Measuring the Value of a Responsive Supply Network

by

Jaime Garza Ramírez

B. S. Industrial and Systems Engineering (2004)
ITSEM (Monterrey Tech), Nuevo Leon, Mexico

and

Subramanian Mambakkam Suryanarayanan

P.G.D.M. (MBA, 1998)
Indian Institute of Management, Ahmedabad, India

B. Eng. Computer Science and Engineering (1995)
University of Madras, Chennai, India

MASSACHUSETTS INSTITUTE OF TECHNOLOGY
JUN 30 2009
LIBRARIES

Submitted to the Engineering Systems Division in Partial Fulfillment of the Requirements for the Degree of

Master of Engineering in Logistics

at the

Massachusetts Institute of Technology

ARCHIVES

June 2009

© 2009 Jaime Garza Ramírez and Subramanian M. Suryanarayanan. All rights reserved.

The authors hereby grant to MIT permission to reproduce and to distribute publicly paper and electronic copies of this thesis document in whole or in part in any medium now known or hereafter created.

Signature of Authors ....
Master of Engineering in Logistics Program, Engineering Systems Division, May 8, 2009

Certified by ....
Dr. Chris Caplice
Executive Director, Center for Transportation and Logistics
Thesis Supervisor

Accepted by ....
Prof. Yossi Sheffi
Professor, Engineering Systems Division
Professor, Civil and Environmental Engineering Department
Director, Center for Transportation and Logistics
Director, Engineering Systems Division

Measuring the Value of a Responsive Supply Network

by

Jaime Garza Ramírez and Subramanian M. Suryanarayanan

Submitted to the Engineering Systems Division on May 8, 2009 in Partial Fulfillment of the Requirements for the Degree of Master of Engineering in Logistics

ABSTRACT

LargeCo, a large multinational consumer goods manufacturing and distribution company, has been pursuing initiatives to improve the responsiveness of its supply network. The impact of these initiatives on the supply network is measured through a few Key Performance Indicators (KPIs). LargeCo has invested in a responsive supply chain so that it can respond swiftly to unpredictable market demand and minimize lost sales. Reduction in lost sales leads to growth in sales. LargeCo is interested in finding out if its responsive supply chain is contributing to sales growth. In particular, LargeCo would like to determine whether improvement in KPIs, driven by improvement in the responsiveness of the supply chain, has a relationship with improvement in sales. LargeCo uses a measure of sales known as Sales Net of Effects (SNE) which measures sales net of the effects of discounts, marketing and promotions. Establishing a relationship between KPIs and sales will help LargeCo measure the value of responsiveness in its supply network.

This research project develops an analytical framework using an econometric model to determine if relationships exist between the KPIs and sales and a causal model to explain the relationships. The econometric model shows that relationships exist between two of the KPIs – Days of Inventory and Supply Chain Cycle Time, and sales. The causal model explains how these KPIs and sales are linked.

Thesis Supervisor: Dr. Chris Caplice
Title: Executive Director, Center for Transportation and Logistics

Acknowledgements

We thank Dr. Chris Caplice, our thesis advisor, for his guidance and support through every step of the thesis process. His insights helped us make progress whenever we struggled to make headway and his encouragement kept us from reaching for MBLOG as our support mechanism.

We thank our sponsor company and key members from the company who provided us the resources and assistance required to tackle this research work.

Our sincere thanks go to Dr. William Haas for working with us to help us understand the difference between writing and writing lucidly.

Rajaram Narayanaswamy, our friend and fellow MLOGGER, was available whenever we needed his assistance as a sounding board to help us make progress with our research work. We acknowledge his patience and support.

And finally we thank the MLOG Class of 2009 for making our MLOG experience an enriching and memorable experience.

