EMPIRICAL STUDIES IN RETAIL OPERATIONS

Juan Santiago Gallino

A DISSERTATION

in

Operations and Information Management

For the Graduate Group in Managerial Science and Applied Economics

Presented to the Faculties of the University of Pennsylvania

in

Partial Fulfillment of the Requirements for the

Degree of Doctor of Philosophy

2013

Supervisor of Dissertation

Gérard P. Cachon, Professor of Operations and Information Management

Co-Supervisor of Dissertation

Marshall Fisher, Professor of Operations and Information Management

Graduate Group Chairperson

Eric Bradlow, Professor of Marketing, Statistics, and Education

Dissertation Committee

Gérard P. Cachon, Professor of Operations and Information Management
Marshall Fisher, Professor of Operations and Information Management
Christian Terwiesch, Professor of Operations and Information Management
Marcelo Olivares, Associate Professor of Decisions, Risk and Operations

Para Flori.

ACKNOWLEDGEMENT

This has been an amazing journey and at the end I only have thoughts of joy and gratitude.

I am very thankful to my advisors Gérard Cachon and Marshall Fisher. Each one, with their own style, gave me invaluable support during the course of my PhD studies. They were a constant example of hard work, passion for research, and respect. Gerard has been wonderful mentor and has had an enormous influence on my professional development. Marshall made a constant effort to help me remember that academic life should be fun and relevant. They both have been a constant source of support at every step and I would not have been able to finish this work without their help. Gerard's and Marshal's contributions in my training as a scholar and, more importantly, as a person have been inestimable, and I am profoundly grateful to them.

Christian Terwiesch has been a wonderful friend during the program and it has been an immense pleasure to have him as the chair of my committee. His door was always open to give me valuable advice and offer help since the first day when I started the program. Thank you!

I am also very thankful to Marcelo Olivares, his generosity to share his knowledge and experience has been absolutely invaluable. His support and advice was always available when I needed.

These four members of my dissertation committee have been extremely generous with their time and advice. I feel very fortunate to have had the chance to work with such exceptional individuals. They have all been a great source of inspiration and one cannot imagine of a better committee.

The faculty and staff of the Operations and Information Management Department have provided a great environment, and I am thankful to all of them, but in particular to Noah Gans, who coordinated the PhD program with attention the details and care.

I have been extremely blessed to share the PhD experience with a fabulous group of peers. I am thankful to all of them for making this experience much more enjoyable. I will always treasure my time at Wharton and the lifelong friends I have made. In particular I would like to thank Alessandro, Bob and Toni, who I can count among my dearest friends.

I am also indebted to my friends at IAE in Argentina who always encouraged me to pursue this journey.

I would also like to thank my parents Alejandro and Graciela, my brother Manuel and my sister Paz for their unconditional love and support. My in-laws were also a source of love and care and the journey would not have been the same without them.

For some weird reason the most important paragraph is always at the end; maybe because it is what we want everyone to remember.

My wife Flori has generously accompanied me in to this crazy journey. These last five years have been for me a constant and silent lesson of what love and generosity really are all about. Flori gave up too many things to make my PhD dream come true. She has been a constant source of support, joy, and inspiration. She is my wife, a gift I do not deserve, and I am so thankful for that. I hope that somehow one day I am able to show her a small portion of the love she gave me during this journey. Our sons Bautista, Pedro, Mateo, Juana and Felipe brought joy and fun to this adventure. Without them the PhD program would have been a lot harder, they remind me every morning where the north was. Without Flori's company and love, this journey would have not been possible. Today a new journey begins without Flori it would not be fun. I dedicate my work to her.

ABSTRACT

EMPIRICAL STUDIES IN RETAIL OPERATIONS

Juan Santiago Gallino

Gérard P. Cachon

Marshall Fisher

This dissertation contains three essays. The first essay, entitled *"Does Inventory Increase Sales? The Billboard and Scarcity Effects in U.S. Automobile Dealerships"* looks into the relationship between inventory and demand beyond the obvious stockout effect. Inventory might signal a popular, and therefore a desirable, product, thereby increasing sale. Or, inventory might encourage a consumer to continue her search, thereby decreasing sales. In this paper we seek to identify these effects in U.S. automobile sales. Our primary research challenge is the endogenous relationship between inventories and 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. Based on this insight, given a fixed set of cars, they should be allocated among a group of dealers so as to maximize each dealer's variety. The second essay, entitled *"Severe Weather and Automobile Assembly Productivity"*, is related to the first one in that presents a detail analysis of the exogenous shock presented there: The weather impact on vehicles assembly lines. It is apparent that severe weather should hamper the productivity of work that occurs outside. But what is the effect of extreme rain, snow, heat and wind on work that occurs indoors, such as the production of automobiles? Using weekly production data from 64 automobile

plants in the United States over a ten-year period, we find that adverse weather conditions lead to a significant reduction in production. Across our sample of plants, severe weather reduces production on average by 1.5%. While it is possible that plants are able to recover these losses at some later date, we do not find evidence that recovery occurs in the week after the event. Our findings are useful both for assessing the potential productivity shock associated with inclement weather as well as guiding managers on where to locate a new production facility. The third essay, entitled *"Integration of Online and Offline Channels in Retail: The Impact of Sharing Reliable Inventory Availability Information"*. In this essay we focus the attention on the impact of inventory information disclosure. Increasingly, retailers are integrating their offline and online channels to reduce costs or to improve the value proposition they make to their customers. Using a proprietary dataset, we analyze the impact of the implementation of a buy-online-pickup-in-store project. Contrary to our expectations, the implementation of this project is associated with a reduction in online sales and an increase in store sales and traffic. We interpret the results in light of recent operations management literature that analyzes the impact of sharing inventory availability information online. The implementation of a buy-online-pickup-in-store project provides an exogenous shock to the verifiability of the inventory information that the firm shows to their customers. Our analysis illustrates the challenges of drawing conclusions about complex interventions using single channel data.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF ILLUSTRATIONS

CHAPTER 1 : Introduction

This dissertation contains three related essays that contribute to the empirical literature in operations and information management. These papers try to better understand the relationship between inventory information and demand. At the same time the papers present different approaches to deal with the challenge of inventory and demand endogeneity in their empirical estimation approach.

The first essay, entitled *"Does Inventory Increase Sales? The Billboard and Scarcity Effects in U.S. Automobile Dealerships"* looks into the relationship between inventory and demand beyond the obvious stockout effect.

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 (Automotive-News (2010)). Were those complaints justified?

Clearly sales could fall if a dealer does not have any inventory - it is hard to sell a car if there is no car to sell in markets, like the U.S., where customers are accustomed to purchase directly from units on the lot rather than a make-to-order process as is more common in Europe. But beyond this *stockout effect*, does carrying more or less inventory influence sales? Traditional inventory theory assumes the answer is "no": demand is generally taken to be independent of inventory, so while sales varies in inventory (due to the stockout effect), demand does not. But in some product categories, including automobiles, there is reason to believe that demand may indeed depend in part on the amount of inventory carried, thereby creating a link between inventory and sales beyond the stockout effect. 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 demand must be slow because the car is not popular, and it must not be popular for a reason, so the customer becomes less likely to purchase. In general, we use the label "*billboard effect*" for any mechanism that assigns a positive relationship between inventory and demand, and "*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 lead either to a billboard or to a scarcity effect, estimating the relationship between inventory and sales is complex primarily because it is reasonable to believe that inventories are endogenously chosen. 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.

The second essay, entitled *"Severe Weather and Automobile Assembly Productivity"*, is related to the first one in that presents a detail analysis of the exogenous shock presented there: The weather impact on vehicles assembly lines.

It is well known that there is a relationship between climate and economic activity. For example, not only are hot countries poorer, temperature even explains variation in economic output within countries (Dell et al., 2009). It is intuitive that climate can impact outdoor activities like agriculture, forestry, construction and tourism. Less clear is the impact on "climate-insensitive" sectors such as manufacturing and services (Nordhaud, 2006).

In this paper we study the relationship between severe weather and weekly automobile production at 64 facilities within the United States over a ten-year period. Although automobiles are made indoors, there are several mechanisms through which bad weather at a plant could influence production. For example, high winds, icy roads or heavy precipitation could cause delays in in-bound delivery of parts from suppliers, possibly due to additional traffic congestion, accidents or cancelled shipments. (See Brodsky and Hakkert (1988); Golob and Recker (2003) for data on precipitation and traffic accidents.). Finished vehicles might be damaged during periods of high wind or hail once they exit the plant. In addition, if a plant operates in a "just-in-time" fashion with relatively little buffer stock of parts, the plant may need to delay the start of a shift or cancel a shift altogether due to the absence of needed parts. The same concern applies to "in-bound" employees – production could be curtailed if workers are unable to (or choose not to) travel to the plant. Finally, even if all of the workers and parts are available, it is possible that bad weather could influence employee productivity. For example, with extreme heat conditions outside, even if the plant has a cooling system, it is possible that the indoor temperature rises to a level that slows down the manual labor associated with automobile production. Alternatively, bad weather outside could influence the affect of employees which in turn may lower their productivity.

In short, it seems reasonable to conclude that weather could influence seemingly sheltered indoor economic activity.

For our study it is safe (we believe) to assume that production does not cause changes in the weather - whether a plant produces more or fewer cars in a week is unlikely to influence its local weather in that week. Of greater concern is whether weather exerts a causal influence on production - are there omitted variables that could lead to an endogeneity bias? For example, maybe automobile production is seasonal for reasons unrelated to local weather. If production seasonality is correlated with a plant's weather (e.g., if fewer cars are made in the summer because demand across the country is lower during the summer), then local weather may only be a proxy for this seasonality. To address this issue we take advantage of the panel structure of our data to include a number of controls: product introduction and ramp-down dummies to account for the possibility that vehicles are introduced at certain times of the year (and their obvious influence on the level of production); plant fixed effects to account for idiosyncratic plant characteristics associated with seasonality; planned shifts to account for known variations in production; weekly dummies to account for national variations in demand, monthly segment dummies (e.g., cars, vans, etc.) to account for segment specific demand seasonality; regional year-month dummies to account for regional differences in weather fluctuations and the possibility that the influence of weather varies by region; and seasonal average weather measures for each plant (e.g., average amount of rain in week t for plant i). In sum, given our extensive set of controls, we believe we have identified a causal impact of severe weather on production.

We also find that weather has a substantial economic impact on automobile production. For example, we estimate that for an average plant, within a week, six or more days with a high temperature of 90^oF or one additional day of heavy winds reduces that week's production by approximately 8%, and six or more days of rain within a week reduces production relative to no rain by 6%. Furthermore, we find that average weekly production losses due to weather events (snow, rain, heat and wind) ranges from a low of 0.5% (Princeton, IN) to a high of 3% (Montgomery, AL), with an overall average of 1.5%. Hence, even though the severe events we identify are not common (e.g., there is only about 2.5 high wind days per year per plant), they are sufficiently common that their collective effect is meaningful.

Our data are suitable for measuring the short term impact of weather on production. An interesting question is how do plants react to the productions shocks we observe? On one extreme, the production could be "lost forever", while at the other extreme, the plant may fully recover the lost production in the same week the weather event occurs. Even if they are able to recover some production in the same week, we find the net impact of severe weather on a week's production to be negative. In addition, we do not find evidence that they are able to recover in the following week - plants are not more likely to schedule overtime in a week that follows bad weather, nor is production higher in weeks following bad weather (all else being equal). Nevertheless, we cannot rule out the possibility that plants recover the production at some point in the future. However, even if they were to fully recover at some point, at the very least, such recovery increases the variability of production (which is costly) and may lead to delayed shipments and stockouts.

Our results could be useful in several ways. First, they are related to the issue of climate change. While there is low confidence of the impact of climate change on wind (Pryor (2009)), the Intergovernmental Panel on Climate Change (Field et al. (2012)) projects that climate change is likely to increase the frequency of extreme weather events, such as heat waves and heavy precipitation. It follows that climate change could have a consequential impact even on indoor economic activities. Second, given that weather varies across the country, our findings should be considered in the location decision for new plants, along with the traditional factors like labor cost and availability, access to suppliers, proximity to markets, etc. Third, our results complements the existing literature on productivity in the automobile industry (Lieberman et al. (1990), Lieberman and Demeester (1999), among others) by presenting evidence of the impact of extreme weather on productivity; plant managers may be unaware of the impact of weather on their output (e.g., attributing variation in output to unexplained causes or to mechanisms that are caused by weather, such as absenteeism or parts shortages), and can use our results to implement policies to counteract these negative effects (e.g., accelerating deliveries in anticipation of weather). Finally, this paper confirms that weather can be used as an exogenous shock in automobile production, which may be useful in the development of valid instruments for other research.

The third essay, entitled *"Integration of Online and Offline Channels in Retail: The Impact of Sharing Reliable Inventory Availability Information"*.

Online retailing has grown steadily over the last few years. Some retailers operate exclusively through online channels, and traditional brick and mortar (B&M) retailers have incorporated online sales channels since the early stages of the commercial Internet (e.g., the Barnes and Noble website launched in May 1997). Today, retailers' online channels are no longer an experiment but a relevant and growing part of their business. Originally, most of the B&M retailers decided to separate the operations of traditional and online channels. Now, some B&M retailers are exploring integration strategies for their online and B&M channels to enrich the customer value proposition and/or reduce costs. Online-offline integration efforts can occur in a variety of configurations. For example, B&M retailers often show in-store inventory availability information online. More advanced integration includes shipping the product ordered from the store closest to its destination, or offering the option to buy products online and pick them up in the store.

In particular, over the last few months, a number of traditional B&M retailers across different categories (e.g., The Home Depot, Apple, Crate & Barrel, Toys "R" Us, among others) have implemented buy-online-pickup-in- store (BOPS) functionality. The retailer shows online viewers the locations at which the item is available, and gives customers the option to close the transaction online and then pick up the product at one of the locations within two hours of closing the purchase.

The integration of online and offline channels provides an opportunity to empirically study issues that have been the subject of theoretical research in operations management. In this paper, we use an online-offline integration project that implements BOPS functionality as a natural experiment to study the impact of sharing reliable inventory availability information with the customers. Implementing a buy-online-pick-up-in-store project provides an exogenous shock to the verifiability of the inventory information that the firm shows to their customers; because the inventory information becomes more credible, the risk that customers face when deciding whether to visit the store is reduced.

We have collected a novel proprietary dataset from a nationwide retailer that has been among the pioneers in implementing BOPS functionality. Using this dataset and a series of natural experiments, we make the following contributions:

First, we evaluate the impact of BOPS implementation on company sales and customer behavior, and give the first piece of empirical evidence on this emerging trend in retailing. We study the impact of the deployment of a BOPS project on both the online and brick and mortar channels. Conventional wisdom within the industry suggests that offering the BOPS functionality will improve online channel revenue (since BOPS transactions are considered online revenue), and that the traditional B&M stores will carry the burden of having the item ready for the customers to pick up, without receiving any significant benefit in their sales. However, as we will describe in detail, a series of natural experiments leads us to conclude that these assumptions are not correct. Our results show that, contrary to what we would expect, sales transacted online *decrease* and B&M sales *increase* when the BOPS functionality is deployed.

Second, we show how the increase of inventory information verifiability affects customer behavior. The impact of availability information and its verifiability on customer behavior has been the subject of recent modeling research in the field of operations management (e.g., Allon and Bassamboo (2011); Su and Zhang (2009)) but to our knowledge, no empirical results were available prior to this paper. Implementing BOPS functionality can be seen as a shock to the verifiability of inventory information online. To implement BOPS functionality, the online system must have access to accurate real-time information about availability of in-store inventory. If the retailer offers the option to pick up an online order at a particular store, the customer knows with very high certainty that the item ordered is available at that store. Therefore, inventory availability information is perceived as very reliable. This contrasts with situations whereby the store simply shows inventory information but does not offer the option to close the transaction online. For example, consider a car dealership showing information online about their inventory. This information is typically unverifiable; if a customer visits the dealer and the product is not available, the dealer can claim that the online information was not updated in real time. We find that increased reliability of in-store availability information increases the probability that customers will visit the store. We present an explanation consistent with empirical evidence we observed regarding the impact of BOPS functionality: Providing BOPS functionality increases the reliability of the inventory information, resulting in an increase in the number of customers visiting the stores to purchase items after checking product availability online. This provides an explanation to the counterintuitive finding described above. We further check the validity of this explanation by presenting further evidence from the shopping cart abandonment behavior.

We use this project as an example of the evaluation of an online-offline strategy, illustrating the complex interactions between the online and offline channels and the challenges of relying on single channel data to evaluate the impact of interventions that affect multiple channels. Retailers often run experiments in their online channel (for example, A/B testing) to evaluate the impact of interventions on their conversion rates or other measures of interest. In our case, an isolated evaluation of the online channel would have considered the impact of the BOPS implementation to yield negative results. Only when closing the loop and looking at the effects in the brick and mortar channel we can quantify the net effects of the BOPS implementation, which are positive.

While the three essays are independent and self-contained studies, each of them focusing on specific research questions, there are many common themes that are present in all of them, which characterize the general contribution of this dissertation. The main common theme is that these three essays study questions that can be characterized as the relationship and interaction between inventory and demand. The first essay is concern with how inventory visible to the consumer can affect their purchase decision and hence demand. The third essay explores the issue of how inventory availability information can affect the consumer decision related to the channel that the customer will use to purchase an item and the impact on the perceived risk of experiencing a stock-out. The second essay gives a solid support to the argument used in the first essay related to the impact of weather in automobile productivity.

Finally, on the methodological side, a common feature that is shared by the three essays is that they are all empirical studies that deal with the endogeneity challenges of the estimation. The three studies use observational data that comes from proprietary sources. The first essay uses data that 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 dealers in our sample and all plants located in the U.S. and Canada. The second essay combines two main data sets. The first data set is weekly vehicle production in the United States at the plant-model level. The second includes daily weather conditions at our sample of vehicle assembly plants. Both cover the period of January 1994 to December 2005. For the third essay we have partnered with one of the leading nationwide retailers in the US that has implemented buy-online-pickup-at-store capabilities. We have obtained data spanning April 2011 to April 2012. Throughout this period, the online store offered information about the availability of inventory at each of the stores. During this period the retailer started to offered the option of placing orders online and picking them up at a store. The period of analysis considered in our analysis covers six months before the store pickup implementation and extends six months after the implementation.

CHAPTER 2 : Does Adding Inventory Increase Sales? The Billboard and Scarcity Effects in U.S. Automobile Dealerships

2.1. 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 (Automotive-News (2010)). Were those complaints justified?

Clearly sales could fall if a dealer does not have any inventory - it is hard to sell a car if there is no car to sell in markets, like the U.S., where customers are accustomed to purchase directly from units on the lot rather than a make-to-order process as is more common in Europe. But beyond this *stockout effect*, does carrying more or less inventory influence sales? Traditional inventory theory assumes the answer is "no": demand is generally taken to be independent of inventory, so while sales varies in inventory (due to the stockout effect), demand does not. But in some product categories, including automobiles, there is reason to believe that demand may indeed depend in part on the amount of inventory carried, thereby creating a link between inventory and sales beyond the stockout effect. 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 demand must be slow because the car is not popular, and it must not be popular for a reason, so the customer becomes less likely to purchase. In general, we use the label "*billboard effect*" for any mechanism that assigns a positive relationship between inventory and demand, and "*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 lead either to a billboard or to a scarcity effect, estimating the relationship between inventory and sales is complex primarily because it is reasonable to believe that inventories are endogenously chosen. 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.

2.2. The stockout, billboard and scarcity effects

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, $E[min(q,d)]$, simply because a stockout occurs when $d > q$. Hence, the stockout effect suggests that sales increase with inventory even though demand is independent of inventory. However, the magnitude of the effect also diminishes with inventory, i.e., the effect is small when q is large.

There are also various mechanisms that we collectively label as billboard effects because they create a positive relationship between demand (the likelihood a customer wants to buy) and inventory, which is observed as a positive relationship between inventory and sales (see Balakrishnan et al. (2004) for a model that assumes this mechanisms). Variety is one example. Suppose 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 also 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. 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)).

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 - as with a literal billboard, a large inventory may 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., Dana and Petruzzi

(2001)). 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.

In contrast to the billboard effect, 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)). 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. 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), Kuksov and Villas-Boas (2010)).

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). 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 lead to a billboard effect (ample inventory enables a better preference match, increases awareness, signals popularity, indicates availability and suggests the potential to obtain a good price) while other mechanisms lead to a scarcity effect (ample inventory signals an unpopular vehicle, creates overwhelming choice, suggests that prices will soon be lowered, and reduces the urgency to purchase immediately while encouraging additional search).

2.3. 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.

2.3.1. 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.[1]

We removed from our sample a limited number of dealerships that opened or closed during the period under analysis.

2.3.2. 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,

[1]If 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 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[2].

Using the website from the National Weather Service Forecast Office (NWSFO) and the website 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.

