POMS PRODUCTION AND OPERATIONS MANAGEMENT

PRODUCTION AND OPERATIONS MANAGEMENT POMS
Vol. 28, No. 9, September 2019, pp. 2338–2364
ISSN 1059-1478 | EISSN 1937-5956 | 19 | 2809 | 2338
DOI 10.1111/poms.13055

Examining the Link between Retailer Inventory Leanness and Operational Efficiency: Moderating Roles of Firm Size and Demand Uncertainty

Howard Hao-Chun Chuang
College of Commerce, National Chengchi University, Taipei 11605, Taiwan, chuang@nccu.edu.tw

Rogelio Oliva*, Gregory R. Heim
Mays Business School, Texas A&M University, College Station, Texas 77843-4217, USA, roliva@tamu.edu, gheim@mays.tamu.edu

Retail inventories have been consistently dropping, relative to sales, since the 1990s. Whether these lean inventory developments translate to better retailer operational performance is still an open question. We empirically examine associations between inventory leanness and operational efficiency for a sample of public US retailers from 2000 to 2013. Via a stochastic frontier analysis that accounts for retailer heterogeneity and time parameters, we find support for the hypothesis that operational efficiency has an inverted U-shape relationship with inventory leanness, suggesting an optimal inventory leanness level beyond which retailer operational efficiency degrades. This relationship, however, is heavily moderated by firm size and demand uncertainty. The former reflects a retailer's abilities to exploit economies of scale and scope, whereas the latter reflects the unpredictability in a firm's operating environment. Our evidence suggests that when increasing inventory leanness, small retailers exhibit efficiency degradation, whereas larger retailers are likely to exhibit efficiency improvement, with diminishing returns. We also find that under high demand uncertainty, being less lean is associated with higher operational efficiency, regardless of firm size. The findings show that depending on firm size and demand uncertainty, retail managers should take special care when pursuing inventory leanness. As part of post hoc robustness tests, we assess how different retail categories vary in their operational efficiency scores and conduct interviews with retail executives who further ground our econometric investigation and point to more nuanced moderators for future studies. We conclude by discussing the implications of our industry model estimation for managers and researchers.

Key words: efficiency; inventory; lean; retailing; stochastic frontier analysis
History: Received: January 2016; Accepted: April 2019 by Edward G. Anderson, after 3 revisions.

1. Introduction

Lean operations principles changed the way many manufacturers operate (Lieberman and Demeester 1999). The perceived success of lean principles in manufacturing has carried over into retailing and other service industries (Corbett 2007). Instead of perceiving inventory as an asset, managers following lean principles tend to view inventory as a waste to minimize. Excessive inventories metaphorically and explicitly represent incoordination and inefficiency of production and distribution systems (Chen 2017). Yet, while considerable efforts are made to achieve zero-excess-inventory objectives in manufacturing, the lean philosophy also is known to have certain limits to its execution (de Haan and Yamamoto 1999).

Inventory decisions in retailing are one of the most significant managerial levers not only in magnitude—inventory represents, on average, 43% of current assets of US retailers in our sample—but also in the consequential effects it has for operational performance. In retail environments, managers may deliberately carry and display abundant inventory, not only to meet marketing requirements (e.g., display ample quantities), but also to achieve operational and financial aims (e.g., reduce lost sales opportunities). Inventory in retailing adds value since consumer demand partially depends on inventory levels and product availability (Balakrishnan et al. 2004, Ton and Raman 2010) and maintaining enough inventories is critical to operational and financial success of retailers (Fisher and Raman 2010). Furthermore, retailers may pursue fast growth strategy based on aggressive commitment of resources (e.g., inventory) to open new stores (Ansoff 1957, Gaur et al. 1999), or when lead times are long, a retailer may buy large inventories to induce herself to work harder to sell products (Kraiselburd et al. 2011).

Even though retail managers have such incentives to maintain high inventories, holding too much excess inventory may not be advantageous. Lean ideas thus also influence retailer inventory decisions. Abernathy et al. (2000) propose the concept of lean retailing and encourage retailers to better manage inventories by employing lean principles. Ideally, lean inventory control enables high service levels, yet reduces stock on hand. Driven by macro-economic forces (e.g., business cycles) and technological factors (e.g., inventory replenishment information systems), today retailers are as lean as ever. Chen et al. (2007) find average inventory days of US retailers decreased since the late 1990s. Solomon (2013) reports US retailer inventory-to-sales ratios declined since 2000 and reached near all-time lows in the post-recession era after 2009. Johnston (2014) finds a strong downward trend (1982–2012) of inventory-to-assets. Recently, big-box retailer Target even deliberately reduced the number of SKUs in stores (Ziobro 2016a).

This evidence of inventory reduction within the US retailing industry is not too surprising, since retail practitioners, motivated by lean retailing visions, exert efforts to eliminate waste and inefficiency while trying not to compromise sales (Cachon 2001). The assumption is that inventory leanness does no harm to retailer performance given the advent of better decision support tools and shorter lead times (Abernathy et al. 1999, Solomon 2013). Yet, an opposite view supports the value of piling up store inventory to promote sales, that is, "Stack them high, let'em fly" (Balakrishnan et al. 2004, Krommyda et al. 2015, Stavrulaki 2011). Having inventory in-store may also help brick-and-mortar retailers compete with online retailers as higher availability may stimulate store visits when customers check availability online. Publicly held retailers are nevertheless motivated to be leaner and to maximize inventory productivity, since retail inventory is a "closely watched statistic by retailers as well as their investors, lenders, and suppliers" (Gaur et al. 2014, p. 55). Hence, whether lean inventory helps improve retailer operational performance is still an open question.

The impact of inventory leanness can be measured via a firm's deviation from the production frontier constituted of best-practice firms in an industry (Lieberman and Dhawan 2005). Vastag (2000) claims firms can move toward a production frontier, improving operational efficiency, by building better operations capabilities without making additional investments in assets such as inventory. Indeed, many retailers strive to develop such capabilities to avoid letting inventories grow faster than sales (Fisher and Raman 2010). However, overly pursuing inventory leanness may harm profitability of retailers, who often face an "earns vs. turns" trade-off (Gaur et al. 2005). That is, retailers who pursue high earns must increase positions of high-value inventory expected to have fewer turns. Moreover, retailers that engage in exuberant inventory reduction could suffer from high stock-out costs, loss of goodwill, frequent restocking efforts, and extra transportation expenses. Despite industry evidence that US retailers are on trajectories to reduce inventories, the relationship between inventory leanness and firm performance remains unclear, and the conditions under which this relationship varies is still not fully understood.

To examine the non-transparent association between inventory leanness and operational performance, we capitalize on microeconomic theory of production frontiers to investigate the inventory-performance link using data on publicly held US retailers. We hypothesize a nonlinear association between inventory leanness and *operational efficiency*, and how the association is moderated by firm size and demand uncertainty—both critical to inventory decisions and firm performance. We use a formal definition of operational efficiency based on a firm's performance relative to an efficiency frontier composed of best-practice companies who convert production inputs into outputs (e.g., gross profit in our case) in the most efficient manner (Lee and Johnson 2013).

While inventory level is one of the most important operational levers that managers have in the retail industry, our two moderators—firm size and demand uncertainty—carry strategic and tactical implications, especially in our study with a partial goal of assessing retailer performance based on an efficiency frontier. On one hand, firm size is an accumulated product of managerial decisions in response to internal/external resources, constraints, and targets (i.e., the firm's growth strategy). It is not uncommon for firms to expand square footage/workforce or downsize sites/employees due to financial motives and environmental changes. The evolution (Angelini and Generale 2008, Kumar et al. 2001) and the direct impact of size on firms' actions, structures, and gains has long been of interest to industrial economists (Amato and Amato 2004, 2012). Yet, the indirect effect of firm size and how it could condition the impact of internal factors on productive efficiency is not well-understood (Halkos and Tzeremes 2007). On the other hand, an industry's environmental dynamism due to its substantial impacts on organization strategies and activities, has long been recognized as an important moderator in strategy research (see Suarez and Oliva 2005 for a review of the strategy literature; Azadegan et al. 2013 and Eroglu and Hofer 2014 explicitly use it in the operation management literature). Since retailers' performance outcomes are largely contingent on changes in consumer demand and use inventory as a buffer against market volatility, we use demand uncertainty as an indicator of environmental dynamism in retailing. While firm size is positively correlated with "scope of operations" and reflects economies of scale, demand uncertainty reflects the unpredictability and rate of change in a firm's operating environment. Jointly, firm size and demand uncertainty impose constraints on the retailers' inventory decision space and affect the inventory leanness-organizational performance relationship. Our study helps managers assess the competitive landscape better and take cautions when leaning out inventory.

Contrary to manufacturing sector studies that find the empirical association between inventory leanness and performance to be an increasing function (Lieberman and Dhawan 2005), we find support for the hypothesis that retailer operational efficiency has an inverted U-shape relationship to inventory leanness, suggesting an optimal inventory leanness level beyond which retailer operational efficiency degrades. While the inverted U-shaped relationship may be intuitive, to date empirical evidence in the retail sector has mainly identified linear relationships between inventory variables and financial ratios (e.g., Chen et al. 2007, Gaur et al. 1999, 2005, 2014, Johnston 2014). This association, however, is heavily moderated by firm size and demand uncertainty. Our evidence suggests that when increasing inventory leanness, small retailers exhibit efficiency degradation whereas larger retailers are capable of extracting efficiency improvement, with diminishing returns. We also find that under high demand uncertainty, being less lean is associated with higher operational efficiency, regardless of firm size.

After establishing the validity of our empirical estimation through multiple robustness checks (e.g., spline regressions, alternative functional forms), we conduct an exploratory analysis of how different retail categories vary in their operational efficiency scores, which illuminate competitive performance within- and across-segments in the retail industry. Finally, we conduct post hoc interviews with retail executives who offer practical perspectives that ground the nonlinear relationship between inventory and performance. The interviews not only confirm the managerial relevance of the two proposed moderators (i.e., firm size and demand uncertainty) but also suggest more nuanced moderators for future studies.

Next, section 2 defines the notion of retailer operational efficiency, investigates the inventory leanness-firm performance link, and formulates research hypotheses. Section 3 introduces our model specification and estimation methods. Section 4 describes data and operationalization of variables. Section 5 shows estimation results and robustness checks. Section 6 reports post hoc response analyses, descriptive statistics of operational efficiency estimates across retail categories, and post hoc interviews with retail interviews. We conclude by discussing practical implications for the retail industry and future research directions.

2. Related Literature and Research Hypotheses

2.1. Retailer Operational Efficiency

Retailing is a large sector of the economy in most developed and developing countries, and it covers many segments (e.g., apparel, grocery, wholesale), in which changes in firms and markets occur frequently. The diverse and volatile nature of retailing creates constant challenges for retailers to maintain efficient operations. Consequently, the notion of retailer *productivity/operational efficiency* has drawn considerable attention from operations management (OM) researchers (e.g., Keh and Chu 2003, Thomas et al. 1998). Early research in the US food retailing industry found total factor productivity rose at a slower rate than in manufacturing during 1959–1979, and that it was stagnant or declining for 1972–1979 (Ratchford and Brown 1985). Industry-level data also suggests labor productivity has dropped in the food retail sector since 1972, despite technological and logistical innovations. This drop is partially explained by the expansion of services (e.g., deli, bakery) within food retail stores (Ratchford 2003). These prior studies, however, do not differentiate between *fixed assets* and inventory, considering them all to be part of the capital input for a retailer's production function. By separating inventory—probably the retailers' most critical *variable asset*—from the fixed capital input factors, our study gauges operational efficiency more precisely and explores the association between retailers' inventory performance and their operational efficiency.

A long research stream evaluates retailer operational efficiency at the store level and examines stores within a single company (Donthu and Yoo 1998, Reiner et al. 2013, Thomas et al. 1998) or multiple companies (Park and King 2007) using cross-sectional data. Few store-level studies investigate operational efficiency of multiple stores using longitudinal data. Keh and Chu (2003) is an exception. All these studies, however, focus mainly on identifying relevant inputs (e.g., labor, capital) and outputs (e.g., sales, profits) such that efficiency scores can be derived to assess relative store performance, thus limiting the generalizability of findings. Banker et al. (2010) is an exception that tries to test the impact of antecedents such as supervisory monitoring, competition, and demographics on operational efficiency in high-end retail stores within a single firm.

In contrast to studies that focus mainly on industry-level productivity or store-level efficiency, our research contributes by examining operational efficiency at the firm-level across retailers from diverse retailing segments. A few prior studies (e.g., Mostafa 2009, Sellers-Rubio and Más-Ruiz 2009) examine firm-level retailing efficiency using cross-sectional data. Those papers give an overview of efficiency differences between the sampled companies. However, variation in retailer operational efficiency over time is not identifiable in a cross-sectional design. We address this issue by performing a panel data analysis. Relative to extant literature, our cross-segment and longitudinal research design better accommodates inter-sector and time heterogeneity, enhancing generalizability of findings.

Our work also contributes by analyzing retailer operational efficiency using stochastic performance frontier methods (Coelli et al. 2005). The analysis methods go beyond deriving efficiency estimates, enabling one to provide parameter estimates for hypothesized drivers of operational efficiency. The estimates have clear economic interpretations that pinpoint directions for performance improvement and might help managers develop operations strategies or adjust resource allocations (Banker et al. 2010). Specifically, we contribute by applying economic theory to empirically estimate how inventory leanness links to retailer operational efficiency, while controlling for other efficiency drivers that are not direct inputs but nevertheless may impact operational efficiency (Banker and Natarajan 2008).

2.2. Inventory Leanness-Firm Performance Link

The association between lean inventory and manufacturer performance at either the plant- or the firm-level is an often-studied topic in OM (e.g., Eroglu and Hofer 2011, 2014, Isaksson and Seifert 2014). With the spread of lean manufacturing and just-in-time production, many manufacturers undertook initiatives to reduce raw materials, work-in-process, and finished goods inventories. Lieberman and Demeester (1999) find inventory reduction is a driver of Japanese automotive manufacturer productivity. Chen et al. (2005) find a decrease of inventory coverage in US manufacturers between 1981 and 2000. Given empirical evidence of how lean inventory control positively associates with manufacturer performance (Huson and Nanda 1995, Koumanakos 2008), inventory reduction has become common practice in the manufacturing industry, typically achieving favorable outcomes.

Unlike with manufacturers, empirical analyses of US retailers provide mixed support for lean inventory initiatives. Examples of cross-firm analyses that are favorable to lean principles include Gaur et al. (2005) who show that inventory turnover is positively associated with above-expectation sales and Gaur and Kesavan (2009) who find that inventory turnover is positively associated with sales growth rate. However, other within-firm analyses (e.g., Koschat 2008, Soysal and Krishnamurthi 2012) report that reducing inventory could reduce sales due to stockouts. While the mixed findings may arise from divergent sampling and measurement approaches, the inconclusiveness about inventory leanness can also be attributed to rapidly changing retail environments and operational strategies. For instance, retailers may choose to hold high inventories not only to reduce lost sales, but also to stimulate sales through *billboard effects*—ample stock quantities may increase customer awareness of items as well as purchase intention (Cachon et al. 2019). This strategy has been analyzed in theoretical inventory models (e.g., Baker and Urban 1988, Balakrishnan et al. 2004, Datta and Paul 2001, Urban and Baker 1997) and suggests retailers need sufficient inventory to reach their latent sales potential.

Despite the strategic motive of holding abundant inventory to lure consumers, strong beliefs seem to persist that lower inventory levels are synonymous with better inventory performance across industries (Mishra et al. 2013). Leading retailers such as Walmart and JC Penney cut back their merchandising stocks and consider the results as significant inventory improvement (Cassidy 2016). This "less-is-more" lean philosophy (i.e., making shelves and backrooms barer) embraced by Home Depot tends to spread across different retail sectors. Contrary to the "piling up inventory" strategy mentioned earlier, numerous managers posit that stacking it high would not make it fly in the era of multi-channel retailing (Ziobro 2016b).

Those altering views suggest that the inventory leanness-firm performance link is a complex issue for retail industry practitioners as well as researchers and it has drawn substantial attention in recent years. Several studies (Chen et al. 2007, Gaur et al. 1999, 2005) empirically assess associations between inventory turnover and financial performance in US. retailers. Gaur et al. (2005) find a negative association between inventory turnover and gross margin, widely known as the "earns vs. turns" trade-off. A number of studies in this research stream (e.g., Gaur et al. 2014, Johnston 2014) use variants of inventory turns to measure retailer inventory performance and consistently find a negative correlation between inventory turns and gross margins, as retailers who aim for high gross margins usually must increase levels of high-value inventory.

