A LONGITUDINAL ANALYSIS OF THE MOTOR CARRIER INDUSTRY—THREE ESSAYS

By

William A. Muir

A DISSERTATION

Submitted to
Michigan State University
in partial fulfillment of the requirements
for the degree of

Business Administration – Logistics – Doctor of Philosophy

2017

ABSTRACT

A LONGITUDINAL ANALYSIS OF THE MOTOR CARRIER INDUSTRY—THREE ESSAYS

By

William A. Muir

Motor carriage is the backbone of supply chains and a major component of the U.S. economy. The industry has endured remarkable levels of external pressure over the last three decades, including major shifts in shipper expectations, volatility and uncertainty in factor markets and changing governmental regulations. These pressures continue to transform the industry. This dissertation examines the motor carrier industry from three distinct perspectives: as an economic agent, as a productive entity and as a principal to public safety.

Essay 1 examines longitudinal change in industry-level aggregate inventories, linking inventory investment to prior-period changes in freight cost. A dynamic model is specified and, using economic indicators and time-series inventory data on forty-five industry sectors, polynomial distributed lags are estimated. The results show that aggregate inventories adjust in the directions expected under optimal lot size ordering, however these adjustments are protracted over several quarters.

Essay 2 links the diversification of motor carrier production activities within the truckload and less-than-truckload segments to operational efficiency. The effect of output diversification on carrier utilization of productive resources is theorized to depend on the degree of interdependence among work flows. A panel is developed with five years of archival data and two-stage data envelopment analysis is used for estimation and theory testing. Results are consistent with theory but also suggest structural differences between carrier classes.

Essay 3 draws on theories of organizational routines to link variability in motor carrier safety performance over time to longitudinal change in performance. A metaroutine, International Organization for Standardization 9000, is hypothesized to simultaneously reduce variability and catalyze change. A regulating effect on variability is supported by a four-year panel of federal safety data and is conditioned by a carrier's average level of safety performance, such that unsafe carriers appear to benefit most from the metaroutine.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

KEY TO ABBREVIATIONS

AIC	Akaike Information Criterion
BASIC	Behavior Analysis and Safety Improvement Category
CFR	Code of Federal Regulations
CPI	Consumer Price Index
CSA	Compliance, Safety, Accountability
DEA	Data Envelopment Analysis
FMCSA	Federal Motor Carrier Safety Administration
HOS	Hours of Service
ICC	Intraclass Correlation Coefficient
ISO	International Organization for Standardization
L-Q	Linear-Quadratic
LTL	Less than Truckload
M3	Manufacturers' Shipments, Inventories and Orders Survey
MCA	Motor Carrier Act of 1980
MRP	Material Requirements Planning
MRTS	Monthly Retail Trade Survey
MWTS	Monthly Wholesale Trade Survey
NAICS	North American Industry Classification System
OMCI	Office of Motor Carrier Information
PPI	Producer Price Index
RFID	Radio-Frequency Identification
SE	Standard Error
TL	Truckload
UD	Unsafe Driving
VM	Vehicle Maintenance

ESSAY 1

LOT SIZE ADJUSTMENTS AND THE DYNAMIC UPDATING OF AGGREGATE INVENTORIES

1.1 Introduction

Business inventories are a critical component of the economy, serving as a lagging indicator of macroeconomic activity (The Conference Board, 2017) and moving procyclically with economic cycles (Khan and Thomas, 2007; Kesavan and Kushwaha, 2014). In the U.S., total business inventories exceeded 1.8 trillion dollars in 2016 (U.S. Census Bureau, 2017), a 30 percent increase from the October, 2006, levels reported by Shah and Shin (2007). At the firm-level, inventory performance is a key driver of customer-focused and financial outcomes. When compared against their peers, firms with poorer than average inventory performance exhibit inferior stock market performance (Chen et al., 2007; Alan et al., 2014). Excess inventories inflate holding expenses, risk product obsolescence and prompt price discounting (Kesavan and Kushwaha, 2014). Out-of-stocks impact service levels directly, resulting in sale deference, product substitution or defection (Musalem et al., 2010).

While normative inventory theory is frequently developed at the stock-keeping-unit level, recent empirical inventory research has greatly improved our understanding of the aggregate behaviors of inventories at the firm and industry levels. Empirical inventory models offer utility as a means to reconcile mathematical models with real-world behavior (Rumyantsev and Netessine, 2007), to investigate between-industry and between-firm heterogeneity in response to phenomena (Bray and Mendelson, 2012; Kesavan and Kushwaha, 2014) and to elucidate indicators of practical and theoretical importance (*e.g.*, Gaur et al., 2005; Eroglu and Hofer, 2011; Kesavan and Mani, 2013). As Rumyantsev and Netessine (2007) note, empirical models of aggregate inventory can directly benefit firm managers and industry analysts since these models may relax modeling assumptions frequently made in analytical models, such as stationarity and exogeneity.

In a recent empirical study, Larson et al. (2015) explore the dynamics of firm-level inventories and observe that current-period inventory investments are a function of prior-period sales growth. This observation is directly at odds with theory underlying the adaptive base stock policy which the authors seek to test (Larson et al., 2015). The delay that they observe in the adjustment of inventory investment to demand has, in fact, been noted elsewhere in the literature (see Bils and Kahn, 2000), but remains understudied. Indeed, researchers have yet to widely incorporate ordering dynamics within empirical inventory models. Nor has a dynamic response to many other key determinants of inventory investment been investigated within the literature.

The purpose of this paper is to further investigate the phenomena identified by Larson et al. (2015). Using a panel of aggregate inventories at the industry-level, we show that business inventories are elastic not only to prior-period changes in sales, but also to prior-period changes in costs associated with placing re-orders and holding inventory. Our results show that aggregate inventories adjust in the directions expected under optimal lot size ordering, however these adjustments are protracted over several quarters. For adjustments resulting from changes to physical inventory holding costs, the updating process is delayed for up to three months. Taken together, these results suggest that institutional and individual processes and biases attenuate timeliness in the acquistion and processing of information and the execution of replenishment orders. We posit that many of the biases underlying this protraction are the same as those identified in laboratory settings by Schweitzer and Cachon (2000) and subsequent studies of single-period newsvendor ordering behavior.

Further, this study shows that the predictive power of a model accounting for dynamic updating is remarkably stronger than that of a model considering only contemporaneous effects of regressors. Competing AR(1) models are estimated using a panel of time-series data from business inventories in the manufacturing, wholesaling and retailing sectors. After removal of seasonal effects, and after controlling for macroeconomic factors, contemporaneous effects of focal regressors (demand, re-order cost, inventory holding cost) explain less than one-tenth of the variation in year-over-year changes in aggregate inventory investments between 1993 and 2016. In contrast, under a dynamic specification with polynomial distributed lags up to twelve months, these same focal predictors explain nearly half of the variation in year-over-year changes in aggregate inventory investments during this same period. Impulse response functions suggest that inventory updating concludes within ten months; we theorize that strategic responses, which are not captured by the model, may extend further in time.

Implications are drawn for inventory managers, policy makers and researchers.

The remainder of the paper is structured as follows. First, the empirical inventory management literature is reviewed. A subsequent section introduces a normative theory of inventory updating behavior, which is followed by a discussion of organizational and individual factors that might attenuate behavior under the normative model. From these, research hypotheses are drawn. Section 4 provides an overview of the data and operationalization of variables. Econometric specifications are developed in Section 5 and results are presented in Section 6. The paper closes with implications for managers and a general discussion.

1.2 Literature Review

Explaining fluctuations in inventory investments remains critical to understanding the business cycle as well as firm behavior. Within the classical production-smoothing theory of inventory (Holt et al., 1960), inventories are held to smooth production under a concave short-run production function (*i.e.*, rising marginal costs), thereby improving average productivity and reducing average production costs. Arising from this theory is the linear-quadratic (L-Q) model, which has been regarded as the workhorse of the inventory literature (Banerjee and Mizen, 2006), especially as it applies to the finished goods inventories of manufacturers (Ramey and West, 1999). The L-Q model assumes that demand is linear and that the costs of production and holding inventory are quadratic. Adjustments to inventories are protracted because production costs are convex. In contrast, the flexible-accelerator variant of the L-Q model, developed by Lovell (1961), assumes that inventory adjustments incur some cost and, thus, firms close only a fraction of the gap between target and actual inventory holdings in any period. Estimates of this fraction suggest that inventories adjust quite slowly to target levels (Blinder, 1986; Bils and Kahn, 2000; Khan and Thomas, 2007); adjustment rates reported in the literature range from three months (Maccini and Rossana, 1981) to more than one year (Feldstein et al., 1976). Speculation as to the cause of slow adjustment under these models include the presence of institutional processes, warehousing and personnel constraints, risk-aversion under uncertainty, contractual inflexibilities, long lead times and the omission of relevant regressors (a thorough discussion is provided by

Feldstein et al., 1976).

Over time, many of the core tenants of production-smoothing theory have been shown to be untenable (Cachon et al., 2007). Under production-smoothing theory and the L-Q model, inventories should respond inversely to demand shocks (*i.e.*, covary negatively) and move countercyclically with the business cycle. These basic assumptions have been called into question. For instance, Blinder (1986) and Khan and Thomas (2007) show that variation in production exceeds that of sales and that the covariance between inventory investments and sales is, in-fact, positive—findings which are directly at odds with a production-smoothing motive. Further, numerous empirical studies have shown that inventory investments move procyclically with the business cycle and not countercyclically as would be expected by a pure production-smoothing motive (*e.g.*, Wen, 2005; Khan and Thomas, 2007; Kesavan and Kushwaha, 2014). In light of these shortcomings, significant efforts have been made to better align the L-Q model and production-smoothing theory with the so-called stylized facts of aggregate inventory behavior in practice. For example, Blinder (1986) argues that supply shocks might account for observations of greater volatility in inventories than in demand (*c.f.*, Wen, 2005). Banerjee and Mizen (2006) relax the assumption of the L-Q model that inventory and sales are linearly cointegrated and propose an I(2) model under the presence of polynomial cointegration. However, given the lack of evidence for a production-smoothing motive, and its narrow application, the extant empirical inventory research has tended to abandon the L-Q model in favor of alternative frameworks. Seminal work is summarized below.

Khan and Thomas (2007) consider inventory investment under generalized (S, s) inventory policies, in which nonconvex delivery (fixed) costs give rise to inventories. Application of this policy assumes that firms seek to minimize inventory costs, ordering some optimal lot size $S - s$ to return an inventory to S only when inventories are at a sufficiently low level (s). A key feature of this theory is that its application is not limited to a narrow subset of manufacturing inventories, as the production-smoothing motive is (Ramey and West, 1999), but may offer a general motive for business inventories including trade inventories held by wholesale and retail firms (Mosser, 1991). In their model economy, Khan and Thomas (2007) are able to reproduce many of the empirical findings that confound the production-smoothing motive, including procyclical inventory investments and positive covariation between inventories and demand. Importantly, Khan and Thomas (2007) provide further evidence that inventories adjust slowly to targets, taking multiple quarters to reach equilibrium in their model.

Gaur et al. (2005) and successive work (*e.g.*, Rumyantsev and Netessine, 2007; Cachon and Olivares, 2010; Larson et al., 2015) appeal to a broader class of management and inventory theory to develop empirical models of inventory behavior. Accordingly, this stream of research departs sharply from the comparatively narrow focus of the L-Q model, investigating a wider set of response variates and covariates. Gaur et al. (2005) demonstrate that a majority of variability in retail inventory turnover can be explained by between-firm differences and within-firm changes over time in three factors: gross margin, capital intensity and deviations of actual sales from sales forecasts. In explaining fluctuations in inventory investments, Rumyantsev and Netessine (2007) consider firm size differences as well as differences in margins, sales growth and demand uncertainty on inventory investments. They further consider inventory holding costs in their model using the three-month treasury bill rate as a proxy to represent the opportunity cost of holding capital, but find the effects of holding cost to be confounded by trend. Cachon and Olivares (2010) show that finished goods inventories of vehicle manufacturers scale with the increases to the downstream (dealership) network but decline with gains in production flexibility of plants. A commonality among this stream of literature is that, with only a few exceptions (*e.g.*, Kesavan and Kushwaha, 2014), inventory adjustments are modeled contemporaneously to covariate shocks.

This study extends the findings of Larson et al. (2015), who examine the behavior of aggregate inventories under an adaptive base stock policy. This policy relaxes an (S, s) policy assumption of independent and identically distributed demands. To examine the degree to which an adaptive base stock policy might explain inventory behavior in the practice, the authors model quarterly inventory investments—purchases scaled by same-period cost of goods sold—as a function of forecasted changes in sales. Underlying this model is a key assumption that inventory investments are not a function of sales growth from prior periods. Instead, Larson et al. (2015) find that greater sales growth in the previous quarter attenuates the relationship between demand and inventory investment. This finding strongly suggests that frictions exist over and above delivery lags that condition the timeliness of inventory investments. We extend the findings of Larson et al. (2015) by explicitely modeling the dynamics of inventory investment. Unlike previous work, we do not limit our dynamic analysis to the demand distribution or assume that the rate of inventory updating is constant over time (*c.f.*, flexible-accelerator model). Instead, we rely on normative theory of optimal ordering behavior and utilize a flexible functional form to investigate non-linear dynamic inventory response to multiple predictors.

1.3 Theory and Hypotheses

1.3.1 Economic Order Quantity

Like the adaptive base stock policy research of Larson et al. (2015) and the (S, s) policy research of Khan and Thomas (2007), we rely on normative theory of optimal ordering behavior to examine inventories in aggregate. Within classical inventory theory, the economic order quantity for an inventory minimizes the costs of re-ordering and holding inventory given some level of demand. For a lot of Q units and demand rate D, the total of these costs per unit of time (C) can be given as the sum of average costs of re-ordering (K) inventory per unit of time (KSQ^{-1}) and the average cost of holding (H) inventory per unit of time $(HQ2^{-1})$. An optimal order quantity (Q^*) is obtained by setting equal to zero the partial derivative of this cost function with respect to quantity, which yields $Q^* = \left(2DKh^{-1}\right)^{1/2}$. In this formulation, demand, re-order costs and holding costs are assumed stationary and exogenous. Given these assumptions, the expected quantity of on-hand cycle stock (Q_c) for an inventory observed at an arbitrary point in time $(\mathbb{E}\left[Q_c\right])$ is $.5Q^*$. For $[D, K, h] \subseteq \mathbb{R}_+$, $\mathbb{E}\left[Q_c\right]$ can be rewritten in a logarithmic specification—in a form that is linear in parameters and suggests optimal partial elasticities of inventory with respect to demand, transportation cost and holding cost:

$$ln\left(\mathbb{E}\left[Q_c\right]\right) = -ln(\sqrt{2}) + .5ln(D) + .5ln(K) - .5ln(h) \tag{1.1}$$

1.3.1.1 Demand for Inventory

According to Equation 1.1, a one percent increase in demand should result in an approximate[1] one-half percent increase in the expected quantity of cycle stock, holding fixed re-order cost and inventory holding cost constant. The association between demand and inventories has been the subject of intense investigation, however the relationship has been specified in log-log form only in a number of studies (Ballou, 1981; Eroglu and Hofer, 2011; Shan and Zhu, 2013; Kesavan and Kushwaha, 2014). The partial elasticity of total inventories with respect to demand depends on the magnitude of aggregate safety stocks held, but recent empirical estimates of this parameter compare quite favorably to the coefficient provided in Equation 1.1. For instance, the average partial elasticity reported by Kesavan and Kushwaha (2014) is .45 for inventories in the United States and Shan and Zhu (2013) report a partial elasticity of .65 for inventories in China.

H$_1$: *The partial elasticity of aggregate inventories with respect to aggregate demand is greater than zero.*

1.3.1.2 Inventory Re-ordering Costs

Similarly, a one percent increase in re-order costs should result in an approximate one-half percent increase in the expected quantity of cycle stock, holding demand and inventory holding cost constant. Blinder and Maccini (1991) cite the existence of fixed re-order costs as a primary piece of evidence against production-smoothing theory, but note a lack of research emphasis on models considering these costs. Indeed, this lack of empirical evidence remains almost three decades later, especially as it pertains to the cost of transporting

[1] For a data generating process given by $Y = X^\beta$, β serves as a reasonable approximation for the percent change in Y associated with a one percent change in X if β is small, as it is here. An exact calculation of the percent change in Y given θ percent change in X is $e^{\beta ln(1+.01\theta)}$.

inventory, arguably the most volatile cost component of re-ordering. The scant evidence available into this relationship can be found within the transportation literature and generally supports a positive association between transportation costs and inventories (Swanson et al., 2016), as would be expected under Equation 1.1. Swanson et al. (2016) demonstrate that a cointegrating relationship exists between aggregate transportation expenditures and aggregate inventory expenditures. They find that changes in transportation expenditures positively influence inventory expenditures, but find no support for a reciprocal relationship, suggesting exogeneity of the transportation cost component within Equation 1.1.

$\mathbf{H}_2$: *The partial elasticity of aggregate inventories with respect to re-order costs is greater than zero.*

1.3.1.3 Inventory Holding Costs

Lastly, Equation 1.1 suggests a negative partial elasticity of inventories with respect to inventory holding cost. When demand and transportation cost are held constant, a one percent increase in holding cost should result in an approximate one-half percent decrease in the expected quantity of cycle stock. Rumyantsev and Netessine (2007) identify two components of a firm's inventory holding cost: a physical cost of carrying inventory and an opportunity cost of committing capital. Firms with inadequate financial resources to carry inventory may also incur financing charges (increasing holding costs) or may place orders at less-than-optimal lot sizes (Buzacott and Zhang, 2004). The relationship between holding cost and inventories has received mixed support in the literature. Overwhelmingly, the real interest rate has been used to proxy for the cost of capital, albeit unsuccessfully (Jones and Tuzel, 2013). This lack of association is perplexing since standard theory of monetary policy predicts that changes to short-term interest rates should impact inventory investment spending (Maccini et al., 2004). In one of these few studies that find evidence of this relationship, Akhtar (1983) reports elasticities of inventory investments with respect to the real interest rate ranging from $-.09$ to $-.11$, directionally correct but lower in magnitude than would be expected under Equation 1.1. Chen et al. (2007) and Rumyantsev and

Netessine (2007) find an overall negative relationship, although neither study is able to produce consistent results across industry sectors. Maccini et al. (2004) find no short-term relationship and conclude instead that only a long-run relationship exists between inventories and the real interest rate, which they argue occurs between changes of persistent regimes. Larson et al. (2015) measure holding cost by risk of product obsolescence and demonstrate that firms alter their inventory ordering behavior when faced with greater obsolescence risk. Several studies, such as Rajagopalan (2013), have been unable to link retail inventories to cost of capital. A search of the inventory literature does not reveal an empirical test of the relationship between the physical cost of holding inventory and inventory investment behavior. For this reason, and because of a lack of consistent findings on other measures, the true overall effect of inventory holding cost on inventory behavior remains unclear.

H$_3$: *The partial elasticity of aggregate inventories with respect to holding costs is less than zero.*

1.3.2 Inventory Adjustment Speed

Normative theories of inventory management largely assume that decision makers act rationally to economize on cost (Equation 1.1) and to maximize the sum of discounted future cash flows (as seen in the objective function of the L-Q model[2]) thus maximizing profit (Ren and Croson, 2013). In reality, decision makers operate in uncertain and changing environments and their decisions are frequently influenced by the presence of institutional

[2]The L-Q model also assumes that economic agents will economize on cost, given that there is a trade-off between the costs of holding inventory and moving to a new level of production. This can be seen in the following objective function, where t is time, b is the discount rate, C is cost, Q is a product quantity, H is inventories, P is price, S is sales and U is a cost shock, a_0, a_1, a_2 are costs associated with changing production, producing and holding inventory, and a_3 is the inventory accelerator:

$$\max \lim_{T\to\infty} \mathbb{E}_t \sum_{j=0}^{T} b^j \left(P_{t+j}S_{t+j} - C_{t+j}\right)$$

$$\text{s.t.} \begin{cases} C_t &= 0.5a_0\Delta Q_t^2 + 0.5a_1Q_t^2 + 0.5a_2(H_{t-1} - a_3S_t)^2 + U_{ct} \\ Q_t &= S_t + H_t - H_{t-1} \end{cases}$$

and individual processes and biases. For example, when placing orders in the single-period newsvendor problem, decision makers frequently exhibit *pull-to-center bias*, anchoring order quantity on prior demand distributions (Schweitzer and Cachon, 2000; Bolton et al., 2012). One manifestation of this bias is a tendency to order at some level between that of mean demand and the optimal, profit-maximizing quantity. In multi-period settings, the L-Q model and flexible-accelerator variant assume that inventories adjust gradually toward target levels, making only a partial adjustment in any period. Under an (S,s) policy an adjustment is delayed at least until an inventory reaches its re-order point. While a certain degree of gradual adjustment may be theoretically admissible, the sluggishness with which aggregate inventories adjust in practice cannot be accounted for by normative models (Bils and Kahn, 2000). Further, the collective evidence suggests that influential processes and biases drive inventory behaviors that deviate from those expected of optimal-seeking organizations and agents.

The speed with which an individual or organization responds to new, important external information depends on their ability to perform a series of tasks including scanning the environment, processing or analyzing information and taking action (Thomas et al., 1993). For instance, an individual's cognitive inability to process information may account for some of the pull-to-center bias observed in newsvendor experiments (Bolton et al., 2012). For complex organizations, however, acting on new information typically infers that organizational mechanisms are in-place and available to transfer information from scanners to analyzers and from analyzers to action takers (Smith et al., 1991). These steps frequently involve multiple organizational actors, each of whom may bring their own set of individual biases. Inventory managers, for instance, may exhibit seemingly-irrational behavior such as avoiding adverse information, procrastinating, interpreting or acting on information inconsistently or overestimating their span of control (Gino and Pisano, 2008). These behaviors may slow or stop the acquisition and processing of information or delay action. It is not surprising, then, that speed of organizational response is not homogeneous across firms or within industries—factors such as firm size (Chen and Hambrick, 1995), external orientation (Smith et al., 1991) and capability to interpret information (Daft and Weick, 1984) may constrain the timeliness of response.

Despite significant advances made in supply chain management over recent decades, it remains unclear if the phenomenon of slow inventory updating has changed much, if at all, since it was popularized by Feldstein et al. (1976). For instance, inventory record inaccuracies continue to persist well-beyond the advent of computerized inventory management (DeHoratius and Raman, 2008). Musalem et al. (2010) contend that these innaccuracies have resulting in defacto-periodic review systems where accurate inventory information is available to firms only at discrete points on time. Managers may also choose not to follow automated ordering systems, due to perceived or actual inadequacies in the systems or a misalignment of incentives (van Donselaar et al., 2010). Further, the complexity of inventory management has increased—proliferation of stocking-keeping-units along with increased demands from businesses (*e.g.*, just-in-time deliveries, real-time tracking) and consumers (*e.g.*, service levels) have placed immense pressure on inventories and the personnel who manage them. Meanwhile, the repercussions of contractual inflexibilities with suppliers and third-party logistics firms are magnified under global sourcing arrangements. Domestically, increased fuel volatility in a now-deregulated trucking industry has led to speculation and risk-taking. Fuel surcharges visible to carriers on invoices lag diesel prices by several weeks (Smith, 2016), presenting conflicting informational clues to inventory managers. Taken together, evidence of individual biases in policy updating and the continued presence of these institutional factors suggest that the updating process should be protracted for aggregate inventories across industries.

$\mathbf{H}_4$: *Updates of aggregate inventories to target levels will be protracted over multiple periods.*

1.4 Data Sources and Variable Operationalization

1.4.1 Overview of Modeling Approach

An overview of the modeling approach is as follows. Contemporaneous and dynamic specifications of Equation 1.1 are modeled within the general linear mixed effect modeling framework. Dynamic effects are specified by constructing polynomial distributed lags up to twelve months. An industry-level panel of aggregate inventories and sales is developed using data from the U.S. Department of Commerce, Bureau of Economic Analysis. Re-ordering costs and inventory holding costs are measured using producer price indices. Controls are introduced to account for the common business cycle and other macroeconomic effects. After removing seasonality and time-invariant effects through seasonal differencing, year-over-year inventory investments are modeled as an AR(1) process.

1.4.2 Sample Selection

Consistent with prior research (*e.g.*, Rajagopalan and Malhotra, 2001; Cachon et al., 2007; Chen et al., 2007; Shah and Shin, 2007), we examine inventories aggregated to the industry-level. Among the advantages of industry-level data is generalizability of findings across industry sectors and the avoidance of small sample bias or idiosyncrasies that might accompany stock-keeping-unit level or firm-level panels. A potential disadvantage to the use of industry-level inventory data is the possbility that certain inventory-level effects might be masked or attenuated (Bray and Mendelson, 2012). For example, it is reasonable to assume that inventories will face both idiosyncratic shocks and structural (aggregate) shocks. Chen et al. (2007) compare industry-level inventory data gathered and reported by the U.S. Census Bureau, like that used in this study, to firm-level inventory data obtained through COMPUSTAT. They find inventory patterns between the two data sources to be "reassuringly similar" (Chen et al., 2007, p. 435). Industries were selected at the three and four digit levels of the North American Industry Classification System (NAICS), resulting in forty-four industry subsectors (Table 1.2). Wholesalers were defined at the four digit level

Table 1.1: Descriptive statistics.

Variable‡	Mean	S.D.	(1)	(2)	(3)	(4)	(5)	(6)	(7)
(1) Inventories	.03	.09	1.00						
(2) Recession probability	.00	.29	.24**	1.00					
(3) Consumer price index	.02	.01	.32**	.32**	1.00				
(4) Trade balance	.07	.20	.25**	−.09**	.47**	1.00			
(5) Establishments	.00	.02	.27**	.12**	−.01	.02†	1.00		
(6) Demand	.03	.11	.66**	−.01	.36**	.45**	.17**	1.00	
(7) Reordering cost	.04	.03	.32**	.28**	.69**	.52**	.04**	.33**	1.00
(8) Holding cost	.00	.02	−.04**	.25**	.21**	.02†	.04**	−.10**	−.06**

Notes: †$p < .10$, *$p < .05$, **$p < .01$ (two-tailed); ‡all variables except *recession probabilities* log-transformed and all variables seasonally-differenced

for consistency to manufacturers and retailers; three-digits wholesaling NAICS aggregate across product types for merchant wholesalers of durable and non-durable goods. All data were obtained in unadjusted form, as recommended by Enders (2014).

1.4.3 Variable Operationalization

This subsection provides an overview of the operationalization of variables used in the study. Descriptive statistics for all variables are provided in Table 1.1.

1.4.3.1 Dependent Variable

Monthly inventory data were obtained from the United States Department of Commerce (Census Bureau) for a twenty-three year period beginning June 1993, and ending June 2016. As this study seeks to test for the presence of a common phenomenon across inventories, observations were collected on total inventories within industries (*i.e.*, manufacturing inventories included finished goods inventories as well as raw materials inventories and work-in-process inventories; see also Cachon et al., 2007). Data on manufacturing inventories were collected from the Census Bureau's Manufacturers' Shipments, Inventories and Orders Survey (M3). Data on wholesale inventories were collected from the Census Bureau's Monthly Wholesale Trade Survey (MWTS). Data on retail inventories were collected from the Census Bureau's

Table 1.2: Industry subsectors.

NAICS	Subsector	NAICS	Subsector
Manufacturing			
311xxx	Food products	326xxx	Plastics and rubber products
312xxx	Beverage and tobacco products	327xxx	Nonmetallic mineral products
313xxx	Textile mills	331xxx	Primary metals
314xxx	Textile products	332xxx	Fabricated metal products
315xxx	Apparel	333xxx	Machinery
316xxx	Leather and allied products	334xxx	Computers and electronic products
321xxx	Wood products	335xxx	Electrical equipment/appliances/components
322xxx	Paper products	336xxx	Transportation equipment
323xxx	Printing	337xxx	Furniture and related goods
324xxx	Petroleum and coal products	339xxx	Miscellaneous durable goods
325xxx	Chemical products		
Wholesaling			
4231xx	Motor vehicle and motor vehicle parts/supplies	4241xx	Paper and paper products
4232xx	Furniture and home furnishings	4242xx	Drugs and druggists' sundries
4233xx	Lumber and other construction materials	4243xx	Apparel, piece goods, and notions
4234xx	Professional and commercial equipment/supplies	4244xx	Grocery and related products
4235xx	Metals and minerals, except petroleum	4245xx	Farm product raw materials
4236xx	Electrical and electronic goods	4246xx	Chemicals and allied products
4237xx	Hardware, plumbing and heating equipment/supplies	4247xx	Petroleum and petroleum products
4238xx	Machinery, equipment, and supplies	4248xx	Beer, Wine, and distilled alcoholic beverages
4239xx	Miscellaneous durable goods	4249xx	Miscellaneous nondurable goods
Retailing			
441xxx	Motor vehicles and motor vehicle parts	445xxx	Food and beverage stores
44^{2}_{3}xxx	Furniture, furnishings, electronics, and appliances	448xxx	Clothing and clothing accessory stores
444xxx	Building materials and garden equipment/supplies	452xxx	General merchandise stores

and Monthly Retail Trade Survey (MRTS). Collectively, inventories held by these forty-four subsectors were valued at 1.22 trillion dollars in June 2016, compared to total U.S. business inventories of 1.81 trillion dollars for this same month.