2.4. Model Specification

We seek to estimate the impact of inventory and variety on sales. The available data was used to construct a panel data-set where the unit of analysis is the sales of a particular vehicle model i at a specific dealership j during a week t ($Sales_{ijt}$). Expected sales during a week are influenced by the total number of vehicles available at the dealership during the week ($Inventory_{ijt}$), the number of varieties of the model that where available ($Variety_{ijt}$, 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.

Figure 1 shows multiple effects between the three key variables. First, there is a direct effect of inventory on sales (labeled with the coefficient β_{13}). An example of this effect is when low levels of inventory signal low future availability of the vehicle model and lead to a "buy frenzy" behavior, or when high levels of inventory signals lower prices and therefore increases sales. Therefore, the sign of β_{13} is ambiguous. Second, there is a direct effect of

[2] 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

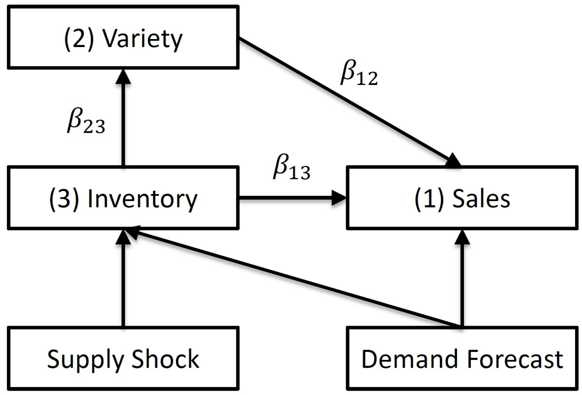


Figure 1: Relationship between Sales, Inventory and Variety

variety on sales (labeled by β_{12}), 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 β_{12} is also ambiguous. Third, there is an indirect effect of inventory on sales through variety (labeled β_{23}): 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 equation 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:

$$
\begin{aligned}
Sales &= \beta_{12} Variety + \beta_{13} Inventory &&+\gamma_1 Z &&+\varepsilon_1 && (2.1)\\
Variety &= \beta_{23} Inventory &&+\gamma_2 Z &&+\varepsilon_2 && (2.2)\\
Inventory &= &&+\gamma_3 Z + \delta_3 W &&+\varepsilon_3 && (2.3)
\end{aligned}
$$

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 $\{\varepsilon_g\}_{g=1,2,3}$ 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 $g = 1, 2, 3$. Next, we discuss identification of the system of equations (1)-(3).

The error term ε_1 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, ε_1 and ε_3 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 δ_3 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. The reason is that, since inventory also affects variety ($Inventory$ is an explanatory variable in equation (2)), $Variety$ is also endogenous in equation (1). Hence, $Variety$ has to 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.,$\delta_2 = 0$). 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 $(\varepsilon_1, \varepsilon_2, \varepsilon_3)$. As previously mentioned, it is likely that ε_1 and ε_3 are positively correlated due to inventory endogeneity. However, it is reasonable to assume that ε_1 and ε_2 are uncorrelated, that is, $E(\varepsilon_1\varepsilon_2) = 0$. 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 $E(\varepsilon_2\varepsilon_3) = 0$: because dealers can only control variety through their inventory levels, other factors that induce variation in variety (captured by ε_2) should be unrelated to factors that affect inventory. These assumptions are sufficient for identification, as shown next.

Proposition 2.1. *If $E(\varepsilon_g|Z, W) = 0$, for all $g \in \{1, 2, 3\}$, $E(\varepsilon_1\varepsilon_2) = 0$ and $E(\varepsilon_2\varepsilon_3)$, then all the parameters of the system of equations (1)-(3) are identified.*

The proof of the proposition can be found in the appendix.

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 $E(\varepsilon_1\varepsilon_2) = 0$, the residual of equation (2) can be used as an instrument for *Variety* in equation (1). This requires a consistent estimator of ε_2. Under the assumption that $E(\varepsilon_2\varepsilon_3)$, the residual of the OLS regression of (2), denoted $\hat{e}_2$, is a consistent estimator of ε_2. 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 $\widehat{Inventory} = \hat{\gamma}_3 Z + \hat{\delta}_3 W$ and the residuals $\hat{e}_2 = Variety - \hat{\beta}_{23} Inventory - \hat{\gamma}_2 Z$.

3. Estimate equation (1) via Two-Stage Least Square using $\hat{e}_2$ and $\widehat{Inventory}$ as instrumental variables for the endogenous variables $Variety$ and $Inventory$.

To obtain the standard errors of this estimation we did 400 bootstraps over these 3 steps, sampling at the dealer-model level.

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 $Seasonal_{r(j)t}$ 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 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 $Variety_{ijt}$ would suggest that the confusion effect dominates the impact on sales. Although $Inventory_{ijt}$ 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 $Variety$, 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.

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 $AvailVar_{ijt}$ 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.

Table 2 summarizes the number of different sub-models observed in our data and the average $Variety_{ijt}$ 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 one more sub-model of a Cobalt (which has many sub-models) can have a smaller impact than adding one more sub-model of an Equinox (for which fewer sub-models were produced). 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 $MarketVar_j$ as the number of sub-models produced for model j in the model-year 2007. Our main measure of variety is defined as:

$$Variety_{ijt} = \frac{AvailVar_{ijt}}{MarketVar_j}. \tag{2.4}$$

For robustness, we considered other definitions of variety; for example, we used $AvailVar_{ijt}$ and its logarithm as alternative measures. The results using these alternative measures, discussed in Section 3.4, were similar.

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.

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.[3] *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.

The thirteen weather variables were reduced to five principal components. By capturing

[3] Cachon et al. (2011) report some measurement problems for fog, but these problems were not observed in 2007, our study period.

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 12. 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 (195) of join significance of the instruments exceed the usual 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.[4]

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 $\{\varepsilon_g\}_{g=1,2,3}$. However, it is possible to estimate the overall effect of inventory on sales – which corresponds to the direct effect β_{13} plus the indirect effect through variety, $\beta_{12}\beta_{23}$ (see figure 1) – under weaker assumptions. To see this, replace $Variety$ from equation (2) into equation (1) :

$$Sales = (\beta_{13} + \beta_{12}\beta_{23})Inventory + \gamma_1' Z + \varepsilon_1', \tag{2.5}$$

[4]Another specification assumed there is no lead-time, hence we include contemporaneous weather. Finally, we considered another specification that included a specific lead-time for each vehicle to the dealers depending on the distance between the dealer and the plant where a particular model was manufactured. When this distance is less than 600 miles we consider that the vehicle will arrive within the week (zero lag). When the distance is between 600 and 1200 miles we consider that the vehicle will arrive with a lead time of one week and when the distance is more than 1200 miles we consider a two weeks lead-time.

where $\gamma_1' = \gamma_1 + \beta_{12}\gamma_2$ and $\varepsilon_1' = \varepsilon_1 + \beta_{12}\varepsilon_2$. Under the exogeneity assumption $E(\varepsilon_g|Z, W)$, $g \in \{1, 2, 3\}$, the coefficient $\beta_{13}' \equiv \beta_{13} + \beta_{12}\beta_{23}$ 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 $(\varepsilon_1, \varepsilon_2, \varepsilon_3)$. 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 2.6.

2.5. Results

Table 11 reports the main estimation results. Column (1) shows the estimates of equation (1), instrumenting the endogenous variables inventory and variety (as defined in equation (2.4)). The estimates suggest that the direct effect of inventory (β_{13} in Figure (1)) is negative and statistically significant, but the effect of variety (β_{12}) 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 small and positive relationship between inventory and variety: the estimated coefficient is $\hat{\beta}_{23} = 0.0054$, with a standard error $SE(\hat{\beta}_{23}) = .0001$. 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 $\beta_{13} + \beta_{12}\beta_{23} = -0.013$ (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 2.6.

Recall that the estimates of column (1) are consistent if the error in equation (2), ε_2, is uncorrelated with ε_1 and ε_3. However, the overall impact of inventory on sales can be obtained by estimating equation (2.5) directly via instrumental variables, without any assumptions on the error term ε_2 (other than the maintained exogeneity assumption $E(\varepsilon_g|Z,W)=0$, for $g=1,2,3$). Column (2) in table 11 shows these estimates. The coefficient of inventory is -0.014, which is close to our previous estimate based on the coefficients of column (1) (which gave -0.013).[5] This provides some support to validate the consistency of the estimates of column (1).

To assess the magnitude of the bias induced by the endogeneity of inventory, we estimated model (2.5) via Ordinary Least Squares (OLS). As mentioned in Section 3.3, if inventory is set in anticipation of demand, then ε_1 and ε_3 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 ε_3 and thereby with ε_1. Hence, ignoring the endogeneity of variety could also lead to a positive bias on the

[5] 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.

estimate of coefficient β_{23}, 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 column (1) of Table 11 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 improves the match between consumer preferences and the available inventory, 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, 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.

Our findings do not suggest that inventory has a strong relationship on how dealer price or how consumers bargains. As shown by Moreno and Terwiesch (2012) 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. Similarly, our estimates cannot simply be explained by a stockout effect - if adding inventory prevents stockouts, then coefficient β_{13} should be positive, not negative.

It is possible that the scarcity effect we observe is due to the information inventory conveys to consumers. For example, a consumer might infer that ample inventory is a signal that the 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 - 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 level rather than based on the inventory of each of 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 influence 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 motivate an immediate sale.

2.5.1. Robustness analysis

In this robustness section we want to look into issues that can potential affect our result or the conclusions we draw from them. We will focus on the following four issues: competition among dealers, alternative measures of variety, transfers between dealers and estimation issues due to sales being a count variable.

Competition among dealers

In our first robustness analysis we want to study to what extent our results are affected by competition among dealers of the different GM brands. 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, inventory level for a model at the dealers vary from one week to another and this variation can potentially change the competitive landscape for the dealers. To explore the impact of these changes we estimate our main model with a subsample of dealers that don't face competition in their local market. This requires defining the relevant market of a dealer. Albuquerque and Bronnenberg (2012) conduct a study of dealership's demand for autos and show that a 15 mile radius covers most of the relevant market of a dealership. Hence, we defined a sub-sample of dealers with no competing GM dealer, of each particular 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.

Alternative measures of variety

To evaluate how robust our analysis is to different specifications of the variety variable, we replicated the analysis described on the previous section considering two alternative measures of variety: (i) $AvailVar_{ijt}$, the total number of different sub-models carried by a dealer on each week (instead of the relative measure of variety considered before); and (ii) the logarithm of $AvailVar_{ijt}$. 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 11 column (1)). Second, the coefficient of variety is positive and significant in both specifications.

Transfers between dealers

A second issue that can potentially affect our conclusions is related to the transfer of vehicles between dealers. When dealers lack of a particular submodel that is being demanded buy a customer they have the possibility to look for that particular vehicle in nearby dealers. When a dealer finds the car she can request a transfer. This transfer needs to be approved by the dealer that is currently carrying the vehicle. If the transfer is approved it can happen in two different ways: the transfer can by a swap of vehicles between dealers or the transfer can be an internal sale between dealers involving only one car. In our sample sales that include a transfers represent 12.9% of the total sales.

In both cases it is possible to argue that sales of those vehicles that include a transfer were not directly influenced by the inventory at the dealer's lot. If this is the case our results can be explain through a different mechanism: when facing low inventory for a particular model, dealers change their behavior by increasing the number of transfers. To find out if this potential explanation is taking place we look in to this issue by exploring the existence of this change in behavior.

We first identify those sales that involved a transfer. To do this we look into vehicles transfers that were sold within ten days of arriving to the dealer's lot. We consider this ten days threshold because we wanted to make sure that we include in the analysis the vehicle that were transfers to be sold and not those that were part of the transfer swap. Then we estimate our model including only those sales that were linked to the transfer. The result we obtain shows that inventory quantity was not a significant variable to explain changes in those "transfer-sales". This results rules out the potential explanation that the increase in sales we observe is driven by a change in the number of transfers that dealers are willing to make when they face low inventory quantity.

Sales being a count variable

Finally a potential estimation issue is that $Sales$, our main dependent variable, is a count variable with frequent zeroes (about 60% of the weekly model sales were zero). A negative binomial regression, which accounts for the counting nature of sales, could increase the model fit and therefore lead to more precise estimates. We argue, however, that our estimation strategy – which is consistent under weaker assumptions – already provide precise estimates of the coefficients of interest. Nevertheless, we estimated a model of sales via a negative binomial regression, including inventory as independent variable. We used a control function approach to account for the endogeneity of inventory (see Wooldridge

(2010) and Hilbe (2011) for details on the implementation of the control function approach in negative binomial regressions). The results are presented in column (4) of table 5. The overall effect of inventory for this model is still negative and statistical significance as in the comparable specification reported in table 11, column (3).

We also considered alternative specification that include inventory with log transformation, variety measured in the actual number of sub-models (as defined by $AvailVar$ rather than $Variety$), its logarithm, and combinations of these. The results obtained were similar in magnitude and statistical significance.

2.6. 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 two different approaches. The first approach estimates the potential sales improvement from reallocating the existing vehicles among the dealers in a small local area. The second approach 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.

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.

2.6.1. Local reallocation among dealers

The analysis in this section partitions dealers into reasonably 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 efficiently 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 a vehicle.

To formulate the problem as a mathematical program, let $i = 1 \dots n$ indexes the dealers within the CBSA and $k = 1 \dots m_j$ 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 the reallocating vehicles among the dealers within the CBSA – in order to solve the following non-linear integer optimization problem:

$$\max_{Q_{ijk}} \left[\sum_{i=1}^{n} \exp \left(\delta_{ij} + \hat{\beta}_{13} \sum_{k=1}^{m_j} Q_{ijk} + \hat{\beta}_{12} \cdot \text{Variety}_{ij} \right) \right] \tag{2.6}$$

s.t.

$$\forall k \qquad \sum_{i=1}^{n} Q_{ijk} = \sum_{i=1}^{n} I_{ijk} \tag{2.7}$$

$$\forall i \qquad \sum_{k=1}^{m_j} Q_{ijk} = \sum_{k=1}^{m_j} I_{ijk} \tag{2.8}$$

$$\text{Variety}_{ij} = \frac{\sum_k^{m_j} \mathbb{I}\left(\mathrm{Q}_{ijk} \geq 1\right)}{m_j} \tag{2.9}$$

$$Q_{ijk} \in 0, 1, 2, \ldots \infty \tag{2.10}$$

where I_{ijk} is dealer i's initial endowment of inventory of model j and sub-model k (i.e., if there is no reallocation). The parameters $\hat{\beta}_{12}$ and $\hat{\beta}_{13}$ are the estimated coefficients of *Inventory* and *Variety,* respectively, and δ_{ij} is the estimated model-dealer fixed-effect (the estimates correspond to the model reported in Table 11, column (1)). Constraint (2.7) ensures that the reallocation does not change the total inventory within the CBSA of sub-model k and constraints (2.8) ensure that dealer $i's$ inventory of model j after the exchanges is identical to its inventory before the exchange. The objective is then to maximize $Variety_{ij}$ (which is a function of the decision variables Q_{ijk} as described by 2.9) while keeping the dealership's model inventory constant.

The first column on Table 6 shows the solution to this math program, measured by the average potential sales improvement for each car model. We find that on average, exchanging inventory among dealers within a CBSA with the objective of maximizing each dealer's offered variety yields an weighted average sales gain of 1.7%.

2.6.2. State-wide reallocation of vehicles

Instead of swapping vehicles after they arrive at dealerships, we now consider changing the allocation of vehicles after they leave the production facility. At that point in time there may be some flexibility with respect to the final destination of vehicle and this flexibility may come with little incremental cost. 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. With this approach there are no transfers among dealers - each dealer's initial inventory remains with that dealer. However, rather than send sub-model k to a dealer who already has some units of sub-model k, it is better to send that vehicle to a dealer who begins the week without any units of sub-model k.

To specify this optimization problem, fix a given week and define A_{jk} as the total number of sub-model jk sent to the state on this week. The decision is to choose Y_{ijk} , the number of units dealer i receives of sub-model jk during the week. For each state and each model j we solved the following integer non-linear optimization problem:

$$\max_{Y_{ijk}} \left[\sum_{i}^{n} \exp\left(\delta_{ij} + \bar{\hat{\beta}}_{13} \cdot \sum_{k}^{m_j} Q_{ijk} + \hat{\beta}_{12} \cdot \text{Variety}_{ij} \right) \right] \tag{2.11}$$

s.t.

$$\sum_{i}^{n} Y_{ijk} = A_{jk} \tag{2.12}$$

$$\sum_{k}^{m_j} Q_{ijk} \leq M_{ij} \tag{2.13}$$

$$Q_{ijk} = I_{ijk} + Y_{ijk} \tag{2.14}$$

$$\text{Variety}_{ij} = \frac{\sum_{k}^{m_j} \mathbb{I}\left(\text{Q}_{ijk} \geq 1\right)}{m_j} \tag{2.15}$$

$$Y_{ijk} \geqslant 0 \tag{2.16}$$

Constraint (2.12) 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. To preclude allocations that result in an unreasonably large amount of inventory allocated to some dealers, constraint (2.13) 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, denoted by M_{ij}. Equation (2.14) merely states that a dealer's inventory of a model equals the dealer's initial endowment, I_{ijk}, 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%.

2.7. 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 billboard effect - a positive relationship between inventory and sales. For example, at a basic level, adding inventory can increase sales by improving the visibility of the product 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 additional search from the customers or may give a sense that the product is not selling due to inferior quality.

In our sample, we find that an increase in inventory at a dealer actually lowers sales. However, it is important to decompose this effect into two parts: increasing inventory of a sub-model does indeed reduces sales, but if the increase in inventory also expands the number of sub-models available, then sales increase. In short, the benefit of expanding variety can dominate 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.

Our findings emphasizes the importance of careful vehicle allocation. The data suggest that vehicles are allocated in a way that does not maximize the heterogeneity of sub-models available to consumers. Dealers may view one sub-model as particularly desirable and then take actions to increase their inventory in that sub-model rather than to expand the set of sub-models offered. Based on our estimates, an allocation policy that is focused on maximizing variety can increase sales by about 2.5%, without changing the number of vehicles produced or the number of vehicles each dealer carries. In some cases this sales improvement may come with relatively little incremental costs - transportation costs are dominated by the frequency and quantity of deliveries, and less so by the composition of each delivery.

2.8. Appendix

Proof Proposition 1

Consider the following system of simultaneous equations:

$$Sales = \beta_{12} Variety + \beta_{13} Inventory \quad + \gamma_1 Z \quad + \varepsilon_1 \tag{2.17}$$

$$Variety = \beta_{23} Inventory \quad + \gamma_2 Z \quad + \varepsilon_2 \tag{2.18}$$

$$Inventory = + \gamma_3 Z + \delta_3 W \quad + \varepsilon_3 \tag{2.19}$$

We show that if $E(\varepsilon_g|Z, W) = 0$, for all $g \in \{1, 2, 3\}$, $E(\varepsilon_1\varepsilon_2) = 0$ and $\mathrm{E}(\varepsilon_2\varepsilon_3) = 0$, then all the parameters of the system of equations (1)-(3) are identified.

Proof. The reduced form of the system of equations (1)-(3) is denoted by:

$$Sales = \pi_1 Z + \psi_1 W + u_1 \tag{2.20}$$

$$Variety = \pi_2 Z + \psi_2 W + u_2 \tag{2.21}$$

$$Inventory = \pi_3 Z + \psi_3 W + u_3 \tag{2.22}$$

Because Z and W are exogenous, the coefficients (π_1, π_2, π_3) and (ψ_1, ψ_2, ψ_3) are identified, as well as the covariance matrix of the reduced form error terms (u_1, u_2, u_3), denoted by Ω.

