a collectible (e.g. Brock (1968); Brehm and Brehm (1981); Worchel et al. (1975); Tereyağölü and Veeraraghavan (2012)). This may apply to some specialty vehicles in the auto industry, but probably not to the sample of mainstream vehicles we consider.

If it is costly for consumers to consider all possible options, then low inventory may imply a low variety of options and higher confidence that a good option has been identified (e.g. Kuksov and Villas-Boas (2010)). Similarly, high inventory and high variety may create confusion or frustration (too many options to know where to begin), thereby leading to lower demand and sales (e.g. Iyengar and Lepper (2000), Schwartz (2004)).

A large inventory may indicate that a product will be available later on at a good price (because the dealer may need to discount the item), thereby encouraging consumers to wait before buying, which lowers current sales (e.g., Aviv and Pazgal (2008), Su and Zhang (2008), Cachon and Swinney (2009)). In contrast, with a low current inventory consumers not only anticipate that the price will not fall, they also anticipate that the item may not be available in the future. This can lead to a “buying frenzy” in which the low current inventory creates a sense of urgency among consumers to buy immediately (DeGraba (1995), Qian and van Ryzin (2008)). A similar effect can materialize in search behavior. Say a consumer finds a vehicle that she likes at a dealership. If the dealer has only one of that type of car, she may be inclined to stop her search and just buy the car - if she continues her shopping at other dealers, then she risks not finding a better car and losing the current car to another customer. But if the dealer has several of her desired cars, she may be more inclined to continue her search, and that search may lead her to make a purchase from some other dealership. See Cachon et al. (2008) for a model in which variety influences the degree of consumer search.

To summarize, there are several mechanisms that suggest more inventory increases sales (ample inventory enables a better preference match, increases awareness, signals popularity, indicates availability and suggests the potential to obtain a good price). For simplicity, we collectively refer to these as variety effects given that variety is likely to be a key factor in consumer purchasing decisions in the auto industry. In contrast, other mechanisms suggest more inventory decreases sales (ample inventory reduces the urgency to purchase immediately while encouraging additional search, signals an unpopular vehicle, creates overwhelming choice, and suggests that prices will soon be lowered). We refer to these as scarcity effects.
3 Data Description and Definition of Variables

As a general reference, during the period of our study six car companies accounted for about 90% of sales in the U.S. auto market. The company we focus on, General Motors (GM), captured 25% of the market. This market share was distributed across several different brands: Chevrolet, GMC, Pontiac, Buick, Saturn, Cadillac and Hummer.

The data used in our analysis can be separated in two groups. The first group includes the inventory and sales information for the dealers in our sample. The second group includes geographic location, weather information for all the GM dealers in our sample and all GM plants located in the U.S. and Canada.

3.1 Dealer’s sales and inventory data.

We obtained, via a web crawler, daily inventory and sales data from a website offered by GM that enables customers to search new vehicles inventory at local dealerships. The data collection was done from August 15, 2006 to February 15, 2007, and includes a total of 1,289 dealers in the following states: California, Colorado, Florida, Maine, Nebraska, Texas and Wisconsin. These states are geographically dispersed and somewhat geographically isolated - they may border with Mexico or Canada or have a substantial coastline. The dealers in the sample are all the GM dealers in those states and they represent approximately 10% of all GM dealers in the U.S. for the period under analysis.

The crawler collected specific information for each vehicle at a dealer’s lot, such as its trim level, options, list price and Vehicle Identification Number (VIN). Our sample of GM vehicles includes all cars and a large portion of light-truck models manufactured and sold in the U.S. and Canada. VINs uniquely identify all vehicles in the U.S. Thus, they provide three key pieces of information. First, the VINs allow us to identify when a new car arrived at a dealer and when a sale happened (a vehicle is removed from a dealer’s inventory). Second, the VIN code identifies the particular plant where the vehicle was produced even if the model is manufactured at multiple plants. Finally, the VINs provide us with information regarding dealer transfers - we can observe when a vehicle is removed from one dealer’s inventory and added to another dealer’s inventory within the state.²

²If a vehicle leaves a dealer in week \( t \) and does not reappear in another dealer’s inventory in week \( t + 1 \), then we code this as a sale. Otherwise, it is coded as a transfer. For example, car A is transferred from dealer 1 to dealer 2 and then sold at dealer 2, a sale is counted only at dealer 2. We can only observe transfers between dealerships within the same state. We anticipate that we observe the majority of transfers because transfers probably occur in a limited geographic area and the isolated states we analyze have fewer borders with other states.
We removed from our sample a limited number of dealerships that opened or closed during the period under analysis.

### 3.2 Geographic location and weather data

For each dealer and all 22 GM plants supplying vehicles in our sample (located in the U.S. and Canada), we obtained their address and exact geographic location (longitude and latitude) from GM’s website.

We identified the closest weather station to each plant and each dealer. The selected weather stations are close to our plants with a mean and median distance of 12 and 10 miles, respectively. No plant is further than 32 miles from its corresponding weather station. To assess whether a station’s weather is likely to be similar to the weather at its nearby plant, we constructed a sample of weather stations that are between 30 and 60 miles apart. In this sample, the correlation in our weather variables is no less than 95%, suggesting that the weather reported at the nearby weather station is representative of the weather at the plant\(^3\).

Using the website from the National Weather Service Forecast Office (NWSFO) and www.wunderground.com, we obtained daily weather information for every dealership and plant location in our sample for the period August 15, 2006 to February 15, 2007. Section 4 describes in detail the weather variables included in our analysis.

Table 1 summarized the number of dealers in each state and Figure 2 shows the geographic location of GM plants and the dealers in our sample.

### 4 Model Specification

We seek to estimate the impact of inventory and variety on sales. The available data were used to construct a panel data-set where the unit of analysis is the log of sales plus one of a particular vehicle model \(i\) at a specific dealership \(j\) during a week \(t\) (\(Sales_{ijt}\)). Expected sales of model \(i\) at dealer \(j\) in week \(t\) is influenced by the total number of vehicles available at the dealership during the week (\(Inventory_{ijt}\)), the available variety of that model (\(Variety_{ijt}\), to be described in more detail shortly), plus other factors that could influence the demand for vehicles at the dealership. Figure 1 illustrates the relationship between the key variables in our analysis – sales, inventory and variety.

\(^3\)The locations consider for this analysis were: Marysville, Ohio and Columbus, Ohio; Washington DC and Baltimore, Maryland; Kansas City, Missouri, and Topeka, Kansas; Lansing, Michigan and Grand Rapids, Michigan
Figure 1 shows multiple effects between the three key variables. First, there is a direct effect of inventory on sales (labeled with the coefficient $\beta_{IS}$). An example of this effect is when a low level of inventory signals low future availability of the vehicle model that leads to a “buying frenzy” behavior, or when a high level of inventory signals lower prices which then increase sales. Therefore, the sign of $\beta_{IS}$ is ambiguous. Second, there is a direct effect of variety on sales (labeled by $\beta_{VS}$), as when more variety leads to a better match of customer preferences, thereby increasing sales. Higher variety could also lead to more confusion in choosing among too many options, lowering sales. Hence, the sign of $\beta_{VS}$ is also ambiguous. Third, there is an indirect effect of inventory on sales through variety (labeled $\beta_{IV}$): adding inventory can lead to an increase in variety, which in turn could affect sales.