1.4.3.2 Focal Predictors

Demand. Monthly sales data were obtained for each subsector in their unajusted form. Consistent with previous studies, sales were used to proxy for demand (*e.g.*, Rumyantsev and Netessine, 2007; Bray and Mendelson, 2012; Kesavan and Kushwaha, 2014).

Re-ordering cost. The fixed portion of costs incurred when re-ordering inventory is predominantly comprised of two parts: personnel-related costs associated with preparing, administering and inspecting re-orders and costs associated with the movement of inventory. Recognizing that the latter of these two costs is the most volatile component of re-order costs,[3] we proxy transportation costs for re-ordering costs. We use the producer price index (PPI) for less-than-truckload shipments of general freight. PPIs have previously been utilized in the empirical inventory work (*e.g.*, Chen et al., 2005) and measure the average change over time in market prices observed by domestic producers for their output. The less-than-truckload PPI was selected as carriers specializing in less-than-truckload shipments face costs in addition to those faced by full truckload carriers (*e.g.*, terminal operations, break-bulk handling, operation of pickup and delivery vehicles). We later replace this PPI with the index for full-truckload shipments and find the statistical results to be qualitatively similar. The 2017 NAICS defines the less-than-truckload industry as follows:

> This U.S. industry comprises establishments primarily engaged in providing long-distance, general freight, less than truckload (LTL) trucking. LTL carriage is characterized as multiple shipments combined onto a single truck for multiple deliveries within a network. These establishments are generally characterized by the following network activities: local pick-up, local sorting and terminal operations, line-haul, destination sorting and terminal operations, and local delivery (U.S. Census Bureau, 2017).

[3] In an ordinary least squares regression, linear trend accounts for 99.39 percent of the variability in the U.S. Social Security Administration's National Average Wage Index (NAWI) between calendar years 1992 and 2015, $F(1, 22) = 3067, p < .01$.

Inventory holding cost. Despite decades of investigation, researchers have been unable to consistently link fluctuations in the cost to finance inventory to aggregate inventory investment behavior (Jones and Tuzel, 2013). Instead, we focus on the physical cost of carrying inventory, a component of holding cost that is distinct from financing cost and an area that remains understudied in the empirical inventory management literature (Rumyantsev and Netessine, 2007). We measure changes in this cost, as faced by the average firm, using the PPI for warehousing and storage services. The 2017 NAICS defines this industry as follows:

> Industries in the Warehousing and Storage subsector are primarily engaged in operating warehousing and storage facilities for general merchandise, refrigerated goods, and other warehouse products. These establishments provide facilities to store goods. They do not sell the goods they handle. These establishments take responsibility for storing the goods and keeping them secure. They may also provide a range of services, often referred to as logistics services, related to the distribution of goods. Logistics services can include labeling, breaking bulk, inventory control and management, light assembly, order entry and fulfillment, packaging, pick and pack, price marking and ticketing, and transportation arrangement. However, establishments in this industry group always provide warehousing or storage services in addition to any logistic services. Furthermore, the warehousing or storage of goods must be more than incidental to the performance of services, such as price marking (U.S. Census Bureau, 2017).

1.4.3.3 Covariates

We introduce covariates into the model to control for potential confounds (*e.g.*, common trend, seasonal patterns, underlying processes and macroeconomic cycles) and to improve statistical efficiency in estimation. Coviariates were selected based on a review of empirical inventory models within the literature and are described in the following section.

Trend. Linear trend is included in the model to control for potential confounds associated with time that might bias parameter estimates. Chen et al. (2005, 2007) and Rajagopalan and Malhotra (2001) find evidence of a long-term decline in U.S. inventories that they attribute to the adoption of new technology and inventory practices, such as just-in-time production.

Seasonality. Month effects are mathematically removed through seasonal differencing. The specification of the pre-differenced model includes month dummy variables with random slopes for exposition.

Business Cycles. Firms actively adjust inventory investments in response to economic shocks (Kesavan and Kushwaha, 2014). We control for macroeconomic conditions using smoothed continuous recession probabilities (Piger and Chauvet, 2017), which were obtained from the U.S. Federal Reserve. Recession probabilities are developed using a Markov-switching dynamic factor model and indicate the probability that the U.S. economy is in a recession or expansion at a given time (Chauvet, 1998).

Inflation. The Consumer Price Index (CPI) is included as an inflationary control since currency inflation may confound the relationship between aggregate inventories and other variables that are measured in dollars. The use of a price index is consistent with previous studies (*e.g.*, Chen et al., 2007; Kesavan and Kushwaha, 2014). Monthly CPI observations were obtained from the U.S. Bureau of Labor Statistics.

Lead Time. Inventory theory suggests that safety stock levels depend on re-order lead time, thus impacting aggregate inventory levels. Global sourcing typically increases re-order lead times by 20 to 50 days, due to increased transit time and customs clearance (Jain et al., 2014). Rajagopalan and Malhotra (2001) argue that, in the long-run, increased importation has not only increased lead times but may prompt firms to add inventory as a buffer against fluctuations in currency and other uncertainties associated with global sourcing. Indeed, evidence suggests that firms importing goods observe longer lead times and hold more inventory than firms who source domestically (Alessandria et al., 2010). To control for these effects, we obtained data on the monthly U.S. trade balance of goods in its unadjusted form from the U.S. Department of Commerce.

Industry Consolidation. Lastly, we control for the level of consolidation within each industry subsector using establishment data obtained from the U.S. Department of Labor's Quarterly Census of Wages and Employment. Firms should enjoy economies of scale in inventory management due to decreased lot sizes (as seen in Equation 1.1), decreased demand variability (Cachon and Olivares, 2010) and network advantages such as inventory pooling (Lim et al., 2016). Establishment data were obtained in unadjusted form and monthly estimates were interpolated separately for each subsector using cubic splines (Forsythe et al., 1977).

1.5 Model Specification

This section specifies mixed-effect models used in hypothesis testing. A initial model (not tested) is described in Equation 1.2, where the inventory level for subsector i at time t depends on a subsector-specific deviation from its sector-level intercept; subsector-specific effects of time and seasonality; and population-average effects of the business cycle (CYCLE), the consumer price index (CPI), the trade balance (TRADE), the number of establishments in the subsector (EST), the level of demand in the subsector (DEMAND), re-ordering cost (REORDER) and holding cost (HOLDING). All continuous variables in the model are log-transformed except time and CYCLE, the latter of which is in probability form and contains non-positive observations. We scale CYCLE by .01 for interpretability in the regression equation.

Level 1

$$
\begin{aligned}
Y_{it} = {} & \beta_{0i} + \beta_{1i} \times t + \sum_{j=1}^{12} \psi_{ji} \times M_t + \sum_{k=2}^{3} \pi_k \times S_i \\
& + \beta_2 \times \text{CYCLE}_t + \beta_3 \times ln(\text{CPI})_t + \beta_4 \times ln(\text{TRADE})_t \\
& + \beta_5 \times ln(\text{EST})_{it} + \beta_6 \times ln(\text{DEMAND})_{it} \\
& + \beta_7 \times ln(\text{REORDER})_t + \beta_8 \times ln(\text{HOLDING})_t + R_{it}
\end{aligned} \tag{1.2}
$$

Level 2

$$\beta_{0i} = \gamma_{00} + U_{0i} \tag{1.3}$$

$$\beta_{1i} = \gamma_{10} + U_{1i} \tag{1.4}$$

$$\psi_{ji} = \gamma_{j0} + U_{ji}, \forall j = \text{Feb}, \ldots, \text{Dec} \tag{1.5}$$

1.5.1 Contemporaneous Model

A contemporaneous model (Equation 1.7) is specified by taking a seasonal difference of both sides of Equation 1.2. Stationarity is not a mathematical necessity in longitudinal analysis of panel data since covariances in the residual variance-covariance matrix can be estimated (Newsom, 2015). Nonetheless, the removal of seasonal components in time series data is often desirable to eliminate common covariation and autocorrelations that may persist at seasonal lags (Enders, 2014). After seasonal differencing, time-invariant effects—those of month (ψ_{ji}), sector (π_k) and starting position (β_{0i})—are removed from the regression equation. Subsector-specific effects of time (linear trend) are estimated as random intercepts in the differenced equation. Since inventories are log-transformed (from Equation 1.1), the seasonally differenced equation explains year-over-year percent changes in inventories.

Level 1

$$\begin{aligned}\Delta_{12}Y_{it} = {} & 12 \times \beta_{1i} + \beta_2 \times \Delta_{12}\text{CYCLE}_t + \beta_3 \times \Delta_{12}ln(\text{CPI})_t \\ & + \beta_4 \times \Delta_{12}ln(\text{TRADE})_t + \beta_5 \times \Delta_{12}ln(\text{EST})_{it} \\ & + \beta_6 \times \Delta_{12}ln(\text{DEMAND})_{it} + \beta_7 \times \Delta_{12}ln(\text{REORDER})_t \\ & + \beta_8 \times \Delta_{12}ln(\text{HOLDING})_t + R_{it}\end{aligned} \tag{1.6}$$

Level 2

$$\beta_{1i} = \gamma_{10} + U_{1i} \tag{1.7}$$

1.5.2 Dynamic Model

To estimate the competing, dynamic specification (Equation 1.9), the contemporaneous effects in Equation 1.7 of demand (β_6), reorder cost(β_7) and holding cost (β_8) are replaced by polynomial distributed lags. Lags up to twelve months are specified, following the advice of Lovell (1961) that "the planning horizon of the firm is surely shorter than a year for decisions involving output adjustment and inventory" (p. 294). Polynomials up to the second degree where constructed as a balance between making potentially limiting assumptions about the distribution of lagged effects and overfitting the model.

Level 1

$$\begin{aligned}\Delta_{12}Y_{it} &= 12 \times \beta_{1i} + \beta_2 \times \Delta_{12}\text{CYCLE}_t + \beta_3 \times \Delta_{12}ln(\text{CPI})_t \\ &+ \beta_4 \times \Delta_{12}ln(\text{TRADE})_t + \beta_5 \times \Delta_{12}ln(\text{EST})_{it} \\ &+ \sum_{p=0}^{2}\sum_{z=0}^{12} \alpha_p \times z^p \times \Delta_{12}ln(\text{DEMAND})_{it-z} \\ &+ \sum_{p=0}^{2}\sum_{z=0}^{12} \delta_p \times z^p \times \Delta_{12}ln(\text{REORDER})_{it-z} \\ &+ \sum_{p=0}^{2}\sum_{z=0}^{12} \omega_p \times z^p \times \Delta_{12}ln(\text{HOLDING})_{it-z} \\ &+ R_{it}\end{aligned} \tag{1.8}$$

Level 2

$$\beta_{1i} = \gamma_{10} + U_{1i} \tag{1.9}$$

1.6 Results

Models of year-over-year inventory investment were estimated using maximum-likelihood with AR(1) residual structures. Four models were estimated: a model with only control variables, the contemporaneous model, the dynamic model and a parsimonious dynamic model. Parameter estimates and measures of model fit are listed in Table 1.3. Prior to estimation, we tested for the presence of a unit-root in the response variable. The Levin-Lin-Chu panel unit-root test (Levin et al., 2002) rejects the null hypothesis that the response panel contains unit roots before subtracting out cross-sectional means ($t^* = -10.41,\ p < .01$) and after subtracting out cross-sectional means ($t^* = -11.97,\ p < .01$), when lag lengths for augmented Dickey-Fuller tests are selected using the Akaike information criterion (AIC). As seen in Table 1.3, estimates of the AR(1) parameter range from .92 to .95; positive log-likelihood values were estimated for all models.

Table 1.3: Results of AR(1) mixed effect model estimation.

Response= $\Delta_{12}ln$(Inventories)		Control		Contemporary		Dynamic (A)		Dynamic (B)	
		Est.	S.E.	Est.	S.E.	Est.	S.E.	Est.	S.E.
Fixed effects									
Intercept†	γ_{10}	.015**	.005	.014**	.005	.002	.004	.002	.004
Δ_{12}CYCLE	β_2	.043**	.003	.044**	.003	.028**	.004	.028**	.004
$\Delta_{12}ln$(CPI)	β_3	.576**	.063	.452**	.070	.179*	.070	.180**	.068
$\Delta_{12}ln$(TRADE)	β_4	.008**	.002	−.002	.002	−.002	.002	−.002	.002
$\Delta_{12}ln$(EST)	β_5	.404**	.087	.412**	.084	.334**	.074	.334**	.074
$\Delta_{12}ln$(DEMAND)	β_6			.073**	.005				
	α_0					.096**	.004	.096**	.004
	α_1					−.013**	.002	−.013**	.002
	α_2					<.001**	<.001	<.001**	.000
$\Delta_{12}ln$(REORDER)	β_7			.057**	.021				
	δ_0					.057**	.017	.057**	.012
	δ_1					−.006	.007	−.006**	.002
	δ_2					<.001	.001		
$\Delta_{12}ln$(HOLDING)	β_8			−.075*	.030				
	ω_0					−.020	.026	−.020	.025
	ω_1					−.015	.011	−.015	.011
	ω_2					.002*	.001	.002*	.001
Variance Components									
Variance (e_{it})	σ_e^2	.008**	<.001	.007**	<.001	.004**	<.001	.004**	<.001
Autoregressive (AR1)	φ	.952**	.003	.947**	.003	.920**	.004	.920**	.004
Measures of fit									
−2 Log likelihood		−48,841.9		−49,069.5		−49,861.3		−49,861.3	
AIC		−48,827.9		−49,049.5		−49,829.3		−49,831.3	
pR^2 (Level 1)		.13		.22		.52		.52	

Notes: $^{*}p < .05$, $^{**}p < .01$ (two-tailed); †intercept scaled by 12; R^2 for Level 1 approximated as $(\sigma^2_{\text{null}} - \sigma^2_{\text{full}})/\sigma^2_{\text{null}}$ (Bryk and Raudenbush, 2002)

1.6.1 Model Testing

The contemporaneous model exhibits improved fit over the control model ($\chi^2_\Delta(3) = 227.6$, $p < .01$) and explains 22 percent of the Level 1 variability in year-over-year inventory investments, when approximated using the method proposed by Bryk and Raudenbush (2002). Coefficients for the effects of demand, re-ordering cost and inventory holding cost are statistically significant and are in their hypothesized directions.

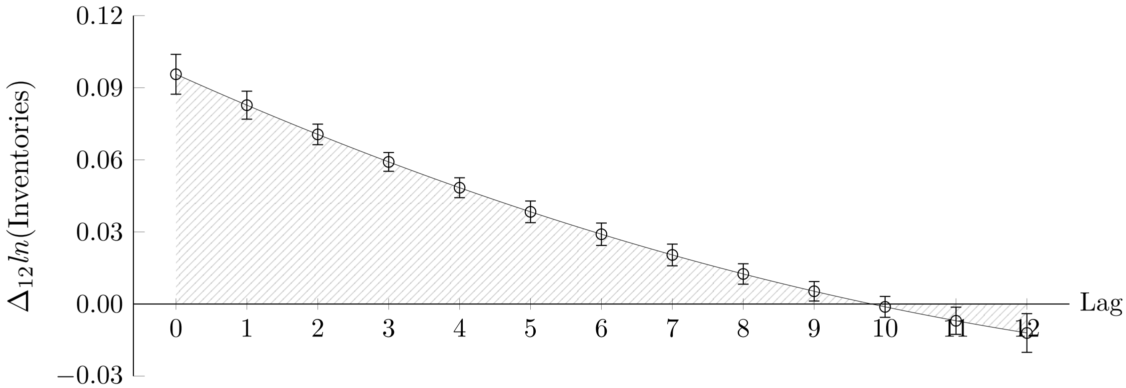
Fit of the dynamic model (A) is superior to that of the contemporaneous model, when compared by AIC ($\Delta = -779.8$). This model explains an additional 30 percent of the Level 1 variability in year-over-year inventory investments, for a total of 52 percent. Thus, the data support Hypothesis 4 which suggest that aggregate inventories exhibit a protracted adjustment to target levels. Prior to calculating partial elasticities, we prune the second-degree polynomial of re-ordering cost since the coefficient (δ_2) is not statistically different than zero. This change results in a 2.0 point decrease in AIC without a change to log-likelihood ($\chi^2_\Delta < .01,\ p > .05$), suggesting improved fit. Model fit of this more parsimonious dynamic model (B) is otherwise qualitatively identical to that dynamic model (A).

1.6.2 Magnitude and Pattern of Adjustment

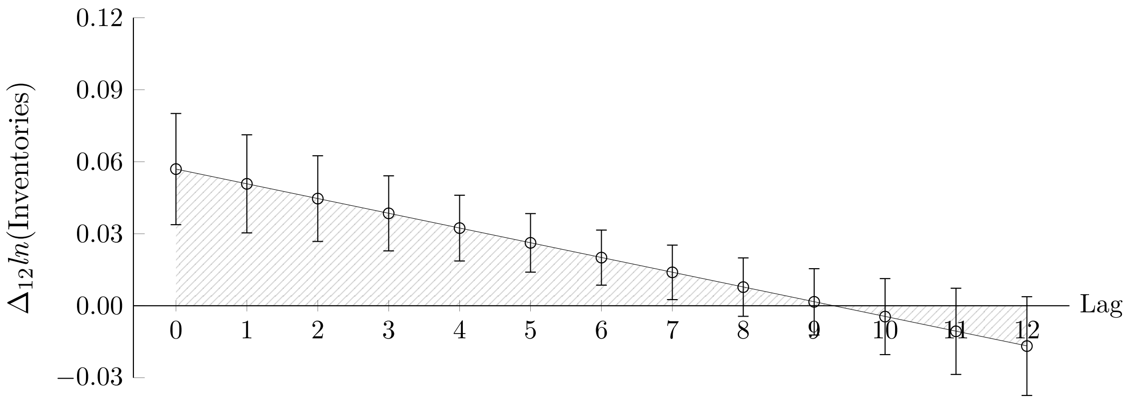
Estimates and confidence intervals for the overall magnitude of adjustment (as partial elasticities) and pattern of adjustment (impulse response) are constructed using the Monte Carlo method. $10,000$ draws were taken from a multivariate normal distribution using the model-produced mean vector and variance-covariance matrix of parameter estimates. Equation 1.10 contains the formula used to recover point estimates of the partial elasticities with respect to demand (α), re-order cost (δ) and holding cost (ω). Corresponding impulse responses with 95% confidence intervals are depicted in Figure 1.1.

$$\varepsilon_i = \sum_{p=0}^{2} \sum_{z=0}^{12} i_p \times z^p, \forall i = \alpha, \delta, \omega \tag{1.10}$$

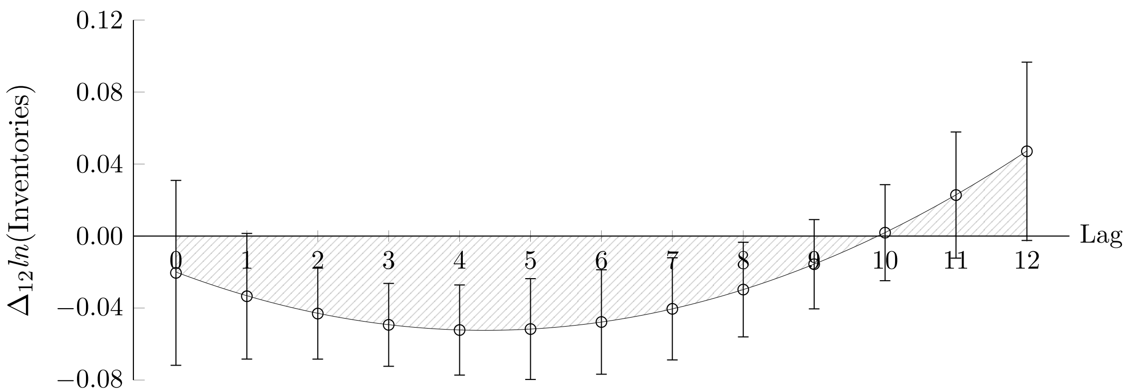
Hypothesis 1 suggests that the partial elasticity of aggregate inventories with respect to aggregate demand is greater than zero. Consistent with this hypothesis, a one percent



(a) Demand



(b) Reorder cost



(c) Holding cost

Figure 1.1: Impulse response for a one percent shock to demand (a), reorder cost (b) and holding cost (c).

change in aggregate demand is associated with an estimated .44 percent adjustment to aggregate inventories [$\mathrm{CI_L} = .40,\ \mathrm{CI_U} = .48$] over the subsequent year. As shown in Figure 1.1, the impulse response to a demand shock is slightly convex, tapering off to zero at the ten month mark.

Hypothesis 2 suggests that the partial elasticity of aggregate inventories with respect to re-order costs is greater than zero. we also find support for this hypothesis. A one percent change in re-order costs, as proxied by the PPI for less-than-truckload freight shipments, is associated with an estimated .26 percent adjustment to aggregate inventories [$\mathrm{CI_L} = .11,\ \mathrm{CI_U} = .41$] over the subsequent year. The impulse response to a re-order cost shocks declines linearly over time and does not differ from zero after the eight months.

Lastly, Hypothesis 3 suggests that the partial elasticity of aggregate inventories with respect to inventory holding costs is less than zero. This hypothesis was also supported by the data. A one percent change in re-order costs, as proxied by PPI for warehousing and storage services, is associated with an estimated $-.31$ percent adjustment to aggregate inventories [$\mathrm{CI_L} = -.52,\ \mathrm{CI_U} = -.10$] over the subsequent year. The impulse response to a holding cost shock is convex, with significant effects between two and eight months after the shock.

1.7 Managerial Implications

These results carry important implications for policy makers, firm-level managers and market analysts. At current inventory levels, a one percent increase in reorder costs results in an additional 4.75 billion dollar investment in U.S. inventories within the following twelve months. Our analysis focuses on the transportation component of these costs, of which fuel costs accounted for 34 percent of the average marginal trucking cost per mile in 2014 (Torrey and Murray, 2015). Public-sector officials considering legislation or policies that may impact the transportation industry (*i.e.*, fuel taxes, adjustments to strategic oil reserves, highway tolls) should be aware of, and consider the potential effects on private inventories. Since inventory is an investment of firm resources, both firms and economies experience opportunity costs related to foregoing alternative resource allocations, such as investments in human captial, physical capital or technology. As an example, a well-intentioned but short-sighted increase to fuel tax by a local or state government to fund infrastructure projects may channel firm resources into inventories and away from the local labor market.

The estimation of a convex impulse response to shocks in physical inventory holding costs also carries significant implications for providers of transportation and warehousing services. First, these results show that inventories decrease by nearly a third of one percent in response to a one percent increase in the physical costs of holding inventory. However, the effects of this price sensitivity may not be immediately apparent to service providers; it is likely that contractual inflexibilities constrain the inventory updating process (Larson et al., 2015). Thus, service providers should be careful not to draw conclusions from initial trends or responses to rate changes until clients have had sufficient time to adjust inventories to a new equilibrium. Further, our results suggest that these service providers must not insulate themselves from rate information in the transportation and warehousing markets, even if they do not provide a full suite of logistics services. Normative theory of optimal ordering behavior posits that client firms seek to jointly economize on the costs of re-ordering and holding inventory, implying a trade-off between these costs. Indeed, the findings of this study are consistent with those expected under trade-off behavior. Thus, the quality of a service provider's forecast for client demand of logistics services may hinge, in part, on the service provider's ability to broadly forecast future rates in markets for logistics services. Service providers offering integrated logistics services may hold an advantage in their ability to control rates on both sides of the trade-off. For instance, an integrated service provider desiring to maintain some maximum level of warehouse utilization may choose to buffer or attenuate rising freight costs through pricing of their services.

A recent stream of research has explored the relationship between a firm's inventory leaness and financial and market performance (Eroglu and Hofer, 2011; Steinker and Hoberg, 2013; Alan et al., 2014). Our findings suggest that inventory leaness should not be viewed narrowly as a relationship between inventory levels and demand, as has frequently been done in the past. For instance, Eroglu and Hofer (2011) model firm's inventory leanness to sales and estimate an average elasticity of .91, nearly twice the value estimated in this study and in other recent studies that consider covariates (*e.g.*, Kesavan and Kushwaha, 2014) and far greater than what would be expected under optimal ordering behavior. Indeed, inventory theory strongly suggests that firms who respond myopically to changes in demand will order at suboptimal quantities if other important factors do not remain stationary. Thus it is possible that many of the firm-level inventory abnormalities found in previous studies (*i.e.*, deviations from inventory productivity curves) can be accounted for by variability in re-order and holding costs—or these abnormalities may simply be transient artifacts of protracted updating processes. This should give caution to firm managers, market analysts and investors who are using firm-level inventory data to compare performance against peer firms. It also suggests that firms who can react quickly to changing market conditions may gain a cost advantage over other firms who are slower to acquire, process and act on information from the external environment.

1.8 Discussion

This research makes several important contributions to the literature. This study extends the finding of Larson et al. (2015) that inventory investments are a function of prior-period sales. In line with the macroeconomic literature, we view this phenomenon as the protraction of inventory updating processes. However, we show that protraction occurs across the broad range of business inventories. Further, the phenomenon is not only driven by the demand distribution, which might be inferred from Schweitzer and Cachon (2000) and subsequent laboratory research in single-period problems—inventory updates are also protracted in response to changes in re-order costs and inventory holding costs. This may explain why previous studies have struggled to find a significant contemperaneous effect of holding cost on inventories: effects on aggregate inventories are not felt until at least two months after a cost shock. New micro-level theororizing is needed to account for these facts since the underlying forces appear to be both cognitive and institutional in nature and are not well-known.

We believe this is the first study to test the partial elasticity of aggregate inventories with respect to transportation costs. While the optimization literature is replete with studies into the relationship between transportation cost and inventories (Ali and O'Connor, 2013), very little empirical evidence exists into the true nature of this relationship. We treat transportation cost as a component of fixed reordering cost and demonstrate that inventories respond strongly to variability in the freight market. These results suggest that transportation costs should no longer be assumed away in inventory models; rather, there may be a need to review and update existing empirical models to account for this effect. This study is limited in that we only examine operational changes to inventories out to twelve months and do not account for strategic change which may have occured at longer intervals (*i.e.*, due to sustained shifts in the transportation market), although Lovell's (1961) cautions on lags of greater than 12 months suggest we are safe in our assumption. Nonetheless, future research should investigate changes to supply and distibution networks in response to the price swings and general volatility in the post-recessionary market for freight tranportation and its major cost components, such as fuel.

The literature has been keenly interested in assessing aggregate inventories for the presence of long-term trend. Managerial practices, such as just-in-time production, and the proliferation of technology (*e.g.*, MRP, RFID) have enabled leaner inventories (Rajagopalan and Malhotra, 2001). However, demands for greater product variety and service levels, decreased product life cycles and globalization have placed conflicting pressures on firms to bolster inventory holdings (Chen et al., 2007). Gaur et al. (2005) find an overall decrease in inventory turnover within the retail sector between 1987 and 2000, after controlling for gross margin, capital intensity and sales surprise. They find this effect to be more pronounced for retailers with low capital intensity. In contrast, Chen et al. (2007) conclude that inventory turnover improved in the retail sector beginning in mid-1990s, and document a steady decline in inventories within other sectors (*i.e.*, manufacturing, wholesaling) since the early 1980s. Rajagopalan and Malhotra (2001) show that, at least for manufacturers, this trend may date as far back as the 1960s. Our results indicate that an overall, long-term trend in aggregate inventories is not present between 1993 and 2016. And, while we do not test individually for the presence of trend within sectors and subsectors, subsector-level variability of linear trend (modeled as a random intercept in the differenced equation) was tested for, and was not found to be statistically significant. These findings do not rule out the presence of non-linear trend, such as quadtratic trend or one or more stuctural breaks as may have occurred during the 2007–2009 recession. However, after macroeconomic factors and the order quantity regressors are held constant, there is no evidence of a level shift over time.

Lastly, these results strongly suggest, as others have noted, that economies of scale exist in inventory management. Holding demand and other factors in the regression equation constant, a one percent consolidation within an industry subsector is associated with a one-third percent decrease in inventories. This is not entirely surprising given the vast amount of research into optimization of supply chain networks. As firms grow, so should their ability to pool inventories (Lim et al., 2016). However, the empirical evidence on scale economies in inventory management is still in its infancy. Future research should consider estimating inventory productivity functions to assess the extent to which economies are enjoyed across product types (*e.g.*, durables, non-durables) and between industry sectors.

REFERENCES

REFERENCES

Akhtar, M. (1983). Effects of interest rates and inflation on aggregate inventory investment in the United States. *The American Economic Review*, 73(3):319–328.

