Abstract#: 015-0148

Evaluating Store Design Responsiveness to Product Line Margin Changes: An Empirical Analysis of U.S. Public Retailers

Jeff Shockley
College of Business and Economics
Radford University
130 Whitt Hall, Radford, VA 24142-6054
tjshockley@radford.edu
Phone: 864-986-9232

Lawrence A. Plummer
Price College of Business
University of Oklahoma
307 W. Brooks, AH6B, Norman, OK 73019
larry.plummer@ou.edu

Aleda V. Roth
College of Business and Behavioral Science
Clemson University
124B Sirrine Hall, Clemson, SC 29634-1305
aroth@clemson.edu

Lawrence D. Fredendall
College of Business and Behavioral Science
Clemson University
123A Sirrine Hall, Clemson, SC 29634-1305
flawren@clemson.edu

POMS 21st Annual Conference
Vancouver, Canada
May 7 to May 10, 2010

Abstract

This paper subjects to rigorous empirical scrutiny the influence of retail store design responsiveness on firm financial performance. We posit that retail store systems that are managed dynamically to be responsive to changes in product line gross margin have higher firm performance (ROA). Using similar terminology and approaches from inventory management theory (e.g., Rumyantsev & Netessine 2005, 2007), "responsiveness" is defined as the firm matching (or aligning) product offering margins and store system design strategies. We employ an econometric model to test our theory using panel data collected from Compustat, 10-K, and S&P industry reports for "bricks and mortar" store retailers for the period 1994 – 2006. We find that firms that are increasing or decreasing labor and capital intensity in their store systems design--*at a faster rate than product line gross margin change*--have worse financial operating performance year-to-year than those firms who actively align store design strategy with product gross margins. The exploratory findings of this study also bring insight as to when retail store systems should be designed to be more responsive with store labor and capital intensity, and indicates when these store design elements need to managed to create a more efficient store design system.

1.0 Introduction

This study investigates and measures 'bricks and mortar' retailers' strategic store system design responses to product line gross margin changes over time. Service system design strategy specifies the roles of people (e.g. service workers), capital, and the specific processes by which services are created and delivered (Chase and Bowen 1991, Goldstein et al. 2002, Roth and Jackson 1995). Borrowing from inventory management research methods and terminology (Rumyantsev and Netessine 2005, 2007b), 'design responsiveness' describes and measures the co-movement (alignment) of key store design decisions with product line gross margins over time. Specifically, measures of responsiveness and dynamic panel data analysis techniques are used to determine if retailers that simultaneously manage product-line margins, store labor and capital intensity (store investment) in their store system design have superior operational performance versus other retail firms.

An example of store system design responsiveness can be seen in the U.S. consumer electronics retail segment. In the 1990's, Best Buy Company, Inc. relied on a predominantly self-service (low labor intensity) store system design strategy. During this period, Best Buy stores specialized in selling accessories, games, and personal computers in a rapidly declining product line margin environment and competed primarily on price and cost-efficiency. By 2001 - 2002, mass merchants Costco and Wal-Mart, as well as web-channels, had moved aggressively into the consumer electronics segment using an even more cost-efficient product selling model. In response, Best Buy's management recognized that it had to reformulate its store system design strategy to sell a portfolio of higher margin goods and services to remain competitive (Lal et al. 2006, p.3). While other segment competitors (e.g. Circuit City, CompUSA) de-emphasized human contact in their stores, Best Buy introduced a higher margin product and service offering (Lal et al. 2006, p.4, O'Donnell 2008), and simultaneously invested in more store labor and capital intensity to sell a wider variety of bundled digital products and supporting store services. While the store system operating changes were initially met with skepticism by outsiders, Best Buy has far outperformed all its segment competitors in recent years (O'Donnell 2008).

Service/product bundle offering strategies do not often align with delivery system design strategies in practice, providing an opportunity for service operations management research (Roth and Menor 2003). Retail analysts, in particular, have struggled to craft meaningful measures to link strategic design-related factors to financial and operational performance measures (Gage, *Forbes* 2007). Linking the service-product offering to the design strategy is a critical determinant of service capabilities and sustainable performance (Roth and Menor 2003). Yet, the alignment of product line margin and store system design strategies is a dynamic process that is not well-understood by either practitioners or academics.

Retailers simultaneously do manage both product offering and service delivery functions (Murray and Schlacter 1990) in their store systems. Because they offer both tangible products and supporting store services, retailers also have different operating and environmental characteristics than other types of services - e.g. hospitals or banks. ‘Bricks and mortar’ store retailing has also received little specialized attention in design strategy literature. Notable exceptions include work by DeHoratius and Raman (2007), examining manager job design and incentive structure at Tweeter Electronics stores; Fisher et al. (2006) who examine retail store execution measures among stores in a single chain retailer; and Fisher, Krishnan, and Netessine (2009) who investigate the profit impact of retail stores abilities to match labor to store traffic over time. In 2001, *Manufacturing and Service Operations Management* published a focused issue (Vol. 3, No. 3) on ‘Retail Operations Management.’ However, the focus of this series of papers is on more tactical applications of operations research techniques to solve assortment, logistic, and inventory optimization problems. Like other more strategic issues surrounding services (Menor et al. 2001, p.275), retail store design strategy has not been a key focus area of empirical operations management research.

Nevertheless, store retailing is a critical component of the U.S. economy, employing the largest number of American workers and making up over $1.3 trillion in domestic economic output (U.S. Bureau of Economic Analysis 2007, http://www.bea.gov/industry, 9/6/2008). It is also a particularly aggressive and dynamic, with any strategic move (like a price cut) requiring an “immediate response” from other retailers to maintain competitive position or to even survive (Ghosh 1990, p.37). Moreover, U.S. retailers spend over $30 billion annually on capital investment (mostly on technology systems or better store locations) to improve internal and external process performance (Fisher and Raman 2001). Yet, the perceived financial return on these strategic capital investments has been mixed. Retail firms are still characterized by high failure rates and low customer service (McGurr and DeVaney 1998), but have seen improvement in some inventory and operating-efficiency measures in recent years (Chen et al. 2007).

Retailers may respond to competitor threats either by imitation (Boyd and Bresser, 2008), or by investing more in store labor or capital resources (like Best Buy) to support the selling of more complex products, or a wider variety products and store services (Menor et al. 2001, Lal et al. 2006).

The research questions for this study are: Can store design responsiveness be measured using publicly accessible secondary data? Do retail firms pursue responsive store design strategies to product line margin changes? Finally, hypotheses are proposed to examine: Does store system design responsiveness indicate better (or worse) firm operating performance? By developing a theoretical and empirical link from strategic store design choices to financial operating performance, this paper examines whether retailers should design their store systems in alignment with product offering margins, or if they should strive to design stores to be more cost-efficient in all cases, reflecting a Wal-Martization of retail store design strategy (Boyd and Bresser 2008).

2.0 Literature Review

Design responsiveness further reflects the importance of aligning customer contact requirements with actual retail store system design strategies. Customer contact theory (Chase 1978, 1981) has arguably become the dominant theoretical lens through which researchers have viewed service operations management (SOM) and design strategy. Generally, design strategies can be organized and positioned around the customer contact requirements anticipated in the service system (Chase and Tansik 1983). Customer contact needs are driven by both the customer-perceived complexity and the information content of the service (Buzacott 2000). More recent interpretations of customer contact theory have focused on management opportunities to use technology capital investment and location accessibility to substitute for human contact requirements (Xue et al. 2007, Froehle and Roth 2004, Boyer et al. 2002).

Customer contact levels are often weighed against a retailer's desire to design more automated delivery channels. While automation in service contexts generally means developing self-service channels for product/service delivery (Buzacott 2000), this is only possible if customers can effectively perform service tasks without help. More labor (human) contact will help to manage additional service encounter complexity (Chase 1978, Kellogg and Chase 1995) in most cases, but investments in technology and/or increased channel accessibility can mediate direct human contact requirements (Froehle and Roth 2004, Boyer et al. 2002). Customer prior product knowledge, location convenience, or information clarity will ultimately determine the effectiveness of the self-service design channels (e.g. Bateson 1985, Xue et al. 2007).

Service strategy literature grounded in the resource based view (RBV) of the firm (Wernerfelt 1984), argues that organizations strategically choose, build, combine, and deploy human and capital resources to design service "architecture" that adapts to customer contact needs. A firm's design "architecture" is built from specific structural (buildings, equipment), infrastructural (policies, job design, and labor management), and coordinative resource choices (Roth and Jackson 1995). While empirical studies of retail design architecture are lacking, the importance of continuously aligning service design capabilities with product/service offerings is generally acknowledged in the extant literature (e.g. Roth and Menor 2003, Fitzsimmons and Fitzsimmons 1999).

In contrast to simply designing service systems to be more cost-efficient (more self-service) in all cases, some research argues that firms follow a progression in aligning resource competencies with product markets (Heskett et al. 1990, Menor et al. 2001). For example, both

Menor et al. (2001) and Roth and van der Velde (1992) find that increasing banking product offering complexity ultimately resulted in higher operating margins, and the need for retail banks to deploy more flexible service design architecture. Moreover, there are immediate negative financial impacts to firms being unresponsive (or not adapting) to customer service needs (Fornell 2007).

The field of operations management (OM) has established important associations between gross margins, product-service offering complexity and variety, and the most appropriate production design strategy (e.g. Gaur et al. 2005, Randall and Ulrich 2001, Hayes and Wheelwright 1979). More human skill (or know-how), not just capital (technology) investment, may also be necessary to manage more dynamic product-service offering environments (Menor et al. 2001). More complex products, higher product margins, and inefficient operating design strategies can also cause operating profit stagnation in many cases (Gottfredson and Aspinall 2005). Nevertheless, Gaur et al. (1999) find generally positive associations between retail product gross margins and firm performance.[1] Retailers can improve profits through effectively manipulating their store labor and capital resources; particularly, if the appropriate incentive structures are in place (DeHoratius and Raman 2007), and if store labor-staffing requirements are met (Fisher et al. 2006).