Table of Contents

List of Figures

List of Tables

List of Abbreviations

Abbreviation	Explanation
DI	Days of Inventory
FR	Fill Rate
KPI	Key Performance Indicator
LC	Logistics Cost
OLS	Ordinary Least Squares
SCCT	Supply Chain Cycle Time
SKU	Stock Keeping Unit
SNE	Sales Net of Effects; Sales net of the effects of discounts, marketing and promotions
TBPR	Time Between Production Runs

1 Introduction

Organizations in various industries such as manufacturing and retail, to name a few, rely on their supply chains to be successful. LargeCo is a large manufacturer and distributor of a wide variety of consumer products. The markets for these products are characterized by low growth rates, high product variety and demand uncertainty. A supply chain strategy founded on responsiveness helps firms compete successfully in such markets (de Groote, 1994; Fisher, 1997; Randall et al., 2003). A responsive supply chain built around short production lead times and small batch sizes allows firms to adapt quickly to markets that have fluctuating demand (Skinner, 1974). A responsive supply chain provides organizations the flexibility to respond not just to changes in market conditions caused by changing customer behavior but also to changes that are internal to organizations and changes that are caused by competitors and suppliers (Beamon, 1999).

LargeCo relies on its supply chain to be successful in the markets that it competes. It has focused on building a responsive supply chain that will allow it to respond rapidly and cost-effectively to changing market conditions. LargeCo has initiated and executed a number of initiatives to improve the responsiveness of its supply chain. Some of the areas these initiatives focus on include reduction of batch sizes, supplier lead times and production lead times.

Organizations that rely on their supply chains to be successful use various key performance indicators (KPI) to measure the performance of their supply chains. KPIs to measure the performance of a responsive supply chain are centered on measures to monitor time-effectiveness of the supply chain (Hausman, 2003; Chan & Qi, 2004). KPIs used by organizations to measure the performance of their supply chains should align not just with the strategic goals of the organizations but also with the strategic goals of their customers (Beamon, 1999). KPIs that track performance with respect to customer goals include fill rate and stockout rate. LargeCo uses various KPIs to track the performance of its supply chains. Some of these KPIs are Fill Rate (FR), Supply Chain Cycle Time (SCCT) and Time Between Production Runs (TBPR). In addition LargeCo also uses other KPIs focused on cost and inventory to measure the performance of its supply chain. Such KPIs include Days of Inventory (DI) and Logistics Cost (LC).

LargeCo believes that the initiatives it executed to improve the responsiveness of its supply chain have had the desired effects. LargeCo also believes that improvement in responsiveness is reflected in improvements in the various KPIs used by it to track the performance of its supply chain.

1.1 Motivation

Organizations build responsive supply networks to respond rapidly to changes in market conditions. The ability to respond rapidly to market changes helps organizations manage unpredictable demand while minimizing lost sales caused by unmet demand (Hoover et al., 2001). Reduction in lost sales is expected to be achieved without building excess inventory. Organizations that have invested in initiatives to improve the responsiveness of their supply chains have done so with the belief that improvement in responsiveness reduces lost sales and contributes to growth in sales. However no known model exists to verify this belief. LargeCo is therefore interested in determining whether improvements in KPIs, driven by improvement in responsiveness of its supply network, contribute to improvements in sales. LargeCo believes that establishing a relationship between KPIs and sales will help measure the value of responsiveness in its supply network. Therefore LargeCo's interest in measuring the value of responsiveness can be translated to the following research question:

Can improvements in KPIs be linked to improvements in sales and if so, how?

This research project develops a model to answer this research question. The model helps LargeCo establish a relationship between improvements in KPIs and improvements in sales.

1.2 Scope of Research

The research project focuses on improvements in the following five KPIs:

1. Days of Inventory (DI)
2. Fill Rate (FR)
3. Logistics Cost (LC)
4. Supply Chain Cycle Time (SCCT)
5. Time Between Production Runs (TBPR)

LargeCo uses a proxy measure for Sales known as Sales Net Effects (SNE). This proxy measure eliminates from Sales the effect of discounts, marketing and promotions thereby providing a measure of sales devoid of these effects. For the purpose of this research SNE is treated as Sales.