Alan, Y., Gao, G. P., and Gaur, V. (2014). Does inventory productivity predict future stock returns? A retailing industry perspective. *Management Science*, 60(10):2416–2434.

Alessandria, G., Kaboski, J. P., and Midrigan, V. (2010). Inventories, lumpy trade, and large devaluations. *The American Economic Review*, 100(5):2304–2339.

Ali, A. I. and O'Connor, D. J. (2013). Using truck-inventory-cost to obtain solutions to multi-period logistics models. *International Journal of Production Economics*, 143(1):144–150.

Ballou, R. H. (1981). Estimating and auditing aggregate inventory levels at multiple stocking points. *Journal of Operations Management*, 1(3):143–153.

Banerjee, A. and Mizen, P. (2006). A re-interpretation of the linear quadratic model when inventories and sales are polynomially cointegrated. *Journal of Applied Econometrics*, 21(8):1249–1264.

Bils, M. and Kahn, J. A. (2000). What inventory behavior tells us about business cycles. *American Economic Review*, 90(3):458–481.

Blinder, A. S. (1986). Can the production smoothing model of inventory behavior be saved? *The Quarterly Journal of Economics*, 101(3):431–453.

Blinder, A. S. and Maccini, L. J. (1991). Taking stock: A critical assessment of recent research on inventories. *The Journal of Economic Perspectives*, 5(1):73–96.

Bolton, G. E., Ockenfels, A., and Thonemann, U. W. (2012). Managers and students as newsvendors. *Management Science*, 58(12):2225–2233.

Bray, R. L. and Mendelson, H. (2012). Information transmission and the bullwhip effect: An empirical investigation. *Management Science*, 58(5):860–875.

Bryk, A. S. and Raudenbush, S. W. (2002). *Hierarchical linear models for social and behavioral research: Applications and data analysis methods*. Sage, Newbury Park, CA.

Buzacott, J. A. and Zhang, R. Q. (2004). Inventory management with asset-based financing. *Management Science*, 50(9):1274–1292.

Cachon, G. P. and Olivares, M. (2010). Drivers of finished-goods inventory in the U.S. automobile industry. *Management Science*, 56(1):202–216.

Cachon, G. P., Randall, T., and Schmidt, G. M. (2007). In search of the bullwhip effect. *Manufacturing & Service Operations Management*, 9(4):457–479.

Chauvet, M. (1998). An econometric characterization of business cycle dynamics with factor structure and regime switching. *International Economic Review*, 39(4):969–996.

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

Chen, H., Frank, M. Z., and Wu, O. Q. (2007). U.S. retail and wholesale inventory performance from 1981 to 2004. *Manufacturing & Service Operations Management*, 9(4):430–456.

Chen, M.-J. and Hambrick, D. C. (1995). Speed, stealth, and selective attack: How small firms differ from large firms in competitive behavior. *Academy of Management Journal*, 38(2):453–482.

Daft, R. L. and Weick, K. E. (1984). Toward a model of organizations as interpretation systems. *Academy of Management Review*, 9(2):284–295.

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

Enders, W. (2014). *Applied econometeric time series*. John Wiley & Sons, Hobuken, NJ.

Eroglu, C. and Hofer, C. (2011). Lean, leaner, too lean? The inventory-performance link revisited. *Journal of Operations Management*, 29(4):356–369.

Feldstein, M., Auerbach, A., Hall, R. E., and Lovell, M. C. (1976). Inventory behavior in durable-goods manufacturing: The target-adjustment model. *Brookings Papers on Economic Activity*, 1976(2):351–408.

Forsythe, G. E., Moler, C. B., and Malcolm, M. A. (1977). *Computer methods for mathematical computations*. Prentice-Hall, Englewood Cliffs, NJ.

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

Gino, F. and Pisano, G. (2008). Toward a theory of behavioral operations. *Manufacturing & Service Operations Management*, 10(4):676–691.

Holt, C. C., Modigliani, F., Muth, J. F., and Simon, H. A. (1960). *Production planning, inventories, and workforce*. Prentice-Hall, Englewood Cliffs, NJ.

Jain, N., Girotra, K., and Netessine, S. (2014). Managing global sourcing: Inventory performance. *Management Science*, 60(5):1202–1222.

Jones, C. S. and Tuzel, S. (2013). Inventory investment and the cost of capital. *Journal of Financial Economics*, 107(3):557–579.

Kesavan, S. and Kushwaha, T. (2014). Differences in retail inventory investment behavior during macroeconomic shocks: Role of service level. *Production and Operations Management*, 23(12):2118–2136.

Kesavan, S. and Mani, V. (2013). The relationship between abnormal inventory growth and future earnings for U.S. public retailers. *Manufacturing & Service Operations Management*, 15(1):6–23.

Khan, A. and Thomas, J. K. (2007). Inventories and the business cycle: An equilibrium analysis of (S, s) policies. *The American Economic Review*, 97(4):1165–1188.

Larson, C. R., Turcic, D., and Zhang, F. (2015). An empirical investigation of dynamic ordering policies. *Management Science*, 61(9):2118–2138.

Levin, A., Lin, C.-F., and Chu, C.-S. J. (2002). Unit root tests in panel data: Asymptotic and finite-sample properties. *Journal of Econometrics*, 108(1):1–24.

Lim, M. K., Mak, H.-Y., and Shen, Z.-J. M. (2016). Agility and proximity considerations in supply chain design. *Management Science*. Forthcoming.

Lovell, M. (1961). Manufacturers' inventories, sales expectations, and the acceleration principle. *Econometrica: Journal of the Econometric Society*, 29(3):293–314.

Maccini, L. J., Moore, B. J., and Schaller, H. (2004). The interest rate, learning, and inventory investment. *The American Economic Review*, 94(5):1303–1327.

Maccini, L. J. and Rossana, R. J. (1981). Investment in finished goods inventories: An analysis of adjustment speeds. *American Economic Review Proceedings*, 71(2):17–22.

Mosser, P. C. (1991). Trade inventories and (S, s). *The Quarterly Journal of Economics*, 106(4):1267–1286.

Musalem, A., Olivares, M., Bradlow, E. T., Terwiesch, C., and Corsten, D. (2010). Structural estimation of the effect of out-of-stocks. *Management Science*, 56(7):1180–1197.

Newsom, J. T. (2015). *Longitudinal structural equation modeling: A comprehensive introduction*. Routledge, New York.

Piger, J. M. and Chauvet, M. (2017). Smoothed U.S. recession probabilities [RECPROUSM156N]. Retrieved from FRED, Federal Reserve Bank of St. Louis; https://fred.stlouisfed.org/series/RECPROUSM156N, February 8, 2017.

Rajagopalan, S. (2013). Impact of variety and distribution system characteristics on inventory levels at U.S. retailers. *Manufacturing & Service Operations Management*, 15(2):191–204.

Rajagopalan, S. and Malhotra, A. (2001). Have U.S. manufacturing inventories really decreased? An empirical study. *Manufacturing & Service Operations Management*, 3(1):14–24.

Ramey, V. A. and West, K. D. (1999). Inventories. In Taylor, J. B. and Woodford, M., editors, *Handbook of Macroeconomics Volume 1b*, pages 863–923. Elsevier Science, North-Holland, Amsterdam.

Ren, Y. and Croson, R. (2013). Overconfidence in newsvendor orders: An experimental study. *Management Science*, 59(11):2502–2517.

Rumyantsev, S. and Netessine, S. (2007). What can be learned from classical inventory models? A cross-industry exploratory investigation. *Manufacturing & Service Operations Management*, 9(4):409–429.

Schweitzer, M. E. and Cachon, G. P. (2000). Decision bias in the newsvendor problem with a known demand distribution: Experimental evidence. *Management Science*, 46(3):404–420.

Shah, R. and Shin, H. (2007). Relationships among information technology, inventory, and profitability: An investigation of level invariance using sector level data. *Journal of Operations Management*, 25(4):768–784.

Shan, J. and Zhu, K. (2013). Inventory management in China: An empirical study. *Production and Operations Management*, 22(2):302–313.

Smith, J. (2016). Trucking industry hits road bump with rising diesel prices. *Wall Street Journal*. Retrieved from https://www.wsj.com.

Smith, K. G., Grimm, C. M., Gannon, M. J., and Chen, M.-J. (1991). Organizational information processing, competitive responses, and performance in the U.S. domestic airline industry. *Academy of Management Journal*, 34(1):60–85.

Steinker, S. and Hoberg, K. (2013). The impact of inventory dynamics on long-term stock returns–an empirical investigation of U.S. manufacturing companies. *Journal of Operations Management*, 31(5):250–261.

Swanson, D., Williams, B. D., Gu, J., and Waller, M. A. (2016). Full steam ahead: Firms in the U.S. economy adjust inventory for changes in transportation costs but not the reverse. *Transportation Journal*, 55(3):282–295.

The Conference Board (2017). *Economics Watch: Monthly Economic Series Report (January, 2017)*. New York City, NY.

Thomas, J. B., Clark, S. M., and Gioia, D. A. (1993). Strategic sensemaking and organizational performance: Linkages among scanning, interpretation, action, and outcomes. *Academy of Management Journal*, 36(2):239–270.

Torrey, W. F. and Murray, D. (2015). *An Analysis of the Operational Costs of Trucking: 2015 Update*. American Transportation Research Institute, Arlington, VA.

U.S. Census Bureau (2017). *Manufacturing and trade inventories and sales: November, 2016 (Release Number CB17-08)*. Suitland, MD.

van Donselaar, K. H., Gaur, V., van Woensel, T., Broekmeulen, R. A., and Fransoo, J. C. (2010). Ordering behavior in retail stores and implications for automated replenishment. *Management Science*, 56(5):766–784.

Wen, Y. (2005). Understanding the inventory cycle. *Journal of Monetary Economics*, 52(8):1533–1555.

ESSAY 2

PRODUCTION POSSIBILITIES, OUTPUT DIVERSIFICATION AND OPERATIONAL EFFICIENCY IN MOTOR CARRIAGE

2.1 Introduction

Corporate diversification has been regarded as one of the most important areas for industrial research (Wan et al., 2011), with strong links to both accounting-based and market-based measures of firm performance (Palich et al., 2000). Corporate incentives for diversification include reduced risk exposure within external markets, presence of anti-trust policies and curtailment of executive employment risk (Hoskisson and Hitt, 1990). Resource-based theorists view managerial incentives to improve the utilization of productive services of resources under the firm's control as the most prominent motive for corporate diversification, arguing that firm resources are only rarely fully utilized due to their indivisibility (Kor et al., 2016). However, market failures (*e.g.*, transaction costs, imperfect mobility) restrict the firm's ability to sell the unused services of resources (Peteraf, 1993). Thus, firms are motivated to diversify into related business activities where resources and the services they provide can be spread across multiple businesses (*i.e.*, productive activities), thereby improving resource utilization and the firm's overall level of efficiency (Wan et al., 2011).

This research examines corporate diversification and its relationship to operational efficiency within the U.S. motor carrier industry. Only a handful of studies have examined motor carriers' diversification of production activities (*e.g.*, Stock, 1988; Smith et al., 1990; Feitler et al., 1997; Hanna and Maltz, 1998; Pettus, 2001); similarly, its linkages to carrier efficiency have received only passing attention (*e.g.*, Silverman et al., 1997; Scheraga, 2011). Yet, operational efficiency is a matter of prominent importance to motor carriers, having been linked to firm survival in both the truckload (TL) segment (Fawcett et al., 2016, p. 23) and in the less-than-truckload (LTL) segment (Nebesky et al., 1995, p. 571). Operational efficiency of firms in the motor carrier industry is also a matter of economic importance. Freight bills in the U.S. for over-the-road shipments totaled more than 700 billion dollars in 2015 (ATA, 2017), accounting for nearly four percent of the nation's gross domestic product. The industry is also one of the nation's largest employers, with more than 1.4 million workers employed in the U.S. alone, of which approximately 900 thousand workers are drivers (BLS, 2017).

Thus, the purpose of this study is to propose and test theory relating a motor carrier's corporate diversification to efficiency in carrier operations. We examine motor carriers that have expanded from one of the two primary market segments for general freight carriage (TL and LTL) into the other segment. We postulate that this diversification is driven by managerial desire to improve the utilization of productive resources, but also suggest that the success of this strategy is contingent on the direction of shifts in organizational interdependence. We test our theory using a panel of 120 medium and large U.S. motor carriers of general freight, with data obtained from governmental archives. Our methodology consists of a two-stage analysis, whereby data envelopment analysis (DEA) is employed in a first stage to estimate annual carrier technical efficiency scores over five measurement occasions and, in a second stage, efficiency scores are regressed on contextual factors in a mixed effects modeling framework. Our statistical model estimates our theorized cubic polynomial relationship between a carrier's segment focus and operational efficiency. Thus, we produce estimates of operational efficiency across the production possibilities curve and use these estimates to assess marginal effects from changes to segment focus.

The remainder of this essay is structured as follows. First, we review the literature on motor carrier productivity, segment focus and operational efficiency. Following this review, we present a theoretical framework for our study and formulate research hypotheses. Our two-stage methodology is then presented in the subsequent section, along with the results from hypothesis testing. Next, we provide a discussion of our findings and implications for theory and motor carrier management. We close by addressing limitations of our study and proposing several potential avenues for future research.

2.2 Literature Review

Our research examines the composition of a motor carrier's core service offering (production of TL and LTL carriage) and the impact of this diversification or focus strategy on carrier efficiency. While issues adjacent to our topic such as motor carrier productivity and returns (economies) of scale have received extensive treatment in the transportation literature, attention to the issue of production efficiency has been very limited (Scheraga, 2011).[1] Further, the literature has tended to bifurcate the motor carrier population such that it has only been in rare instances that carrier production of TL and LTL service has been treated as a continuum—indeed, we will later show that a non-trivial portion of large U.S. motor carriers concurrently produce both TL and LTL carriage. To gain an understanding of these two market segments and the differences between them, we review the literature on the deregulated market for TL and LTL movements of general freight and strategic focus by motor carriers on one or both of these sectors. In doing so, we pay special attention to literature pertaining to output production and carrier utilization of productive resources (*i.e.*, labor, capital assets), since these areas form the basis for our research.

2.2.1 Industry Deregulation

Regulatory reform under the Motor Carrier Act of 1980 (MCA) sharply changed the U.S. trucking industry, eliminating pricing controls, decreasing organization of labor (unionization) among drivers, easing entry for new carriers and greatly increasing competition between carriers (Belzer, 2000). Industry deregulation also reduced or eliminated many regulatory roadblocks and restrictions that hindered carriers from further developing their resource bases (Pettus, 2001) and entering new markets (Silverman et al., 1997). In doing so, a greater degree of polarization emerged in the deregulated marketplace between carriers offering truckload (TL) service–carriage of shipments generally weighing 10,000 or more pounds directly from origin to destination—and those offering less-than-truckload (LTL) service—carriage of shipments of less than 10,000 pounds typically via a hub-and-spoke

[1] The terms *productivity* and *efficiency* are frequently conflated in the popular press but are distinct concepts (Coelli et al., 2005). Consistent with the literature, we define productivity as the ratio of outputs produced to inputs used in production (McMullen and Okuyama, 2000). Our use of the term *efficiency* refers to the technical efficiency—not allocative efficiency—in production, represented by distance of some observed production level from an efficient frontier (Farrell, 1957). If scale economies exist, a firm that is perfectly technically efficient may be able to improve productivity by increasing scale. Conversely, if diseconomies of scale exist then a firm that is technically efficient and increases scale may realize a lower level of productivity while remaining technically efficient.

network.[2] For instance, while the total number of motor carriers nearly tripled within the ten years following industry deregulation (Silverman et al., 1997), the LTL segment experienced a sizeable consolidation. For the portion of the LTL market comprised of medium and large sized carriers, estimates placed the level of consolidation as high as 60 percent (Smith et al., 1990; Nebesky et al., 1995). This LTL segment consolidation came in stark contrast to the expectations of deregulation advocates, many of whom had anticipated market entry and proliferation by small trucking firms within both industry segments (Rakowski, 1988; Belzer, 2000).

2.2.2 The 'Returns to Scale' Controversy

A common explanation for diverging market structures post-regulatory reform has been the differences between the two segments in returns to scale (Rakowski, 1988; Belzer, 2000; Giordano, 2008). *Returns to scale* refers to economies gained (or, if diseconomies exist, relinquished) as carriers grow their operations. If economies exist (*i.e.*, marginal productivity increases and marginal cost decreases with size) then carriers are incentivized to grow and a consolidation would be expected in the market (Belzer, 2000). Alternatively, if diseconomies exist then carriers will have fewer incentives to scale their operations. The presence or absence of scale economies in the trucking industry has been a long-standing matter of debate among practitioners, industry analysts, regulators and academic researchers (Rakowski, 1988; Belzer, 2000; Fawcett et al., 2016). Beyond the differences that exist between the two industry segments, the discussion is further complicated by an often-implicit assumption that any economies or diseconomies might be classified as being linearly homogeneous across levels of carrier size—an assumption inconsistent with economic theory (Johnston and Ozment, 2013). Unfortunately, findings from the estimation of economic production and cost functions have

[2]Prior to the MCA, the U.S. Interstate Commerce Commission avoided classifying carriers between TL and LTL and instead categorized individual shipments (McMullen and Tanaka, 1995). The U.S. Department of Transportation, which now regulates the U.S. trucking industry, still categorizes intercity shipments of general freight as TL or LTL but also classifies carriers by their *majority* revenue commodity group, which includes TL or LTL classifications for carriers of general freight (DOT, 2017).

been mixed even when this assumption is relaxed, as is most commonly accomplished within a transcendental logarithm form or first-order Taylor series expansion of the Cobb-Douglas production function.[3] For example, in the LTL segment, researchers have reached differing conclusions that carriers enjoy scale economies (*e.g.*, Rakowski, 1988; McMullen and Tanaka, 1995; Giordano, 2008), are faced with diseconomies (*e.g.*, Xu et al., 1994) or that neither economies nor diseconomies exist (*e.g.*, Nebesky et al., 1995). Conclusions from the TL segment offer a similarly mixed picture (Fawcett et al., 2016).

Over time, research into scale economies in transport became increasingly nuanced; analysts began fixing regressors such as operational and network characteristics and redefined economies using terms such as *returns to density* (economies gained when operations are scaled by networks remain fixed; Caves et al., 1984), *economies of scope* (economies gained from network expansion; Jara-Díaz and Basso, 2003) and *economies of integration* (economies of density combined with economies from vertical integration; Keeler, 1989). Given that the underlying issue of scale returns has been (and remains) controversial and unresolved, the motor carrier literature has occasionally disregarded these nuances as irrelevant (Rakowski, 1988) or simply confusing (Johnston and Ozment, 2013) to the core issue at hand.

Despite this long-standing lack of consensus on the topic, recent studies such as Fawcett et al. (2016) and Jin et al. (2017) have tended to yield to the positive evidence that scale economies might be available to at least some carriers within each of the segments. Anecdotally, a recent flurry of mergers and acquisitions within the trucking industry also seems to support this. For instance, in April 2017, the boards of directors for Knight Transportation and Swift Transportation unanimously approved a merger of the two carriers, creating the industry's largest long-haul TL carrier (Tyler, 2017). Swift Transportation's Chairman stated that the merger will allow the two companies to "capitalize on economies of scale to achieve substantial synergies" (Business Wire, 2017). A band of other mergers, such as XPO Logistics's 2015 acquisition of Con-way Freight, suggests that similar size

[3]We refer readers interested in the properties and restrictions of this function to Beattie et al. (2009) or, alternatively, to Creel and Farell (2001) for a discussion within the transportation literature of the function's assumptions and limitations.

advantages might also be present in the LTL segment.

2.2.3 Operational inefficiency and the Role of Carrier Segment Focus

One potential explanation for the mixed results reported in the returns-to-scale literature lies in the literature's pre-occupation with assessing scale economies using econometric functions that estimate a conditional mean response. There are several strong, underlying assumptions which, if violated, bring into question the validity of statistical conclusions drawn from estimation of these models. First, evidence suggests that exogeneity assumptions may not hold for input quantities (Creel and Farell, 2001) or for output quantities (Nebesky et al., 1995) in these models, suggesting the potential for bias in estimation. Second, the duality theorom used to relate cost and production functions relies on an assumption of cost minimization under perfect competition, which is a weak, if not unrealistic, assumption in many markets (Bairam and Kahya, 1998). Most importantly, econometric production and cost functions, such as the commonly-employed transcendental logarithm, implicitly assume that all firms are technically efficient (Coelli et al., 2005). Scheraga (2011) shows that not all carriers in the trucking industry are efficient; in fact, more than 90 percent of the sampled carriers exhibited some degree of inefficiency. Thus, estimates from prior production and cost functions may not have presented sound evidence regarding the presence returns to scale exist in trucking. Rather, these models have provided an estimate of the *average* economies or diseconomies realized when a firm at one level of output adjusts to some new level of output. Accordingly, the presence of scale economies could readily be masked within a sample of carriers who are, on average, inefficient and unable to effectively leverage their scale.

If this is true then the issue of carrier inefficiency is arguably more relevant to the trucking industry (across the population of carriers) than the presence or absence of scale economies which might only be accessible to some small proportion of efficient firms. Indeed, in their analysis of the TL segment, Jin et al. (2017, p. 23) contend that "operating at an exceptional level of efficiency has allowed modern TL carriers to survive decades of intense competition post-industry deregulation." Nebesky et al. (1995) make a similar argument for carriers in the LTL segment, asserting that superior efficiency of surviving firms may have driven much of the large-scale consolidation seen in the segment after regulatory reform. If carriers are to succeed in the hyper-competitive trucking market then they must identify and address sources of inefficiency in their operations and exploit potential opportunities for increased efficiency.

Unfortunately, only a handful of studies have examined operational (in)efficiency in trucking. McMullen and Okuyama (2000) examine changes in technical efficiency of motor carriers before and after the MCA using Malmquist productivity indices. Prior to the MCA, proponents of regulatory reform argued that deregulation of the trucking industry would move carriers closer to production frontiers, since institutional rules enacted by the Interstate Commerce Commission to govern the trucking industry imposed certain technical inefficiencies and, in some cases, impaired carrier adoption of new technologies (Belzer, 2000; McMullen and Okuyama, 2000). McMullen and Okuyama (2000) indeed find evidence for an overall increase in technical efficiency after regulatory reform, but also identified a technological regress which they interpret as an increased focus by surviving firms on delivering quality service. Like Jin et al. (2017), McMullen and Okuyama (2000, p. 343) postulate that surviving carriers were those who were "best able to increase technical efficiency and productivity in response to regulatory reform," forcing out less efficient carriers who were unable to adapt to the post-MCA environment.

Examining the motor carrier industry, Scheraga (2011) extends the concept of production competence (Cleveland et al., 1989) to carrier operations and utilizes Malmquist productivity indices to examine changes in carrier efficiency over time based on the seven dimensions of strategic orientation for motor carriers (see Feitler et al., 1997): cost, efficiency, price, risk, service, size and niche. The last of these dimensions, *niche* or *product niche*, refers to the degree to which a motor carrier focuses production within the LTL segment (Scheraga et al., 1994; Feitler et al., 1997). For clarity, we adopt an alternative term for this focus dimension, *LTL Concentration*, as proposed by Smith et al. (1990). In Scheraga's (2011) analysis only the *LTL concentration* dimension was found to influence the Malmquist index scores; motor carriers with higher *LTL concentration* (greater production in the LTL segment) showed improvements in operating efficiency. In contrast, the other six dimensions appeared to offer little explanatory power in Scheraga's (2011) model or, in the case of the carrier size dimension, served primarily to establish boundary conditions around the effects of *LTL concentration*. Accordingly, *LTL concentration* of carriers appears to be uniquely influential among the seven dimensions of carrier strategic orientation, holding a critical association with efficiency in production.

Our study extends the results of Scheraga's (2011) analysis, along with findings by Fawcett et al. (2016) who show, using data solely from the TL segment, that asset leanness is positively associated with performance. We rely on theories of firm growth through diversification (Penrose, 1959) and organizational interdependence (Thompson, 1967) to argue that effective utilization of carrier assets in multiple markets underlies the positive relationship between *LTL concentration* and operational efficiency that was found by Scheraga (2011). In the following section, we present our reasoning for this mechanism along with research hypotheses and a conceptual model of our hypothesized relationship between *LTL concentration* and a carrier's operational (in)efficiency.

2.3 Theory and Hypothesis Development

Firms are collections of productive resources of which managers are incentivized to utilize fully through expansion within existing markets and diversification into new markets (Penrose, 1959). Firm growth and a firm's "productive opportunity" (Penrose, 1959, p. 31) are functions of an existing resource base and are limited by managerial ability; external markets for a firm's products and markets for the factors used in production; and uncertainty and risk (Kor and Mahoney, 2000). Desirable firm outcomes (*e.g.*, profitability, efficiency) are achieved through effectively managing, bundling and deploying firm resources to render valuable services, such as those used as inputs in production (Sirmon et al., 2007; Kor et al., 2016). However, only in rare and transitory cases might a firm's resources and the services they produce be fully utilized (Penrose, 1959)—underutilization is a product of the indivisibility of resources as well as continuous organizational learning (Teece, 1982). Thus, firm managers remain perpetually motivated to more effectively employ unused and underutilized productive services of resources toward improving firm outcomes (Pettus, 2001; Kor et al., 2016). A firm's diversification of production activities is one approach that can be used toward this end.

In a production context, *diversification* refers to the addition of new productive activities which differ sufficiently from existing activities (Penrose, 1959). Diversification need not occur within a new area of specialization; firms may diversify by increasing the variety of items produced or may diversify through vertical integration (Penrose, 1959). Resource-based theorists widely recognize the advantages of diversifying into related areas, since firms can realize economies of scope from sharing of resources across operations and from combining related production activities (Teece, 1982; Palich et al., 2000; Wan et al., 2011). Diversification may not always be advantageous for firms, however, such as when existing productive resources are transferred to highly dissimilar markets (Wernerfelt and Montgomery, 1988), when a firm enters a market where they hold a technological inferior position (Penrose, 1959) or when a firm's business portfolio becomes so diverse that opportunities for resource sharing become limited or managerial capability is strained (Palich et al., 2000).

Organizational theory further suggests that a firm's efficient utilization of existing productive resources within a new activity depends on complementarity of the new activity's work flow to that of existing activities. Thompson (1967) considers forms of organizational work flow by intensity of linkages between constituent units and identifies three generic patterns of organizational interdependence: pooled interdependence (low intensity or absence of work flow between units), sequential interdependence (moderate intensity of work flow) and reciprocal interdependence (high intensity and dependence of work flow). As intensity and/or complexity of work flow increases, so do requirements for coordination and communication among constituent units. Thus, effective resource-sharing and coordination of activities become increasingly difficult as an organization transitions into a higher degree of interdependence (Thompson, 1967), suggesting that diversification's benefits (*i.e*, improved employment of unused or underutilized productive resources) might be best realized when new firm activities do not drastically increase the level of intensity in organizational work flows.

A motor carrier's expansion and diversification into additional service areas may benefit the firm by improving market presence, stabilizing output, improving customer service levels and reducing the risks associated with single-product production (Stock, 1988). Indeed, after regulatory reform the trucking industry saw a significant increase in the number of carriers offering expanded services (Hanna and Maltz, 1998; Feitler et al., 1997; Scheraga, 2011). In many cases, motor carriers have diversified production activities by offering transportation service in both the TL and LTL segments (Smith et al., 1990; Ying and Keeler, 1991; McMullen and Tanaka, 1995). However, the transportation literature has noted significant differences between these segments in the level of coordination required for production. Freight flows in an LTL network are highly interdependent (Silverman et al., 1997); the effects of a single failure (*e.g.*, breakdown of a truck) can ripple through a LTL carrier's entire network (Giordano, 2008; Han et al., 2008). Thus, the efficient completion of any single shipment is predicated upon smooth operations in the carrier's hub and spoke network, termed here as *intra-haul interdependency*. In contrast, efficient completion of truckload shipments necessitates that TL carriers schedule drivers and equipment against individual shipments in such a way as to meet shipper requirements while minimizing deadheading (empty hauls). We refer to this lower level of interdependency seen in the TL segment as *inter-haul sequential interdependency*.