The triangular structure of the system (17)-(19) facilitates its inversion into the reduced form system (2.20)-(2.22). First, equations (19) and (2.22) are identical, so $\pi_3 = \gamma_3$, $\psi_3 = \delta_3$ and $\Omega_{33} = \Sigma_{33}$ Hence, equation (2.22) alone identifies γ_3, δ_3 and $\Sigma_{33} = Var(\varepsilon_3)$. For equation (2.21) we have:

$$\pi_2 = \beta_{23}\gamma_3 + \gamma_2 \qquad \psi_2 = \beta_{23}\delta_3$$

so $\beta_{23} = \psi_2/\delta_3$ and $\gamma_2 = \pi_2 - \frac{\psi_2}{\delta_3}\gamma_3$ are also identified. The variance of u_2 is given by:

$$\Omega_{22} = \Sigma_{22} + \beta_{23}^2 \Sigma_{33}$$

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

For equation (2.20):

$$\begin{aligned} \pi_1 &= \beta_{12}\beta_{23}\gamma_3 + \beta_{12}\gamma_2 + \beta_{13}\gamma_3 + \gamma_1 \\ \psi_1 &= \beta_{12}\beta_{23}\delta_3 + \beta_{13}\delta_3 \end{aligned} \tag{2.23}$$

with unknowns β_{12}, β_{13} and γ_1. Additional identifying equations can be obtained from Ω, the covariance matrix of the reduced form error u. By definition, the reduced form errors of equations (2.21) and (2.22) imply:

$$u_2 - \beta_{23} u_3 = \varepsilon_2$$

$$u_1 = \varepsilon_1 + \beta_{12}\varepsilon_2 + (\beta_{12}\beta_{23} + \beta_{13})\varepsilon_3$$

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

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

This identifies $\beta_{12} = (\Omega_{12} - \beta_{13}\Omega_{13})/\Sigma_{22}$, which replacing in (2.23) identifies the remaining parameters β_{13} and γ_1. $\square$

Table 1: Dealers by state in our sample

State	Number of Dealers
California	355
Colorado	67
Florida	237
Main	31
Nebraska	50
Texas	366
Wisconsin	183
TOTAL	1,289

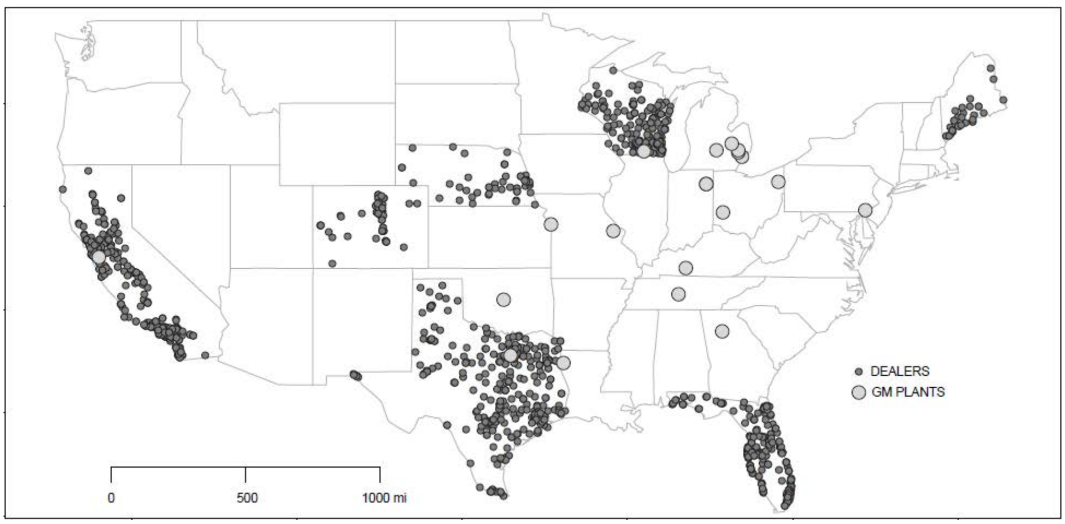


Figure 2: Dealer and plant locations in our sample

Table 2: Model Variety for the top ten selling models

	Total Model Variety (*MarketVar*)	Average Variety Available (*AvailVar*)
Cobalt	18	3.5
Equinox	4	2.2
G6	37	6.1
HHR	4	2.9
Impala	10	3.7
Suburban	18	4.5
Tahoe	13	4.0
TrailBlazer	10	2.1
Saturn VUE	5	4.6
Yukon	30	8.6
AVERAGE	14.9	4.2

Model Variety is the maximum number of variants that could be produced for the model.
Available variety is the number of variants with at least one unit during a particular week.

Table 3: Weather variables included in the empirical study

Variable	Description
Wind	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.
Cloud	Average cloud cover during the week (0 = no clouds; 8 = sky completely covered).
Fog 1	Weeks with 1 days with fog during the week.
Fog 2-3	Weeks with 2 or 3 days of fog during the week.
Fog 4-7	Weeks with more than 3 days of fog during the week.
Rain 1-2	Weeks with 1 or 2 days of rain during the week.
Rain 3-5	Weeks with 3 to 5 days of rain during the week.
Rain >5	Weeks with more than 5 days of rain during the week.
Snow 1	Weeks with 1 day of snow during the week.
Snow 2-4	Weeks with 2 to 4 days of snow during the week.
Snow >4	Weeks with more than 4 days of snow during the week.
High Temp 1	Weeks with 1 day of high temperature, above 90 degrees Fahrenheit, during the week.
High Temp 2-5	Weeks with 2 to 5 days of high temperature, above 90 degrees Fahrenheit, during the week.
High Temp 6-7	Weeks with more than 5 days of high temperature, above 90 degrees Fahrenheit, during the week.

Table 4: Main Model Results - Log Linear Model

	(1)	(2)	(3)	(4)
Inventory	-0.0147***	-0.0140***	0.0130***	-0.0189***
	(0.0030)	(0.0030)	(0.0002)	(0.0034)
Variety	0.2958***			0.7223***
	(0.0094)			(0.0487)
Fixed Effects	YES	YES	YES	YES
Week - Season	YES	YES	YES	YES
Dealer's Local Weather	YES	YES	YES	YES
N	293776	293776	293776	293776
N_g	12969	12969	12969	12969

Standard errors in parentheses

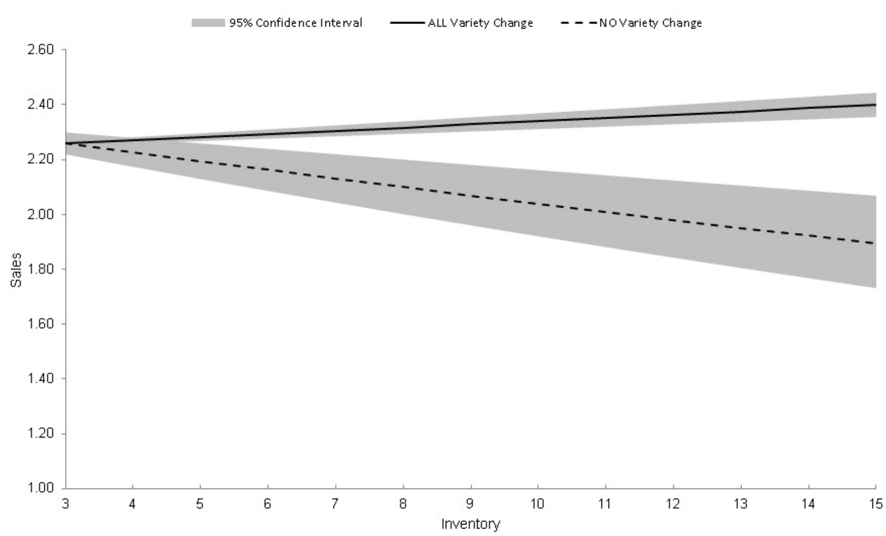
* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

(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.



The Figure 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), for a dealer that starts with 3 vehicles of a particular model.

Figure 3: Sensitivity Analysis

Table 5: Robustness Checks

	(1)	(2)	(3)	(4)
Inventory	-0.0171*	-0.0151***	-0.0156***	-0.0752***
	(0.0074)	(0.0030)	(0.0030)	(0.0114)
Variety	0.2847***			
	(0.0123)			
$AvailVar$		0.0421***		
		(0.0012)		
$Log(AvailVar)$			0.1310***	
			(0.0034)	
Fixed Effects	YES	YES	YES	YES
Week - Season	YES	YES	YES	YES
Dealer's Local Weather	YES	YES	YES	YES
N	150619	274399	274399	274399
N_g	6803	11879	11879	11879

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 submodels.

(3) Analysis including the logarithm of variety as a count of different submodels.

(4) Analysis considering a negative binomial specification.

Table 6: The impact of inventory allocation

	Average CBSA Improvement Vehicle Swap	Average State Improvement Reallocation
Cobalt	0.7%	2.0%
Equinox	0.5%	1.5%
G6	0.5%	2.0%
HHR	0.5%	4.8%
Impala	0.9%	3.9%
Suburban	1.4%	1.2%
Tahoe	0.9%	1.7%
TrailBlazer	0.7%	0.6%
Saturn VUE	5.6%	3.0%
Yukon	4.5%	2.6%
WEIGHTED AVERAGE	1.7%	2.5%

CHAPTER 3 : Severe Weather and Automobile Assembly Productivity

3.1. Introduction

It is well known that there is a relationship between climate and economic activity. For example, not only are hot countries poorer, temperature even explains variation in economic output within countries (Dell et al., 2009). It is intuitive that climate can impact outdoor activities like agriculture, forestry, construction and tourism. Less clear is the impact on "climate-insensitive" sectors such as manufacturing and services (Nordhaud, 2006).

In this paper we study the relationship between severe weather and weekly automobile production at 64 facilities within the United States over a ten-year period. Although automobiles are made indoors, there are several mechanisms through which bad weather at a plant could influence production. For example, high winds, icy roads or heavy precipitation could cause delays in in-bound delivery of parts from suppliers, possibly due to additional traffic congestion, accidents or cancelled shipments. (See Brodsky and Hakkert (1988); Golob and Recker (2003) for data on precipitation and traffic accidents.). Finished vehicles might be damaged during periods of high wind or hail once they exit the plant. In addition, if a plant operates in a "just-in-time" fashion with relatively little buffer stock of parts, the plant may need to delay the start of a shift or cancel a shift altogether due to the absence of needed parts. The same concern applies to "in-bound" employees – production could be curtailed if workers are unable to (or choose not to) travel to the plant. Finally, even if all of the workers and parts are available, it is possible that bad weather could influence employee productivity. For example, with extreme heat conditions outside, even if the plant has a cooling system, it is possible that the indoor temperature rises to a level that slows down the manual labor associated with automobile production.[1] Alternatively, bad weather outside could influence the affect of employees which in turn may lower their productivity.[2]

[1] It has been established that thermal heat stress has a non-linear impact on productivity – the impact of increased temperature begins around 25° C (Ramsey and Morrissey (1978) and Wyon (2001)). Internal temperatures in a automobile plant may exceed this threshold, especially if the outside temperature is high.

[2] Simonsohn (2010) finds that the decisions of admissions officers at an academically oriented college are

In short, it seems reasonable to conclude that weather could influence seemingly sheltered indoor economic activity.

For our study it is safe (we believe) to assume that production does not cause changes in the weather - whether a plant produces more or fewer cars in a week is unlikely to influence its local weather in that week. Of greater concern is whether weather exerts a causal influence on production - are there omitted variables that could lead to an endogeneity bias? For example, maybe automobile production is seasonal for reasons unrelated to local weather. If production seasonality is correlated with a plant's weather (e.g., if fewer cars are made in the summer because demand across the country is lower during the summer), then local weather may only be a proxy for this seasonality. To address this issue we take advantage of the panel structure of our data to include a number of controls: product introduction and ramp-down dummies to account for the possibility that vehicles are introduced at certain times of the year (and their obvious influence on the level of production); plant fixed effects to account for idiosyncratic plant characteristics associated with seasonality; planned shifts to account for known variations in production; weekly dummies to account for national variations in demand, monthly segment dummies (e.g., cars, vans, etc.) to account for segment specific demand seasonality; regional year-month dummies to account for regional differences in weather fluctuations and the possibility that the influence of weather varies by region; and seasonal average weather measures for each plant (e.g., average amount of rain in week t for plant i). In sum, given our extensive set of controls, we believe we have identified a causal impact of severe weather on production.

We also find that weather has a substantial economic impact on automobile production. For example, we estimate that for an average plant, within a week, six or more days with a high temperature of 90^oF or one additional day of heavy winds reduces that week's production by approximately 8%, and six or more days of rain within a week reduces production relative to

influenced by cloud cover even though admission decisions are not made outside nor should they objectively be influenced by the weather. However, Lee and Staats (2012) argue that bad weather may increase productivity as it eliminates cognitive distractions associated with good weather.

no rain by 6%. Furthermore, we find that average weekly production losses due to weather events (snow, rain, heat and wind) ranges from a low of 0.5% (Princeton, IN) to a high of 3% (Montgomery, AL), with an overall average of 1.5%. Hence, even though the severe events we identify are not common (e.g., there is only about 2.5 high wind days per year per plant), they are sufficiently common that their collective effect is meaningful.

Our data are suitable for measuring the short term impact of weather on production. An interesting question is how do plants react to the productions shocks we observe? On one extreme, the production could be "lost forever", while at the other extreme, the plant may fully recover the lost production in the same week the weather event occurs. Even if they are able to recover some production in the same week, we find the net impact of severe weather on a week's production to be negative. In addition, we do not find evidence that they are able to recover in the following week - plants are not more likely to schedule overtime in a week that follows bad weather, nor is production higher in weeks following bad weather (all else being equal). Nevertheless, we cannot rule out the possibility that plants recover the production at some point in the future. However, even if they were to fully recover at some point, at the very least, such recovery increases the variability of production (which is costly) and may lead to delayed shipments and stockouts.

Our work is related to a growing literature on the impact of climate and weather on economics. A number of studies focus on agriculture (e.g., Crocker and Horst (1981); Mendelsohn et al. (1994); Olesen and Bindi (2002); Deschenes and Greenstone (2007)). Others include more (or all) sectors of the economy. For example, Dell et al. (2008) find in a long time horizon sample that a 1° C increase in temperature in a given year decreases economic growth in a sample of poor countries by 1.1 percentage points. However, they do not find evidence that annual shocks in temperature or precipitation have an impact on growth of "rich" countries. Using import and export data, Jones and Olken (2010) report results that are consistent with those from Dell et al. (2008). Andersen et al. (2010) report that at the state level, the incidence of lightning strikes influences growth rates in the United

States over the period of 1990-2007. Also with U.S. data, Bansal and Ochoa (2009) report a substantial negative correlation (-0.79) between 10-year changes in temperature and 10-year GDP growth. Hsiang (2010) finds that a 1° C temperature increase in a year's average temperature decreases output in 28 Caribbean-basin countries. The largest negative impact is in the "wholesale, retail, restaurants and hotel" sector (-6.5%) and the smallest is in "manufacturing" (+1.4%). Our study differs in that we focus on a single industry (automobile production), we measure the short term effect (weekly data) of local weather on local productivity (a single plant) and we expand the array of observed weather variables beyond temperature and precipitation (e.g., wind). The fine granularity of our data allows us to identify meaningful effects that could be masked in more aggregated data (regional, annual data).

There is a considerable literature on supply chain disruptions. For example, a number of papers investigate sourcing strategies when suppliers have varying reliability (e.g. Tomlin (2006), Wang and Tomlin (2010), Dong and Tomlin (2012)). Some work investigates disruptions empirically (e.g., Hendricks and Singhal (2005)) but in none of these cases is a connection made between the disruption and severe weather. Furthermore, the focus is generally on upstream disruptions whereas we investigate the impact of local disruptions (i.e., local weather).

Our results could be useful in several ways. First, they are related to the issue of climate change. While there is low confidence of the impact of climate change on wind (Pryor (2009)), the Intergovernmental Panel on Climate Change (Field et al. (2012)) projects that climate change is likely to increase the frequency of extreme weather events, such as heat waves and heavy precipitation. It follows that climate change could have a consequential impact even on indoor economic activities. Second, given that weather varies across the country, our findings should be considered in the location decision for new plants, along with the traditional factors like labor cost and availability, access to suppliers, proximity to markets, etc. Third, our results complements the existing literature on productivity in the automobile industry (Lieberman et al. (1990), Lieberman and Demeester (1999), among others) by presenting evidence of the impact of extreme weather on productivity; plant managers may be unaware of the impact of weather on their output (e.g., attributing variation in output to unexplained causes or to mechanisms that are caused by weather, such as absenteeism or parts shortages), and can use our results to implement policies to counteract these negative effects (e.g., accelerating deliveries in anticipation of weather). Finally, this paper confirms that weather can be used as an exogenous shock in automobile production, which may be useful in the development of valid instruments for other research.

3.2. Data

Our study combines two main data sets. The first is weekly vehicle production in the United States at the plant-model level. The second includes daily weather conditions at our sample of vehicle assembly plants. Both cover the period of January 1994 to December 2005.

3.2.1. Production data

For the period January 1994 to December 2005, we obtained from Wards Auto weekly production of each model produced at all 64 U.S. vehicle assembly plants making light-passenger vehicles, including cars, sport utility vehicles, mini-vans, and pick-up trucks. (We exclude heavy-truck production.) Manufacturers report these data to market analysts. In addition, for each plant and each week we obtained data on the number of shifts and hours scheduled from Automotive News, The Harbour Report and the Interuniversity Consortium for Political and Social Research (ICPSR). Similar data were used by (Bresnahan and Ramey, 1994).

As the production data is reported at the model level, we are able to infer when a plant was closed during a particular week (i.e., zero production), when a particular model was introduced (first week of reported production) or discontinued (last week of reported production). Naturally, we also can infer when a plant is opened or permanently closed.

Table 1 provides descriptive statistics on the production data for the plants in our sample and Figure 1 shows their geographic location.

3.2.2. Geographic location and weather

For each of the 64 plants in our sample we obtained its address and exact geographic location (longitude and latitude). We identified the closest weather station to each plant. Using the National Weather Service Forecast Office (NWSFO) and the weather.com website, we obtained from these weather stations daily data for the period January 1994 to December 2005. Included in the sample are for each day the day's maximum, mean and minimum values for the following weather variables: temperature, wind speed, humidity, pressure, visibility and dew point. We also obtained information on the type of event during a day (rain, thunderstorm, snow, etc.). Finally, we obtained historical weather data for each day: the historical average high, low and mean temperature and the record high and low temperature.

The selected weather stations are close to our plants with a mean and median distance of 13 and 10 miles, respectively. No plant is further than 36 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]

3.3. Model Specification

Using the collected data on plant production and weather, we constructed a panel dataset that relates weekly plant production to weather-related factors and other control variables. We use i to index a plant (e.g. Fort Wayne, Indiana) and t to index a specific week

[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.

(e.g. 3rd week of 2002). Because there is substantial heterogeneity in the production volume across plants, we define the dependent variable in the regression as the logarithm of weekly production ($logProd_{it}$). Hence, the impact of weather on production is measured in relative terms (percent of total production) rather than in absolute terms. The covariates in the regression can be grouped into three categories: (i) factors related to local plant weather (denoted $WEATHER_{it}$); (ii) variables related to seasonality, which could potentially vary across plants ($SEASONAL_{it}$); and (iii) other factors that affect plant productivity ($PRODFACTORS_{it}$). The linear regression model can be summarized as follows:

$$logProd_{it} = \beta WEATHER_{it} + \gamma_1 SEASONAL_{it} + \gamma_2 PRODFACTORS_{it} + \delta_i + \varepsilon_{it}. \quad (3.1)$$

The term δ_i is a fixed-effect that captures the plant's average production, and ε_{it} is the error term. In what follows, we describe the covariates included in *WEATHER, SEASONAL* and *PRODFACTORS.*

Using daily weather data, we constructed several measures capturing weather conditions at each plant for every week in our sample period. The literature in climate and weather research uses two main approaches to measure severe weather: (1) based on the likelihood of occurrence of the event, typically measured as percentiles of the probability distribution for a given time period and location; (2) number of days above specific absolute thresholds of temperature or precipitation (Field et al. (2012) Box 3-1). An advantage of the second (the absolute threshold) approach is that it facilitates the comparison across regions. For this reason, we use this approach in most of our analysis. The main disadvantage is that the impact of an event exceeding a fixed threshold may depend on its location and the time of the year. As a robustness analysis, we also estimated and discuss specifications that allow the impact of weather to vary across geographic regions.

Table 8 defines the main weather variables used in our analysis. *Wind* is the number of days in a week in which a wind advisory was issued by the National Weather Service Forecast

Office. A wind advisory is issued when maximum winds in a area achieve a threshold defined for that area, typically in excess of 40 miles per hour. *Rain* and *Snow* are the number of days in which the respective event is recorded in the week. We include *Wind*, *Rain* and *Snow* because each may influence travel to and from a plant. Although foggy conditions may also affect travel, we found some inconsistencies on how this weather variable was recorded and therefore decided to leave it out of the analysis.[4] *Heat* and *Cold* are the number of days in a week in which the extreme temperature for the day exceeds a threshold: 90 degrees Fahrenheit for *Heat* and 15 degrees Fahrenheit for *Cold*. *Heat* is included because it could influence ambient temperature within the plant or employees that must work outside, such as at the loading docks (e.g. Soper (2011)). *Cold* may proxy for hazardous road conditions (e.g., ice). Many of the variables, such as *Wind, Heat* and *Cold*, directly capture extreme weather shocks. To capture extreme events related to the other weather variables, we estimated specifications including multiple levels of the variable to capture potential non-linear effects on production (described in Section 3.4).

Table 9 shows summary statistics for the weather variables. We defined four regions that cover the locations of the plants in the study: Lakes, Central, Gulf and East, which are illustrated in Figure 8. (The plant in California, not shown in the figure, is included in the Gulf region.) The weather statistics are shown by region, and for some weather variables there are marked differences across regions (e.g. *Snow*). Table 10 shows a correlation matrix for the weather variables. Except for the higher correlation between *Cold* and *Snow*, all the correlations are less than 0.4 in magnitude. To check for potential multicollinearity, we regressed each weather variable on the others; the maximum R-square was less than 0.35, suggesting that multicollinearity is not a major concern in identifying the effect of the multiple weather measures in our study.

