

What can be learned from classical inventory models: a cross-industry empirical investigation¹

Serguei Roumiantsev Serguei Netessine
roumiant@wharton.upenn.edu netessine@wharton.upenn.edu
The Wharton School
University of Pennsylvania

April 2005, revised March 2006 and August 2006

Abstract: Classical inventory models offer a variety of insights into the optimal way to manage inventories of individual products. However, top managers and industry analysts are often concerned with the aggregate macroscopic view of a firm's inventory rather than with the inventories of individual products. Given that classical inventory models often do not account for many practical considerations that a company's management faces (e.g., competition, industry dynamics, business cycles, the financial state of the company and of the economy, etc.) and that they are derived at the product and not at the firm level, can insights from these models be used to explain the inventory dynamics of entire companies? This exploratory study aims to address this issue using empirical data.

We analyze absolute and relative inventories using a quarterly data panel that contains 722 public US companies for the period 1992 to 2002. We have chosen companies that are not widely diversified and whose business in large part relies on inventory management in order to concentrate on empirically testing hypotheses derived from a variety of classical inventory models (EOQ, (Q,r), newsvendor, periodic review, etc.). We find empirical evidence that firms operating with more uncertain demand, longer lead times and higher gross margins have higher inventory levels. Furthermore, larger companies appear to benefit from economies of scale and therefore have relatively less inventory than smaller companies. We obtain mixed evidence on the relationship between inventory levels and inventory holding costs. We also analyze the breakdown of data into eight segments—oil and gas, electronics, wholesale, retail, machinery, hardware, food and chemicals—and find that, with a few notable exceptions, our hypotheses are supported within the segments as well. Overall, our results demonstrate that many of the predictions from classical inventory models extend beyond individual products to the more aggregate firm level; hence, these models can help with high-level strategic choices in addition to tactical decisions.

¹ The authors gratefully acknowledge financial support from the Fishman-Davidson Center for Service and Operations Management. We are indebted to Vishal Gaur, Lorin Hitt, Taylor Randall, Justin Ren, Christian Terwiesch and Anita Tucker, who provided detailed and thoughtful comments on the previous version of this paper. Helpful comments from the special issue editor Aleda Roth, senior editor and two referees are gratefully acknowledged.

1. Introduction

Inventory management theory is mostly concerned with finding the optimal way to manage inventories given exogenous and usually static business environments. As a result, most of the classical inventory models are normative in nature, since they prescribe how rational agents must behave. Modeling and analysis typically occur at the microscopic level of a representative product or group of products, so that the entire firm is analogous to a “black box” and the analysis is conducted “inside the black box.” The potential flaw of the inside-the-black-box approach is that simplified assumptions made by the modeler may not reflect business realities because they may not (and typically do not) account for competition, business cycles, industry trends, a company’s financial distress, etc. Additionally, most of the models are derived for specific product(s)/echelon(s) and ignore the complexity of supply chains encountered in practice. Hence, it is not clear if high-level managers and industry analysts can benefit from understanding classical inventory models (which are typically taught to MBA students), because their concern is aggregate inventory behavior at the firm level, which can be dominated by effects other than those accounted for in inventory models.

The field of macroeconomics takes an entirely different view: it looks at firms from “outside the black box” by analyzing the surrounding environment and studying the aggregate behavior of the economy/industry. The goal is to analyze the partial or general equilibrium in an industry or in the economy and to link various macroeconomic indicators by looking at the joint dynamics of their time series. However, without looking inside the black box the macroeconomic approach is less useful in describing drivers of individual firms’ inventory behavior. For example, the macroeconomic approach cannot predict the relationship between lead times and inventories, whereas classical inventory models do offer the necessary intuition.

Our goal in this paper is to analyze whether the insights from inside-the-black-box analytical models developed by operations researchers are consistent with outside-the-black-box macroeconomic data. While classical inventory models are successfully applied in practice at the tactical level, the strategic decisions of high-level managers are often segregated from the tactical and data-driven operational decisions at the product and plant levels. By exploring the link between macroeconomic data and classical inventory models, we contribute to the operations literature by demonstrating empirically that many insights from classical inventory models continue to hold at the company level, and therefore understanding these models can aid in the managerial decision-making process.

We begin by describing our research methodology and by pointing out two crucial aspects of firm-level empirical inventory analysis: the issue of time and space aggregation and the difference between the prescriptive nature of underlying inventory models (“how much to order?”) and the descriptive nature of observed parameters in practice (“what is the inventory level for a firm?”). We then formulate six hypotheses postulating monotone relationships between aggregate inventory levels and environmental variables such as mean demand and demand uncertainty, lead times, margins, inventory holding costs and the extent of economies of scale. All these hypotheses are formulated using well-known results from classical inventory models such as EOQ (Economic Order Quantity), newsvendor, (Q,r) , periodic review and other models. We use quarterly data containing 44 time points for each of 722 companies from 1992 to 2002, representing about 30% of US manufacturing and retailing inventories. We find that our data is consistent with predictions from these models that higher inventory levels are associated with higher demand uncertainty, longer lead times, higher margins and lower economies of scale. We also find mixed evidence on the relationship between inventory levels and inventory holding costs. In addition to the aggregate analysis, we test our hypotheses across eight segments of the economy. We find that most of the hypotheses are robustly supported in split regressions as well, although there are a few notable exceptions.

Our second contribution is to quantify the association between the environmental variables and inventories at the firm level. Namely, we estimate the elasticity of inventory levels to changes in mean demand, demand uncertainty, earnings uncertainty, lead times, margins, inventory holding costs, and several other control variables. Hence, we evaluate empirically the magnitude of the comparative statics that are often conducted in the inventory management literature. We find that, for the absolute inventory model, COGS, demand uncertainty, margins, lead times and the time trend explain approximately 70%, 7%, 5%, 2% and 1% of variance in inventories, respectively. Whereas classical inventory models only quantify the impact of environmental variables in a stylized setting that abstracts away many real-world effects, our analysis estimates the impact of these variables upon inventory across segments, as likely to be observed by management in practice. Our results can be used by management to budget inventory requirements by, for example, taking into account the relationship between scale economies and inventory. Although our study is exploratory in nature, it offers several fruitful directions for future research that might shed additional light on the behavior of inventories and the predictive power of classical inventory models at the firm level.

The rest of the paper is organized as follows. In Section 2 we review the literature on related research in operations management and economics. In Section 3 we formulate hypotheses to be tested,

and in Section 4 we describe data. In Section 5 we define the variables of the regressions, and in Section 6 we specify the econometric model. Results are presented in Section 7, and Section 8 concludes the paper with a discussion of results, limitations and directions for future research.

2. Literature review

Inventory modeling has been an area of intense inquiry in operations management and operations research. Starting from a simple deterministic Economic Order Quantity (EOQ) model dating from over a century ago, the field of operations management has developed much more advanced inventory models that incorporate stochastic and correlated demands, multiple products and multiple echelons of inventory. Some widely used inventory models are described in Silver et al. (1998), Zipkin (2000), Porteus (2002) and Cachon and Terwiesch (2005).

In an anniversary article, Wagner (2002) underscores the issues of aggregation and data availability. He points out that product-level theoretical modeling historically dominated the field of operations research, whereas industry-level data modeling was widely used in economics. As a result, there are few if any implementable solutions at the product level and especially at the firm level of aggregation. He also suggests that the field of inventory management started getting more attention when more product-level and firm-level data became available at the SKU level and when effects that are very often beyond the scope of classical inventory models became obvious (e.g., skewness and discontinuity of demand, promotions, competition, etc.). Wagner suggests that the pure operations research model “is blind to data issues” and that “the objective should be stocking and replenishing logic that is driven by such data.”

At the same time, only a few papers in operations management analyze inventories empirically and try to reconcile the inventory behavior observed in practice with the behavior predicted by the models. Rajagopalan and Malhotra (2001) study trends in inventory levels at US firms over time to test the widely held belief that inventory management has improved due to the introduction of just-in-time (JIT) practices and IT system implementations. Using a large sample of firms from the US Census Bureau including both private and public companies, they find that material and work-in-process inventories decreased in the majority of the two-digit SIC industries from 1961 to 1994. Furthermore, in some segments there were greater improvements after the 1980s, when JIT practices were adopted. Rajagopalan and Malhotra (2001) also find that inventories of finished goods decreased in only a few industry segments. Chen et al. (2005a, 2005b) analyze inventory trends for US public companies in the manufacturing and retailing sectors. They find that between 1981 and 2001 median manufacturing

inventory levels declined from 96 days to 81 days, with an average rate of reduction of 2% per year. In the retail and wholesale segment, the median inventory levels decreased from 72 days to 52 days. Moreover, Chen et al. (2005a, 2005b) show that public companies with abnormally high inventory levels have experienced abnormally low levels of financial returns, but on average lower inventory levels are not associated with higher financial returns. Lai (2006) reports evidence that, when the market discounts high inventory firms, firms decrease inventory, and vice versa.