Figure 2.1 depicts our hypothesized relationship between a carrier's segment focus (*LTL concentration*) and operational inefficiency. Consistent with the arguments of Thompson (1967), we postulate that a carrier's movement in segment focus toward (*i.e.*, diversification or expansion into) a higher level of interdependence should make the coordination of resources and production activities increasingly difficult (Thompson, 1967). Following the resource-based perspective of diversification as a corporate mechanism to improve resource utilization, resource coordination challenges across a business portfolio should translate

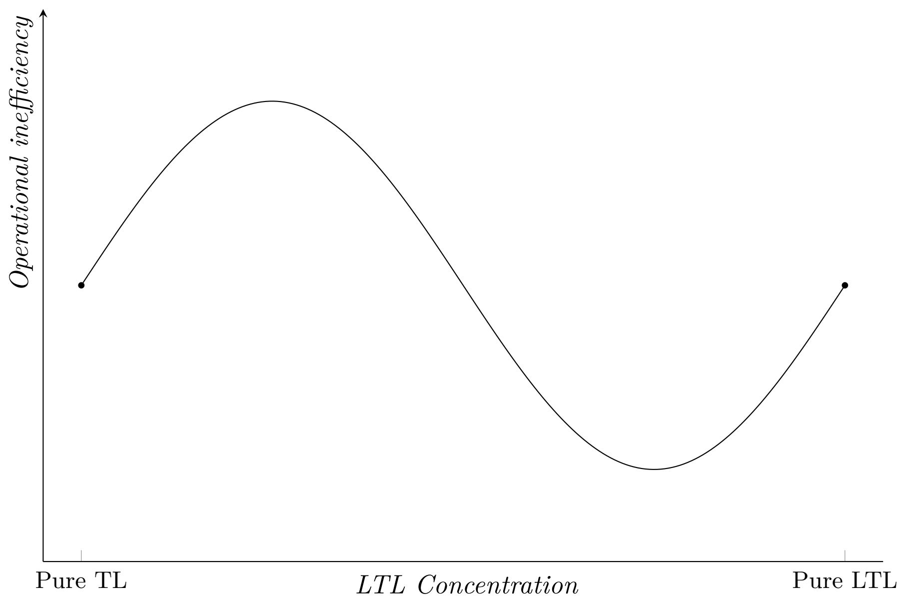


Figure 2.1: Conceptual model.

into lower overall levels of operating efficiency. This relationship can be seen in Figure 2.1 where, for pure TL carriers, the marginal effect of *LTL concentration* on inefficiency is positive (**Hypothesis 1**). Conversely, we hypothesize that pure LTL carriers who diversify into the TL segment—moving to a lower overall level of interdependence—will be able to better employ underutilized productive resources, resulting in a lower degree of operational inefficiency (**Hypothesis 2**). Accordingly, we theorize the overall relationship between *LTL concentration* and operational inefficiency as a polynomial with two inflection points.

2.4 Methodology

2.4.1 Data and Sample

We test our research hypotheses using a sample obtained from governmental archives of motor carrier production, operating and financial statistics over the five-year period of 1999 through 2003. The use of archival data from this period is appropriate to our test of theory. Most importantly, we seek to examine the motor carrier industry during a period of relative stability, void of much of the transient—but still lingering—volatility that has occurred in factor markets since 2003. For instance, escalation of the well-publicized driver shortage began in 2005 and, with some ebbs and flows in the trucking labor market over the past decade, has artificially constrained motor carriers' ability to grow and expand (Costello and Suarez, 2015) while also impacting carrier productivity (Min and Lambert, 2002; Suzuki, 2007; Cantor et al., 2011). Additionally, transient shifts and volatility in the price of fuel have, to some extent, influenced the operational decisions of motor carriers (Winebrake et al., 2015) and have led to speculation, hedging or even bankruptcy filings in the industry (Smith, 2016). Further, carriers continue to adjust to governmental regulations imposed over the past two decades, such as changes to hour-of-service requirements, and these regulations continue to influence carrier productivity (Costello and Suarez, 2015). Nonetheless, both the fundamental nature of the TL and LTL segments (*e.g.*, levels of market competition, productive resources in use, contrast in coordination requirements and complexity in operations) and our underlying theory remain durable over time. Our use of historical data allows us to research corporate diversity by motor carriers in a more 'controlled' setting—one where threats to the validity of our findings from transient environmental factors are significantly reduced.

Data used in this study were obtained from archives published by the U.S. Department of Transportation, Office of Motor Carrier Information (OMCI). These data contain Form M filings required by carriers under 49 U.S.C. § 14123. Consistent with previous studies into the industry (*e.g.* Xu et al., 1994; McMullen and Okuyama, 2000; Giordano, 2008), we limit our sample to large (Class I) and medium (Class II) carriers of general freight. Thus, annual revenues for all carriers in our sample exceed three million dollars after adjustment to 1994 dollars using the procedure within 49 C.F.R. § 369. To ensure that our analysis includes only established carriers in the industry we limit our sample to those motor carriers who filed a Form M during each of the five years. Motor carriers are uniquely identified in our dataset by their motor carrier number.

A central premise of our theorizing is that carrier management is incentivized to improve the utilization of productive resources under their control (Penrose, 1959). Managers have limited control over owner-operators (Corsi and Grimm, 1987)—drivers that move freight under contract with carriers using vehicles that the drivers own and maintain (Cantor et al., 2013). The use of owner-operators has been ubiquitous in the U.S. trucking industry since contractual relationships with these drivers can offer carriers operational flexibility (*i.e.*, in level of output) while reducing carrier investments in employee labor and capital equipment (Corsi and Grimm, 1987). Han et al. (2008) found that, around the time of our period of analysis, nearly 70 percent of all general freight carriers used owner-operators to some extent. For these carriers, expenses attributed to hiring owner-operators accounted for an average of 30 percent of annual operating expenses (Han et al., 2008).

Previous studies have linked the use of owner-operators to key carrier outcomes. For instance, Miller and Saldanha (2016) find that motor carrier safety declines as carriers rely more heavily on owner-operators. While empirical evidence is lacking on the relationship between owner-operator use and carrier operational efficiency, Scheraga (2011) notes that a carrier's use of owner-operators may bias input measures in studies of carrier productivity. We adopt two approaches to minimize our potential for exposure to this bias. First, we restrict our sample to carriers whose owner-operator expenses accounted for no more than 15 percent of total operating expenses, thus removing carriers who rely more heavily on owner-operators from our sample. A balance between sample size and internal validity were primary criteria in establishing this threshold, which is half that of the industry average. For instance, lowering the threshold to zero percent of total operating expenses (no owner-operator use) would result in a sample consisting of only one carrier. As an additional safeguard, we employ a statistical control for owner-operator use in our analysis; our operationalization of this variable is later discussed.

We performed an audit of the resulting sample to identify and filter anomalous observations, as is typical when working with transportation archives (*e.g.*, Han et al., 2008; Fawcett et al., 2016). For 10 of the 130 carrier in our dataset, Form M filings contained reporting errors (*e.g.*, implausible or non-positive values) and, out of caution and a desire to maintain balanced data, we filtered these carriers from our sample. Thus, our final sample consisted of five years of annual observations on 120 carriers, for a total sample size of 600.

Table 2.1: Descriptive statistics.

Variable	Mean	S.D.	ICC†	Pearson Product-Moment Correlation (1)	(2)	(3)	(4)	(5)	(6)
(1) Carrier class	.56	.50	1.00						
(2) Drivers ($\times 10^3$)	.74	2.23	.98	$-.34^*$					
(3) Power units ($\times 10^3$)	.54	1.48	1.00	$-.37^*$	$.97^*$				
(4) Total miles ($\times 10^9$)	.04	.11	.99	$-.38^*$	$.95^*$	$.97^*$			
(5) LTL concentration	.29	.41	.94	$-.17^*$	$.27^*$	$.29^*$	$.23^*$		
(6) Outsourcing	.05	.04	.81	$-.25^*$	$.19^*$	$.20^*$	$.22^*$	$.09^*$	
(7) Operational inefficiency	1.85	1.09	.74	$.19^*$	$-.16^*$	$-.18^*$	$-.21^*$	$.16^*$	$-.14^*$

Notes: $^*p < .05$ (two-tailed); †ICC calculated as $\sigma^2_{\mathrm{b}} / \left(\sigma^2_{\mathrm{b}} + \sigma^2_{\mathrm{w}}\right)$

2.4.2 Variable Operationalization

This subsection provides an overview of the operationalization of variables used in the study. Descriptive statistics for all variables are provided in Table 2.1.

2.4.2.1 Dependent Variable

We use data envelopment analysis (DEA) to estimate technical efficiency scores for each carrier at each of the five measurement occasions. DEA is an appropriate analytic strategy to assess relative efficiency of productive entities (*i.e.*, decision-making units) over time in longitudinal settings (Bogetoft and Otto, 2010). Since its introduction to the logistics literature (see Clarke and Gourdin, 1991; Kleinsorge et al., 1991), DEA has been widely employed to assess efficiency of productive entities in the transportation sector (*e.g.*, Mejza and Corsi, 1999; Scheraga, 2004; Weber and Weber, 2004; Ha et al., 2010; Rogers and Weber, 2011; Wanke, 2014; Min and Ahn, 2017). We select DEA because it is suited to multiple-input, multiple-output models and can estimate efficiency scores without requiring distributional assumptions. Further, commonly used econometric production or cost models (*e.g.*, those based on least-squares estimation) assume that all productive entities are technically efficient (Coelli et al., 2005). In contrast, this paper seeks to directly examine a determinant of technical inefficiency, necessitating the use of a benchmarking technique

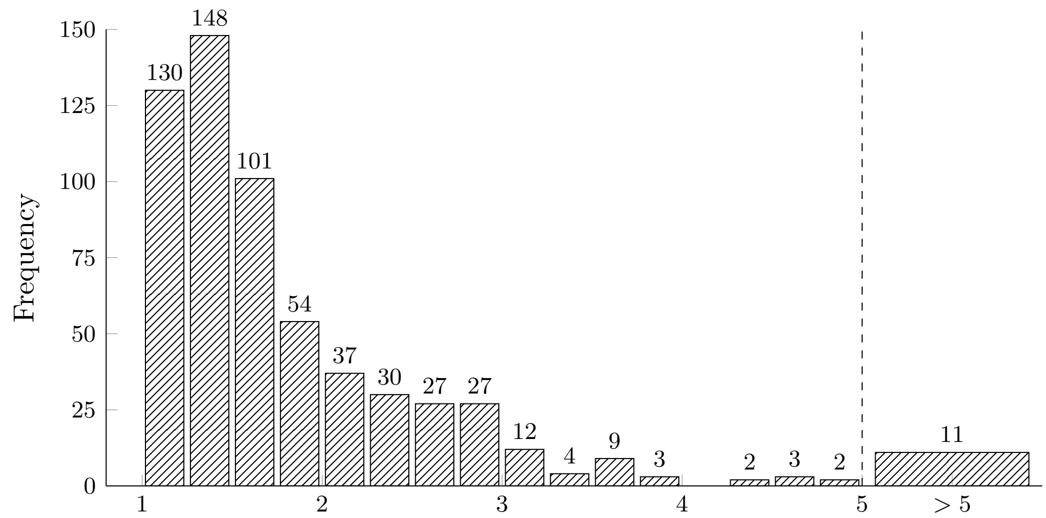


Figure 2.2: Distribution of estimated output-oriented DEA efficiency (operational inefficiency) scores.

to derive efficiency scores. Malmquist productivity indices are not created, as has often been done in prior carrier efficiency research (*e.g.*, McMullen and Okuyama, 2000; Scheraga, 2011), since we do not seek to examine changes in productivity between time periods. A histogram of estimated efficiency scores is provided in Figure 2.2; we discuss our specification of the DEA model which produced these scores next.

We specify an output-oriented DEA model since the focus of our theorizing is on improvements to the utilization of productive resources and, thus, output efficiency or enhancement (Cook et al., 2014). Our use of an output-oriented model is consistent with previous work on motor carrier efficiency (*e.g.*, McMullen and Okuyama, 2000; Scheraga, 2011). We view carriers as having a common production possibility set $T^t = \{(x^t, y^t) \in \mathbb{R}_+^m \times \mathbb{R}_+^n \mid x^t \text{ can produce } y^t\}$, where m is a vector of inputs consisting of the number of power units (x_1) and drivers (x_2) for carrier i during period t and n is a vector of outputs consisting of the miles produced[4] of TL service (y_1) and LTL service (y_2) for the

[4]The use of annual miles is consistent with Feitler et al. (1998), who view efficiency for less-than-truckload firms as a (ratio) relationship between annual miles and number of power units. However, we also consider a labor input, recognizing that classical economic theory of productivity (*e.g.*, the Cobb-Douglas production function) and the literature on transportation economics in the motor carrier industry considers both labor and capital assets as inputs to production.

same carrier and period. TL and LTL miles are treated separately given differences in the nature of output between the two segments (McMullen and Tanaka, 1995; McMullen and Okuyama, 2000); solving for a frontier using aggregates of these outputs would be clearly inappropriate given our hypotheses. Thus, given the two outputs, a production possibilities frontier (concave output isoquant) is solved, representing the various combinations of outputs that a carrier might efficiently produce given some level of input. We allow for variable returns to scale (Banker et al., 1984), an assumption that is appropriate for most production contexts (Talluri et al., 2013) and that is consistent with recent literature on carrier efficiency. For instance, Fawcett et al. (2016, p. 48) state, "it is clear that it would be overly simplistic to assume the complete absence of economies of scale in the motor carrier industry." The distance function for our output-oriented model is defined in Equation 2.1. Output-based distances for each carrier at each year are evaluated against the efficiency frontier and these observations are assigned a Farrell output efficiency score θ which is lower-bounded at a value of one (corresponding to efficient carriers). Thus, we refer to this variable as *operational inefficiency*.

$$\max_{\theta,\lambda} \theta, \tag{2.1}$$

$$\text{s.t.} \quad \begin{cases} \sum_{i=1}^{n} \lambda_i x_{ji} \leq x_{jp} \quad \forall j \\ \sum_{i=1}^{n} \lambda_i y_{ki} \geq \theta y_{kp} \quad \forall k \\ \sum_{i=1}^{n} \lambda_i = 1 \\ \lambda_i \geq 0 \quad \forall i \end{cases} \tag{2.2}$$

2.4.2.2 Focal Predictor

We operationalize our focal predictor, *LTL concentration*, as the ratio of LTL miles produced by a carrier in a given year to total miles (LTL + TL) produced by that carrier in the same year. Thus, possible values of *LTL concentration* range from a zero, corresponding to occasions in which a carrier produced only TL carriage, to one, which corresponds to occasions in which a carrier produced only LTL carriage. The observed distribution of *LTL*

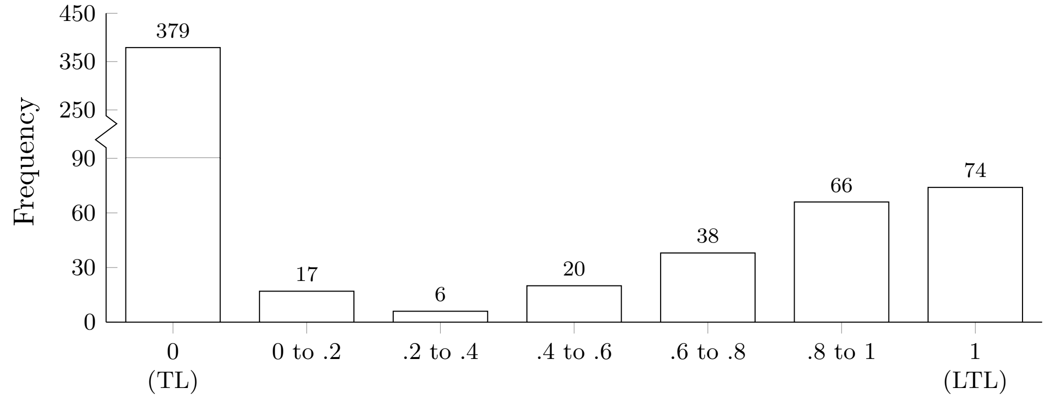


Figure 2.3: Histogram of carrier output concentration within the less-than-truckload segment (LTL concentration).

concentration is shown in Figure 2.3. The transportation literature has tended to bifurcate carriers between the two segments (Scheraga, 2011); 31 of the 120 carriers in our sample (25.8%) concurrently produced both TL carriage and LTL carriage, thus participating in both segments. Further, carrier *LTL concentration* was fairly stable over our period of analysis with an intraclass correlation (ICC) of .94, indicating that the overwhelming majority of variability in *LTL concentration* exists between (and not within) carriers. The sample data does not suggest that average levels of *LTL concentration* differ in the population between Class I carriers ($m = .37, \text{sd} = .43$) and Class II carriers ($m = .23, \text{sd} = .38$), based on estimation within a linear mixed model ($\chi^2_\Delta(1) = 3.74, p > .05$). Nor is carrier class (I or II) a significant predictor of a carrier's participation in both market segments; the odds ratio associated with carrier class did not differ significantly from a ratio of one when estimated within a binomial (logit link) generalized estimating equation ($OR = .72, \text{SE(robust)} = 1.21, p > .05$).

2.4.2.3 Covariates

Carrier Class. We control for carrier class to hold constant any residual, conditional differences in carrier inefficiency due to size (*i.e.*, after variable returns to scale is accounted for in the DEA model). Carrier class is assigned using the procedures at 49 C.F.R. § 369. Class I carriers are defined as those carriers having adjusted annual carrier operating revenues (after revenue deflation to 1994 dollars) of more than 10 million dollars. Class II carriers are defined as those carriers having adjusted annual carrier operating revenues of more than 3 million dollars but less than 10 million dollars. Our sample of 120 carriers is comprised of 53 Class I carriers and 67 Class II carriers; there were no instances of class promotion or demotion during our sampling period. We code *carrier class* using weighted effects coding such that groups are compared with the aggregate population mean (Cohen et al., 2003); Class I carriers are used as the base group.

Year. We include year fixed effects to control for unobserved, carrier-invariant sources of variability in production inefficiency that occur over time (*e.g.*, macroeconomic conditions, regulatory changes). *Year* dummy variables were constructed using unweighted effects coding such that measurement occasions (years) are compared against the overall unweighted mean. 1999 is used as the base year.

Outsourcing. We operationalize *outsourcing* as the proportion of a carrier's total operating expenses at a measurement occasion arising from equipment rentals (with and without drivers) and carrier-procured transportation services. Thus, this variable accounts for carrier expenses incurred under contract with owner-operators and other carriers for the movement of shipments, as well as contracts with third-party firms for the use of revenue vehicles. In each of these instances, the carrier has limited control over drivers and/or capital equipment used in production. In their analysis of carrier utilization of owner-operators, Han et al. (2008) show that carriers who participate in the LTL market segment are less likely to outsource freight movements. Thus, if *outsourcing* is also related to *operational inefficiency*, its omission from our regression equation may bias coefficient estimates for *LTL concentration*. We include *outsourcing* as a covariate on these grounds and mean-center the variable prior to entry into the regression equation for interpretability of the intercept.

2.4.3 Statistical Analysis

Statistical analysis is performed using the mixed effect modeling framework. A mixed effects model is selected since it offers flexibility in the specification of regression models for longitudinal data, accounting for both population characteristics and subject-specific effects (Fitzmaurice et al., 2012). Our model is defined in Equation 2.3. In this model, output-oriented DEA efficiency scores (*operational inefficiency*; Y_{it}) for carrier i at a measurement occasion t are a function of the carrier's class, the year, the extent of outsourcing and a cubic polynomial of *LTL concentration*. We allow for a randomly varying subject effect and specify three models: a model which includes only covariates while excluding the *LTL concentration* polynomial (Control Model); a model which includes all variables and assumes no within-group residual correlations (Model A); and a model with includes all variables and allows for an AR(1) autoregressive process in the within-group residual correlation structure (Model B). Estimation is performed by maximum likelihood.

Level 1

$$\begin{aligned} Y_{it} = \beta_{0i} + \beta_1 \times \text{CLASS}_i + \sum_{j=2}^{5} \beta_j \times \text{YEAR}_t + \beta_6 \times \text{OUTSOURCING}_{it} \\ + \beta_7 \times \text{LTL\_CONCENTRATION}_{it} + \beta_8 \times (\text{LTL\_CONCENTRATION}_{it})^2 \\ + \beta_9 \times (\text{LTL\_CONCENTRATION}_{it})^3 + R_{it} \end{aligned} \tag{2.3}$$

Level 2

$$\beta_{0i} = \gamma_{00} + U_{0i} \tag{2.4}$$

2.5 Results

The results of estimation for the three models are presented in Table 2.2. Model B, the AR(1) model, offered superior fit to the data over the control model ($\chi^2_\Delta(4) = 41.11, p < .01$) and over Model A, the variance-components model ($\chi^2_\Delta(1) = 22.96, p < .01$). The effects from the cubic polynomial were jointly significant ($\chi^2_\Delta(3) = 16.89, p < .01$), and predictions of operational inefficiency in Model B explained 81.34 percent of the total variability in DEA-estimated operational inefficiency scores and 15 percent of Level 1 (within carrier) variability in DEA-estimated operational inefficiency scores when approximated using Bryk and Raudenbush's (2002) proposed technique.

Table 2.2: Results of mixed effect model estimation.

		Control Model ($R = \text{VC}$)		Model A ($R = \text{VC}$)		Model B ($R = \text{AR(1)}$)	
Response=Operational inefficiency		Est.	S.E.	Est.	S.E.	Est.	S.E.
Fixed effects							
Intercept	γ_{10}	1.85^{**}	.09	1.98^{**}	.11	1.98^{**}	.12
CLASS:II	β_1	$.19^{*}$	.08	$.15^{\dagger}$	.10	.15	.09
YEAR:2000	β_2	$-.01$	.05	$<.01$	.04	$<.01$	.04
YEAR:2001	β_3	.01	.05	$<.01$	.04	$<.01$	.04
YEAR:2002	β_4	$-.06$	.05	$-.06$	.04	$-.06$	.04
YEAR:2003	β_5	$-.07$	.05	$-.06$	.04	$-.06$	.05
OUTSOURCING	β_6	.54	1.19	.92	1.17	.48	1.25
LTL_CONCENTRATION	β_7			$4.03^{\dagger}$	2.12	$3.88^{\dagger}$	2.33
LTL_CONCENTRATION^2	β_8			-15.91^{**}	5.32	-15.46^{**}	5.77
LTL_CONCENTRATION^3	β_9			11.93^{**}	3.38	11.62^{**}	3.62
Variance Components							
Standard Deviation (b_{0i})	σ_b	$.92^{**}$	.07	1.04^{**}	.10	1.05^{**}	.11
Standard Deviation (e_{it})	σ_e	$.55^{**}$	.02	$.52^{**}$	.02	$.60^{**}$	.04
Autoregressive (AR1)	φ					$.41^{**}$	.08
Measures of fit							
-2 Log likelihood		$1,307.00$		$1,288.85$		$1,265.89$	
AIC		$1,325.00$		$1,312.85$		$1,291.89$	
AICc		$1,325.31$		$1,313.38$		$1,292.52$	
pR^2 (Level 1)		.02		.11		.15	

Notes: $^{\dagger}p < .10$, $^{*}p < .05$, $^{**}p < .01$ (two-tailed); R^2 for Level 1 approximated as $(\sigma^2_{\text{null}} - \sigma^2_{\text{full}})/\sigma^2_{\text{null}}$ (Bryk and Raudenbush, 2002), where the covariance structure of the R matrix in the null model is of the same type (*e.g.*, VC, AR(1)) as the model being tested.

The model-estimated cubic relationship between *LTL concentration* and *operational inefficiency* is shown in Figure 2.4. Polynomial inflections estimated by the model are comparable to those theorized within Figure 2.1. Hypothesis 1 suggested that, for pure TL carriers, *LTL concentration* will have a positive marginal effect on *operational inefficiency*. Indeed, a positive marginal effect was estimated at this level of *LTL concentration*, but was significant under a two-tailed hypothesis test only at an α of .10 ($\beta = 3.88, t = 1.66, p < .10$). Hypothesis 2 suggested that, for pure LTL carriers, *LTL concentration* will also have a positive marginal effect on *operational inefficiency* (thus, a decrease *LTL concentration* will be associated with greater operational efficiency). Consistent with this hypothesis, a positive marginal effect of *LTL concentration* was estimated for pure LTL carriers ($\beta = 7.81, t =$

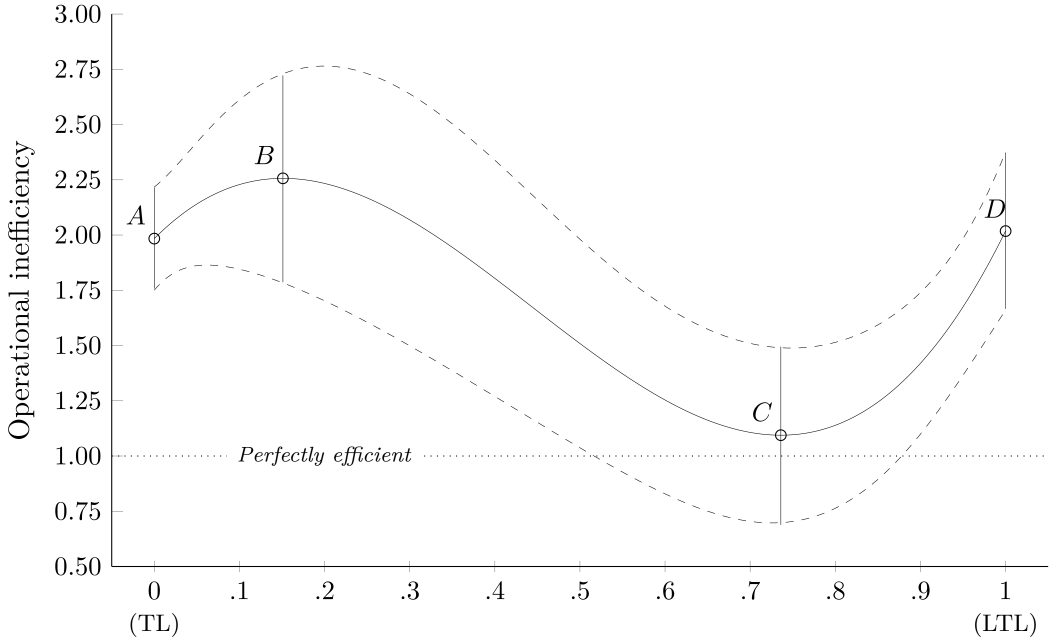


Figure 2.4: Model-estimated relationship between LTL concentration and production inefficiency, with 95% confidence bands.

$4.11, p < .01$).

Prior to estimating regions of significance for the estimated polynomial, we test for mean differences between the four points identified in Figure 2.4 (pure TL, local maximum, local minimum, pure LTL) within a Monte Carlo simulation. $10,000$ draws were taken from a multivariate normal distribution using the model-produced mean vector and variance-covariance matrix of parameter estimates; Bonferroni corrections are made to maintain a family-wise error rate of .05. From Model B, a local maximum in the polynomial was found at an *LTL concentration* of .15 (85% truckload), where *operational inefficiency* was estimated to be the highest across all levels of *LTL concentration* at a value of 2.26. A local minimum was found at an *LTL concentration* of .74, where *operational inefficiency* was estimated to be the lowest across all levels of *LTL concentration* at a value of 1.09. The results of comparisons between model-derived estimates at these two points and those at levels of *LTL concentration* corresponding to Pure TL and Pure LTL are listed in Table 2.3. Notably estimates of *operational inefficiency* do not differ significantly between pure TL carriers ($m = 1.98$) and pure LTL carriers ($m = 2.02$).

Table 2.3: Monte Carlo simulation results.

Point	LTL Concentration	Est.	S.E.	Absolute Mean Difference (A)	(B)	(C)
(A)	Pure Truckload	1.98	.12			
(B)	85% Truckload†	2.26	.24	.28		
(C)	74% Less-than-truckload‡	1.09	.21	.89*	1.17*	
(D)	Pure Less-than-truckload	2.02	.18	.04	.24	.93*

Notes: *mean difference differs from zero at $\alpha = .05/6$ (two-tailed); †local maximum $\hat{\theta}^*$ for ≥50% TL; †local minimum $\hat{\theta}^*$ for ≥50% LTL.