Responsive firms adapt quickly to changes in market demand conditions, resulting in higher profits (Randall et al. 2003). Retail firms, in particular, need responsive store design architectures to facilitate quick changes because the industry is so highly competitive. However, this study's concept of design responsiveness is only a proxy measure for the degree of operational alignment between product line margin and retail store design strategy over time, and

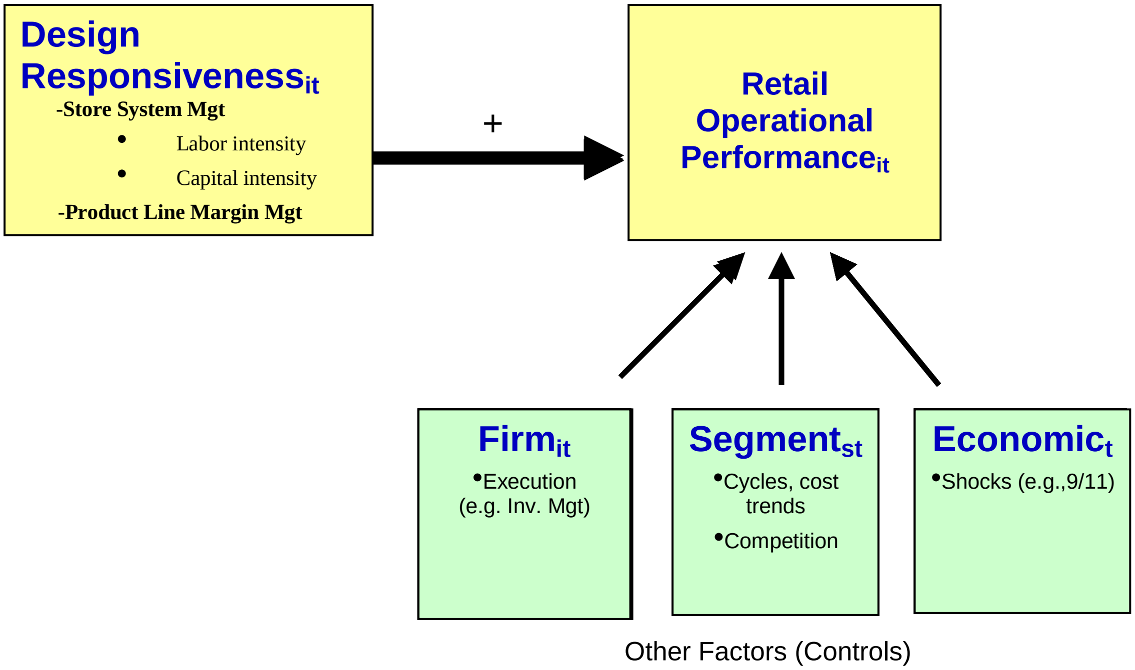
[1] Gaur, Fisher, and Raman (1999) do not examine the co-movement of gross margin and design strategy. Rather, they examine a measure of GMROI to show the positive association of gross margins on retail stock returns.

is not a direct measure of a firm's agility or strategic intent. Therefore, several assumptions are made: First, the assumption that gross margins and design strategies are linked is grounded in the understanding of the retail investment analyst community and academic research that product margin peaks indicate strategic design shifts in retailing (Gage, *Forbes* 2007). Second, little retail industry literature examines retail store system design strategy, or develops much empirical measurement related to service strategy at all (Menor et al. 2001). Finally, the systematic and dynamic relationships between service design strategy, product line or service context, and performance measurement are not well developed (Soteriou and Zenios 1999). This paper contributes to a greater understanding of these relationships by analyzing the dynamic changes in retail design strategies (Boyd and Bresser 2008) using dynamic panel analytical techniques.

3.0 Model Development and Hypotheses Formulation

Retail design responsiveness is the simultaneous management of the store system design strategy with product line margins over time (Figure 1). This study posits that design responsive firms will outperform unresponsive firms, after controlling for other firm-specific, segment-specific, and time effects (McGahan and Porter 2002, p.835). Therefore, retail firm strategic design planning and decision-making is seen in year to year relationships among key operational variables such as gross margin and in resource investment decisions about store system employee labor and store capital. To identify strategic design shifts, the model measures the amount of 1) store labor intensity, and 2) store capital intensity required to expand firm profits. Retailing firms choose the level of customer contact their store systems provide, and what form this contact will take (Chase 1978).

Figure 1: Conceptual Model – Factors Affecting Operational Performance in Store Retailing



3.1 Measuring Design Responsiveness

The proxy measures used to evaluate store design responsiveness are grounded in both current and classical inventory management research methodologies to measure the elasticity of inventory supply and sales demand (Rumyantsev and Netessine 2005, 2007b, O'Glove 1987). This logic is adapted in this study to analyze the annual change in product line gross margin versus annual change to store system designs by using store labor and capital intensity as proxy variables to understand those strategic design relationships. The basic equation for measuring design responsiveness is stated as follows for the store labor intensity design responsiveness variable:

$$SL_t = \frac{L_t - L_{t-1}}{L_{t-1}} - \frac{GM_t - GM_{t-1}}{GM_{t-1}} \quad \text{(Eq. 1)}$$

where L_t stands for period t store labor intensity, and GM_t stands for period t product line gross margins. A positive result (> 0) indicates store system labor intensity is increasing at a faster rate than gross margins, while a negative (< 0) result indicates that store system labor intensity is declining relative to gross margins. A score of zero would indicate complete design responsiveness, as changes in gross margin were matched with store system labor intensity shifts in the given year.

The degree of human contact (or conversely self-service level) used in the store delivery system strategy is measured with a *store labor intensity* (*L*) ratio, which is simply the number of store employees per selling square foot. While self-service store design strategies requiring less human contact will typically require lower labor intensity to deliver the service and maintain profitability, firms increasing human contact levels will also increase the labor intensity ratio in their store systems. Because research finds that more complex product offerings are associated with higher gross margins (Randall and Ulrich 2001), high store labor intensity will need to correspond with high gross margins to stay in alignment. Alternatively, store system designs that fail to provide adequate human labor contact will ultimately see negative performance impacts (Menor et al. 2001) from the lack of sales support. Therefore, the following hypotheses are stated for both the positive and negative responsiveness measures for store labor intensity:

H1: When SL is <u>positive</u>, a higher measure of design responsiveness in store labor intensity will be associated with worse operational performance.

H2a: When SL is <u>negative</u>, a lower measure of design responsiveness in store labor intensity will be associated with worse operational performance.

While some operations research has advocated that retail firms should pursue self-service design strategies only when selling simple, lower-margin products (Buzacott 2000); other research challenges this notion by indicating that retailers have achieved disproportionate financial benefits through self-service designs to increase economies of scale and per unit costefficiencies (Boyd and Bresser 2008, Chen et al. 2007). Therefore, the following alternative hypothesis for negative store labor intensity responsiveness is examined.

H2b: When SL is <u>negative</u>, a lower measure of design responsiveness in store labor intensity will be associated with better operational performance.

Similarly, substituting *K* for *L* provides the baseline equation for calculating design responsiveness for store capital intensity vis-à-vis product line gross margin changes:

$$\text{(Eq. 2)} \quad SK_t = \frac{K_t - K_{t-1}}{K_{t-1}} - \frac{GM_t - GM_{t-1}}{GM_{t-1}}$$

where K_t stands for period *t* store capital intensity, and GM_t stands for period *t* product line gross margins. Store design strategies that leverage technology, store fixtures, or location investment are represented by the variable *store capital intensity* (*K*) – which is the ratio of store-invested capital per selling square foot. Operations management research states that a retail firm may wish to use technology capital to manage complexity or product variety (Gaur et al. 1999), or may wish to invest in new store locations that are more convenient for customers to access (Xue et al. 2007). Increasing store capital intensity also helps the firm substitute for human contact, or helps provide greater economies of scale and a more cost-efficiency. Retail firms either purchase or enter into lease agreements for buildings, technology, or store fixtures to achieve customer contact objectives. A positive responsiveness measure for store capital intensity (> 0) indicates that a retail firm may be over-investing in store capital. Conversely, a negative capital intensity responsiveness measure (< 0) suggests that a retail firm was under-investing in store capital, possibly leaving itself vulnerable to more adaptive retailers with more robust selling systems, financing, or better store locations.

H3: When SK is <u>positive</u>, a higher responsiveness measure in store capital intensity will be associated with worse operational performance.

H4: When SK is <u>negative</u>, a lower responsiveness measure in store capital intensity will be associated with worse operational performance.

4.0 Database Sample Description

Financial data was collected for the time-period 1994-2006 for all "bricks and mortar" U.S. public retailers listed on the U.S. stock exchanges in the Standard and Poor's COMPUSTAT Annual Fundamentals database using Wharton Research Data Services (WRDS, http://wrds.wharton.upenn.edu). The fiscal year 1994 was chosen as the starting date because it is the first full year of data after the end of the last retail recession. Product-selling retailers and their product line categories were identified based on the four digit Standard Industry Classification (SIC) sample selection criteria for "retail trade" outlined in Gaur et al. (2005), excluding wholesalers, e-commerce retailers, retail holding companies, bankruptcy years, and American depository receipts (ADRs). There were 487 retailers that report at least one year of financial data to the Securities and Exchange Commission (SEC) during the study time period.

Selected firms had five or more years of complete financial data during the period (Peterson and Fabozzi 2006, Kremer et al. 2000), which further reduced the industry field sample to 320 retailers. Data was then manually collected on the number of stores and the gross selling space (square feet) for each retailer in each year from 10-K (annual report) statements accessed through the SEC database (http://www.sec.gov/edgar). This data was supplemented and validated using retail statistics purchased from the U.S. Business Reporter Database (http://www.usbrn.com), and S&P Retail Industry Reports (Also available in WRDS, http://wrds.wharton.upenn.edu). Most retailers do report aggregate store-level operating data in their annual statements, and only 88 retailers (out of the 320 selected) did not report any store-level information during the study time-period. This left 232 retail firms and 2,039 observations for the final sample (Table 1).

Table 1:
Frequency table showing number of years of reported store level data (1994 – 2006)

Number years of complete store-level information	**Number of retailers**	**Number of observations**
Retailers 5+ yrs Financial Data	*320*	
0 (dropped)	-88	------
1-3	**21**	**48**
4-6	**58**	**287**
7-9	**38**	**301**
10-12	**44**	**480**
13	**71**	**923**
Final Industry Field Sample	***232***	***2,039***

Four digit primary SIC codes are assigned to each retail firm to identify the primary industry or product line segment in which it operates. Because retail firms may span several product line segments or have multiple (secondary) SIC codes, the guidelines of Gaur et al. (1999, 2005) were followed to avoid small sample bias in these cases. Retailers were put in product line segments based on these relatively distinct 12 product line / primary SIC code groupings. Table 2 lists the segments, groupings, and example firms.