Note that *Rain* and *Snow* are measured in the number of days with rain and snow in that

[4] Between 1994 and 1996, several plants exhibited a frequency of fog that was orders of magnitude higher, which cannot be explained by changes in the weather patterns. This made the estimates of the effect of fog unstable, leading some plants to be highly influential in the estimation.

week. Alternatively, one could use cumulative precipitation to measure the intensity of rain and snow. However, our weather data only includes information about total precipitation, aggregating snow and rain precipitation together. Moreover, total precipitation was unavailable for some weeks in our sample, usually at the smaller weather stations. The precipitation data also appears to be subject to more measurement error: for example, the correlation for precipitation (measured in inches) across a sample of weather stations located 30-60 miles away is between 0.47 and 0.85, substantially lower than the other weather variables in our data (see footnote 3 for the sample). To summarize, we feel that the number of days of rain and snow is a more reliable measure to capture the effect of these weather shocks.

We include weekend observations in each weather variable even though plants are often (though not always) closed on weekends. This is appropriate if weather may have an effect on production that extends a few days before or a few days after the day in which it occurs. For example a weekend snow storm could influence deliveries both on Friday and especially on Monday. In addition, we are using the number of days of an event to proxy for the intensity of an event. A week with 7 days of rain is likely to be more extreme than a week with 5 days of rain. Similarly, a week with snow Friday through Sunday (i.e., three days of snow in our coding) may be more like a week with snow Wednesday through Friday (again, three days of snow in our coding) than a week with snow only on Wednesday (which is one day of snow in our coding, as the first example would be if we ignored the weekend). In addition, plants may attempt to recover lost production during weekdays by working on days off, but this recovery strategy would be limited if bad weather continues through the weekend (see Detroit (2011) for an example on how auto plants attempt to recover production in days off).

$PRODFACTORS$ includes covariates that capture adjustments to the production schedule and changes in productivity. Gopal et al. (2012) show that productivity is lower during the launch of a new model, so we include the dummy variable, *New Model,* that indicates the first 9 weeks during which a plant is producing a new model. We also include the dummy variable, *Drop Model,* to indicate the last 9 weeks before the production of the model is phased out. While *New Model* and *Drop Model* control for changes in productivity during the life-cycle of a model, temporary production stoppages of a model could also affect productivity. Assembly plants can be temporarily closed for several reasons, for example, due to holidays, plant re-tooling and also to adjust inventories of finished vehicles in the supply chain (Bresnahan and Ramey (1994)). Two dummy variables, *Prod Start* and *Prod Stop* indicate the week following and preceding a full stoppage of the plant, respectively. Note that all time-invariant factors affecting the productivity of the plant, such as plant capacity and proximity to suppliers, are captured by the fixed effect δ_i.

Using our data on scheduled production, we constructed a new variable capturing the total planned labor hours per week:

$$PLANHRS = Number\ Of\ Shifts \times Hours\ Per\ Shift$$

PLANHRS controls for scheduled shifts in production that may be associated with an anticipated reaction to weather. For example, *PLANHRS* controls for cases in which a plant schedules maintenance in a week in which they expect heavy snow. This may be viewed as a conservative approach as one could argue that if production is reduced due to scheduled maintenance in anticipation of bad weather, then there is indeed a causal effect of bad weather on production. However, it is possible that *PLANHRS* captures seasonality in production schedules that are not due to weather but still correlated with weather. (For example, the plant shuts down for a week in August for vacation no matter if that week turns out to be hot or not.) Hence, we include *PLANHRS* in our regressions.

As just mentioned, seasonality is an important potential confounder in our estimation. For example, seasonality in demand for new vehicles can lead to seasonal production patterns. If these seasonal production patterns are correlated with weather, then we cannot interpret the effect of weather in regression (3.1) as a causal effect on production. Hence, it is important to include controls in $SEASONAL_{it}$ that capture seasonality patterns in weather and production. These seasonal controls are discussed next.

The first set of controls for seasonality includes weekly dummy variables, τ_t, which control for seasonal production patterns and macro-economic effects affecting production of plants nation-wide. For example, this controls for differences in nation-wide plant productivity during different weeks of the year. But τ_t also controls for any nationwide-trends in production – such trends may be caused by economic shocks affecting aggregate demand for vehicles (e.g. oil prices). The weekly dummies also control for reduced working hours during national holidays. Note that if weather is perfectly correlated across plant locations, we cannot identify its effect separately from the weekly dummy τ_t. However, weather patterns vary substantially across regions. Figures 5 and 6 show two example that illustrate differences in local weather patterns across geographic regions– there is clearly more snow in the Lakes region than in the Gulf region. There is also some variation across plants within the same region – for example, there are differences in the number of *Wind* events among different plants in the East region. Hence, the inclusion of weekly dummies doesn't preclude the identification of the weather effects.

Because τ_t is common to all plants, it does not control for differences in seasonality or trends across plants. Therefore, the second set of controls that we use capture potential differences in seasonality across plants. In particular, we include region-specific year-month dummy variables, $\rho_{r(i)m(t)}$, where $r(i)$ is the pre-defined region where plant i is located, and $m(t)$ is the month of week t. This controls for monthly seasonality that could differ across regions (e.g., Spring arrives earlier in the year in the Gulf than in the Lakes region). We chose these regions because they have marked differences in their weather patterns; if regional production seasonality is correlated with weather patterns, omitting $\rho_{r(i)m(t)}$ from the regression would lead to biased estimates. In addition, we also include controls that capture potential differences in *demand* seasonality, which could thereby lead to different production patterns across plants. Specifically, we classified the production of each plant into one of the following segments: cars, vans, sport vehicles and pick-ups. If a plant is producing vehicles on multiple segments, we used the segment with higher production volume to classify the plant. The dummy variables $\psi_{s(i)m(t)}$, where $s(i)$ is the segment of plant i, control for these potential differences in production across plants.[5]

Two plants located in the same region and classified within the same segment could still have differences in their production patterns. If these patterns are related to weather then this could generate a bias in the causal effect we seek to estimate. To mitigate this kind of bias, we propose a third set of controls which captures seasonal average weather patterns specific to a plant. To explain the construction of these controls, let W_{it} be a weather-related variable (e.g. *Wind*) for plant i in week t and let $w(t)$ be week t's number within its year (e.g. the 54^{th} week in the sample is in week 2 of the second year). We define $\bar{W}(i, w(t))$ as the average weather at plant i during a 5 week time window around week $w(t)$ across all of the years in our sample:

$$\bar{W}(i, w(t)) = \frac{1}{5 \cdot N} \left(\sum_{y=0}^{N-1} \sum_{u=-2}^{2} W_{i,w(t)+52y+u} \right)$$

where $N = 10$ is the number of years in our sample. Hence, if there is correlation between production seasonality at a plant and the seasonality of any of our weather variables at that plant, this should be captured by $\bar{W}(i, w(t))$. We calculated these average weather measures for *Wind, Rain* and *Snow*. Notice that when we include this third set of controls in the model, the β coefficients for these weather variables are estimated using deviation from the weekly average at each plant.

3.4. Main results

Table 11 presents the first set of estimation results of regression (3.1). This specification includes all the seasonality controls: weekly dummies (τ_t), segment-month and region-month dummies ($\rho_{r(i)m(t)}$ and $\psi_{s(i)m(t)}$), and the average weather variables at each location

[5]Only four plants in our sample shifted their production from one segment to another.

$(\bar{W}(i, w(t)))$. (The estimates associated to these controls are not reported in the table for space considerations.) Among the weather variables, *Heat*, *Wind* and *Snow* are negative and statistically significant *(Cold* and *Rain* are not significant). We also estimated a specification with fewer seasonal controls – only the weekly dummies – and the results were similar, suggesting that the estimated effect of the weather variables is not driven by potential confounders related to seasonality. In addition, the controls for other production-related factors (grouped as $PRODFACTORS$ in regression (3.1)) are highly statistically significant; these suggest productivity drops associated with production ramp-ups and ramp-downs, and new model introductions. As expected, the total schedule hours ($PLANHRS$) during a week has a positive and significant effect on the number of vehicles produced.

A second specification, reported in Table 12, includes the weather variables in multiple levels to analyze the extent to which extreme weather events impact productivity. This specification also includes all the seasonality and production factor controls; all the coefficients of the production factors were similar to those in Table 11 and so they are omitted in the table. For *Rain* and *Snow* we include three levels based on the number of days the weather event occurred during the week. The cut-off points are indicated in the variable name and correspond to the 50^{th} and 95^{th} percentile of each measure, conditional on having at least one day of precipitation that week. In both cases, the effect of each level is relative to weeks with zero days of the respective precipitation (i.e. the excluded dummies are *Rain=0* and *Snow=0*). For example, *Snow[1]* indicates weeks with one day of snow and *Rain [3,5]* indicates weeks with 3,4 or 5 days of rain. We find empirical evidence that the effect of precipitation is non-linear for both rain and snow. One day of snow has no significant effect on production, but the effect is significant for 2 to 4 days of snow. The highest level of snow is also negative and larger in magnitude, but is not statistically significant at the 5% level (p-value=0.1), possibly due to the small number of observations for this extreme event (see Table 13). Nevertheless, we expect its impact should be as severe and we cannot reject the null hypothesis that the effect of *Snow[5,7]* is larger than *Snow[2,4]*. For rain, the effect is statistically significant for 6 or more days of rain, but not significant for fewer days of rain.

We defined three levels for *Heat*, and *Cold*. The highest level for heat, 6 or 7 days with a high temperature exceeding $90°F$, is closely related to the definition of a heat wave.[6] We find strong evidence of a non-linear effect of *Heat*, but the effect of *Cold* is still insignificant at all levels.

Because days with *Wind* advisory alert are relatively infrequent (See Table 13), levels for this variable were defined based on thresholds of wind-speed. Two levels were defined with cut-offs at 34 and 44 mph, and each level counts the number of days with maximum wind speed on each level's range. For example, *Wind[35,44]* counts the number of days with wind speed between 34 and 44 mph. The results suggest evidence of non-linear effects of *Wind.* Next, we describe the economic significance of these results.

For all the variables reported in Table 12, except for the *Wind* variables, the coefficient represents (approximately) the percentage drop in weekly production when the corresponding weather event occurs during a given week (net of any production recovery that might occur in that week). For *Wind,* the coefficient measures the percentage drop in weekly production of an *additional day* with the indicated wind speed. To put the effect of weather in perspective, the productivity reduction during the first week a vehicle model is introduced is 32%, similar in magnitude to the combined effect of one day of high wind, a heat wave with 6 or more days of high temperature and 6 or more days of rain during a week. But such extreme weather incidents are also rare – for example, weeks with wind-speeds above 44 mph have a frequency of 0.6% in the sample. To estimate the economic impact, we measure the expected production reduction which combines the likelihood of the weather incident with the impact estimated in Table 12. Table 13 reports these calculations for the weather variables that have a statistically significant effect on production as reported in Table 12 along with *Snow[5,7]*, as we cannot reject the null hypothesis that the effect of *Snow[5,7]* is larger than *Snow[2,4]*. Here, we see that snow and rain tend to have the largest economic

[6]The Warm Spell Duration Index (WSDI) – commonly used to characterize the frequency of heat waves – is defined as the fraction of days belonging to spells of at least 6 days with maximum temperature exceeding the 90th percentile (Field et al. (2012)).

effect on weekly production.

Based on the average weather variables observed at each location, we calculated the average percent drop in productivity due to weather shocks for each plant location (this calculation considers all the weather variables included in regression (3.1)). Table 14 shows the results for the 49 cities in our sample. (Plants in the same city have the same weather and therefore the same effect.) Table 16 shows the average percent loss in productivity due to weather for each of the four regions. While the average loss is not statistically different across regions it is possible to observe a statistically significant difference for the impact of snow and heat across the different regions.

Regression Diagnostics and Robustness Analysis

We conducted a series of regression diagnostics to analyze the robustness of our results. To check the generalizability of the results to other time periods, we expanded our dataset to include production from 2006 to 2009 using data provided by Automotive News. In 2006, manufacturers stopped reporting weekly production and moved to monthly production reports. Automotive News interpolated weekly production based on monthly production and information on shift patterns, parts shortages, etc. Because we view these data as less reliable we do not use them for our main results, but they are useful as a robustness check. All of the results are qualitatively similar to the period 1994-2009, but some of the coefficients are estimated with less precision and are not significant (specifically *Heat*). This is consistent with the larger measurement error associated with the dependent variable for those additional years.

Because some of the weather events are infrequent, we checked for influential points in the data. To do this, we re-estimated the model removing each of the plants (one-at-a-time), and found no significant difference in our results. We conclude that the estimates are not driven by influential locations in the sample.

As the nearest weather stations to the plants are located on average 13 miles away from the plants, a potential concern is measurement error with our weather variables. To address this issue, we estimated the regressions using only plants with corresponding weather stations within 25 miles of the plant. The sample size in this regressions drops to about 26,000 observations. All the results are similar in magnitude and statistical significance, and all the point estimates are within the 95% confidence interval of the results reported in Table 12. This analysis alleviates concerns with potential measurement error in the dependent variable due to the location of weather stations.

The measures of weather used in our main analysis include weather events on weekends even though most plants do not work on weekends. This is reasonable if adverse weather can have an impact on production just before or after the actual weather event. Furthermore, including weekends allows us to better use the duration of the event for a proxy of its intensity. Nevertheless, we estimated our model with weekend weather excluded and found that the results were consistent with those reported in Table 12.

Plant Heterogeneity

Our results provide a measure of the average impact of weather on automobile production. It is possible that individual plants may experience different effects depending on their idiosyncratic features, such as the location of the parts suppliers, or inventory management practices or other operating procedures. For example, while we have measured the impact of severe heat on plants A and B, given the same level of heat, plant A may experience less of an adverse reaction than plant B. As long as the magnitude of the impact of a weather event on a plant is uncorrelated with the frequency of the event at that location, the estimates of the economic impact reported in Tables 13 to 16are unbiased. However, if plant location decisions are endogenous so that plants for which the effect of a weather event is larger are located in areas with lower frequency of these events, then our estimates would overestimate the average economic impact on production even though the estimated impact conditional on an event occurring, i.e., the β coefficients, remain unbiased. This potential bias can be corrected by accounting for the heterogenous effect of weather across plants.

Since it is not possible to estimate a separate coefficient for each plant (the estimates would be too imprecise), we instead categorize plants into groups and estimate a different vector of coefficients for each group. The idea is to group plants based on their weather similarities, so that weather patterns are similar within group but different across groups. If there is any selection based on the incidence of weather events, then one should observe differences in the estimated coefficients across groups.

We conducted a hierarchical cluster analysis to segment plants into groups. Let $\bar{X}_{kiq}$ denote the average incidence of weather variable k at plant i during quarter q (using the weather variables defined in Table 8), and $\bar{X}_i$ the vector containing all these weather metrics that characterize a plant i. The cluster analysis calculates the distance between plants based on these metrics, generating a partition of plants into groups. We used Ward's hierarchical clustering method to create the groups (see Johnson and Wichern (1992) for details of the method). For the regression analysis, we considered using two clusters which are shown in Figure 7. There is a clear geographic segmentation of the two groups, which we name the North and South clusters.

We estimated regression (3.1) including interactions of the weather variables and an indicator variable for the South cluster. The results of this analysis are presented on Table 17 (the first column, "Main Results", shows the original estimates for comparison; the interactions are labeled "SC"). Given the larger number of coefficients to estimate, the standard errors increase and many of the variables are no longer statistically significant. We focus in testing the null hypotheses of equal coefficients between the North and South clusters, which can be done by testing the significance of the interactions. These results show a difference on the coefficients estimated for the highest level of *Rain* – the effect tends to be higher in magnitude for the South cluster – and no significant difference for the other coefficients. A possible explanation for this difference is that on average the South locations receive 10 inches more of rain per year than the North locations (44 vs 34 inches) even though rain in the South is about as frequent as in the North. Overall, the differences in the coefficients are observed for weather variables whose frequency is similar across the two groups: about 60% of the *Rain[6,7]* events happened in the south cluster. Although there is some heterogeneity across plants, it is not systematically related to the frequency of extreme weather events, and so we conclude that the average economic impact deduced from Tables 13 and 16 are correct.

To the extent that these differences between plants exist, it is worthwhile to know if they are associated with managerial decisions. Unfortunately, while our data is well suited for measuring the average impact, because we have heterogeneity in weather across different plants, it is not particularly well suited for identifying practices that are more or less robust to weather disruptions. To explain, to understand if plant A is more robust to weather than plant B, ideally we want them to have similar weather, or at least weather that is uncorrelated with the practices that make them different. Most of the plants in our sample are located far away from other plants, so few plants have similar weather. Furthermore, we lack data on the specific relevant operational characteristics that could be used to infer differences across plants. Put another way, our panel data is appropriate for identifying the average impact of weather of automobile production, but to understand differences across plants requires a cross-section analysis and that introduces a host of identification challenges. Nevertheless, we can make some initial exploration based on our data.

It is possible that plants owned by GM, Ford and Chrysler (labeled the US group) operate in a different way than all other plants (non-US group). For example, they may be more unionized or use fewer "lean manufacturing" techniques (Bennett et al. (2011) provides some anecdotal evidence on how lean plants may be more prone to disruptions). To test for differences in these two groups, we estimated regression (3.1) including interactions of the weather variables and a binary variable indicating the non-US group. Again, the estimated coefficients on this analysis are measured with less precision. Interestingly, the results seem to replicate the North/South segmentation reported in Table 17: the *Rain* appears to have a larger impact on plants in the non-US group. Nevertheless, we do not wish to conclude that U.S. plants are better able to cope with *Rain* because of their managerial practices. U.S. plants tend to be located more in northern regions and non-US plants are more prevalent in southern regions (see Figure (8)). Consequently, the differences we observe could be due to differences in the nature of weather in the north relative to the south. For example, six days of rain in Tennessee (which has a non-US plant) may be more intense than six days of rain in Michigan, which is dominated by US plants (as reported earlier, rain tends to be more intense in the south). Therefore, those results may suggests a north/south difference rather than a US/non-US difference.

To further explore this issue, we identified sets of plants which are collocated within 100 miles and have different ownership (US vs non-US). Four pairs of US/non-US plants were identified. We estimated regression (3.1) with interactions with the non-US group indicator. This regression has little power due to the small sample size, and none of the coefficients are statistically significant. Hence, we believe that with our data and estimation strategy it is not possible to determine if US plants are differentially robust to weather relative to non-US plants or if southern weather is different than northern weather in ways that our main regression does not capture. Put another way, we cannot provide evidence that our average effects are different between US and non-US plants.

Short-term production recovery

Another question of interest is the extent to which plants are able to recover from the short term productivity losses we observe due to weather shocks. At one extreme, plants may be able to recover all of their lost production at some point in the future. Even if this is true, the short term productivity losses would be costly as they can lead to stockouts at dealerships and to volatile production (which could require costly overtime). To further explore the extent of recovery, we analyzed how weather incidents could impact production in the week after the time the incident occurs. Specifically, we estimated regression (1) using “lagged” weather variables. Table 18 shows the results of this analysis. For reference, column (1) reports the estimates of Table 12 and column (2) includes the weather variables that were significant on the main analysis with one week of lag. For the most part, the results when we include the lag variables are similar in sign, magnitude and significance relative to the weekly analysis. In addition, the lagged effects for *Rain [6,7]*, *Snow [5,7]* and *Wind >44* are negative and significant. Not only does this contradict the hypothesis that plants are able to recover their production in the following week of bad weather, it suggests that bad weather may have an impact beyond the week it occurs. Alternatively, it may due to how we code weather events - a six day period of rain that straddles two weeks is probably one weather event, but because we divide time into weeks, it is viewed as two weather events in our analysis. Either way, we do not find evidence suggesting that firms recover their lost production in the week immediately following an adverse weather event. We also considered specifications that added further lags, but these were jointly insignificant.

To further analyze production recovery after a weather incident, we estimated the impact of weather on the likelihood of scheduling overtime during the weeks after the incident, as overtime is a likely mechanism to recover lost production. We defined an indicator variable that is equal to one if the plant scheduled overtime during the three weeks following any week t. We estimated a Probit regression of this indicator variable, including the weather factors and all the other independent variables of regression (3.1) as covariates. The estimates suggest that the none of the factors have a significant influence on the probability of overtime (p-values<0.05) The evidence suggest that the production schedule is rarely affected in the weeks immediately following a weather incident.

Although we do not find evidence of a short term recovery, we cannot rule out that recovery occurs with a greater lag. However, plants may choose different lags for recovery - e.g., some may recover in four weeks while others within eight weeks, and even the same plant may take a different amount of time to recover from different events. Hence, it is difficult to use our data to identify this recovery, if it occurs. One possible approach is to aggregate data over time. For example, regress quarterly production on quarterly weather. If recovery occurs within the quarter, we should not observe significant relationships between weather and production. However, this substantially reduces the sample size, thereby complicating the interpretation of the results - failing to reject that null that weather has no adverse impact on production is not the same as accepting the null. Clearly, further research is needed on the long term impact of weather on production. But even if long run recovery is possible, we can conclude that adverse weather has a short term impact on productivity, and therefore leads to a costly increase in the volatility of production.