The estimation can be viewed as a system of simultaneous equations with three endogenous variables – $Sales_{ijt}$, $Variety_{ijt}$ and $Inventory_{ijt}$. Let $Sales$, $Variety$ and $Inventory$ be vectors containing the observations for these three variables, respectively (indexes $i, j, t$ are therefore suppressed). The system is given by:

\[
\begin{align*}
Sales &= \beta_{VS} Variety + \beta_{IS} Inventory + \gamma_S Z + \varepsilon_S \\
Variety &= \beta_{IV} Inventory + \gamma_V Z + \varepsilon_V \\
Inventory &= \gamma_I Z + \delta_I W + \varepsilon_I
\end{align*}
\]

The matrix of covariates $Z$ is a set of exogenous controls to be specified in detail later. The matrix of covariates $W$ is a set of weather shocks at the plant that produces a specific model. The error vectors $\{\varepsilon_g\}_{g=S,V,I}$ represent unobservable factors that affect each of the endogenous variables. Throughout we assume that $Z$ and $W$ are predetermined in the three equations, in the sense that $E(\varepsilon_g | Z, W) = 0$, for $g = S, V, I$. Next, we discuss identification of the system of equations (1)-(3).

The error term $\varepsilon_S$ represent factors that affect sales which are unobservable in the data. Dealerships and manufacturers may predict some of these factors in advanced and use them in their demand forecast to choose inventory levels (see Figure 1 for an illustration). Hence, $\varepsilon_S$ and $\varepsilon_I$ are likely to be positively correlated, making $Inventory$ endogenous in the sales equation (1).

$W$ is excluded from the sales equation but included in the inventory equation. If weather at the plant affects its productivity, then weather shocks at the plant affect the inventory level at the dealerships; this effect is captured by the coefficient $\delta_I$ in equation (3). A dealer’s local weather
is included in $Z$, but because most of the plants are located far away from the dealerships in our study, weather shocks at the plants should be unrelated to the local demand for autos. Hence, $W$ is excluded from equation (1). Consequently, the explanatory variables in $W$ are valid instrumental variables for $Inventory$ in equation (1).

Nevertheless, this exclusion restriction on $W$ is insufficient to identify the parameters of the system of equations; in fact, the parameters of the first equation are not identified without additional assumptions because as inventory also affects variety ($Inventory$ is an explanatory variable in equation (2)), $Variety$ is also endogenous in equation (1). Hence, $Variety$ must be instrumented to obtain consistent estimates of the coefficient in equation (1). Note that although $W$ affects inventory, it does not have any further effect on the variety of vehicles; that is, $W$ is excluded from equation (2) (i.e., $\delta_V = 0$). Weather at the plant is a productivity shock that affects total production at the plant but not the mix of vehicles that are produced at the plant. Hence, $W$ is not a valid instrument for $Variety$.

In the absence of further exclusion restrictions of the exogenous variables ($Z, W$), identification of the system (1)-(3) requires assumptions about the covariance structure of the errors ($\varepsilon_S, \varepsilon_V, \varepsilon_I$). As previously mentioned, it is likely that $\varepsilon_S$ and $\varepsilon_I$ are positively correlated due to inventory endogeneity. However, it is reasonable to assume that $\varepsilon_S$ and $\varepsilon_V$ are uncorrelated, that is, $E(\varepsilon_S \varepsilon_V) = 0$. Although dealerships can control to some extent the number of vehicles of a particular model that they receive, they typically have little control on the exact sub-models that are allocated to them. Therefore, the variations in variety after controlling for inventory levels should be unrelated with the demand forecasts or other unobservable factors related to demand. Moreover, it is also reasonable that $E(\varepsilon_V \varepsilon_I) = 0$: because dealers can only control variety through their inventory levels, other factors that induce variation in variety (captured by $\varepsilon_V$) should be unrelated to factors that affect inventory. These assumptions are sufficient for identification, as shown next.

**Proposition 1.** If $E(\varepsilon_g|Z, W) = 0$, for all $g \in \{S, V, I\}$, $E(\varepsilon_S \varepsilon_V) = 0$ and $E(\varepsilon_V \varepsilon_I)$, then all the parameters of the system of equations (1)-(3) are identified.

See the appendix for the proof.

We need instrumental variables to estimate the parameters of equation (1) because $Variety$ and $Inventory$ are endogenous. As noted earlier, the exogenous plant weather variables $W$ are excluded from (1) and can therefore be used as instruments for $Inventory$. Moreover, under the assumption $E(\varepsilon_S \varepsilon_V) = 0$, the residual of equation (2) can be used as an instrument for $Variety$ in equation (1).
This requires a consistent estimator of $\varepsilon_V$. Under the assumption that $E(\varepsilon_V \varepsilon_I) = 0$, the residual of the OLS regression of (2), denoted $\hat{e}_V$, is a consistent estimator of $\varepsilon_V$. Thus, the following method can be used to obtain consistent estimates of the coefficients of equation (1):

1. Estimate regressions (2) and (3) via OLS.

2. Compute the fitted values $\hat{\text{Inventory}} = \hat{\gamma}_I Z + \hat{\delta}_I W$ and the residuals $\hat{e}_V = \text{Variety} - \beta_{IV} \hat{\text{Inventory}} - \hat{\gamma}_V Z$.

3. Estimate equation (1) via Two-Stage Least Square using $\hat{e}_V$ and $\hat{\text{Inventory}}$ as instrumental variables for the endogenous variables Variety and Inventory.

The use of instrumental variables as an empirical strategy to identify causal effects has become prevalent in the Operations Management literature—e.g., see Kc and Terwiesch (2012) in a healthcare application and Parker et al. (2012) in the context of IT adoption. See Kesavan et al. (2010) for an alternative approach to implement instrumental variables for aggregate company-wide inventory levels in cross-sectional studies using financial data.

**Controls**

$Z$ includes model-dealership fixed-effects which control for the invariant characteristics of each dealer: each dealer location, the average popularity of a model at a particular dealership, the intensity of competition a model faces at each dealer, the average discount policy a dealer offers for a particular model, etc. $Z$ also includes a seasonality dummy variable to account for changes in the sales across weeks. This is implemented by grouping dealers into four geographic regions: {Florida, Texas}, {Colorado, Nebraska}, {Maine, Wisconsin}, and {California}. Let $r(j)$ be the region containing dealership $j$. We include the set of dummy variables $\text{Seasonal}_{r(j)}t$ to control for different seasonal patterns across geographic regions, e.g., a different weekly sales pattern in Texas than in Wisconsin. Finally, as already mentioned, $Z$ includes measures of local weather at each dealership to control for the effect of local weather on sales and demand forecasts. (See Perdikaki et al. (2012), Steele (1951) and Murray et al. (2010) for examples of how local weather affects retail sales. There is also anecdotal evidence of this relationship in the public press, e.g. BloombergTV (2012)).
Measuring variety

To identify which of the main effects of inventory on sales described earlier dominates, we identify separately the impact of our two measures of availability – inventory and variety. For example, a negative effect of $\text{Variety}_{ijt}$ would suggest that the confusion effect dominates the impact on sales. Although $\text{Inventory}_{ijt}$ can be objectively defined as the number of vehicles available for a model, variety could be defined in many different ways depending on the relevant product characteristics that are considered by customers when making their purchase decision. For example, a customer wanting to buy a Chevrolet Malibu may consider two vehicles with different horsepower as two different products, but could be indifferent on the color of the car. To measure $\text{Variety}$, it is necessary to define a set of attributes that describes relevant differences across vehicle options within a model. See Hoch et al. (1999) for a framework on how customers perceive variety and Li and Netessine (2012) for empirical approach to identify customer perceptions of variety from online search data.