Gaur et al. (2005) examine firm-level inventory behavior in retailing companies. They propose a model explaining differences in inventory turns across companies and create an adjusted measure of inventory turns that is better suited to gauging the operational metrics of retailers. Gaur et al. (2005) also find that inventory turnover for retailing firms is positively correlated with capital intensity and sales surprise and, similar to our result, is negatively correlated with gross margins. Gaur et al. (1999) demonstrate that the financial excellence of retailing companies comes from various operational strategies that may involve low or high product margins and low or high inventory turns in different retailing segments. Although all of these studies deal with aggregate companies' inventories, they pursue different goals and, with the exception of one hypothesis in Gaur et al. (2005), do not test the implications stemming from classical inventory models.

Several studies analyze inventories in the automotive industry. Fisher and Ittner (1999) study the impact of product variety on automotive assembly plant operations and find that increased option content variability in car assembly has an adverse effect on plants' operational performance, which is manifested in higher total labor hours, overhead hours, downtime hours, rework and inventory levels. Lieberman and Asaba (1997) and Lieberman and Demeester (1999) study inventory management and JIT practices in the Japanese auto industry and compare these to the US automotive segment. They find that Japanese companies are leaner and that the introduction of JIT systems in US companies has helped them to become leaner as well. Lieberman et al. (1999) study the dynamics of inventory levels for automotive suppliers in North America. They use a combined survey and secondary plant-level data to show that inventory levels are affected both by technological and managerial factors in a manner consistent with classical inventory theory. Namely, they show that inventory levels at selected plants are increasing with setup costs, per unit item costs, and production lead times and that inventories are lower for plants in which the workforce engages in making process improvements. Surprisingly, the plants of Japanese companies in North America hold no less inventory than the plants of American companies. This study is, perhaps, the closest to ours in that some of the hypotheses tested follow directly from classical inventory models. We, however, study a more diverse set of companies that belong to eight

different industrial segments and test several relationships not investigated in Lieberman et al. (1999). We also extend our analysis to quantifying the elasticities of inventory as a decision variable with respect to other operating variables and find differences across segments that can be used to aid in managerial decision-making.

Several other papers do not study inventories directly but focus on issues related to supply chain management and are therefore relevant to our work. Hendricks and Singhal (2005) show that supply chain disruptions are very costly to public companies, since they cause a substantial loss in market value. Randall and Ulrich (2001) investigate the relationships among product variety, supply chain structure and firm performance. They show that matching the supply chain structure to the type of product variety results in the higher financial performance of companies in the American bicycle industry. Randall et al. (2006) find that Internet retailers selecting supply chain structures in accordance with classical inventory models are less likely to go bankrupt. Cachon et al. (2004) find evidence that demand variability is often higher in the lower echelons of the supply chain than in the higher echelons, which is contrary to predictions based on the widely cited bullwhip effect.

The last related stream of operations literature attempts to estimate empirically the parameters that companies use to make decisions using the classical newsvendor inventory model. Cohen et al. (2003) use data on demand forecasting, actual orders, production lead times, and delivery dates for a semiconductor supply chain to show that the supplier perceives the cost of order cancellation to be about two times higher and the holding cost to be about three times higher than the delay cost. Olivares et al. (2004) show how to estimate overage and underage costs using data for operating room capacity reservations at hospitals. In our paper we do not attempt to estimate the parameters used by managers in practice, for we believe that the aggregate company data with which we are working is too crude for these purposes.

Finally, inventory has long been a focus of research in the field of macroeconomics, mostly from the viewpoint of aggregate economy dynamics. Kahn (1987, 1992), Kahn and Bils (2000) and Blinder and Maccini (1991) indicate that inventory investments (changes in inventory levels) have been the main source for US GDP fluctuations and hence understanding inventories would help us to understand the behavior of the economy in general. Amihud and Mendelson (1989) examine the empirical relation between the market power expressed either by the market share or by Lerner's index and demonstrate that both are positively related to mean inventory levels and their volatility. Reassuringly, this finding is consistent with predictions from economic models. (We refer to economic models here, because classical inventory models ignore competition.)

Blinder and Maccini (1991) provide a detailed summary of inventory research in economics, which dried out somewhat in the late 1980s. As Ramey and West (1999) summarize, economists have proposed stock-adjustment, production-smoothing, and other models to link inventory with production, sales and GDP to explain two main stylized facts about inventory behavior, namely (i) the procyclical nature of aggregate economy-level inventory dynamics and (ii) the persistence in the relationship between sales, production and inventory in the form of production smoothing (i.e., lower variability of production than variability of sales and the stabilizing role of inventory). To date, little evidence consistent with these models has been found. Instead, data seem to indicate that (i) based on the de-seasonalized data, in the US production has been more volatile than sales (see Kahn 1992), which contradicts the linear-quadratic stock-adjustment model in Ramey and West (1999) and (ii) the imputed speed of inventory adjustments based on linear-quadratic models is unrealistically low (see Blinder 1986 and Blinder and Maccini 1991). Moreover, there is some evidence that the (S,s) inventory model (introduced in Arrow et al. 1951 and extended in numerous operations research papers) may better explain inventory behavior (see Blinder and Maccini 1991).

Caplin (1985) demonstrates that the (S,s) model, even after aggregation over multiple firms, captures the fact that production is more volatile than sales in the U.S. retail industry. Mosser (1991) finds empirical support for the (S,s) model by showing that industry-level sales and inventory data support the (S,s) model and not the linear-quadratic model. Both Caplin (1985) and Mosser (1991) make important contributions by addressing the data aggregation issue. In most economics papers the analysis is done at the industry level using the linear-quadratic model, with few attempts to disaggregate this approach to the firm-level data. On the contrary, most of the classical inventory models are derived at the product level and therefore there is a challenge in product-to-firm aggregation, the issue that Caplin (1985) partially addresses.

Finally, several studies consider the relationship between product variety and other variables. We do not account for product variety in our work but, since its impact is potentially important, it is useful to understand the effects that omitting product variety may impose on our results. Extant theory postulates that larger product variety allows firms to increase sales and profit margins but results in larger inventory and production costs. Empirical evidence related to product variety is mixed and limited due to the difficulties inherent in measuring product variety. Kekre and Srinivasan (1990) find that larger product variety is associated with a larger sales/market share and do not find any evidence that larger product variety translates into higher costs. Bayus and Putsis (1999), on the other hand, find that the costs (including inventory and production costs) of wider product lines outweigh any benefits.

Cachon et al. (2005) find that product variety is associated with greater finished goods inventory and increased sales but not with higher profit margins.

3. Formulation of research hypotheses

Firms hold inventories for at least one of the following reasons: production typically does not occur at the same time and place as demand (in other words, there are procurement lead times); production capacity might be rigid but demand is typically variable; there are economies of scale in handling inventories; or there is non-stationarity (seasonality, stochasticity) in demand and/or supply. In this paper we touch upon each of these reasons to hold inventories, although our analysis is limited by data availability, so we might be unable to account for all drivers of inventories that exist in practice.

To be able to use classical inventory models to guide our analysis of inventories at the firm level, we need to address at least two major methodological challenges. First, we need to understand what insights from classical inventory models, if any, might continue to hold after aggregating data across time and space (e.g., multiple products/echelons). We need to address this issue, because classical inventory models are typically formulated at the single-product level, although there are multi-product extensions. We argue that many structural properties of inventory models will not survive aggregation, because there is no way to identify statistically product-specific lead times and cost and revenue parameters as well as demand parameters unless we make an unrealistic assumption that products are perfectly homogenous (in the space of the model parameters) and that the inventory control system is synchronized across products.

On the other hand, we argue that *monotone* (or comparative statics) properties of inventory models will survive aggregation across time (from daily inventory decisions to quarterly data) and space (from the product to the firm-level). For example, if we believe that the inventory of every product decreases in inventory holding costs, then the same will be true for a firm's inventory that combines millions of products. Thus, we hope to see those monotone properties in the aggregate data despite all the noise of aggregating across product lines that are managed differently. Conceptually, this argument is similar to that of Caplin (1985), who demonstrates that a particular property of the (S,s) model (i.e., amplification of demand uncertainty) survives aggregation, as later confirmed empirically by Mosser (1991) using aggregate industry-level analysis.

Inventory, unlike sales, is one of the few variables that are largely (but not completely) determined by internal decision makers. Thus, the second methodological challenge is to understand the relationship between observations of operational variables (inventory, sales, etc.) over the course of a

firm's operations and the decision-making process within the firm. Indeed, prescriptive inventory models typically provide us with the optimal inventory positions at the beginning of the replenishment cycle whereas firm-level data only allows us to observe inventories at the end of certain time periods (months, quarters, years), which do not necessarily coincide with either the beginning or the end of the inventory replenishment cycle. However, we argue that we can use even these snapshots of aggregate inventories to analyze monotone properties derived from the classical inventory models, because these properties apply equally to the optimal inventory decisions as well as to observations of inventory positions at random points during the review period. For example, longer lead times are associated with higher order-up-to levels in the periodic review model and hence with a higher inventory position throughout the review period. Thus, it is methodologically feasible to measure stochastic time series realizations of various variables over discrete points of time and to relate them to each other. Similar logic applies to many other empirical studies in the social sciences, such as those that observe stock price dynamics to judge investor expectations in finance.