To date, these analyses on the effect of inventory leanness on firm performance have been based on accounting or financial metrics like return on assets or return on sales. These metrics, however, ignore effects of different operational strategies involving various mixtures of basic production factors. To correct for this, we adopt a performance metric—operational efficiency—based on the transformation of inputs to output in a neoclassical production function. This metric has the added benefit that it is assessed relative to the efficiency frontier created by best performers in the industry. Moreover, under classical inventory models, numerous retailers' ordering decisions are made to avoid lost sales and maintain service levels, without referring to labor and capital, thus adding nuance to the link between inventory leanness and operational performance. Hence, our study further evaluates the effect that the retailers' internal resources and demand uncertainty have on moderating the inventory-performance link.

2.3. Inventory and Performance Metrics

Our empirical investigation into the link between inventory leanness and operational efficiency is distinctly different from prior studies in that both the leanness and efficiency metrics in our study consider *peer effects* and *returns-to-scale*. As mentioned earlier, inventory level is the managerial lever of interest. While inventory level (e.g., days of supply) (Rajagopalan 2013) or inventory turns are volume-adjusted metrics, that is, they measure inventory relative to the sales volume, they are calculated solely based on a firm's own outcomes and are not useful for industry comparisons as firms face different demand patterns and segments differ in inventory requirements. Instead, we build on Eroglu and Hofer's (2011) empirical leanness indicator (ELI) to capture how well a retailer *converts its inventory to sales* relative to competitors in the same industry segment.[1] The competition-sensitive ELI not only addresses the foregoing limitations and but also allows us to compare the firm's performance relative to the competitors. Eroglu and Hofer (2011, 2014) provide detailed theoretical foundations of the ELI and explain why ELI is superior to other inventory metrics that also consider peers in the same segment, for example, adjusted inventory turnover (Gaur et al. 2014) and standardized sales-to-inventory ratio (Mishra et al. 2013, Modi and Mishra 2011). They argue that ELI is "easily interpretable, comparable across industries, effectively addresses concerns of potential attenuation bias, and it is based on flexible and robust regression analyses" (Eroglu and Hofer 2014, p. 350). Furthermore, ELI estimation is based on an empirically validated inventory throughput function that explicitly considers a parameter that captures returns-to-scale (Ballou 2005, Eroglu and Hofer 2011, Waller and Esper 2014). Despite being an aggregate post hoc measure, the focal variable—inventory leanness measured by ELI—can be viewed as a sales-adjusted relative inventory level, which is highly reflective of firm-level inventory decisions.

With regard to the measure of retailer performance, in our study, we define operational efficiency—a latent distance that reflects how well a retailer *utilizes its labor and capital to create economic value added*—as the performance metric. Like the performance measure used by related literature in retailer efficiency in §2.1, the operational efficiency is essentially technical efficiency in production economics (Coelli et al. 2005) and thus it reflects a retailer's distance to the efficiency frontier composed of *best-performing peers* in the industry. In the estimation processes, returns-to-scale are accommodated by output elasticities of input factors. Hence, our metric captures operational efficiency, which from an OM standpoint is a superior metric than ratio indices (e.g., ROA, ROI) and comparable across industries (Chen et al. 2015, Lam et al. 2016).

Moreover, since we use gross profit to measure economic value added, the efficiency estimates from the stochastic frontier analysis reflect not only operational efficacy but also financial profitability. Even though the outputs of production functions are usually financial metrics such as sales or gross profit, as a distance estimate, technical efficiency captures how productive a firm is in operating (i.e., turning inputs to outputs). Thus, the measure is labeled *operational efficiency* in the production economics literature (e.g., Lee and Johnson 2013, Yu and Ramanathan 2008, 2009) and OM research (e.g., Lam et al. 2016, Saranga 2009, Sarkis 2000), although it is an estimation based on operational inputs and financial outcomes. Compared to other oft-used retailer performance measures (e.g., revenue, operating income, stock market returns, Tobin's *q*), our frontier-based efficiency metric can be deemed as a more operational and holistic performance measure that accounts for input factors, contextual factors, and firm heterogeneity, and is more stable to variations in financial markets.

Note that while leanness might be often understood as operational efficiency and vice versa, in our case, *inventory leanness* and *operational efficiency*, as explained above, have specific technical definitions that are not equivalent. While the ELI proposed by Eroglu and Hofer (2011) measures inventory level (variable asset) relative to sales and other competitors, it does not consider other resources available to the organization. Operational efficiency, on the other hand, captures the effect of inputs (labor and fixed assets) as well as technology change on economic value added (Lam et al. 2016, Li et al. 2010). Although we would not argue that *inventory leanness* is orthogonal to *operational efficiency*, the two metrics differ by construction and they capture complementary aspects of retailer performance while considering peer competition.

2.4. Research Hypotheses

Theoretically, it is reasonable to expect the existence of an unknown *optimal level* of inventory leanness with the objective of maximizing firm performance, leading to a commonly accepted inverted-U relationship between inventory leanness and firm performance (Eroglu and Hofer 2014). However, the seemingly self-evident inverted-U shape should not be rationalized by simply stating that "too much of a good thing can be harmful;" a specific account of the mechanisms that create that U-shape must be articulated (Haans et al. 2016).

In the retailing context, low inventory leanness (i.e., a firm achieves its sales level with relatively high inventory levels in its industry sector) has negative implications for firm performance. First and foremost, excess inventory locks down financial resources, thus imposing constraints on resources for improvement initiatives and capital investment, both of which are expected to improve operational efficiency. Second, retailers who possess high inventory levels are more likely to activate clearance sales (e.g., higher inventories make them less responsive to changing preferences) and inventory write-offs that hamper profitability (Kesavan and Mani 2013). Finally, high inventory levels make instore logistics more prone to execution errors, causing shelf stockouts, for example, the item is in the store (e.g., backroom) but unavailable to customers (Fisher and Raman 2010).

Higher inventory levels, however, may also improve the retailers' performance. First, higher inventories make retailers less likely to experience stockouts and avoid the accumulation of unsatisfied latent shopper demand due to bare shelves that demotivate customers (Ziobro 2016b). Second, retailers with higher inventories are more probable to offer rich assortments that stimulate sales through variety effects (Cachon et al. 2019) or effects from high display quantities (Balakrishnan et al. 2004, Krommyda et al. 2015, Stavrulaki 2011). Nevertheless, those abundance effects saturate above certain levels as inventory availability is not enough to generate demand by itself. The negative implications of low inventory leanness, and the marginal benefits—after saturation—of higher inventories, suggest that retailers could achieve performance gains by taking inventory reduction initiatives, either by lowering inventory level or by reducing the number of in-store SKUs.

As relative inventory level drops, however, negative effects on operational efficiency could take place. Specifically, at high levels of inventory leanness, retailers demand more frequent in-bound/out-bound replenishment and transportation, which require labor and/or capital investments that may reduce operational efficiency. Also, limited inventory levels reduce chances of up- or cross-selling and in turn reduce the utilization of other inputs as idle labor and capital. Moreover, the elimination of the negative consequences of high inventory identified above (i.e., locking in of financial resources, increased probability of clearances and write-offs, and volume-related operational errors) will also exhibit diminishing benefits as financial returns of subsequent investments drop, clearances are less frequently required, and lean operations become less error-prone. All these negative implications, combined with the diminishing benefits of leaner inventory level, suggest that the negative effects can outweigh performance gains when inventory leanness goes beyond certain levels, pointing to the downward side in the inventory leanness-operational efficiency link.

The foregoing reasoning suggests that there is a smooth continuum of operational performance with either low or high inventory leanness having detrimental effects on performance after the benefits of the relative inventory position have maxed out. Those tradeoffs suggest the link between inventory leanness and operational efficiency to manifest itself in an inverted U-shape.

HYPOTHESIS 1. *The level of retailer inventory leanness will have an inverted U-shaped association with operational efficiency.*

The first moderator we examine is *firm size*, which is positively correlated with "scope of operations" (Porter 2008) and reflects economies of scale (e.g., fixed operating expenses) that are critical for operational efficiency. The fact that returns-to-scale are explicitly taken into account in both inventory leanness and operational efficiency metrics, allows us to isolate the direct/moderating effects of firm size itself on operational efficiency. Amato and Amato (2004) posit that internal firm strategies and external market changes will affect retailer profitability differently at various firm sizes, particularly due to market power and strategic advantage possessed by large retailing firms. Gaur and Kesavan (2009) provide an in-depth discussion about how large retailers benefit from their size from an inventory control point of view. Their basic premise, supported by safety stock reduction through risk pooling (Eppen and Schrage 1981) and fixed cost reduction in the EOQ model, is that inventory grows less than linearly in stores, products, or sales.

Extending their viewpoint, we posit from different perspectives that large retailers could benefit more from inventory leanness than small retailers. First, when both a large retailer and a small one try to become leaner, the former has more stores and/or products to leverage demand pooling effects and reach the same service level with fewer inventories. Also, to facilitate pooling operations and fully realize performance gains from pooling, large retailers usually have more financial slack (Amato and Amato 2004) to invest in advanced material handling (e.g., warehouse management system) and in-bound/out-bound logistics capabilities (e.g., transportation management system, yard management system) (Mason et al. 2003). Those resources/capabilities are crucial for fulfilling consumer demand with lower inventories in a cost-effective and timely manner.

Second, it is justifiable that perfect competition does not hold in the US retailing market since giant retailers like Walmart do have strong market power and lower input prices (Bloom and Perry 2001). Thus, large retailers might extract financial benefits from suppliers (e.g., bulk purchasing, bargaining power, vendor-managed inventory) (Mottner and Smith 2009) and even make suppliers accountable for inventory ownership until inventories are sold (Cetinkaya and Lee 2000). With their bargaining power in supply networks, large retailers could set up favorable contracts and payment terms (Serrano et al. 2018). Since in general retailers, regardless of firm size, have immediate receivables from customers, deferring payables to suppliers and potential rebates from suppliers allow large retailers to optimize their working capital without compromising operational efficiency (reflective of operational and financial performance).

Third, in addition to the advantages listed above, strategic initiatives such as supplier auditing/monitoring, collaborative forecasting, and trade credit are more likely to be taken by large retailers with higher leverage in supplier relationships (Murfin 2014, Serrano et al. 2018). It is intrinsically more difficult for small retailers to implement supplier evaluation and integration programs, since small retailers may account for only a small portion of large suppliers' business. Those evaluation and integration programs (e.g., certify or even finance suppliers) may help large retailers to become leaner and reduce cost of goods sold without compromising their sales performance.

In sum, all of the afore-mentioned scale advantage, market power, and strategic initiatives tend to give large retailers a potential edge in exploiting inventory leanness and reducing cost of goods sold (and hence improving profitability). Thus, we expect large size retailers to be less vulnerable to the negative implications of pursuing inventory leanness for operational efficiency detailed in *H1*. Hence, we hypothesize:

HYPOTHESIS 2. *The association between inventory leanness and operational efficiency is positively moderated by the firm size. That is, size will shift the turning point of the inverted U-shape toward higher inventory leanness.*

The second moderator we examine is *demand uncertainty*, which at a strategic level reflects the unpredictability and stochastic rate of change in a firm's operating environment (Azadegan et al. 2013, Eroglu and Hofer 2014, Gligor 2016, Suarez and Oliva 2005). Operationally, demand uncertainty refers to the degree to which a firm can forecast sales (Germain et al. 2008). Demand uncertainty is a major contributor to overall uncertainty and perceived to have negative influences on firm performance, because it compromises firms' capabilities to predict and react to changing conditions in consumer markets (Beckman et al. 2004). In addition to interfering with operational decisions, demand uncertainty can further jeopardize strategic decision-making for people, process, and technology, as judgment biases tend to increase when managers face higher uncertainty (Eroglu and Hofer 2014). The effects of demand uncertainty on stability of operations are expected to manifest themselves in retailer efficiency as well as in the link between inventory leanness and operational efficiency.

In the presence of demand uncertainty, we posit that the positive implications of inventory leanness for operational efficiency would be more limited due to the following reasons. First, random fluctuations in demand cause instability in store and warehouse operations (e.g., stockouts, transaction errors) for nearly all retailing firms. Given the intrinsic difficulty of predicting the timing and magnitude of demand uncertainty, retailers under higher uncertainty are expected to have irregular schedules of in-store replenishment and warehouse delivery. Those retailers probably also have extra difficulties in optimizing/stabilizing decisions regarding labor, assortment, and pricing, among others. All the afore-mentioned irregularities and difficulties can harm operational efficiency and make it less probable for a retailer to sustain efficiency gains in the midst of pursuing leanness. Second, the central tenet of related literature is that demand uncertainty hinders retailers from being leaner (Rajagopalan 2013). Volatile demand will force a retailer to buffer against uncertainty by carrying additional inventory (e.g., safety stock increases in demand variation) (Hancerliogullari et al. 2016). When retailers pursue inventory leanness or destock inventory under high demand uncertainty, they will be more likely to lose sales opportunities and underutilize labor and capital investments when demand spikes. Such degrading of customer service levels can make the negative implications of inventory leanness outweigh its positive implications for operational efficiency. Thus, we hypothesize:

HYPOTHESIS 3. *The association between inventory leanness and operational efficiency is negatively moderated by the demand uncertainty. That is, demand uncertainty will shift the turning point of the inverted U-shape toward lower inventory leanness.*

3. A Stochastic Production Function Model

We apply the stochastic frontier analysis (SFA) technique (Aigner et al. 1977) to estimate retailer operational efficiency and examine our research hypotheses. SFA models can exploit the panel data structure to accommodate firm heterogeneity (Greene 2005). Given the existence of potential errors in firm-level measures due to store-level data aggregation and retail industry market uncertainties (Sellers-Rubio and Más-Ruiz 2009), the fact that SFA explicitly incorporates modeling of random noise components in the production function makes SFA appropriate for our longitudinal efficiency analysis. SFA also addresses non-symmetric error terms caused by inefficient operations, which regular regression models (e.g., OLS, fixed effects) do not. Our model is adapted from Lieberman and Dhawan (2005), but we make several changes to address firm heterogeneity and improve estimation stability and consistency.

We specify the stochastic production frontier as follows:

$$\begin{aligned} Y_{it} &= F(L_{it}, K_{it}) OE_{it} e^{V_{it}} \\ &= F(L_{it}, K_{it}) e^{-U_{it}} e^{V_{it}} \end{aligned} \tag{1}$$

where Y_{it} is the output for the ith firm in year t and F (L_{it}, K_{it}) is the production function in terms of labor (L_{it}) and capital (K_{it}). OE_{it} is the operational efficiency, a scaling factor ranging in (0, 1]. We operationalize OE_{it} as $e^{-U_{it}}$, where U_{it} is a non-negative random variable that reflects operational *inefficiency* (Jondrow et al. 1982) and V_{it} is an i.i.d. $N(0, \sigma_V^2)$ error term that captures random shocks affecting the outputs.