2.5.1 Regions of Significance

We estimate regions of significance of the simple slope (rate of change) of the cubic polynomial *LTL concentration* on *operational inefficiency* using the Johnson-Neyman Technique (see Miller et al., 2013). We provide a derivation for regions of significance of a cubic polynomial here. Given the linear combination of regression coefficients, an estimate of the variance of a simple slope is given by:

$$s_b^2 = \boldsymbol{w'} \boldsymbol{S_b} \boldsymbol{w} \tag{2.5}$$

The weight column vector $\boldsymbol{w}$ is the first partial derivative of the design matrix with respect to the explanatory variable of interest (X), here $\boldsymbol{w'} = \begin{bmatrix} 1 & 2X & 3X^2 \end{bmatrix}$. $\boldsymbol{S_b}$ is the estimated 3×3 variance-covariance matrix of regression coefficients. In this case of a curvilinear relationship between X and Y, where raw polynomials up the third degree (cubic) are included in the regression equation and where the effect of X on Y does not depend on the value of some other variable(s), the standard error for the simple slope of X on $\hat{Y}$ is:

$$s_b = \sqrt{s_{11} + 4Xs_{12} + 6X^2 s_{13} + 4X^2 s_{22} + 12X^3 s_{23} + 9X^4 s_{33}} \tag{2.6}$$

The confidence bands of the simple slope are then solved for zero, representing those values of the explanatory variable where the slope crosses the threshold of significance (Miller et al., 2013). This can be simplified to solving the absolute value of the t-value for the critical value from the t distribution:

$$t_{\text{crit}} = \pm \frac{\beta_1 + 2\beta_2 X + 3\beta_3 X^2}{\sqrt{s_{11} + 4Xs_{12} + 6X^2 s_{13} + 4X^2 s_{22} + 12X^3 s_{23} + 9X^4 s_{33}}} \tag{2.7}$$

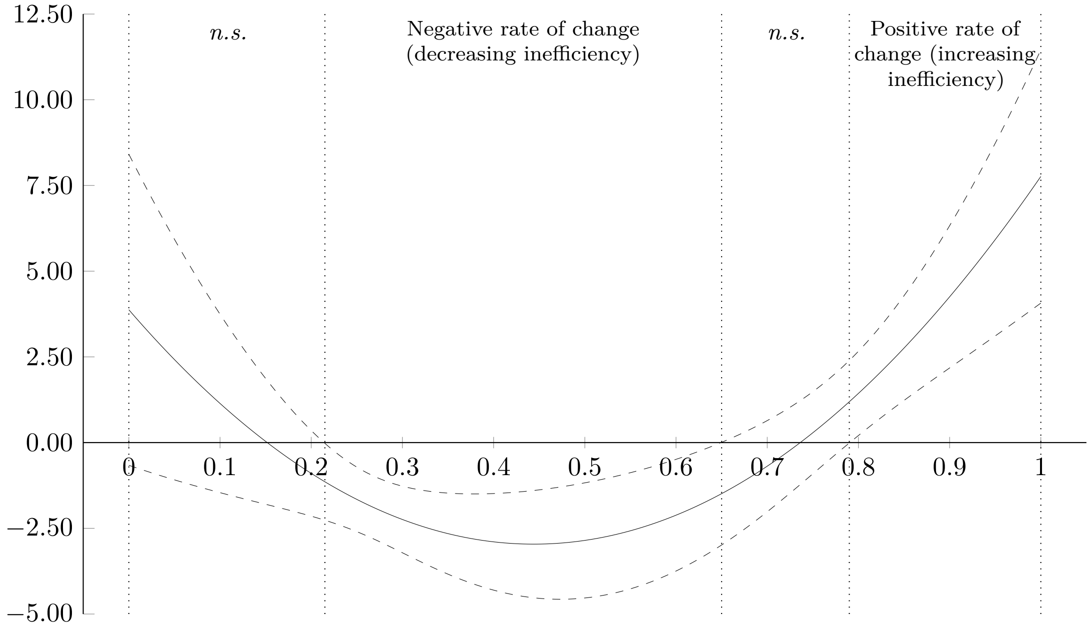


Figure 2.5: Johnson-Neyman plot of the regions of significance for the simple slope of LTL concentration on operational inefficiency.

The resulting Johnson-Neyman plot is shown in Figure 2.5. Two regions of significance were found. The marginal effect of *LTL concentration* on *operational inefficiency* is significantly negative between *LTL concentration* values of .22 and .65. This effect turns significantly positive between *LTL concentration* values of .79 and 1.00. The marginal effect of *LTL concentration* on *operational inefficiency* is nonsignificant at all other values of *LTL concentration*.

2.5.2 Alternative Specifications

Two alternative specifications to the AR(1) cubic model (Model B) are considered. First, we reduce the degree of the polynomial for *LTL concentration* to a quadratic (Model C), thus testing for significance of a second inflection.[5] Next we impose a harmonic form on the non-linear relationship between *LTL concentration* and *operational inefficiency* using

[5]An AR(1) model including controls and only a linear effect of *LTL concentration* on *operational inefficiency* (*i.e.*, polynomial degree = 1) does not differ significantly from an AR(1) control model $\left(\chi^2_\Delta(1) = 0.60, p = .44\right)$.

sinusoidal regression (Model D), in order to estimate a more parsimonious model with symmetry in the curvatures. This is accomplished first using a single sine wave (Equations 2.8 and 2.9), matching the harmonic in Figure 2.1 with respect to phase (φ, implicitly zero) and frequency (f, implicitly one), but leaving amplitude (β_{10}) free for estimation. Estimates of amplitude in harmonic regression are known to be asymptotically efficient (*e.g.*, under ordinary least squares estimator; see Durbin, 1960); we continue to allow AR(1) covariances in the R-matrix, recognizing that these within-carrier residuals are correlated and may otherwise bias standard errors.

Level 1

$$Y_{it} = \beta_{0i} + \beta_1 \times \text{CLASS}_i + \sum_{j=2}^{5} \beta_j \times \text{YEAR}_t + \beta_6 \times \text{OUTSOURCING}_{it} \tag{2.8}$$

$$+ \beta_{10} \times \sin\left(2\pi \times \text{LTL\_CONCENTRATION}_{it}\right) + R_{it}$$

Level 2

$$\beta_{0i} = \gamma_{00} + U_{0i} \tag{2.9}$$

The results of estimation for the two alternative specifications are presented in Table 2.4. Estimation of Model C resulted in a convex quadratic relationship between *LTL concentration* and *operational inefficiency*; coefficents of the polynomial are jointly significant ($\chi^2_\Delta(2) = 6.82, p = .03$). However, consistent with our theoretical reasoning for a cubic relationship, the quadratic model offers inferior fit to the sample ($\chi^2_\Delta(1) = 10.07, p < .01$). Similarly, the less flexible (but more parsimonious) sinusoidal model (Model D) improved fit over an AR(1) control model ($\chi^2_\Delta(1) = 6.50, p = .01$), with an estimated sine wave amplitude of 0.42 ($t(475) = 2.64,\ p < .01$). Comparisons between non-nested models B and D are made using Akaike information criterion (AIC), as recommended by Singer and Willett (2003) and Fitzmaurice et al. (2012). Both AIC and second-order AIC (AICc) give preference to Model B, the cubic model.

Table 2.4: Results of estimation for alternative models.

Response=Operational inefficiency		Model B (R = AR(1)) Est.	S.E.	Model C (R = AR(1)) Est.	S.E.	Model D (R = AR(1)) Est.	S.E.
Fixed effects							
Intercept	γ_{10}	1.98**	.12	1.99**	.12	1.90**	.10
CLASS:II	β_1	.15	.09	.16†	.09	.17†	.09
YEAR:2000	β_2	<.01	.04	−.01	.04	−.01	.04
YEAR:2001	β_3	<.01	.04	<.01	.04	<.01	.04
YEAR:2002	β_4	−.06	.04	−.05	.04	−.06	.04
YEAR:2003	β_5	−.06	.05	−.06	.05	−.06	.05
OUTSOURCING	β_6	.48	1.25	.31	1.27	.25	1.26
LTL_CONCENTRATION	β_7	3.88†	2.33	−2.77**	.95		
LTL_CONCENTRATION^2	β_8	−15.46**	5.77	2.63**	.96		
LTL_CONCENTRATION^3	β_9	11.62**	3.62				
sin ($2\pi \times$ LTL_CONCENT)	β_{10}					.42**	.16
Variance Components							
Standard Deviation (b_{0i})	σ_b	1.05**	.11	1.03**	.09	.97**	.08
Standard Deviation (e_{it})	σ_e	.60**	.04	.62**	.05	.62**	.04
Autoregressive (AR1)	φ	.41**	.08	.45**	.08	.42**	.08
Measures of fit							
−2 Log likelihood		1,265.89		1,275.97		1,276.29	
AIC		1,291.89		1,299.97		1,298.29	
AICc		1,292.52		1,300.50		1,298.74	
pR^2 (Level 1)		.15		.09		.10	

Notes: $^\dagger p < .10$, $^* p < .05$, $^{**} p < .01$ (two-tailed); R^2 for Level 1 approximated as $(\sigma^2_{\text{null}} - \sigma^2_{\text{full}})/\sigma^2_{\text{null}}$ (Bryk and Raudenbush, 2002).

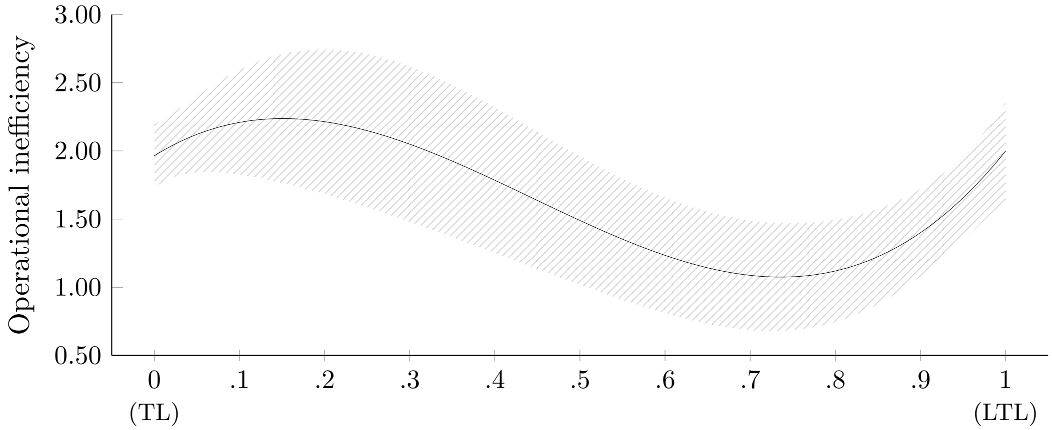
2.5.3 Robustness Tests

2.5.3.1 Subsamples

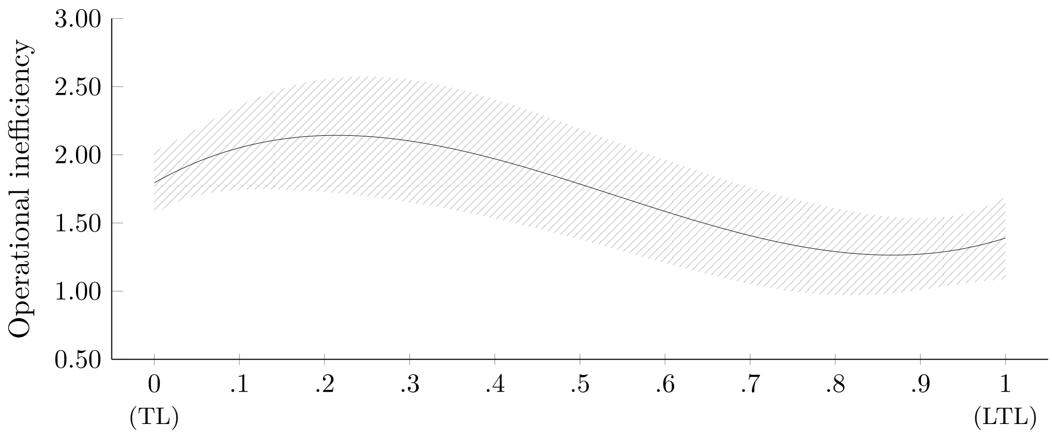
We first test for robustness of the regression results using subsamples by carrier class. The full sample of carrier data ($n = 600$) was divided into two subsamples consisting of Class I motor carriers ($n = 265$) and Class II motor carriers ($n = 335$). Model B (the cubic model) was estimated separately for each subsample; estimation results are presented in Table 2.5. The analyses suggest a structural difference between carrier classes; a $\chi^2(11)$ difference of 116.63 between the aggregated model (Model B) and disagreggated (subsample) models is statistically significant at $\alpha = .05$ ($p < .01$). Parameter estimates of the cubic function in the subsample models were jointly significant for Class I carriers ($\chi^2_\Delta(3) = 19.25, p < .01$) and Class II carriers ($\chi^2_\Delta(3) = 17.22, p < .01$). For pure truckload carriers, the instantaneous rate of change in *operational inefficiency* for a unit increase in *LTL concentration* is significantly positive at $\alpha = .05$ for both Class I carriers ($m = 3.54, t = 1.98, p = .05$) and Class II carriers ($m = 8.61, t = 2.14, p = .03$). Thus, while Hypothesis 1 was only supported at $\alpha = .10$ when carriers classes were aggregated within Model B, disaggregation to subsamples offers stronger statistical evidence. Lastly, as with the aggregated model (Model B), Hypothesis 2 was supported at $\alpha = .05$ for Class II carriers, where the instantaneous rate of change in *operational inefficiency* for a unit decrease in *LTL concentration* is significantly positive ($m = 13.20, t = 4.27, p < .01$); however, this effect is weaker and nonsignificant for Class I carriers ($m = 1.99, t = 1.25, p = .21$).

2.5.3.2 Censored Regression

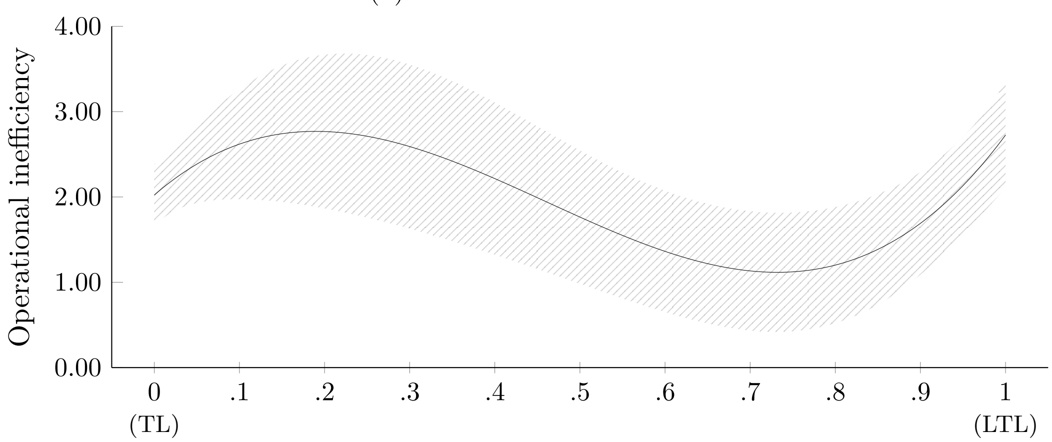
Previous motor carrier efficiency studies using two-stage DEA analysis (*e.g.*, McMullen and Okuyama, 2000; Scheraga, 2011) have argued that DEA efficiency scores are censored and, as such, have employed Tobit regression to estimate the second-stage regression model. The appropriateness of this statistical strategy (and other statistical strategies when DEA efficiencies are regressed) is the subject of intense debate within the literature (Liu et al., 2016). For instance, McDonald (2009) contends that the use of Tobit models is inappropriate since DEA efficiency scores are not generated by a censoring process and, instead, represent fractional data. Nonetheless, recognizing that our data contains a lower bound, we estimate a random-effects Tobit regression to assess the extent to which our results might be impacted by any censoring in the data generating process. We estimate two models—one with time fixed effects (Tobit A) and one without time fixed effects (Tobit B)—since the inclusion of fixed effects within non-linear panel data models using finite samples may produce biased and inconsistent estimates under the maximum likelihood estimator (Greene, 2004). The results of estimation are presented in Table 2.6; within each model, coefficients for the cubic polynomial are comparable to those previously estimated within a linear mixed



(a) Full sample



(b) Class I Motor Carriers



(c) Class II Motor Carriers

Figure 2.6: Plots of the effects of LTL concentration on operational inefficiency for the full sample (a) and subsamples of Class I (b) and Class II (c) motor carriers.

Table 2.5: Results of estimation for subsamples.

		Full Sample (R = AR(1))		Class I (R = AR(1))		Class II (R = AR(1))	
Response=Operational inefficiency		Est.	S.E.	Est.	S.E.	Est.	S.E.
Fixed effects							
Intercept	γ_{10}	1.98**	.12	1.79**	.12	2.02**	.19
CLASS:II	β_1	.15	.09				
YEAR:2000	β_2	<.01	.04	−.02	.05	<.01	.06
YEAR:2001	β_3	<.01	.04	.01	.05	−.01	.06
YEAR:2002	β_4	−.06	.04	−.02	.05	−.08	.06
YEAR:2003	β_5	−.06	.05	−.05	.04	−.07	.09
OUTSOURCING	β_6	.48	1.25	−.11	1.10	.91	1.99
LTL_CONCENTRATION	β_7	3.88†	2.33	3.54*	1.79	8.61*	4.02
LTL_CONCENTRATION^2	β_8	−15.46**	5.77	−10.28*	4.33	−28.61**	10.07
LTL_CONCENTRATION^3	β_9	11.62**	3.62	6.34*	2.75	20.70**	6.30
Variance Components							
Standard Deviation (b_{0i})	σ_b	1.05**	.11	.71**	.09	1.22**	.17
Standard Deviation (e_{it})	σ_e	.60**	.04	.34**	.02	.80**	.11
Autoregressive (AR1)	φ	.41**	.08	−.33**	.10	.60**	.08
Measures of fit							
−2 Log likelihood		1,265.89		351.52		797.75	
AIC		1,291.89		375.51		821.75	
AICc		1,292.52		376.75		822.72	
pR^2 (Level 1)		.15		.12		.17	

Notes: $^\dagger p < .10$, $^* p < .05$, $^{**} p < .01$ (two-tailed); R^2 for Level 1 approximated as $(\sigma^2_{\text{null}} - \sigma^2_{\text{full}})/\sigma^2_{\text{null}}$ (Bryk and Raudenbush, 2002).

model (Model B). Consistent with Hypothesis 1, the marginal effect of *LTL concentration* on *operational inefficiency* for pure TL carriers was positive and significant for Tobit A ($m = 4.53, t = 2.08, p = .03$) and for Tobit B ($m = 4.46, t = 1.88, p = .02$). Similarly, a positive marginal effect of *LTL concentration* on *operational inefficiency* for pure LTL carriers was positive and significant in Tobit A ($m = 7.56, t = 1.17, p < .01$) and Tobit B ($m = 8.07, t = 1.14, p < .01$), as predicted under Hypothesis 2.

2.6 Discussion

Our study bridges Thompson's (1967) conceptualization of interdependence in organizational workflows with the Penrosian theoretical perspective of firm growth and utilization of productive resources. Using archival data from the motor carrier industry, we find evidence

Table 2.6: Results of estimation for random-effects Tobit regression models.

Response=Operational inefficiency		Tobit A (Year Effects) Est.	Tobit A (Year Effects) S.E.	Tobit B (No Year Effects) Est.	Tobit B (No Year Effects) S.E.
Fixed effects					
Intercept	γ_{10}	1.96^{**}	.18	1.97^{**}	.18
CLASS:II	β_1	.15	.11	.14	.11
YEAR:2000	β_2	$<.01$	.10		
YEAR:2001	β_3	.01	.09		
YEAR:2002	β_4	$-.06$	.13		
YEAR:2003	β_5	$-.06$	.11		
OUTSOURCING	β_6	.72	1.88	.54	1.94
LTL_CONCENTRATION	β_7	4.53^{*}	2.08	4.46^{*}	1.88
LTL_CONCENTRATION^2	β_8	-16.58^{**}	4.30	-16.93^{*}	3.98
LTL_CONCENTRATION^3	β_9	12.06^{**}	2.41	12.49^{*}	2.27
Variance Components					
Standard Deviation (b_{0i})	σ_b	1.03^{**}	.13	1.05^{**}	.13
Standard Deviation (e_{it})	σ_e	$.53^{**}$	.03	$.54^{**}$	.02
-2 Log likelihood		$1,141.33$		$1,149.56$	

Notes: $^{\dagger}p < .10$, $^{*}p < .05$, $^{**}p < .01$ (two-tailed).

that carrier diversification into a market segment with a higher degree of internal interdependence reduces operational efficiency. We also find evidence that carrier diversification into a market segment with a lower degree of internal interdependence increases operational efficiency. Relying on the work of Penrose (1959), we postulate that the mechanism underlying these relationships is carrier management's ability (or inability) to improve the utilization of unused and underutilized productive resources in the face of changing levels of interdependence among organization workflows. For motor carriers participating only in the TL market segment, our results suggest that diversification of production into the LTL market segment may lead to a decrease in overall operational efficiency. Production in the LTL market segment requires a substantially greater degree of coordination among carrier activities than does production in the TL market segment.

In LTL carriage, workflows exhibit a reciprocal pattern of interdependence, which we refer to as *intra-haul interdependency*. The results of our subgroup analysis suggest that this decrease in overall operational efficiency as carriers shift from *inter-haul sequential* *interdependency* is stronger for Class II (medium-sized) carriers than it is for Class I (large) carriers. When data across the two classes were aggregated, we found mixed support for this effect, which was statistically significant at α of .05 in a censored regression but was only statistically significant at α of .10 when data was viewed as fractional (not censored) and within-group residuals were allowed to covary as an AR(1) process.

We contend that this evidence regarding Hypothesis 1, although mixed in level of significance across statistical techniques, is compelling. Specifically, economic theory and management theory jointly suggest that firms should seek efficiency in their operations. Penrose (1959), for instance, argues that firm management will strive to most efficiently utilize productive resources. Our hypotheses suggested that, for a pure TL carrier, the marginal effect of LTL output production would be an increase in overall operational inefficiency. Thus, consistent with observed frequencies of carrier concentration in our dataset (Figure 2.3), one should expect to find, and does find, few carriers operating in this region of inefficient production. The lackluster level of statistical support offered by the AR(1) mixed effect model of aggregated carrier data may simply be a reflection of consistency between carrier behavior and theory. Given an absence of carrier concentration at this point of inefficiency, wider confidence intervals were estimated by the statistical model.

Our second Hypothesis suggested that motor carriers participating solely in the LTL market segment could better employ underutilized productive resources and, thus, improve overall operating efficiency, by diversifying production into the TL market segment. This hypothesis was broadly supported in our analyses; while it did not hold for one of our subsamples, the effect was prominent in the full data set. Our analysis of regions of significant for the simple slope of the effect of *LTL concentration* on *operational inefficiency* suggests that carriers producing between 79 percent and 100 percent of their output as LTL carriage might achieve additional gains to overall operational efficiency by increasing production in the TL market segment. As before, we rely on resource-based theory (Penrose, 1959) to argue that the underlying mechanism is carrier management's ability to improve the employment of unused and underutilized productive resources.

Our analysis carries important insights into the role of carrier size on operational efficiency. Mean levels of carrier operational efficiency only vary weakly, if at all, across the two carrier classes. This suggests that, in the motor carrier population, large carriers appear no more efficient than medium-sized carriers. However, our subgroup analysis demonstrates that structural differences do exist between these two classes. Specifically, while average efficiency levels may not differ between the classes, operational efficiency for large (Class I) carriers appears to be more robust than that of medium (Class II) carriers to changes in the focus of production between the two market segments. However, this also suggests that medium size carriers may be better able to capitalize on opportunities to improve the utilization of productive resources through diversification, but may also be more susceptible to the potentially damaging effects of increased interdependence in operations. Thus, our findings are consistent with those of Scheraga (2011), showing that carrier size attenuates the effects of *LTL concentration* (concentration of production on LTL carriage) on operational efficiency, while also providing a more fine-grained view into the nature of the moderated relationship.

2.6.1 Theoretical Contributions

A primary theoretical contribution of our study is the link between interdependence and a firm's ability to drive improvements in resource utilization through diversification. In the context of the motor carrier industry, we delineate between two types of interdependence: *intra-haul interdependency* and, for the TL market segment where coordination requirements are substantially lower, *inter-haul sequential interdependency*. This distinction in level of interdependence between segments, though subtle, has important risk-sharing implications for carriers, in addition to the production efficiency implications examined in this study. For instance, Silverman et al. (1997, p. 36) state that "the hub-and-spoke nature of LTL carriage makes drivers throughout the network highly interdependent, requiring a degree of coordination for which hierarchy (company driver) offers a lower cost than market (independent owner-operators). TL carriage requires far less coordination, making independents a more feasible contractual form." However, the implications of increased interdependence for LTL carriage extend well beyond the decision to outsource some portion of shipments to owner-operators. For instance, higher interdependence may impact capital investment decisions (*e.g.*, equipment maintenance, lease/purchase tradeoffs, investments in monitoring technologies) and the nature of shipper-carrier relationships.

For researchers, we believe that our study provides strong evidence against bifurcation of motor carriers by market segment. Clearly the market segments for TL carriage and LTL carriage are distinct, but many large and medium-sized motor carriages—more than 25 percent in our sample—participate concurrently within both segments. Our data also suggests that carriers of both classes (I and II) are equally likely to concurrently participate in both segments. The literature has viewed the output decision as a critical dimension of a motor carrier's strategic orientation (Feitler et al., 1998). We demonstrate the complex, non-linear relationship between this dimension and an important firm outcome—operational efficiency. Bifurcation of carriers into these segments forces an unnatural dichotimization on the industry and may significantly impair the ability to detect associations with critical carrier outcomes.

Lastly, our analysis paints a rich picture of the nature of variability in operational efficiency that exists within the motor carrier industry. Our DEA estimates of output-oriented efficiency indicate that carriers in our sample were perfectly efficient (on the efficiency frontier) during only four percent of the measurement occasions. While operational efficiency estimates were generally high across the sample (Figure 2.2), nearly 75 percent of the variability in operational efficiency was due to between-carrier differences—suggesting stability in efficiency levels over time. Much of the research into the motor carrier industry after regulatory reform has been focused on the issue of economies of scale; most frequently, this research has utilized econometric functions which assume firms to be perfectly efficient (Coelli et al., 2005). We provide evidence that this assumption my be unrealistic. Further, given that inefficient firms may be unable to capitalize on economies of scale (if economies do exist), we contend that researchers should first seek to better explain the stable, between-carrier differences that exist in operational efficiency. Thus, new theory is needed to explain the determinants of carrier inefficiency, along with empirics to asses its impact to operations and key business outcomes.

2.6.2 Managerial Implications

A carrier's decision of how much carriage to produce in each segment (TL and LTL) is an important, potentially even vital matter of strategy (Feitler et al., 1998), impacting the utilization of productive resources, altering interdependencies among workflows, changing coordination needs and coordination-related investments and influencing risk-sharing decisions. Market-driven rewards may even exist for carriers who can offer shippers a wider range of transportation services. Our study shows that the diversification decision impacts a carrier's overall level of operational efficiency. Our analysis provides managers with direct estimates of industry average levels of efficiency across the production possibilities curve (Figure 2.4) as well as estimated effects of altering the production decision (Figure 2.5). Our results offer separate conclusions for carriers that predominantly participate in the LTL segment and for those that predominantly participate in the TL segment.

Core LTL carriers committing sufficient production capacity (*i.e.*, resources) toward the TL carriage may be able to realize sizable gains in efficiency. Among the carriers in our sample, operational efficiency was greatest for those allocating 74 percent of production output to LTL carriage (where output is measured in miles) and 26 of production output to TL carriage. We theorize that, at this point, carriers are best able to leverage underutilized productive resources (*e.g.*, labor, capital equipment) toward TL production. An overcommitment of production capacity to TL carriage may have undesirable consequences; operational inefficiency increases, on average, once TL production exceeds 26 percent. Thus, managers seeking to improve resource utilization and overall efficiency through a diversification strategy into the TL segment should be mindful of the balance between a sufficient commitment and over-commitment of productive resources and production capacity.

Core TL carriers diversifying (by entering the LTL segment) will not experience the immediate efficiency gain their LTL peers diversifying into TL enjoy. Instead, these core TL carriers appear to face significant coordination challenges associated with interdependence in the production of LTL carriage. Specifically, core TL carriers who fail to sufficiently penetrate the LTL market—those who allocate less than 35 percent of production output toward the LTL carriage—risk a substantial increase in production inefficiency. McMullen and Okuyama (2000) suggest that carriers who expand into the LTL segment may do so to pursue a differentiation strategy. While differentiation or even full-line diversification (Penrose, 1959) may indeed be the underlying impetus for core TL carriers to expand into the LTL market, managers should be cognizant of the level of investment required to remain competitive. Barring sufficient resource investment and market penetration, expansion challenges are likely to manifest themselves in the form of operational inefficiency.

Accordingly, it is imperative to note that achieving a high level of production efficiency (*i.e.*, technical efficiency) is not an end in itself. Cost reductions realized from gains in production efficiency may be overshadowed by revenues forgone when a carrier's product mix is suboptimal in the marketplace. Thus, carriers must jointly consider allocative efficiency, including shipper preferences and marginal revenues in product lines, within their production decisions. As one example, it may be the case that carriers offering a complete range of services can demand a price premium from certain classes of shippers, such those who prefer fewer but closer relationships with their transportation service providers. Given the general competitiveness of the industry it is unclear how widespread these premiums are in the transportation marketplace. Existing evidence has instead emphasized the importance of lean and efficient production. For example, recent research by Fawcett et al. (2016) provides insight into the relationship between resource leanness, resource slack and financial outcomes for motor carriers operating in the TL segment. Their research indicates that leanness is a key driver of financial performance for motor carriers, as has been found in several other industries. Thus, Fawcett et al.'s (2016) findings provide underlying support to the importance of fully utilizing carrier resources, as we argue here, while drawing a link to firm financial performance. We caution managers, however, not to neglect the marketplace and the need for allocative efficiency within their production decisions.