Throughout the paper we refer to our findings as *associations* and we do not try to test/impose causality, because it is hard to do so using only realizations (sample paths) of variables instead of the true population and the underlying data available for the decision-making process. Our hypotheses rely on the assumption that firms behave rationally and make inventory decisions to maximize expected profit (or to minimize expected cost, as is the case in many classical inventory models). Hence, firms' behavior is manifested in the relationship among such operational variables as inventory, sales, margins, demand uncertainty, lead times and inventory holding costs, which should be consistent with classical inventory models. Clearly, there are many variations of classical inventory models accounting for various phenomena observed in practice. Understandably, some of our hypotheses may not be consistent with certain variations of these models. Thus, when discussing results following from classical inventory models, we refer to the most traditional formulations. For example, it is assumed throughout the paper that demand is stationary and independent across time periods in the periodic review models, etc. We do, however, introduce control variables and proxies that account for behavioral aspects of inventory management and do not rely on classical inventory models.

Our first hypothesis comes from the combination of the classical EOQ model and stochastic inventory models (see Silver et al. 1998) and relates absolute inventories and mean demand. The EOQ model evaluates a trade-off between fixed ordering costs and inventory holding costs, resulting in the expression for the optimal order quantity (which is proportional to inventory) that exhibits a square-root dependence with respect to demand. This simple inventory model has proven its reliability in a variety

of applications in part because it is very robust to parameter changes and in part because it has been adopted in environments with stochastic demand. For example, Johnson and Montgomery (1974, p. 59) demonstrate that when demand is stochastic (stationary), inventory is reviewed continuously and the policy is to order the same batch Q when inventory drops below reorder point r , the dependence between inventory and mean demand is still governed by the square root formula (both under lost-sales and back-ordering assumptions). On the other hand, if inventory ordering costs are minimal (as in, e.g., the newsvendor or the order-up-to models in Cachon and Terwiesch 2005), the relationship between inventory and the optimal order quantity is linear. When we aggregate over various products utilizing variants of models with and without ordering costs, we expect to see a concave relationship between expected demand and average inventory levels (i.e., somewhere between linear and square root functions). Therefore:

H1: Inventory level is positively associated with mean demand through a concave function.

Hypothesis 1 can be generalized somewhat. Inventory management is typically subject to economies of scale because manufacturing and distribution tasks involve significant fixed costs. The EOQ model accounts for fixed inventory ordering costs and demonstrates that inventory is increasing more slowly than demand due to the fixed costs of placing replenishment orders. Moreover, there are also statistical economies of scale in inventory management. For example, in a stochastic environment (e.g., in the newsvendor model), higher mean demand typically corresponds to a higher standard deviation of demand. However, if we increase both mean demand and standard deviation of demand while holding their ratio fixed, it is easy to verify that inventory will increase at a rate lower than mean demand. Hence, a larger firm (a firm facing larger mean demand) should enjoy economies of scale in inventory management. Eppen (1979) generalized this notion under the umbrella of risk pooling: a larger firm can pool together demand from many locations/stores/warehouses/products, resulting in less inventory in relative terms. Hence, whether Hypothesis 1 holds or not, it should still be the case that larger companies are able to hold relatively less inventory. Thus:

H2: Relative inventory level (i.e., the ratio of inventory to sales) is negatively associated with company size.

Our third hypothesis postulates the relationship between inventory and demand uncertainty. Many classical inventory models lead to the well-known result that inventory levels increase with demand uncertainty, i.e., that firms buffer inventories against demand uncertainty. The single-period newsvendor inventory model leads to this result when the cost of having too little inventory is lower

than the cost of having too much inventory, the situation encountered in the vast majority of practical situations.² This result holds irrespective of problem parameters for the end-of-period inventory in periodic review inventory models (see Cachon and Terwiesch 2005, p. 259), multi-echelon inventory models and many others. The underlying intuition for this result is that firms need more slack (in the form of spare inventory or safety stock) to be able to respond to large swings in demand. Hence, in line with the extant inventory theory we hypothesize that:

H3: Inventory level is positively associated with demand uncertainty.

Similar to buffering inventory against demand uncertainty, companies should buffer inventory against longer lead times. Longer lead times increase the intervals between product deliveries and hence require more inventory to satisfy demand while awaiting replenishment. Several classical inventory models demonstrate this effect. Cachon and Terwiesch (2005) discuss periodic review inventory models with deterministic lead times and demonstrate that the longer the lead time, the higher the inventory level. Similarly, in models with stochastic lead times (see, e.g., Zipkin 2000) both the mean and variance of the lead time increase the amount of inventory the company needs to hold. Lieberman et al. (1999) propose the same hypothesis. In line with these results, we hypothesize that:

H4: Inventory level is positively associated with procurement lead times.

Most classical stochastic inventory models focus on the trade-off between underage costs (the cost of having too little inventory) and overage costs (the cost of having too much inventory). This trade-off is best reflected in the solution to the newsvendor model postulating that the optimal order quantity is located at the fractile of demand distribution that is equal to the ratio of underage cost to the sum of underage and overage costs (Cachon and Terwiesch 2005). Many of the more elaborate inventory models including periodic review (S,s) , (Q,r) and multi-echelon models follow the same logic. Hence, all else being equal, larger underage costs lead to higher inventory levels in these models (see Silver et al. 1998). In the newsvendor model the underage cost is typically taken as the product's gross margin. The same assumption is often made in multi-period models with lost sales. Even in multi-period models with full back-ordering the penalty for stocking out is likely to be correlated with the product's margin (see Cachon and Terwiesch 2005): customers buying high-margin products are likely to be more sensitive to stock-outs. Hence, we hypothesize that more inventory will be associated with

² In the newsvendor model with symmetric demand distribution, the sufficient condition for this relationship is the critical ratio that is higher than 0.5.

higher product margins, because higher margins correspond to a higher cost of underage (see Gaur et al. 2005 for a similar hypothesis with respect to inventory turnover). Thus:

H5: Inventory level is positively associated with product margins.

Our final hypothesis concerns inventory holding costs. In the EOQ model the level of inventory is decreasing in the inventory holding cost parameter, an intuitive proposition, because the amount of inventory should be decreased as holding inventories becomes more expensive. In stochastic multi-period inventory models, the inventory holding cost appears as the cost of overage. Hence, once again, the higher the cost of overage, the less inventory should be stocked. Consistent with these observations, we hypothesize that:

H6: Inventory level is negatively associated with inventory holding costs.

4. Data description

We use a representative sample of public US companies obtained from the COMPUSTAT financial database through Wharton Research Data Services. Public companies are obliged to provide operational and financial information following GAAP standards to ensure that investors have access to data regarding their performance dynamics. The choice of public companies restricts our findings from being representative of the whole US economy. However, due to the lack of reliable operational data for private companies, we focus on public companies alone.

We use quarterly data containing 44 time points between 1992 and 2002 for every company in our sample. The 1992 cutoff is somewhat arbitrary, but it allows us to analyze the most recent data that is less affected by such factors as price inflation and changing industry structure. This time period also helps us to minimize panel attrition. We utilize quarterly rather than annual data to account for seasonal inventory fluctuations within a given year (i.e., demand/inventory shifting across quarters), which has a major and statistically significant impact in many industries. Moreover, quarterly data allows us to obtain more accurate estimates of demand uncertainty than annual data. We use calendar quarters instead of fiscal quarters since companies have different fiscal periods (i.e., fiscal periods starting in different quarters, which leads to different codifying). Using quarterly data we cannot obtain separate information on different inventory types (raw materials, work in process, finished goods), whereas this information is available in the annual data (see Rajagopalan and Malhotra 2001). However, we do not perceive this issue to be significant, because our goal is to study inventories at the aggregate company level, and our hypotheses would not differ by inventory types. Although more frequent (say, monthly or

weekly) data may seem a good alternative to quarterly data, the monthly survey data provided by the US Census are insufficient for our analysis, since they do not track revenues and costs.

Working with a panel of data allows us to be certain that the statistical relations we obtain are neither applicable at only a single point of time (the cross-sectional aspect of analysis) nor driven by a single company (the time series aspect of analysis). We use pooled and segment-specific estimations to test whether our hypotheses hold within certain segments, and we ensure that possible biases are reduced through proper panel data estimation.