Following economic theory, we employ a Cobb–Douglas production function for $F(\cdot)$. One may estimate such a function in several ways, depending on the specification of the exponential intercept term. We examine several specifications, with the most sophisticated model as follows:

$$F(L_{it}, K_{it}) = e^{\alpha_i + \beta\tau} L_{it}^{\gamma} K_{it}^{\delta} \tag{2}$$

where the year index τ is a continuous time trend variable. The equation parameterization captures the unobserved heterogeneity α_i and the time parameter β, both of which could affect production outcomes. This "true fixed-effects" (TFE) model (Greene 2004) improves on the Lieberman and Dhawan (2005) specification by incorporating firm heterogeneity to capture more than just labor and capital. Accommodating firm fixed effects α_i not only alleviates endogeneity concerns attributed to unobservable factors, but also relaxes the assumption of homogeneous production units, which is not fully realistic for retailers, given that assortments and services provided by retailers can differ substantially across different retailers' store formats (e.g., supermarkets, drug stores). The time parameter β accounts for technology change (e.g., improved IT infrastructure) and macro-economic volatility across years (Bloom and Perry 2001). Note that the parameters γ and δ provide information about returns-to-scale. Specifically, $\gamma + \delta = 1$ indicates constant returns-to-scale, $\gamma + \delta < 1$ refers to decreasing returns-to-scale, and $\gamma + \delta > 1$ implies increasing returns-to-scale. Combining Equations (1) and (2), and taking logarithms, we derive the production model to be estimated as follows.

$$\ln(Y)_{it} = \alpha_i + \beta\tau + \gamma \ln(L)_{it} + \delta \ln(K)_{it} - U_{it} + V_{it} \tag{3}$$

Finally, we use the Wang and Ho (2010) model to capture the operational inefficiency U_{it} as a function of a vector of explanatory variables $\mathbf{Z}_{it}$ and an i.i.d. half-normal random variable with mean μ and variance σ_U^2.

$$U_{it} = f(\mathbf{Z}_{it}\boldsymbol{\omega}) \times N^+(\mu, \sigma_U^2) \tag{4}$$

where $f(\cdot)$ is a positive *scaling* function such as exp ( ) to ensure the non-negativity of U_{it} and the half-normal random variable $N^+(\mu, \sigma_U^2)$ captures the time-invariant component of inefficiency. Essentially, firm-specific factors $\mathbf{Z}_{it}$ and their coefficients ω shrink or stretch the scale of the underlying inefficiency distribution. Alvarez et al. (2006) discuss advantages of this *scaling model* over earlier SFA models (e.g., Aigner et al. 1977, Battese and Coelli 1995) by allowing $\mathbf{Z}_{it}$ to affect the whole distribution (as opposed to only the mean) and allowing U_{it} to be correlated over time (via time-invariant $N^+(\mu, \sigma_U^2)$).

We use the following specification of the scaling model $f(\mathbf{Z}_{it}\boldsymbol{\omega})$ to empirically examine the hypothesized inventory leanness-operational efficiency relationship.

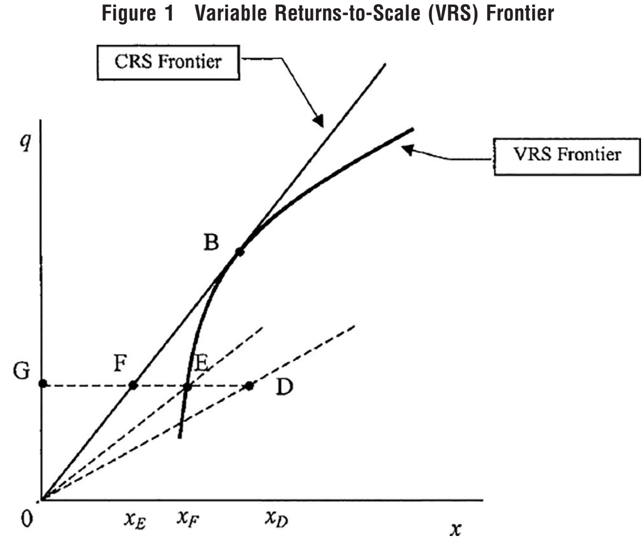
$$\begin{aligned} f(Z_{it}\omega) = \exp\big(&\omega_1 InvLean_{it} + \omega_2 InvLean_{it}^2 + \omega_3 Size_{it} \\ &+ \omega_4 DUnc_{it} + \omega_5 SalesGR_{it} + \omega_6 StoreGR_{it} \\ &+ \omega_7 LaborGR_{it} + \omega_8 CapitalGR_{it} \\ &+ \omega_9 InvLean_{it} \times Size_{it} + \omega_{10} InvLean_{it} \times DUnc_{it}\big) \end{aligned} \tag{5}$$

An intercept is not included as it is econometrically unidentifiable (Wang and Ho 2010). As articulated in §2, the association between inventory leanness and firm performance is expected to be non-linear, thus the quadratic term of *InvLean* is included in the model. We also include two moderators associated with retailer performance—firm size (*Size*) (Gaur and Kesavan 2009) and demand uncertainty (*DUnc*) (Rajagopalan 2013). In line with Shockley et al. (2015), who empirically show that retailers' growth/decline in sales, stores, labor, and capital from period to period has substantial impacts on their operational performance (ROA and ROS), our model also controls for a firm's sales growth (*SalesGR*), growth of the number of stores within the retail firm (*StoreGR*), labor growth (*LaborGR*), and capital growth (*CapitalGR*). Those growth factors affect how retailers act and perform since they are related to macro-economic trends and their financial health. Gaur and Kesavan (2009) posit retailers act and perform differently during expansion or contraction periods. Also, when more resources become available to retailers due to *SalesGR*, the difference in their resource allocation to store, labor, and capital would affect operational efficiency. Hence, despite moderate inter-correlations (~0.5), the set of growth rates need to be included. Note that it is fairly common for SFA models to have correlated input factors and efficiency drivers. Even when an input factor is identical to an efficiency driver, the distributional assumptions on inefficiency terms permit the effects to be identified in both production and inefficiency functions (Battese and Coelli 1995). Finally, the two interaction terms (ω_9 and ω_{10}) in Equation (5) are sufficient for us to test *H2* and *H3* on the shift of turning points. Interaction terms between *Size*/*DUnc* and the quadratic term of *InvLean* are necessary only when testing whether moderators flatten or steepen the U-shape (Haans et al. 2016). A detailed explanation of how we operationalize all variables is provided in the next section.

It is worth clarifying at this point the differences between economies of scale and returns-to-scale, as the terminologies are often improperly used interchangeably in applied research (Beattie et al. 1985). The two terms are related but represent different concepts. Returns-to-scale relates to the response of production output to a proportional expansion of all input factors so that it captures merely the technological aspects of scale economies. Mathematically, returns-to-scale is the sum of partial elasticities of production. On the other hand, economies of scale include operational or cost advantages gained from size (e.g., bulk purchasing, market power). Technically, economies of scale refer to "returns-to-size that has more to do with a proportional change in output as input factors are expanded in least-cost fashion" (Beattie et al. 1985). In particular, firms having lower input prices can attain economies of scale even when they experience decreasing returns-to-scale (Cohn 1992, Gelles and Mitchell 1996). Also, since our model imposes no prior assumptions on returns-to-scale, the estimated frontier function exhibits variable returns-to-scale (VRS). As a result, a firm can lay on the frontier (i.e., having no operational inefficiency) but still have scale inefficiency (Coelli et al. 2005). We illustrate this idea via Figure 1, in which firms B and E are operationally efficient (i.e., lying on the VRS frontier). However, only firm B has no scale inefficiency as firm E deviates from the constant returns-to-scale (CRS) frontier. Thus, the idea of including firm size in Equation (5) is to understand the association between returns-to-size and operational efficiency (from parameter ω_3) rather than to make inferences about returns-to-scale, which is assessed through parameters γ and δ.

Together, the production model in Equation (3) and the inefficiency model in Equations (4) and (5) jointly constitute the stochastic production function to estimate. Despite a substantial improvement over traditional SFA panel data models with firm-invariant intercept α, the introduction of firm-specific fixed effects α_i in our model creates the possibility of an incidental parameter problem (Greene 2005). We address this possibility by adopting the Wang and Ho (2010) technique that makes our model specification much more robust than prior SFA studies (Jorge-Moreno and Carrasco 2015, Lieberman and Dhawan 2005, Zhang et al. 2012) that assume fixed effects to be a firm-invariant constant. Wang and Ho (2010) analytically show that first-difference and demean transformations (i.e., the elimination of α_i) can be performed

Figure 1 Variable Returns-to-Scale (VRS) Frontier



Source: Adapted from Coelli et al. (2005).

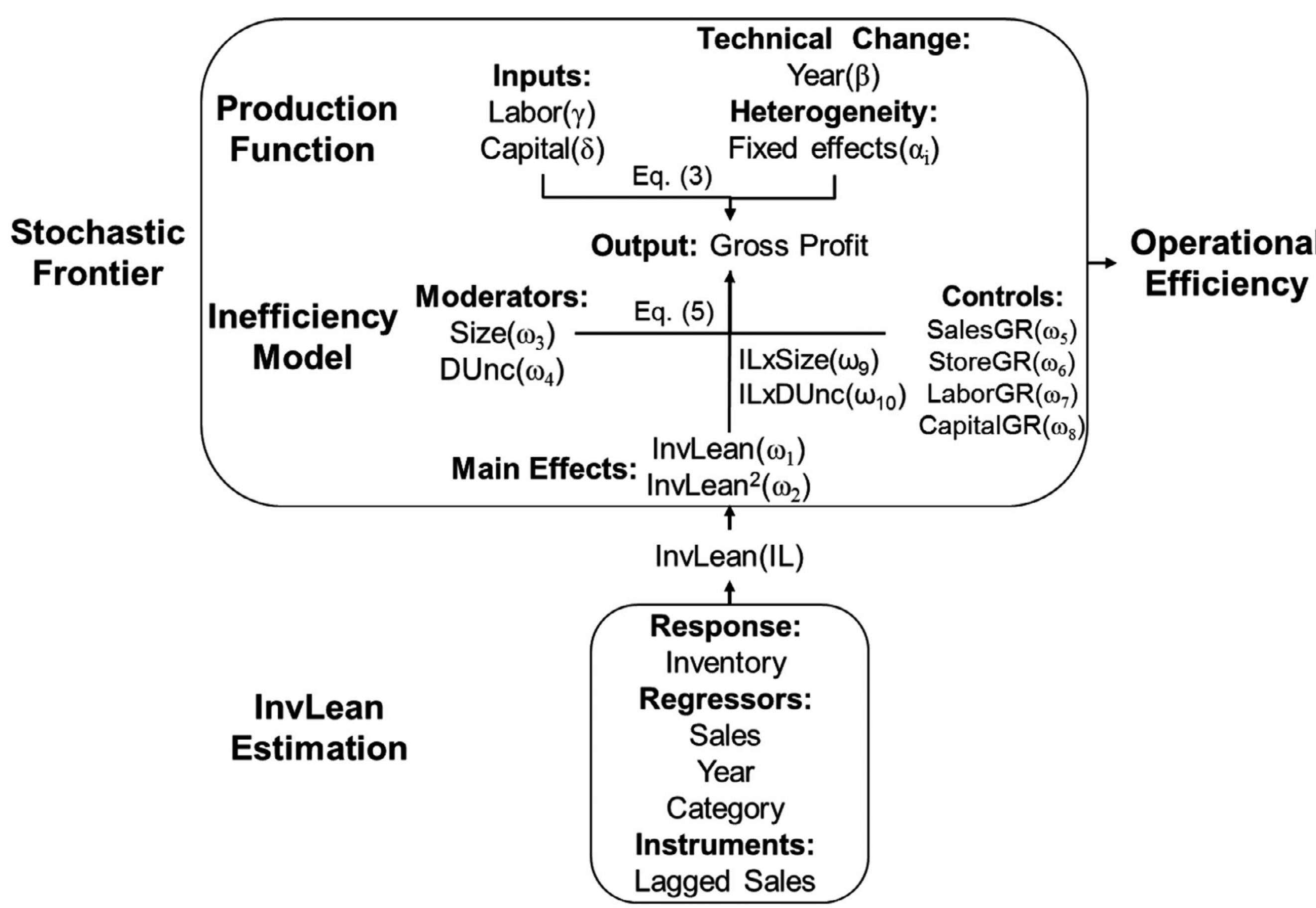
on this scaling model to avoid directly estimating α_i (typically not feasible for other SFA panel data models). After model transformation, estimation consistency for the remaining parameters is then achieved via maximum likelihood estimators that are immune to the incidental parameter issue (Kumbhakar et al. 2015). Using maximum likelihood estimation (MLE) to estimate the stochastic frontier model in one stage ensures consistent estimates and also avoids the conceptual inconsistencies of two-stage procedures, in which operational efficiency is estimated first without considering $\boldsymbol{Z}_{it}$ and then efficiency estimates are regressed on $\boldsymbol{Z}_{it}$ (Wang and Schmidt 2002).

Figure 2 provides an integrated view of our stochastic frontier panel data model and analysis strategy. The stochastic frontier analysis in the upper part is composed of the Cobb–Douglas production function (Equation (3)) and the operational inefficiency model (Equation (5)). As discussed above, Equations (3) and (5) jointly constitute the stochastic production function and the two equations are simultaneously estimated in a joint log-likelihood function. The lower part of the figure shows the components of the empirical inventory leanness estimation, to be discussed in detail in the next section. Figure 2 makes it clear that operational efficiency is a function of the estimated production function, its inputs and contextual factors, and the factors captured in the inefficiency model.

4. Data Description

We use data from publicly traded US retailers (Gaur and Kesavan 2009, Gaur et al. 2005, Kesavan and Mani 2013, Shockley et al. 2015). The US Department of Commerce identifies retailer categories using two-digit standard industrial classifications (SIC) from 52 to 59. Due to our interest in associations of inventory with operational efficiency, we omit retailers that have limited control over their own inventory levels. We eliminate non-store retailers (SIC 5960, SIC 5961) and credit card companies (SIC 5900) in the miscellaneous retail category (Kesavan et al. 2010). We exclude automotive dealers and gasoline service stations (SIC 55) and eating and drinking places (SIC 58), as they have sizable service business components and single-source suppliers that may intervene in inventory decisions (Kesavan et al. 2010).

Figure 2 Stochastic Frontier Analysis Model and Procedures



Our initial sample contains retailers that were still active in 2013. We obtain yearly data for 167 retailers for the 16-year period 1998–2013 from the COMPUSTAT database using Wharton Research Data Services (WRDS). The data in year 1998 and year 1999 are used for construction of some variables and lagged measures. Hence, we only analyze the time period from 2000 to 2013. To enable longitudinal analysis of operational efficiency over time, we only consider retailers that had at least 5 years of consecutive data. We choose 5 years following Greene's empirical examples of SFA when frontier models explicitly consider fixed effects (Greene 2004, 2005). Having at least 5 years of observations increases the consistency of the mean estimates of each firm to be used in the Wang and Ho (2010) within-transformation for SFA. Even though increasing the time period threshold may improve the firm-specific mean estimates for within-transformation, doing so would reduce the number of retailers included and result in a sample less representative of the US retail industry. Retailers that did not report information about inventory and the number of employees during 2000–2013 are omitted. Since foreign retailers might have different cost structures that could bias estimates for the production function, we also exclude non-US retailers listed as *American Depositary Receipt* (ADR) firms in the US stock exchange (Kesavan et al. 2010) and retailers whose *Foreign Incorporation Code* (FIC) is not for the USA.

Our overall potential sample contains 118 US retailers, representing 71% of the eligible COMPUSTAT sample. The sample for analysis contains 1531 observations. The effective sample size for stochastic frontier estimation is below 1531 due to using 1-year lagged variables and since certain observations are eliminated due to missing data. The panel is unbalanced since some firms do not have full information for all variables across all years. To make monetary measures comparable across years and absorb some macro-economic trends, we deflate all monetary variables into US dollars for the year 2000 using the consumer price deflator calculated by the Bureau of Labor Statistics (Banker et al. 2010). Table 1 summarizes, by category, the information about the firms in our sample. While we cannot fully rule out the possibility that managers may manipulate financial data, it is reasonable to assume that the secondary data accurately reflects firm characteristics since the numbers in COMPUSTAT have been audited. Also, the Sarbanes–Oxley Act of 2002 requires executive managers to testify (sign-off) that the reported numbers are correct. Accordingly, COMPUSTAT data has been used in many published analyses (e.g., Chen et al. 2005, 2007, Gaur et al. 2005, Kesavan et al. 2010).