3.5. Conclusion

Based on our sample of U.S. automobile assembly plants over a ten-year period, we find that a plant's local weather can have a substantial impact on production, ranging from a reduction of 0.5% to 3.0%, with an average of 1.5%. The immediate follow-on question is "Can automobile companies do a better job managing this problem?". The answer depends on the underlying mechanisms. Given that we find heavy winds, snow and rain are associated with production losses, it is possible that disruptions to in-bound deliveries is a major cause. If this is the case, firms could mitigate this factor by carrying more inventory of parts or at least increasing deliveries of parts in anticipation of bad weather. This approach goes against the "just-in-time" philosophy of carrying lean inventory and ensuring a smooth production flow, but avoiding the productivity losses due to weather may justify a more flexible operating strategy. If, on the other hand, bad weather is problematic because it increases employee absenteeism, then mitigating strategies may be more difficult to develop. For example, it would be costly to "pre-position" workers in anticipation of bad weather - people are not likely to want to live at the plant for an extensive period. However, it may be possible to provide employees with alternative transportation options (company operated shuttles), as long as these transportation options are available during poor weather.

We find that high temperatures reduce production. The obvious mitigating strategy for heat is to provide cooling systems. It is possible that heat is influencing worker productivity in "interface" areas between the outside and inside environments, such as on loading and unloading areas, because these areas may be difficult to cool. Alternatively, if the ambient temperature outside is significant, then it is possible that existing cooling systems are unable to maintain the interior temperature under $77^\circ\ F$ (a threshold for heat stress). If this is the case, then maybe an investment in higher capacity cooling systems could be justified.

It is not clear the extent to which automobile companies are aware of the impact of weather on their productivity beyond obvious effects like "a blizzard can disrupt production". About half of companies in a survey, Staff (2011), report that they experienced a weather related disruption to their supply chain, but magnitudes were not estimated and our results suggest that nearly all facilities may experience some form of weather disruption. If they are indeed not aware, then it is possible that the mitigating strategies discussed above (or others) could improve productivity. But if they are already aware of these effects, then they may have already implemented all cost effective mitigating strategies. That would leave only the option to move production to a more weather friendly location. Of course, moving production is costly and raises a host of other issues - labor costs, access to suppliers, etc.

Our study focuses on the automobile industry, which offers several advantages: it is an economically significant industry, there are many geographically dispersed assembly plants operated by a number of different companies, and detailed production data is available over a long period of time (ten years) at the weekly level (rather than monthly, quarterly or annually). However, it is not clear to what extent these results carry over to other industries. Again, the answer depends on the underlining mechanism. If disruptions in in-bound parts deliveries are the cause of the productivity loss, then these effects are likely to occur in any manufacturing industry that operates with limited buffer stocks of inventory. Industries that carry substantial inventory are probably more robust. But if the cause is due to disruptions in in-bound employees, then these effects are likely to be common across many industries, including services. Additional data are needed to tease out which of the mechanisms we have identified (or others) are responsible for these effects.

Our findings provide an interesting contrast with the existing literature on climate change and economic activity. For example, Dell et al. (2008) find that hot years only impact poor countries, but we find that hot temperatures impact production in a "rich" country. Furthermore, they find that rainy years neither impact poor nor rich countries but we find that intense periods of rain do negatively affect productivity. Similarly, Hsiang (2010) find that adverse weather actually increases manufacturing output in Caribbean basin countries. But those studies work with annual shocks (e.g., a hot year) and annual output measures across a wide range of industries. It is possible that their level of aggregation masks productivity losses in specific industries. Furthermore, because their estimation is based on annual shocks, they are unable to measure short term shocks (e.g., weekly shocks) that nevertheless add up to a substantial annual impact - if the frequency of short term shocks is relatively constant, then there may not be enough variation in annual data to identify their effect (e.g., if there are 5 windy weeks each year and every year, the effect of wind cannot be estimated with annual data).

Finally, our work provides additional evidence on the impact of climate change on economic output. Climate change is forecasted to be associated with increases in severe weather (Field et al. (2012)), in particular with heat and rain, and we find a direct link between severe weather (high winds, high heat, and extensive periods of snow or rain) and productivity losses. Long run forecasts of extreme weather are challenging and there can be uncertainty in the direction of the change (e.g., wind) as well as the magnitude of the change (e.g., temperature). Hence, even though we are not comfortable combining our estimates of productivity losses with extreme weather forecasts to yield a long fun forecast of potential losses in the North American automobile industry due to climate change, we believe the impact of weather on manufacturing productivity is likely to be a growing concern.

Data Sources

Automotive News, Crain Communications, 1995-2009.

National Oceanic and Atmospheric Administration, NOAA, 1994-2009.

The Weather Channel, www.weather.com, 1994-2009.

Ward's AutoWorld, Ward's Auto, 1994–2005.

Weather Underground, www.wunderground.com, 1994-2009.

Table 7: Descriptive statistics of assembly plants in the study.

Company	Number of plants	Average weekly production (vehicles/plant) (1)	Minimum weekly production (vehicles/plant) (2)	Maximum weekly production (vehicles/plant) (3)	Average utilization (4)
GM	20	4048	231	13155	74%
FORD	16	4547	202	12400	75%
CHRYSLER	9	4666	560	9359	74%
TOYOTA	5	4769	663	12165	76%
HONDA	4	5273	698	11100	74%
ISUZU	2	4031	609	6798	76%
MAZDA	2	3372	874	7382	75%
BMW	1	1640	201	3932	73%
HYUNDAI	1	2516	800	4520	56%
MB	1	1423	223	1990	77%
MITSUBISHI	1	3410	614	5821	75%
NISSAN	1	4800	1619	9165	65%
SUZUKI	1	8270	1814	12972	79%

(1) The average is taken over the companies plants' average weekly production.

(2) This is the minimum number of units produced during a week among all of the company's plant.

(3) This is the maximum number of units produced during a week among all of the company's plant.

(4) To calculate this value, we first obtained the utilization for each plant during each year in our sample as the average production divided by the maximum production value. Then we average across each plant and finally obtain the average across each company.

Table 8: Weather variables included in the empirical study

Variable	Description
Wind	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 typically in excess of 40 miles per hour.
Rain	Number of days with rain during the week.
Snow	Number of days with snow during the week.
Heat	Number of days with a high temperature above 90 degrees Fahrenheit.
Cold	Number of days with low temperature below 15 degrees Fahrenheit.

Table 9: Mean and standard deviation (in parentheses) of the weather variables, by geographic region.

	Central	East	Gulf	Lakes	Total
Wind	0.006	0.010	0.011	0.007	0.007
	(0.076)	(0.101)	(0.106)	(0.083)	(0.087)
Rain	2.508	2.804	2.678	2.334	2.507
	(1.848)	(1.804)	(1.892)	(1.793)	(1.841)
Snow	0.490	0.264	0.065	0.870	0.518
	(1.128)	(0.712)	(0.336)	(1.524)	(1.193)
Heat	0.480	0.428	1.001	0.211	0.483
	(1.299)	(1.128)	(1.997)	(0.708)	(1.326)
Cold	0.382	0.176	0.038	0.652	0.390
	(1.168)	(0.724)	(0.294)	(1.557)	(1.206)

Table 10: Correlation matrix of weather variables.

	Wind	Rain	Snow	Heat	Cold
Wind	1.000				
Rain	0.021	1.000			
Snow	-0.009	-0.357	1.000		
Heat	0.009	0.013	-0.161	1.000	
Cold	-0.009	-0.314	0.599	-0.117	1.000

Table 11: Estimation results of regression (3.1).

Production factors		Weather		Additional Controls
Prod. Start	-0.1006***	*Heat*	-0.0127***	*Week*
	(0.0257)		(0.0037)	*Region-Month*
Prod. Stop	-0.0578*	*Cold*	0.0004	*Segment-Month*
	(0.0240)		(0.0043)	*Avg. Weather*
New Model	-0.3236***	*Wind*	-0.0800*	
	(0.0202)		(0.0337)	
Drop Model	-0.0145	*Rain*	-0.0033	
	(0.0121)		(0.0023)	
PLANHRS	0.7272***	*Snow*	-0.0138***	
	(0.0167)		(0.0043)	

Number of observations=31,174. R-square=0.61. Robust Standard errors in parentheses.
$^{*}\ p < 0.05$, $^{**}\ p < 0.01$, $^{***}\ p < 0.001$

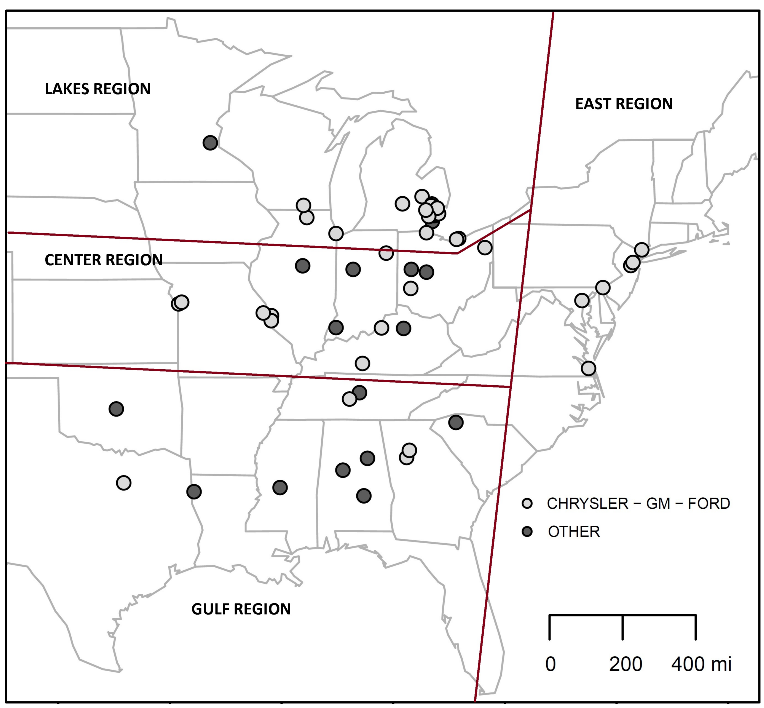


Figure 4: Plant locations and geographic regions (The plant in Fremont, CA, — not shown — is classified within the Gulf region)

Table 12: Estimation results of regression (3.1) including levels of the weather variables

Precipitation		Temperature		Wind		Additional Controls
Snow [1]	0.0019	*Heat [1]*	-0.0065	*Wind [35,44]*	-0.0196	*Week*
	(0.0106)		(0.0163)		(0.0128)	*Region-Month*
Snow [2,4]	-0.0278*	*Heat [2,5]*	-0.0273	*Wind ¿44*	-0.0791*	*Segment-Month*
	(0.0133)		(0.0155)		(0.0339)	*Avg. Weather*
Snow [5,7]	-0.0429	*Heat [6,7]*	-0.0875**			*Production Factors*
	(0.0270)		(0.0291)			
Rain [1,2]	-0.0053	*Cold [1]*	-0.0035			
	(0.0101)		(0.0176)			
Rain [3,5]	0.0039	*Cold [2,5]*	-0.0020			
	(0.0117)		(0.0183)			
Rain [6,7]	-0.0590**	*Cold [6,7]*	-0.0212			
	(0.0182)		(0.0297)			

Number of observations=31,174. R-square = 0.61. Robust Standard errors in parentheses.

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table 13: Frequency and economic impact of weather variables

Weather incident	Frequency (per week)	Average production reduction (weekly)
Snow [2,4]	11.8%	0.34%
Snow [5,7]	2.1%	0.09%
Rain [6,7]	6.3%	0.37%
Heat [6,7]	1.7%	0.14%
Wind ¿44	0.6%	0.05%

Table 14: Ranking of average productivity reduction due to weather by location

Rank	*City*	*State*	*Total productivity loss (%)*	*Snow loss (%)*	*Rain loss (%)*	*Temp. loss (%)*	*Wind loss (%)*
1	Montgomery	AL	2.88%	0.00%	0.34%	2.45%	0.10%
2	Arlington	TX	2.41%	0.01%	0.52%	1.71%	0.17%
3	Shreveport	LA	2.18%	0.02%	0.48%	1.56%	0.12%
4	Canton	MS	1.93%	0.00%	0.53%	1.40%	0.00%
5	Avon Lake	OH	1.83%	0.74%	0.54%	0.22%	0.33%
6	St Paul	MN	1.81%	1.03%	0.23%	0.41%	0.14%
7	Oklahoma City	OK	1.81%	0.10%	0.17%	1.23%	0.30%
8	Lorain	OH	1.80%	0.78%	0.45%	0.25%	0.33%
9	Warren	OH	1.78%	1.00%	0.44%	0.21%	0.13%
10	Roanoke	IN	1.77%	0.79%	0.43%	0.29%	0.25%
11	Hazelwood	MI	1.70%	0.35%	0.50%	0.70%	0.16%
12	Lansing	MI	1.66%	0.93%	0.38%	0.27%	0.08%
13	Toledo	OH	1.65%	0.70%	0.42%	0.35%	0.18%
14	Vance	AL	1.63%	0.03%	0.35%	1.10%	0.16%
15	Wayne	MI	1.63%	0.76%	0.35%	0.26%	0.26%
16	Edison	NJ	1.61%	0.28%	0.70%	0.42%	0.21%
17	Linden	NJ	1.59%	0.28%	0.67%	0.38%	0.26%
18	Fenton	MO	1.58%	0.36%	0.32%	0.73%	0.17%
19	Smyrna	TN	1.57%	0.19%	0.54%	0.74%	0.10%
20	Flint	MI	1.55%	0.92%	0.26%	0.28%	0.09%
21	Spring Hill	TN	1.52%	0.17%	0.54%	0.73%	0.08%
22	Lake Orion	MI	1.50%	0.87%	0.26%	0.28%	0.10%
23	Baltimore	MD	1.50%	0.17%	0.67%	0.49%	0.18%
24	Wentzville	MO	1.48%	0.37%	0.27%	0.65%	0.19%
25	Sterling Heights	MI	1.45%	0.98%	0.20%	0.27%	0.00%

Table 15: Ranking of average productivity reduction due to weather by location (continued)

Rank	*City*	*State*	*Total productivity loss (%)*	*Snow loss (%)*	*Rain loss (%)*	*Temp. loss (%)*	*Wind loss (%)*
26	Norfolk	VA	1.44%	0.09%	0.70%	0.47%	0.17%
27	Moraine	OH	1.42%	0.56%	0.38%	0.29%	0.19%
28	Wixom	MI	1.41%	0.92%	0.20%	0.25%	0.03%
29	Belvidere	IL	1.40%	0.58%	0.36%	0.33%	0.13%
30	Spartanburg	SC	1.39%	0.02%	0.62%	0.69%	0.05%
31	Janesville	WI	1.36%	0.62%	0.29%	0.30%	0.15%
32	Kansas City	MO	1.36%	0.28%	0.33%	0.60%	0.15%
33	Louisville	KY	1.33%	0.32%	0.36%	0.53%	0.12%
34	Kansas City	KS	1.33%	0.28%	0.35%	0.55%	0.14%
35	Bowling Green	KY	1.31%	0.19%	0.36%	0.60%	0.16%
36	Pontiac	MI	1.30%	0.83%	0.17%	0.26%	0.04%
37	Lafayette	IN	1.30%	0.43%	0.30%	0.36%	0.20%
38	Lincoln	AL	1.30%	0.03%	0.31%	0.93%	0.03%
39	Georgetown	KY	1.29%	0.30%	0.51%	0.39%	0.08%
40	Normal	IL	1.27%	0.42%	0.29%	0.48%	0.08%
41	Chicago	IL	1.22%	0.60%	0.11%	0.39%	0.12%
42	Marysville	OH	1.19%	0.47%	0.20%	0.28%	0.23%
43	Atlanta	GA	1.15%	0.01%	0.42%	0.55%	0.17%
44	Warren	MI	1.15%	0.67%	0.16%	0.26%	0.06%
45	Wilmington	DE	1.15%	0.14%	0.36%	0.38%	0.27%
46	Dearborn	MI	1.15%	0.66%	0.16%	0.26%	0.06%
47	Detroit	MI	1.14%	0.65%	0.17%	0.26%	0.06%
48	Fremont	CA	0.81%	0.00%	0.61%	0.16%	0.04%
49	Princeton	IN	0.46%	0.05%	0.05%	0.33%	0.03%
	Average		1.50%	0.43%	0.37%	0.56%	0.14%

Table 16: Average productivity reduction due to weather by region

Region	Average productivity loss *(%)*	Average productivity loss due to Snow *(%)*	Average productivity loss due to Rain *(%)*	Average productivity loss due to Heat *(%)*	Average productivity loss due to Wind *(%)*
Central	1.40%	0.45%	0.34%	0.45%	0.16%
East	1.46%	0.19%	0.62%	0.43%	0.22%
Gulf	1.72%	0.05%	0.45%	1.10%	0.11%
Lakes	1.45%	0.79%	0.26%	0.29%	0.11%

Table 17: Estimation results considering two weather clusters

	Main Results	Weather Clusters Results		
		Main Effects	Interactions	
Snow [1]	0.0019	0.0011	*SC*Snow [1]*	0.0143
	(0.0106)	(0.0139)		(0.0189)
Snow [2,4]	-0.0278*	-0.0242	*SC*Snow [2,4]*	-0.0038
	(0.0133)	(0.0162)		(0.0231)
Snow [5,7]	-0.0429	-0.0457	*SC*Snow [5,7]*	-0.0036
	(0.0270)	(0.0296)		(0.0627)
Rain [1,2]	-0.0053	-0.0046	*SC*Rain [1,2]*	-0.0007
	(0.0101)	(0.0142)		(0.0191)
Rain [3,5]	0.0039	-0.0017	*SC*Rain [3,5]*	0.0133
	(0.0117)	(0.0160)		(0.0199)
Rain [6,7]	-0.0590**	0.0013	*SC*Rain [6,7]*	-0.0963***
	(0.0182)	(0.0263)		(0.0307)
Heat [1]	-0.0065	0.0025	*SC*Heat [1]*	-0.0163
	(0.0163)	(0.0260)		(0.0297)
Heat [2,5]	-0.0273	0.0061	*SC*Heat [2,5]*	-0.0477
	(0.0155)	(0.0288)		(0.0289)
Heat [6,7]	-0.0875**	-0.0212	*SC*Heat [6,7]*	-0.0727
	(0.0291)	(0.1181)		(0.1203)
Cold [1]	-0.0035	0.0090	*SC*Cold [1]*	-0.0298
	(0.0176)	(0.0207)		(0.0307)
Cold [2,5]	-0.0020	0.0102	*SC*Cold [2,5]*	-0.0378
	(0.0183)	(0.0215)		(0.0259)
Cold [6,7]	-0.0212	-0.0057	*SC*Cold [6,7]*	-0.1825
	(0.0297)	(0.0294)		(0.1284)
Wind [35,44]	-0.0196	-0.0162	*SC*Wind [35,44]*	-0.0081
	(0.0128)	(0.0178)		(0.0239)
Wind ¿44	-0.0791*	-0.0779	*SC*Wind ¿44*	-0.0041
	(0.0339)	(0.0438)		(0.0660)
Additional controls				
Region-Month	YES	YES		
Segment-Month	YES	YES		
Avg. Weather	YES	YES		
Cold (in levels)	YES	YES		
Production Factors	YES	YES		

Number of observations=31,174. R-square = 0.61. Robust Standard errors in parentheses.
* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

"SC" = South Cluster

Table 18: Estimation results including lagged effects for the weather variables.