The sample itself was selected as follows. We excluded service, construction and transportation industries and focused only on inventory-rich economy segments: minerals and mining (SIC codes 1000-1499), manufacturing (SIC codes 2000-3999) and wholesale and retail segments (SIC codes 5000-5999). First, we decided to focus on eight industry segments (oil and gas, wholesale, retail, consumer electronics, food, chemicals, hardware, and machinery) in which the importance of inventory is particularly obvious. We then selected at random several two-digit SIC codes within these eight segments of the economy. From those SIC codes we further selected all companies that were active in the period between 1992 and 2002 and have not merged and have not been acquired. Next, we excluded companies that had fewer than \$5M in sales cumulatively over 10 years and those that had zero sales and inventory data for the first three years of data, even if they were otherwise active. The purpose of the filtering process was to ensure that the final sample contained only companies that had been actively operating in inventory-related business activities: retailing, distribution or manufacturing. While it is possible to explicitly control for exits and entrances and to analyze their effects on inventories, these effects are outside the scope of our study.

We obtained a final sample of 722 companies including 233 S&P500 companies. To make sure that our sample was representative of the US economy as a whole, we verified that the total inventory in our sample represented 30% of the total US manufacturing and retailing business inventory and, moreover, that it is strongly correlated with the total US inventory (Pearson $r=0.91$, $p<0.001$). We could also have used the total COMPUSTAT inventory-related data population, but COMPUSTAT itself is only a sample of public companies out of the universe of companies operating in the US and is subject to such potential flaws as survivorship bias and size bias. We proceeded with the analysis of the 722 sampled companies in the hope of obtaining robust and repeatable results across different yet not-too-numerous segments that would indicate that the sample properties are representative of the population properties. The usage of a limited number of segments allows us to compare segment-specific results; otherwise the potential number of segments would be too large to fit within one paper. For that reason,

previous studies often focused on retailing/wholesaling/manufacturing segments only (see Gaur et al. 2005, Chen et al. 2005a, 2005b).

The obvious disadvantage of using COMPUSTAT as a source of information for testing hypotheses about inventory models is that financial accounting (especially after aggregation) may only crudely reflect actual processes within a company. For example, at the industry level one can use consumer price indexes to express everything in constant dollar terms, an approach that is not applicable for firm-level data. It is a typical practice in empirical research (see Blinder et al. 1981, Kahn 1992) to use deflators while working with economy-level or industry-level data but not to use common deflators (e.g., inflation indexes for specific industries) for firm-level analysis, since it is not a valid proxy for capturing firm-level price effects. We partially control for inflation by using variables expressed in the same price terms—inventory and COGS (sales in inventory prices). Price inflation has historically been a smaller source of dollar output fluctuations compared to unit output fluctuations in the US (see Blinder and Maccini 1991).

As we discussed earlier, another potential problem is that the data is aggregated across product lines and production units and time within a company, meaning that different processes with potentially different operational structures are merged. Although this may lead to biases due to a difference in product variety, we, as explained above, expect that the direction of effects should be preserved even at the aggregate level so that conceptually it is possible to capture product-level operational effects in the firm-level data. Follow-ups on our exploratory study might use our basic model but incorporate other variables such as product variety.

Table 2 summarizes the sample. Companies in the sample hold on average \$396M of inventory and have on average \$527M of quarterly sales expressed in input prices. From Table 2 we also see that companies vary in size across segments, with the oil and gas and retail segments having larger companies on average. We also see that relative inventory levels vary by segment: the chemicals, computer hardware and electronics segments have the largest relative inventory levels (1.40, 1.25 and 1.22 respectively for quarterly relative inventory levels), whereas the oil and gas segment appears to be the leanest, with an average relative inventory ratio of only .42. Such preliminary observations point out the heterogeneity of operational and technological factors across industries that impose different conditions on the ways companies operate and make inventory decisions.

5. Description of variables

We use three subscripts to account for time-specific ($t=1, \dots, 44$), company-specific ($i=1, \dots, 722$) and segment-specific ($s=1, \dots, 8$) effects as in Gaur et al. (2005). For *dependent variables*, we use both absolute and relative inventory. Although classical inventory models typically work with absolute inventory levels, we also conduct cross-company analysis and comparison, which is more meaningful for relative inventories. We denote absolute inventory for the firm i in the segment s by Inv_{its} which is measured by the inventory expressed in monetary terms at the end of the corresponding quarter t as reported in quarterly balance statements. Relative inventory is measured by the inventory-to-cost-of-goods-sold (denoted by $COGS_{its}$) ratio for the corresponding quarter taken from the income statement and is denoted by $InvCOGSratio_{its} = Inv_{its} / COGS_{its}$.

For *explanatory variables* we use the following proxies (see summary statistics in Table 4 and definitions in Table 1). To measure *product margins* we use gross margins. Several other margins are available (operating margin EBIT and EBITDA, net margin, returns on assets, equity and sales), but we believe that gross margin is more appropriate, because it does not include fixed costs and other items such as the effects of taxes, amortization, etc. that are not directly related to inventory management. Gross margin is defined as a percentage difference between net sales and the cost of goods sold:

$$GrossMargin_{its} = (Sales_{its} - COGS_{its}) / Sales_{its} .$$

Inventory holding costs typically include two components: the physical cost of holding inventory and the opportunity cost of tying up capital in inventory. Data for the physical cost of holding inventory is not publicly available and is generally difficult to estimate due to several fixed and variable components. Therefore, our proxy for inventory holding costs only accounts for the opportunity cost of capital, so our results on the association between holding costs and inventories may be diluted. However, capital costs tend to account for a larger portion of holding costs and, moreover, as indicated in Irvine (1981), the physical holding cost is much more stable over time than the opportunity cost of capital, so arguably we account for the more important part of the inventory holding cost. We use two proxies for the opportunity cost of holding inventory. The first proxy is the three-month T-bill rate (coded as $TBillRate_t$), which is known to be a good proxy for risk-free interest rates and industry-wide capital opportunity costs and has been used in economics for that reason (Irvine 1981). This is an admittedly crude proxy, as it does not account for differences across companies (in that some companies may borrow and raise capital more cheaply than others). The second proxy we use accounts for firm-specific cost of capital and is denoted by $CostCapital_{its}$. To estimate this proxy we follow the realized returns approach as suggested by Lieberman (1980) and Botosan and Plumbee (2005). Namely, we

estimate historical non-compounded realized quarterly returns on equity by using historical price and dividend data: the last month of the quarter closing price, the first month of the quarter price and dividends paid during the period. Next, we estimate the historical period-average cost of debt outstanding for a company expressed as the ratio of interest paid to debt outstanding. Finally, we estimate the weighed average of the two measures using a company capital (debt/equity) structure.

There are many other ways to estimate the cost of debt/equity and the cost of capital for a firm. This topic has been an area of extensive research in finance and accounting (see Fama and French 1999 and Botosan and Plumbee 2005 for summaries and references). An alternate approach to the backward-looking historically realized returns is typically to use the cost of capital imputation based either on direct expectation-based imputation from both the realized and expected times series of stock prices and dividends or on model-based imputation from CAPM or APT models. Expectation-based or model-based approaches may well serve their roles in predicting future expected stock returns or in analyzing additional factors affecting stock prices, since they are forward-looking by construction, but for our research purposes we believe that realized return usage based on historical stock, dividend and interest paid data is the right choice. The correlation between the two inventory holding cost proxies is $r=.66$ ($p<.01$).

Lead time data is not publicly available and in practice varies both by product and by supplier. As a proxy for a procurement lead time, we use the average days of accounts payable outstanding, such that $LeadTime_{its} = 365 / (4 \times COGS_{its} / AP_{its})$ where AP_{its} stands for accounts payable. Justification for this proxy comes from financial accounting definitions. Production cycle time is defined as the average days of inventory outstanding, sourcing lead time for inputs is defined as the average days of accounts payable outstanding, and cash collection (or output delivery time, or days of sales outstanding) is defined as the average days of accounts receivable outstanding (see Stickney and Weil 1999). Together, these measures define a cash conversion cycle—the average time it takes a dollar of investment to buy inputs, produce, sell outputs and collect cash. Although these measures are only proxies for the physical production cycle and lead times, they provide the right direction of logic; accounts payable are credited, then input product is shipped by the supplier and is typically debited, then it is received and cash is paid for it. Hence, we argue, financial transactions should be correlated with times of shipment and delivery of inputs and therefore should be correlated with the lag a company has to respond to changing market environment by adjusting inventories. It is worth noting that accounts receivables cannot be used as a proxy for lead times because they represent an after-sale process and should not affect a company's ability and need to source/produce units in advance to have inventory ready for sale.