The standard output metric used in the SFA literature is economic value added, and sales revenue and gross profit are both popular output measures in retailer efficiency studies (Fu et al. 2015, Sellers-Rubio and Mas-Ruiz 2006, Thomas et al. 1998). We use annual *gross profit*—net annual sales minus the cost of goods sold (COGS)—as the measure of output (i.e., Y_{it}) for our analysis following Parmeter and Kumbhakar (2014) who suggest that it is possible to model an efficiency frontier in the context of profit maximization as a function of observed covariates and unobserved inefficiency. While empirical literature in the retail space has shown that inventory is positively related to sales (Baumol and Ide 1962, Dubelaar et al. 2001), a retailer's sales revenue does not fully capture the expected costs and benefits of inventory leanness (e.g., lower inventory levels, less warehouses, lower spoilage). For instance, some lean retailers may pay higher prices to have suppliers hold inventories on their behalf. Gross profit avoids the potential for the effect of inventory leanness to be overstated by absorbing such price inflations that are not captured by sales dollars. Using gross profit as the output variable also controls for the influence of cost elements/input prices that are not included as production factors. Finally, even though net profit also considers profitability, it includes noisy administrative expenses (e.g., attorney fees, insurance) and tends to violate monotonicity in production (i.e., outputs need to be non-decreasing in inputs). In line with prior SFA

Table 1 Retailer Category and Sample Descriptors

SIC code	Category	N	Examples	Average gross profit (10^6\$)	Average sales (10^6\$)	Average inventory (10^6\$)	Average employees (10^3)	Average capital (10^6\$)
52	Building Materials	5	Home Depot; Lowe's	7060	20,909	3216	114	7800
53	General Merchandise	19	Macy's; Sears	7431	27,861	3215	185	8508
54	Food	17	Kroger; Safeway	3047	11,560	745	70	3051
56	Apparel and Accessories	40	DSW; GAP	875	2313	319	25	556
57	Home Furniture	13	Conn's, Pier 1 Imports	1575	4325	614	27	727
59	Miscellaneous	24	Chalet; Petsmart	2214	7691	1065	40	2114

studies (e.g., Ge and Huang 2014), using gross profit for economic value-added results in a model equivalent to a production function, but estimates inefficiency using a dependent variable that captures most of the inventory related costs as part of the cost of goods sold.

We operationalize labor and capital similar to previous studies on firm productivity (Imrohoroglu and Tuzel 2014) and operational efficiency (Jacobs et al. 2016, Mostafa 2009). We approximate labor by the number of employees. We measure capital as fixed assets (i.e., total assets minus total current assets).

Following Eroglu and Hofer (2011), we construct the inventory leanness *InvLean* as the negative of the Studentized residuals of the regression of total inventories on sales. We first consider the regression for each category (j) in each year (t) using ordinary least square (OLS).

$$\ln\left(InvAvg_{ijt}\right) = \alpha_{jt} + \beta_{jt}\ln\left(Sales_{ijt}\right) + e_{ijt}, \quad \forall j = 52,\ldots,59,\ \forall t = 2000,\ldots,2013 \tag{6}$$

where $InvAvg_{ijt}$ is average total inventory and $Sales_{ijt}$ is the net annual sales of retailer i in category j and year t. To address our sample characteristics, we make two modifications to the Eroglu and Hofer (2011) model. First, to avoid large standard errors that would result from 84 regressions (6 categories × 14 years) with limited observations, we introduce category dummies to capture the differences across categories. The category-specific coefficients estimate the heterogeneity across categories and provide us with estimates having smaller standard errors. In addition, inventory decisions arguably are endogenously chosen by management based on sales performance, and sales are driven by inventory as well (Balakrishnan et al. 2004). Specifically, retailers tactically use inventory to increase sales, and sales in turn provides an input to retailers' decisions about inventory (Kesavan et al. 2010). If the simultaneity between inventory and sales violates the orthogonality assumption $cov(Sales_{ijt}, e_{ijt}) = 0$, the OLS estimates of *InvLean* will be inaccurate (Wooldridge 2010). To tackle this issue, we use 2-year lagged sales (i.e., $Sales_{ijt-2}$) as an instrument to perform a two-stage least square (2SLS) estimation:

$$\left.\begin{aligned} ln(Sales_{ijt}) &= \theta_0 + \boldsymbol{\vartheta}_j + \theta_1 \ln(Sales_{ijt-2}) + e_{ijt} \\ ln(InvAvg_{ijt}) &= \tau_0 + \boldsymbol{\varphi}_j + \tau_1 \ln(\widehat{Sales}_{ijt}) + \varepsilon_{ijt} \end{aligned}\right\}\forall t = 2000,\ldots,2013 \tag{7}$$

where $\boldsymbol{\vartheta}_j$ and $\boldsymbol{\varphi}_j$ are vectors of the category dummies.

Table 2 summarizes the model fit and endogeneity test results of sales in the average inventory annual regressions. The 14 annual models (for years 2000–2013) are significant and explain over 88% of the observed variance. The F statistics of all first-stage regression models (fourth column in Table 2) are much higher than the threshold of 10 (Larcker and Rusticus 2010), indicating that 2-year lagged sales (i.e., $Sales_{ijt-2}$) is a good instrumental variable. The Wu–Hausman endogeneity test (Cameron and Trivedi 2009) rejects the null hypothesis that sales are exogenous for all but one of the years 2000 through 2013 (last column in Table 2). Consequently, we retain the 2SLS formulation.

Since we found no evidence that the coefficients for $ln(\widehat{Sales}_{it})$ for the 14 annual models are statistically different, we integrated the fourteen models into a single model by adding year dummies ($\boldsymbol{\rho_t}$ and $\boldsymbol{\lambda_t}$):

$$\left.\begin{aligned} ln(Sales_{ijt}) &= \theta_0 + \boldsymbol{\vartheta_j} + \boldsymbol{\rho_t} + \theta_1 \ln(Sales_{ijt-2}) + e_{ijt} \\ ln(InvAvg_{ijt}) &= \tau_0 + \boldsymbol{\varphi_j} + \boldsymbol{\lambda_t} + \tau_1 \ln(\widehat{Sales}_{ijt}) + \epsilon_{ijt} \end{aligned}\right\} \tag{8}$$

The integrated model confirms there is no annual variance, as the year dummies are not significant. The integrated model is highly significant ($F = 1130.52$, $p < 0.001$) and explains 94% of the observed variance. The inventory leanness metric ($InvLean_{it}$) is calculated, following Eroglu and Hofer (2011), by studentizing the residuals (ε_{ijt}) of Equation (8) and multiplying by (−1). This transformation ensures that positive deviations (i.e., above expected average inventory levels) result in negative *InvLean* and vice versa.

For the hypothesized moderating variables, we use the number of employees with 1-year lag (i.e., $Labor_{it-1}$) as the proxy for a retailer's *Size* (Angelini and Generale 2008, Gligor 2016, Kumar et al. 2001, Laeven and Woodruff 2007).[2] The operationalization

Table 2 Model Fit and Endogeneity Test Results of *Sales on Average Inventory Regression*

Year	N	R^2 (1st stage)	F-value (1st stage)	R^2 (2nd stage)	H_0 p-value
2000	88	0.94	221	0.95	0.000
2001	89	0.98	673	0.96	0.000
2002	93	0.99	1004	0.95	0.001
2003	92	0.99	1039	0.95	0.002
2004	96	0.99	1134	0.94	0.004
2005	98	0.99	1221	0.94	0.005
2006	100	0.98	910	0.95	0.046
2007	103	0.98	988	0.94	0.006
2008	102	0.99	1211	0.95	0.000
2009	105	0.99	1796	0.94	0.023
2010	109	0.99	2245	0.94	0.198
2011	113	0.99	1784	0.94	0.013
2012	119	0.99	1562	0.94	0.002
2013	119	0.99	1510	0.93	0.038
All	1,425	0.98	67,843	0.94	0.000

Note: H_0: $Sales_{it}$ is exogenous.

of demand uncertainty *DUnc* is not as straightforward and involves a few steps. In line with prior work (e.g., Eroglu and Hofer 2014, Rajagopalan 2013), using quarterly sales data, we estimate a regression model having an intercept, a trend term, and seasonality dummy variables for each retailer. We use the fitted model to generate predictions as sales $forecast_{iqt}$ (where q denotes quarter). Then, we approximate *DUnc* for the i_{th} retailer in year t as:

$$\sum_q |sales_{iqt} - sales\ forecast_{iqt}| / \sum_q sales\ forecast_{iqt}.$$

Our variable *DUnc* is conceptually similar to the demand uncertainty measure in Rumyantsev and Netessine (2007a), except we use a sum of absolute (instead of squared) errors and add the denominator such that *DUnc* is properly scaled and unit free. For the four growth-related controls in Equation (5)—*SalesGR*, *StoreGR*, *LaborGR*, and *CapitalGR*—we define each metric as the ratio of its value in the current year over its value in the previous year. Table 3 defines and describes each variable. Table 4 reports summary statistics and Pearson correlation coefficients.

5. Empirical Findings

5.1. Estimation Results

We substitute Equations (5) and (4) into Equation (3) to estimate the stochastic frontier model in one stage via the MLE method using *sf_fixeff* (Kumbhakar et al. 2015) and *ml max* in STATA 15.1 (StataCorp 2017). The reported results are based on the final estimation sample of 118 firms with 1292 observations (without missing data values). Table 5 presents our final models, which differ in their specification of fixed individual firm heterogeneity as well as operational inefficiency.

In addition to the Wang and Ho (2010) model (Models 3 and 4 W&H 2010), we estimate two popular and representative stochastic frontier panel data models—Battese and Coelli (1995) adopted by Lieberman and Dhawan (2005) (Model 1 B&C 1995) and true random effects (Model 2 TrueRE)—for the sake of comparing goodness-of-fit and results.[3] In addition to the estimated coefficients, we report the models' log-likelihood, Akaike Information Criterion (AIC), and AIC corrected for overfitting (AICc). Note the Wang and Ho (2010) model differs from the others in several aspects, including distributional assumption, scaling property, and variable transformation. Nonetheless, AIC and AICc enable us to compare different models in a theoretically robust fashion such that we can identify the best fit model for hypotheses testing. Also, for stochastic frontier models with highly nonlinear structures, the AIC and AICc are favored over pseudo R^2, which tends to have severe bias for evaluating nonlinear models (Spiess and Neumeyer 2010). That said, the Wang and Ho (2010) model with best fit (smallest AIC/AICc) shows high correlations (Pearson's $\rho > 0.88$) between Y and Y-hat (Wooldridge 2010), suggesting our model specification explains a fairly large fraction of variance in the production output. We also tested for collinearity among the regressors of the inefficiency models and found that the highest VIF was 5.74, well within the recommended range.

As expected, for each stochastic production function model (Models 1 through 4), we find labor and capital have positive and significant influences on retailer production outcomes (*GrossProfit*). For models 2 through 4, the coefficients for labor and capital are relatively stable across model specifications. Labor

Table 3 Description and Definition of Variables

Variable	Measurement	Description	Units
Y	*GrossProfit*	Net annual sales minus cost of goods sold	\$
L	*NumberOfEmployees*	Average number of employees during the year	Employee count
K	*FixedAssets*	Annual total assets minus total current assets	\$
U	*Inefficiency*	A non-negative random variable	Dimensionless
ε	*InvLean*	Negative of Studentized residual of linear regression of *InvAvg* on *Sales*, as defined by Eroglu and Hofer (2011)	Dimensionless
	Size	One-year lagged total employment	Employee count
	DUnc	Ratio of sum of absolute sales forecast error to sum of sales forecast	Dimensionless
	SalesGR	$Sales_{i,t}/Sales_{i,t-1}$	Dimensionless
	StoreGR	$Stores_{i,t}/Stores_{i,t-1}$	Dimensionless
	LaborGR	$Employees_{i,t}/Employees_{i,t-1}$	Dimensionless
	CapitalGR	$FixedAssets_{i,t}/FixedAssets_{i,t-1}$	Dimensionless
	InvAvg	Average total inventory during the year	\$
	Sales	Net annual sales	\$

Note: We deflate all variables measured in monetary terms into US dollars for the year 2000 using the consumer price deflator from the United States Bureau of Labor Statistics.

Table 4 Summary Statistics and Correlation Coefficients

Variables	Units	N	Mean	SD	Min	Max
L	10^3	1443	66.13	194.16	0.27	2200.00
K	$\$10^3$	1492	2891.70	9366.15	3.16	107,642.90
Y	$\$10^3$	1492	3442.15	9758.36	15.53	125,060.00
InvLean	NA	1425	0.00	1.00	−5.20	3.76
Size	10^3	1412	65.44	189.53	0.15	2200.00
DUnc	NA	1488	0.10	0.14	0.00	2.78
SalesGR	NA	1370	1.05	0.13	0.47	2.50
StoreGR	NA	1488	1.07	0.18	0.58	5.45
LaborGR	NA	1405	1.06	0.29	0.27	9.80
CapitalGR	NA	1370	1.09	0.50	0.05	13.48

Variables are in natural units (i.e., not log transformed).

	L	*K*	*Y*	*ELI*	*Size*	*DUnc*	*SalesGR*	*StoreGR*	*LabGR*
L	1.00								
K	0.97***	1.00							
Y	0.98***	0.98***	1.00						
InvLean	0.04	0.05*	0.05*	1.00					
Size	1.00***	0.97***	0.98***	0.04*	1.00				
DUnc	−0.07***	−0.06**	−0.07***	0.03	−0.07**	1.00			
SalesGR	−0.01	0.00	−0.02	−0.10***	−0.03	0.13***	1.00		
StoreGR	−0.01	−0.01	−0.02	−0.04	−0.03	0.18***	0.59***	1.00	
LaborGR	−0.01	−0.01	−0.02	−0.13***	−0.03	0.06**	0.64***	0.61***	1.00
CapitalGR	0.02	0.03	0.01	−0.05*	-0.01	0.17***	0.56***	0.53***	0.50***

Note: ***$p < 0.01$, **$p < 0.05$, *$p < 0.1$.

elasticity is significantly larger than capital elasticity with respect to gross profit. Without explicitly addressing individual firm heterogeneity, Model 1 based on the Battese and Coelli (1995) specification ($\alpha_i = \alpha$, $\forall i$) overestimates capital elasticity and underestimates labor elasticity. Model 2 based on TrueRE (that treats α_i as a random variable and estimates its mean), and Models 3 and 4, based on the Wang and Ho (2010) technique that cancels out α_i using a demean or first-difference prior to estimation have more stable estimates of the labor and capital elasticity parameters. The significance of the positive elasticity estimates further supports the implicit assumption that production inputs must be technical complements according to the Cobb–Douglas functional form. When performing a χ^2 test for H_0: $\gamma + \delta = 1$ on each model, we reject the null hypothesis that retailers in the sample have constant returns-to-scale technology ($p < 0.001$). We find evidence that retailers face decreasing returns-to-scale, consistent with prior work (Chuang et al. 2016, Perdikaki et al. 2012).

Table 5 also shows Model 3 has significantly lower AIC than Models 1 and 2. This suggests the Wang and Ho (2010) scaling model specification makes better use of available information and provides superior fit over the B&C 1995 and TrueRE models. As described above, to test *H1*, we incorporate linear and squared versions of *InvLean* to explain operational inefficiency in all of the models. To test *H2* and *H3* on the moderating associations of firm size and demand uncertainty, we introduce the interaction terms—*InvLean* × *Size*, *InvLean* × *DUnc*—into the operational inefficiency portion of Model 4.[4] Model 4 significantly outperforms Model 3 in terms of AIC and AICc, thus, we base our discussion on Model 4.