The VIN of a vehicle contains information about vehicle characteristics, including the model, body style, engine type and restraint type. We use all these relevant characteristics reported in the VIN to define the different possible variants of a model and we refer to each variant as a *sub-model*. The variable $\text{AvailVar}_{ijt}$ is the number of sub-models of a model $i$ available at dealership $j$ during week $t$. The assumption being that the variety information included in the VIN describes relevant differences across vehicle options from the customer perspective. However, $\text{AvailVar}_{ijt}$ may not capture all relevant differences. For example, VINs do not report color, so while two vehicles in the same sub-model are similar, they need not be identical.

Table 2 summarizes the number of different sub-models observed in our data and the average $\text{Variety}_{ijt}$ observed at the dealerships for a sample of models. The table reveals that there is variation in the number of sub-models available across the set of models. Hence, it is plausible that the impact of variety is different across models: for example, adding to a dealer’s inventory one more sub-model of a Cobalt (which has 18 sub-models) can have a smaller impact than adding one more sub-model of an Equinox (which has 4 sub-models). To account for this, the amount of available variety can be measured relative to the number of sub-models that exist for that model. Denote $\text{MarketVar}_j$ as the number of sub-models produced for model $j$ in the model-year 2007.
Our main measure of variety is defined as:

\[ Variety_{ijt} = \frac{AvailVar_{ijt}}{MarketVar_j}. \]  

(4)

**Weather Instrumental Variables**

There are multiple mechanisms by which plant weather can influence dealership inventory. Bad weather can affect the supply of parts to the production line slowing the production process. In addition, weather conditions can affect employee behavior both in their task performance and by increasing absenteeism. Alternatively, weather can delay shipments of vehicles to dealers. Consistent with these mechanisms, Cachon et al. (2011) provide evidence that weather in the vicinity of an assembly plant affects its productivity.4

Our weather variables are defined as in Cachon et al. (2011) and are described in detail in Table 3. We included Wind, Fog, Rain and Snow variables because each of these weather events may influence travel to and from a plant. Cloud could proxy for other inclement weather and could influence employee behavior. High Temp is included because it could influence ambient temperature within the plant or employees that must work outside (e.g., loading docks). Low Temp may proxy for hazardous road conditions (e.g. ice). Some of the variables, such as Wind and Cloud, directly capture weather shocks. For other measures –specifically for Fog, Rain, High Temp, Low Temp and Snow– we estimated specifications including multiple levels of the variable to capture potential non-linear effects on production.

Some of these weather variables have a weak impact on dealership inventory, in part because of the high correlation between the many alternative measures of weather that we considered. Using a large number of instruments in a two-stage least square estimation can induce bias on the estimates (Buse (1992)). There is also a rich literature that discusses other challenges that can arise when dealing with multiple instruments, in particular when some of these instruments might be weak (Bekker (1994), Donald and Newey (2001), Chao and Swanson (2005)).

Kloek and Mennes (1960) proposed a practical solution to solve the shortcomings of dealing with a large number of (possibly weak) instruments. The idea is to use a reduced number of principal components of the original set of instruments as the instrumental variables in the estimation. We follow a similar approach.

---

4Weather has also been shown to impact productivity in other industries, including agriculture (Schlenker and Roberts (2006), Deschenes and Greenstone (2007)) and back-office processes in banking (Lee et al. (2012)).
The thirteen weather variables were reduced to five principal components. By capturing more than fifty percent of the variance on the original variables, the components obtained contain a good portion of the information in our instruments. The OLS regression of equation (3) shows that the five principal components coefficients are significant with an average p-value for the five factors of less than $10^{-4}$. In addition, to validate the strength of our instrument, we observe that both the R-squared (0.9) of this regression and the F-test of join significance of the instruments (p-value less than $10^{-4}$) meet the standards to rule out weak instruments. For robustness, we also estimated our model using all of the the original weather variables as instruments in $W$, and all the main results continued to hold. But the estimation with the five principal components is more efficient (i.e. smaller standard errors), so we use those as our main results.

Although plant productivity is affected during the same week of a weather incident (as reported in Cachon et al. (2011)), the impact on dealership inventory is lagged due to delivery lead-times. We used a one-week lag based on anecdotal evidence reporting one-week delivery lead-times, but we also tested other specifications and obtained similar results.\footnote{One specification assumes there is no lead-time, hence contemporaneous weather is included. Another specification assumes the lead-time for a vehicle to a dealer depending on the distance between the dealer and the plant where the model was manufactured: we assume a vehicle arrives within the week (zero lag) when the plant is within 600 miles of the dealer; the vehicle arrives within one week when the plant is between 600 and 1200 miles from the dealer; and the vehicle arrives within two weeks when the plant is 1200 miles or more from the dealer.}

The use of the weather as instrumental variables has also been recently used in the economics and political science literature to analyze the impact of poverty on crime (Mehlum et al. (2006)), the effect of riots in property values (Collins and Margo (2007)), the influence of protests on the growth of political parties (Madestam et al. (2011)) and the relationship between economic shocks and civil conflict (Miguel et al. (2004)). Our work is the first empirical study to use weather instruments as supply shocks in manufacturing.

\textbf{An alternative estimation approach of the overall effect of inventory}

Proposition 1 establishes sufficient conditions to estimate the system of equations (1)-(3) consistently. This requires assumptions about the covariance structure of the error terms $\{\varepsilon_g\}_{g=S,V,I}$. However, it is possible to estimate the overall effect of inventory on sales – which corresponds to the direct effect $\beta_{IS}$ plus the indirect effect through variety, $\beta_{VS}\beta_{IV}$ (see figure 1) – under weaker assumptions. To see this, replace Variety from equation (2) into equation (1):

\[ Sales = (\beta_{IS} + \beta_{VS}\beta_{IV})Inventory + \gamma_s'Z + \varepsilon_s', \]  

(5)
where $\gamma'_S = \beta_{VS} \gamma_V$ and $\varepsilon'_S = \varepsilon_S + \beta_{VS} \varepsilon_V$. Under the exogeneity assumption $E(\varepsilon_g | Z, W) = 0$, $g \in \{S, V, I\}$, the coefficient $\beta'_{IS} \equiv \beta_{IS} + \beta_{VS} \beta_{IS}$ can be estimated via instrumental variables, instrumenting Inventory with the weather variables $W$. This provides an alternative estimate of the overall effect inventory on sales without making assumptions about the covariance structure of the error terms $(\varepsilon_S, \varepsilon_V, \varepsilon_I)$. The drawback of this approach is that it doesn’t identify separately the effect of inventory and variety on sales. In particular, this precludes analyzing the allocation strategies described in Section 6.

5 Results

Table 4 reports the main estimation results. Column (1) shows the estimates of equation (1), instrumenting the endogenous variables inventory and variety (as defined in equation (4), which captures variations in vehicles across sub-models). The estimates suggest that the direct effect of inventory ($\beta_{IS}$ in Figure (1)) is negative and statistically significant, but the effect of variety ($\beta_{VS}$) is positive and also statistically significant. This suggests that sales increase if new sub-models are made available to customers, but sales decrease if inventory is added to a sub-model that is already available at the dealership.