Table 2:
Retailers reporting store-level information for square feet and # of stores (1994 – 2006)

SIC (4 digit)	Segment Group Name	# of Retailers	Examples
5211	Lumber and building materials stores	6	Home Depot, Lowes, National Home Centers
5311	Department stores	17	Sears, Macy's, Dillards, J.C. Penny
5331, 5399	Variety stores	25	Wal-Mart, Target, Warehouse Clubs
5400 – 11	Grocery stores	35	Albertsons, Kroger, Safeway
5600 – 99	Apparel and accessory stores	64	Ann Taylor, Gap, Limited
5700 – 11	Home furnishings and equip stores	14	Bed, Bath, and Beyond, Linens-N-Things
5731	Radio, TV, and appliance stores	14	Best Buy, Circuit City, RadioShack
5734, 5735	Computer and computer software stores, Records and tapes	9	Babbages, CompUSA, Gamestop
5912	Drug and proprietary stores	7	CVS, Rite Aid, Walgreens
5940	Misc. stores- other	24	Staples, Barnes and Noble, Sports Authority, etc
5944	Jewelry stores	7	Tiffany, Zale
5945	Hobby, toy, and game	10	Toy's R Us, Zany Brainy, Michaels, etc.
Sample Total		**232**	

5.0 Variable Definitions

The following notation is used for the model variables. From the Compustat Annual Fundamentals data, for firm *i* in year *t*, let S_{it} be the total sales for the firm (Compustat Fundamentals field 'SALE'); $COGS_{it}$ be the cost of goods sold (COGS); AT_{it} ending total assets for the period (AT); $LIFO_{it}$ be the LIFO reserve (LIFR); $INVT_{it}$ be the ending total inventory for the period (INVT); $OIBD_{it}$ be the operating income before depreciation (OIBDP); PPE_{it} be the ending net property, plant, and equipment for the period (PPENT); and EMP_{it} be the average number of employees for firm *i* calculated by averaging the ending number of employees (EMP) for year *t-1* and year *t* for each year *t*. From the 10-K and S&P collected data, let $SQFT_{it}$ be the average gross selling square feet for firm *i* calculated by averaging SQFT for year *t-1* and year *t*.

Several adjustments must be made to make firm-level performance variables and ratios comparable. Retail firms use different inventory valuation methods (e.g. FIFO versus LIFO methods) and this accounting practice produces differences in firm to firm reporting of period-ending inventory (INVT) and cost of goods sold (COGS). To account for inventory valuation method differences, the LIFO reserve (Compustat field 'LIFR') is added into the ending inventory calculation for of a given fiscal year and the change in LIFO reserve from year to year was subtracted out of period-ending COGS (e.g., see Kesavan et al. 2008). This practice ensures that resulting ratios calculated from these variables for the sample firms are comparable.

5.1 Dependent Variables

The retail firm's return on assets (*ROA*) is the primary measure of operational performance used in this study (Barber and Lyon 1996, Rumyantsev and Netessine 2005, 2007a,b, Gaur et al 1999). Because performance measures using operating income ratios can vary based on firm scale or accounting treatments (Barber and Lyon 1996, p.397), any potential performance measurement bias is controlled for by using return on sales (*ROS*) as an alternative performance dependent variable. Examining the carryover associations of strategic design decisions on forward firm operational performance is also important, so forwarded ROA (*ROAF*) for *t+1* year period is examined as a dependent variable (Rumyantsev and Netessine 2005, 2007b). The basic formula for ROA in year *t* is calculated[2]:

$$\text{(Eq. 3)} \quad ROA_{its} = \frac{OIBD_{its}}{\left(AT_{i(t-1)s} + AT_{its}\right)/2}$$

5.2 Independent Variables

The joint movement of product gross margins and store system design strategies for labor and capital intensity is measured using the design responsiveness elasticity measures SL_{its} and SK_{its} introduced in section 3.1 and Table 3 shows the components and each design responsiveness measure. GM_{its} is Sales minus COGS, adjusted for the inventory valuation

---

[2] Return on Sales (ROS) and Forwarded ROA (ROAF) are calculated in a similar fashion as regular ROA using OIBD in the numerator, and sales and forward average assets respectively in the denominator.

method. L_{its} is the ratio of the number of employees (*EMP*) to total gross selling square feet (*SQFT*) for all stores during the time-period. K_{its} is the sum total of PPE_{its} and the net present value of five-year lease contracts (capitalized leases) using the notation $LC_{it,1}$ (MRC1), ….., $LC_{it,5}$ (MRC5) in Compustat[3]. To simplify the capitalized lease analysis, the discount rate $r = 9.3\%$ was used based on the average weighted average cost of capital (WACC) of the retailing industry reported from Value Line (A. Damodaran, Damodaran Online, http://pages.stern.nyu.edu/~adamodar, 9/3/2008).

Table 3:
Definition of Component Measures and Independent Variables

Component Measures	
Product Line Margin (Gross Margin)	$GM_{its} = \frac{S_{its} - COGS_{its}}{S_{its}}$
Store Labor Intensity	$L_{its} = \frac{EMP_{its}}{SQFT_{its}}$
Store Capital Intensity	$K_{its} = \left[PPE_{its} + \sum_{\tau=1}^{5} \frac{LC_{its}}{(1+r)^{\tau}} \right] \Big/ SQFT_{its}$
Store System Design Responsiveness Measures	
Design Responsiveness – Store Labor Intensity	$SL_{its} = \frac{L_{its} - L_{i(t-1)s}}{L_{i(t-1)s}} - \frac{GM_{its} - GM_{i(t-1)s}}{GM_{i(t-1)s}}$
Design Responsiveness– Store Capital Intensity	$SK_{its} = \frac{K_{its} - K_{i(t-1)s}}{K_{i(t-1)s}} - \frac{GM_{its} - GM_{i(t-1)s}}{GM_{i(t-1)s}}$

** Note that COGS is adjusted for the LIFO reserve as stated above*

A positive (or negative) result for any store design responsiveness measure would indicate that a firm was trying to increase (decrease) labor or capital intensity in their store

[3] Both Gage, *Forbes* (2007) and Kesavan, Gaur, and Raman (2008) discuss the importance and techniques for adjusting for capitalized leases when conducting capital analyses among different store retailers.

systems at a faster (slower) rate than changes to product line gross margins. Since the hypotheses predict negative relationships between these variables and the operating performance measures, the following inventory co-movement methodology is used (adapted from Rumyantsev and Netessine 2007b) to further define variables to capture both directional positive (increasing) and negative (decreasing) co-movements for each design responsiveness variable:[4]

For *Store Design Responsiveness – Labor Intensity*,

$$SLinc_{its} = SL_{its} \times 1_{(SL \geq 0)} \text{ ; } SLdec_{its} = SL_{its} \times -1_{(SL \leq 0)}$$

For *Store Design Responsiveness – Capital Intensity*,

$$SKinc_{its} = SK_{its} \times 1_{(SK \geq 0)} \text{ ; } SKdec_{its} = SK_{its} \times -1_{(SK \leq 0)}$$

5.3 Control Variables

The following additional notation is used to operationalize the control variables. From the 10-K and S&P collected data, let N_{it} be the total number of stores open for firm i at the end of year t. Inv_{it} for a given year is calculated based on the adjusted average of the prior ($INVT_{i(t-1)}$) and current year ($INVT_{it}$) period-ending inventory balances (Adjusted for the LIFO reserve).

From this data, the following additional independent variables are defined in Table 4:

[4] The notation used here assumes the variable is "0" otherwise. So, for "labor intensity" responsiveness: *SLinc* = SL*1 if SL >= 0; *SLinc* = 0 if SL <= 0; *SLdec**-1 if SL < =0; *SLdec* = 0 if SL >= 0, etc.

Table 4:
Definition of Control Measurement Variables

Control Variables	Calculations	Related References
Firm-specific		
Firm size	$\log S_{its}$	Barber and Lyon (1996), Gaur, Fisher, and Raman (2005, 1999), Rumyantsev and Netessine (2005, 2007a,b)
Sales (revenue) growth rate (firm sales growth)	$RG_{its} = S_{its} / S_{i,t-1,s}$	Gaur, Fisher, and Raman (2005, 1999), Rumyantsev and Netessine (2007a,b)
Store growth rate (firm store growth)	$NG_{its} = N_{its} / N_{i,t-1,s}$	Gaur, Fisher, and Raman (2005, 1999)
Inventory management* (relative inventory)	$I_{its} = Inv_{its} / COGS_{its}$	Gaur, Fisher, and Raman (2005, 1999), Rumyantsev and Netessine (2005, 2007a,b), Fisher, Ramdas, and Zheng, (2001), Jayanthi, Roth, Kristal, and Venu (2009)
Industry Segment		
Segment margin*	$SM_{ts} = [S_{ts} - COGS_{t-1,s}] / S_{(t-1),s}$	Rumyantsev and Netessine (2005, 2007b), Cheng (2005)
Segment sales growth rate	$SG_{ts} = S_{ts} / S_{(t-1),s}$	Rumyantsev and Netessine (2007b), Cheng (2005)
Competitive intensity (segment diversification or entropy)	$E_{ts} = \sum_{s=1}^{S} \rho_{ist} \ln\left(1 / \rho_{ist} \right)$	Jayanthi, Roth, Kristal, and Venu (2009), Palepu (1985)
Economic (shocks)		
Fiscal Year	*fyear* = yearly dummy	Roodman (2006), McGahan and Porter (2002) STATA-‘xi: …DV IV i.fyear, …’

** Note that here both COGS and Inv are adjusted for LIFO reserve as stated above*

Firm-specific control variables

Controls for *firm size* (log of firm Sales*)*, *sales growth rate*, *store growth rate*, and *inventory management* were used based on guidance from the extant literature. *Inventory management* (I_{it}) is operationalized by using the ratio of average inventory (*Inv*) to cost of goods sold (*COGS*) to evaluate inventory management effectiveness and supply chain performance. Controlling for inventory management effectiveness is important to account for any association between retail firm inventory position and ROA, an association that has been already established in the OM literature (e.g. Gaur et al., 2005, 1999, Rumyantsev and Netessine 2005, 2007a, Fisher et al. 2001, Jayanthi et al. 2009). Furthermore, inventory ratios are proven measures of retail inventory management effectiveness on an annual basis (e.g., Gaur et al. 1999, Rumyantsev and Netessine 2005).