	Main Results	Including Lags		
		Main Effects	Lagged Variables	
Snow [1]	0.0019	0.0015		
	(0.0106)	(0.0105)		
Snow [2,4]	-0.0278*	-0.0287*	*Lagged* Snow *[2,4]*	-0.0199
	(0.0133)	(0.0130)		(0.0121)
Snow [5,7]	-0.0429	-0.0416	*Lagged Snow [5,7]*	-0.0608*
	(0.0270)	(0.0272)		(0.0280)
Rain [1,2]	-0.0053	-0.0015		
	(0.0101)	(0.0102)		
Rain [3,5]	0.0039	0.0088		
	(0.0117)	(0.0118)		
Rain [6,7]	-0.0590**	-0.0448*	*Lagged Rain [6,7]*	-0.0529***
	(0.0182)	(0.0183)		(0.0159)
Heat [1]	-0.0065	-0.0063		
	(0.0163)	(0.0164)		
Heat [2,5]	-0.0273	-0.0238		
	(0.0155)	(0.0158)		
Heat [6,7]	-0.0875**	-0.0759*	*Lagged Heat [6,7]*	-0.0431
	(0.0291)	(0.0292)		(0.0265)
Wind [35,44]	-0.0196	-0.0195		
	(0.0128)	(0.0127)		
Wind > 44	-0.0791*	-0.0793*	*Lagged Wind >44*	-0.1364**
	(0.0339)	(0.0338)		(0.0503)
Additional controls				
Region-Month	YES	YES		
Segment-Month	YES	YES		
Avg. Weather	YES	YES		
Cold (in levels)	YES	YES		
Production Factors	YES	YES		
Observations	31174	30712		
R-square	0.6126	0.6166		

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

Figure 5: Wind map (The scale on the map corresponds to the total number of high wind days at each location during a 10 year period)

Figure 6: Snow map (The scale on the map corresponds to the total number of weeks with more than five days of snow at each location during a 10 year period)

Figure 7: Weather-based clusters (The plant in Fremont, CA, — not shown — is classified within the South cluster)

CHAPTER 4 : Store pick-up

4.1. Introduction

Online retailing has grown steadily over the last few years. Some retailers operate exclusively through online channels, and traditional brick and mortar (B&M) retailers have incorporated online sales channels since the early stages of the commercial Internet (e.g., the Barnes and Noble website launched in May 1997). Today, retailers' online channels are no longer an experiment but a relevant and growing part of their business. Originally, most of the B&M retailers decided to separate the operations of traditional and online channels. Now, some B&M retailers are exploring integration strategies for their online and B&M channels to enrich the customer value proposition and/or reduce costs. Online-offline integration efforts can occur in a variety of configurations. For example, B&M retailers often show in-store inventory availability information online. More advanced integration includes shipping the product ordered from the store closest to its destination, or offering the option to buy products online and pick them up in the store.

In particular, over the last few months, a number of traditional B&M retailers across different categories (e.g., The Home Depot, Apple, Crate & Barrel, Toys "R"Us, among others) have implemented buy-online-pickup-in- store (BOPS) functionality. The retailer shows online viewers the locations at which the item is available, and gives customers the option to close the transaction online and then pick up the product at one of the locations within two hours of closing the purchase[1].

The integration of online and offline channels provides an opportunity to empirically study issues that have been the subject of theoretical research in operations management. In this paper, we use an online-offline integration project that implements BOPS functionality as a natural experiment to study the impact of sharing reliable inventory availability infor-

[1]Most retailers announce that they need a two hour window to have the item ready to pick up. In some cases this time can be less but two hours is representative of the most common commitment. This short lead-time restricts the retailer to fulfill the order with in store inventory

mation with the customers. Implementing a buy-online-pick-up-in-store project provides an exogenous shock to the verifiability of the inventory information that the firm shows to their customers; because the inventory information becomes more credible, the risk that customers face when deciding whether to visit the store is reduced.

We have collected a novel proprietary dataset from a nationwide retailer that has been among the pioneers in implementing BOPS functionality. Using this dataset and a series of natural experiments, we make the following contributions:

First, we evaluate the impact of BOPS implementation on company sales and customer behavior, and give the first piece of empirical evidence on this emerging trend in retailing[2]. We study the impact of the deployment of a BOPS project on both the online and brick and mortar channels. Conventional wisdom within the industry suggests that offering the BOPS functionality will improve online channel revenue (since BOPS transactions are considered online revenue), and that the traditional B&M stores will carry the burden of having the item ready for the customers to pick up, without receiving any significant benefit in their sales. However, as we will describe in detail, a series of natural experiments leads us to conclude that these assumptions are not correct. Our results show that, contrary to what we would expect, sales transacted online *decrease* and B&M sales *increase* when the BOPS functionality is deployed.

Second, we show how the increase of inventory information verifiability affects customer behavior. The impact of availability information and its verifiability on customer behavior has been the subject of recent modeling research in the field of operations management (e.g., Allon and Bassamboo (2011); Su and Zhang (2009)) but to our knowledge, no empirical results were available prior to this paper. Implementing BOPS functionality can be seen as a shock to the verifiability of inventory information online. To implement BOPS functionality, the online system must have access to accurate real-time information about availability of

[2]See, for example, http://operationsroom.wordpress.com/2012/05/16/macys-warehouse-at-the-mall/ and http://operationsroom.wordpress.com/2010/08/25/pooling-inventory-at-nordstrom/

in-store inventory. If the retailer offers the option to pick up an online order at a particular store, the customer knows with very high certainty that the item ordered is available at that store. Therefore, inventory availability information is perceived as very reliable. This contrasts with situations whereby the store simply shows inventory information but does not offer the option to close the transaction online. For example, consider a car dealership showing information online about their inventory. This information is typically unverifiable; if a customer visits the dealer and the product is not available, the dealer can claim that the online information was not updated in real time. We find that increased reliability of in-store availability information increases the probability that customers will visit the store. We present an explanation consistent with empirical evidence we observed regarding the impact of BOPS functionality: Providing BOPS functionality increases the reliability of the inventory information, resulting in an increase in the number of customers visiting the stores to purchase items after checking product availability online. This provides an explanation to the counterintuitive finding described above. We further check the validity of this explanation by presenting further evidence from the shopping cart abandonment behavior.

Finally, we use this project as an example of the evaluation of an online-offline strategy, illustrating the complex interactions between the online and offline channels and the challenges of relying on single channel data to evaluate the impact of interventions that affect multiple channels. Retailers often run experiments in their online channel (for example, A/B testing) to evaluate the impact of interventions on their conversion rates or other measures of interest. In our case, an isolated evaluation of the online channel would have considered the impact of the BOPS implementation to yield negative results. Only when closing the loop and looking at the effects in the brick and mortar channel we can quantify the net effects of the BOPS implementation, which are positive.

The rest of the document is organized as follows: Section 2 reviews the literature related to our problem of interest; Section 3 describes the empirical setting and data; Section 4 shows the impact of the deployment of BOPS functionality on the online channel and the brick-an-mortar channel. Section 5 provides an interpretation of the results based on information verifiability and tests the validity of this interpretation with additional analyses. Finally, Section 6 concludes by highlighting the managerial insights that can be drawn from our analysis.

4.2. Context and Related Literature

Integration of online and offline retail channels is a very recent phenomenon. In the early stages of online business, many traditional B&M retailers developed online branches of their traditional businesses. In some cases, they saw in online stores a new version of their traditional catalog channel since there were, and still are, several similarities.

Today, the online channel has developed characteristics of its own. The relevance of this channel in the retail sector and the pressure from customers that want to interact with the company in a cohesive way have pushed B&M retailers to consider channel integration efforts with varying characteristics. Integration is not always evident to the customer, as is the case when a retailer ships an online purchase from a store rather than the warehouse. In other cases, integration is driven by the need to offer a homogenous and more rewarding online-offline customer experience. Examples include offering customers the possibility to return to a store items that were bought online, place online orders from the store and have the products shipped to the customer address, buy items online and pick them up later at the store in which they are stocked, or buy an item online and pick it up at the store once it has been delivered to the store.

Online-offline integration efforts are challenging for companies. The retailer must integrate inventory systems, warehouses, marketing campaigns, pricing strategies, etc. Even before these integration attempts are made, retailers often struggle to discern what is really available at their stores or warehouses, as has been studied in previous empirical research documenting substantial inventory record inaccuracy (DeHoratius and Raman, 2008). Another challenge faced in the implementation of some of these integration efforts is an increased complexity in store execution (Fisher et al., 2006). Store processes are designed to sell and not necessarily to support the quick delivery or shipment of goods, activities that these integration strategies allocate to physical stores.

Given that online-offline integration is a recent phenomenon, it is not surprising that there is limited literature that studies it. Some recent work in marketing and information systems has explored related issues, such as the difference in price elasticity between the online and brick and mortar channels (Chu et al., 2008, Granados et al., 2011), customer channel migration (Ansari et al., 2008), the choice between online and offline channels in grocery stores (Chintagunta et al., 2012), the impact of product returns on a multi-channel retailer (Ofek et al., 2011), or customer behavior in multi-channel customer service (Jerath et al., 2012). To our knowledge, no previous work has considered a buy-online-pickup-at-store channel. The competition between brick and mortar and online channels has been studied by Brynjolfsson et al. (2009) and Forman et al. (2009), among others.

In operations management, some work has examined fulfillment and supply chain choice on the Internet. For example, Netessine and Rudi (2006) study the effects of inventory ownership in online channels, and Randall et al. (2006) empirically study the decision to invest in fulfillment capabilities, although we are not aware of any work that has explored the integration of online and offline channels. We contribute to the literature by studying the impact of the implementation of an online-offline integration strategy, namely the "buy-online-pickup in store" functionality, on the online and brick and mortar channel.

When consumers decide to visit a physical store to buy a product, they face the risk that the product is out of stock. Fitzsimons (2000) and Anderson et al. (2006) study how customers respond to stockouts and how to measure and mitigate stockout costs. Substitution effects and its consequences for demand estimation are studied by Kok and Fisher (2007) and Musalem et al. (2012).

Some models in operations management consider the costs that of visiting a store (Dana and Petruzzi, 2001). Recent work has modeled the impact of inventory availability information in attracting consumer demand. In this stream, Su and Zhang (2009) study the value of commitment and availability guarantees when selling to strategic consumers. In a related work, Allon and Bassamboo (2011) explore the issue of cheap talk when the information shared is not verifiable. Our paper contributes to this stream of literature by providing the first empirical analysis of the impact of sharing *verifiable* inventory information. In our case, implementing the BOPS functionality can be interpreted as providing a commitment device to the inventory availability information, which is perceived by the customer as more credible. In this context, customers are able to "reserve" inventory that exists in the store. We can establish an analogy between the "inventory reservation" aspect of BOPS functionality and a restaurant reservation system; Alexandrov and Lariviere (2012) show that a reservation system reduces the uncertainty that customers face and may attract more people to the restaurants in certain situations.

4.3. Empirical Setting and Data

We have partnered with one of the leading nationwide retailers in the US that has implemented buy-online-pickup-at-store (BOPS) capabilities. This retailer specializes in housewares, furniture (indoor and outdoor), and home accessories, and has more than 80 B&M stores in the US and Canada. In addition to traditional B&M stores, this retailer has an online store that ships to ship to anywhere in the US.

We have obtained data spanning April 2011 to April 2012. Throughout this period, the online store offered information about the availability of inventory at each of the stores. After October 11, 2011, the retailer offered the option of placing orders online and picking them up at a B&M store. Under the BOPS mode of interaction, customers pay for the items through the online store (and therefore sales are considered *online* sales), but the order is fulfilled using inventory from the store. The pickup option was available simultaneously for every store in the US, but was not implemented for stores in Canada. The period of analysis considered in our analysis covers six months before the store pickup implementation (since April 11, 2011) and extends six months after the implementation (until April 11, 2012).

The information used in our analysis comes from two main data streams, one related to the online channel and the other related to the brick and mortar channel.

Data related to the online channel

We obtained daily data from the online channel at the designated market area (DMA)[3] level. For our main analysis of the impact of BOPS, we used data on total number of transactions, total dollar sales, and total number of unique visitors in the US for each day. Our data includes a total of 210 DMAs that completely cover the US populated areas. Table 19 shows the main summary statistics for these variables. We also obtained data about online shopping cart abandonment behavior for each DMA and day, using them in Section 4.5 to validate our interpretation of the findings.

Data related to the brick and mortar channel

We obtained daily data for each of the stores in the US and Canada. During the period of analysis, the retailer had a total of 83 stores in the US and Canada. For our analysis of the impact of the BOPS implementation on the B&M channel, we collected data on the total number of transactions, total dollar sales, and total visitors for each day and store in the US and Canada. Table 19 shows relevant summary statistics for these variables.

In addition, we collected data specifically related to the BOPS orders. We obtained information on the date that a BOPS transaction was placed online and the date and store where each one of these pickup transactions was collected by the customer. We use this data in Section 4.4.3.

[3] A designated market area (DMA) is a region where the population can receive the same (or similar) television and radio station offerings, and may also include other types of media including newspapers and Internet content. They can coincide or overlap with one or more metropolitan areas, though rural regions with few significant population centers can also be designated as markets. They are widely used in audience measurements, which are compiled in the United States by Nielsen Media Research (television) and Arbitron (radio) (from Wikipedia)

4.4. Evaluating the Impact of BOPS

A naive approach to evaluating the impact of BOPS would look at the difference in the variables of interest between the pre-implementation period and the post-implementation period. Clearly, this approach would be flawed; many things can differ in the pre-implementation and post-implementation period that are completely unrelated to the BOPS implementation. For example, there might be seasonal factors that cause a change in sales. In order to deal with this challenge, we consider a difference-in-differences approach (DiD).

In general, to implement a DiD approach we need to identify a portion of the population that is not affected by the intervention for which we are trying to estimate the causal effect (the BOPS implementation in our case). In other words, we need a control group. After identifying a control group, we can measure the effect of the treatment by comparing the differences between treatment and control groups before and after the treatment is applied. Figure 9 shows a schematic summary of the DiD approach. For a more detailed discussion on this topic, see Angrist and Pischke (2008).

The rest of this section applies a DiD approach to evaluate the impact of BOPS in the online and brick and mortar channels. Subsection 4.4.1 estimates the effect of BOPS on the online channel using a control group based on the distance between the online customers and the closest store. Customers that visit the website from locations that are very far from a store are used as a control group in the DiD framework. Subsection 4.4.2 uses Canadian stores, where the BOPS functionality was not deployed, as a control group for the DiD framework. Subsection 4.4.3 studies whether BOPS transactions lead to cross-selling of additional in-store purchases by customers who ordered goods online.

4.4.1. Impact on the Online Channel

We start our analysis of the impact of BOPS by focusing on its effects on traffic and sales observed in the online channel. For this purpose, we use data from the online business that covers the six months preceding the implementation of BOPS and the six months following the implementation.

As mentioned before, if we simply compared what happened before and after the intervention, we would not be able to find a causal effect of the intervention, because the pre and post-implementation periods might differ in aspects other than the intervention. For example, the post-intervention period includes the Christmas season, which we can expect to have higher sales independent of the BOPS project. In order to control for differences not related to the BOPS implementation, we define two different groups in our population. The first group includes the portion of the population that was affected by the BOPS implementation (the treatment group); the second group includes the portion of the population that was not affected by this decision (the control group). In the definition of a control group, we take into account the fact that customers who live far from physical B&M stores will be unaffected by the deployment of the BOPS capabilities.

More specifically, we conduct our analysis for the online channel at the DMA level. The retailer has a total of 79 B&M stores in the U.S.; this relatively small number of stores helps us to identify a treatment and control group in our population. Our treatment group is defined to include those DMAs within the area of influence of a B&M store. The control group includes DMAs that are not within the area of influence of a B&M store. As a baseline, we assume that the area of influence of a B&M store covers a radius of 50 miles, but our results are robust to choosing different distances. This classification is used because customers visiting the online store from DMAs that are not within the area of influence of a B&M store will find no use for the pickup implementation. The store's inventory information shown online and the option to pick up online purchases at a store should not affect customer behavior within those DMAs, as it is not practical for customers to visit a physical store. Hence, it is reasonable to assume that customers within these DMAs can behave as a control group, in the sense that they will be affected by the general dynamics of behavior of the online channel (for example, they will respond to the seasonal Christmas period), but not by the BOPS implementation. In contrast, online customers who visit the website from DMAs that are in the area of influence of a store can benefit from this new alternative. It is possible for customers in those DMAs to actually visit a store to pick up the items they bought online, or decide to go to the store shown online to have the item desired in order to make their purchase.

Figure 8 shows the location of the B&M stores and the geographic center of each of the DMAs. From the total of 210 DMAs, 162 of those DMAs do not include a B&M store within their geographic area and the other 48 DMAs have at least one store within their geographic area[4] . In our analysis, following the company's practice, we considered all the pickup sales as online sales.

In the first place, we study whether changes in online traffic can be attributed to the BOPS implementation. To do this, we consider the number of unique visitors ($NUMVISITORS_{it}$) from a DMA i in a day t as our dependent variable. Our independent variables include a dummy variable that indicates whether or not the DMA i is within the area of influence of a store ($CLOSE_i$), a dummy variable that indicates if the observation corresponds to the period after the pickup implementation ($AFTER_t$), and the interaction between these two terms ($CLOSE_i * AFTER_t$), which is our variable of interest.

In addition to defining our treatment and control groups and the independent variables described in the previous paragraph, we include an exhaustive number of control variables, taking advantage of the panel structure in our data. Our model includes fixed effect for each DMA i, week and day of the week in our sample. Our model specification is the following:

$$\begin{aligned} NUMVISITORS_{it} =& \mu_i + \alpha_1 CLOSE_i + \alpha_2 AFTER_t + \\ & \alpha_3 CLOSE_i * AFTER_t + \alpha_4 CONTROLS_{it} + \epsilon_{it} \end{aligned} \tag{4.1}$$

[4]We defined a DMA as being within the area of influence of a store if a 50 miles radius circle centered at a store overlaps with the DMA area. We consider a distance of 50 miles as this is the distance that the retailer's management team estimates as the area of influence of their stores in their business analysis. We tested other distance specifications (e.g., 40 and 60 miles) and our results were robust to these alternatives.

Since we have DMA fixed effects μ_i, it is not possible to identify α_1 separately from μ_i, because there is no variation in $CLOSE_i$ for a given store in the period of analysis. This is not problematic, because we are interested in the value of the coefficient α_3 in this specification, which is identified[5]. The results of the estimation of this model are presented on the first column of Table 21. We can observe that after the BOPS implementation, for those DMAs that are in the area of influence of a store (i.e., $CLOSE_i * AFTER_t = 1$), there was a positive and significant effect on the number of unique website visitors, relative to those DMAs that were not within the area of influence of a store ($CLOSE_i * AFTER_t = 0$). In other words, visits from DMAs that were in the catchment area of a B&M store increased more than visits from DMAs that were not close to a B&M store.

The next step is to study the impact on online sales, before and after the implementation. For this analysis we implement a similar model to the one described above, but now our dependent variable, $SALES_{it}$, corresponds to the total dollar sales from online visits from DMA i on day t. In addition, we also tested the specification including the number of unique visitors ($NUMVISITORS_{it}$) as an additional control in our model.

$$\begin{aligned} SALES_{it} = &\mu_i + \beta_1 CLOSE_i + \beta_2 AFTER_t + \\ &\beta_2 CLOSE_i * AFTER_t + \beta_4 * CONTROLS + \epsilon_{it} \end{aligned} \tag{4.2}$$

The results of this analysis are presented on the second and third columns of Table 21. The third columns presents the results when the online traffic is included as an additional control variable. After the BOPS implementation, for those DMAs in the area of influence of a store (i.e., $CLOSE_i * AFTER_t = 1$), we observe a negative and significant effect on sales, relative to DMAs that are not within the area of influence of a store ($CLOSE_i * AFTER_t = 0$). That is, BOPS implementation reduces online sales.

Although the brick and mortar store locations are decided endogenously, we do not believe

[5]This also happens in equations 4.2, 4.3, 4.4, 4.6 and 4.7. We leave the unidentified variables in the model specification for clarity, but we focus our analysis on the interaction terms that are identified.

that this poses a serious concern for the validity of our results. During the period of analysis, no stores were opened or closed, and the store locations were determined many years before the BOPS implementation. The BOPS implementation was executed at the same time at every location. In addition, the panel structure of the data gives us the ability to add fixed effects that fully control for all the invariant characteristics across the DMAs. For further robustness of our analysis we include the results obtained when a previous sample matching is used. For this analysis, we consider DMAs that have a comparable traffic pattern (less than 900 daily visits to the online store), to make sure that our results are not driven by fundamental differences between the treatment and control groups. These results are presented on Table 26, where the second column includes the online traffic as an additional control.

The results presented in this section go against intuition and conventional wisdom, which suggest that online sales should increase after the BOPS implementation, because customers are given more options to order online. We find that, although online traffic *increases* at the DMAs affected by the BOPS implementation, sales *decrease* at those DMAs, relative to the DMAs that were unaffected by the change. Section 4.5 gives a holistic interpretation of this unexpected phenomenon.

4.4.2. Impact on the Brick and Mortar Channel

We now tackle the analysis of the impact of BOPS on the B&M stores. In particular, we want to understand how the implementation of the store pickup option impacts traffic and sales at the B&M stores.

Again, we face the challenge that an appropriate answer cannot be obtained simply by observing what happened with store sales before and after the pickup implementation. To answer this question properly, we propose a difference-in-differences approach with new treatment and control groups.

In this case, the key variation that allows us to identify the effect of BOPS comes from the fact that while the retailer owns B&M stores both in the US and Canada, the BOPS option was not deployed for stores in Canada. This situation allows us to use the B&M stores in Canada as a control group for our analysis. The treatment group includes all the B&M stores in the US. The reasoning behind this definition is that customers visiting the stores in Canada were not influenced by the BOPS pickup implementation while customers in the U.S. were exposed to this new alternative and could benefit from it.

After BOPS implementation, the retailer had a total of 79 stores in the US that offered the pickup option and 4 stores in Canada that did not offer this alternative.