Given how inventory is allocated to dealerships in our sample, there is a positive relationship between inventory and variety: the estimated coefficient is $\hat{\beta}_{IV} = 0.0054$, with a standard error $SE(\hat{\beta}_{IV}) = .0001$. This estimate together with the estimated coefficients of equation (1) can be used to estimate the overall average impact of inventory on variety, which is given by $\beta_{IS} + \beta_{VS} \beta_{IV} = -0.013$ (with a standard error of 0.003, obtained from a bootstrap of 400 samples). Hence, our estimates suggest that, given how vehicles were allocated to dealerships in our sample, the overall impact of inventory on sales is negative and statistically significant - adding inventory increases variety, but not by much, so the negative effect of adding inventory to an existing sub-model dominates the sales benefit of the (limited) expanded variety. However, different vehicle allocation policies can give different results. Figure 3 illustrates the overall impact of inventory on sales with the vehicle allocation policy that maximizes the expansion of variety (black line) compared to the allocation policy that expands inventory without increasing the number of sub-models available (dashed line). As is apparent from the figure, whether adding inventory increases or decreases overall sales depends on how vehicles are allocated to dealerships. For example, with a vehicle allocation policy that maximizes variety by adding new sub-models to a model’s inventory, the
overall impact of each additional unit of inventory on sales would be 0.5%. A more precise analysis of alternative vehicle allocation policies is described in Section 6.

Recall that the estimates of column (1) are consistent if the error in equation (2), \( \varepsilon_V \), is uncorrelated with \( \varepsilon_S \) and \( \varepsilon_I \). Although this assumption is reasonable in the context of our application, we can also estimate under less restrictive assumptions the overall impact of inventory on sales (i.e., the direct impact \( \beta_{IS} \) plus the indirect impact through Variety, \( \beta_{IV} \beta_{VS} \)). In particular, we estimated equation (5) using weather variables as instruments for Inventory, reported in column (2) of Table 4. Recall that this approach is consistent even when the error terms \( \varepsilon_S \) and \( \varepsilon_I \) are correlated. The coefficient of inventory – which is an estimate of \( \beta_{IS} + \beta_{IV} \beta_{VS} \) – is -0.014, which is close to our previous estimate based on the coefficients of column (1) (which gave -0.013).\(^6\)

To assess the magnitude of the bias induced by the endogeneity of inventory, we estimated model (5) via Ordinary Least Squares (OLS). As mentioned in Section 4, if inventory is set in anticipation of demand, then \( \varepsilon_S \) and \( \varepsilon_I \) are likely to be positively correlated and therefore the OLS estimate of the inventory coefficient could be biased upward. The OLS estimates reported in column (3) show evidence of this endogeneity bias: in fact, the bias is so severe that the coefficient on inventory changes sign and becomes positive with statistical significance.

Column (4) estimates equation (1) instrumenting Inventory but treating Variety as exogenous. As inventory increases variety, variety is positively correlated with \( \varepsilon_I \) and thereby with \( \varepsilon_S \). Hence, ignoring the endogeneity of variety could also lead to a positive bias on the estimate of coefficient \( \beta_{IV} \), which is what we find: the variety coefficient in column (4) more than doubles that of column (1). This highlights the importance of treating both inventory and variety as endogenous in the estimation.

To repeat, the estimates in columns (1)-(2) of Table 4 and the sensitivity analysis presented on Figure 3 suggest that (i) adding inventory decreases sales if variety is held constant (a scarcity effect), (ii) although increasing inventory expands variety and variety has a positive impact on sales, the overall effect of increasing inventory is negative given the way vehicles are allocated in our sample, and (iii) adding inventory while simultaneously expanding variety can increase sales. Some of the mechanisms discussed earlier are consistent with these findings and several are not.

For example, our findings are consistent with the notion that more variety across sub-models (e.g., trim, engine size) improves the match between consumer preferences and the available inventory.

\(^6\)A non-parametric bootstrapping method (based on 400 re-samples of the original data) gives an average difference of 0.0008, with standard error 0.0002. Although the difference is statistically significant at the 99%, the difference is quite small in practical terms.
thereby increasing the likelihood that a customer makes a purchase. In contrast, the results are not consistent with the notion that more variety creates confusion or higher evaluation costs, thereby reducing demand - in some categories it is possible that the confusion effect is real and sufficiently strong, but with automobiles it appears that consumers are more likely to buy when they have more options to choose from.\footnote{A confusion or evaluation cost effect could apply to the variety within a sub-model as we observe increasing inventory of a sub-model reduces sales. However, most dealers have relatively few vehicles within a sub-model (e.g., generally fewer than 10), so while we cannot rule out this explanation, we do not believe it is the most likely.}

Our findings suggest that dealer pricing or consumer bargaining do not have a strong impact on the relationship between inventory and sales. As shown by Moreno and Terwiesch (2012) and Zettelmeyer et al. (2006) one would expect that a dealer is more likely to offer a better price when the dealer has an above average amount of inventory because the dealer would want inventory to return to a more normal level. We observe that sales decrease as inventory increases (holding variety constant) - if this is to be explained by pricing, then one needs to be willing to assume that dealers increase their prices when they have more inventory. We also observe that sales increase as variety increases, holding inventory constant - if this is to be explained by pricing, then dealers would have to decrease their prices when they offer more selection but increase their prices when they offer less, which also does not seem plausible.\footnote{See Gilbert and Cvsa (2003) for a detailed discussion on manufacturer-dealer strategic interactions related to pricing.} Similarly, our estimates cannot simply be explained by a stockout effect - if adding inventory prevents stockouts, then coefficient $\beta_{IS}$ should be positive, not negative.

Although our results indicate the presence of a scarcity effect, they are not consistent with all mechanisms that lead to a scarcity effect. For example, a scarcity effect can occur if consumers infer that ample inventory is a signal that a car is not popular, possibly due to poor design or quality. For this to explain our data, the inventory signal would have to be at the sub-model level rather than at the model level - a consumer would have to believe that ample inventory of two-door Malibus is a bad signal for two-door Malibus, but the overall number of Malibus is not a negative signal. While we cannot rule this out, it does not seem plausible. We suspect that a consumer would infer quality, popularity and design based on the total inventory of a model rather than based on the inventory of each of the various sub-models. If that is the case, then inferences of popularity cannot explain the negative relationship between sales and inventory, controlling for variety.

The scarcity effect we observe is consistent with the notion that inventory influences consumer search. Consumers are likely to desire a particular sub-model. If there is only one unit available of
their desired sub-model, then they may discontinue their search for a new vehicle and purchase the vehicle. However, if the dealer has several units that fit the consumer’s preference, the consumer may continue her search, feeling confident that if she does not find a better match, she can return to the dealership. If the consumer continues her search, then at the very least it delays the sale, but worse, it risks losing the sale - the consumer might discover a better match at another dealership. Thus, we find evidence that low inventory reduces consumer procrastination and motivates an immediate sale.

A similar effect was found by Soysal and Krishnamurthi (2012) in the apparel industry. In that industry, markdowns during the season can be anticipated by customers, which provides them incentives to delay purchases to pay lower prices. Reducing inventories increases the risk of not finding the product later in the season, thereby lowering the customer’s incentive to delay the purchase.\textsuperscript{9}

The remaining of this section considers several robustness checks. In particular, we consider competition among dealers, different specifications for variety and transfers between dealers.