Industry control variables

Segment gross margin (SM_{ts}) is controlled for to make sure that a firm's product line gross margins are measured relative to the average gross margins of its industry (Rumyantsev and Netessine 2005, 2007b). Furthermore, *segment sales growth* (SG_{ts}) ratio, is used to control for sales trends within product line industry segments (Rumyantsev and Netessine 2005, Cheng 2005). Lastly, the *competitive intensity* in the industry is controlled by using a measure of segment diversification or entropy (E_{ts}), which is stated as simply the transformed ratio of total sales for the market share leader in a given industry segment for a given year (Jayanthi et al. 2009). A higher score for E_{ts} would indicate that a firm operated in a more diverse and competitive segment.

Time control variables

Yearly dummies (*fyear*) are used to control for possible trends in profitability over time due to one-time economic shocks or industry cycles (Roodman 2006, McGahan and Porter 2002). This is done using the "xi:…. i.*fyear*" procedure in STATA.

5.4 Empirical Model Specification

The following base empirical model (Eq. 4) is used to examine a retail firm's financial operating performance with store design responsiveness measures, while simultaneously controlling for other firm-specific, industry segment, and timing variables that may be present:

$$\text{(Eq. 4)} \quad \mathrm{ROA}_{it} = \mu_{it} + \varepsilon_{it} + b^{i}\,\mathrm{ROA}_{i,t\text{-}1} + b^{1}SLinc_{it} + b^{2}SLdec_{it} + b^{3}SKinc_{it} + b^{4}SKdec_{it} + b^{5}I_{it} + b^{6}NG_{it} + b^{7}RG_{it} + b^{8}logS_{it} + b^{9}SM_{ts} + b^{10}SG_{ts} + b^{11}E_{ts} + d^{1}fyear$$

where μ_i indicates the firm-specific error, ε_{it} is the remaining random model error**,** b^i is the coefficient for the temporal lag of the "ROA" dependent variable, b^1, b^2 , b^3 and b^4 are the directional coefficients for the firm-specific design responsiveness variables for store labor (*SL*) and capital intensity (*SK*) changes, b^5, b^6, b^7 and b^8 are the coefficients for other firm-specific control variables, b^9, b^{10}, and b^{11} are the coefficients for segment-specific control variables, and d^1 represents the time control dummy variables (Rumyantsev and Netessine, 2007b, Roodman, 2006, McGahan and Porter, 2002). So, the hypotheses are confirmed if negative and statistically significant coefficients are seen for b^1, b^2, b^3 and b^4. The same model (Eq. 4) is estimated substituting for a forwarded ROA dependent variable (*ROAF*) for the *t+1* forward time period, as well as for current period return on sales (*ROS*).

It is also necessary that the results be examined for sensitivity (Kennedy, 2003) to the different assumptions and variables included in the model. While the model's statistical power in excess of 98%, many of the underlying statistical assumptions are sensitive to the number of variables entered into the model and the number of instruments used versus the number of variables. Therefore, multiple statistical analyses were used to "test up" and "test down" the model by adding some variables and removing those that are redundant or may not be necessary (Plummer 2007, Kennedy 2003). In addition, alternative model specifications were examined to determine if the number of instruments used in the model was necessary or appropriate (Roodman 2008).

6.0 Research Design: Analytical and Methodological Approach

The longitudinal research design uses dynamic panel data analysis techniques in STATA v10 to test the hypotheses. The use of dynamic panel models is "part of broader historical trend in econometric practice toward estimators that make fewer assumptions about the underlying data-generating process and use more complex techniques to isolate useful information" (Roodman 2006, p.13) from large longitudinal panel data sets. A dynamic panel data model is "one containing a (temporal) lagged dependent variable (and possibly other regressors), particularly in the 'small T, large N' context" (Baum 2006, p.232). The lagged dependent variable term is assumed to be correlated with the error term in the overall model, and this persistence bias becomes more acute as the number of observations in each time-period sample increases. Profit-derived ratios have been found to exhibit high levels of persistence in prior literature (Plummer 2007, Oei et al. 2008, Roberts 2001, Waring 1996). Therefore, it is both important and necessary to account for the persistent effects of the dependent variable using established generalized method of moments (GMM) estimation techniques (Hansen 1982). All the dynamic models are estimated using the "xtabond2" command in STATA v10 (Roodman 2006).

One method to account for persistence of the dependent variable lagged term (b^i) in the model (Eq. 4) is to use GMM techniques to create a system of equations (one per period) that allow for applicable instruments to be created for each equation term (Baum 2006, p.233). The first differencing in this process effectively removes the individual (fixed) effects for each model variable, and the lagged difference term instruments for additional correlation with the overall model residual error. System GMM gives more "reasonable and precise estimates" (Baltagi 2001, pp.143-144) versus Difference GMM (Arellano and Bond 1995) when instruments are weak, by combining levels terms with differences terms to create new system of equations using all available instruments. This procedure allows System GMM to instrument for the lagged dependent variable term and "any other endogenous variables with variables thought uncorrelated (orthogonal) with the fixed effects" (Roodman 2006, p.16). Because it takes more full advantage of all future moment conditions (Arellano and Bover 1995), System GMM is also more efficient with degrees of freedom, but may not strictly eliminate all the firm fixed-errors.

System GMM estimation can also be sensitive to the number of instruments used versus the number of parameters estimated in the model, and is subject to misinterpretation (Roodman 2006, 2008). Unfortunately, the extant literature provides little guidance on how many instruments is "too many" (Roodman 2006, p.13, 2008), but it does provide some guidance on how to test if assumptions embedded in these techniques are being violated (Plummer 2007, Roodman 2006). To validate the underlying assumptions in the data and to reduce the size of the model, a series of diagnostic tests and sensitivity analyses were conducted (See Online Appendix for a full accounting of these tests). The main findings of interest in these diagnostics were that collinearity concerns could be alleviated by dropping any one or all of the segment control variables. As is typical of financial panel data, heteroscedasticity was found in the sample which required the use of robust estimation techniques to adjust for these scalar differences.

7.0 Analysis

7.1 Descriptive Statistics Analysis

First, descriptive statistics were used to analyze the dependent variable and gross margin trends for the retail trade industry sample (Figure 2). While gross margin and ROA results vary among retail industry segments, is it interesting to note that despite increased spending on technology and supply chain management capital during the period, the retail industry sample in aggregate did not see any real increase in ROA. However, Figure 3 shows that retailers have been active and quite volatile in their annual labor and capital investment decisions vis-à-vis gross margin changes during the period. While the tendency for retailers was to decrease store labor intensity and increase store capital intensity at a faster rate than gross margin changes, there is no discernable or directional trend in either area.

Figure 2: Plot of Industry Gross Margin and Return on Assets by Fiscal Year

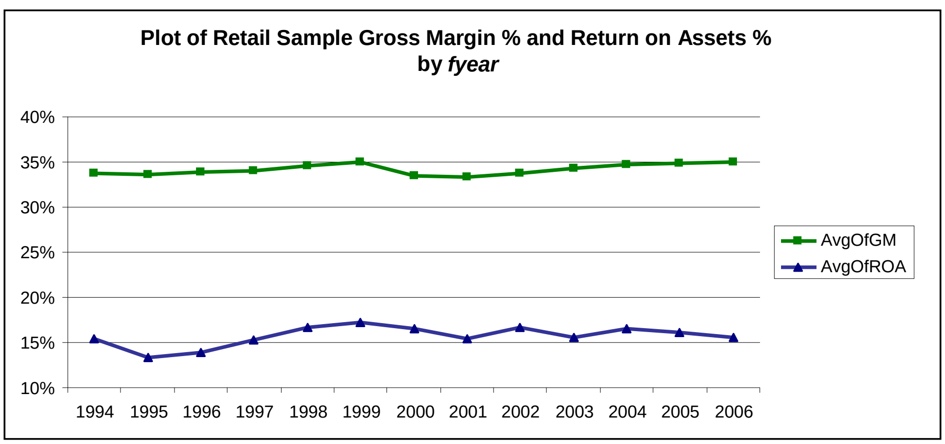
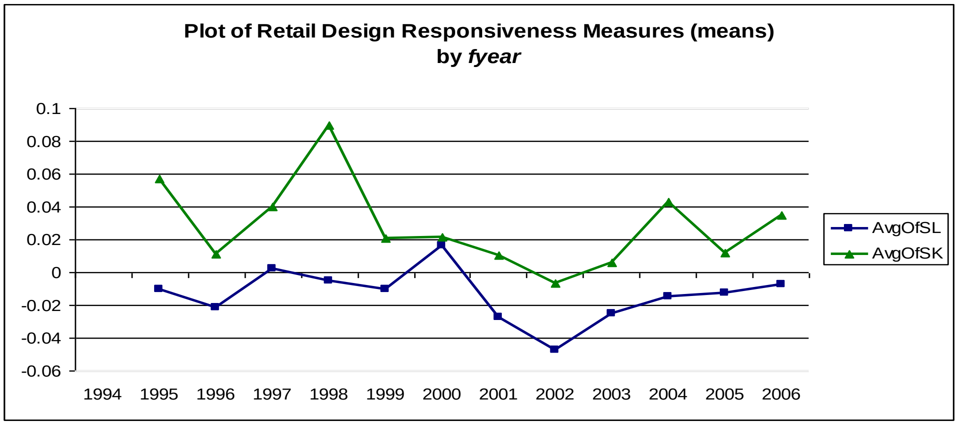


Figure 3: Plot of Design Responsiveness Measures by Fiscal Year



The descriptive data – mean, standard deviation, minimum and maximum value – was examined for all of the model variables and their various components (Table 5). This was done to check for any outliers in the data and to validate and empirically ground the variable calculations. Observations at either tail of the distribution were analyzed to verify if the result was "real" or that it is not the result of measurement error. As a result of this review, no additional observations were dropped or needed to be transformed. The average firm in the sample reported a mean ROA of 16% and an ROS of 7%. Annual firm sales averaged $5 billion dollars. Revenue and store growth averaged 11% and 10% per year respectively. Additional evidence of the heterogeneity of the firms within the retail trade industry is seen when examining store system labor and capital intensity, number of stores, and selling square feet.