We remind the reader that the Equation (4) dependent variable is the retailer's *inefficiency* relative to the production frontier for firms in our sample. Thus, for the operational inefficiency models in Table 5, a positive coefficient represents a negative association with retailer operational efficiency. In order for the inverted U-shape between inventory leanness and operational efficiency in *H1* to be supported, the quadratic term (ω_2) should be positively significant, and, according to Lind and Mehlum (2010), the slope should be sufficiently steep at both ends of the data range, and the turning point needs to be located well within the data range (Haans et al. 2016). For our case, the coefficient of the quadratic term of *InvLean* is indeed positive and significant ($\omega_2 = 0.043$, $p = 0.03$). We determined the slope at the end of the data range, by finding the derivative of the inefficiency estimating equation with respect to *InvLean* and evaluating it at the extremes of the *InvLean* data range (i.e., −4 and 4), using the median value for *Size* and *DUnc*. The slopes at the two extremes have the correct sign and are significant ($s(-4) = -0.465$, $p = 0.06$ and $s(4) = 0.223$, $p = 0.04$, χ^2 test). While the slope at −4 is only marginally significant for the median values of *Size and DUnc*, it becomes significant when testing for the full

Table 5 Parameter Estimates of the Stochastic Frontier Model†

Variables	Model 1 B&C 1995	Model 2 TrueRE	Model 3 W&H 2010	Model 4 W&H 2010
Production function				
Constant α_i‡	0.163	0.326	—	—
	(4.304)	(1.595)		
Year β	0.001	0.002**	0.002*	0.002
	(0.002)	(0.001)	(0.001)	(0.001)
Labor γ	0.474***	0.731***	0.724***	0.750***
	(0.017)	(0.014)	(0.022)	(0.029)
Capital δ	0.426***	0.183***	0.196***	0.196***
	(0.013)	(0.010)	(0.012)	(0.012)
Operational inefficiency†				
InvLean ω_1	1.292***	0.509***	−0.033	0.234**
	(0.406)	(0.169)	(0.087)	(0.119)
InvLean^2 ω_2	−0.115	−0.006	0.067**	0.043**
	(0.0884)	(0.038)	(0.028)	(0.020)
Size ω_3	−0.526***	−0.180**	−0.086	0.100
	(0.180)	(0.076)	(0.098)	(0.107)
DUnc ω_4	4.590***	1.978***	0.922**	0.079
	(1.277)	(0.658)	(0.430)	(0.354)
SalesGR ω_5	−5.591***	−3.705***	−5.395***	−4.257***
	(1.792)	(1.109)	(0.948)	(1.130)
StoreGR ω_6	1.165**	0.669***	0.613***	0.810***
	(0.525)	(0.235)	(0.171)	(0.228)
LaborGR ω_7	1.840***	1.584***	1.602***	1.237***
	(0.663)	(0.412)	(0.329)	(0.305)
CapitalGR ω_8	0.356	0.113	0.154**	0.117*
	(0.280)	(0.131)	(0.070)	(0.063)
InvLean × *Size* ω_9	—	—	—	−0.138***
				(0.042)
InvLean × *DUnc* ω_{10}	—	—	—	0.547**
				(0.237)
Log-likelihood	−233.150	731.218	881.676	889.188
AIC	498.299	−1432.435	−1737.352	−1748.376
AICc	498.726	−1432.008	−1736.976	−1747.896
No. firms	118	118	118	118
No. observations	1292	1292	1292	1292

Notes: Standard errors are in parentheses; $^{***}p < 0.01$, $^{**}p < 0.05$, $^{*}p < 0.1$.
†Note that the dependent variable in Equation (4) is *Inefficiency*. For *InvLean*, a positive coefficient represents a negative association with operational efficiency, and vice versa.
‡Model 2 estimates the mean of α_i, whereas Models 3 and 4 have no α_i estimates due to demean/first-difference transformation.

range of the moderating variables, see §6.1. Finally, the unique turning point for the relationship (1.408, found by equating the slope of the estimated response to zero) is well within the data range of inventory leanness. Thus, we find no evidence to reject the hypothesized inverted U-shape relationship between *InvLean* and *operational efficiency* (*H1*). We further explore the shape of the relationships and the impact of the moderating factors in §6.1.

The first-order effect of the moderator *Size* is not significant. Nonetheless, the negative and significant *InvLean* × *Size* ($\omega_9 = -0.138$, $p < 0.01$) does shift the turning point of the inverted U-shaped relationship toward greater *InvLean*,[5] indicating that large size retailers benefit more from the increases in inventory leanness. Similarly, the first-order effect of the moderator *DUnc* is not significant, but the positive and significant *InvLean* × *DUnc* ($\omega_{10} = 0.547$, $p = 0.02$) shifts the turning point of the inverted U-shape relationship towards lower *InvLean*, suggesting retailers who pursue inventory leanness under high demand uncertainty are likely to achieve lower operational efficiency. That is, retailers under higher demand uncertainty are associated with deviations further away from the performance frontier. As hypothesized (*H2* and *H3*), *Size* and *DUnc* are found respectively to enhance and weaken the association between inventory leanness and operational efficiency. We explore the combined effect of these two factors on the relationship between inventory leanness and operational efficiency in §6.1.

As for the growth-related controls, *SalesGR* ($\omega_5 = -4.257$, $p < 0.01$) is positively associated with operational efficiency, whereas *StoreGR* ($\omega_6 = 0.810$, $p < 0.01$), *LaborGR* ($\omega_7 = 1.237$, $p < 0.01$), and *CapitalGR* ($\omega_8 = 0.117$, $p = 0.06$) are negatively associated with operational efficiency. The findings are similar to Shockley et al. (2015), who use ROA and ROS as operational performance measures. The positive parameter of *SalesGR* is reasonable, since sales revenue growth contributes to the production output and should come with more resources for retailers to improve operational capabilities and technological systems. On the other hand, *StoreGR*, *LaborGR*, and *CapitalGR* are coupled with input expansion, and rapid expansion usually comes with consequences of undermining operational efficiency (Hayes and Clark 1986, Oliva et al. 2003).

5.2. Robustness Checks

To ensure the findings did not simply arise from functional forms or distributional assumptions, we conducted five robustness checks—full results for all robustness checks are available in the electronic supplement (Appendix S1) of this study. First, as a validation of the inverted U-shape, we ran spline-type regressions in the context of SFA by creating knots of inventory leanness. After introducing one and two knots into the inefficiency effects Equation (5), we re-estimated the model to check if operational efficiency first increases and then drops as inventory leanness. One knot splits inventory leanness into the lower 50% and the higher 50%, whereas two knots split inventory leanness into three groups of the lower 25%, the interquartile, and the upper 25%. In the one knot case, the *InvLean* coefficient of the first spline is significantly negative ($p < 0.01$) and the coefficient of the second is significantly positive ($p < 0.01$), supporting a U-shape between inventory leanness and *inefficiency* (i.e., an inverted U-shape between inventory leanness and operational efficiency).

The results are similar when we consider the two knots scenario.

Second, we replaced the single time parameter β with 13 dummy variables for years 2000 through 2013 (one dummy dropped due to collinearity) and re-estimated Model 4 from Table 5. The model with eleven extra parameters resulted in AIC (−1747.089) and AICc (−1745.803) values that suggest a poorer use of information than the more succinct model. Thus, the commonly adopted specification with single time parameter β for technology change and economic volatility is more parsimonious and preferred.

Third, despite its popularity, the Cobb–Douglas production function in Equation (2) is sometimes accused of being simplistic. To ensure our results are not an artifact of this assumed functional form, we instead estimated the Translog production function often used by economists (Coelli et al. 2005, Kumbhakar et al. 2015):

$$\begin{aligned}\ln(Y)_{it} &= \alpha_i + \beta\tau + \gamma \ln(L)_{it} + \delta \ln(K)_{it} \\ &\quad + \eta\frac{1}{2}\ln(L)_{it}{}^2 + \zeta\frac{1}{2}\ln(K)_{it}{}^2 \\ &\quad + \psi \ln(L)_{it}\ln(K)_{it} - U_{it} + V_{it}\end{aligned}$$

The resulting estimates for ζ and ψ were not statistically different from zero, suggesting over-parameterization in the Translog model. Thus, the Cobb–Douglas production function is a reasonable and valid choice in this case.

Fourth, to ensure that the moderate correlations between the growth-related controls do not compromise the model fit, we dropped the least significant *CapGR* and re-estimated Model 4, leading to the qualitatively same results. Retaining *SalesGR*, the most significant of the growth controls, and testing with permutations of one or two additional controls resulted in models with inferior fit and increases of up to 107 in AIC. Models with permutations without *SalesGR* had AIC values increasing by more than 350, indicating poor model fitting performance. The results suggest the full Model 4 with all controls indeed minimizes information loss relative to models with fewer growth factors included.

Lastly, to assess whether our findings are sensitive to distributional assumptions, we replaced the half normal random variable in Equation (4) with a truncated normal random variable (Kumbhakar et al. 2015). The significance and sign of model coefficient estimates derived from an alternative inefficiency distribution are similar to results shown in Table 5. Neither AIC nor AICc shows significant difference attributed to alternative distribution assumptions. *H2* and *H3* are consistently supported. Taken together, the robustness analyses provide strong support for the estimated model.

6. Post-Hoc Analysis

In this section, we explore the implications of our findings. We first analyze the combined impact of the moderating factors on the relationship between *InvLean* and operational efficiency. We then explore the differences in operational efficiency scores across the six retailer categories in our sample. The section concludes by reporting the findings from a set of post-hoc interviews with retail executives with inventory responsibilities.

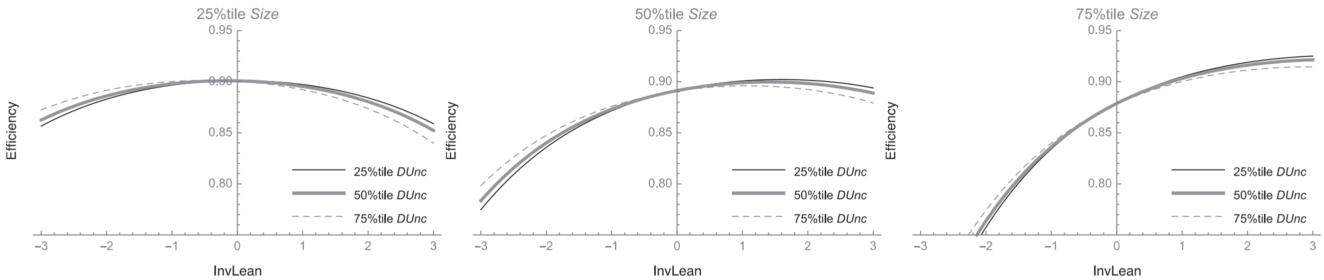
6.1. Combined Effect of Moderators

As explained in §5.1, *InvLean* is present in four terms with significant coefficients in Model 4—linear, quadratic, and interaction terms with *Size* and *DUnc* —making the interpretation of its association with operational efficiency rather complex. Moreover, the significant moderating effects of firm size (*H2*) and demand uncertainty (*H3*) force us to consider several issues: To what extent can *Size* and *DUnc* weaken or strengthen the overall association between *InvLean* and operational efficiency? What are the relative magnitudes and combined effects of *Size* and *DUnc* on the relationship?

To explore these issues, we conduct a post-hoc response analysis of the link between inventory leanness and operational efficiency after performing the transformations from inefficiency estimates to the hypothesized *operational efficiency*. Based on the parameter estimates of Model 4, we estimate the response of the operational inefficiency (U) to *InvLean* (from −3 to 3) using the extended inefficiency model, under the 25th, 50th, and 75th percentiles of *Size* and *DUnc* in our sample. The four growth-related controls are set to the sample medians. After computing the inefficiency U, we estimate operational efficiency using $\exp(-U)$ based on Equation (1), bounding the efficiency estimate to the (0, 1] range (Coelli et al. 2005). Figure 3 shows the estimated operational efficiency response to *InvLean*.

The formal results described in §5.1 are confirmed in Figure 3. The central panel of Figure 3 shows the hypothesized inverted U-shape for the median-sized retailer, indicating that there is an optimal level of inventory leanness beyond which operational efficiency begins to degrade. Next, the turning point of the inverted U-shape of the response curve shifts to the left as *DUnc* increases—the maximum of the dashed line (75 percentile) in the central panel is at *InvLean* = 1.08; this relationship holds across all firm sizes, that is across the three panels. Finally, the moderating effect of firm *Size* can be assessed by comparing across the three panels. The left panel shows the response curve for small retailers (25 percentile). For all levels of demand uncertainty, the turning point of

Figure 3 Operational Efficiency Response to *InvLean* by *Size* and *DUnc*



the inverted U-shape has been shifted to the left (*InvLean* = −0.15), such that the response curve for small retailers exhibits efficiency degradation when increasing inventory leanness beyond the neutral point. The right panel shows the response for large retailers (75 percentile). In this plot, the turning point of the inverted U-shape is to the right of the *InvLean* sample range (*InvLean* = 3.27), such that for large retailers the response curve shows operational efficiency gains as inventory leanness increases for most of the sample range.

Two additional observations can be made from inspection of Figure 3. First, higher demand uncertainty not only shifts the optimal point to the left, but it also tilts the response curve as retailers with relatively more inventory (i.e., *InvLean* < 0) have higher operational efficiency. Reducing inventory beyond the neutral inventory leanness hurts performance under demand uncertainty. Note that, for all firm sizes, retailers with small demand uncertainty (i.e., the 25 percentile of *DUnc*; thin solid line), have improvements and degradations in the expected directions, but the response curve does not differ much from the median case, as would be suggested by the fact that the direct effect of *DUnc* is not significant ($\omega_4 = 0.079$, $p = 0.82$). Second, despite the fact that the direct effect of *Size* on *operational efficiency* was found to be non-significant ($\omega_3 = 0.10$, $p = 0.35$), the maximum expected operational efficiency increases with *Size* and the responsiveness to inventory leanness also increases with *Size*; the response curve for large retailers shows significantly lower efficiency with large inventories (*InvLean* < 0).

By assessing the range of responses and the relative strength of the moderating factors, our analysis not only sheds light on how the link alters the turning point of the inverted U-shape (i.e., optimal leanness) depending on firm size and demand uncertainty, but also can serve as a qualitative guideline for retailers as they can assess their relative size and demand uncertainty prior to undertaking lean initiatives.

6.2. Differences across Retail Categories

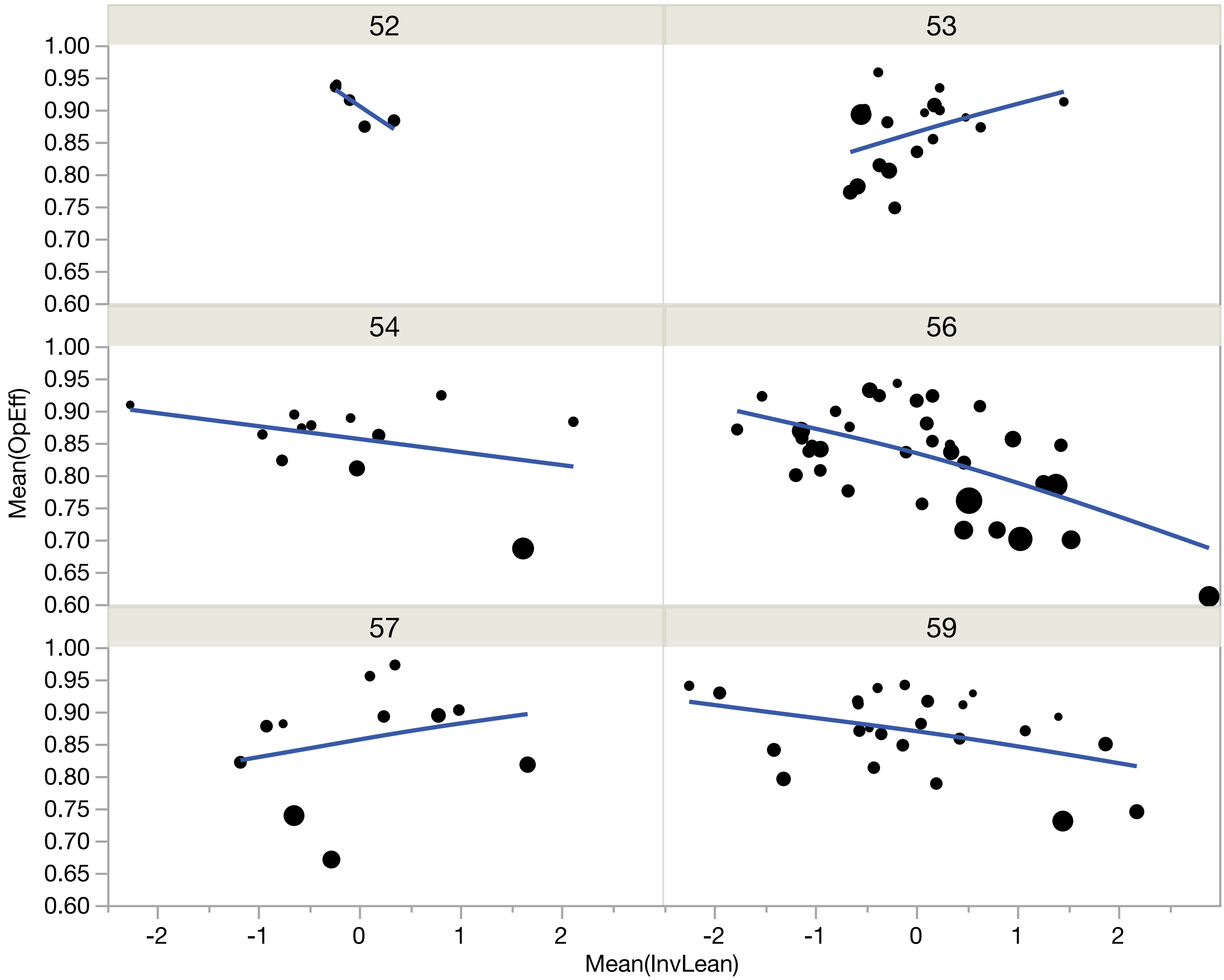
Although it is not realistic to use an industry-wide model for operational/managerial decision support, we believe that an industry model might be useful for managerial understanding of industry status quo and to explore strategic implications across the industry. In this sub-section, we make use of the estimated *InvLean* and operational efficiency (*OpEff*) scores across the six retail categories in our sample (see Table 1) to explore how different categories, facing different product and demand characteristics, differ in their operational strategies.