As before, we want to focus first on the impact on customer traffic at the B&M stores. To do this, we consider the traffic count at a store i on day t ($TRAFFIC_{it}$) as our dependent variable. Our independent variables in the model include a dummy variable that indicates if store i is located in the U.S. or not (US_i); a dummy variable that indicates if the observation corresponds to the period after the pickup in the store implementation ($AFTER_t$); and the interaction between these two terms ($US_i * AFTER_t$), which is our main variable of interest. Our model also includes also fixed effects for each store i, week and day of the week in our sample. The model specification is the following:

$$TRAFFIC_{it} = \mu_i + \alpha_1 US_i + \alpha_2 AFTER_t + \alpha_3 US_i * AFTER_t + \alpha_4 CONTROLS + \epsilon_{it} \tag{4.3}$$

The results for the estimation of this model are presented in the first column of Table 22. We observe that there was a positive and significant effect on the traffic of the US stores (i.e., $US_i * AFTER_t = 1$) compared to the traffic in the Canada stores, after the store pickup implementation. In other words, stores that were affected by the BOPS implementation saw a higher increase in their store traffic.

The next step is to study the impact on store sales. To do this, we define the following model:

$$SALES_{it} = \mu_i + \beta_1 US_i + \beta_2 AFTER_t + \beta_3 US_i * AFTER_t + \beta_4 CONTROLS + \epsilon_{it} \quad (4.4)$$

where $SALES_{it}$, the dependent variable, corresponds to the total dollar sales at store i during day t. In addition, some of our specifications include the total traffic at store i on day t ($TRAFFIC_{it}$) as a control variable (recent work in retail operations has considered how store traffic affects conversion rates, see (Perdikaki et al., 2012).. The results of this analysis are presented on the second and third column of Table 22. We observe a positive and significant effect on sales for the US stores after the pickup implementation (i.e., $US_i * AFTER_t = 1$) compared to the stores in Canada. This result suggests that the B&M stores that were affected by the BOPS implementation (i.e., stores in the US) saw an increase in sales compared to the control group (i.e., stores in Canada).

It is important to notice that our analysis includes, as an additional control, a daily measure for currency exchange rate between the US and Canada. This variable can capture changes in the economic situation between the two countries during the period of analysis that could potentially affect our results.

For additional robustness checks of our results, we restrict our attention to different subsamples of the data, such as a shorter period before and after the implementation of BOPS (e.g., considering only one month before and after). The results do not qualitatively change.

The analysis of the impact of BOPS on the B&M stores indicates that B&M stores received *more traffic* and *increased their sales* as a consequence of the BOPS implementation.

4.4.3. Positive Externality of Store Pickup

There is an additional mechanism by which B&M sales can benefit from the pickup in the store implementation. Customers who visit the B&M stores to pick up an item they bought online might decide to buy extra items during their visit. Also, some customers may go to pick up the items with someone else who may make additional purchases while visiting the store. In any case, we are interested in testing whether there is evidence about this positive externality effect in our data.

Our data does not allow us to uniquely identify purchases made by customers when visiting the store to pick up an item they have ordered online. If this were possible, we could evaluate whether BOPS customers add extra items to their purchase while at the store. The main limitation here is that the pickup transactions are done; hence, if the customer makes a purchase while at the store, the retailer records this as a different transaction. We do not have access to the data that could potentially enable the link between these two transactions.

To overcome this challenge, we test whether the total number of customers that picked up items at a store on a particular day ($NUMPICKUPTX_{it}$) is correlated with an increase in store sales on that day. By doing this analysis we can observe if more pickup customers at the stores generate more store sales. We propose the following model:

$$\begin{aligned} SALES_{it} =& \mu_i + \beta_1 NUMPICKUPTX_{it} + \beta_2 TRAFFIC_{it} + \\ & + CONTROLS + \epsilon_{it} \end{aligned} \tag{4.5}$$

where $SALES_{it}$ corresponds to the total dollar sales at store i during day t. The independent variables in the model include total traffic at each store ($TRAFFIC_{it}$), store fixed effects, and week and day of the week dummies. Controlling for traffic mitigates the potential bias that we would have if people choose to pick up items on days that have particular characteristics that might be correlated with sales. Obviously, this analysis only applies to the period that follows the BOPS implementation, since no customers were picking up items before that.

The results of this analysis are presented on Table 23. In the first column we observe that the number of pick up transactions has a positive and significant effect on the number of transactions. The second and third columns on Table 23 present the analysis for the impact of the pickup transactions on the store sales. This result is consistent with the hypothesis that there are positive externalities on store sales due to those customers that visit a store to pick up their online order.

4.5. Interpretation of Results: The Role of Verifiability

The results presented on the previous section are somewhat perplexing. Conventional wisdom suggests that when the retailer decides to offer a new service to online customers (the store pickup option), if anything, this service should benefit the retailer's online sales. However, the results presented on Section 4.4.1 tell us the opposite: online sales decrease after the pickup implementation[6]. On the other hand, as presented on Section 4.4.2, B&M sales benefit from this new online service. This is also surprising since, a priori, we could have assumed that a service level increase on the online store could have hurt B&M sales.

As discussed in Section 4.4.3, part of the increase in B&M sales can be explained by the positive externality generated by customers picking up their online orders at the stores. However, this does not give the entire picture. For example, this does not explain why online sales go down in areas close to a store or why traffic increases online and at the stores subject to the BOPS implementation. Section 4.5.1 proposes an explanation based on how the implementation of BOPS provides a shock to the credibility of the inventory information shown online. Based on this explanation, two additional hypotheses are developed and tested in Sections 4.5.2 and 4.5.3.

4.5.1. Credibility of Inventory Information

Our explanation draws from recent models proposed by the operations management literature to study how strategic consumers react to inventory information. Following Su and

[6]To make our exposition more clear we will refer to our results in absolute terms. However, it is important to note that the results obtained on section 4.4.1 should be interpreted in relative terms, since we are always comparing a treatment and a control group and their relative differences. For example, it is possible that both the treatment and the control group sales increased during the period under analysis, but at different rates.

Zhang (2009), there is a cost for customers to visit a store. Customers form rational expectations of product availability and make patronage decisions based on their expectations. Our retailer was sharing inventory availiability online during the entire period of analysis. However, customers do not necessarily perceive inventory information shown to them as reliable. For example, in Su and Zhang (2009) the seller has an incentive to convince customers that inventory will be available, because that increases the probability that they visit the store. Accordingly, buyers should ignore the seller's claims. Similarly, Allon and Bassamboo (2011) provide a model in which a retailer shares unverifiable inventory information with strategic consumers. In equilibrium, the information becomes cheap talk and consumers ignore it. Furthermore, DeHoratius and Raman (2008) report a considerable amount of inventory record inaccuracy, and it is possible that some consumers consider any inventory information as unreliable.

Su and Zhang (2009) show that committing to an inventory level is valuable. In our case, offering "buy-online, pick up at store" provides an exogenous shock to the credibility of inventory information and works as a commitment device. Offering an online customer the option to pick up her online purchase at a nearby store can be beneficial to the customer for several reasons. For example, she can get the item in a couple of hours (not a couple of days) or avoid the payment of shipping costs. However, there is an additional benefit to the customer that might not be evident right away: She now knows that the item she wants is available in a nearby store. Hence, the customer can decide to check the availability online and drive to the store to pick up the item without closing the transaction online. This behavior will allow her the benefit of getting the item fast with no shipping cost, and additionally, let her evaluate the item at the store before actually paying for it (while avoiding the risk of making a trip to the store and not finding the item she is looking for).

We want to emphasize that the customer is facing reliable availability information. Everyone has suffered the unpleasant experience of cheap talk regarding inventory information; car dealers are a good example of this type of cheap talk. However, the information released by the retailer under the BOPS alternative is different. After BOPS implementation, online customers do not observe a vague promise of availability; they can actually buy the item online and pick it up at the store two hours later. They know that the item is actually there. Furthermore, customers can anticipate that offering the BOPS functionality requires increased accuracy in the retailer's internal inventory records, lending even more credibility to the inventory information.

In addition, some customers will obtain additional information by touching the fabric, seeing the color, or evaluating the actual size of the product they want to purchase. There is a benefit to seeing and touching the type of items that this retailer sells before committing to the purchase. This first-hand experience is less relevant for customers when buying other types of products that are more standardized (e.g. books or electronics).

Hence, our results can be explained by the following behavior: After implementing BOPS, some customers (more than before) visit the online store to browse the catalog, find the item they want to buy, check its availability and travel to their local store to close the sale knowing the item is there. The customers, after observing reliable availability information, decide to visit the store without closing the sale online.

We predict that a shock to the credibility of inventory information would reduce the cost of visiting the stores and would result in:

1. *Increase in store visits* to stores that share credible inventory information, relative to those that do not.

2. *Increase in store sales* in stores that share credible inventory information, relative to those that do not.

3. *Decrease in online sales* in DMAs that are within the area of influence of the stores, relative to those DMAs that not.

These are precisely the results that we found in Section 4.4. In order to confirm that our explanation is indeed valid, we develop two additional tests in Sections 4.5.2 and 4.5.3.

4.5.2. Cart Abandonment

In order to validate that the aforementioned mechanism is in place, we have established a hypothesis that would be supported if the described mechanism occurs. This hypothesis is related to shopping cart abandonment. A number of customers *abandon* their online shopping carts before finishing the transaction. After BOPS was deployed, customers could place an order for an item in their shopping cart and pick it up from the store. The fact that a product is available in a local store can prompt some customers to abandon their shopping carts and buy the items directly from the store without closing the transaction online. If, as our explanation suggests, the inventory availability information is perceived as more reliable after the BOPS project has been implemented, we can hypothesize that the rate of shopping cart abandonment will increase after BOPS is implemented.

Once again, we follow a DiD approach. We want to compare what happened before and after the BOPS implementation. To do this we consider the group of DMAs that were affected by the BOPS implementation (our treatment group) and the group of DMAs that were not affected (our control group). These groups are the same as those described on Section 4.4.1.

Our dependent variable ($ABANDON_{it}$) corresponds to the fraction of customers that, after placing an item in their cart, did not close the sale[7]. As described before, if our inventory credibility story holds, we would expect a higher increase in cart abandonment rate after BOPS implementation for those DMAs that are in the area of influence of a store (some customers in those DMAs may decide to rely on the inventory availability information and visit the store to make a purchase), relative to those DMAs that are not. This is captured

[7]The abandonment rate in our sample has a mean of 51% with a standard deviation of 24%.

by the $CLOSE_i * AFTER_t$ variable in the following model:

$$\begin{aligned} ABANDON_{it} = &\mu_i + \beta_1 CLOSE_i + \beta_2 AFTER_t + \\ &\beta_3 CLOSE_i * AFTER_t + \beta_4 CONTROLS + \epsilon_{it} \end{aligned} \tag{4.6}$$

The results of this analysis are presented in Table 24. Our hypothesis is validated. Shopping cart abandonment rate increased after BOPS was deployed for those DMAs where the BOPS was available compared to those DMAs where customers could not take advantage of this new feature.

4.5.3. Online Sales in the Area of Influence of a Store

In this subsection we present an additional analysis that attempts to further validate the explanation presented on Section 4.5.1. If our explanation is correct, we expect that, for those DMAs that are within the area of influence of one or more stores, online sales will decrease compared to the B&M store sales in those DMAs. In other words, we expect that in DMAs within driving distance to the store, the sales share of B&M stores will increase after the pickup implementation.

Note that this analysis is different but complementary to the analysis shown in Section 4.4. The model developed in Section 4.4.1 compares the lift in online sales that occur in *DMAs far from stores* with the lift in online sales that occur in *DMAs that are close to stores*. The model developed in Section 4.4.2 compares the lift in *store sales in the US* with the lift in *store sales in Canada*. Although consistent with a shift from the online to the brick and mortar channel, the models in Section 4.4 do not provide a direct test. The present analysis looks at the area of influence of a store, and compares the lift in *online sales in the area of influence of a store* to the lift in *sales of the store*. Combined with the previous evidence, finding that sales in a store increase more than online sales in the area of influence of the store would suggest that there is a shift of customers from the online channel to the brick and mortar channel.

To implement this analysis, we consider daily online and B&M sales ($SALES_{it}$) aggregated at the DMA level as our dependent variable. Our independent variables consist of a dummy that indicates whether the daily sales observation corresponds to an online or B&M sales ($ONLINE_i$), a dummy variable that indicates whether the observation corresponds to the period after the pickup implementation ($AFTER_t$), and the interaction of these two terms ($ONLINE_i * AFTER_t$), our variable of interest. :

$$\begin{aligned} SALES_{it} = & \mu_i + \beta_1 ONLINE_i + \beta_2 AFTER_t + \\ & \beta_3 ONLINE_i * AFTER_t + CONTROLS + \epsilon_{it} \end{aligned} \tag{4.7}$$

The results of this analysis are presented on Table 25. As expected, the coefficient for our variable of interest ($ONLINE_i * AFTER_t$) has a negative and significant effect. This tells us that, after the pickup implementation, the share of online sales decreases with respect to the share of B&M sales.

4.6. Conclusions

Our analysis of the impact of an online-offline integration strategy offering the option of buying items online and picking them up in a physical store shows that, contrary to conventional wisdom, online sales do not increase with the implementation of the BOPS functionality. We find that the BOPS implementation results in lower online sales, higher store sales and higher store traffic. We explain these results in light of recent models from the operations management community that study sharing inventory information with strategic consumers. The implementation of BOPS provides a natural experiment that gives a positive shock to the credibility of the inventory information shared with the customers. As this information becomes more credible, more customers use the online channel to browse store inventory availability but make their purchases online.

One question unaddressed so far is whether the increase in store sales compensates for the decrease in online sales. It turns out that, in our particular case, it does. Based on our results, we can give an estimation of the net effect of the BOPS implementation. Using representative values from our sample, the decrease in online sales that can be attributed to the BOPS implementation is approximately 1.8% of the total retailer sales, while the increase in the store sales amounts to around 3.6% of the total retailer sales, giving a net increase of 1.8% of the total sales. Interestingly, without the holistic interpretation of online and store sales, an evaluation of the effects on the online channel might have suggested that BOPS was not a good idea.

One of the consequences of the increase in online activity in retail is the availability of richer data that can be used to evaluate the impact of operational interventions and to discover relationships between different operational aspects of the business. However, the available data is often channel dependent. Incentives for managers are also often channel dependent, and making decisions based on own-channel data is often tempting. Our results show that decisions affecting one channel should be evaluated holistically. A partial analysis of this project based on online sales might give the wrong conclusions, as illustrated in Section 4.4.1. For example, in online settings it is customary to evaluate potential changes by conducting A/B testing (which consists of showing one condition to a fraction of the visitors and a different condition to the rest of the visitors) and monitor how conversion rates differ across conditions. Typically, interventions that A/B testing identifies as negatively affecting the channel conversion rate are ruled out. Evaluating the impact of online-offline integration strategies requires a holistic view of company operations, since an intervention that might be detrimental in one channel can yield substantial benefits in the other one. Our results show that when evaluating actions that perform integration between channels, it is very important to close the loop and evaluate their impact on all channels.

To our knowledge, this is the first academic study of the impact of implementing a buy-online-pickup-at-store channel. Contrary to the a priori expectation that a BOPS intervention would result in an increase of transactions closed online, we find that the additional reliability of the inventory information prompts some customers to actually visit the stores, which in turn increases the relative importance of the brick and mortar stores in terms of share of sales. Our explanation of the observed results is related to the credibility of the inventory information that retailers share. When retailers share reliable inventory information, the perceived risk that customers experience in visiting the stores is reduced; they visit stores more frequently and stores sell more.

Our results raise questions about where the value of BOPS for customers comes from. The value proposition for the customer was originally to increase the speed of the delivery for online customers and to allow them to save shipping costs. Interestingly, if we examine the time that it takes BOPS customers to pick up their items, we see that speed does not seem such an important concern for the average customer (see Figure 3). Besides saving shipping costs, the BOPS project seems to offer a positive externality to customers in the form of more reliable inventory information that can be used before going to the store. For the company, BOPS results in additional traffic and sales in the stores and cross-selling of products for customers that visit stores to pick up their orders.

Our results might depend substantially on the type of products transacted. In our empirical setting, most of the products have an experiential component and shopping in the store is a pleasant experience. For highly standardized products, or for products which lead to a less pleasant store experience, it is possible that BOPS implementation does not result in such a substantial shift from the online channel to the store channel.

Randomized field experiments are an ideal way to cleanly identify the impact of retail online-offline strategies implementations. In situations where randomized field experiments are not feasible, we believe that the framework we have presented, consisting of identifying natural experiments and performing DiD estimation of the effect of the treatment, can help retailers achieve a more precise estimation of the effect of their online-offline integration strategies.

4.7. Appendix: Tables and Figures

In this appendix, we provide tables and figures for the empirical studies in Chapter 4.

Table 19: Summary Statistics

B&M Stores - Daily Parameters per Store

Traffic		Transactions		Sales	
Mean	StDev	Mean	StDev	Mean	StDev
1018.1	801.8	285.5	205.9	27424.2	20716.1

Online Store - Daily Parameters per DMA

Visitors		Orders		Sales	
Mean	StDev	Mean	StDev	Mean	StDev
912.4	2485.2	24.7	71.2	3722.7	12907.7

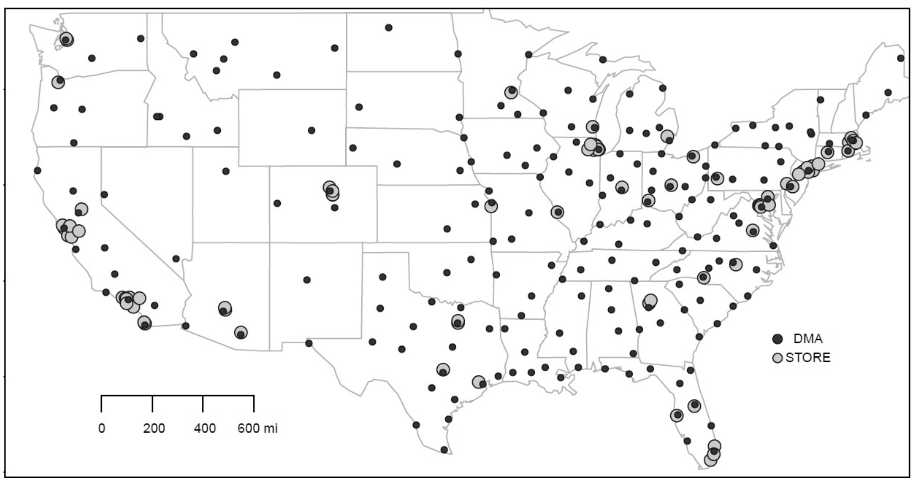


Figure 8: Brick and mortar stores and DMAs

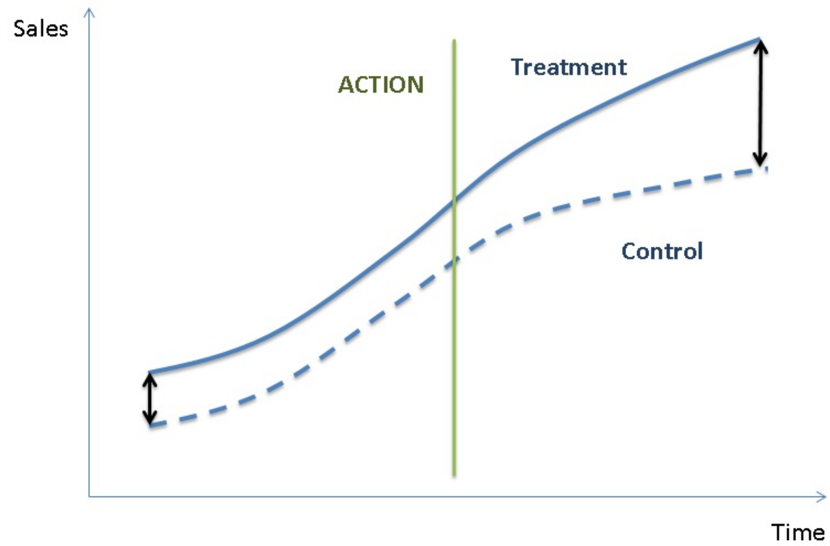


Figure 9: Natural Experiment

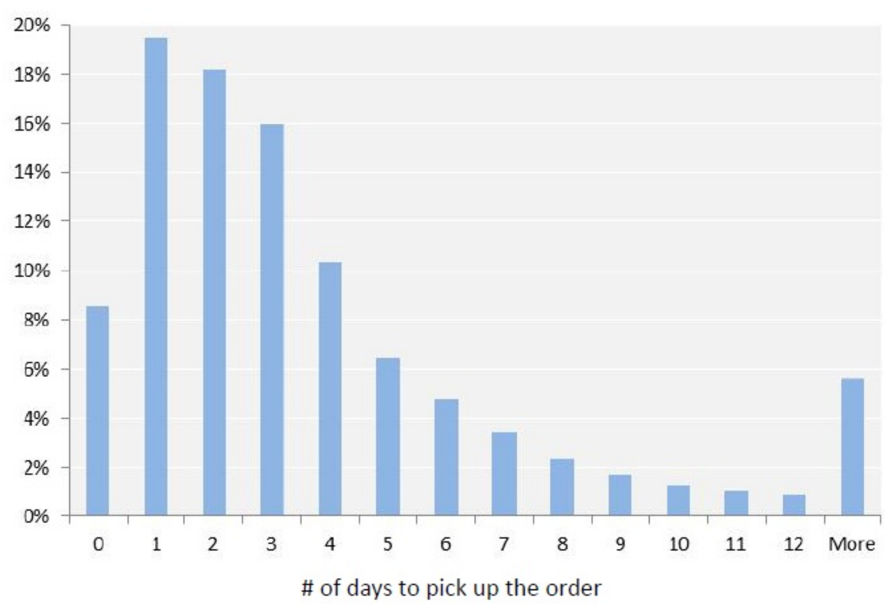


Figure 10: Pickup Delay

Table 20: Variable Definitions

Analysis of the Impact on the Online Channel

Variable Name	Definition
$NUMVISITORS_{it}$	Total number of unique visitor to the CB website coming from DMA i on day t.
$CLOSE_i$	Dummy variable that is 1 if the DMA i has a BM store within its geographic area.
$AFTER_t$	Dummy variable that is 1 if on date t the BOPS option was offered.
$SALES_{it}$	Total dollar sales at the online store coming from DMA i on day t.