Table 5: Descriptive Statistics

Model Descriptive Variables	Mean	S.D.	Min	Max	Obs.
Dependent Variables					
Return on Assets (ROA)	.16	.11	-.82	0.75	2039
Return on Sales (ROS)	.07	.05	-.23	0.26	2039
Firm-Specific Controls					
Sales[1] (S)	$5,830	$19,171	$12	$345,977	2039
Firm Revenue Growth ratio (RG)	1.11	.18	.34	2.92	2039
Firm Store Growth ratio (NG)	1.10	.32	.28	4.81	2039
Relative Inventory (Inventory/COGS) ratio (I)	.29	.18	.04	1.37	2037
Segment Controls					
Segment GM% (SM)	.34	.05	.25	0.56	2039
Segment Revenue Growth ratio (SG)	1.07	.09	.55	1.57	2039
Competitive Intensity (E)	1.52	.59	.08	2.8	2039
Component Measures					
Gross Margin% (GM)	.34	.10	-.05	.70	2039
Operating Income before Depreciation[1] (OIBD)	$443	1388	-728	$23,283	2039
Number of Employees[2] (EMP)	38.0	110.9	.1	39	2023
Number of Stores (N)	730	1177	7	8,079	2039
Labor Intensity[3] (L – adj.)	3.2	3.1	.1	39	2023
Capital Intensity[4] (K)	$130.21	$153.05	$2.72	$3,489.95	2038
Gross Selling Square Feet[2] (SQFT)	18,242	50,817	27	782,287	2039

1 $ millions
2 stated in 000's
3 Labor intensity multiplied by 1000 in table to aid interpretation (e.g. 3.2 employees/ thousand square feet, L=.0032emp/sqft)
4 Interpreted as $ of capital investment per selling square foot

The correlations of the model variables are listed in Table 6. ROS is highly correlated with ROA in the sample (r=0.85). The diagnostics and sensitivity analyses revealed that any of the segment control variables Segment GM% (SM), Segment growth (SG), or Competitive intensity (E) could be removed to reduce collinearity issues, without having any major impact on the coefficient results of interest (see Appendix A.1). All other descriptive statistics for the focal firm variables in Table 6 appear to be within the expected ranges.

Table 6: Correlations of Model Variables

	Variable	1	2	3	4	5	6	7	8	9	10	11	12
1	Return on Assets (ROA)	1											
2	Return on Sales (ROS)	.85	1										
3	SLinc - Increasing store labor responsiveness	-.02	-.02	1									
4	SLdec - Decreasing store labor responsiveness	-.10	-.08	-.12	1								
5	SKinc - Increasing store capital responsiveness	.00	-.00	.28	.04	1							
6	SKdec - Decreasing store capital responsiveness	-.19	-.15	-.06	.48	-.13	1						
7	Size (logS)	.16	.16	-.01	-.11	-.01	-.16	1					
8	Firm Revenue Growth (RG)	.42	.34	.29	.00	.32	-.16	.00	1				
9	Firm Store Growth (NG)	.23	.21	.05	.17	.16	-.05	.00	.54	1			
10	Relative Inventory (I)	-.20	.01	.00	.03	.03	.07	-.30	-.08	-.05	1		
11	Segment GM% (SM) **– dropped**	.04	.24	.00	.02	.04	.06	-.33	.04	.05	.56	1	
12	Segment Revenue Growth (SG) **– dropped**	.00	-.05	.02	.01	.05	.01	-.05	.14	.06	.05	.04	1
13	Competitive Intensity (E)	.14	.11	-.03	-.08	.00	-.01	-.10	.00	-.03	-.18	.06	-.05

7.2 Model Estimation using System GMM

According to the tests for serial dependence (see Online Supplement A.1), the ROA profit lag term is strongly associated with the dependent variable error term. As such, System GMM estimation provides some advantages over Difference GMM, particularly as the lagged dependent variable term coefficient becomes more persistent ($\boldsymbol{b^i} \rightarrow 1$). The panel data also has several characteristics that make the use of System GMM attractive. First, the model generalizes to the universe of retailers using N=232 selected retail firms with sufficient data, so System GMM avoids losing degrees of freedom. Second, Difference GMM may over-fit models by using more instruments than is necessary if the number of variables is high (as in this case). Finally, the dependent variables in the model have been shown to exhibit highly persistent properties in prior economics and operations management research (e.g. ROA).

Table 7: Model Estimation Using the System GMM estimator, DV=ROA

ROA	1	2	3	4	5	6
time lag t-1	0.37 **	0.43 **	0.60 **	0.60 **	0.49**	0.50**
	[3.21]	[4.01]	[7.43]	[7.40]	[4.28]	[4.38]
Firm						
SLinc	0.09	0.06	-0.05 **		-0.10**	
	[0.45]	[0.33]	[2.04]		[2.92]	
SLdec	0.38	0.64 **	0.08		0.09	
	[1.57]	[2.70]	[1.21]		[0.93]	
SL (non-directional)				-0.06 **		-0.11**
				[2.39]		[3.17]
SKinc	-0.27 **	-0.30 **	-0.04 **	-0.04 *	-0.01	-0.02
	[2.18]	[2.30]	[2.04]	[1.74]	[0.21]	[0.22]
SKdec	-0.43 **	-0.49 **	-0.13 **	-0.13 **	-0.24**	-0.25**
	[3.00]	[3.48]	[2.45]	[2.67]	[2.51]	[2.72]
Size (logS)	0.02	0.04	0.00	0.00	0.00	0.00
	[0.64]	[1.17]	[0.46]	[0.44]	[0.08]	[0.11]
Revenue Growth (RG)	0.62 **	0.58 **	0.18 **	0.18 **	0.25**	0.25**
	[5.39]	[5.89]	[5.13]	[5.15]	[3.80]	[3.82]
Store Growth (NG)	-0.11	-0.10	-0.03 **	-0.03 **	-0.04	-0.05
	[1.00]	[1.05]	[2.40]	[2.51]	[1.42]	[1.44]
Relative Inventory (I)	-0.29 *	-0.33 **	-0.09 **	-0.08 **	-0.17**	-0.17**
	[1.79]	[1.97]	[2.18]	[2.18]	[2.23]	[2.24]
Segment						
Competitive Intensity (E)	0.00	0.00	0.03 **	0.03 **	0.02	0.02
	[0.12]	[0.21]	[2.41]	[2.41]	[0.71]	[0.73]
Time						
Time dummies (included)	Yes	No	Yes	Yes	Yes	Yes
Constant	-0.46 *	-0.46 **	-0.13 **	-0.14 **	-0.12	-0.12
Observations	1784	1784	1784	1784	1784	1784
Number of Firms	226	226	226	226	226	226
Hansen Test (p-value)	.359	.411	1.00	1.00	.698	.710
Arellano-Bond AR(1)	-3.1 **	-3.5 **	-4.6 **	-4.5 **	-4.2**	-4.2**
Arellano-Bond AR(2)	1.6	0.8	-0.4	-0.4	0.5	0.5
F Test	13.2 **	19.8 **	26.7 **	28.0 **	22.0**	22.9**

System GMM estimates (Stata, xtabond2); the lag of dependent variable is endogenous; absolute value of t statistics are in brackets; robust standard errors;
Model 1 – 2 treat IVs as exogenous;
Model 3 – 4 treat IVs as follows: (I, logS - endogenous; NG, RG – predetermined);
Model 5 – 6 treat IVs as follows: (I, logS – endogenous);
One-tailed tests: * Significant at 10%; ** Significant at 5%

Table 7 reports the six alternative models using System GMM estimation. Models in column 1 and 2 treat the lagged dependent variable as endogenous and the rest of the independent variables as exogenous. The difference-in-Hansen, Hansen, and Arellano-Bond test statistics all support the model instrumentation. H2b, H3, and H4 are all statistically supported, while the other hypotheses were not supported. Next, the analysis was repeated for the *ROS* dependent variable and found that the results were consistent (see Online Supplement). Therefore, there is empirical evidence in Model 1 and 2 that decreasing store labor responsiveness is associated with better operational performance, and that either increasing or decreasing store capital responsiveness is associated with worse operational performance.

Column 3 and 4 models also use System GMM. However, additional instruments for the independent variables are included that are suspected of being either highly endogenous or predetermined. This is done to test the robustness of the findings to alternative model specifications. Some economics and operations management empirical research suggests that both inventory (*I*) and sales (*logS*) variables may be highly persistent (Ramey and West 1999, Rumyantsev and Netessine 2005, 2007b). In addition, the growth rate variables (store growth and sales growth) are specified as predetermined variables.[5] Although there is limited research using these techniques in operations management, it appears appropriate to instrument for these conditions by treating *logS* and *I* as endogenous and *NG* and *RG* as predetermined variables.

Column 3 and 4 report the results of the analysis using instruments for the specified endogenous and predetermined variables. Column 3 results show that while H1 is now supported in the respecified model, and H2b is not supported. Store capital intensity

[5] There is debate in the econometrics literature on how to instrument growth rates. Generally, growth rates are instrumented as predetermined variables, but no specific instances were found where retail revenue growth and store growth were used as proxy variables. However, this treatment appears reasonable based on GDP studies, etc.

responsiveness hypotheses (H3 and H4) continue to be supported. The model fit statistics also show improvement, as the AR(2) statistics have a serial correlation closer to 0, and the Hansen and difference-in-Hansen results both indicate acceptable instrumentation. Because the change in the store labor intensity responsiveness results was interesting, the model was re-specified (column 4) by substituting the directional variables for store labor intensity (*SLinc*, *SLdec*) for the base non-directional store labor intensity elasticity variable (*SL*). This produced a statistically significant and negative coefficient ($p<.05$), suggesting that increases in the variable resulted in negative financial operating performance. The control variable relationships to the DV also change in these models, as all coefficients except that for *logS* are statistically significant and directionally consistent with what is found in the literature. Since the Hansen test statistic was equal to 1.0 (Roodman 2008)[6], the sensitivity was examined using only the specified endogenous variables (Column 5 and 6). These models confirm the results for the store labor intensity responsiveness variables (H1 supported) and the decreasing store capital intensity responsiveness variable (H4 supported), but do not statistically support H3, which suggests that increasing store capital responsiveness worsens firm operational performance. Collectively, the evidence shows that many of the findings in base model, where the lagged DV is instrumented, are very sensitive once instruments for the other firm-specific control variables that may be endogenous are included. Therefore, the findings should be viewed with caution and within the context of how aggressively one specifies the firm-specific model control variables.