To aggregate firms' performance, we calculated each firm's average and standard error of their annual *InvLean* and *OpEff* scores. We dropped from the sample those firms for which we had less than five annual scores, thus reducing the sample to 105 firms. Figure 4 shows six scatter plots (one per retail category) of each firm's average *OpEff* vs. *InvLean*. The size of each dot is proportional to the firm's *OpEff* standard error.

Note that since the scatter plot does not consider size, demand uncertainty, or the control variables included in the regression, the inverted U-shape found in the estimation is not evident from these scatter plots. The grouping of firms by retail category, however, allows us to derive some industry sector-specific insights. First, operational efficiency is not equal across retail categories. Indeed, considering all annual readings, the average *OpEff* for category 52 (building materials), 91%, is significantly higher than all the other categories ($p < 0.001$) and that of category 56 (apparel and accessories), 83%, is significantly lower ($p < 0.001$); the other four categories are not statistically different from each other. In terms of dispersion, the variance of the *OpEff* scores of categories 56 (apparel and accessories) and 57 (home furniture) are significantly higher than the variance in the other categories (Welch's test $p < 0.001$). On the other hand, as is expected from an inventory metric calculated based on relative performance, the *InvLean* scores across the different retail categories are not statistically different in their means nor variances.

While tempting to explore potential explanations for variations across sectors, this is not feasible as *operational inefficiency* is an unobserved random variable (by construction a function of input factors in Equation (3) and efficiency drivers in Equation (5)) as

Figure 4 Scatter Plots of Firms' Mean *InvLean* and *OpEff* by Retail Category [Color figure can be viewed at wileyonlinelibrary.com]



Note: Retail category is defined by the two-digit SIC code of the firm. SIC codes are listed on top of each scatter plot.

opposed to a metric that can be estimated a priori and directly used as a dependent variable in regressions. Nevertheless, the above-reported differences improve the credibility and usability of our findings. Consider, for example, the two categories at the extremes of operational efficiency listed above (i.e., SICs 52 and 56). A quick analysis reveals that these two categories are also at the extremes of the allocated mix of production factors—as measured by labor intensity (i.e., the ratio of labor to fixed assets). Specifically, the apparel and accessories category's average labor intensity is 38% higher than the closest category ($p < 0.001$)—home furniture—and 50% higher than the sample mean. The high labor intensity of the category is expected for a purchasing experience where staff plays an important role (see, e.g., Chuang et al. 2016, for an estimation of the staffing effect on the sales function of an apparel retailer). Conversely, the labor intensity of the building materials category, while not statistically different from the other four categories, is the lowest in the full sample, as the vast majority of building material sales are in large quantities and to knowledgeable contractors (Duois and Gadde 2000).

Considering the two sectors with the lowest and most volatile efficiency scores—apparel and accessories (56) and home furniture (57)—we note that these sectors have the smallest size (60% below the sample mean) and the largest demand uncertainty (15% above the sample mean). Thus, while the two segments are not different from other sectors in average *InvLean* scores, the two focal sectors probably are *least benefited from positive size moderations* and *most compromised by negative uncertainty moderations* in the inventory leanness and operational efficiency relationship. While such a conjecture is somewhat ad-hoc, it also resonates with our theorizing efforts (*H2* and *H3*). Although informal, the above analyses confirm that the *OpEff* scores derived using our estimation equation remain realistic and yield insights regarding strategic choices across retail category. Assessment of specific strategies, however, is beyond the scope of this study.

6.3. Post shoc Interviews with Retail Executives

To assess the relevance and potential impact of our study results, we followed up our empirical analysis with a series of structured post hoc interviews of real-world supply chain executives. In doing so, we attempted to identify firms at both ends of the inverted-U relationship (i.e., low inventory leanness, high inventory leanness) as well as in the middle of this relationship (i.e., middle leanness). We developed a questionnaire protocol based on the findings of our study to further delve into the relevance of the constructs we empirically evaluated via our econometric stochastic frontier analysis models. We also wanted to further explore what dimensions or decisions we might have missed through focusing on this firm-level data set analysis. Given our line of inquiry, interviewees were not required to stay within the standard questions. In line with the process of grounded theory development (Glaser and Strauss 1967), this semi-structured protocol changed over time as each subsequent interview was used to triangulate the responses from previous interviews and expanded the list of questions as we uncovered more related topics.

Our sample included a variety of executives from large and small retailing corporations across different sectors (e.g., jewelry, grocery, rent-to-own, fashion/apparel). The individuals included top-level managers (e.g. Chief Supply Chain Officer, VP/Director of supply chain planning, SVP of merchandising) directly focused on determining inventory levels for their firms, whether from a procurement and warehouse management perspective or from a retail merchandising perspective. While our sampling was discriminate (i.e., we were purposefully looking for variance among the respondent firms), we sampled until we reached saturation.

In asking questions pertaining to this study, we attempted to divine certain dimensions that managers find meaningful when thinking about how their firms (and competing firms) go about evaluating inventory management decisions and deciding upon related inventory management tactics. To the best of our ability, we tried not to proactively ask about the moderators discussed in our study—only after the discussants brought up those dimensions (i.e., firm size, sources of demand uncertainty) would we probe further into those dimensions. After starting with initial warm-up questions about the managers' background and employment history, we asked questions about the managers' firm and the products sold by that firm. Next, we asked how the firm's managers approached setting retail inventory levels, who makes such decisions, the regularity (i.e., time bucket) of such decisions, what information is considered for these decisions, and the extent to which the firm's managers try to focus attention on balancing between too little or too much inventory. We also probed for exogenous and endogenous drivers that managers considered to be salient factors affecting how they set their retail inventory levels.

Overall, we find that both retail merchandising and supply chain managers do think very carefully about the balance between having too much vs. too little inventory on hand. With regard to the managerial salience of inventory leanness, we generally found that managers carefully think about how to achieve lean inventory positions. For example, a fashion apparel retail manager suggested that inventory levels/leanness are both corporate high-level decisions as well as bottom-up decisions, linked directly to the financial planning process. A jewelry retailer decides inventory centrally at brand level, making plans strategically with top leadership, so that tactical use of forward logistics and reverse logistics can move and reposition inventories as needed by region and store. A manager from a large grocery retailer suggested his firm's employees "manage the inventory from both ends" (i.e., headquarters and store), and that managers at both ends are held accountable for inventory stocks. This firm's focus was on protecting against too much risky, high-value, or high-shrink inventory. From a growth perspective, holding excessive inventory was thought of in terms of an opportunity cost, in terms of how many fewer stores that retailer might open in the future. Trading off against numbers of future stores, that manager also was very cognizant of the role of inventory in terms of generating customer disappointment if the inventory was not there when needed. This grocery manager does not think of leanness in terms of a focus on inventory reduction, but rather on "correction of inventory"—the firm only should tie dollars up in the right type of product, and with 170,000 different SKUs across its many stores, correction of inventory drives better inventory utilization and better selling potential of all products. To accomplish this aim, local managers need to be motivated, inspired, and held accountable for making good inventory decisions.

While merchandiser concern is largely to satisfy store needs, in order to make sure that the store experience is exceptional, we found a variety of ways in which internal or external factors might limit the ability to control this balance. A manager from a rent-to-own retailer discussed how a CFO managing cash efficiently, and thus forcing a reduced inventory, might risk having a terrible experience at a store. A manager brought up how suppliers may require a firm to buy a minimum number of truckloads of goods, and "won't budge," forcing the firm to carry high inventory on that item. Banks can also play a role in the inventory balance: "Do the banks want us to have inventory, or do the banks want us to go out of business?"

Technology was a consistent theme of these managers, but not always viewed as the solution. One manager suggested a need for "Having the technology. Understanding the technology. Working the technology." That is, it is not sufficient to simply install the technology—employees and inventory managers need to understand how to work the technology to manage the appropriate time bin of retailing activities. Technology was seen as a facilitator of both centralized and decentralized inventory decisions wherein local managers were better able to match inventories to the local population's needs. But technology sophistication has only recently realized the vision of inventory leanness: "Ten years ago, we couldn't do what we do today." A retail jewelry manager suggested that lean inventory management requires a sophisticated technology capability, enabling a very complex process involving individual SKU-level picks and fulfillment. With technology changing very fast today, continuing to become more predictive of individual consumer needs, the impact of this trend upon inventory levels and leanness is certain to continue, suggesting a major research opportunity to evaluated impacts of technological initiatives carried out at major retailers.

With respect to our moderator variable for demand uncertainty, the managers generally always discussed a variety of ways in which forms of demand uncertainty entered into their inventory decisions. At a rent-to-own retailer, uncertainty had affected product-line launching, with intense unanticipated demand signals leading to orders of huge inventories, but not in a manner timely enough to take advantage of the selling opportunity. This manager characterized a certainty arising from a philosophy of "We Will Not Lose A Sale," combined with high demand uncertainty, to be guaranteed to deliver pain.

With respect to a retailer's firm size, again managers often independently brought up how larger size firms had certain advantages when undertaking a lean inventory tactic. We observed comments such as: "Volume does help," "It's huge," and "Impact of scale of the operation is enormous." The benefits of retailer size arise from purchasing power, leverage over suppliers, and other factors. Size often was used to refer to a firm's systems—with larger size, a firm generally can build sophistication into its logistics systems. One manager said that what his organization does would be very costly to do if they only had a couple stores; with scale, it makes sense to have complex inventory management processes. Another manager gave an example of how Walmart can cross-dock very efficiently, whereas smaller firms are unable to do so.

We asked the managers, outside of firm size and demand uncertainty, what might be considered in the link between their inventory decision and associated performance impact. Their comments included a myriad of potential factors that researchers might examine in future studies. Among the most frequent responses were:

- Firms that are joint manufacturer/retailer/wholesaler involve many more channels, which makes the inventory decision incredibly more complicated.
- Speed vs. Customization. While customers are okay with some custom products being slow, they increasingly are wanting custom products much faster.
- Changing location stocking strategies are affecting how managers stock inventory.
- Sophistication of systems—The more sophisticated you can get with your systems, the more aggressive you can be with inventory leanness. If you believe in your planning tools, you are more willing to act on them. In contrast, if the decision models are known to be less than perfect, then you start adding buffers.
- Competitive threats/defending turf from outside threats

Taken together, the interviewees report a deliberate calculus of balancing multiple tradeoffs between too much and too little inventory, thus supporting the empirically-found inverted U-shape relationship between leanness and efficiency. Given the substantial impact of inventory on operational performance, management seems to deploy an important amount of effort and coordination across organizational levels to find the right level of inventory. While all interviewees mentioned unprompted "demand uncertainty" as an important consideration when making the inventory decisions, scale as a moderator of the relationship between inventory and performance was only mentioned by three quarters of the interviewees. All interviewees, however, mentioned organizational capabilities (e.g., information technology, logistic infrastructure) as direct enablers of leaner inventories without affecting performance. While our data did not give us access to detailed retailer capabilities, these capabilities are, as we argue in the hypotheses section above, related to firm size. Indeed, it seems like some of the managers are looking for and studying models of other retailers that have successfully journeyed toward being a leaner retailer. Technology-enabled and finance-driven (e.g., Wall Street valuation) movement toward inventory leanness seems inevitable to retailers. Hence, they are interested in knowing, depending on contextual factors (including but not limited to firm size and demand uncertainty), what is the optimal level or industry position of inventory leanness. In sum, the post-hoc interviews consolidate the grounding of our econometric investigation, establish relevance of the proposed moderators, and point to more nuanced moderators of the inventory-performance link than those available in our data sample. We discuss implications for managers and researchers in the following section.

7. Discussion and Conclusions

In this article, we contribute to retailing industry research by using a stochastic production function approach to examine potential associations between lean inventory and retailer operational efficiency. Microeconomic theory provides a basis for using stochastic frontier analysis to estimate operational efficiency (Coelli et al. 2005) and to provide insights regarding operational confounders of lean inventory positions. The inventory leanness metric in our model avoids the potential pitfalls of inventory turns, which simply compares a firm's inventories to its sales in isolation of its competition and can be misleading due to its exclusion of scale economies and inventory-sales relationships for an industry segment (Eroglu et al. 2015).

Our analysis uses panel data on US public retailers to estimate in a single stage the production and inefficiency models. We find that inventory leanness exhibits a complex relationship to retailer operational efficiency, one that is heavily moderated by firm size and demand uncertainty. The inverted-U associations observed in some manufacturing industries carry over to a majority of retailers within the practical range of inventory leanness. However, in contrast to extant findings in manufacturing contexts that suggest being lean positively influences operational performance (Fullerton et al. 2014, Shah and Ward 2003, Womack et al. 1991, Yang et al. 2011), we find that under high demand uncertainty, not being lean is associated with higher operational efficiency, regardless of firm size. This contrast is sensible as manufacturing production plants are intrinsically different from retail stores, in which realized shopper demand could be inventory-dependent.

Our findings that large retailers see efficiency gains as inventory leanness increases might be explained by the additional operational resources available to large retailers. For example, Walmart has a higher proportion of part-time labor force, which requires only minimum wages and less insurance coverage (Christopherson 2007). As such, Walmart may be more flexible in allocating its labor force to fit lean practices. Interestingly, our finding seems to confirm management anecdotes on a tale of two inventory scenarios—"The big retailers are lean because they've learned how to do it. By contrast, smaller companies may appear lean due to management's cutting back on orders..." (Solomon 2013). For medium and particularly small sized retailers, the empirical evidence raises cautions against the practice of "de-stocking" (i.e., proactively and haphazardly cutting inventory levels) in the US retail sector (Cassidy 2015). Moreover, large firms benefit from economies of scale in inventory management since they may pool demands from many store locations, resulting in relatively lower inventory-ordering and -holding costs and higher efficiencies (Chen et al. 2007, Gaur and Kesavan 2009, Rumyantsev and Netessine 2007b). In a nutshell, our findings for retailer size/scale and the decreasing returns-to-scale ($\gamma + \delta < 1$) jointly verify the theoretical argument that large firms in a non-perfectly competitive market could have scale inefficiency in production, while benefitting from economies of scale (Gelles and Mitchell 1996).

Demand uncertainty, which reflects environmental dynamism and market instability (Azadegan et al. 2013), is also found to significantly moderate the association between inventory leanness and operational efficiency. Demand uncertainty is usually found to be positively associated with aggregate inventory levels (Rajagopalan 2013, Rumyantsev and Netessine 2007a) or negatively associated with inventory turns (Hancerliogullari et al. 2016) as retailers use inventories to buffer against demand uncertainty. In practice, the rapid growth of product variety in retailing sectors over the past few decades has forced store-based retailers to carry more inventories. The increased product variety complicates consumer choices and amplifies demand uncertainty, increasing the probability of stock-outs. Consequently, some retailers (e.g., Target) have decided to cull products in stores (variety reduction) to resolve out-of-stock issues and pursue inventory leanness (Ziobro 2016a). While high inventory leanness seems to imply anecdotally that retailers make good inventory decisions, our finding—that as demand uncertainty increases retailers with relatively less inventory (i.e., *InvLean* > 0) are associated with lower operational efficiency—suggests that those seemingly lean retailers may actually need additional slack (e.g., spare inventory or safety stock) in order to respond to large demand fluctuations, grasp lost sales opportunities, and improve operational efficiency. Taken together, the moderating associations of firm size and demand uncertainty discussed above appear to be useful in guiding managers' intuition with respect to pursuing inventory leanness.