2.6.3 Limitations and Directions for Future Research

We identify two primary limitations of our study. First, our sample, while advantageous to our test of theory, may not provide an accurate representation of the current or future population of the U.S. motor carrier industry. Thus, caution should be exercised in the generalization of our findings. Second, we utilized a two-stage data analysis strategy where we estimated DEA efficiency scores in a first stage and then, in a second stage, regressed these estimates on a set of predictors. DEA scores represents benchmarks of efficiency and may contain error, since they are only estimates of a productive unit's level of efficiency relative to a sampled population. While we believe that our sample of 600 observations is sufficiently large to avoid severe errors in our solution of a production possibilities frontier, we recognize that non-random error in the resulting efficiency estimates might bias statistical results of our second-stage analysis.

We contend that explaining stable, between-carrier differences in production efficiency remains a fruitful avenue for future research. Such research would not only increase our theoretical understanding of the motor carriers, but could be impactful to the motor carrier industry, shippers and larger economies. Additionally, we believe that greater insight is warranted into the role of interdependence in the motor carrier industry. For instance, future research could examine implications for risk-sharing, as well as coordination challenges associated with greater interdependence.

2.7 Conclusion

Our study examined the relationship between a motor carrier's diversification of production into a new market segment and operational efficiency. Using panel data obtained from transportation archives, we provide evidence that the effects of diversification on operational efficiency depend on shifts in organizational interdependence, impacting a carrier's ability to efficiently employ existing, underutilized resources in the new segment.

REFERENCES

REFERENCES

American Trucking Association (ATA) (2017). *Reports, Trends & Statistics.* Retrieved from http://www.trucking.org/News_and_Information_Reports_Industry_Data.aspx [Accessed 5-May-2017].

Bairam, E. I. and Kahya, E. (1998). Production versus cost functions: Unreliability of the duality theorem in accounting and economics. In Bairam, E. I., editor, *Production and Cost Functions*, chapter 3, pages 42–53. Ashgate, Brookfield, VT.

Banker, R. D., Charnes, A., and Cooper, W. W. (1984). Some models for estimating technical and scale inefficiencies in data envelopment analysis. *Management Science*, 30(9):1078–1092.

Beattie, B. R., Taylor, C. R., and Watts, M. J. (2009). *The economics of production.* Krieger Publishing Company, Malabar, Florida.

Belzer, M. H. (2000). *Sweatshops on wheels: Winners and losers in trucking deregulation.* Oxford University Press, New York, NY.

Bogetoft, P. and Otto, L. (2010). *Benchmarking with DEA, SFA, and R*, volume 157 of *Internation Sereis in Operations Research & Management Science.* Springer Science & Business Media, New York, NY.

Bryk, A. S. and Raudenbush, S. W. (2002). *Hierarchical linear models for social and behavioral research: Applications and data analysis methods.* Sage, Newbury Park, CA.

Bureau of Labor Statistics (BLS) (2017). *Industries at a glance: Truck transportation: NAICS 484.* Retrieved from https://www.bls.gov/iag/tgs/iag484.htm [Accessed 5-May-2017].

Business Wire (2017). *Knight Transportation and Swift Transportation announce all stock transaction with a combined enterprise value of $6 billion.* Retrieved from http://www.businesswire.com/news/home/20170410005557/en/Knight-Transportation-Swift-Transportation-Announce-Stock-Transaction [Accessed 5-May-2017].

Cantor, D. E., Celebi, H., Corsi, T. M., and Grimm, C. M. (2013). Do owner–operators pose a safety risk on the nation's highways? *Transportation Research Part E: Logistics and Transportation Review*, 59:34–47.

Cantor, D. E., Macdonald, J. R., and Crum, M. R. (2011). The influence of workplace justice perceptions on commercial driver turnover intentions. *Journal of Business Logistics*, 32(3):274–286.

Caves, D. W., Christensen, L. R., and Tretheway, M. W. (1984). Economies of density versus economies of scale: Why trunk and local service airline costs differ. *The RAND Journal of Economics*, 15(4):471–489.

Clarke, R. L. and Gourdin, K. N. (1991). Measuring the efficiency of the logistics process. *Journal of Business Logistics*, 12(2):17.

Cleveland, G., Schroeder, R. G., and Anderson, J. C. (1989). A theory of production competence. *Decision Sciences*, 20(4):655–668.

Coelli, T. J., Rao, D. S. P., O'Donnell, C. J., and Battese, G. E. (2005). *An introduction to efficiency and productivity analysis*. Springer, New York, NY, 2 edition.

Cohen, J., Cohen, P., West, S. G., and Aiken, L. S. (2003). *Applied multiple regression/correlation analysis for the behavioral sciences*. Routledge, New York, NY, 3 edition.

Cook, W. D., Tone, K., and Zhu, J. (2014). Data envelopment analysis: Prior to choosing a model. *Omega*, 44:1–4.

Corsi, T. M. and Grimm, C. M. (1987). Changes in owner-operator use, 1977–1985: Implications for management strategy. *Transportation Journal*, 26(3):4–16.

Costello, B. and Suarez, R. (2015). *Truck driver shortage analysis*. Technical report, American Trucking Associations: Arlington, VA.

Creel, M. and Farell, M. (2001). Economies of scale in the U.S. airline industry after deregulation: a Fourier series approximation. *Transportation Research Part E: Logistics and Transportation Review*, 37(5):321–336.

Department of Transportation (DOT) (2017). *Form M: Annual Report Form for Class 1 & Class 2 Motor Carriers of Property and Household Goods*. Retrieved from https://www.fmcsa.dot.gov/mission/form-m [Accessed 5-May-2017].

Durbin, J. (1960). The fitting of time-series models. *Revue de l'Institut International de Statistique*, 28(3):233–244.

Farrell, M. J. (1957). The measurement of productive efficiency. *Journal of the Royal Statistical Society*, 120(3):253–290.

Fawcett, A. M., Jin, Y. H., Hofer, C., Waller, M. A., and Brazhkin, V. (2016). Sweating the assets: Asset leanness and financial performance in the motor carrier industry. *Journal of Business Logistics*, 37(1):43–58.

Feitler, J. N., Corsi, T. M., and Grimm, C. M. (1997). Measuring firm strategic change in the regulated and deregulated motor carrier industry: An 18 year evaluation. *Transportation Research Part E: Logistics and Transportation Review*, 33(3):159–169.

Feitler, J. N., Corsi, T. M., and Grimm, C. M. (1998). Strategic and performance changes among LTL motor carriers: 1976-1993. *Transportation journal*, 37(4):5–12.

Fitzmaurice, G. M., Laird, N. M., and Ware, J. H. (2012). *Applied longitudinal analysis*. John Wiley & Sons, Hoboken, NJ.

Giordano, J. N. (2008). Economies of scale after deregulation in LTL trucking: A test case for the survivor technique. *Managerial and Decision Economics*, 29(4):357–370.

Greene, W. (2004). The behaviour of the maximum likelihood estimator of limited dependent variable models in the presence of fixed effects. *The Econometrics Journal*, 7(1):98–119.

Ha, H.-K., Yoshida, Y., and Zhang, A. (2010). Comparative analysis of efficiency for major Northeast Asia airports. *Transportation Journal*, 49(4):9–23.

Han, C., Corsi, T. M., and Grimm, C. M. (2008). Why do carriers use owner operators in the U.S. for-hire trucking industry? *Transportation Journal*, 47(3):22–35.

Hanna, J. B. and Maltz, A. (1998). LTL expansion into warehousing: A transaction cost analysis. *Transportation Journal*, 38(2):5–17.

Hoskisson, R. E. and Hitt, M. A. (1990). Antecedents and performance outcomes of diversification: A review and critique of theoretical perspectives. *Journal of Management*, 16(2):461–509.

Jara-Díaz, S. R. and Basso, L. J. (2003). Transport cost functions, network expansion and economies of scope. *Transportation Research Part E: Logistics and Transportation Review*, 39(4):271–288.

Jin, Y. H., Swanson, D., Waller, M., and Ozment, J. (2017). To survive and thrive under hypercompetition: An exploratory analysis of the influence of strategic purity on truckload motor-carrier financial performance. *Transportation Journal*, 56(1):1–34.

Johnston, A. and Ozment, J. (2013). Economies of scale in the U.S. airline industry. *Transportation Research Part E: Logistics and Transportation Review*, 51:95–108.

Keeler, T. E. (1989). Deregulation and scale economies in the U.S. trucking industry: An econometric extension of the survivor principle. *The Journal of Law and Economics*, 32(2):229–253.

Kleinsorge, I. K., Schary, P. B., and Tanner, R. D. (1991). The shipper-carrier partnership: A new tool for performance. *Journal of Business Logistics*, 12(2):35.

Kor, Y. Y. and Mahoney, J. T. (2000). Penrose's resource-based approach: The process and product of research creativity. *Journal of Management Studies*, 37(1).

Kor, Y. Y., Mahoney, J. T., Siemsen, E., and Tan, D. (2016). Penrose's 'The Theory of the Growth of the Firm': An exemplar of engaged scholarship. *Production and Operations Management*, 25(10):1727–1744.

Liu, J. S., Lu, L. Y., and Lu, W.-M. (2016). Research fronts in data envelopment analysis. *Omega*, 58:33–45.

McDonald, J. (2009). Using least squares and tobit in second stage DEA efficiency analyses. *European Journal of Operational Research*, 197(2):792–798.

McMullen, B. S. and Okuyama, K. (2000). Productivity changes in the U.S. motor carrier industry following deregulation: a Malmquist index approach. *International Journal of Transport Economics*, 27(3):335–354.

McMullen, B. S. and Tanaka, H. (1995). An econometric analysis of differences between motor carriers: Implications for market structure. *Quarterly Journal of Business and Economics*, 34(4):16–29.

Mejza, M. M. and Corsi, T. M. (1999). Assessing motor carrier potential for improving safety processes. *Transportation Journal*, 38(4):36–50.

Miller, J. W. and Saldanha, J. P. (2016). A new look at the longitudinal relationship between motor carrier financial performance and safety. *Journal of Business Logistics*, 37(3):284–306.

Miller, J. W., Stromeyer, W. R., and Schwieterman, M. A. (2013). Extensions of the Johnson-Neyman technique to linear models with curvilinear effects: Derivations and analytical tools. *Multivariate Behavioral Research*, 48(2):267–300.

Min, H. and Ahn, Y.-H. (2017). Dynamic benchmarking of mass transit systems in the United States using data envelopment analysis and the Malmquist productivity index. *Journal of Business Logistics*, 38(1):55–73.

Min, H. and Lambert, T. (2002). Truck driver shortage revisited. *Transportation journal*, 42(2):5–16.

Nebesky, W., McMullen, B. S., and Lee, M.-K. (1995). Testing for market power in the U.S. motor carrier industry. *Review of Industrial Organization*, 10(5):559–576.

Palich, L. E., Cardinal, L. B., and Miller, C. C. (2000). Curvilinearity in the diversification–performance linkage: An examination of over three decades of research. *Strategic Management Journal*, 21(2):155–174.

Penrose, E. T. (1959). *The theory of the growth of the firm.* John Wiley, New York, NY.

Peteraf, M. A. (1993). The cornerstones of competitive advantage: A resource-based view. *Strategic management journal*, 14(3):179–191.

Pettus, M. L. (2001). The resource-based view as a developmental growth process: Evidence from the deregulated trucking industry. *Academy of Management Journal*, 44(4):878–896.

Rakowski, J. P. (1988). Marketing economies and the results of trucking deregulation in the less-than-truckload sector. *Transportation Journal*, 27(3):11–22.

Rogers, M. M. and Weber, W. L. (2011). Evaluating CO2 emissions and fatalities tradeoffs in truck transport. *International Journal of Physical Distribution & Logistics Management*, 41(8):750–767.

Scheraga, C. A. (2004). The relationship between operational efficiency and customer service: A global study of thirty-eight large international airlines. *Transportation journal*, 43(3):48–58.

Scheraga, C. A. (2011). Strategic fit in the general freight motor carrier industry. *Transportation Research Part E: Logistics and Transportation Review*, 47(4):490–506.

Scheraga, C. A., Haslem, J. A., and Corsi, T. M. (1994). The financial mobility of motor carriers: Pre-and post-deregulation. *Transportation Research Part A: Policy and Practice*, 28(5):401–414.

Silverman, B. S., Nickerson, J. A., and Freeman, J. (1997). Profitability, transactional alignment, and organizational mortality in the U.S. trucking industry. *Strategic Management Journal*, 18(Summer):31–52.

Singer, J. D. and Willett, J. B. (2003). *Applied longitudinal data analysis: Modeling change and event occurrence*. Oxford University Press, New York, NY.

Sirmon, D. G., Hitt, M. A., and Ireland, R. D. (2007). Managing firm resources in dynamic environments to create value: Looking inside the black box. *Academy of Management Review*, 32(1):273–292.

Smith, J. (2016). Trucking industry hits road bump with rising diesel prices. *Wall Street Journal*. Retrieved from https://www.wsj.com.

Smith, R. D., Corsi, T. M., and Grimm, C. M. (1990). Motor carrier strategies and performance. *Transportation Research Part A: General*, 24(3):201–210.

Stock, J. R. (1988). The maturing of transportation: An expanded role for freight carriers. *Journal of Business Logistics*, 9(2):15–31.

Suzuki, Y. (2007). Truck driver turnover: What rate is good enough? *International Journal of Physical Distribution & Logistics Management*, 37(8):612–630.

Talluri, S. S., Kull, T. J., Yildiz, H., and Yoon, J. (2013). Assessing the efficiency of risk mitigation strategies in supply chains. *Journal of Business Logistics*, 34(4):253–269.

Teece, D. J. (1982). Towards an economic theory of the multiproduct firm. *Journal of Economic Behavior & Organization*, 3(1):39–63.

Thompson, J. D. (1967). *Organizations in action: Social science bases of administrative theory*. Transaction Publishers, New Brunswick, NJ.

Tyler, E. (2017). Knight and Swift get speedy antitrust clearance to merge. *Bloomberg: The Bureau of National Affairs*. Retrieved from https://www.bna.com/knight-swift-speedy-n57982087453/ [Accessed 5-May-2017].

Wan, W. P., Hoskisson, R. E., Short, J. C., and Yiu, D. W. (2011). Resource-based theory and corporate diversification: Accomplishments and opportunities. *Journal of Management*, 37(5):1335–1368.

Wanke, P. F. (2014). Efficiency drivers in the Brazilian trucking industry: A longitudinal study from 2002-2010. *International Journal of Physical Distribution & Logistics Management*, 44(7):540–558.

Weber, M. M. and Weber, W. L. (2004). Productivity and efficiency in the trucking industry: Accounting for traffic fatalities. *International Journal of Physical Distribution & Logistics Management*, 34(1):39–61.

Wernerfelt, B. and Montgomery, C. A. (1988). Tobin's q and the importance of focus in firm performance. *The American Economic Review*, 78(1):246–250.

Winebrake, J. J., Green, E. H., Comer, B., Li, C., Froman, S., and Shelby, M. (2015). Fuel price elasticities in the U.S. combination trucking sector. *Transportation Research Part D: Transport and Environment*, 38:166–177.

Xu, K., Windle, R., Grimm, C., and Corsi, T. (1994). Re-evaluating returns to scale in transport. *Journal of Transport Economics and Policy*, 28(3):275–286.

Ying, J. S. and Keeler, T. E. (1991). Pricing in a deregulated environment: The motor carrier experience. *The Rand Journal of Economics*, 22(2):264–273.

ESSAY 3

CONSISTENTLY SAFE: HOW ISO 9000 CERTIFICATION PROMOTES CONSISTENCY IN MOTOR CARRIER SAFETY

3.1 Introduction

The safe operation of large trucks on public roads is a matter of importance to motor carriers, to shippers and to the welfare of the general public (Cantor et al., 2017; Miller, 2017). Vehicle accidents involving large trucks are inherently unpredictable and account for a disproportionate share of total accidents on U.S. roads (Monaco and Redmon, 2012). These accidents are also costly. The economic cost of fatal large truck accidents in the U.S. exceeded 15 billion dollars in 2014—nearly 4.5 million dollars per fatal accident (Cantor et al., 2017). Because of the financial and safety-related implications of these accidents, motor carrier safety is routinely viewed as an issue of public policy (Cantor et al., 2009).

In the U.S., the Federal Motor Carrier Safety Administration (FMCSA) is charged with reducing the number of crashes involving large, commercial vehicles. The FMCSA collects and compiles data pertaining to crash investigations, moving violations and carrier performance during roadside inspections as part of the agency's safety compliance and enforcement program, titled *Compliance, Safety, Accountability* (CSA). These carrier safety data are used to "identify unsafe motor carriers and prioritize FMCSA enforcement resources on those that pose the greatest safety risk" (FMCSA, 2017, p. 9). A central assumption of the CSA program and, more broadly, public policy regarding motor carrier safety, is that a carrier's safety performance is influenced by its behavior and its management practices (Mejza et al., 2003; Miller and Saldanha, 2016). This assumption has been widely supported in the literature: some carriers tend to be safer than others and a carrier's level of safety performance tends to persist over time (Miller et al., 2017b).

Accordingly, a primary focus of the motor carrier safety literature has been to identify systematic explanations for the presence of these stable, between-carrier differences in safety performance. Identifying factors that contribute to motor carrier safety is important. Regulatory agencies such as the FMCSA can utilize these factors when formulating new policy and when evaluating existing policy, such as those pertaining to driver logs (Cantor et al., 2009) or the effectiveness of the FMCSA's new entrant program (Cantor et al., 2017). Regulatory agencies can also use these factors to identify high-risk carriers and to prioritize enforcement resources. For example, motor carrier safety tends to decline as turnover increases in the driver workforce (Miller et al., 2017c) and when carriers experience greater financial strain (Miller and Saldanha, 2016). Identifying factors that contribute to safety also remains important to carriers—who risk human, equipment and financial loss—as well as shippers, brokers and numerous other stakeholders.

Our research examines the relationship between motor carrier adoption of an organizational routine for quality management, International Organization for Standardization (ISO) 9000, and consistency in a carrier's safety performance. The ISO 9000 series is a family of international quality standards and quality guidelines that specifies what organizations should do to produce products and services that consistently meet the expectations of the organization's customers as well as statutory and regulatory requirements (ISO, 2016). Prior research has linked motor carrier adoption of ISO 9000 to reductions in the number of vehicle accidents as well as improvements in FMCSA-collected safety performance measures (Naveh and Marcus, 2007). Our research focuses instead on variability that occurs in a carrier's safety performance over time. The quality literature refers to a state of low outcome variability as *consistency* and acknowledges that attainment of a high level of performance does not imply consistency in performance (Su and Linderman, 2016). Consistency has long been considered to be among the most critical of motor carrier attributes (Stephenson and Willett, 1969; Bardi et al., 1989; Morash and Clinton, 1997; Premeaux, 2002; Dobie, 2005; Voss et al., 2006; Robinson et al., 2013). Yet, to date, the concept of consistent performance remains unexplored within the motor carrier safety literature.

It is important to identify factors that promote consistency in the safety performance of motor carriers, just as it is important to identify factors that promote a high level of safety performance. First, the organizational competencies required to achieve a high level of performance differ from those required to achieve consistency (Su and Linderman, 2016). Second, the concept of consistency in safety performance carries unique implications for carriers and policy makers. Intuitively, a carrier's latent profile of safety performance might be influenced by the carrier's history regarding consistency in safety performance over time, since variability indicates predictability at any given measurement occasion. From this perspective, minimizing variability (maximizing consistency) in safety performance would be beneficial to carriers and stakeholders—if carriers differ systematically in the variability of their safety performance, uncovering the systemic causes of this variability can assist carriers in improving the predictability of safety outcomes, such as vehicle accidents and safety violations. These factors may also identify to regulators carrier-level traits or competencies that reduce the likelihood of extreme instances of safety-related events, such as carrier involvement in a large number of vehicle accidents within a given measurement occasion. A less intuitive implication of safety consistency, and one that we also advance here, is that some level of variability in a carrier's safety performance might be beneficial, since variability may catalyze change. To relate consistency and change we rely on the theory of routine dynamics (Feldman and Pentland, 2003). We envision a motor's carrier behavior as being a function of its organizational routines, as has been posited regarding the behavior of organizations in the general case (Knott, 2001). We then theorize the effects of ISO 9000 on consistency and change given ISO 9000's role as a higher-level organizational routine, or metaroutine (Pentland et al., 2012).

The remainder of this essay proceeds as follows. First, we review the literature on motor carrier safety and ISO 9000 quality assurance. Following these reviews, we present our theoretical framework and formulate research hypotheses relating ISO 9000 to consistency and change in motor carrier safety. We then provide our methodology and present the results of hypothesis testing. The essay closes with a discussion of our findings and implications for theory and managers.

3.2 Literature Review

3.2.1 Motor Carrier Safety

Table 3.1 summarizes recent motor carrier safety research. Existing work is organized around several research streams including carrier adoption of monitoring technology (*e.g.*, Cantor et al., 2006, 2008, 2009; Hickman and Hanowski, 2011); labor topics, such as the use of owner-operators, unionization and the effects of driver turnover on carrier safety (*e.g.*, Corsi et al., 2012; Monaco and Redmon, 2012; Miller et al., 2017c); truck driver behavior (*e.g.*, Braver et al., 1992; Häkkänen and Summala, 2001; Morrow and Crum, 2004; Cantor et al., 2010; Swartz et al., 2017); regulatory effectiveness and carrier behavior under governmental regulation (*e.g.*, Moses and Savage, 1992; Saltzman and Belzer, 2002; Chen, 2008; Cantor et al., 2017; Miller, 2017; Miller et al., 2017b); and, lastly, the relationship between carrier safety and firm-level characteristics such as firm size and financial performance (*e.g.*, Naveh and Marcus, 2007; Britto et al., 2010; Cantor et al., 2016; Miller and Saldanha, 2016). However, as seen in Table 3.1, carrier consistency in safety performance has not been incorporated into recent causal models of motor carrier safety. Indeed, a review of the motor carrier safety literature across multiple disciplines (see Miller, 2017, for a discussion) did not reveal any prior investigation into the topic of consistency in safety over time. This absence may not be surprising, however, given that data on motor carriers' longitudinal safety performance was far less granular until only recently, in late 2010, when the FMCSA launched its CSA program (Miller et al., 2017a). Since this time, research into carrier-level safety performance has shifted toward longitudinal frameworks.

When viewed longitudinally, total variability in the safety performance of the motor carrier population appears to dominated by time invariant, between-carrier differences. In Miller and Saldanha (2016), between-carrier differences accounted for more than fifty percent of the variation in each of three CSA Behavior Analysis and Safety Improvement Category (BASIC) safety measures, *vehicle maintenance* (VM), *hours of service compliance* (HOS) and *unsafe driving* (UD). However, carrier safety performance is not static; the nature of longitudinal change in safety performance has also been the subject of research. Hickman and Hanowski (2011) employ a difference-in-difference quasi-experimental design to examine the effects of in-cab video monitoring on safety-related driver behaviors, finding that adoption led to a significant decrease in at-risk behavior. Miller et al. (2017a) examine carrier improvement in safety after implementation of the CSA program and ensuing disclosure of carrier safety information. Miller et al. (2017b) estimate autoregressive relationships and

Table 3.1: Summary of the motor carrier safety literature, 2007–2017.

Carrier-level study	Key finding(s)	Motor Carrier Safety Performance Aspect: Performance Level	Performance Change	Performance Consistency
Naveh and Marcus (2007)	Accident rates are lower for ISO 9000 certified motor carriers than for non-ISO certified carriers	R/E	R	.
Cantor et al. (2008)	Motor carrier adoption of safety technologies is highest among carriers with poor records of safety performance	E	.	.
Cantor et al. (2009)	Electronic logbook adoption is associated with fewer hours of service violations and fewer crashes	R/E	.	.
Britto et al. (2010)	Carrier financial performance is positively associated with carrier safety	R	.	.
Hickman and Hanowski (2011)	Carrier implementation of in-cab video monitoring technology decreases at-risk driving behaviors	.	R	.
Corsi et al. (2012)	Motor carriers who employ unionized drivers tend to be safer than those utilize non-union drivers	R	.	.
Monaco and Redmon (2012)	Owner-operator use leads to improved carrier safety outcomes, but is unrelated to average crash severity	R	.	.
Cantor et al. (2016)	Firm size is positively related to carrier safety, with varying effect sizes across industry segments	R	.	.
Miller and Saldanha (2016)	Carrier safety is positively associated with both financial performance improvement and static financial performance	R	R	.
Cantor et al. (2017)	Motor carriers who successfully complete the FMCSA's new entrant program are safer than those carriers who fail	R	.	.
Miller (2017)	Likelihood of a carrier exiting warning flag status for HOS compliance is a function of elapsed time and carrier size	E	R	.
Miller et al. (2017a)	Safety improvement tapered over time in response to CSA information disclosure; tapering influenced by carrier traits	E	R	.
Miller et al. (2017b)	FMCSA carrier safety measures (BASICs) serve as stable, leading indicators of other safety measures	R/E	R	.
Miller et al. (2017c)	Carrier safety declines with increases in driver turnover; effect is stronger for carriers with low activity control	R	.	.
This study	*ISO certification attenuates carrier safety performance-variability relationship among large U.S. motor carriers*	E	R	R/E

Notes: R = modeled as a response variable; E = modeled as a explanatory variable; R/E = modeled as both response and explanatory longitudinal inter-relationships among CSA measures, identifying unique (but temporally stable) reflexive changes in the measures in response to prior-period changes in other CSA measures. Our research argues that temporally stable differences exist between carriers in the consistency of their safety performance, and that these differences can be explained by carrier-level factors (*i.e.*, traits, competencies). We examine the effects of one such factor, carrier certification to ISO 9000 quality assurance standards, since the principal purpose of these standards is to promote uniformity and consistency in product and service quality (Anderson et al., 1999; Benner and Veloso, 2008).

3.2.2 ISO 9000 Series Quality Assurance

The ISO 9000 series is a family of international quality standards and quality guidelines that specifies what organizations should do to produce products and services that consistently meet the expectations of the organization's customers as well as statutory and regulatory requirements (ISO, 2016). Early versions of the series (prior to 2000) were applicable only to the manufacturing sector, however the scope of the ISO 9000 series was later expanded to address quality assurance for service providers (Martínez-Costa et al., 2009). Since its inception, companies across a wide variety of industries have sought to obtain certification of compliance to ISO standards (Gray et al., 2015). To become ISO 9000 certified, organizations must map their operating processes associated with quality production, implement replicable sets of routines and procedures to assure a consistent level of quality and demonstrate conformance to documented processes (Naveh and Marcus, 2005; Benner and Veloso, 2008). In addition to promoting an improved organizational focus on quality, ISO certification serves as a public signal of effective quality management (Anderson et al., 1999) and may provide competitive benefits to firms, such as improved financial performance (Corbett et al., 2005; Levine and Toffel, 2010; Lo et al., 2013). It is not uncommon for industrial purchasers to require suppliers to become ISO 9000 certified, since a supplier's ISO 9000 compliance should reduce the buyer's need to monitor supplier behavior (Terlaak and King, 2006) while promoting uniformity in product or service quality across multiple, ISO 9000 certified suppliers (Su et al., 2015).

ISO 9000 series certification has been linked to improvements in the management of workplace safety as well as improved safety outcomes for employees. Employees of ISO 9000 certified organizations report higher levels of employee involvement in safety management, enhanced effectiveness of safety communication and feedback within their organizations and increased comprehensiveness of safety-related training (Vinodkumar and Bhasi, 2011). Levine and Toffel (2010) find that ISO certified employers experience lower on-the-job death rates than non-ISO certified employers and are more likely to have zero filings for injury-related workers' compensation than their non-certified counterparts. The existence of ties between ISO 9000 and workplace safety are not entirely surprising. Veltri et al. (2013, p. 121) note that ISO 9000 series standards "point to the same basic set of best practices" as other leading sets of formal standards used for safety management. Further, ISO 9000 certified organizations tend to be more compliant with documented organizational processes (Gray et al., 2015) and ISO 9000 quality management tools such as those used for root cause analysis and continuous improvement can readily be applied to manage health and safety in the workplace (Levine and Toffel, 2010).