7.3 Forward Impact of Design Strategy Shifts

The final step of the analysis was to examine if any of the design responsiveness measures have carryover effects to the following period (*ROAF*). There is no evidence that any

[6] Roodman (2008) states while Hansen statistics of p=1.0 indicate acceptable instrumentation, these results should be viewed with caution and checked for sensitivity, as they may indicate that the model is over-specified.

responsiveness measures have particularly strong carryover effects on forward ROA (*ROAF*) – see Table 8. This finding indicates that the financial benefits of being design responsive are generally realized over the short-term timeframe, and that managing design responsiveness is an ongoing, year-to-year process for store retailers.

Table 8: Alternative Model Specification Using System GMM estimator for ROAF

Dependent Variable	ROAF	ROAF
Model (column)	1	2
time lag t-1	0.82 **	0.64 **
	[4.90]	[4.47]
Firm		
SLinc - Increasing store labor responsiveness	0.00	0.14
	[0.03]	[1.08]
SLdec - Decreasing store labor responsiveness	0.01	0.38
	[0.17]	[1.52]
SKinc - Increasing store capital responsiveness	-0.00	-0.12
	[1.18]	[1.51]
SKdec - Decreasing store capital responsiveness	0.08	0.15
	[1.33]	[0.80]
Sales (logS)	-0.00	0.02
	[0.18]	[0.96]
Revenue Growth (RG)	0.02	0.32 **
	[4.50]	[2.24]
Store Growth (NG)	-0.01	-0.22 **
	[0.93]	[2.37]
Relative Inventory (I)	-0.07 **	-0.11
	[2.21]	[1.06]
Segment		
Competitive Intensity (E)	0.00	0.02
	[0.63]	[0.49]
Time		
Time dummies (included)	Yes	Yes
Constant	.01	-.18
Observations	1562	1562
Number of Firms	220	220
Hansen Test (p-value)	1.00	.838
Arellano-Bond AR(1)	-4.6 **	-4.3 **
Arellano-Bond AR(2)	-0.46	1.1
F Test	93.1 **	21.1 **

System GMM estimates (Stata, xtabond2); the lag of dependent variable is endogenous; model treats IVs as follows: Column 1 = (I, logS - endogenous; NG, RG – predetermined); Column 2 = (all IVs exogenous); absolute value of t statistics are in brackets; robust standard errors; One-tailed tests: * Significant at 10%; ** Significant at 5%

[1]Sensitivity analysis for the ROS DV revealed that reported coefficient patterns were similar to those observed for an ROA DV.

8.0 Discussion

This study contributes to retail store research and practice in a number of ways. First, a statistical proxy for measuring store system design responses to gross margin changes in the retail trade industry for both labor and capital intensity is developed. Also, it provides an empirical model to analyze the impact of strategic design shifts on operational performance, while simultaneously controlling for other firm-specific, segment, and temporal effects. A major benefit of the analytical approach over other studies is that this model controls for the persistent effects of the dependent variable (e.g. ROA). Furthermore, the findings confirm that controlling for the persistence of *ROA* is both useful and necessary to fully understand the impact that relevant model variables are having on retail firm profits.

The first research question asked if it was possible to develop an empirical method to measure strategic store design shifts in the retailing industry. Drawing upon the inventory management studies of Rumyantsev and Netessine (2005, 2007b), four elasticity measures were constructed to examine the directional co-movements of strategic design changes with gross margins for store retailers. The results of this analysis not only validate findings from earlier studies on the impact of these control variables have on retail firm operating performance, but complement this previous work by demonstrating the importance of store labor and capital intensity design management to operational performance.

The second research question asked if retailers actually pursue responsive store design strategies. The descriptive analysis revealed that retail firms do not manage the co-movements of design strategy and margins as often as might be expected. While the mean responsiveness score for both store labor and capital intensity measures was centered on zero for all observations, there was a wide standard deviation and range of design responsiveness scores across the field sample (see online Supplement A.5). Directionally, the tendency for retail firms was to reduce store labor intensity and increase store capital intensity at a faster rate than gross margins. This finding provides some insight into the internal motivations of reported retail design strategic shifts towards greater cost-efficiency and economies of scale (Boyd and Bresser 2008).

For the final research question, four hypotheses were tested to discover the impact each of the measures would have on operational performance. The first series of hypotheses (H1, H2a, H2b) stated that increasing or decreasing store labor responsiveness measures would have negative performance impacts. The results suggests that that when one instruments for the possible endogenous and predetermined variables, increasing store labor intensity responsiveness does have a negative short-term effect on retail operating performance (Support H1). This finding is consistent with literature arguing that retailers are becoming more like Wal-Mart in that they are increasingly relying on low-contact/self-selection store environments, more automation, and supply chain management to deliver products to customers, regardless of margin changes (or in recognition that margins will continue to decrease). These findings also confirm the general importance of actively managing labor intensity to be efficient in retail store systems.

Generally strong support was found for H3 and H4 stating that both increasing and decreasing store capital responsiveness measures are associated with worse financial operating performance in retail firms. This finding suggests that retail store systems need to be deliberate and flexible in managing their property commitments in conjunction with product margin changes. This may also indicate that retail firms need to coordinate margin shifts tightly with capital planning and forecasting efforts.

Finally, increasing or decreasing responsiveness had no carryover effect on future financial operating performance. Because none of the design responsiveness measures have any association with forward ROA (*ROAF*), it appears that the financial benefits of design responsiveness are only realized in the short-term. This indicates managers should monitor the responsiveness measures on an ongoing basis.

It is also important to point out the limitations of this study. While the results are fairly robust to different panel data analysis techniques that control for the lagged dependent variable term (*ROA* or *ROS*), they are very sensitive to excluding or changing the assumptions about the endogeneity of different firm-specific control variables. No evidence of significant endogeneity in the model diagnostic tests was found, but it is possible that the large number of independent variables in this study could bias those tests. Nevertheless, each control variable was chosen because of its documented effects on retail operational performance, labor, and capital is grounded in the extant literature and the differences that changing the assumptions about variable endogeneity was tested for sensitivity and reported. Nevertheless, more work needs to be done to establish the degree of endogeneity among retail sales growth, inventory management, and margin-related variables. The findings are also limited because of how certain model proxy variables were formulated. Store capital intensity ratio, for example, may include investments in technology, store locations, fixtures, warehouses, or other items. It is difficult to determine if these items collectively influence the results, or if only certain capital should be managed responsively. However, it might be possible to separate each of these capital items into separate responsiveness variables in future work if one could get access to more detailed capital data than is typically reported in company financial statements.

9.0 Conclusions and Future Applications

Both the academic and investment analyst community suggest a deficiency in the area of retail store system design strategy measurement and theory. In this research, a statistical means for practitioners, industry analysts, and academics to evaluate the effectiveness of strategic store design shifts in retailing on financial accounting returns is created. Strong support is found that increasing/decreasing store capital intensity responsiveness measures are associated with worse operational performance in most cases. For increasing/decreasing store labor intensity responsiveness measures, decreasing store labor intensity responsiveness year to year may have positive financial benefits, and increasing labor intensity responsiveness may have negative associations with ROA. None of these responsiveness measures appear to have statistically significant carryover effects on ROA (*ROAF*).

By developing a design responsiveness measurement and modeling approach, this study offers a new way to evaluate the performance of store system design strategies year to year using publicly available data. In contrast, more traditional measures used to evaluate retail design performance, such as same store sales, sales or gross margin per square foot, and profit per store, do not indicate the important strategic shifts, incorporate customer contact implications, or evaluate capital resource investment decisions that are critical to understanding firm financial performance in dynamic store retailing environments (Gage, *Forbes* 2007).

In conclusion, this paper improves both practitioner and academic understanding of the dynamics of retail design strategy shifts and its effects on operational performance. By focusing on the retail trade industry, it concentrates its performance measurement efforts on retail responsiveness metrics that are of direct relevance to retail design strategy. In developing an empirical means to show how customer contact can be managed through design decisions regarding store capital and labor intensity management, it provides a direct link to evaluate how retail design strategy responses affect financial operating performance in an ever important industry in the U.S. economy.

REFERENCES

Arellano, M., S. Bond. 1991. Some tests of specification for panel data: Monte Carlo evidence and an application to employment equations. Review of Economic Studies, 58, 277-297.

Arellano, M., O. Bover. 1995. Another look at the instrumental variables estimation of error component models. Journal of Econometrics, 68, 29-51.

Baltagi, B. H. 2005. Econometric Analysis of Panel Data. John Wiley and Sons, Ltd., New York

Barber, B. M., J.D. Lyon. 1996. Detecting abnormal operating performance: The empirical power and specification of test statistics. Journal of Financial Economics, 41, 359-399.

Baum, C.F. 2006. An Introduction to Modern Econometrics Using STATA. STATA Press, College Station, TX.

Boyd, J.L., R.K.F. Bresser. 2008. Performance implications of delayed competitive responses: Evidence from the U.S. retail industry. Strategic Management Journal, 29, 1077-1096.

Boyer, K.K., R. Hallowell, A.V. Roth. 2002. E-services: Operations strategy – a case study and a method for analyzing operational benefits. Journal of Operations Management, 20, 175-188.

Buzacott, J.A. 2000. Service system structure. International Journal of Production Economics, 68, 15-27.

Chase, R.B., D.E. Bowen. 1991. Service quality and the service delivery system. In: Service Quality: Multidisciplinary and Multi-National Perspectives. Lexington Books, Lexington, MA 157-178.

Chase, R.B., D.A. Tansik. 1983. The customer contact model for organization design. Management Science, 29, 1037-1050.

Chase, R.B. 1981. The customer contact approach to services: Theoretical bases and practical extensions. Operations Research, 21, 98-105.

Chase, R.B. 1978. Where does the customer fit in a service operation? Harvard Business Review, 56, 137-142.