**Competition among dealers**

The dealers in our sample face different levels of competition from GM and non-GM dealerships. As mentioned earlier, the dealer-model fixed effects included in our main specification account for the average competition intensity for a particular model at a dealer. However, the inventory level for a model could vary across dealers from one week to another and this variation may be known to the dealers. To explain our results, low inventory at dealer A in a market would need to be correlated with reduced competition from the other dealerships in the same market, thereby allowing dealer A to increase his sales. Although we do not view this as likely, to remove the effect of competition we estimate our main model with a subsample of dealers that do not face competition in their local market. Based on empirical work defining the relevant market for a dealership (Albuquerque and Bronnenberg (2012)), we defined a sub-sample of dealers with no competing GM dealer (of any brand) within a 15 mile radius. The analysis with this sub-sample is reported on the first column of Table 5. This result is consistent with the results obtained with the complete sample and suggests that our main results are not confounded by the impact of competition patterns between GM dealers.

\textsuperscript{9}See Lazarev (2011) and Li et al. (2011) for empirical studies in the airline industry that show forward-looking customer behavior; these studies, however, do not account for availability risk.
Alternative specifications for variety

We replicated the analysis described in the previous section with two alternative measures of variety: (i) $\text{AvailVar}_{ijt}$, the total number of different sub-models carried by a dealer in each week (instead of the relative measure of variety considered before); and (ii) the logarithm of $\text{AvailVar}_{ijt}$. The results for these two alternative specifications are reported in columns (2) and (3) of table 5. First, we note that the coefficient of inventory barely changes (compared to that of Table 4 column (1)). Second, the coefficient of variety is positive and significant in both specifications.

Transfers between dealers

If a dealer lacks a sub-model that a consumer wants, the dealer can try to convince the consumer to purchase a different sub-model or the dealer can try to find the desired sub-model at a nearby dealer. If the desired vehicle is found at another dealer, a transfer can occur between the two dealerships if they can agree to the transfer. In many cases this transfer involves a swap of vehicles rather than an exchange of cash - the requesting dealer has a customer that will purchase the donating dealer’s vehicle, and in return, the requesting dealer offers one of its vehicles to the donating dealer, who may not have an customer at the time of the swap willing to purchase the vehicle it receives. In our sample, 12.9% of total sales are for vehicles that were transferred from one dealership to another.

The use of transfers may provide an alternative explanation for our results - when a dealer has fewer vehicles, it compensates by making more frequent use of transfers and thereby increase sales. Through this mechanism, lower inventory does not induce higher sales because of influencing consumer search or preferences, but rather by influencing dealer behavior (i.e., they do more transfers). This alternative mechanism would imply a negative relationship between inventory and transfer sales. Since we observe transfer sales, we tested this hypothesis by estimating a regression with the log of transfers as the dependent variable and inventory as an independent variable, plus the same set of controls $Z$ included in 1. The OLS estimates of this regression does not find a significant relationship between transfers sales and the inventory of a model. This suggests that the scarcity effect that we estimate is not driven by the additional use of transfers when inventory is low.

6 The impact of inventory allocation

Our empirical estimation reveals that adding inventory to a dealer is only beneficial if the added vehicle expands the dealer’s set of sub-models - increasing the inventory of a particular sub-model
actually lowers sales. This sections explores the potential sales benefit of using this result to better allocate vehicles to dealers. We take three approaches. The first estimates the potential sales improvement from reallocating the existing vehicles among the dealers in a small local area. The second is similar, except it allows for some reduction in the total number of vehicles. The third considers only the incoming vehicles to a larger region (e.g., a state) and attempts to maximize sales by allocating those vehicles to the dealers in the area while leaving the dealers’ existing inventory intact. Note that the three approaches take production volume and variety as given and focus in improving the allocation of the existing pool of vehicles to dealerships.\(^\text{10}\)

Given the size of our data-set (1289 dealers, 30 weeks, etc) we focus our analysis on a particular week (the week with the median number of total cars) and the ten most popular models. These models represent approximately sixty percent of the weekly sales across all the GM models in our sample: Cobalt, Equinox, G6, HHR, Impala, Suburban, Tahoe, TrailBlazer, Saturn, VUE, and Yukon.

### 6.1 Local reallocation among dealers

The analysis in this section partitions dealers into small local markets. For each model we know each dealer’s available inventory in our chosen week. Some dealers may have multiple units within a sub-model and other dealers within the same local market might not have any vehicles of that sub-model. Hence, based on our results, both dealers could benefit from a vehicle transfer - moving a vehicle from the dealer with multiple units to the dealer with no units increases sales at both dealers. Thus, we evaluate for each model the total sales gain across all markets that could be achieved by intelligently transferring vehicles so as to maximize the variety each dealer offers and to minimize the duplication of units within sub-models. We do not model the cost of actually transferring these vehicles - any sales improvement from reallocation would have to be compared with the cost of achieving the better balance of variety across dealers.

We group dealers as part of the same local market if they are in the same core based statistical area (CBSA) - a CBSA is a U.S. geographic area defined by the Office of Management and Budget based around an urban center of at least 10,000 people and adjacent areas that are socioeconomically tied to the urban center by commuting. We consider vehicle swaps only between dealers in the same CBSA. Hence, the total inventory within each CBSA remains constant. In addition, we require that each dealer’s total inventory remains constant - each dealer that gains a vehicle must also give up

\(^{10}\text{See Fisher and Ittner (1999) for an empirical study on the effect of variety on automobile productivity.}\)
a vehicle. The average number of dealers per CBSA is 4.5; the minimum is 2 and the maximum is 28.

To formulate the problem as a math program, let \( i = 1 \ldots n \) index the dealers within the CBSA and \( k = 1 \ldots m_j \) index the sub-models of model \( j \). The problem can be formulated as choosing \( Q_{ijk} \) – the number of vehicles at dealer \( i \) of model \( j \) and sub-model \( k \) after reallocating vehicles among the dealers within the CBSA – in order to solve the following non-linear integer optimization problem:

\[
\max_{Q_{ijk}} \left[ \sum_{i=1}^{n} \exp \left( \delta_{ij} + \hat{\beta}_{IS} \sum_{k=1}^{m_j} Q_{ijk} + \hat{\beta}_{VS} \cdot \text{Variety}_{ij} \right) \right] \tag{6}
\]

subject to the constraints

\[
\sum_{k=1}^{m_j} Q_{ijk} = \sum_{k=1}^{m_j} I_{ijk} \quad \forall i, j \tag{7}
\]

\[
\sum_{i=1}^{n} Q_{ijk} = \sum_{i=1}^{n} I_{ijk} \quad \forall j, k \tag{8}
\]

\[
Q_{ijk} \leq T_{ij} \quad \forall i, j, k \tag{9}
\]

\[
\text{Variety}_{ij} = \frac{\sum_{k=1}^{m_j} I_{ijk} (Q_{ijk} \geq 1)}{m_j} \quad \forall i, j \tag{10}
\]

\[
Q_{ijk} \in \{0, 1, 2, \ldots \infty\} \quad \forall i, j, k \tag{11}
\]

where \( I_{ijk} \) is dealer \( i \)'s observed inventory of model \( j \) and sub-model \( k \) (i.e., before the reallocation). The parameters \( \hat{\beta}_{IS} \) and \( \hat{\beta}_{VS} \) are the estimated coefficients of Inventory and Variety, respectively, and \( \delta_{ij} \) is the estimated model-dealer fixed-effect (the estimates correspond to the model reported in Table 4, column (1)). \( T_{ij} \equiv \max_k \{I_{ijk}\} \) is the maximum number of vehicles dealer \( i \) carried across all the sub-models \( k \) of model \( j \) during the selected week.