The recognition of shipments/payments is linked to a company's policy of recognizing revenues/expenses and is known to vary by company to some extent. However, since we study public companies that are closely monitored by investors and the Securities and Exchange Commission, in most cases these companies will have practices that are relatively consistent if not in the aggregate then within the industry segment. There are at least three more issues associated with the lead time proxy: payment terms can be dominated by standard arrangements existing in the industry, the timing of recognition of shipments/payments may correlate with the size/power of a company and cash payments might be used instead of accounts. To verify that the lead time proxy is not dominated by standard payment terms (e.g., 30 or 60 days), we analyzed the distribution of $LeadTime_{its}$ and found that it is well approximated by the log-normal distribution with coefficient of variation 3. Hence, we did not find any evidence of "spikes" in payables around some standard values. To ensure that lead times are not entirely driven by the size/power of the company, we also checked for correlation between the inventory level and the lead time proxy and the company's market power as measured by the market share within a four-digit industry segment. We found that both correlations are insignificant and negative. We further verified this result by constructing a regression model with the lead time proxy as a dependent variable. We did not find statistically significant evidence that the lead time proxy is affected by company size as measured both by sales and by market power. Therefore, the lead time proxy does not appear to be affected on average by the ability of a company to change the terms of payments due to market power. Finally, we believe that cash payments are rare relative to payments through accounts payable and receivable, since most public companies operate with suppliers through accounts rather than through cash transactions. We verified that the average ratio of accounts payable to cash holdings for our sample companies exceeds 13, indicating that the use of accounts payable dominates cash payments. In closing, we acknowledge the possibility that, to some extent and for some companies, the LeadTime proxy might be capturing payment lead time rather than sourcing lead time, so that our results with respect to lead times can be diluted. Nevertheless, while payment lead times are not typically a part of classical inventory models, we expect them to have the same directional effect: longer payment lead times to suppliers should lead to more inventories (in essence, suppliers finance buyer's inventories). From the publicly available data it appears impossible to separate out sourcing lead times from payment lead times. However, we do believe that these two variables are strongly correlated, which suffices for our exploratory analysis. Future research might be able to suggest a better proxy for lead times.

Estimation of *demand uncertainty* is typically of first-order importance in any finance-related research, since variability is priced in financial markets via derivatives and there is a "market" for

uncertainty. Various techniques and estimations have been proposed for very frequent financial trading data that often do not exhibit any deterministic trends. In non-financial markets, however, the situation is quite different—the major concern is typically to predict sales, profits and expenses. Therefore, it is important to forecast trends in data and deal with first-order rather than second-order estimated moments. Variance estimation techniques are much less sophisticated in both the theory and practice of operations management, and this simplicity is driven both by the scarcity of data and by the absence of markets for uncertainty analogous to derivatives markets. Nevertheless, the impact of demand uncertainty on inventories is a focal point of many classical inventory papers and we thus ensure that demand uncertainty is captured adequately.

We assume that our data is decomposable in an additive way into trend, seasonal and noise components. Additive techniques are by far the most common and are used by the US Census (e.g., X-12ARIMA for monthly Census data) as well as by other statistical agencies. Additive decomposition implies that variance of sales is determined by variance of noise only. To estimate noise, we run individual regressions with a polynomial capturing trend and with seasonal (quarterly) dummies and we take residuals as demand noise. We tried several other specifications (e.g., higher-degree polynomials) and finally conducted the estimation using the first-degree polynomial (linear trend) with seasonal dummies, because this specification minimizes in-sample average mean squared forecasting error (Greene 1997). We conduct this estimation for all 772 firms in our panel and thereafter we estimate the variance of residuals, again using a four-quarter moving window as follows:

$$R_t = X_t - a_1 t - a_2 - b_1 q_1 - b_2 q_2 - b_3 q_3,$$

$$\text{Sigma}_t(X) = \text{Sigma}_t(R) = \sqrt{\frac{\sum_{i=0}^3 \left(R_{t-i} - \sum_{j=0}^3 R_{t-j} / 4 \right)^2}{4}}, t = 3, \dots, 44. \quad (2)$$

We note that, to our knowledge, there is no commonly accepted way to estimate demand uncertainty using past data in the operations management literature. We provide three arguments to motivate employing four past quarters of sales data: (i) similar approaches appear in popular textbooks and cases, indicating that companies employ similar techniques in practice (see, for example, Cachon and Terwiesch 2005 and the LLBean case by Schleifer 1992); (ii) it is advantageous to use recent data in an uncertainty estimation; and (iii) longer rolling-back estimation horizons create more missing or quasi-missing data points in estimations. It should be noted that for the sake of removing seasonality the exact number of past quarters for model-based de-seasoned and de-trended data is of secondary importance, since the model fitted over the total time horizon ensures that the residuals are seasonality-corrected.

We acknowledge that other proxies for demand uncertainty can prove to be less biased or more efficient when analyzed using other criteria (AIC, BIC, out-of-sample MSE; see Greene 1997): for example, when performing detailed, firm-specific demand forecasting one could use additional explanatory variables or use parametric truncated distributions to impute demand forecast from sales data. However, for the purposes of our exploratory cross-sectional study that involves data aggregation, we believe that using the proxy described above endows us with a sense of firm-level demand uncertainty. Moreover, we repeated our analysis with several other demand uncertainty proxies. We measured both uncertainty in sales expressed in inventory prices (so-called top-line uncertainty following the accounting terminology, see Stickney and Weil 1999—sales are recorded at the top of the income statement) and uncertainty in earnings (EBITDA or operating income, the so-called “bottom-line uncertainty,” since it is recorded at the bottom of the income statement). Furthermore, we used another technique to estimate the variance of sales and earnings using raw uncertainty (rather than de-trended and de-seasonalized uncertainty). We found the results to be very robust as to both the model complexity and model specification. Therefore, we report our results only for the demand uncertainty proxy defined above.

As a proxy for *company size*, we use the fixed assets variable FixedAssets_{it} . We repeated our analysis for sales as a proxy for company size (not reported), and the results were robust. As previously mentioned, we use seasonal dummies q_t to control for season-specific effects and yearly dummies (from 1 to 11) to control for the time trend. Two more control variables are introduced to capture the second-order relations between sales and inventory in addition to the first-order relation captured by the absolute level of demand. Most classical inventory models rely on the assumption that demand is independent and identically distributed across time periods. In practice, however, there are typically well-known deterministic trends in demand that could result from the expansion of the company or of the industry or of the whole economy (as opposed to the stochastic component captured by the proxy for demand uncertainty). The variable SalesGrowth_{it} measures a percentage growth in COGS_{it} from quarter to quarter, which is essentially the speed of change in sales. The variable $\text{PositiveSalesSurprise}_{it}$ takes the value of 1 if the realized demand is higher than forecasted and takes the value of 0 otherwise. This last dummy variable is included to account for a lower-than-expected inventory level in case the demand exceeds the forecast (which is somewhat similar to the approach of Gaur et al. 2005). It should be noted the two proxies differ from each other: the first one captures the total growth, whereas the second one captures the potential non-symmetry of inventory reaction to sales shocks. Summary statistics for all variables are found in Table 4. Table 3 contains correlations among independent variables.

6. Model specification

In line with Gaur et al (2005), we use two multiplicative regression models (“multiplicative” in that independent variables are in multiplicative form before the log transformation) to capture inventory dynamics and to test the proposed hypotheses: one model for absolute inventory levels and the other one for relative inventory levels. The model for absolute inventory levels is:

$$\begin{aligned} \log(\text{Inv}_{its}) = & a_s + b_s^1 \log(\text{COGS}_{its}) + b_s^2 \log(\text{GrossMargin}_{its}) + b_s^3 \log(\text{LeadTime}_{its}) \\ & + b_s^4 \log(\text{SigmaSales}_{its}) + b_s^5 \log(\text{TBillRate}_t) + b_s^7 \log(\text{PositiveSalesSurprise}_{its}) \\ & + b_s^8 \log(\text{SalesGrowth}_{its}) + c_s^1 q_1 + c_s^2 q_2 + c_s^3 q_3 + \text{year} + \varepsilon_{its}, \end{aligned} \quad (3)$$

$t = 1, \dots, 44$ - time index,

$i = 1, \dots, 722$ - company index,

$s = 1, \dots, 8$ - segment index.

The disturbance term ε_{its} additively includes firm-specific and time-specific effects along with random noise to ensure that unobserved variables in the panel such as management quality, marketing decisions, etc. are controlled for so as to avoid biased estimators (Greene 1997). We model firm-specific and time-specific effects using both random and fixed-effect models. When reporting final results we use fixed-effects coefficients estimation because the Hausman test (Greene 1997) rejects the hypothesis that random-effects estimators are efficient and because the large size of our sample as well as our attention to the eight specific segments indicates the need for fixed effects in the research design. Finally, we control for possible autocorrelation in the data by conducting estimation with the AR(1) disturbance term. We do not report autocorrelation coefficients that are on average 0.1 or smaller, and we also checked that the regression results are robust with respect to autocorrelation of residuals specification. We use log-log model specification because (i) it helps deal with heteroskedasticity; (ii) the distribution of inventory data is not normal but rather is skewed to the right; and (iii) such specification has shown the best results in previous studies (Gaur et al. 2005).