Chen, H., M.Z. Frank, O.Q. Wu. 2007. U.S. retail and wholesale inventory performance from 1981 to 2004. Manufacturing and Service Operations Management, 9, 430-456.

Cheng Q. 2005. What determines residual income? The Accounting Review, 80, 85-112.

DeHoratius, N., A. Raman. 2007. Store manager incentive design and retail performance: An exploratory investigation. Manufacturing and Service Operations Management, 9, 518-534.

Fisher, M.L., J. Krishnan, S. Netessine. Labor planning, execution, and financial performance: the exploratory investigation in the retail industry. The Wharton School, University of Pennsylvania, Working Paper, February, 2009.

Fisher, M.L., J. Krishnan, S. Netessine. 2006. Retail store execution: An empirical study. Knowledge@Wharton, Operations and Information Management Department, The Wharton School, University of Pennsylvania, Published December, 2006.

Fisher, M.L., A. Raman. 2001. Introduction to focused issue: Retail operations management. Manufacturing and Service Operations Management, 3, 189-190.
Fisher, M.L., K. Ramdas, K., Y.-S.Zheng. 2001. Ending Inventory Valuation in Multiperiod Production Scheduling. Management Science, 47, 679-692.
Forester, B. 2002. “What killed Service Merchandise?” Nashville Business Journal, January 14, 2002, 1.
Fornell, C. 2007. The Satisfied Customer: Winners and Losers in the Battle for Buyer Preference. Palgrave Macmillan, New York.
Froehle, C.M., A.V. Roth. 2004. New measurement scales for evaluating perceptions of the technology-mediated customer service experience. Journal of Operations Management, 22, 1-21.
Gage, J. 2007. “The fortunes of retailers can turn on a dime. We offer some unusual metrics for investors to stay a step ahead.” Forbes, August 13, 2007.
Gaur, V., M.L. Fisher, A. Raman. 2005. An econometric analysis of inventory turnover performance in retail services. Management Science, 51, 181-194.
Gaur, V., M.L. Fisher, A. Raman. 1999. What explains superior retail performance? Archive@NYU OM-2005-3, Published October, 1999.
Goldstein, S.M., R. Johnston, J. Duffy, J. Rao. 2002. The service concept: the missing link in service design research? Journal of Operations Management, 20, 121-134.
Gottfredson, M., K. Aspinall. 2005. Innovation versus complexity: what is too much of a good thing? Harvard Business Review, 83, 62-71.
Greene, W. 2000. Econometric Analysis, 4th Ed. Prentice-Hall, Upper Saddle River, NJ.
Hansen, L. 1982. Large sample properties of generalized method of moments estimators. Econometrica, 50, 1029-1054.
Heskett, J.L., W.E. Sasser, W.L. Hart. 1990. Service Breakthroughs. Free Press, New York.
IBM Corporation. 2005. Guided selling for complex products. IBM Corporation Products for the Retail Industry, http://ibm.com/industries/retail:1-2.
ifM and IBM. 2007. Succeeding through service innovation: A discussion paper. University of Cambridge Institute for Manufacturing, Cambridge, U.K.. ISBN: 978-1-902546-59-8.
Jayanthi, S., A.V. Roth, M.M. Kristal, L.C.R. Venu. 2009. Strategic resource dynamics of manufacturing firms. Forthcoming, Management Science manuscript #MS-00775-2006.R2.
Kellogg, D.L. and Chase, R B. 1995. Constructing an empirically derived measure for customer contact. Management Science, 41, 1734-1749.
Kennedy, P. 2003. A Guide to Econometrics. MIT Press, Cambridge, MA.

Kesavan, S., V. Gaur, A. Raman. 2008. Incorporating price and inventory endogeneity in firm-level sales forecasting. Working Paper, Under Journal Revision.
Kremer, C, R. Rizzuto, with J. Case, J. 2000. Managing by the Numbers: A Common Sense Guide to Understanding and Using Your Company’s Financials. Perseus, Cambridge, MA
Lal, R., C.I. Knoop, I. Tarsis. 2006. Best Buy Co., Inc.: Customer-centricity. Harvard Business School Case #9-506-055.
McGahan, A.M., M.E. Porter. 2002. What do we know about variance in accounting profitability? Management Science, 48, 834-851.

McGurr, P.T., S.A. DeVaney. 1998. Predicting business failure of retail firms: An analysis using mixed industry models. Journal of Business Research, 43, 169-176.

Menor, L. J., A.V. Roth, C.H. Mason. 2001. Agility in retail banking: A numerical taxonomy of strategic service groups, Manufacturing and Service Operations Management, 3, 273-292.

Oei, R., A. Ramsay, P. Mather. 2008. Earnings persistence, accruals and managerial share ownership. Accounting and Finance, 48, 475-502.

O'Donnell, J. "Electronics retailers find service sells" USA Today, July 22, 2008, D1.

Palepu, K., 1985. Diversification strategy, profit performance and entropy measure. Strategic Management Journal, 6, 239–255.

Peterson, P.P., F.J. Fabozzi. 2006. Analysis of Financial Statements, 2nd Ed. Wiley, Hoboken, NJ.

Plummer, L. 2007. Good neighbors or bad? A Time-space model of the effects of entrepreneurial entry on the profits of post-IPO firms. Doctoral Dissertation, Leeds School of Business, University of Colorado.

Randall, T.R., R.M. Morgan, A.R. Morton. 2003. Efficient versus responsive supply chain choice: An empirical examination of influential factors. Journal of Production and Innovation Management, 20, 430-443.

Randall, T.R., K. Ulrich. 2001. Product variety, supply chain structure, and firm performance: Analysis of the U.S. bicycle industry. Management Science, 47, 1588-1604.

Ramey, V., K. West. 1999. Ch. 13: Inventories. In The Handbook of Macroeconomics, eds. J.B. Taylor and M. Woodford. Elsevier Science, Amsterdam 863-923.

Roberts, P.W. 2001. Innovation and firm-level persistent profitability: A Schumpeterian framework. Managerial and Decision Economics, 22, 239-250.

Roodman, D. 2008. A Note on the Theme of Too Many Instruments. Working Paper No. 125, Center for Global Development, Washington, DC. Forthcoming in the Oxford Bulletin of Economics and Statistics.

Roodman, D. 2006. How to Do Xtabond2: An Introduction to "Difference" and "System" GMM in STATA. Working Paper No. 103, Center for Global Development, Washington, DC.

Roth, A.V., L.J. Menor. 2003. Insights into service operations management: A research agenda. Production and Operations Management, 12, 145-164.

Roth, A.V., W.E. Jackson. 1995. Strategic determinants of service quality and performance: Evidence from the banking industry. Management Science, 41, 1720-33.

Roth, A.V., M. van der Velde. 1991. Operations as marketing: A competitive service strategy. Journal of Operations Management, 10, 303-328.

Rumyantsev, S., S. Netessine, 2007a. What can be learned from classical inventory models: A cross-industry exploratory investigation. Manufacturing and Service Operations Management, 9, 409-429.

Rumyantsev, S., S. Netessine. 2007b. Linking Inventory Management and Profitability. Paper and Presentation at the 2007 INFORMS Annual Meeting, Seattle, WA, November 5, 2007.

Rumyantsev, S. and Netessine, S. 2005. Should inventory policy be lean or responsive? Evidence for U.S. public companies. Knowledge@Wharton, University of Pennsylvania, December 1, 2005: 1-32.

Sargan, J. 1958. The estimation of economic relationships using instrumental variables. Econometrica, 26, 393-415.

Soteriou, A., S. Zenios. 1999. Operations, quality, and profitability in the provision of banking services. Management Science, 45, 1221-1238.
Voss, C., A.V. Roth, R.B. Chase. 2008. Experience, service operations strategy, and services as destinations: Foundations and exploratory investigation. Production and Operations Management, 17, 247-266.
Waring, G. F. 1996. Industry differences in the persistence of firm-specific returns. American Economic Review, 86, 1253-1265.
Wernerfelt, B. 1984. A resource-based view of the firm. Strategic Management Journal, 5, 171-180.
Xue, M., L.M. Hitt, P.T. Harker. 2007. Customer efficiency, channel usage and firm performance in retail banking. Manufacturing and Service Operations Management, 9, 535-558.

APPENDIX A – MODEL DIAGNOSTIC TESTING & VALIDATION ANALYSES

A.1 Tests for Serial Dependence (Autocorrelation of Errors) in the Dependent Variable

Serial dependence of the dependent variable can be tested by using any of a wide range of tests (e.g. Baltagi, 2003, pp. 81-102, discusses many of these tests including the popular Lagrange multiplier (LM) and Likelihood ratio (LR) tests). We use an easy (and robust) test for serial correlation recommended by Drukker (2003), who discusses Wooldridge's (2002) method for testing serial autocorrelation by "using the residuals from a regression in first differences" of the specified model (p.169). Drukker's (2003) simulation analysis finds that Wooldridge's (2002) test for serial correlation removes individual effects by taking the first differences of the model, and then it compares the correlation of the differenced error term to the lagged differenced error term. Wooldridge's reports that, if the error terms of the dependent variable "are not serially correlated," then the "coefficient on the lagged residuals" should be equal to -.5 (Drukker, 2003, p.169). Like the Durbin-Watson test statistic and the Breusch-Godfrey LM test statistic, this method tests for autocorrelation in the model under the null hypothesis of no autocorrelation. The f-test is executed in STATA as follows:

. xtserial *variablename*

Wooldridge test for autocorrelation in panel data
H0: no first order autocorrelation

Dependent Variable	F (df) =	Prob > F =
ROA	309.13 (df=219)	.000
ROS	244.59 (df=219)	.000

A.2 Collinearity Diagnostics

Collinearity is assessed using the "collin" procedure in STATA following the procedures discussed in Plummer (2007, p.83) and Baum (2006, p.85). The table statistics are explained as follows: R-squared (R-Sq) is the independent variable regressed on the other independent variables, the "tolerance" value equals one minus the reported r-squared, the variance inflation factor (VIF) equals the reciprocal of the tolerance, and the model condition index is the square root of the ratio of the largest to smallest eigenvalue in the model matrix. A condition index >30 or a VIF > 10 are often used as cutoffs.