Analysis of the Impact on the Brick and Mortar Channel

Variable Name	Definition
$TRAFFIC_{it}$	Total number visitors to BM store i on day t.
US_i	Dummy variable that is 1 if the BM store i is located in the US.
$AFTER_t$	Dummy variable that is 1 if on date t the BOPS option was being offered.
$SALES_{it}$	Total dollar sales at the BM store i on day t.

Analysis of Channel Marketshare

Variable Name	Definition
$ONLINE_i$	Dummy variable that is 1 if the observation referes to the online store.
$AFTER_t$	Dummy variable that is 1 if on date t the BOPS option was being offered.
$SALES_{it}$	Total dollar sales from the BM or the online store i on day t.

Analysis of BOPS Positive Externality

Variable Name	Definition
$NUMPICKUPTX_{it}$	Total number of customers that visit the BM store i on day t to pick up their BOPS order.
$TRAFFIC_{it}$	Total number visitors to BM store i on day t.
$SALES_{it}$	Total dollar sales at the BM store i on day t.

Analysis of Cart Abandonment

Variable Name	Definition
$ABANDONMENT_{it}$	Percentage of customers that place at least one item in their shopping cart and left without closing the sale at DMA i on day t.

Table 21: DMA's Online Store

	TRAFFIC	SALES	SALES
AFTER	-50.30***	366.48**	686.99
	(13.05)	(111.80)	(173.91)
CLOSE*AFTER	135.56**	-410.72*	-1274.55***
	(42.39)	(188.81)	(355.15)
ONLINE TRAFFIC			6.37***
			(0.36)
Fixed Effects	YES	YES	YES
Week	YES	YES	YES
Day of the Week	YES	YES	YES
N	76903	76903	76903
DMA	210	210	210
R^2	0.96	0.77	0.83

Robust standard errors in parentheses

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table 22: BM Stores

	TRAFFIC	SALES	SALES
AFTER	-375.00***	7995.71***	12486.33***
	(53.63)	(1598.99)	(1526.24)
US*AFTER	246.71***	4857.41***	2210.21**
	(39.33)	(754.02)	(711.54)
TRAFFIC			10.93***
			(0.17)
Fixed Effects	YES	YES	YES
Macroeconomic controls	YES	YES	YES
Week	YES	YES	YES
Day of the Week	YES	YES	YES
N	28138	28133	28133
Stores	83	83	83
R^2	0.75	0.66	0.70

Robust standard errors in parentheses

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table 23: Pickup Visits in the Stores

	TRANSACTIONS	SALES	SALES
PICKUP VISITS	5.13***	318.03***	139.68**
	(0.91)	(61.95)	(46.20)
TRAFFIC	0.22***		10.83***
	(0.02)		(1.28)
Fixed Effects	YES	YES	YES
Week	YES	YES	YES
Day of the Week	YES	YES	YES
N	12093	12547	12088
Stores	79	79	79
R^2	0.92	0.61	0.68

Robust standard errors in parentheses

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table 24: Cart Abandonment

	ABANDONMENT
AFTER	0.004
	(0.013)
CLOSE*AFTER	0.009***
	(0.007)
Fixed Effects	YES
Week	YES
Day of the Week	YES
N	73500
DMA	210
R^2	0.136

Robust standard errors in parentheses

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table 25: Store sales vs online sales in the area of influence of a store

	SALES
AFTER	21759.35***
	(1862.40)
ONLINE*AFTER	-3685.37***
	(572.74)
Fixed Effects	YES
Week	YES
Day of the Week	YES
N	55412
DMA	161
R^2	0.41

Robust standard errors in parentheses

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table 26: DMA's Online Store with Average Traffic less than 900 daily visits

	SALES	SALES
AFTER	34.54	24.02
	(27.58)	(28.29)
CLOSE*AFTER	-65.20*	-51.08*
	(25.45)	(24.49)
ONLINE TRAFFIC		4.03***
		(0.21)
Fixed Effects	YES	YES
Week	YES	YES
Day of the Week	YES	YES
N	42405	42405
DMA	116	116
R^2	0.18	0.22

Robust standard errors in parentheses

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

BIBLIOGRAPHY

P. Albuquerque and B. J. Bronnenberg. Measuring the impact of negative demand shocks on car dealer networks. *Marketing Science*, 31(1):pp. 4–23, 2012. ISSN 07322399.

A. Alexandrov and M. Lariviere. Are reservations recommended? *Manufacturing and Service Operations Management*, 2012.

G. Allon and A. Bassamboo. Buying from the babbling retailer? the impact of availability information on customer behavior. *Management Science*, 57(4):713, 2011.

T. Andersen, J. Bentzen, C. Dalgaard, and P. Selaya. Lightning, it diffusion and economic growth across us states. *Discussion Papers of Business and Economics*, 2010.

E. Anderson, G. Fitzsimons, and D. Simester. Measuring and mitigating the costs of stock-outs. *Management Science*, pages 1751–1763, 2006.

J. D. Angrist and J.-S. Pischke. *Mostly Harmless Econometrics: An Empiricist's Companion*. Princeton University Press, 1 edition, Dec. 2008. ISBN 0691120358.

A. Ansari, C. Mela, and S. Neslin. Customer channel migration. *Journal of Marketing Research*, 45(1):60–76, 2008.

AutomotiveNews. Dealers: We want more cars! in quest for profits, factories keep inventory tight. *http://www.autonews.com/apps/pbcs.dll/article?AID=/20100823/OEM01/308239949/1262*, page David Barkholz, 2010.

S. Balachander, Y. Liu, and A. Stock. An empirical analysis of scarcity strategies in the automobile industry. *Management Science*, 55(10):1623–1637, 2009. doi: 10.1287/mnsc.1090.1056.

A. Balakrishnan, M. S. Pangburn, and E. Stavrulaki. "stack them high, let 'em fly": Lot-sizing policies when inventories stimulate demand. *Management Science*, 50(5):630–644, 2004.

R. Bansal and M. Ochoa. Temperature, growth, and asset prices. Technical report, working paper, 2009.

P. Barwise, A. Elberse, and K. Hammond. Marketing and the internet: a research review. 2002.

P. A. Bekker. Alternative approximations to the distributions of instrumental variable estimators. *Econometrica*, 62(3):pp. 657–681, 1994. ISSN 00129682.

E. Bendoly, J. Blocher, K. Bretthauer, S. Krishnan, and M. Venkataramanan. Online/in-store integration and customer retention. *Journal of Service Research*, 7(4):313–327, 2005.

J. Bennett, S. Terlep, and M. Ramsey. Weather dents auto production. *Wall Street Journal*, February 2 2011.

BloombergTV. Warm weather a boost to retail sales? *TV show*, March 2012.

S. S. Brehm and J. W. Brehm. *Psychological Reactance: A Theory of Freedom and Control.* Academic Press, New York, 1981.

T. Bresnahan and V. Ramey. Output fluctuations at the plant level. *The Quarterly Journal of Economics*, 109(3):593, 1994.

T. C. Brock. Implications of commodity theory for value change. In *A. G. Greenwald, T. C. Brock, & T. M. Ostrom (Eds.). Psychological foundations of attitudes.*, pages 243–275. New York: Academic Press, 1968.

H. Brodsky and A. Hakkert. Risk of a road accident in rainy weather. *Accident Analysis & Prevention*, 20(3):161–176, 1988.

E. Brynjolfsson, Y. Hu, and M. Rahman. Battle of the retail channels: How product selection and geography drive cross-channel competition. *Management Science, forthcoming*, 2009.

A. Buse. The bias of instrumental variable estimators. *Econometrica*, 60(1):pp. 173–180, 1992. ISSN 00129682.

G. P. Cachon, C. Terwiesch, and Y. Xu. On the effects of consumer search and firm entry in a multiproduct competitive market. *Marketing Science*, 27(3):461–473, 2008.

G. P. Cachon, S. Gallino, and M. Olivares. Severe weather and automobile assembly productivity. *Working Paper*, 2011.

J. C. Chao and N. R. Swanson. Consistent estimation with a large number of weak instruments. *Econometrica*, 73(5):pp. 1673–1692, 2005. ISSN 00129682.

P. K. Chintagunta, J. Chu, and J. Cebollada. Quantifying transaction costs in online/off-line grocery channel choice. *Marketing Science*, 31(1):96–114, 2012.

J. Chu, P. Chintagunta, and J. Cebollada. Research note: A comparison of within-household price sensitivity across online and offline channels. *Marketing Science*, 27(2):283–299, 2008.

T. Crocker and R. Horst. Hours of work, labor productivity, and environmental conditions: A case study. *The Review of Economics and Statistics*, 63(3):361–368, 1981.

J. Dana, James D. and N. C. Petruzzi. Note: The newsvendor model with endogenous demand. *Management Science*, 47(11):pp. 1488–1497, 2001. ISSN 00251909.

P. DeGraba. Buying frenzies and seller-induced excess demand. *The RAND Journal of Economics*, 26(2):331–342, 1995. ISSN 07416261.

N. DeHoratius and A. Raman. Inventory record inaccuracy: an empirical analysis. *Management Science*, 54(4):627–641, 2008.

M. Dell, B. Jones, and B. Olken. Climate change and economic growth: evidence from the last half century, 2008.

M. Dell, B. Jones, and B. Olken. Temperature and income: reconciling new cross-sectional and panel estimates, 2009.

O. Deschenes and M. Greenstone. The economic impacts of climate change: evidence from agricultural output and random fluctuations in weather. *The American Economic Review*, 97(1):354–385, 2007.

C. N. Detroit. Auto production resuming after big storm, February 2 2011. URL `http://detroit.cbslocal.com/2011/02/02/snowstorm-idles-auto-plants/`.

R. Dholakia, M. Zhao, and N. Dholakia. Multichannel retailing: A case study of early experiences. *Journal of Interactive Marketing*, 19(2):63–74, 2005.

S. G. Donald and W. K. Newey. Choosing the number of instruments. *Econometrica*, 69 (5):pp. 1161–1191, 2001. ISSN 00129682.

L. Dong and B. Tomlin. Managing disruption risk: The interplay between operations and insurance. *Management Science*, 58(10):1898–1915, 2012.

C. Field, V. Barros, T. Stocker, D. Qin, D. Dokken, K. Ebi, M. Mastrandrea, K. Mach, G.-K. Plattner, S. Allen, M. Tignor, and P. Midgley. *Managing the Risks of Extreme Events and Disasters to Advance Climate Change Adaptation.* Cambridge University Press, 32 Avenue of the Americas, New York, NY 10013-2473, USA, 2012.

M. Fisher, J. Krishnan, and S. Netessine. Retail store execution: An empirical study. *University of Pennsylvania, the Wharton School and Research Center: Operations and Information Management Department. Retrieved December*, 1:2006, 2006.

G. Fitzsimons. Consumer response to stockouts. *Journal of Consumer Research*, 27(2): 249–266, 2000.

C. Forman, A. Ghose, and A. Goldfarb. Competition between local and electronic markets: How the benefit of buying online depends on where you live. *Management Science*, 55 (1):47–57, 2009.

I. Geyskens, K. Gielens, and M. Dekimpe. The market valuation of internet channel additions. *The Journal of Marketing*, pages 102–119, 2002.

T. Golob and W. Recker. Relationships among urban freeway accidents, traffic flow, weather, and lighting conditions. *Journal of Transportation Engineering*, 129:342, 2003.

A. Gopal, M. Goyal, S. Netessine, and M. Reindorp. The impact of new product introduction on plant productivity in the north american automotive industry. INSEAD Working Paper No. 2012/15/TOM. Available at SSRN: http://ssrn.com/abstract=1750123, February 2012.

N. Granados, A. Gupta, and R. Kauffman. Online and offline demand and price elasticities: Evidence from the air travel industry. *Information Systems Research Forthcoming*, 2011.

K. Hendricks and V. Singhal. Association between supply chain glitches and operating performance. *Management Science*, 51:695–711, 2005.

J. Hilbe. *Negative Binomial Regression*. Cambridge University Press, 2011. ISBN 9780521198158.

S. J. Hoch, E. T. Bradlow, and B. Wansink. The variety of an assortment. *Marketing Science*, 18(4):pp. 527–546, 1999. ISSN 07322399. URL `http://www.jstor.org/stable/193241`.

S. Hsiang. Temperatures and cyclones strongly associated with economic production in the caribbean and central america. *Proceedings of the National Academy of Sciences*, 107 (35):15367, 2010.

S. S. Iyengar and M. R. Lepper. When choice is demotivating: Can one desire too much of a good thing? *Journal of Personality & Social Psychology*, 79:995–1006, 2000. ISSN 00223514.

K. Jerath, A. Kumar, and S. Netessine. An information stock model of customer behavior in multichannel customer support services. 2012.

R. Johnson and D. Wichern. *Applied multivariate statistical analysis*, volume 4. Prentice hall Englewood Cliffs, NJ, 1992.

B. Jones and B. Olken. Climate shocks and exports, 2010.

T. Kloek and L. B. M. Mennes. Simultaneous equations estimation based on principal components of predetermined variables. *Econometrica*, 28(1):pp. 45–61, 1960. ISSN 00129682.

A. Kok and M. Fisher. Demand estimation and assortment optimization under substitution: Methodology and application. *Operations Research*, 55(6):1001–1021, 2007.

D. Kuksov and J. M. Villas-Boas. When more alternatives lead to less choice. *Marketing Science*, 29(3):507–524, 2010. ISSN 07322399.

F. G. Lee, J.L. and B. Staats. Rainmakers: Why bad weather means good productivity. Technical report, Harvard Business School working paper 13-0005, 2012.

M. Lewis, V. Singh, and S. Fay. An empirical study of the impact of nonlinear shipping and handling fees on purchase incidence and expenditure decisions. *Marketing Science*, pages 51–64, 2006.

E. Lieber and C. Syverson. Online vs. offline competition, 2011.

M. B. Lieberman and L. Demeester. Inventory reduction and productivity growth: Linkages in the japanese automotive industry. *Management Science*, 45(4):pp. 466–485, 1999. ISSN 00251909. URL `http://www.jstor.org/stable/2634818`.

M. B. Lieberman, L. J. Lau, and M. D. Williams. Firm-level productivity and management influence: A comparison of u.s. and japanese automobile producers. *Management Science*, 36(10):1193–1215, 1990. doi: 10.1287/mnsc.36.10.1193. URL `http://mansci.journal.informs.org/content/36/10/1193.abstract`.

R. Mendelsohn, W. Nordhaus, and D. Shaw. The impact of global warming on agriculture: a ricardian analysis. *The American Economic Review*, pages 753–771, 1994.

A. Moreno and C. Terwiesch. Pricing and production flexibility: An empirical analysis of the u.s. automotive industry. *Working Paper*, 2012.

K. B. Murray, F. D. Muro, A. Finn, and P. P. Leszczyc. The effect of weather on consumer spending. *Journal of Retailing and Consumer Services*, 17(6):512 – 520, 2010. ISSN 0969-6989.

A. Musalem, M. Olivares, E. Bradlow, C. Terwiesch, and D. Corsten. Structural estimation of the effect of out-of-stocks. *Management Science*, 52(1):113, 2012.

S. Neslin, D. Grewal, R. Leghorn, V. Shankar, M. Teerling, J. Thomas, and P. Verhoef. Challenges and opportunities in multichannel customer management. *Journal of Service Research*, 9(2):95–112, 2006.

S. Netessine and N. Rudi. Supply chain choice on the internet. *Management Science*, pages 844–864, 2006.

W. Nordhaud. Geography and macroeconomics: New data and new findings. *Proceedings of the National Academy of Sciences*, pages 3510–17, 2006.

E. Ofek, Z. Katona, and M. Sarvary. Bricks and clicks: The impact of product returns on the strategies of multichannel retailers. *Marketing Science*, 30(1):42–60, 2011.

J. Olesen and M. Bindi. Consequences of climate change for european agricultural productivity, land use and policy. *European Journal of Agronomy*, 16(4):239–262, 2002.

K. Pauwels, S. Neslin, and M. S. Institute. *Building with bricks and mortar: The revenue impact of opening physical stores in a multichannel environment.* Marketing Science Institute, 2008.

O. Perdikaki, S. Kesavan, and J. Swaminathan. Effect of traffic on sales and conversion rates of retail stores. *Manufacturing and Service Operations Management*, 14(1):145, 2012.

R. B. D. Y. E. T. R. A. D. F. W. G. A. N. J. R. Pryor, S.C. Wind speed trends over the contiguous united states. *Journal of Geophysical Research*, 114:D14105, 2009.

L. Qian and G. J. van Ryzin. Strategic capacity rationing to induce early purchases. *Management Science*, 54(6):1115–1131, 2008. ISSN 00251909.

J. Ramsey and S. Morrissey. Isodecrement curves for task performance in hot environments. *Applied Ergonomics*, 9(2):66–72, 1978.

T. Randall, S. Netessine, and N. Rudi. An empirical examination of the decision to invest in fulfillment capabilities: A study of internet retailers. *Management Science*, 52(4): 567–580, 2006.

S. Rao, S. Griffis, and T. Goldsby. Failure to deliver? linking online order fulfillment glitches with future purchase behavior. *Journal of Operations Management*, 2011.

B. Schwartz. *The Paradox of Choice: Why More Is Less.* Ecco/HarperCollins Publishers, 2004.

U. Simonsohn. Weather to go to college. *The Economic Journal*, 120(543):270–280, 2010.

S. A. Smith and D. D. Achabal. Clearance pricing and inventory policies for retail chains. *Management Science*, 44(3):285–300, 1998. ISSN 00251909.

S. Soper. Inside amazon's warehouse: Lehigh valley workers tell of brutal heat, dizzying pace at online retailer. *The Morning Call*, 2011.

Staff. Weather is the leading culprit for supply chain disruptions. *Supply Chain Quarterly*, Q4, 2011.

A. T. Steele. Weather's effect on the sales of a department store. *Journal of Marketing*, 15 (4):pp. 436–443, 1951. ISSN 00222429.

A. Stock and S. Balachander. The making of a "hot product": A signaling explanation of marketers' scarcity strategy. *Management Science*, 51(8):1181–1192, 2005. ISSN 00251909.

X. Su and F. Zhang. On the value of commitment and availability guarantees when selling to strategic consumers. *Management Science*, 55(5):713–726, 2009.

R. Swinney. Inventory pooling with strategic consumers: Operational and behavioral benefits. 2011.

K. Talluri and G. van Ryzin. Revenue management under a general discrete choice model of consumer behavior. *Management Science*, 50(1):pp. 15–33, 2004. ISSN 00251909.

B. Tomlin. On the value of mitigation and contingency strategies for managing supply chain disruption risks. *Management Science*, 52(5):639–657, 2006.

K. E. Train. *Discrete Choice Methods with Simulation.* Cambridge University Press, 2 edition, June 2009.

W. G. Wang, Y. and B. Tomlin. Mitigating supply risk: Dual sourcing or process improvement? *Manufacturing & Service Operations Management*, 12(3):489–510, 2010.

J. M. Wooldridge. *Econometric Analysis of Cross Section and Panel Data*. The MIT Press, Cambridge, Massachusetts, second edition edition, 2010.

S. Worchel, J. Lee, and A. Adewole. Effects of supply and demand on ratings of object value. *Journal of Personality and Social Psychology*, 32(5):906 – 914, 1975. ISSN 0022-3514.

D. Wyon. *Thermal effects on performace. Indoor Air Quality Handbook.* McGraw-Hill Professional, 2001.

F. Zettelmeyer. Expanding to the internet: pricing and communications strategies when firms compete on multiple channels. *Journal of Marketing Research*, 37(3):292–308, 2000.

F. Zettelmeyer, F. S. Morton, and J. Silva-Risso. Scarcity rents in car retailing: Evidence from inventory fluctuations at dealerships. *NBER*, 2006.