The model for relative inventory levels is defined analogously, the only exception being that we control for firm size by excluding COGS and by including FixedAssets in the regression to avoid possible collinearity issues. We run a pooled regression first and then run the same model within segments to see how the variables are related within specific segments and whether the hypotheses are supported therein. Our relatively short and wide panel (with large cross-sectional dimension and short time-series dimension) has more similarities to cross-sectional analysis than to time-series analysis. Therefore, we did not analyze such aspects as the stationarity of data or vector autoregression properties

(which is in line with Greene 1997). We did include an explicit time trend in the model even though inventory management theory does not indicate a need to do so: time trends should affect sales and sales forecasts in the first place, and companies should react by adjusting their operating policies. Hence, one could assume that the time trend is captured by sales dynamics and that inventory is managed rationally. To check the robustness of our results, we conducted analysis with and without the time trend. We find that our assumption about rationally managed inventory cannot be rejected: estimated coefficients do not change significantly compared to the no-time-trend approach, the only exception being inventory holding cost coefficients. The time trend adds little explanatory power to the model, as indicated by the levels of F-statistics. We use analysis with the time trend when reporting final results.

To ensure the robustness of our results we also used alternative econometric specifications. We employed linear regression instead of multiplicative regression and added several interaction terms to see if they could improve the overall fit expressed in the adjusted goodness-of-fit measure. Linear models were rejected, since they performed significantly worse, and interaction terms between COGS, margins and lead times did not incrementally increase the explanatory power of the model, while the directions of main effects did not change. We do not include interaction terms in our final analysis, because we could not find theoretical or empirical justifications for them in either operations management or economics literature. To test Hypothesis 1 we used a polynomial model (with respect to inventory as a function of COGS) with up to 5th-degree polynomials and down to 1/5th-degree polynomials and found that the fit of such a model was inferior to multiplicative models. We did try to introduce additional and nonlinear terms and found that those terms did not add explanatory power while the directions of main results did not change.

7. Results

We use panel data modules “xtreg” and “xtregar” in STATA to perform our analysis. The STATA package, similar to the SAS package, provides embedded tools to analyze panel data and to estimate fixed and random effects regressions with possible autoregression structure.

Table 5 provides results for the model with absolute inventory level as a dependent variable (the adjusted goodness of fit is 85%), and Table 6 provides results for the model with relative inventory level as a dependent variable (the adjusted goodness of fit is 12%). We report results only with $TBillRate_t$ used as a proxy for inventory holding costs, because the alternative measure, $CostCapital_{its}$, yields qualitatively similar results. Although it appears that the absolute inventory model results in a better adjusted goodness of fit, the two non-nested models should not be directly compared: after the log-

transform, the domain of the absolute inventory is much smaller than that of the relative inventory and hence the difference in explanatory power. Finally, Table 7 provides results split by industry segment for the model with relative inventory (the adjusted goodness of fit ranges from 4% to 29%). We utilize the relative inventory model in Table 7 because comparisons across segments are rather meaningless for absolute inventory levels.

The data are *consistent* with Hypothesis 1: elasticity of inventory to COGS is .68 such that inventory is a sub-linear function of sales, implying a concave relationship and hence hinting at economies of scale in inventory management.

The data are *consistent* with Hypothesis 2: relative inventory levels are negatively associated with company size as measured by fixed assets, indicating the presence of economies of scale in inventory management. We find that a 1% increase in the fixed assets of a company is, on average, associated with a .12% decrease in relative inventory. From Table 7, the (absolute) elasticity of relative inventory to company size is highest for retail and machinery. The food segment shows the lowest degree of economies of scale in inventory management, and from the Table 7 we observe that Hypothesis 2 is supported across all segments.

The data at the aggregate level is *consistent* with Hypothesis 3: both absolute and relative inventory levels increase with demand uncertainty. We find that a 1% increase in sales uncertainty increases absolute inventory on average by .02%. When we run segment-specific regressions, it turns out that the relationship is not statistically significant for the chemicals segment and, moreover, for the retail segment the relationship is reversed. Given the difficulty in estimating uncertainty in actual demand and the variety of methods used in practice, it is not surprising that the expected result is not supported in all industry segments. The statistically significant negative association between uncertainty and inventory in the retail segment is truly interesting and requires further investigation.

The data at the aggregate level is *consistent* with Hypothesis 4: both absolute and relative inventory levels are increasing in lead times. Specifically, a 1% increase in lead time leads to a .11% increase in absolute inventory and a .13% increase in relative inventory. This relationship is statistically significant for all industry segments.

The data at the aggregate level is *consistent* with Hypothesis 5: both absolute and relative inventory are increasing in product margins, which is consistent with the finding of Gaur et al. (2005) that inventory turns and gross margins are negatively related in retailing. The elasticity of absolute and relative inventory levels to gross margin is about .03. The relationship between margins and inventory is also not significant for the electronics segment, and it is reversed for the machinery segment.

The data are *not consistent* with Hypothesis 6: the relationship between holding costs and inventory appears to be positive rather than negative, and it is not significant in the pooled regressions. The relationship turns out to be of the right sign for the oil and gas, wholesale and food segments, but the opposite is true for all other segments. We check for the robustness of results for Hypothesis 6 by using both the risk-free interest rate $TBillRate_t$ and the firm-specific proxy $CostCapital_{its}$. The latter gives very similar results (and therefore we do not report additional tables) with respect to both signs and significance. We also observe that inventory holding cost proxies are the only ones among dependent variables that are sensitive to linear and quadratic time trends, implying that they are impacted by the cyclical nature of interest rate dynamics during the period from 1992 to 2002. If the time trend is excluded, the relationship between inventory holding costs and inventory levels becomes negative (as predicted by Hypothesis 6) and significant in most tests.

We note that, for the absolute inventory model, COGS, demand uncertainty, margins, lead times and the time trend explain approximately 70%, 7%, 5%, 2% and 1% of variance in inventories, respectively. Seasonality and other dummy variables explain a very small proportion of variances in inventories. In addition to the results of the tests of the hypotheses, we also make several observations on the significance of various dummy variables included in the analysis. Variable $SalesGrowth$ (the percentage increase in sales from the previous quarter) exhibits a statistically significant negative association with absolute and relative inventory levels which also persists in all eight industry segments. Hence, as demand increases, companies do not react immediately by increasing inventory. Variable $PositiveSalesSurprise$ (valued at 1 if demand is higher than forecasted) exhibits statistically significant positive associations with absolute and relative inventory levels which also persists in seven out of eight industry segments (except machinery). These two results do not contradict but rather complement each other—while inventories, on average, decrease with higher overall sales growth, inventories also, on average, increase when sales increase unpredictably. This observation supports the idea of both nonlinearity and non-symmetry at the level of association between sales activities and inventory levels: inventory dynamics lag somewhat behind sales dynamics in terms of growth rates, which can be partially explained by managers' adaptive expectations and predictions that are typically based on historical data. For example, if changes in sales speed up, adaptive expectations and adaptive demand forecasting will underestimate sales and result in lower than optimal planned inventory levels. Moreover, inventory levels are adjusted more rapidly when there is a mismatch in the demand forecast. This effect makes sense if managers become more optimistic when demand realization is higher than forecasted. It is interesting to note that such empirical observations are not grounded in existing

inventory models, and they provide directions for further theoretical inventory modeling that would attempt to link the nonlinear and non-stationary nature of most of the demand processes with the behavioral aspects of inventory management decisions. Finally, seasonal and yearly dummies are statistically significant in most cases, demonstrating the importance of accounting for business seasonality and time-trends, something that is rarely incorporated into inventory models. It is interesting to notice that the time trend is positive in our models after controlling for other effects. This does not necessarily contradict the results in Chen et al. (2005a, 2005b) and Rajagopalan and Malhotra (2001), who show that US public inventories have been declining for the past 20 years. The reason is that we do not directly study the behavior of inventories over time (inventory, indeed, has a negative and significant time trend in the absence of other variables). Instead, we capture both cross-sectional and time-series inventory fluctuations and show that, after controlling for significant cross-sectional effects (as predicted by the inventory models), the unpredictable portion (partial residuals without accounting for the time trend) of both absolute and relative inventory is positively associated with the time trend. Namely, the positive time trend in our studies (despite the fact that inventories do decrease over time in our sample) is an indication that other explanatory variables (COGS, margins, lead times) also change over time and explain a significant portion of inventory variance, whereas the time trend by itself can explain only around 1% of inventory variance. We thus believe that there is a need to capture the time trend directly, since other explanatory variables are also time-dependent and change along with inventory over time in our panel.

8. Summary

A rich variety of inventory models have been created in operations management based on the cost-minimizing approach for price-taking monopolistic companies. In this exploratory study we have identified the need for conducting empirical research to test predictions from classical inventory models to understand whether predictions from these models help explain aggregate inventories of entire companies. We tested for several predictions using a large quarterly panel data set from the COMPUSTAT financial database of US public companies operating during the 1992-2002 period. In our sample we included eight segments of the US economy: oil and gas, retail, wholesale, electronics, hardware, chemicals, food, and machinery. Nearly 45% of S&P500 companies are present in our sample, making a total of 722 companies that have continuously been in operation and have had nonzero financial results. We tested proposed associations for pooled data as well as across segments.