STATA: 'collin independent variables'

Variable	VIF	SQRT VIF	Tolerance	R-Sq	Order	Condition Index
SLinc	1.20	1.09	0.83	.16	1	2.5
SLdec	1.43	1.20	0.69	.30	2	2.9
SKinc	1.21	1.10	0.82	.17	3	3.4
SKdec	1.41	1.19	0.70	.29	4	4.7
Log'S"	1.22	1.10	0.82	.17	5	5.4
RG	1.77	1.33	0.56	.43	6	8.4
NG	1.52	1.23	0.65	.34	7	13.4
I	1.69	1.30	0.59	.40	8	19.1
SM	1.63	1.28	0.61	.39	9	25.8
SG	1.05	1.02	0.95	.04	10	28.6
E	1.12	1.06	0.89	.10	11	58.6

*****Final Iteration[1]

Variable	VIF	SQRT VIF	Tolerance	R-Sq	Order	Condition Index
SLinc	1.19	1.09	0.84	.16	1	2.1
SLdec	1.43	1.20	0.69	.30	2	2.6
SKinc	1.21	1.10	0.82	.17	3	3.0
SKdec	1.41	1.19	0.71	.29	4	4.1
Log'S'	1.17	1.08	0.85	.15	5	4.7
RG	1.72	1.31	0.58	.41	6	7.5
NG	1.52	1.23	0.65	.34	7	11.7
I	1.18	1.09	0.84	.15	8	20.1
E	1.08	1.04	0.92	.07	9	29.4

[1]Removal of IVs for SM and SG were shown to exhibit collinearity, so we respecified the model without these most problematic variables and it showed acceptable properties. The sensitivity analysis revealed that including any of industry segment variable had no impact on the design responsiveness coefficients of interest.

A.3 Panel-level Heteroscedasticity Diagnostics

Greene (2003, pp.230-232) reports several methods to test for normality of errors in panel data, including the Likelihood ratio (LR), Lagrange multiplier (LM), and the Wald test (p.230). The Wald test is a particularly easy and robust procedure to run in STATA. The standardized Wald test statistic (Greene, 2003) tests the null hypothesis of homoscedasticity of errors by comparing the maximum likelihood results of two covariance matrices of data (e.g. homoscedastic versus heteroscedastic error structures would be compared). A rejection of the null hypothesis of homoscedasticity, suggests that the data is heteroscedastic and that robust estimation of errors (and related Hansen test statistics) is needed to adjust for scalar differences in the data structure. We run a series of linear models in STATA, followed by the 'xttest3' postestimation:

Modified Wald Statistic[1]
STATA postestimation command: 'xttest3'

Cross-sectional time-series generalized least squares regression:
xi: xtgls ROA SLinc SLdec SKinc SKdec logS xRG NG xI xE i.fyear

Modified Wald test for groupwise heteroscedasticity
in cross-sectional time-series FGLS regression model

H0: sigma(i)^2 = sigma^2 for all i

chi2 (226) = 1.6e+05
Prob>chi2 = 0.0000

Fixed-effects (within) regression:
xtreg ROA Slinc Sldec Skinc Skdec 41ogs xRG NG xI xE, fe

Modified Wald test for groupwise heteroscedasticity
in fixed effect regression model

H0: sigma(i)^2 = sigma^2 for all i

chi2 (226) = 4.0e+32
Prob>chi2 = 0.0000

[1] Both models show strong evidence of heteroscedasticity across panels as p<.01 (Baum, 2006, p.222)

A.4 Dependent Variable Stationarity* Diagnostics

A key assumption of System GMM is that of the dependent variable is stationary (Plummer 2007, p.82; Baltagi, 2001, p.143). The dependent variable is assumed to be stationary if the mean, distribution, and variance do not change over time periods (Plummer 2007, p.82). Following Plummer's (2007, pp.82-83) and Baltagi's (2001, pp. 235, 240) guidelines, we report Fisher's test, which was further developed by Maddala and Wu (1999, p.636), to test for stationarity of each dependent variable. Fisher's test procedure examines summed log of individual panel unit root tests (p-values) and combines them into a common test statistic using the Fisher test command "xtfisher" in STATA. This statistic tests the null hypothesis of non-stationarity, and works well with unbalanced panel designs (Maddala and Wu 1999, pp.636-637). We run this procedure for each of the dependent variables used in our model and find evidence (see Table below) that they are all stationary across time periods (p<.01).

STATA: xtfisher *dependentvariable*, drift lags(1)

Fisher Test for panel unit root using an augmented Dickey-Fuller test (1 lags)
Ho: unit root is non-stationary[1]

Dependent Variable	X^2	Prob > X^2
ROA	905.92(df=344)	.000
ROS	881.96(df=344)	.000

[1]Null (Ho) of non-stationarity is rejected in all cases.

Table A.5: Distribution Statistics for Baseline Design Responsiveness Variables

Variable	*Mean*	*Std. Dev.*	*Obs*[1] *(Δt)*	*5%*	*25%*	*50%*	*75%*	*95%*
SL (labor)	-0.01	0.17	1793	-0.16	-0.06	-0.02	0.02	0.11
SK (capital)	0.03	0.27	1803	-0.17	-0.05	0.001	0.02	0.11

[1] The number of reported observations represents a loss of one degree of freedom to calculate the elasticity variable

Table A.6: Alternative Model Specifications Using System GMM estimator, DV=ROS[1] & ROAF

Dependent Variable	ROS	ROAF
Model (column)	1	2
time lag t-1	0.60 **	0.64 **
	[4.90]	[4.47]
Firm		
SLinc - Increasing store labor responsiveness	0.01	0.14
	[0.25]	[1.08]
SLdec - Decreasing store labor responsiveness	0.23 **	0.38
	[1.96]	[1.52]
SKinc - Increasing store capital responsiveness	-0.11 **	-0.12
	[1.97]	[1.51]
SKdec - Decreasing store capital responsiveness	-0.14 **	0.15
	[2.41]	[0.80]
Sales (logS)	0.00	0.02
	[0.35]	[0.96]
Revenue Growth (RG)	0.15 **	0.32 **
	[4.50]	[2.24]
Store Growth (NG)	0.01	-0.22 **
	[0.41]	[2.37]
Relative Inventory (I)	-0.06	-0.11
	[1.51]	[1.06]
Segment		
Competitive Intensity (E)	-0.03	0.02
	[0.02]	[0.49]
Time		
Time dummies (included)	Yes	Yes
Constant		
Observations	1784	1562
Number of Firms	226	220
Sargan / Hansen Test (p-value)	.334	.838
Arellano-Bond AR(1)	-4.5 **	-4.3 **
Arellano-Bond AR(2)	0.67	1.1
F Test	19.6 **	21.1 **

System GMM estimates (Stata, xtabond2); the lag of dependent variable is endogenous; all the independent variables entered as exogenous; absolute value of t statistics are in brackets; robust standard errors;
One-tailed tests: * Significant at 10%; ** Significant at 5%

[1]Sensitivity analysis for the ROS DV revealed that reported coefficient patterns were similar to those observed for an ROA DV.

APPENDIX B: SENSITIVITY ANALYSIS

Summary of Sensitivity Analysis

While there are no established rules for conducting sensitivity analysis with dynamic panel data models, it is important to acknowledge how the findings react to inputting different variables or making certain assumptions when modeling the data (Plummer 2007, Kennedy 2003). We constructed a STATA "Do-file" and examined a series of alternative model specifications (A '.pdf' of the Log-file is available upon request) for each analytical approach and dependent variable (*ROA, ROAF, ROS*) used in the paper tables. First we examined a series of liner (non-dynamic) pooled and fixed-effect regression models. We find that these models generally confirm our findings on the negative performance impact of increasing store labor intensity responsiveness (*SLinc*) and also for both store capital intensity responsiveness measures (*SKinc* & *Skdec*). We also find that the firm-specific results are robust when controlling segment membership – e.g. the results hold up when segment effects are included (using the dummy variable code *i.Seg* and the xi: command in STATA).

In general, our findings related to the elasticity measure of store labor responsiveness (*SL*) are robust to different model specifications provided that firm-specific control variables for size (*logS*), sales growth (*RG*), store growth (*NG*) and inventory management (*I*) and store capital intensity responsiveness (*SKinc* & *SKdec*) are all included. As discussed in the text of the paper, the decision to include any or all of the segment control variables really has no bearing on any of the dynamic model results where ROA is the DV. This may be because the lagged dependent variable used in each equation term is incorporating much of the random error in the firm's segment. This finding was useful because it allowed us to remove any or all of these three variables to reduce model multicollinearity issues without affecting the model.

However, we do find that the results (particularly for increasing store labor and capital intensity responsiveness) become quite sensitive once we start instrumenting for any endogeneity and persistence bias in several firm-specific control variables, particularly those related to Size (logS) – for store labor intensity responsiveness - and revenue (*RG*) and store growth (*NG*) rates – for store capital intensity responsiveness. This is not particularly surprising given the role that each variable plays in firm staffing and capital planning models for populating stores in new locations, or in providing specific revenue support for revenue planning and capital management efforts. It is also clear that being overly responsive with capital intensity (*SKinc*) is not as robust, and does not have nearly the same negative performance effect as does under investing in store capital (*SKdec*) year to year. This finding may speak to both the importance to profits of a retail firm's ongoing investment in good store locations and the short-term significance of in-store capital investment in internal store systems.

Finally, we examined the alternative dependent variables (*ROS*) and (*ROAF*) to see if using different performance measures made a difference in interpreting the model findings and to find out if there were carryover affects on forward profits. The results for the *ROS* models largely mirror the findings using *ROA*, except that the findings for both store labor and capital responsiveness (*SLinc*, *SLdec*, *SKinc*, *SKdec*) are weaker ($p<.10$ or greater) in the ROS models. As reported in the paper text, for forward ROA (*ROAF*), we see no significant ($p<.05$) results for any of our design responsiveness measures.

Supplemental Citations for the Appendix (Not in Paper Text):

Drukker, D.M. 2003. Testing for serial correlation in linear panel-data models. *The STATA Journal* 3 168-177.

Maddala, G.S., S. Wu. 1999. A comparative study of unit root tests with panel data and a new simple test. *Oxford Bulletin of Economics and Statistics* 61 631-652.

Sargan, J. 1958. The estimation of economic relationships using instrumental variables. *Econometrica* 26 393-415.

Wooldridge, J.M. 2002. *Econometric Analysis of Cross-section and Panel Data*. MIT Press, Cambridge, MA.