Constraint (7) ensures that dealer \( i \)'s inventory of model \( j \) remains the same after the allocation. Constraint (8) ensures that the reallocation does not change the total inventory within the CBSA of model \( j \). Thus, these two constraints ensure that each dealer that gains a vehicle must also give up a vehicle of the same model. Constraint (9) ensures that at the end of the swaps the maximum number of units of a particular sub model \( k \) at dealer \( i \) is less than or equal to the maximum number of units of any sub model \( k \) that dealer \( i \) was carrying at the beginning of the swaps. Given constraint (7), the objective is to maximize \( \text{Variety}_{ij} \) (10) while keeping each dealership’s inventory
constant. As the objective function in non-linear in $\textit{Variety}_{ij}$, we used a non-linear optimization solver to find the optimal solution.\textsuperscript{11}

The first column on Table 6 shows the optimal solution, measured by the expected sales improvement for each car model relative to the expected sales with the actual allocation observed in the data (that is, with $Q_{ijk} = I_{ijk}$). We find that on average, exchanging inventory among dealers within a CBSA with the objective of maximizing each dealer’s offered variety yields a weighted average sales gain of 1.7%. (The results for each model are also weighted averages. Hence, this is the overall increase across all dealerships due to the inventory exchanges.)

6.2 Local reallocation among dealers with inventory reduction

Following on the analysis of subsection 6.1, in addition to the vehicles swaps among nearby dealers (within the same CBSA) we allow dealers to reduce their inventory of each model by up to ten percent. To do so, we solved a mathematical program similar to (6)-(11), but constraints (7) and (8) are replaced by:

\[
\begin{align*}
&\sum_{k=1}^{m_j} Q_{ijk} \geq 0.9 \sum_{k=1}^{m_j} I_{ijk} \\
&\sum_{i=1}^{n} Q_{ijk} \leq \sum_{i=1}^{n} I_{ijk}
\end{align*}
\]

The first requires that a dealer’s inventory of a model cannot be reduced by more than 10% and the second that the total inventory of a sub-model in a CBSA cannot increase (but it can decrease).

The second and third columns on Table 6 shows the solution to this math program. The second column shows the average inventory reduction at each dealer for each model. Column 3 presents the results for the average potential sales improvement for each model. We find that exchanging inventory among dealers within a CBSA while giving the dealers the option to reduce their inventory results in an weighted average inventory reduction of 2.9% and a weighted average sales gain of 5.0%. (Again, the inventory reductions and the sales increases are the overall effect rather than the average percent change across dealers.)

\textsuperscript{11}Since $\hat{\beta}_{IS} < 0$ and $\hat{\beta}_{IS} + \hat{\beta}_{VS} > 0$, the first unit of a sub-model always increases (expected) sales but additional units decrease sales. If constraint (9) were removed, because the objective is convex decreasing in $Q_{ijk}$, the optimal solution would allocate one unit of each sub-model to each dealer (to increase variety) and all the remaining model inventory to the dealer with the smallest sales rate $\delta_{ij}$ (to minimize the sales reduction induced by higher inventories). Constraint (9) prevents this, giving a more reasonable solution to the problem. The model was implemented using GAMS with the KNITRO non-linear solver.
### 6.3 State-wide reallocation of vehicles

Instead of swapping vehicles after they arrive at dealerships, we now consider changing the allocation of vehicles coming from the plants to the dealers. In particular, we estimate the sales gain that can be achieved through smarter allocation of vehicles that arrive to a particular state in a given week. Let \( R \) be the set of dealers in the focal state and \( A_{ijk} \) be the number of vehicles of sub-model \( k \) of model \( j \) arriving to dealer \( i \) in the given week. Hence, the observed dealer inventory \( I_{ijk} \) during this week is the sum of the remaining inventory from last week, denoted \( I_{ijk}^- \), plus the incoming vehicles \( A_{ijk} \). To avoid costly transfers between dealerships, we consider reallocating the total incoming vehicles to the state, \( \sum_{i \in R} A_{ijk} \), without changing the units already in the lots (i.e. keeping \( I_{ijk}^- \) fixed). It is optimal then to allocate first the incoming units of a sub-model \( k \) to a dealer with no inventory of that sub-model, \( I_{ijk}^- = 0 \). Denote \( Y_{ijk} \) the decision variable corresponding to the number of units of sub-model \( k \) (of model \( j \)) to be allocated to dealer \( i \). For a given state and model \( j \), the problem can be stated as the following non-linear integer program:

\[
\max_{Y_{ijk}} \left[ \sum_i^n \exp \left( \delta_{ij} + \hat{\beta}_{IS} \cdot \sum_k^{m_j} Q_{ijk} + \hat{\beta}_{VS} \cdot \text{Variety}_{ij} \right) \right] 
\]

subject to

\[
\sum_{i \in R} Y_{ijk} = \sum_{i \in R} A_{ijk}, \quad \forall j, k
\]  
(13)

\[
\sum_{k=1}^{m_j} Q_{ijk} \leq M_{ij}, \quad \forall i, j
\]  
(14)

\[
Q_{ijk} = I_{ijk}^- + Y_{ijk}, \quad \forall i, j, k
\]  
(15)

\[
\text{Variety}_{ij} = \frac{\sum_{k=1}^{m_j} \mathbb{1}(Q_{ijk} \geq 1)}{m_j}, \quad \forall i, j
\]  
(16)

\[
Y_{ijk} \in \{0, 1, 2, \ldots, \infty\}, \quad \forall i, j, k
\]  
(17)

where \( M_{ij} \) is the maximum number of vehicles dealer \( i \) carried of model \( j \) across our 30 week sample. All the other parameters are defined as before.

Constraint (13) ensures that the state receives the same number of vehicles of model \( j \) and sub-model \( k \) as we observed in our data for the chosen week. Constraint (14) ensures that dealer \( i \)'s inventory of model \( j \) after the assignment is not greater than the maximum number of vehicles of model \( j \) that dealer \( i \) had in any week of our sample. This precludes allocations that result in
some dealers having an unreasonably large amount of inventory. Equation (15) merely states that a dealer’s inventory of a model equals the dealer’s inventory from the previous week, $I_{ijk}^\text{−}$, plus the dealer’s allocation, $Y_{ijk}$.

The second column on Table 6 shows average results for each model in this state-wide allocation problem. On average, we find that routing vehicles to dealers in a state so as to minimize overlap within a dealer’s inventory while maximizing variety across dealers yields an average sales increase of 2.5%.

7 Conclusion

We develop an econometric model to estimate the effect of inventory on sales at U.S. automobile dealerships. Theory is ambiguous with respect to the impact of inventory on sales. There are several mechanisms that lead to a positive relationship between inventory and sales. For example, at a basic level, adding inventory can increase sales by reducing stockouts, or by expanding the variety of sub-models available. However, there are mechanisms that lead to a scarcity effect - a negative relationship between inventory and sales. For instance, adding inventory may encourage consumers to engage in additional search, which may reduce a dealer’s sales (i.e., once a customer leaves the dealership, they may not return).

In our sample, we find that an increase in a dealer’s inventory of a model actually lowers sales. However, it is important to decompose this effect into two parts: increasing inventory of a sub-model does indeed reduce sales, but if the increase in inventory also expands the number of sub-models available, then sales increase. In short, the benefit of expanding the number of available sub-models dominates the negative effect of increasing inventory within a sub-model. This is consistent with two mechanisms relating inventory to demand: (i) expanded variety enables a better fit to consumer preferences, thereby increasing demand, and (ii) too many of the same sub-model encourages consumers to procrastinate the purchase decision, thereby lowering sales. To maximize sales a dealer wants to have one unit of each sub-model (to generate an urgency to “buy now before they are all gone”) while also having as many sub-models available as possible, to cater to the heterogeneous tastes of consumers. As vehicles within the same sub-model can be different (e.g., different color), our results suggest that if there are benefits from expanding variety within a sub-model, those benefits are dominated by the negative effects of increased inventory within a sub-model. This suggests that variety across sub-models (e.g., engine sizes, number of doors) is of
stronger importance to consumers in terms of finding a vehicle that matches their preferences than the variety that exists within a sub-model.