We find that multiplicative models best capture inventory dynamics. Our model explains 85% of the dynamics for absolute inventory level. We believe it is fair to say that classical inventory models continue to explain inventory behavior well, even at the company level. We proposed several proxies for the independent variables. Average days of accounts payable outstanding are used to measure lead times, and we find that companies do buffer inventory against longer lead times. The treasury bill rate and the weighted-average cost of capital are used to measure the opportunity cost of carrying inventory, but our findings regarding holding costs are inconclusive and are affected by the inclusion of the time trend. It is possible that findings in Lai (2006) indicating the need to find new ways of calculating the cost of capital for operations models will be helpful in the future research. We obtain consistent results that companies do buffer inventory against demand uncertainty. We also find that higher gross margins are associated with higher inventory and that large companies enjoy economies of scale and hence stock less inventory in relative terms.

It is reassuring to see evidence that predictions from classical inventory models appear to work at the aggregate (company) level. Hence, we believe that these models should be taught to business students and executives. Even though these models do not account for competition, business cycles, and industry dynamics and do not endogenize many other decisions (e.g., pricing), they still appear to be useful in guiding our intuition with respect to inventory behavior. We do not, however, claim that there is a causal relation between operational variables such as lead time, demand uncertainty, margin, and inventory. The relationships we tested are based only on statistical associations that are stable over time. One could try to go further and notice that in the case of operational models there is a clear model-driven causality: “inputs” such as lead time, margins, uncertainty and sales might affect “outputs” such as profits and inventories. That is, inventory may always be a function of both exogenous variables (mean demand, uncertainty) and a firm’s decisions (when and how much to order given these variables). In this setting, some of the statistical statements tested may also impose causality on the operational relations. For example, longer lead times, higher demand uncertainty and higher margins may cause both relative and absolute inventory to increase, while the statement that “larger companies are leaner” does not impose causality. These distinctions could be addressed in follow-up work.

Our study comes with several potential limitations. We use aggregated COMPUSTAT financial data that can cause both space (product variety and different units) and time aggregation biases, and therefore we focus on the monotonic properties of inventories. Other explanatory variables (such as product variety, product life cycle lengths, quantity discounts, and forecasts of supply chain disruptions) are not included in our models. More research is needed to find adequate proxies for these variables.

Limitations of this kind are quite standard in the empirical literature working with inventories (see Rajagopalan and Malhotra 2001 and Gaur et al. 2005). Furthermore, we are using imperfect proxies for some of the independent variables such as lead times and inventory holding costs. If better data become available, it might be prudent to reevaluate our findings.

Future studies could look more closely at various segments of the economy and attempt to explain why some of our hypotheses do not hold for certain segments by using better proxies, possibly obtained through a combination of surveys and secondary data. A lot of research has been done to explain the behavior of inventories in the automotive industry (see a series of papers by Lieberman et al.) and in retailing (see a series of papers by Gaur et al.) but not in other segments that we capture in our study and not with the relevant set of all potentially economically important variables.

References

- Amihud Y. and H. Mendelson. 1989. Inventory Behavior and Market Power: An Empirical Investigation, *International Journal of Industrial Organization*, Vol. 7, 269-280.
- Arrow K., T. Harris and J. Marschak. 1951. Optimal Inventory Policy, *Econometrica*, Vol. 19, No. 3, 250-272.
- Bayus, B.L. and W.P. Putsis. 1999. Product Proliferations: an Empirical Analysis of Product Line Determinants and Market Outcomes. *Marketing Science*, Vol. 18, No. 2, 137-153.
- Blinder A. 1986. Can the Production Smoothing Model of Inventory Behavior Be Saved? *The Quarterly Journal of Economics*, Vol. 101, No. 2, 431-454.
- Blinder A., M. Lovell and L. Summers. 1981. Retail Inventory Behavior and Business Fluctuations. *Brookings Papers on Economic Activity*, Vol. 1981, No. 2, 443-520.
- Blinder A. and L. Maccini. 1991. Taking Stock: A Critical Assessment of Recent Research on Inventories, *Journal of Economic Perspectives*, Vol. 5, 73-96.
- Botosan C. and M. Plumbee. 2005. Assessing Alternative Proxies for the Expected Risk Premium. *The Accounting Review*, Vol. 80, No. 1, 21-53.
- Cachon, G., G. Gao and L. Hitt. 2005. Product Variety, Inventory Management and Firm Performance. Working paper, University of Pennsylvania.
- Cachon G., T. Randall and G. Schmidt. 2004. In Search of the Bullwhip Effect. Presentation at the Wharton School, University of Pennsylvania.
- Cachon G. and C. Terwiesch. 2005. *Matching Supply with Demand: An Introduction to Operations Management*. Second edition, The McGraw-Hill Companies.
- Caplin A. 1985. The Variability of Aggregate Demand with (S,s) Inventory Policies. *Econometrica*, Vol. 53, No. 6, 1395-1410.
- Chen H, M. Frank, and O. Wu. 2005a. What Actually Happened to the Inventories of American Companies Between 1981 and 2000? *Management Science*, Vol.51, 1015-1031.

- Chen H., M. Frank and O. Wu. 2005b. US Retail and Wholesale Inventory Performance from 1981 to 2003. Working paper, University of British Columbia.
- Cohen M., T. Ho, Z. Ren and C. Terwiesch. 2003. Measuring Imputed Costs in the Semiconductor Equipment Supply Chain, *Management Science*, Vol. 49, 1653-1670.
- COMPUSTAT North America User's Guide, 2003.
- Eppen, G.D. 1979. Effects of centralization on expected costs in a multi-location newsboy problem. *Management Science*, Vol. 25, No. 5, 498-501.
- Fama E. and K. French. 1999. The Corporate Cost of Capital and the Return on Corporate Investment. *The Journal of Finance*, Vol. 54, 1939-1967.
- Fisher M. and C. Ittner. 1999. The Impact of Product Variety on Automobile Assembly Operations: Empirical Evidence and Simulation Analysis. *Management Science*, Vol. 45, 771-786.
- Gaur V., M. Fisher and A. Raman. 2005. An Econometric Analysis of Inventory Turnover Performance in Retail Services. *Management Science*, Vol. 51, 181-194.
- Gaur V., M. Fisher and A. Raman. 1999. What Explains Superior Retail Performance? Working paper, New York University.
- Greene W. 1997. *Econometric Analysis*. Third edition. Prentice-Hall, NJ.
- Hendricks K. and V. Singhal. 2005. Association Between Supply Chain Glitches and Operating Performance. *Management Science*, Vol. 51, No. 5, 695-711.
- Irvine, F.O. Jr. 1981. Retail Inventory Investment and the Cost of Capital. *The American Economic Review*, Vol. 71, 633-68.
- Johnson L.A. and D.C. Montgomery. 1974. *Operations Research in Production, Planning, Scheduling, and Inventory Control*. John Wiley and Sons, New York, NY.
- Kahn J. 1987. Inventories and the Volatility of Production. *The American Economic Review*, Vol. 77, 667-679.
- Kahn J. 1992. Why Is Production More Volatile than Sales? Theory and Evidence on the Stockout-Avoidance Motive for Inventory Holding. *The Quarterly Journal of Economics*, Vol. 107, No.2, 481-510.
- Kahn J. and M. Bils. 2000. What Inventory Behavior Tells Us about Business Cycles. *The American Economic Review*, Vol. 90, No. 3, 458-481.
- Kekre, S. and K. Srinivasan. 1990. Broader Product Lines: A Necessity to Achieve Success? *Management Science*, Vol. 36, No. 10, 1216-1231.
- Lai, R. 2006. Inventory and the Stock Market. Working paper, Harvard University.
- Lieberman C. 1980. Inventory Demand and the Cost of Capital Effects. *The Review of Economics and Statistics*, Vol. 62, No. 3, 348-356.
- Lieberman M.B. and S. Asaba. 1997. Inventory Reduction and Productivity Growth: A Comparison of the Japanese and U.S. Automotive Sectors. *Managerial and Decision Economics*, special issue on Japanese technology management. Vol. 18, 73-85.
- Lieberman M.B. and L. Demeester. 1999. Inventory Reduction and Productivity Growth: Linkages in the Japanese Automotive Industry. *Management Science*, Vol. 45, 466-476.