As a complement to our econometric modeling effort, our post-hoc interviews with retail executives provide practical support for the inverted U-shape and the proposed moderators. Combined with theory development and model estimation, the field work enables us to instill theoretical, statistical, and practical underpinnings in our empirical analysis. The interviews also reveal that executives across retail sectors all acknowledge the importance of inventory decisions and consider inventory management to be increasingly complex in spite of technology advancements. Managers' responses to our inquiries show that the constraints, contingencies, and consequences of lean inventory control are of high interest to retail practitioners. The extra factors coupled with inventory decisions but not included in our analysis are definitely worth exploring.

As with all empirical studies, our study exhibits potential limitations. First, we adopt a classical two-input (i.e., labor, capital) productivity model specification without information on input prices. We address this by using gross profit, as opposed to sales revenue, as the production output. Moreover, our model specification isolates firm-specific fixed effects from operational inefficiency associations. Second, our list of drivers of retailer efficiency is by no means exhaustive. Yet we note that adding more efficiency drivers may cause convergence problems in MLE due to the highly non-linear structure of stochastic frontier models. Nonetheless, our models cover theoretically relevant efficiency drivers and achieve reasonable fit. Third, we could not perform endogeneity tests pertaining to reverse causality due to the intrinsic limitations of the methodology. In the SFA literature, it was not until recently that researchers reported a formal investigation into endogeneity (Amsler et al. 2016). To our knowledge, the investigation is exclusively for cross-sectional stochastic frontier models and such development for panel data stochastic frontier models is scant. Nonetheless, compared to OLS regression, our SFA approach with fixed effects is an improvement in that the unobservable inefficiency component and the effects of its covariates are empirically identified and decoupled from random noise that affects retailer performance. This error term decomposition and the one-stage estimation to some extent alleviate threats of endogeneity. Moreover, both inventory leanness and operational efficiency are aggregate and *relative* metrics with post-hoc natures. Thus, it is extremely difficult to identify valid instruments for inventory leanness. Such scarcity of good instruments is common in operations management research and using invalid instruments do more harm than good to estimation efficiency (Lu et al. 2018). Fourth, unlike prior studies covering longer time periods from the 1980s (e.g., Chen et al. 2005, Johnston 2014), our study focuses on the period of 2000–2013. This deliberate choice of post-dot-com boom time periods allows us to better control and observe the impact of most lean inventory initiatives since the early 2000s (Abernathy et al. 2000). Last, we recognize that the rapidly growing e-retailing should affect the dynamics of inventory and inventory productivity. However, the online fraction of a retailer's sales is not available to be used as a control variable. Having said that, we reduce the potential effect of the online retail channel in our analysis by eliminating retailers with most sales from online sources, merger/acquisition periods, bankruptcy years, and any suspicious values (e.g., zero inventory holdings). The fact that our model explains almost 80% of the observed variance suggests that the potential impact of this control variable is limited.

Despite potential limitations, our analysis articulates potential impacts of retailers being leaner and carries important managerial implications. Based on our findings, managers need to realize that the benefits of being lean in retailing are not as widely attainable as in manufacturing, where it is more simple and straightforward to visually detect excessive inventory or starved production processes (e.g., idle equipment). In the retail industry, it is more difficult to link specific inventory levels to SKU- or basket-level sales, and to observe consequences of aggressive inventory reduction practices. Based on the findings, we learn that instead of uniformly enhancing retailer operational performance, lean-motivated inventory reductions may potentially backfire on small and medium retailers at certain points.

So, what do lean principles mean to retailers? The lean vision of streamlining processes, reducing waste, and increasing productivity is still valuable and applicable to retail operations, especially where efficient in-bound and out-bound logistics processes are of paramount importance. However, managers ought to avoid taking a myopic view of lean retailing, as reduction in inventory holdings is not synonymous with improvement of operational efficiency. Taken together, the moderating associations of firm size and demand uncertainty discussed above appear to be useful in guiding managers' intuition with respect to pursuing inventory leanness. More formal investigations into the effect of environmental and managerial factors (e.g., innovation, competition) on operational performance response to inventory leanness would allow practitioners to see a bigger picture of how retailers act and perform under various contingencies.

Our work raises interesting questions that motivate additional research. First, even though Eroglu and Hofer (2011, 2014) provide detailed elaboration on why ELI is superior to other metrics of inventory productivity (e.g., inventory turnover, inventory-to-sales ratio), ELI does not explicitly differentiate between raw materials, work-in-process, and finished goods in manufacturers' inventory holdings. Arguably the three types of inventories do not contribute to sales generation equally. For retailers who typically hold only finished goods, we improve the efficiency of ELI estimation by controlling for SIC-level differences in a pooled 2SLS regression, which also addresses potential endogeneity between inventory and sales. How to further tackle endogeneity (e.g., better instruments, structural equations) and accommodate inventory type differences when estimating ELI or its variants is a promising area for research.

Second, subsequent studies ought to analyze antecedents of retailer inventory leanness. After analyzing a cross-sectional sample of manufacturers, Hofer et al. (2012) find a negative association between internal lean practices and observed inventory leanness, questioning whether the finding is merely a statistical artifact of the data. It will be important for retail managers to understand the causal effects of internal lean practices on observed inventory leanness as well. Also, formal tests on reverse causality for SFA approaches are still in their infancy. We are thus cautious about interpreting the observed associations. Examining the causal effects of inventory leanness on operational efficiency requires more sophisticated research design and advancement in econometric methods for frontier estimation.

Third, even though we find inventory leanness is associated with lower operational efficiency for small retailers in our sample, it would be interesting to see whether some of them are going through a "worse before better" dynamic (Forrester 1961) as they push for ever leaner inventories. For instance, during our post hoc interviews, one manager suggested we look at the recent case study of Kohls, which that manager claimed to presently be a Wall Street darling due to its right-sizing of inventory a few years ago. By removing up to 8% of its inventory, Kohls took an immediate hit to its top line sales, which of course was looked upon unfavorably by the market. Yet 3 years later, the strategy is starting to pay off in increased top line revenues. Potentially, by taking a hit for a few years to reset the firm's baseline inventory, the Kohls case study provides an indicator of how much top line revenue is at risk, and for how long it is at risk, when managers try to lean out inventory. Also, technological innovation may help small retailers better match supply with demand while fulfilling the lean inventory prospect of higher operational efficiency, but perhaps only after a painful period of technology installation and adaptation. Researchers might re-examine findings of associations between firm performance (e.g., ROA, ROS, and other financials) and inventory turnover/inventory leanness while considering technology adoption.

Lastly, in an econometric analysis of inventory performance of Chinese companies, Shan and Zhu (2013) find the relationship between inventories and holding costs in Chinese public companies is quite different from that in US public companies. An interesting extension of our work might examine whether the inverted-U association between inventory leanness and operational efficiency, contingent on firm size and demand uncertainty, holds for other nations' retailers that face different labor and capital costs and where shoppers exhibit potentially different purchase dependencies on inventory levels.

Acknowledgments

We thank Yating Feng for help in data collection and analysis in earlier versions of this work. We also thank Edward Anderson, the Senior Editor, and four referees for their helpful comments and suggestions, which greatly helped improve this study.

Notes

[1]For instance, in our sample, the SIC sector 54 has an average of 13 inventory turns per year, whereas other SIC sectors have averages of ~4 turns per year. A retailer in SIC 54 with an average of 10 inventory turns per year is not lean within its sector but would be considered lean within the regression sample. Adopting the ELI that considers within-sector effects avoids the misperception described above.

[2]Number of employees and sales are the two most popular measures of *Size*. In performance and efficiency analysis research, sales is used for measuring production outputs and thus inappropriate for efficiency inputs. Researchers have a tradition of using employment to approximate size in industrial economics, as it captures coordination costs, financial constraints, firm age, and even legal systems that are tied with firm size (e.g., Kumar et al. 2001, Laeven and Woodruff 2007).

[3]For completeness, we also attempted to fit the true fixed effects (TFE) model with 117 firm dummies. The TFE model, however, failed to converge due to incidental parameters.

[4]Note that we only hypothesize shifts in the turning points and hence the quadratic interaction terms—$InvLean^2 \times Size$ and $InvLean^2 \times DUnc$—are excluded from the model on theoretical grounds (Haans et al. 2016). When included in the model, both $InvLean^2 \times Size$ and $InvLean^2 \times DUnc$ are not statistically significant at the 0.05 level, suggesting

no significant flattening/steepening of the inverted U-shape, nor do they significantly improve model fit.

[5]As above, we found the turning point by equating the first derivative of the inefficiency estimating equation with respect to *InvLean* to zero. The turning point is given by a function of $\omega_9 \times Size$ and $\omega_{10} \times Dunc$. That is, the sign of the coefficients ω_9 and ω_{10} determine the effect of the moderating factor on the turning point.

References

Abernathy, F. H., J. T. Dunlop, J. H. Hammond, D. Weil. 1999. *A Stitch in Time: Lean Retailing and the Transformation of Manufacturing-Lessons from the Apparel and Textile Industries*. Oxford University Press, New York.

Abernathy, F. H., J. T. Dunlop, J. H. Hammond, D. Weil. 2000. Control your inventory in a world of lean retailing. *Harv. Bus. Rev.* **78**(6): 169–176.

Aigner, D., C. K. Lovell, P. Schmidt. 1977. Formulation and estimation of stochastic frontier production function models. *J. Econom.* **6**(1): 21–37.

Alvarez, A., C. Amsler, L. Orea, P. Schmidt. 2006. Interpreting and testing the scaling property in models where inefficiency depends on firm characteristics. *J. Prod. Anal.* **25**(3): 201–212.

Amato, L. H., C. H. Amato. 2004. Firm size, strategic advantage, and profit rates in US retailing. *J. Retail. Consum. Serv.* **11**(3): 181–193.

Amato, L. H., C. H. Amato. 2012. Retail philanthropy: Firm size, industry, and business cycle. *J. Bus. Ethics* **107**(4): 435–448.

Amsler, C., A. Prokhorov, P. Schmidt. 2016. Endogeneity in stochastic frontier models. *J. Econom.* **190**(2): 280–288.

Angelini, P., A. Generale. 2008. On the evolution of firm size distributions. *Am. Econ. Rev.* **98**(1): 426–438.

Ansoff, H. I. 1957. Strategies for diversification. *Harv. Bus. Rev.* **35** (5): 113–124.

Azadegan, A., P. C. Patel, A. Zangoueinezhad, K. Linderman. 2013. The effect of environmental complexity and environmental dynamism on lean practices. *J. Oper. Manag.* **31**(4): 193–212.

Baker, R. C., T. L. Urban. 1988. A deterministic inventory system with an inventory-level dependent demand rate. *J. Oper. Res. Soc.* **39**(9): 823–831.

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

Ballou, R. H. 2005. Expressing inventory control policy in the turnover curve. *J. Bus. Log.* **26**(2): 143–164.

Banker, R. D., R. Natarajan. 2008. Evaluating contextual variables affecting productivity using data envelopment analysis. *Oper. Res.* **56**(1): 48–58.

Banker, R. D., S. Y. Lee, G. Potter, D. Srinivasan. 2010. The impact of supervisory monitoring on high-end retail sales productivity. *Ann. Oper. Res.* **173**(1): 25–37.

Battese, G. E., T. J. Coelli. 1995. A model for technical inefficiency effects in a stochastic frontier production function for panel data. *Empir. Econ.* **20**(2): 325–332.

Baumol, W., E. Ide. 1962. *Variety in Retailing, Mathematical Models and Methods in Marketing*. Richard D. Irwin, Inc, Homewood, IL.

Beattie, B. R., C. R. Taylor, M. J. Watts. 1985. *The Economics of Production*. Wiley, New York.

Beckman, C. M., P. R. Haunschild, D. J. Phillips. 2004. Friends or strangers? Firm-specific uncertainty, market uncertainty, and network partner selection. *Organ. Sci.* **15**(3): 259–275.

Bloom, P. N., V. G. Perry. 2001. Retailer power and supplier welfare: The case of Wal-Mart. *J. Retail.* **77**(3): 379–396.

Cachon, G. P. 2001. Stock wars: Inventory competition in a two-echelon supply chain with multiple retailers. *Oper. Res.* **49**(5): 658–674.

Cachon, G., S. Gallino, M. Olivares. 2019. Does adding inventory increase sales? Evidence of a scarcity effect in U.S. automobile dealerships. *Management Sci.* **65**(4): 1469–1485.

Cameron, A. C., P. K. Trivedi. 2009. *Microeconometrics Using Stata*. Stata Press, College Station, TX.

Cassidy, W. B. 2015. Widespread inventory "de-stocking" softens US freight markets. Available at http://www.joc.com/international-logistics/distribution-centers/widespread-inventory-de-stocking-softens-us-freight-markets_20150824.html (accessed date April 29, 2019).

Cassidy, W. B. 2016. US retailers reducing inventory, but e-commerce stock balance still tricky. Available at https://www.joc.com/international-logistics/distribution-centers/can-retailers-reduce-inventory-and-expand-online_20160818.html (accessed date April 29, 2019).

Cetinkaya, S., C. Y. Lee. 2000. Stock replenishment and shipment scheduling for vendor-managed inventory systems. *Management Sci.* **46**(2): 217–232.

Chen, C. M. 2017. Supply chain strategies and carbon intensity: The roles of process leanness, diversification strategy, and outsourcing. *J. Bus. Ethics* **143**(3): 603–620.

Chen, H., M. Z. Frank, O. Q. Wu. 2005. What actually happened to the inventories of American companies between 1981 and 2000? *Management Sci.* **51**(7): 1015–1031.

Chen, H., M. Z. Frank, O. Q. Wu. 2007. U.S. retail and wholesale inventory performance from 1981 to 2004. *Manuf. Serv. Oper. Manag.* **9**(4): 430–456.

Chen, C. M., M. Delmas, M. Lieberman. 2015. Production frontier methodologies and efficiency as a performance measure in strategic management research. *Strateg. Manag. J.* **36**(1): 19–36.

Christopherson, S. 2007. Barriers to "US style" lean retailing: The case of Wal-Mart's failure in Germany. *J. Econ. Geog.* **7**(4): 451–469.

Chuang, H. H. C., R. Oliva, O. Perdikaki. 2016. Traffic-based labor planning in retail stores. *Prod. Oper. Manag.* **25**(1): 96–113.

Coelli, T. J., D. S. P. Rao, C. J. O'Donnell, G. E. Battese. 2005. *An Introduction to Efficiency and Productivity Analysis*. Springer Science & Business Media, New York.

Cohn, E. 1992. Returns to scale and economies of scale revisited. *J. Econ. Educ.* **23**(2): 123–124.

Corbett, S. 2007. Beyond manufacturing: The evolution of lean production. *McKinsey Q.* **3**: 94.

Datta, T. K., K. Paul. 2001. An inventory system with stock-dependent, price-sensitive demand. *Prod. Plan. Control* **12**(1): 13–20.

Donthu, N., B. Yoo. 1998. Retail productivity assessment using data envelopment analysis. *J. Retail.* **74**(1): 89–105.

Dubelaar, C., G. Chow, P. D. Larson. 2001. Relationships between inventory, sales and service in a retail chain store operation. *Int. J. Phys. Distrib. Log. Manag.* **31**(2): 96–108.

Duois, A., L. Gadde. 2000. Supply strategy and network effects—purchasing behavior in the construction industry. *Eur. J. Purch. Supply Manag.* **6**(3–4): 207–215.

Eppen, G., L. Schrage. 1981. Centralized ordering policies in a multi-warehouse system with lead times and random demand. L. Schwarz, ed. *Multi-Level Production/Inventory Control Systems: Theory and Practice*, vol 16. North Holland, Amsterdam, 51–67.

Eroglu, C., C. Hofer. 2011. Lean, leaner, too lean? The inventory-performance link revisited. *J. Oper. Manag.* **29**(4): 356–369.

Eroglu, C., C. Hofer. 2014. The effect of environmental dynamism on returns to inventory leanness. *J. Oper. Manag.* **32**(6): 347–356.

Eroglu, C., C. Hofer, A. R. Hofer, A. Dedeke. 2015. A new way to measure inventory leanness. Available at http://www.supplychainquarterly.com/topics/Strategy/20151027-a-new-way-to-measure-inventory-leanness/ (accessed date April 29, 2019).

Fisher, M., A. Raman. 2010. *The New Science of Retailing: How Analytics Are Transforming the Supply Chain and Improving Performance*. Harvard Business Review Press, Boston, MA.

Forrester, A. T. 1961. Photoelectric mixing as a spectroscopic tool. *J. Opt. Soc. Am.* **51**(3): 253–259.