Our findings emphasizes the importance of careful vehicle allocation. Based on our estimates, an allocation policy that focuses on maximizing variety while minimizing duplication of sub-models can increase sales by about 2.5%, without changing the number of vehicles produced or the number of vehicles each dealer carries. Sales can be improved further if the company is willing to somewhat lower total inventory - during our sample period, we estimate that a 5.0% sales increase can be achieved with the combination of variety maximizing allocation and an overall 2.9% reduction in the number of vehicles. In sum, in the auto industry it appears there are substantial advantages from a “lean but not too lean” inventory approach that attempts to maximizes the variety of sub-models available to consumers at dealerships while minimizing the duplication of vehicles of the same sub-model.

A Appendix

Proof Proposition 1

Consider the following system of simultaneous equations (1)-(3). We show that if $E(\varepsilon_g|Z,W) = 0$, for all $g \in \{S,V,I\}$, $E(\varepsilon_S\varepsilon_V) = 0$ and $E(\varepsilon_V\varepsilon_I)$, then all the parameters of the system of equations (1)-(3) are identified.

Proof. To facilitate the exposition of the proof, we replace indexes $(S,V,I)$ by $(1,2,3)$, in that respective order (the indexes correspond to the number of the equation in the system). The reduced form of the system of equations (1)-(3) is denoted by:

\begin{align*}
    Sales &= \pi_1 Z + \psi_1 W + u_1 \\
    Variety &= \pi_2 Z + \psi_2 W + u_2 \\
    Inventory &= \pi_3 Z + \psi_3 W + u_3
\end{align*}

(18) (19) (20)

Because $Z$ and $W$ are exogenous, the coefficients $(\pi_1, \pi_2, \pi_3)$ and $(\psi_1, \psi_2, \psi_3)$ are identified, as well as the covariance matrix of the reduced form error terms $(u_1, u_2, u_3)$, denoted by $\Omega$ (we use $\Omega_{ij}$ to index the elements of the covariance matrix).

The triangular structure of the system (18)-(20) facilitates its inversion into the reduced form system (18)-(20). First, equations (1) and (20) are identical, so $\pi_3 = \gamma_3$, $\psi_3 = \delta_3$ and $\Omega_{33} = \Sigma_{33}$.
Hence, equation (20) alone identifies $\gamma_3$, $\delta_3$ and $\Sigma_{33} = \text{Var}(\varepsilon_3)$. For equation (19) we have:

$$
\pi_2 = \beta_{32}\gamma_3 + \gamma_2 \quad \psi_2 = \beta_{32}\delta_3
$$

so $\beta_{32} = \psi_2/\delta_3$ and $\gamma_2 = \pi_2 - \psi_2\gamma_3$ are also identified. The variance of $u_2$ is given by:

$$
\Omega_{22} = \Sigma_{22} + \beta_{32}^2\Sigma_{33}
$$

which identifies $\text{Var}(\varepsilon_2) = \Sigma_{22}$.

For equation (18):

$$
\begin{align*}
\pi_1 &= \beta_{21}\beta_{32}\gamma_3 + \beta_{21}\gamma_2 + \beta_{31}\gamma_3 + \gamma_1 \\
\psi_1 &= \beta_{21}\beta_{32}\delta_3 + \beta_{31}\delta_3
\end{align*}
$$

(21)

with unknowns $\beta_{21}$, $\beta_{31}$ and $\gamma_1$. Additional identifying equations can be obtained from $\Omega$, the covariance matrix of the reduced form error $u$. By definition, the reduced form errors of equations (19) and (20) imply:

$$
\begin{align*}
&u_2 - \beta_{32}u_3 = \varepsilon_2 \\
&u_1 = \varepsilon_1 + \beta_{21}\varepsilon_2 + (\beta_{21}\beta_{32} + \beta_{31})\varepsilon_3
\end{align*}
$$

Taking covariance between these two equations, assumptions $E(\varepsilon_1\varepsilon_2) = E(\varepsilon_2\varepsilon_3) = 0$ imply:

$$
\Omega_{12} - \beta_{32}\Omega_{13} = \beta_{21}\Sigma_{22}.
$$

This identifies $\beta_{21} = (\Omega_{12} - \beta_{32}\Omega_{13})/\Sigma_{22}$, which replacing in (21) identifies the remaining parameters $\beta_{31}$ and $\gamma_1$. \qed
References


David Barkholz.


BloombergTV. 2012. Warm weather a boost to retail sales?


Li, Jun, Serguei Netessine. 2012. Who are my competitors?-let the customer decide.


Figure 1: Relationship between Sales, Inventory and Variety

Figure 2: Dealer and plant locations in the selected sample
Expected weekly sales as a function of inventory with the vehicle allocation policy that maximizes the expansion of variety (black line) and the allocation policy that expands inventory without increasing the number of sub-models available (dashed line), for a dealer that starts with 3 vehicles of a particular model.

Figure 3: Sensitivity Analysis

<table>
<thead>
<tr>
<th>State</th>
<th>Number of Dealers</th>
</tr>
</thead>
<tbody>
<tr>
<td>California</td>
<td>355</td>
</tr>
<tr>
<td>Colorado</td>
<td>67</td>
</tr>
<tr>
<td>Florida</td>
<td>237</td>
</tr>
<tr>
<td>Main</td>
<td>31</td>
</tr>
<tr>
<td>Nebraska</td>
<td>50</td>
</tr>
<tr>
<td>Texas</td>
<td>366</td>
</tr>
<tr>
<td>Wisconsin</td>
<td>183</td>
</tr>
<tr>
<td>Total</td>
<td>1,289</td>
</tr>
</tbody>
</table>

Table 1: Dealers by state in the selected sample
<table>
<thead>
<tr>
<th>MarketVar</th>
<th>AvailVar</th>
</tr>
</thead>
<tbody>
<tr>
<td>Cobalt</td>
<td>18</td>
</tr>
<tr>
<td>Equinox</td>
<td>4</td>
</tr>
<tr>
<td>G6</td>
<td>37</td>
</tr>
<tr>
<td>HHR</td>
<td>4</td>
</tr>
<tr>
<td>Impala</td>
<td>10</td>
</tr>
<tr>
<td>Suburban</td>
<td>18</td>
</tr>
<tr>
<td>Tahoe</td>
<td>13</td>
</tr>
<tr>
<td>TrailBlazer</td>
<td>10</td>
</tr>
<tr>
<td>Saturn VUE</td>
<td>5</td>
</tr>
<tr>
<td>Yukon</td>
<td>30</td>
</tr>
<tr>
<td>Average</td>
<td>14.9</td>
</tr>
</tbody>
</table>

*MarketVar* is the number of sub-models that could be produced for the model.

*AvailVar* is the number of sub-models with at least one unit during a particular week.