Initial evidence suggests a positive link between ISO 9000 certification and safety performance in the motor carrier industry. Naveh and Marcus (2007) contend that a motor carrier's adoption of ISO 9000 standards can lead to increased safety performance, primarily through improvements in three areas: carrier driving procedures, maintenance of vehicle and overall safety management. Using matched samples of certified and non-certified carriers, Naveh and Marcus (2007) associate ISO 9000 certification with improved safety performance. However, further evidence is needed to corroborate, update and extend their findings. First, the ISO 9000 series in use by their sample of carriers—the 1994 version of the ISO series—has since undergone three iterations of revisions (2000, 2008, 2015). Additionally, this 1994 ISO 9000 series was directed at quality assurance in manufacturing (Martínez-Costa et al., 2009) as the series was not expanded for applicability to service providers until the 2000 revision. More recent updates include changes related to safety. For instance, the current

ISO 9000 series (2015) takes a more systematic approach to risk management, focusing on preventative action through risk-based thinking. Next, the SafeStat indicators utilized in Naveh and Marcus's (2007) analysis are now obsoleted under the CSA program, having been replaced by more granular measures of carrier safety performance (Miller et al., 2017a). Lastly, as Miller and Saldanha (2016) note, Naveh and Marcus (2007) leave several lingering questions regarding longitudinal components of their data collection and analysis. One such question is if ISO 9000 series certification by motor carriers is associated with improved consistency in safety performance, as would be expected under a management system focused upon controlling quality. We present theory for this association next.

3.3 Theory and Hypothesis Development

An organization's behavior is defined by its routines (Nelson and Winter, 1982). Organizational routines are "repetitive, recognizable patterns of interdependent actions, carried out by multiple actors" (Feldman and Pentland, 2003, p. 95). Theories of organizational behavior commonly distinguish between two classes of routines: operational routines and metaroutines (Knott, 2001). Operational routines guide day-to-day organizational behavior. Metaroutines are higher-level routines—routines to govern and modify routines (Nelson and Winter, 1982). The ISO 9000 series is regarded as an organizational metaroutine (Pentland et al., 2012). Metaroutines such as ISO 9000 serve as integrating frameworks between operational routines and the environment, regulating changes in operational routines (Lewin et al., 2011), systematizing creative processes (Adler et al., 1999), identifying the need for new routines (Schroeder et al., 2008) and providing a conduit for managerial intervention (Knott, 2001). The theory of routine dynamics (Feldman and Pentland, 2003) posits that organizational routines can promote stability through *structure* and, simultaneously, promote change through *agency*. We rely on this theory to hypothesize unique effects of a motor carrier's adoption of ISO 9000 on consistency and continuous improvement in safety performance.

3.3.1 Structure: How ISO 9000 Promotes Consistency in Carrier Safety

Structure refers to the abstract idea of a routine, viewing a routine as a static pattern of action, a habit, a heuristic or a standard operating procedure (Feldman and Pentland, 2003). In structure, routines serve as a form of organizational memory that direct recurring activities, thereby promoting efficiency and reducing variation in organizational response to stimuli (Huber, 1991; Knott, 2001). When viewed in this frame, routines are considered slow to change and do so only when prompted by some exogenous event, such as adoption of new technology, experience of failure or attainment of a milestone (Feldman, 2000; Gersick and Hackman, 1990). This feature of stability, along with the efficiency and legitimacy that stable routines provide, is prized by organizations (Feldman and Pentland, 2003).

Viewing ISO 9000 from this structural perspective, a motor carrier's adoption of the ISO 9000 series metaroutine should lead to increased stability in operational routines and consistency in organizational outcomes. The concepts of stability and consistency are interrelated. In the safety and risk literature, stability is a condition signaled by the return of a system to some original state after perturbation (Wildavsky, 1988). This is to say, a stable system is mean-reverting (trend-stationary) in the presence of shocks. Consistency describes the pattern of variability (*i.e.*, of an outcome) when the system is under perturbation (Su and Linderman, 2016). Accordingly, achieving a high level of consistency requires stability, and systems that are more stable should also tend to produce outcomes that are more consistent. Routines promote consistency by enabling an efficient and predictable organizational response to stimuli (Nelson and Winter, 1982; Knott, 2001). Additionally, metaroutines promote consistency by regulating changes in existing operational routines and by turning non-routine tasks into routine tasks through the development of new operational routines (Lewin et al., 2011). Indeed, ISO contends that the 9000 series directly promotes consistency by requiring organizations to implement standardized and replicable sets of routines and procedures (Naveh and Marcus, 2005).

Motor carriers are regularly exposed to perturbations within the operational and regulatory environments in which they operate. Mello and Hunt (2009) describe the business environment for motor carriers as “highly competitive, uncertain and demanding” (p. 21). Carriers must contend with turbulence in highway conditions and congestion, in the demand market (Dobie, 2005) and within factor markets (Torry and Murray, 2015). A carrier’s implementation of routines which formalize organization rules, procedures and processes has been linked to improvements in carrier outcomes. Saldanha et al. (2013) link carrier use of formalized routines that detail rules, behavior and standards for drivers with operational and market performance. Miller et al. (2017c) find that carriers who routinize driver-related activities mitigate the effects of high driver turnover on unsafe driving behavior. ISO 9000 should promote the standardization and formalization of processes associated with the production of motor carriage by regulating changes in existing routines and inventing new routines. As the carrier’s production system becomes more stable to perturbations in the environment, safety outcomes should vary less, improving consistency.

H_1: *ISO 9000 certified motor carriers exhibit greater consistency in safety performance than non-certified carriers.*

3.3.2 Agency: How ISO 9000 Promotes Continuous Improvement in Carrier Safety

Agency recognizes that action in the performance of routines is situated and that human performance of a routine in practice may diverge from the performance of that routine in principal (Feldman and Pentland, 2003). For instance, organizational actors may need to adjust their performance of a routine to account for unique or changing contexts. There are also practical constraints on the degree of completeness with which organizational routines, even if formalized in writing, can routinize human action while anticipating and addressing all possible contingencies. Other variations in routine performance may simply be unintentional or may be a function of differences between actors in their understanding and subjective interpretations of an organizational routine. Routine dynamics considers organizational actors to be “knowledgeable and often reflective” (p. 506) in their performance of routines (Feldman et al., 2016). Accordingly, actors and their organizations learn from variations introduced in routine performance—purposeful or not—and choose to retain those variations that offer added utility to the organization, appeal to its actors or improve fit with other organizational routines (Feldman and Pentland, 2003). It is because of this sequence of variation, reflection and selection that routines become dynamic, capable of endogenous change. As routines change endogenously so may organizational outcomes, resulting in continuous change (Feldman, 2000).

Routine dynamics suggests that the ISO 9000 series metaroutine can be a source for continuous change in motor carrier safety. In production contexts, continuous (or constant) positive change is referred to as continuous improvement (Anand et al., 2009). Zangwill and Kantor (1998) view the process of organizational learning that underlies continuous improvement as a cyclical pattern of trial and error, where managers retrospectively assess outcomes from trials of variations in strategy. The literature on safety and risk has also advocated for a trial and error approach. For instance, Wildavsky (1988) likens the process of attaining increased safety to trial-and-error entrepreneurial activity, arguing that safety is a condition to be discovered and continuously improved. The ISO 9000 series directly encourages continuous improvement (Terziovski et al., 2003), thereby promoting the evolution of operational routines under its management. While the motor carrier literature has examined longitudinal change in safety (see Table 3.1), it has not documented the role of metaroutines in promoting continuous safety improvement. Routine dynamics suggests that this relationship should exist. As a dynamic routine, ISO 9000 should encourage change as carriers select and retain variations associated with performance of the ISO 9000 metaroutine and underlying routines.

H_2: *ISO 9000 certified motor carriers exhibit greater improvement in safety performance than non-certified carriers.*

3.4 Methodology

3.4.1 Data and Sample

We test our research hypotheses using a combination of archival and secondary data on large, for-hire motor carriers in the combination (articulated vehicle) segment. By focusing our analysis on a relatively homogeneous subsection of the carrier population we are able to reduce our exposure to uncontrolled variance on safety outcomes that might be attributed to factors such as latent differences in competitive pressures between for-hire and private fleets (Miller et al., 2017b) and differences in carrier operational characteristics between the combination and straight-truck (non-articulating) segments. Our development of a sample is discussed next. Monthly observations of carrier safety performance were obtained from the FMCSA's Safety Measurement System database on the following BASICs: *vehicle maintenance*, *hours-of-service compliance* and *unsafe driving* (for a review, see Cantor et al., 2017). Information on carrier characteristics were obtained from the FMCSA's Motor Carrier Census and from the results of an annual survey data collected on large U.S. motor carriers by *Inbound Logistics*, a leading practitioner-oriented publication for the logistics industry.

Motor carriers were identified by their presence in *Inbound Logistics*'s annual list of leading carriers ("Top 100 Truckers", 2013, 2014, 2015, 2016). These lists publish carrier-reported information regarding company demographics and operational characteristics, such as each carrier's operating area and the types of services provided. We utilize these lists for our research since they contain carrier-reported information on ISO certification status and since they are comprised of only established carriers in the industry, who have a history of safety performance. Using the compiled information, we identified for-hire carriers in the combination segment and linked these carriers to the Safety Measurement System database and carrier census data using a DOT number or, for parent firms, DOT numbers. Since some initial, transient effects from ISO adoption might be associated with the maturing of an organization's quality culture (Terziovski et al., 2003) or dwindle from a lack of long-term organizational commitment (Sroufe and Curkovic, 2008), we examined only carriers whose ISO certification status (certified or not certified) remained stable for all four years and, thus, were also present in each of the carrier lists.

Our data collection resulted in a sample of 78 carriers, of which 22 were ISO 9000 certified. Consistent with prior research into outcomes of ISO 9000 certification (*e.g.*, Corbett et al., 2005; Naveh and Marcus, 2007; Gray et al., 2015), we use a matching technique to create an "artificial control group" (Wangenheim and Bayón, 2007, p. 39). To do so, we paired certified carriers to non-certified carriers by propensity score. The use of propensity score matching is an appropriate method to remove bias in observational studies and quasi-experiments where confounds may be introduced through non-randomized assignment or self-selection on treatment conditions (Rosenbaum and Rubin, 1983). Propensity scores for ISO 9000 certification were obtained using logistic regression. A logistic regression of ISO certification status on a series of covariates (carrier age, size, segment, labor profile and initial safety performance) fit significantly better to the data than a null model ($\chi^2_\Delta(9) = 28.30, p < .01$) and correctly predicted certification status for 80.77 percent of carriers (50.00 percent of ISO-certified carriers). We used nearest neighbor caliper matching (Althauser and Rubin, 1970) to match with replacement (Dehejia and Wahba, 1999) ISO 9000 certified carriers with non-certified carriers by propensity score. Calipers were specified at a tolerance of .20 of the standard deviation of the propensity score, as recommended by Austin (2011). Of the 22 ISO-certified carriers in our original sample, 19 successfully paired with a similar non-certified carrier, resulting in a matched sample of 38 carriers.

3.4.2 Variable Operationalization

This subsection provides an overview of the operationalization of variables used in the study. Descriptive statistics for all variables are provided within Table 3.2.

3.4.2.1 Dependent Variables

Inconsistency. Consistency is measured as variability around a trend (Stephenson and Willett, 1969), and a greater degree of variability reflects greater inconsistency. We operationalize *inconsistency* as the magnitude of fluctuation in a carrier's safety performance, measured as the scaled standard deviation of first derivative estimates of a given BASIC with respect to time. Researchers commonly operationalize measures of variability within a unit (*e.g.*, consistency, volatility) using the standard deviation of raw scores (Hoffman, 2007; Deboeck et al., 2009; Matta et al., 2017). However, for a set of ordered observations

Table 3.2: Descriptive statistics.

Variable	Mean	S.D.	Pearson Product-Moment Correlation											
			(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
(1) VM Performance†	3.12	.93												
(2) HOS Performance†	.31	.33	.46*											
(3) UD Performance†	1.30	.65	.28	.22										
(4) VM Change	−.47	.37	−.57*	−.26	−.10									
(5) HOS Change	−.12	.20	−.11	−.90*	−.09	.04								
(6) UD Change	−.07	.35	−.04	.22	−.63*	−.07	−.31							
(7) VM Inconsistency	.04	.03	.15	.51*	−.07	.03	−.45*	.13						
(8) HOS Inconsistency	.01	.01	.28	.86*	.10	−.20	−.76*	.19	.81*					
(9) UD Inconsistency	.05	.07	.28	.82*	.18	−.21	−.70*	.06	.74*	.93*				
(10) ISO Certification	.50	.51	.23	−.02	.04	−.18	.09	−.09	−.04	−.04	−.05			
(11) Union Labor	.53	.49	.25	$<$.01	−.29	−.19	.22	.26	.25	.21	.14	.10		
(12) Inspections‡	7.56	1.18	−.01	−.34*	.20	−.12	.29	−.16	−.89*	−.65*	−.56*	.05	−.28	
(13) Carrier Age‡	3.53	.36	−.38*	−.48*	−.11	.13	.45*	−.13	−.56*	−.51*	−.53*	−.06	−.16	.46*

Notes: $n = 38$;* $p < .05$ (two-tailed); †initial (2012) safety performance; ‡log-transformed.

containing fluctuation in the presence of a linear time trend, the standard deviation of raw scores (*i.e.*, zeroth derivatives) is a positively-biased estimate of the true standard deviation of fluctuations. We adopt, instead, to measure the inconsistency as the average deviation in the rate of change, as proposed by Deboeck et al. (2009), using first differences of the raw scores. We then divide this measure by the square root of two since the variance of the ith difference of a set of independent, random observations and the variance of raw values of those observations differ by a scaling factor equal to the ith central binomial coefficient (a value of two for the first difference).

Figure 3.1 illustrates the potential for measurement bias when the standard deviation of raw scores is taken in the presence of linear trend. Data used within the figure came from observations of unsafe driving measures on two carriers within our original sample. Average unsafe driving performance differed little between the two carriers during this period, with a mean of 1.91 for Evans Delivery Company (leftmost plot) and a mean of 1.89 for C.R. England (rightmost plot). In comparing the two plots, longitudinal performance for Evans Delivery Company is characterized by a relatively small amount of upward trend but a large degree of variability. In contrast, longitudinal performance for C.R. England is characterized by a large amount of downward trend but only a small-to-moderate degree of variability. A researcher comparing the two carriers using the standard deviations of the raw values might mistakenly conclude that Evans Delivery Company was twice as consistent on unsafe driving as C.R. England was during this period. However, upon examining the scaled standard deviations of first derivative estimates, the exact opposite, and seemingly more correct conclusion would be made that C.R. England was the more consistent carrier.

Change. We operationalize *change* in safety performance using the observed linear change in a carrier's BASIC score throughout the four year period. This is measured as the sum of first differences of observations on the BASIC score. When examining BASIC scores, a greater level of safety is indicated by a lower BASIC score. Accordingly, positive change reflects a decline in carrier safety while negative change reflects a safety improvement.

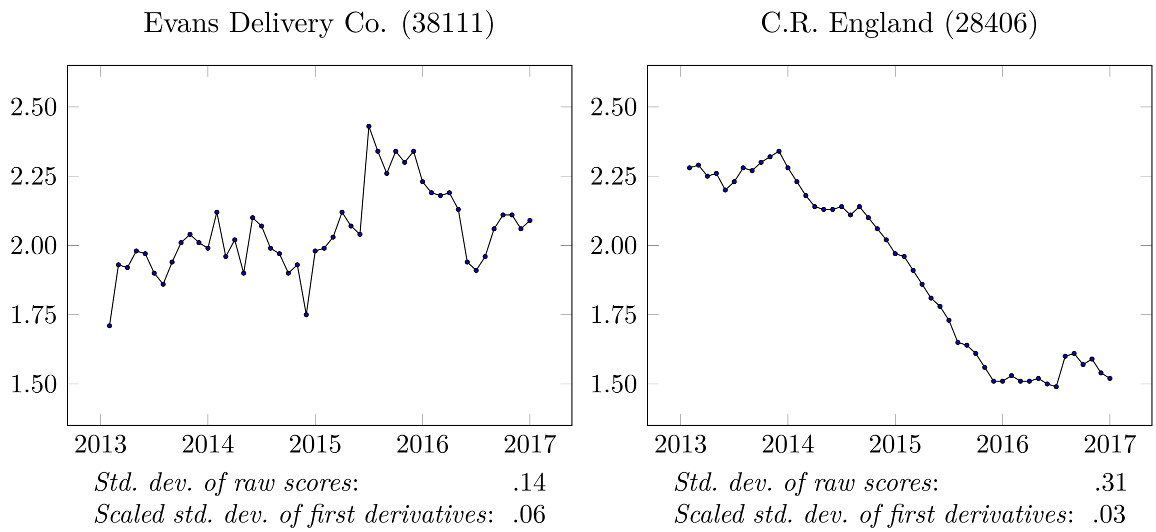


Figure 3.1: Comparison of alternative consistency measures using unsafe driving data for two carriers.

3.4.2.2 Focal Predictor

Carrier observations on our focal predictor, *ISO 9000 Certification*, were obtained from survey data published by *Inbound Logistics*. Half of the 38 motor carriers in our sample held ISO 9000 certification after matching. In the matched sample, ISO 9000 is statistically unrelated to the series of covariates utilized for matching ($\chi^2_\Delta(11) = 10.77, p = .46$). While this suggests a partial mitigation of threats due to confounds by covariates such as carrier size, age and initial level of safety performance, we retain the covariates used for matching to improve efficiency in estimation and to reduce confounds between contextual variables and safety outcomes.

3.4.2.3 Covariates

Initial Performance. Intuitively, a carrier's level of performance should relate to their level of consistency and likelihood of committing resources toward change. To avoid possible confounds, we control for carrier performance on a BASIC using the initial performance of that carrier on the BASIC. We reduce our potential exposure to measurement error by operationalizing initial performance using the carrier's annual performance in 2012, the year prior to the start of our sample. For any of the three BASICs used in this research (VM, HOS and UD), higher scores reflect poorer safety performance.

Labor Unionization. We include a dummy variable in our model to account for carrier use of unionized employees. The literature is mixed with regard to the effects of unionization on carrier safety outcomes. Miller et al. (2017c) do not find unionization to be a significant predictor of carrier safety performance on vehicle maintenance, hours-of-service compliance or unsafe driving. In contrast, Corsi et al. (2012) find that carriers utilizing unionized labor perform better on FMCSA safety metrics and have fewer vehicle accidents than their non-union counterparts. While the true nature of the relationship between carrier use of unionized labor and safety outcomes remains unclear, we include the covariate in our model to guard against potential confounding effects.

Carrier Size/Inspections. Cantor et al. (2016) theorize that large carriers may hold a safety advantage over smaller carriers due to superior access to human, technological and financial resources. Within the recent motor carrier safety literature, size is operationalized using one of two forms: as the number of power units (Corsi et al., 2012; Cantor et al., 2016, 2017) or as the natural logarithm of the number of power units (Britto et al., 2010; Miller and Saldanha, 2016; Miller, 2017; Miller et al., 2017a). A significant relationship between size and safety has been found in several studies (*e.g.*, Cantor et al., 2016, 2017), but results were mixed or inconclusive in others (*e.g.*, Britto et al., 2010; Miller et al., 2017b,c).

A carrier's size and the number of times they are inspected over a period of time are naturally correlated. We log-transform both *carrier size* and *number of inspections* in our sample (for skewness; Britto et al., 2010) and find a high correlation of .88 ($p < .01$) between the two. When examined annually, each of the two measures also exhibit very high stability, both having an intra-class correlation of .98. To avoid multicollinearity we select one of the two measures, *number of inspections*, for inclusion as a covariate in our model. Most importantly, data from the FMCSA on carrier size are based on carrier filings that do not occur on a set schedule (Miller et al., 2017a); carrier size updates do not correspond directly with FMCSA updates regarding carrier safety performance. In contrast, carrier inspection data is updated in the Safety Measurement System database monthly along with updates to carrier safety performance. While endogeneity concerns from measurement bias on *carrier size* are minimal (*i.e.*, given stability of carrier size over time), we select *number of inspections* out of an abundance of caution.

Carrier Age. We include carrier age in our model as a covariate since older carriers may have had more time to mature their organizational routines. An estimate of carrier age was obtained from the FMCSA's motor carrier census, using the carrier's date of entry into the Motor Carrier Management Information System database. Carrier age in years was log-transformed for skewness.

Carrier Segment. Consistent with Cantor et al. (2017) and others, we include dummy variables in the regression equation to control for potential differences due to carrier segment. We identify the carriers in our sample by their primary segment: *less-than-truckload* (16), *truckload* (11), *specialized/flatbed* (7) and *bulk* (4). We use *less-than-truckload* as the base category.

3.4.3 Model Specification

3.4.3.1 Contextual Factors

The literature on ISO 9000 wide recognizes that potential benefits from ISO 9000 adoption differ based on contextual factors. Lo et al. (2013) show that poor performing firms tend to benefit most from ISO 9000 adoption. Their findings are consistent with the ISO 9000 literature: firms with below-average quality management stand to benefit the most from implementing a quality management program. Accordingly, we allow for the effects of ISO 9000 certification on motor carrier safety outcomes to differ based on the carrier's initial level of safety performance. Given the findings of Lo et al. (2013), we would expect a negative coefficient to be estimated on the interaction term, such that higher initial BASIC scores (greater unsafety) amplifies a negative effect of ISO certification on inconsistency and change in safety performance.

3.4.3.2 Regression Equations

The model of safety consistency is defined within Equation 3.1 and the model of safety change is defined within Equation 3.2. In each of these these models, observations on set $M = \{\text{VM}, \text{HOS}, \text{UD}\}$ of correlated response variables, given by Y^M, are made on carrier i. In Equation 3.1, the response of *inconsistency* is given by $\dot{Y}$. In Equation 3.2 the response of *change* is given by $\ddot{Y}$. We control for *inconsistency* in Equation 3.2 since variability associated with trial-and-error should manifest in safety outcomes and, given Equation 3.1, may otherwise confound the relationship between ISO 9000 and *change* if *inconsistency* is also associated with *change*.

$$\begin{aligned}\dot{Y}_i^M = \beta_0^M &+ \sum_{k=1}^{3} \beta_k^M \times \text{SEGMENT}_i + \beta_4^M \times \text{UNION\_LABOR}_i \\ &+ \beta_5^M \times ln(\text{CARRIER\_AGE}_i) + \beta_6^M \times ln(\overline{\text{INSPECTIONS}}_i) \\ &+ \beta_7^M \times \text{ISO}_i + \beta_8^M \times \text{INITIAL\_PERFORMANCE}_i^M \\ &+ \beta_9^M \times \text{ISO}_i \times \text{INITIAL\_PERFORMANCE}_i^M\end{aligned} \tag{3.1}$$

$$\begin{aligned}\ddot{Y}_i^M = \beta_0^M &+ \sum_{k=1}^{3} \beta_k^M \times \text{SEGMENT}_i + \beta_4^M \times \text{UNION\_LABOR}_i \\ &+ \beta_5^M \times ln(\text{CARRIER\_AGE}_i) + \beta_6^M \times ln(\overline{\text{INSPECTIONS}}_i) \\ &+ \beta_7^M \times \text{ISO}_i + \beta_8^M \times \text{INITIAL\_PERFORMANCE}_i^M \\ &+ \beta_9^M \times \text{ISO}_i \times \text{INITIAL\_PERFORMANCE}_i^M \\ &+ \beta_{10}^M \times \text{INCONSISTENCY}_i^M\end{aligned} \tag{3.2}$$

3.5 Results

Estimation of Equations 3.1 and 3.2 was performed using seemingly unrelated regression (Zellner, 1962). We selected seemingly unrelated regression because the technique can improve estimation efficiency (beyond that of ordinary least squares) by accounting for the cross-equation residual covariance structure. We test for the appropriateness of this technique in our setting using Breusch and Pagan's (1980) Lagrange multiplier test. We are able to reject a null hypothesis of diagonality of the residual correlation matrix ($\chi^2(15) = 55.11, p < .01$), and conclude that the equations are more efficiently estimated as a system. The results of estimation are listed in Table 3.3. Overall fit of the system of equations was measured using the goodness-of-fit measure proposed by McElroy (1977), which suggests satisfactory fit ($R_z^2 = .93$). The effects of ISO certification were jointly significant in the system ($\chi_\Delta^2(12) = 58.64, p < .01$). Traditional R^2 values are presented within Table 3.3 for each of the six equations. For *inconsistency*, these values ranged from .55 to .94 and, for *change*, they ranged from .44 to .92.

3.5.1 Hypothesis Testing

Hypothesis 1 suggests that ISO 9000 certified motor carriers will exhibit greater consistency in safety performance than non-certified carriers. We previously discussed that our review of the ISO 9000 literature led us to include a contextual factor, *initial performance*, as a moderator, since poor performing organizations appear to benefit most from ISO 9000 certification. This interaction term is significant in the VM regression ($\beta = -.024, t = -4.792, p < .001$) and in the HOS regression ($\beta = -.007, t = -3.070, p < .005$), but not in the UD regression ($\beta = .031, t = 1.379, p = .179$). For carriers with a poor initial VM performance (one standard deviation above the mean on the VM BASIC in 2012), ISO certification is associated with a decrease in VM inconsistency ($\beta = -.031, t = -4.344, p < .001$). However, for carriers with strong initial VM performance (one standard deviation below the mean on the VM BASIC in 2012), ISO certification is associated with an increase in VM inconsistency ($\beta = .014, t = 2.382, p = .024$). Similarly, for carriers with strong initial HOS performance, ISO certification is associated with a decrease in HOS inconsistency ($\beta = -.003, -2.640, p = .013$). However, an effect of ISO certification on HOS inconsistency is not present for carriers with strong initial HOS performance ($\beta = .002, t = 1.477, p = .151$). ISO certification was not significantly associated with UD inconsistency; after removing the

Table 3.3: Regression results for inconsistency and change in motor carrier safety performance.

	Inconsistency						Change					
	Vehicle Maintenance		Hours of Service		Unsafe Driving		Vehicle Maintenance		Hours of Service		Unsafe Driving	
	Est.	S.E.	Est.	S.E.	Est.	S.E.	Est.	S.E.	Est.	S.E.	Est.	S.E.
Intercept	.194**	.028	.026**	.006	.534**	.109	.739	1.170	−.420*	.182	.907	.621
SEGMENT:LTL	−.011†	.006	−.001	.001	−.026	.027	−.122	.155	−.002	.031	−.336*	.122
SEGMENT:BULK	−.019*	.009	−.002	.002	−.078†	.039	−.290	.227	.035	.049	−.358†	.178
SEGMENT:SPEC/FB	−.017*	.007	−.004*	.001	−.040	.031	.140	.191	.082†	.041	.159	.137
UNION_LABOR	−.004	.006	.001	.001	−.016	.025	.035	.146	.099*	.029	.202†	.107
ln(CARRIER_AGE)	−.009	.007	.001	.001	−.049	.030	.014	.185	.058	.038	−.107	.132
ln($\overline{\text{INSPECTIONS}}$)	−.024**	.003	−.003**	.001	−.036**	.011	−.064	.111	.024	.018	.006	.052
ISO	.068**	.016	.002	.001	−.053	.035	−.025	.563	−.077*	.031	−.012	.186
INITIAL_PERFORMANCE	.026**	.005	.024**	.001	.003	.018	−.242	.179	−.865**	.109	−.485*	.101
ISO × INITIAL_PERFORM.	−.024**	.005	−.007**	.002	.032	.023	.008	.190	.371**	.080	−.009	.130
INCONSISTENCY							.409	3.808	9.625*	4.059	.789	.753
R^2	.860		.941		.546		.442		.921		.664	

Notes: $^{\dagger}p < .10$, $^{*}p < .05$, $^{**}p < .01$ (two-tailed).

interaction term from the regression equation, estimated coefficients remained non-significant for ISO certification ($\beta = -.012, t = -.650, p = .521$) and for initial UD performance ($\beta = .020, t = 1.556, p = .131$).

Hypothesis 2 suggests that ISO 9000 certified motor carriers will exhibit greater improvement in safety performance than non-certified carriers. As before, *initial performance* was included as a moderating contextual factor. First, we examined interaction terms between ISO certification and initial performance to determine if moderation was present. These terms are non-significant in the VM regression ($\beta = .008, t = .041, p = .968$) and in the UD regression ($\beta = -.009, t = -.072, p = .942$). After removal of the interaction term the effect of ISO certification remains non-significant in the VM regression ($\beta = .02, t = .784, p = .440$) and in the UD regression ($\beta = -.026, t = -.333, p = .742$). In the HOS regression, the interaction term is significant but the sign is opposite of our a priori expectations ($\beta = .37, t = 4.646, p < .01$). Next, we examined the relationship between *inconsistency* and *change*. While secondary to our hypothesized direct effect, *inconsistency* provides a possible mediated pathway from ISO 9000 certification to change. We found non-significant effects of *inconsistency* in the VM regression ($\beta = .409, t = .107, p = .915$) and in the UD regression ($\beta = .789, t = 1.048, p = .304$). Again, the coefficient in the HOS regression is significant but the sign is opposite of our a priori expectations ($\beta = 9.625, t = 2.372, p = .025$). We employ an auxiliary regression test of endogeneity (Davidson and MacKinnon, 1993) to test for possible bias on the estimate of *inconsistency*, since bias may be present due to simultaneity. The coefficients associated with the residual terms from the auxiliary regressions are non-significant within the VM regression ($\beta = -2.963, t = -.365, p = .718$), HOS regression ($\beta = 8.309, t = 1.01, p = .321$) and UD regression ($\beta = 0.050, t = .030, p = .976$). Accordingly, we are unable to conclude that bias exists on estimates of *inconsistency* due to endogeneity.