Fu, H. P., T. H. Chang, L. Shieh, A. Lin, S. W. Lin. 2015. Applying DEA–BPN to enhance the explanatory power of performance measurement. *Syst. Res. Behav. Sci.* **32**(6): 702–720.

Fullerton, R. R., F. A. Kennedy, S. K. Widener. 2014. Lean manufacturing and firm performance: The incremental contribution of lean management accounting practices. *J. Oper. Manag.* **32**(7): 414–428.

Gaur, V., S. Kesavan. 2009. The effects of firm size and sales growth rate on inventory turnover performance in the U. S. retail sector. N. Agrawal, S. Smith, eds. *Retail Supply Chain Management*. Springer, New York, 25–52.

Gaur, V., M. Fisher, A. Raman. 1999. What explains superior retail performance? Working paper, New York University, New York.

Gaur, V., M. Fisher, A. Raman. 2005. An econometric analysis of inventory turnover performance in retail services. *Management Sci.* **51**(2): 181–194.

Gaur, V., S. Kesavan, A. Raman. 2014. Retail inventory: Managing the canary in the coal mine. *Calif. Manage. Rev.* **56**(2): 55–76.

Ge, C., K. Huang. 2014. Analyzing the scale economies of software-as-a-service software firms: A stochastic frontier approach. *IEEE Trans. Eng. Manage.* **61**(4): 610–622.

Gelles, G. M., D. W. Mitchell. 1996. Returns to scale and economies of scale: Further observations. *J. Econ. Educ.* **27**(3): 259–261.

Germain, R., C. Claycomb, C. Droege. 2008. Supply chain variability, organizational structure, and performance: The moderating effect of demand unpredictability. *J. Oper. Manag.* **26**(5): 557–570.

Glaser, B. G., A. L. Strauss. 1967. *The Discovery of Grounded Theory*. AldineTransaction, New Brunswick, NJ.

Gligor, D. M. 2016. The role of supply chain agility in achieving supply chain fit. *Decis. Sci.* **47**(3): 524–553.

Greene, W. 2004. Distinguishing between heterogeneity and inefficiency: Stochastic frontier analysis of the World Health Organization's panel data on national health care systems. *Health Econ.* **13**(10): 959–980.

Greene, W. 2005. Reconsidering heterogeneity in panel data estimators of the stochastic frontier model. *J. Econom.* **126**(2): 269–303.

de Haan, J., M. Yamamoto. 1999. Zero inventory management: Facts or fiction? Lessons from Japan. *Int. J. Prod. Econ.* **59**(1): 65–75.

Haans, R. J., C. Pieters, Z. He. 2016. Thinking about U: Theorizing and testing U- and inverted U-shaped relationships in strategy research. *Strateg. Manag. J.* **37**(7): 1177–1195.

Halkos, G. E., N. G. Tzeremes. 2007. Productivity efficiency and firm size: An empirical analysis of foreign owned companies. *Int. Bus. Rev.* **16**(6): 713–731.

Hancerliogullari, G., A. Sen, E. A. Aktunc. 2016. Demand uncertainty and inventory turnover performance: An empirical analysis of the US retail industry. *Int. J. Phys. Distrib. Log. Manag.* **46**(6/7): 681–708.

Hayes, R. H., K. B. Clark. 1986. Why some factories are more productive than others. *Harv. Bus. Rev.* **64**(5): 66–73.

Hofer, C., C. Eroglu, A. R. Hofer. 2012. The effect of lean production on financial performance: The mediating role of inventory leanness. *Int. J. Prod. Econ.* **138**(2): 242–253.

Huson, M., D. Nanda. 1995. The impact of just-in-time manufacturing on firm performance in the US. *J. Oper. Manag.* **12**(3): 297–310.

Imrohoroglu, A., S. Tuzel. 2014. Firm-level productivity, risk, and return. *Management Sci.* **60**(8): 2073–2090.

Isaksson, O. H., R. W. Seifert. 2014. Inventory leanness and the financial performance of firms. *Prod. Plann. Control* **25**(12): 999–1014.

Jacobs, B. W., R. Kraude, S. Narayanan. 2016. Operational productivity, corporate social performance, financial performance, and risk in manufacturing firms. *Prod. Oper. Manag.* **25**(2): 2065–2085.

Johnston, A. 2014. Trends in retail inventory performance: 1982–2012. *Oper. Manage. Res.* **7**(3–4): 86–98.

Jondrow, J., C. K. Lovell, I. S. Materov, P. Schmidt. 1982. On the estimation of technical inefficiency in the stochastic frontier production function model. *J. Econom.* **19**(2): 233–238.

Jorge-Moreno, J. D., O. R. Carrasco. 2015. Technical efficiency and its determinants factors in Spanish textiles industry (2002–2009). *J. Econ. Stud.* **42**(3): 346–357.

Keh, H. T., S. Chu. 2003. Retail productivity and scale economies at the firm level: A DEA approach. *Omega* **31**(2): 75–82.

Kesavan, S., V. Mani. 2013. The relationship between abnormal inventory growth and future earnings for US public retailers. *Manuf. Serv. Oper. Manag.* **15**(1): 6–23.

Kesavan, S., V. Gaur, A. Raman. 2010. Do inventory and gross margin data improve sales forecasts for US public retailers? *Management Sci.* **56**(9): 1519–1533.

Koschat, M. A. 2008. Store inventory can affect demand: Empirical evidence from magazine retailing. *J. Retail.* **84**(2): 165–179.

Koumanakos, D. P. 2008. The effect of inventory management on firm performance. *Int. J. Prod. Perf. Manag.* **57**(5): 355–369.

Kraiselburd, S., R. Pibernik, A. Raman. 2011. The manufacturer's incentive to reduce lead times. *Prod. Oper. Manag.* **20**(5): 639–653.

Krommyda, I. P., K. Skouri, I. Konstantaras. 2015. Optimal ordering quantities for substitutable products with stock-dependent demand. *Appl. Math. Model.* **39**(1): 147–164.

Kumar, K., R. G. Rajan, L. Zingales. 2001. What determines firm size? Working paper, University of Chicago Graduate School of Business.

Kumbhakar, S. C., H. J. Wang, A. Horncastle. 2015. *A Practitioner's Guide to Stochastic Frontier Analysis Using Stata*. Cambridge University Press, New York.

Laeven, L., C. Woodruff. 2007. The quality of the legal system, firm ownership, and firm size. *Rev. Econ. Statist.* **89**(4): 601–614.

Lam, H. K. S., A. C. L. Yeung, T. C. E. Cheng. 2016. The impact of firms' social media initiatives on operational efficiency and innovativeness. *J. Oper. Manag.* **47–48**: 28–43.

Larcker, D. F., T. O. Rusticus. 2010. On the use of instrumental variables in accounting research. *J. Account. Econ.* **49**(3): 186–205.

Lee, C. Y., A. L. Johnson. 2013. Operational efficiency. A. B. Badiru, ed. *Handbook of Industrial and Systems Engineering*. CRC Press, Boca Raton, FL, 17–44.

Lieberman, M. B., L. Demeester. 1999. Inventory reduction and productivity growth: Linkages in the Japanese automotive industry. *Management Sci.* **45**(4): 466–485.

Li, S., J. Shang, S. A. Slaughter. 2010. Why do software firms fail? Capabilities, competitive actions, and firm survival in the software industry from 1995 to 2007. *Inform. Syst. Res.* **21**(3): 631–654.

Lieberman, M. B., R. Dhawan. 2005. Assessing the resource base of Japanese and US auto producers: A stochastic frontier production function approach. *Management Sci.* **51**(7): 1060–1075.

Lind, J. T., H. Mehlum. 2010. With or without U? The appropriate test for a U-shaped relationship. *Oxford Bull. Econ. Stat.* **72**(1): 109–118.

Lu, G., X. Ding, D. X. Peng, H. H. Chuang. 2018. Addressing endogeneity in operations management research: Recent developments, common problems, and directions for future research. *J. Oper. Manag.* **64**: 53–64.

Mason, S. J., P. M. Ribera, J. A. Farris, R. G. Kirk. 2003. Integrating the warehousing and transportation functions of the supply chain. *Transp. Res. Part E* **39**(2): 141–159.

Mishra, S., S. B. Modi, A. Animesh. 2013. The relationship between information technology capability, inventory efficiency, and shareholder wealth: A firm-level empirical analysis. *J. Oper. Manag.* **31**(6): 298–312.

Modi, S. B., S. Mishra. 2011. What drives financial performance–resource efficiency or resource slack?: Evidence from US Based Manufacturing Firms from 1991 to 2006. *J. Oper. Manag.* **29**(3): 254–273.

Mostafa, M. M. 2009. Benchmarking the US specialty retailers and food consumer stores using data envelopment analysis. *Int. J. Retail Distrib. Manag.* **37**(8): 661–679.

Mottner, S., S. Smith. 2009. Wal-Mart: Supplier performance and market power. *J. Bus. Res.* **62**(5): 535–541.

Murfin, J. 2014. Big box retailers squeeze smaller suppliers by borrowing from them. Available at https://insights.som.yale.edu/insights/big-box-retailers-squeeze-smaller-suppliers-by-borrowing-from-them (accessed date April 29, 2019).

Oliva, R., J. D. Sterman, M. Giese. 2003. Limits to growth in the new economy: Exploring the "get big fast" strategy in e-commerce. *Syst. Dynam. Rev.* **19**(2): 83–117.

Park, T. A., R. P. King. 2007. Evaluating food retailing efficiency: The role of information technology. *J. Prod. Anal.* **27**(2): 101–113.

Parmeter, C. F., S. C. Kumbhakar. 2014. Efficiency analysis: A primer on recent advances. *Found. Trends® Econometrics* **7**(3–4): 191–385.

Perdikaki, O., S. Kesavan, J. M. Swaminathan. 2012. Effect of traffic on sales and conversion rates of retail stores. *Manuf. Serv. Oper. Manag.* **14**(1): 145–162.

Porter, M. E. 2008. *Competitive Advantage: Creating and Sustaining Superior Performance*. The Free Press, New York.

Rajagopalan, S. 2013. Impact of variety and distribution system characteristics on inventory levels at US retailers. *Manuf. Serv. Oper. Manag.* **15**(2): 191–204.

Ratchford, B. T. 2003. Has the productivity of retail food stores really declined? *J. Retail.* **79**(3): 171–182.

Ratchford, B. T., J. R. Brown. 1985. A study of productivity changes in food retailing. *Market. Sci.* **4**(4): 292–311.

Reiner, G., C. Teller, H. Kotzab. 2013. Analyzing the efficient execution of in-store logistics processes in grocery retailing—The case of dairy products. *Prod. Oper. Manag.* **22**(4): 924–939.

Rumyantsev, S., S. Netessine. 2007a. What can be learned from classical inventory models? A cross-industry exploratory investigation. *Manuf. Serv. Oper. Manag.* **9**(4): 409–429.

Rumyantsev, S., S. Netessine. 2007b. Should inventory policy be lean or responsive? Evidence for US public companies. Working paper, The Wharton School, University of Pennsylvania, Philadelphia, PA.

Saranga, H. 2009. The Indian auto component industry—Estimation of operational efficiency determinants using DEA. *Eur. J. Oper. Res.* **196**(2): 707–718.

Sarkis, J. 2000. The Indian auto component industry—Estimation of operational efficiency determinants using DEA. *J. Oper. Manag.* **18**(3): 335–351.

Sellers-Rubio, R., F. Mas-Ruiz. 2006. Economic efficiency in supermarkets: Evidences in Spain. *Int. J. Retail Distrib. Manag.* **34**(2): 155–171.

Sellers-Rubio, R., F. J. Más-Ruiz. 2009. Technical efficiency in the retail food industry: The influence of inventory investment, wage levels, and age of the firm. *Eur. J. Mark.* **43**(5/6): 652–669.

Serrano, A., R. Oliva, S. Kraiselburd. 2018. Risk propagation through payment distortion in supply chains. *J. Oper. Manag.* **58–59**: 1–14.

Shah, R., P. T. Ward. 2003. Lean manufacturing: Context, practice bundles, and performance. *J. Oper. Manag.* **21**(2): 129–149.

Shan, J., K. Zhu. 2013. Inventory management in China: An empirical study. *Prod. Oper. Manag.* **22**(2): 302–313.

Shockley, J., L. A. Plummer, A. V. Roth, L. D. Fredendall. 2015. Strategic design responsiveness: An empirical analysis of US retail store networks. *Prod. Oper. Manag.* **24**(3): 451–468.

Solomon, M. B. 2013. The end of inventory. Available at http://www.dcvelocity.com/articles/20130308-the-end-of-inventory/ (accessed date April 29, 2019).

Soysal, G. P., L. Krishnamurthi. 2012. Demand dynamics in the seasonal goods industry: An empirical analysis. *Market. Sci.* **31**(2): 293–316.

Spiess, A., N. Neumeyer. 2010. An evaluation of R^2 as an inadequate measure for nonlinear models in pharmacological and biochemical research: A Monte Carlo approach. *BMC Pharmacol.* **10**(6): 1–11.

StataCorp. 2017. Stata Statistical Software: Release 15.1. StataCorp LLC, College Station, TX.

Stavrulaki, E. 2011. Inventory decisions for substitutable products with stock-dependent demand. *Int. J. Prod. Econ.* **129**(1): 65–78.

Suarez, F. F., R. Oliva. 2005. Environmental change and organizational transformation. *Ind. Corp. Change* **14**(6): 1017–1041.

Thomas, R. R., R. S. Barr, W. L. Cron, J. W. Slocum. 1998. A process for evaluating retail store efficiency: A restricted DEA approach. *Int. J. Res. Mark.* **15**(5): 487–503.

Ton, Z., A. Raman. 2010. The effect of product variety and inventory levels on retail store sales: A longitudinal study. *Prod. Oper. Manag.* **19**(5): 546–560.

Urban, T. L., R. C. Baker. 1997. Optimal ordering and pricing policies in a single-period environment with multivariate demand and markdowns. *Eur. J. Oper. Res.* **103**(3): 573–583.

Vastag, G. 2000. The theory of performance frontiers. *J. Oper. Manag.* **18**(3): 353–360.

Waller, M. A., T. L. Esper. 2014. *The Definitive Guide to Inventory Management*. Pearson Education, Upper Saddle River, NJ.

Wang, H., C. Ho. 2010. Estimating fixed-effect panel stochastic frontier models by model transformation. *J. Econom.* **157**(2): 286–296.

Wang, H., P. Schmidt. 2002. One-step and two-step estimation of the effects of exogenous variables on technical efficiency levels. *J. Prod. Anal.* **18**(2): 129–144.

Womack, J. P., D. T. Jones, D. Roos. 1991. *The Machine that Changed the World: The Story of Lean Production*. Harper Collins Publishers, New York.

Wooldridge, J. M. 2010. *Econometric Analysis of Cross Section and Panel Data*. MIT Press, Cambridge, MA.

Yang, M. G. M., P. Hong, S. B. Modi. 2011. Impact of lean manufacturing and environmental management on business performance: An empirical study of manufacturing firms. *Int. J. Prod. Econ.* **129**(2): 251–261.

Yu, W., R. Ramanathan. 2008. An assessment of operational efficiencies in the UK retail sector. *Int. J. Retail Distrib. Manag.* **36** (11): 861–882.

Yu, W., R. Ramanathan. 2009. An assessment of operational efficiency of retail firms in China. *J. Retail. Consum. Serv.* **16**(2): 109–122.

Zhang, J. J., N. Joglekar, R. Verma. 2012. Publishing the frontier of sustainable service operations management: Evidence from US hospitality industry. *J. Serv. Manag.* **23**(3): 377–399.

Ziobro, P. 2016a. Target corp. culling products in its stores to help resolve out-of-stock issues. Available at http://www.wsj.com/articles/target-corp-culling-products-in-its-stores-to-help-resolve-out-of-stock-issues-1456954473 (accessed date April 29, 2019).

Ziobro, P. 2016b. Retailers embrace barer shelves. Available at https://gcalhoun.files.wordpress.com/2016/06/16-06-28-wsj-retailers-embrace-barer-shelves.pdf (accessed date April 29, 2019).

Supporting Information

Additional supporting information may be found online in the Supporting Information section at the end of the article.

Appendix S1: Details on Robustness Checks.