Table 2: Model Variety for the top ten selling models

<table>
<thead>
<tr>
<th>Variable</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Wind</strong></td>
<td>Number of days in which a wind advisory is issued by the National Weather Service Forecast Office. A wind advisory is issued when maximum wind speed exceeds a threshold for the area which is typical in excess of 40 miles per hour.</td>
</tr>
<tr>
<td><strong>Cloud</strong></td>
<td>Average cloud cover during the week (0 = no clouds; 8 = sky completely covered).</td>
</tr>
<tr>
<td><strong>Fog 1</strong></td>
<td>Weeks with 1 days with fog during the week.</td>
</tr>
<tr>
<td><strong>Fog 2-3</strong></td>
<td>Weeks with 2 or 3 days of fog during the week.</td>
</tr>
<tr>
<td><strong>Fog 4-7</strong></td>
<td>Weeks with more than 3 days of fog during the week.</td>
</tr>
<tr>
<td><strong>Rain 1-2</strong></td>
<td>Weeks with 1 or 2 days of rain during the week.</td>
</tr>
<tr>
<td><strong>Rain 3-5</strong></td>
<td>Weeks with 3 to 5 days of rain during the week.</td>
</tr>
<tr>
<td><strong>Rain &gt;5</strong></td>
<td>Weeks with more than 5 days of rain during the week.</td>
</tr>
<tr>
<td><strong>Snow 1</strong></td>
<td>Weeks with 1 day of snow during the week.</td>
</tr>
<tr>
<td><strong>Snow 2-4</strong></td>
<td>Weeks with 2 to 4 days of snow during the week.</td>
</tr>
<tr>
<td><strong>Snow &gt;4</strong></td>
<td>Weeks with more than 4 days of snow during the week.</td>
</tr>
<tr>
<td><strong>High Temp 1</strong></td>
<td>Weeks with 1 day of high temperature, above 90 degrees Fahrenheit, during the week.</td>
</tr>
<tr>
<td><strong>High Temp 2-5</strong></td>
<td>Weeks with 2 to 5 days of high temperature, above 90 degrees Fahrenheit, during the week.</td>
</tr>
<tr>
<td><strong>High Temp 6-7</strong></td>
<td>Weeks with more than 5 days of high temperature, above 90 degrees Fahrenheit, during the week.</td>
</tr>
</tbody>
</table>

Table 3: Weather variables included in the empirical study
### Table 4: Main Model Results

<table>
<thead>
<tr>
<th></th>
<th>(1)</th>
<th>(2)</th>
<th>(3)</th>
<th>(4)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Inventory</td>
<td>-0.0147***</td>
<td>-0.0140***</td>
<td>0.0130***</td>
<td>-0.0189***</td>
</tr>
<tr>
<td></td>
<td>(0.0030)</td>
<td>(0.0030)</td>
<td>(0.0002)</td>
<td>(0.0034)</td>
</tr>
<tr>
<td>Variety</td>
<td>0.2958***</td>
<td></td>
<td></td>
<td>0.7223***</td>
</tr>
<tr>
<td></td>
<td>(0.0094)</td>
<td></td>
<td></td>
<td>(0.0487)</td>
</tr>
<tr>
<td>Fixed Effects</td>
<td>YES</td>
<td>YES</td>
<td>YES</td>
<td>YES</td>
</tr>
<tr>
<td>Week - Season</td>
<td>YES</td>
<td>YES</td>
<td>YES</td>
<td>YES</td>
</tr>
<tr>
<td>Dealer’s Local Weather</td>
<td>YES</td>
<td>YES</td>
<td>YES</td>
<td>YES</td>
</tr>
<tr>
<td>Total number of obs.</td>
<td>293776</td>
<td>293776</td>
<td>293776</td>
<td>293776</td>
</tr>
<tr>
<td>Number of dealer-models</td>
<td>12969</td>
<td>12969</td>
<td>12969</td>
<td>12969</td>
</tr>
</tbody>
</table>

Standard errors in parentheses
* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

(1) Main estimation results where the estimates are obtained instrumenting the endogenous inventory and variety.
(2) Estimation results for the overall impact of inventory on sales instrumenting the endogenous inventory.
(3) Estimation results for the overall impact of inventory on sales without instrumenting the endogenous inventory.
(4) Estimation results where the estimates are obtained instrumenting the endogenous inventory and without instrumenting variety.

### Table 5: Robustness Analysis

<table>
<thead>
<tr>
<th></th>
<th>(1)</th>
<th>(2)</th>
<th>(3)</th>
<th>(4)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Inventory</td>
<td>-0.0171*</td>
<td>-0.0151***</td>
<td>-0.0156***</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.0074)</td>
<td>(0.0030)</td>
<td>(0.0030)</td>
<td></td>
</tr>
<tr>
<td>Variety</td>
<td>0.2847***</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.0123)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>AvailVar</td>
<td></td>
<td>0.0421***</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>(0.0012)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Log(AvailVar)</td>
<td></td>
<td></td>
<td></td>
<td>0.1310***</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>(0.0034)</td>
</tr>
<tr>
<td>Fixed Effects</td>
<td>YES</td>
<td>YES</td>
<td>YES</td>
<td></td>
</tr>
<tr>
<td>Week - Season</td>
<td>YES</td>
<td>YES</td>
<td>YES</td>
<td></td>
</tr>
<tr>
<td>Dealer’s Local Weather</td>
<td>YES</td>
<td>YES</td>
<td>YES</td>
<td></td>
</tr>
<tr>
<td>Total number of obs.</td>
<td>150619</td>
<td>274399</td>
<td>274399</td>
<td></td>
</tr>
<tr>
<td>Number of dealer-models</td>
<td>6803</td>
<td>11879</td>
<td>11879</td>
<td></td>
</tr>
</tbody>
</table>

Standard errors in parentheses
* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

(1) Analysis excluding dealers that have another GM dealer within a 15 miles radius.
(2) Analysis including variety as a count of different sub-models.
(3) Analysis including the logarithm of variety as a count of different sub-models.
<table>
<thead>
<tr>
<th>Model</th>
<th>CBSA reallocation</th>
<th>State reallocation</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>No inventory</td>
<td>10% Inventory</td>
</tr>
<tr>
<td></td>
<td>reduction allowed</td>
<td>reduction allowed</td>
</tr>
<tr>
<td></td>
<td>Sales increase</td>
<td>Inventory reduction</td>
</tr>
<tr>
<td>Cobalt</td>
<td>4.1%</td>
<td>1.7%</td>
</tr>
<tr>
<td>Equinox</td>
<td>2.3%</td>
<td>2.3%</td>
</tr>
<tr>
<td>G6</td>
<td>3.3%</td>
<td>0.8%</td>
</tr>
<tr>
<td>HHR</td>
<td>3.4%</td>
<td>3.0%</td>
</tr>
<tr>
<td>Impala</td>
<td>3.5%</td>
<td>2.2%</td>
</tr>
<tr>
<td>Suburban</td>
<td>2.4%</td>
<td>2.8%</td>
</tr>
<tr>
<td>Tahoe</td>
<td>2.3%</td>
<td>5.0%</td>
</tr>
<tr>
<td>TrailBlazer</td>
<td>1.4%</td>
<td>0.7%</td>
</tr>
<tr>
<td>Saturn VUE</td>
<td>0.5%</td>
<td>3.4%</td>
</tr>
<tr>
<td>Yukon</td>
<td>2.1%</td>
<td>3.7%</td>
</tr>
<tr>
<td>Weighted average</td>
<td>2.5%</td>
<td>2.9%</td>
</tr>
</tbody>
</table>

Table 6: The impact of inventory allocation