- Lieberman M.B., S. Helper and L. Demeester. 1999. The Empirical Determinants of Inventory Levels in High-Volume Manufacturing. *Production and Operations Management*, Vol. 8, 44-55.
- Mosser P. 1991. Trade Inventories and (S,s). *The Quarterly Journal of Economics*, Vol. 106, No.4, 1267-1286.
- Olivares M., C. Terwiesch and L. Cassorla. 2004. Reserving Operating Room Capacity: A Structural Estimation using the Newsvendor Model. Working paper, University of Pennsylvania.
- Porteus E.L. 2002. *Foundations of Stochastic Inventory Theory*. Stanford University Press.
- Rajagopalan S. and A. Malhotra. 2001. Have US Manufacturing Inventories Really Decreased? An Empirical Study. *Manufacturing and Service Operations Management*, Vol. 3, 14-24.
- Ramey V. and K. West. 1999. Inventories. Chapter 13 in the *Handbook of Macroeconomics*, Vol. 1, Elsevier Science.
- Randall T., S. Netessine and N. Rudi. 2006. An Empirical Examination of the Decision to Invest in Fulfillment Capabilities: A Study of Internet Retailers. *Management Science*, Vol. 52, No. 4, 567-580.
- Randall T. and K. Ulrich. 2001. Product Variety, Supply Chain Structure, and Firm Performance: Analysis of the US Bicycle Industry. *Management Science*, Vol. 47, 1588-1604.
- Schleifer A., Jr. 1992. L.L. Bean, Inc.: Item Forecasting and Inventory Management. Case Study 9-893-003, Harvard Business School Publishing.
- Silver E., D. Pyke and R. Paterson. 1998. *Inventory Management and Production Planning and Scheduling*. Third edition, John Wiley and Sons.
- Stickney C.P. and R.L. Weil. 1999. *Financial Accounting: An Introduction to Concepts, Methods, and Uses*, Ninth edition. Dryden Press, Orlando, FL.
- Wagner H. 2002. And Then There Were None. *Operations Research*, Vol. 50, No. 1, 217-226.
- Zipkin P.H. 2000. *Foundations of Inventory Management*. McGraw-Hill.

Table 1. Definitions of explanatory variables

COGS	Cost of goods sold, as reported in quarterly income statements
FixedAssets	Long-lived property owned by a firm, as reported in quarterly income statements
GrossMargin	(Sales-COGS)/Sales
LeadTime	365/(4*COGS/AP), where AP is accounts payable
SigmaSales	$\sqrt{\frac{\sum_{i=0}^3 \left(\widehat{Sales}_{t-i} - \sum_{i=0}^3 \widehat{Sales}_{t-i} / 4 \right)^2}{4}}$ where $\widehat{Sales}_t = Sales_t - a_1 t - a_2 - b_1 q_1 - b_2 q_2 - b_3 q_3$
SalesGrowth _{t+1}	(COGS _{t+1} -COGS _t)/COGS _t
TBillRate	3-month T-bill rate
PositiveSalesSurprise	1 is Sales exceeds forecast obtained using extrapolation of Sales: $\widehat{Sales}_t = \hat{a}_0 Sales_{t-1} - \hat{a}_1 t - \hat{a}_2 - \hat{b}_1 q_1 - \hat{b}_2 q_2 - \hat{b}_3 q_3$
q ₁ ,q ₂ ,q ₃	Quarter dummies
Year	1992:Year=1; ... 2002:Year = 11

Table 2. Sample description (quarterly in \$M, 1992-2002 period)

Segment	Segment name	SIC codes	# of companies	Mean Inventory	Mean COGS	Mean Inv/COGS	St.Dev./Mean, Inv/COGS	Median, Inv/COGS
1	oil and gas	1311, 1381, 2911	86	559	1343	.42	5.48	.23
2	electronics	3630, 3651, 3663, 3672, 3674	190	168	173	1.22	2.51	.89
3	wholesale	5045, 5047, 5122, 5140, 5172	61	254	502	.39	1.23	.29
4	retail	5311, 5331, 5411, 5412, 5940, 5944, 5945	95	968	1057	1.02	1.02	.67
5	machinery	3523, 3531, 3532, 3537	22	578	573	1.11	.82	.91
6	computer hardware	3571, 3571, 3575, 3576, 3577	117	141	181	1.25	4.38	.84
7	food and beverages	2000, 2011, 2030, 2080, 2082	35	736	872	.64	1.11	.51
8	chemicals	2800, 2820, 2834, 2844	116	388	314	1.40	1.70	1.01
Total:			722	396	527	1.03	2.94	.74

Table 3. Correlation table (all variables in Log, all correlations are significant at the .01 level)

	COGS	Fixed Assets	Gross Margin	Lead Time	Sales Growth	Sigma Sales	TBill rate	Cost Capital
COGS								
FixedAssets	.46							
GrossMargin	-.30	.11						
LeadTime	.09	.31	.17					
SalesGrowth	-.01	.07	.00	.05				
SigmaSales	.62	.12	-.20	.09	.01			
TBillRate	-.02	.03	.00	-.01	-.07	-.03		
CostCapital	.01	.02	.00	-.00	-.06	-.00	.66	
PositiveSales Surprise	.61	.40	-.25	.12	.72	.70	-.03	.00

Table 4: Sample summary statistics

	Mean	25% percentile	Median	75% percentile
FixedAssets	309.5	4.1	31.6	390
GrossMargin	0.16	0.15	0.31	0.48
LeadTime	101.9	34.1	54.9	105.3
SigmaSales	180.1	3.1	43.8	92.3
SalesGrowth	0.54	-0.18	0.02	0.21
TBillRate	1.12	0.79	1.06	1.24
PositiveSalesSurprise	0.44	0	0	1

Table 5. Absolute inventory pooled regression

	Log(Inv)
Log(COGS)	.68*** (.01)
Log(GrossMargin)	.03*** (.01)
Log(LeadTime)	.11*** (.00)
Log(SigmaSales)	.02*** (.00)
SalesGrowth	-.001*** (.00)
Log(TBillRate)	.03 (.02)
PositiveSalesSurprise	.15*** (.02)
q1	.10*** (.01)
q2	.07*** (.01)
q3	.10*** (.01)
Year	.02*** (.00)
Constant	.06*** (.03)
Adjusted goodness of fit	85%

Table 6. Relative inventory pooled regression

	Log(InvCOGSratio)
Log(FixedAssets)	-.12*** (.03)
Log(GrossMargin)	.03*** (.01)
Log(LeadTime)	.13*** (.002)
Log(SigmaSales)	.01*** (.001)
SalesGrowth	-.001*** (.0001)
Log(TBillRate)	.02 (.02)
PositiveSalesSurprise	.17*** (.02)
q1	.16*** (.02)
q2	.08*** (.02)
q3	.13*** (.01)
Year	.03*** (.001)
Constant	.19*** (.02)
Adjusted goodness of fit	12%

Note: ***, ** and * denote statistical significance at the 1%, 5% and 10% level. Standard errors are reported in brackets.

Table 7. Relative inventory regression by segment

Log(InvCOGSratio)	oil and gas	electronics	wholesale	retail	machinery	hardware	food	chemicals
Log(FixedAssets)	-.17*** (.04)	-.13*** (.02)	-.14*** (.03)	-.19*** (.02)	-.19*** (.03)	-.11*** (.02)	-.10*** (.02)	-.12*** (.03)
Log(GrossMargin)	.05*** (.02)	-.02 (.01)	.06*** (.02)	.09*** (.01)	-.11*** (.02)	.06*** (.03)	.04** (.02)	.03*** (.01)
Log(LeadTime)	.07*** (.02)	.08*** (.02)	.11*** (.03)	.05*** (.01)	.05*** (.01)	.03*** (.01)	.06*** (.02)	.12*** (.01)
Log(SigmaSales)	.02*** (.01)	.02*** (.01)	.02*** (.01)	-.001*** (.001)	.005** (.001)	.02** (.01)	.04*** (.02)	-.02 (.03)
SalesGrowth	-.001* (.0004)	-.02** (.003)	-.001*** (.0001)	-.002*** (.0006)	-.001*** (.0001)	-.002*** (.0004)	-.003*** (.0008)	-.001* (.0006)
Log(TBillRate)	-.11*** (.03)	.05*** (.02)	-.07** (.04)	.05*** (.01)	.08** (.02)	.11*** (.02)	-.08*** (.02)	.05* (.03)
PositiveSalesSurprise	.13*** (.05)	.13*** (.05)	.12*** (.05)	.03*** (.01)	.03 (.05)	.15*** (.04)	.15** (.05)	.07*** (.03)
q1	.12*** (.03)	.06*** (.01)	.14*** (.03)	.12*** (.03)	.13*** (.03)	.14*** (.04)	.09*** (.02)	.01 (.03)
q2	.12*** (.03)	.06*** (.01)	.09*** (.03)	.12*** (.01)	.13*** (.04)	.08*** (.02)	.02 (.02)	.06*** (.01)
q3	.18*** (.03)	.07*** (.02)	.11*** (.02)	.16*** (.04)	.06** (.02)	.08*** (.01)	.03*** (.01)	.08*** (.02)
Year	.02*** (.001)	.02*** (.004)	-.02*** (.01)	.07*** (.006)	.07*** (.001)	-.001 (.002)	.03*** (.002)	.05*** (.004)
Constant	.47*** (.12)	-.31* (.14)	-1.34*** (.23)	1.48*** (.27)	1.69*** (.37)	-.67*** (.39)	1.15*** (.24)	.48 (.12)
Adjusted R²	4%	17%	5%	19%	29%	29%	9%	9%

Note: ***, ** and * denote statistical significance at the 1%, 5% and 10% level. Standard errors are reported in brackets.