Table 3.4: Bootstrapped intra-class correlations of annualized safety performance measures.

	Performance Level		Performance Consistency		Performance Change	
BASIC	ICC	95% C.I.	ICC	95% C.I.	ICC	95% C.I.
Vehicle Maintenance (VM)	.91	[.77, .96]	.80	[.64, .89]	.01	[.00, .05]
Hours of Service (HOS)	.88	[.73, .95]	.73	[.58, .84]	.18	[.06, .28]
Unsafe Driving (UD)	.89	[.78, .95]	.58	[.35, .81]	.01	[.00, .06]

Notes: $R = 2,500$.

3.5.2 Post-Hoc Analysis

The results of our analysis lead us to revisit several assumptions central to our theorizing. First, we made an explicit case that differences exist between carriers in the consistency of their safety performance and that these differences persist over time. Similarly, our theorizing of safety change assumed that carriers differ in the amount that they change their safety performance that these differences also persist over time (*i.e.*, for some carriers, safety performance undergoes continuous improvement). Lastly, consistent with prior findings in the literature, we assumed that persistent between-carrier differences exist in the level of safety performance within the carrier population. We test our assumptions regarding persistence of between-subject differences by developing bootstrap confidence intervals for intra-class correlations (ICCs). We follow the procedure given by Ukoumunne et al. (2003). For each ICC to be estimated, $2,500$ bootstrap data sets were created by random sampling at the carrier level, with replacement. We estimate ICCs for each dataset within the mixed effect modeling framework and use the resulting empirical distributions of bootstrapped estimates to construct percentile confidence intervals. The results of this procedure are presented in Table 3.4. Higher values of ICC imply that more of the total variation in the outcome of interest resides between carriers. The data supports our assumption of persistent differences between carriers in *level* and *consistency*, but does not widely support this assumption regarding *change*.

3.6 Discussion

Ours is the first study that we are aware of to examine longitudinal consistency in a motor carrier's safety performance. We do so by conceptualizing motor carrier behavior as a function of organizational routines and rely on a theory of dynamic routines (Feldman and Pentland, 2003) to hypothesize inter-relationships between a metaroutine (ISO 9000), consistency in safety performance and change in safety performance. We tested our hypotheses using a matched sample of certified and non-certified U.S. carriers with carrier-level safety data from the FMCSA on the VM, HOS and UD BASICs. The results of these tests are discussed next.

Our first hypothesis linked ISO 9000 certification to consistency in motor carrier safety performance, suggesting that ISO certification would promote greater consistency. ISO 9000 series quality assurance standards specify basic requirements for organizations to "consistently provide products and services that enhance customer satisfaction and meet applicable statutory and regulatory requirements" (ISO, 2016, p. 2). Thus, improved consistency in organizational outcomes should be a principal benefit from adoption of ISO 9000 series standards. We tested this proposition directly, examining differences in consistency of motor carrier safety performance between a sample of carriers who are ISO 9000 certified and a matched sample of carriers who are not. We found support for the hypothesis for the VM and HOS BASICs, but not for UD. In our sample, ISO 9000 certification was a time-invariant carrier trait and, accordingly, would best explain between-carrier differences in BASICs. When the data were disaggregated in our post-hoc analysis to a lower level (annual safety performance), we found UD consistency to differ more within carriers than did VM consistency or HOS consistency. Thus, UD consistency appears to be a less stable phenomenon over time than its VM and HOS counterparts, and would be less predictable by a stable carrier trait such as ISO 9000 certification.

Arguably, carriers may have more control over the maintenance of vehicles and enforcing compliance with hours-of-service regulations (*i.e.*, via logbooks) than on-the-road behavior of drivers. This provides one potential explanation for the lower ICC found on UD and the lack of significant findings regarding its relationship with ISO 9000. In contrast, we found a link between ISO 9000 certification and VM consistency, however the strength of the relationship depended heavily on a carrier's initial VM performance. Unsafe carriers appeared to benefit the most, with ISO 9000 certification leading to an increase in VM consistency of nearly one standard deviation. However, for safe carriers, ISO 9000 certification was associated with a small but significant decline in VM consistency. Lastly, for HOS, we found that ISO 9000 certification of unsafe carriers was associated with a one standard deviation improvement in HOS consistency. However, no effect was present for safe carriers. Taken together, these findings suggest that unsafe carriers stand to benefit the most from ISO 9000 certification—at least when it comes to maintaining a consistent level of safety performance.

Our second hypothesis suggested that ISO certification would promote change in a carrier's safety performance. ISO 9000 certification requires organizations to enact procedures for the continuous improvement of organizational processes, and has been linked to improved occupational safety outcomes (Levine and Toffel, 2010). However, this hypothesis was not supported by the data. Upon examining the results of our post-hoc analysis, this lack of support is not surprising: change in motor carrier safety performance is dominated by within-subject variation for all three BASICs (Table 3.4). Thus, our sample offers very little support for the notion that carriers differ in the degree that they continuously improve their safety performance. On average, our sample of carriers improved across all three BASICs during our period of analysis (Table 3.2). However, for VM and UD, we are unable to conclude that *any* stable between-carriers differences exist in safety change over time. For HOS these between-carrier differences appear to be minimal. While incentives to improve may be greatest for unsafe carriers—our analysis suggests that carriers who were initially unsafe were the most likely to improve—carriers may experience a rapid decline in marginal returns from the allocation of additional resources toward improving safety. Our lack of evidence linking ISO 9000 to safety improvement is not unprecedented. Despite the enormity of the body of research into ISO 9000, findings have been inconclusive regarding its on-going benefits to level firm outcomes (Singh et al., 2011). We conclude from our assessment of the findings regarding our second hypothesis that changes in motor carrier safety appear to be transient adjustments or even movements "with the tide," but our data offer little evidence to suggest significant differences between carriers in their ability to continuously improve safety performance.

A primary contribution of our research is the introduction of the concept of performance consistency to the motor carrier safety literature. To date, this body of literature has largely focused on modeling levels of safety and changes in safety at the carrier and driver levels. The concept of consistency in safety performance has gone unstudied despite its sizeable implications on predictability of a carrier's safety behavior. We show that carriers differ in the consistency of their safety performance and that these differences persist over time. We also offer a factor that contributes to safety consistency, ISO 9000. Su and Linderman (2016) suggest that the organizational capabilities needed to achieve consistent performance differ from those needed to achieve a high level of performance. We find that ISO 9000 certification contributes to carrier consistency in the VM and HOS BASICs, conditional on a carrier's level of performance, but does not widely influence the presence of level changes in BASICs. Nor are we able to link ISO 9000 certification to consistency in UD, which might be influenced by other managerial initiatives such as carrier adoption of in-cab monitoring technologies. Our findings hold several implications for carrier management and policymakers. We discuss these implications next.

3.6.1 Managerial and Policy Implications

Our results imply that, for carrier management, ISO 9000 may be an effective mechanism to improve the consistency of safety performance. We found that these effects were strongest for unsafe carriers. Inconsistency is related to predictability. Inconsistent carriers may be more likely to experience extreme safety events such as a large number of accidents or violations at a given measurement occasion. This holds certain financial implications for carriers. First, improvements in predictability of a carrier's safety performance may reduce their level of assessed risk by insurers. Competitive and economic pressures have increasingly forced insurers to rely on risk assessments when establishing insurance rates and actuaries now commonly consider a motor carrier's safety record (among other factors) when assigning premiums (Hymel et al., 2012). Predictability in safety performance may also improve a carrier's ability to attract advantageous financing terms and reduce their own financial exposure. Factors attributed to motor carrier safety performance are important to investors, investment brokers and financial analysts who hold, trade or assess securities of public carriers. Not only might poor safety performance impact a carrier's reputation and future business, large carriers routinely risk earnings by self-insuring a substantial portion of their potential liability (*e.g.*, Werner Enterprises, Inc., 2016).

For shippers, selection of an unsafe carrier can have severe consequences including disrupting the efficient flow of goods through the supply chain. Because of this, shippers and brokers increasingly consider safety performance within their carrier selection decisions (Cantor et al., 2013). While we were unable to link ISO 9000 certification to longitudinal change in safety performance, our evidence suggests that ISO 9000 certification serves as a credible public signal of quality control in a carrier's safety performance. Prior research has shown customer demands to be a leading cause for organizations to pursue ISO 9000 certification. For shippers entering into high-volume, long-term contracts with carriers, demanding ISO 9000 certification may be viable mechanism to improve predictability and reduce variability in carrier behavior. This could be especially important to shippers who utilize just-in-time manufacturing and are especially sensitive to disruptions in the supply chain. Other shippers holding large inventories may benefit from a more predictable supply chain by reducing safety stock, since safety stock levels are tied to variability in the fulfillment of inventory reorders.

The most important implications of our research are for policy makers and regulators, such as the FMCSA. Motor carrier safety policy is largely based on the assumption that carriers differ in their level of safety and that these differences are attributable to carrier management practices. Our findings support this, but also show that motor carriers differ substantially in the consistency of their safety performance over time. We examine safety data from large carriers over four years and find these between-carrier differences in level of safety consistency are persistent. Most importantly, we show that carrier management practices can influence consistency. In our analysis, unsafe carriers exhibited significantly greater consistency in their VM and HOS safety performance when they were ISO 9000 certified. Other carrier management practices, such as the adoption of monitoring technologies, may also lead to consistency improvements for UD.

Unfortunately, current FMCSA policy under the CSA program is silent regarding consistency of carrier safety performance. Yet, the carriers posing the greatest threat to public safety may be those who exhibit a combination of low average safety performance but high variability in performance. Not only are these carriers generally unsafe, but their safety behavior at any time is unpredictable. The FMCSA has been criticized for its inability to consistently identify high-risk carriers under its current methodology (Government Accountability Office, 2014). Indeed, the CSA methodology has led the FMCSA to disproportionately target their intervention resources toward small carriers with few inspections and vehicles (Government Accountability Office, 2015). If low-consistency carriers are those that lack the sort of internal controls needed to self-regulate safety behavior then FMCSA methodologies for determining risk and interventions should account for carrier inconsistency in safety performance.

3.6.2 Limitations and Future Research

Our research is limited in that we examined long-term effects of ISO 9000 certification on motor carrier safety, but did not explore any effects that may occur prior to certification or short-term changes to carrier safety outcomes which may occur just after certification. Rather, we utilized a carrier's initial level of performance on FMCSA BASICs in developing a matched sample, thus avoiding certain confounds that might be attributable to carrier performance at the start of our analysis. While our assessment of long-term effects does not impact our conclusions regarding carrier consistency, it may impact certain conclusions regarding carrier changes in safety performance that are associated with ISO 9000. Future research might revisit the findings of Naveh and Marcus (2005) and determine when ISO 9000-related effects on level of carrier safety performance occur within the ISO 9000 adoption and certification process and if these effects are stable over time.

Further, our research examined directionality of carrier change in motor carrier safety performance and assumed that ISO 9000 certification and variability in safety outcomes should exclusively result in positive change (safety improvements). Routine dynamics theorizes that routines evolve when useful variations in routine performance are adopted, or selected for retention, by managers and other organizational agents (Feldman and Pentland, 2003). However, organizational routines may constitute "competing scripts" (Aroles and McLean, 2016, p. 542) when incentives and goals differ across levels of organizational hierarchy or between business units. Additionally, with regard to trial-and-error induced improvements in safety outcomes, it might be the case that some ideal level of variability (trials) exists. Instances of too little or too much variability might uniquely constrain a carrier's ability to enact beneficial change, due to information underload in the former case and information overload in the latter (see O'Reilly, 1980, for a discussion). In extreme instances, a high level of variability may lead to an excess of "noise" in the informational environment and cloud managerial judgment regarding the utility of any individual routine variation. Since agents select on routine variations retrospectively, extreme levels of variability may lead to the selection of "lemon" variations—those that inadvertently result in non-beneficial changes to routines and organizational outcomes. Given this, we believe that future research into non-linearities in the relationship between variability in routines and the nature of ensuing change to those routines may contribute significantly to the understanding of routine dynamics. Such research might also reveal the presence of a more nuanced relationship between variability and change in motor carrier safety performance than the linear relationship we theorized.

Lastly, future research should explore alternate causal mechanisms. For instance, our research advances a perspective of organizational routines and their underlying variability as a source of change. An alternative theory might suggest the presence of a simultaneous effect of change on variability and examine these relationships as a system of equations (*i.e.*, using three-stage least squares estimation). We tested for simultaneity in our analysis but were unable to conclude that it was present. However, if simultaneous, positive effects exist between intra-carrier variability and absolute intra-carrier change—a downward spiral—then it might be critical to uncover steps that carriers could take to identify and temper the variability-change pattern.

REFERENCES

REFERENCES

Adler, P. S., Goldoftas, B., and Levine, D. I. (1999). Flexibility versus efficiency? A case study of model changeovers in the Toyota Production System. *Organization Science*, 10(1):43–68.

Althauser, R. P. and Rubin, D. (1970). The computerized construction of a matched sample. *American Journal of Sociology*, 76(2):325–346.

Anand, G., Ward, P. T., Tatikonda, M. V., and Schilling, D. A. (2009). Dynamic capabilities through continuous improvement infrastructure. *Journal of Operations Management*, 27(6):444–461.

Anderson, S. W., Daly, J. D., and Johnson, M. F. (1999). Why firms seek ISO 9000 certification: Regulatory compliance or competitive advantage? *Production and Operations Management*, 8(1):28–43.

Aroles, J. and McLean, C. (2016). Rethinking stability and change in the study of organizational routines: Difference and repetition in a newspaper-printing factory. *Organization Science*, 27(3):535–550.

Austin, P. C. (2011). An introduction to propensity score methods for reducing the effects of confounding in observational studies. *Multivariate Behavioral Research*, 46(3):399–424.

Bardi, E. J., Bagchi, P. K., and Raghunathan, T. (1989). Motor carrier selection in a deregulated environment. *Transportation Journal*, 29(1):4–11.

Benner, M. J. and Veloso, F. M. (2008). ISO 9000 practices and financial performance: A technology coherence perspective. *Journal of Operations Management*, 26(5):611–629.

Braver, E. R., Preusser, C. W., Preusser, D. F., Baum, H. M., Beilock, R., and Ulmer, R. (1992). Long hours and fatigue: A survey of tractor-trailer drivers. *Journal of Public Health Policy*, 13(3):341–366.

Breusch, T. S. and Pagan, A. R. (1980). The Lagrange multiplier test and its applications to model specification in econometrics. *The Review of Economic Studies*, 47(1):239–253.

Britto, R. A., Corsi, T. M., and Grimm, C. M. (2010). The relationship between motor carrier financial performance and safety performance. *Transportation Journal*, 49(4):42–51.

Cantor, D. E., Celebi, H., Corsi, T. M., and Grimm, C. M. (2013). Do owner-operators pose a safety risk on the nation's highways? *Transportation Research Part E: Logistics and Transportation Review*, 59:34–47.

Cantor, D. E., Corsi, T. M., and Grimm, C. M. (2006). Safety technology adoption patterns in the U.S. motor carrier industry. *Transportation Journal*, 45(3):20–45.

Cantor, D. E., Corsi, T. M., and Grimm, C. M. (2008). Determinants of motor carrier safety technology adoption. *Transportation Research Part E: Logistics and Transportation Review*, 44(5):932–947.

Cantor, D. E., Corsi, T. M., and Grimm, C. M. (2009). Do electronic logbooks contribute to motor carrier safety performance? *Journal of Business Logistics*, 30(1):203–222.

Cantor, D. E., Corsi, T. M., and Grimm, C. M. (2017). The impact of new entrants and the new entrant program on motor carrier safety performance. *Transportation Research Part E: Logistics and Transportation Review*, 97:217–227.

Cantor, D. E., Corsi, T. M., Grimm, C. M., and Özpolat, K. (2010). A driver focused truck crash prediction model. *Transportation Research Part E: Logistics and Transportation Review*, 46(5):683–692.

Cantor, D. E., Corsi, T. M., Grimm, C. M., and Singh, P. (2016). Technology, firm size, and safety: Theory and empirical evidence from the U.S. motor-carrier industry. *Transportation Journal*, 55(2):149–167.

Chen, G. X. (2008). Impact of federal compliance reviews of trucking companies in reducing highway truck crashes. *Accident Analysis & Prevention*, 40(1):238–245.

Corbett, C. J., Montes-Sancho, M. J., and Kirsch, D. A. (2005). The financial impact of ISO 9000 certification in the United States: An empirical analysis. *Management Science*, 51(7):1046–1059.

Corsi, T. M., Grimm, C. M., Cantor, D. E., and Sienicki, D. (2012). Safety performance differences between unionized and non-union motor carriers. *Transportation Research Part E: Logistics and Transportation Review*, 48(4):807–816.

Davidson, R. and MacKinnon, J. G. (1993). *Estimation and inference in econometrics*. Oxford University Press, New York, NY.

Deboeck, P. R., Montpetit, M. A., Bergeman, C., and Boker, S. M. (2009). Using derivative estimates to describe intraindividual variability at multiple time scales. *Psychological Methods*, 14(4):367.

Dehejia, R. H. and Wahba, S. (1999). Causal effects in nonexperimental studies: Reevaluating the evaluation of training programs. *Journal of the American Statistical Association*, 94(448):1053–1062.

Dobie, K. (2005). The core shipper concept: A proactive strategy for motor freight carriers. *Transportation Journal*, 44(2):37–53.

Feldman, M. S. (2000). Organizational routines as a source of continuous change. *Organization Science*, 11(6):611–629.

Feldman, M. S. and Pentland, B. T. (2003). Reconceptualizing organizational routines as a source of flexibility and change. *Administrative Science Quarterly*, 48(1):94–118.

Feldman, M. S., Pentland, B. T., D'Adderio, L., and Lazaric, N. (2016). Beyond routines as things: Introduction to the special issue on Routine Dynamics. *Organization Science*, 27(3):505–513.

FMCSA (2017). *Safety Measurement System Methodology: Behavior Analysis and Safety Improvement Category Prioritization Status*. Technical report, Federal Motor Carrier Safety Administration (FMCSA).

Gersick, C. J. and Hackman, J. R. (1990). Habitual routines in task-performing groups. *Organizational Behavior and Human Decision Processes*, 47(1):65–97.

Government Accountability Office (2014). *Federal Motor Carrier Safety: Modifying the Compliance, Safety, Accountability Program Would Improve the Ability to Identify High Risk Carriers.* Technical Report GAO-14-114, United States Government Accountability Office, Washington, DC.

Government Accountability Office (2015). *Federal Motor Carrier Safety: Improvements to Data-Driven Oversight Could Better Target High Risk Carriers.* Technical Report GAO-15-443T, United States Government Accountability Office, Washington, DC.

Gray, J. V., Anand, G., and Roth, A. V. (2015). The influence of ISO 9000 certification on process compliance. *Production and Operations Management*, 24(3):369–382.

Häkkänen, H. and Summala, H. (2001). Fatal traffic accidents among trailer truck drivers and accident causes as viewed by other truck drivers. *Accident Analysis & Prevention*, 33(2):187–196.

Hickman, J. S. and Hanowski, R. J. (2011). Use of a video monitoring approach to reduce at-risk driving behaviors in commercial vehicle operations. *Transportation Research Part F: Traffic Psychology and Behaviour*, 14(3):189–198.

Hoffman, L. (2007). Multilevel models for examining individual differences in within-person variation and covariation over time. *Multivariate Behavioral Research*, 42(4):609–629.

Huber, G. P. (1991). Organizational learning: The contributing processes and the literatures. *Organization Science*, 2(1):88–115.

Hymel, K., Lee, D. B., Pearlman, J., Pritchard, R., and Rainville, L. (2012). *Financial Responsibility Requirements for Commercial Motor Vehicles.* Technical Report FMCSA-RRA-12-045, John A. Volpe Transportation Systems Center, U.S. Department of Transportation.

ISO (2016). *Selection and use of the ISO 9000 family of standards.* International Organization for Standardization, Geneva, Switzerland.

Knott, A. M. (2001). The dynamic value of hierarchy. *Management Science*, 47(3):430–448.

Levine, D. I. and Toffel, M. W. (2010). Quality management and job quality: How the ISO 9001 standard for quality management systems affects employees and employers. *Management Science*, 56(6):978–996.

Lewin, A. Y., Massini, S., and Peeters, C. (2011). Microfoundations of internal and external absorptive capacity routines. *Organization Science*, 22(1):81–98.

Lo, C. K., Wiengarten, F., Humphreys, P., Yeung, A. C., and Cheng, T. (2013). The impact of contextual factors on the efficacy of ISO 9000 adoption. *Journal of Operations Management*, 31(5):229–235.

Martínez-Costa, M., Choi, T. Y., Martínez, J. A., and Martínez-Lorente, A. R. (2009). ISO 9000/1994, ISO 9001/2000 and TQM: The performance debate revisited. *Journal of Operations Management*, 27(6):495–511.

Matta, F. K., Scott, B. A., Colquitt, J. A., Koopman, J., and Passantino, L. G. (2017). Is consistently unfair better than sporadically fair? An investigation of justice variability and stress. *Academy of Management Journal*, 60(2):743–770.

McElroy, M. B. (1977). Goodness of fit for seemingly unrelated regressions: Glahn's $r^2_{y,x}$ and Hooper's $\bar{r}^2$. *Journal of Econometrics*, 6(3):381–387.

Mejza, M. C., Barnard, R. E., Corsi, T. M., and Keane, T. (2003). Driver management practices of motor carriers with high compliance and safety performance. *Transportation Journal*, 42(4):16–29.

Mello, J. E. and Hunt, C. S. (2009). Developing a theoretical framework for research into driver control practices in the trucking industry. *Transportation Journal*, 48(4):20–39.

Miller, J. (2017). Discrete time hazard modeling of large motor carriers' longitudinal safety performance. *Transportation Journal*, 56(2):107–139.

Miller, J., Fugate, B., and Golicic, S. (2017a). How organizations respond to information disclosure: Testing alternative longitudinal performance trajectories. *Academy of Management Journal*, 60(3):1016–1042.

Miller, J. W., Golicic, S. L., and Fugate, B. S. (2017b). Developing and testing a dynamic theory of motor carrier safety. *Journal of Business Logistics*, 28(2):96–114.

Miller, J. W. and Saldanha, J. P. (2016). A new look at the longitudinal relationship between motor carrier financial performance and safety. *Journal of Business Logistics*, 37(3):284–306.

Miller, J. W., Saldanha, J. P., Rungtusanatham, M., and Knemeyer, M. (in press, 2017c). How does driver turnover affect motor carrier safety performance and what can managers do about it? *Journal of Business Logistics*.

Monaco, K. and Redmon, B. (2012). Does contracting with owner operators lead to worse safety outcomes for U.S. motor carriers? Evidence from the Motor Carrier Management Information System. *Accident Analysis & Prevention*, 45:654–659.

Morash, E. A. and Clinton, S. R. (1997). The role of transportation capabilities in international supply chain management. *Transportation Journal*, 36(3):5–17.

Morrow, P. C. and Crum, M. R. (2004). Antecedents of fatigue, close calls, and crashes among commercial motor-vehicle drivers. *Journal of Safety Research*, 35(1):59–69.

Moses, L. N. and Savage, I. (1992). The effectiveness of motor carrier safety audits. *Accident Analysis & Prevention*, 24(5):479–496.

Naveh, E. and Marcus, A. (2005). Achieving competitive advantage through implementing a replicable management standard: Installing and using ISO 9000. *Journal of Operations Management*, 24(1):1–26.

Naveh, E. and Marcus, A. (2007). Financial performance, ISO 9000 standard and safe driving practices effects on accident rate in the U.S. motor carrier industry. *Accident Analysis & Prevention*, 39(4):731–742.

Nelson, R. R. and Winter, S. G. (1982). *An evolutionary theory of economic change*. Harvard Business School Press, Cambridge, MA.

O'Reilly, C. A. (1980). Individuals and information overload in organizations: Is more necessarily better? *Academy of Management Journal*, 23(4):684–696.

Pentland, B. T., Feldman, M. S., Becker, M. C., and Liu, P. (2012). Dynamics of organizational routines: A generative model. *Journal of Management Studies*, 49(8):1484–1508.

Premeaux, S. R. (2002). Motor carrier selection criteria: perceptual differences between shippers and motor carriers. *Transportation Journal*, 42(2):28–38.

Robinson, J. L., Thomas, R. W., and Manrodt, K. B. (2013). Food for thought in the transportation carrier-selection decision. *Transportation Journal*, 52(2):277–296.

Rosenbaum, P. R. and Rubin, D. B. (1983). The central role of the propensity score in observational studies for causal effects. *Biometrika*, 70(1):41–55.

Saldanha, J. P., Shane Hunt, C., and Mello, J. E. (2013). Driver management that drives carrier performance. *Journal of Business Logistics*, 34(1):15–32.

Saltzman, G. M. and Belzer, M. H. (2002). The case for strengthened motor carrier hours of service regulations. *Transportation Journal*, 41(4):51–71.

Schroeder, R. G., Linderman, K., Liedtke, C., and Choo, A. S. (2008). Six Sigma: Definition and underlying theory. *Journal of Operations Management*, 26(4):536–554.

Singh, P. J., Power, D., and Chuong, S. C. (2011). A resource dependence theory perspective of ISO 9000 in managing organizational environment. *Journal of Operations Management*, 29(1):49–64.

Sroufe, R. and Curkovic, S. (2008). An examination of ISO 9000:2000 and supply chain quality assurance. *Journal of Operations Management*, 26(4):503–520.

Stephenson, P. R. and Willett, R. P. (1969). Consistency: The carrier's ace in the hole. *Transportation Journal*, 8(3):28–33.

Su, H.-C., Dhanorkar, S., and Linderman, K. (2015). A competitive advantage from the implementation timing of ISO management standards. *Journal of Operations Management*, 37:31–44.

Su, H.-C. and Linderman, K. (2016). An empirical investigation in sustaining high-quality performance. *Decision Sciences*, 47(5):787–819.

Swartz, S. M., Douglas, M. A., Roberts, M. D., and Overstreet, R. E. (2017). Leavin' on my mind: Influence of safety climate on truck drivers' job attitudes and intentions to leave. *Transportation Journal*, 56(2):184–209.

Terlaak, A. and King, A. A. (2006). The effect of certification with the ISO 9000 Quality Management Standard: A signaling approach. *Journal of Economic Behavior & Organization*, 60(4):579–602.

Terziovski, M., Power, D., and Sohal, A. S. (2003). The longitudinal effects of the ISO 9000 certification process on business performance. *European Journal of Operational Research*, 146(3):580–595.

"Top 100 Truckers" (2013). *Inbound Logistics*, 33(9):56–65.

"Top 100 Truckers" (2014). *Inbound Logistics*, 34(9):58–67.

"Top 100 Truckers" (2015). *Inbound Logistics*, 35(9):56–65.

"Top 100 Truckers" (2016). *Inbound Logistics*, 36(9):64–73.

Torry, IV, W. F. and Murray, D. (2015). *An Analysis of the Operational Costs of Trucking: 2015 Update*. Technical report, American Transportation Research Institute, Arlington, VA.

Ukoumunne, O. C., Davison, A. C., Gulliford, M. C., and Chinn, S. (2003). Non-parametric bootstrap confidence intervals for the intraclass correlation coefficient. *Statistics in Medicine*, 22(24):3805–3821.

Veltri, A., Pagell, M., Johnston, D., Tompa, E., Robson, L., Amick III, B. C., Hogg-Johnson, S., and Macdonald, S. (2013). Understanding safety in the context of business operations: An exploratory study using case studies. *Safety Science*, 55:119–134.

Vinodkumar, M. N. and Bhasi, M. (2011). A study on the impact of management system certification on safety management. *Safety Science*, 49(3):498–507.

Voss, M. D., Page Jr, T. J., Keller, S. B., and Ozment, J. (2006). Determining important carrier attributes: A fresh perspective using the theory of reasoned action. *Transportation Journal*, 45(3):7–19.

Wangenheim, F. v. and Bayón, T. (2007). Behavioral consequences of overbooking service capacity. *Journal of Marketing*, 71(4):36–47.

Werner Enterprises, Inc. (2016). *Form 10-K 2016*. Retrieved from http://investor.werner.com/financial-information/sec-filings/ [Accessed 17-July-2017].

Wildavsky, A. B. (1988). *Searching for safety*. Transaction Books, New Brunswick, CA.

Zangwill, W. I. and Kantor, P. B. (1998). Toward a theory of continuous improvement and the learning curve. *Management Science*, 44(7):910–920.

Zellner, A. (1962). An efficient method of estimating seemingly unrelated regressions and tests for aggregation bias. *Journal of the American Statistical Association*, 57(298):348–368.