This research project develops an analytical framework to assess and explain relationships between the KPIs and sales. The framework includes an econometric model and a causal model. The econometric model uses correlation, regression, t-Test and F-test to examine if relationships exist between various KPIs and sales. The causal model uses causal diagrams and mathematical formulations to explain these relationships.

The initial scope of research for LargeCo is restricted to using this framework for LargeCo's sales in the US for five of its product lines for the period July of Year01 to February of Year04 and is based on the monthly data available from LargeCo for this period. The five product lines are:

1. ProductLine1
2. ProductLine2
3. ProductLine3
4. ProductLine4
5. ProductLine5

Each product line is made up of multiple Stock Keeping Units (SKUs). An SKU uniquely identifies an item in a product line. Aggregated demand across all SKUs in a single product line may not show significant uncertainty and variation; but demand at the level of an SKU within each product line may show uncertainty and variation due to a high level of product variety and low product differentiation. This behavior at the SKU level has led LargeCo to focus on a responsive supply chain to manage demand uncertainty.

1.3 Thesis Roadmap

The next chapter - Chapter 2, presents a review of previous research in the areas of responsiveness, supply chain performance measures and development of models to assess relationships between sets of variables. Chapter 3 discusses the methodology developed to address LargeCo's research question. The discussion on the methodology begins with an overview of the analytical framework developed to address the research question followed by details on the how the framework was applied to answer the research question in the particular context of LargeCo. Chapter 4 discusses the analysis performed for LargeCo using the framework and key findings from the analysis. Chapter 5 summarizes the findings, recommendations derived from the findings, limitations of the analysis and suggestions for future research.

2 Review of the Literature

Developing a framework to measure the value of responsiveness presented the following challenges:

1. Understanding what is a responsive supply network and assessing whether LargeCo needed a responsive supply network
2. Determining if KPIs used by LargeCo were appropriate for a responsive supply network
3. Designing an analytical framework to measure the relationship between KPIs and sales for LargeCo

Review of literature was a key source in understanding characteristics of responsive supply networks, obtaining information on appropriate KPIs to analyze performance of responsive supply networks and building an analytical framework founded on econometric and causal models to assess relationship between KPIs and sales. Details of publications, research papers and research work found in literature relevant to the areas listed above are presented in the following sections.

2.1 Responsive Supply Networks and Performance Measures

Markets for LargeCo's products are characterized by low growth rates, demand uncertainty and product variety. Randall et al., (1997) in their work on responsive vs. efficient supply networks highlight that a responsive supply network is better suited for organizations that operate in such markets. Therefore it comes as no surprise that LargeCo focuses on a responsive supply network. Randall et al., also point out that a responsive supply chain is characterized by short production lead times and small batch sizes to allow the supply chain to be responsive. LargeCo carries out initiatives to improve the responsiveness of its supply chain and some of these initiatives are designed to reduce production lead times and batch sizes. Hausman (2003) and Chan & Qi (2004) argue for the use of metrics to measure the performance of supply chains. Hausman recommends metrics that measure, among other elements, the time and cost effectiveness of the supply chain. Some of the metrics highlighted by the author include Fill Rate, Inventory Turnover (inverse of Days of Inventory) and Supply Chain Cycle Time. These exact metrics are among a basket of metrics used by LargeCo to monitor the performance of its supply network.

2.2 Relationship between KPIs and Sales

LargeCo believes that initiatives it carries out to improve the responsiveness of the supply chain helps respond rapidly to market changes and reduce lost sales. Hoover et al., (2001) highlight that organizations build market-responsive supply chains to respond rapidly to unpredictable demand. The ability to respond rapidly to unpredictable demand through improved responsiveness minimizes lost sales without building excess inventory. Reduction in lost sales results in growth in sales. Therefore LargeCo is interested in linking the improvements in responsiveness to improvements in sales. Linking KPIs, which measure responsiveness of a supply chain, to sales helps measure the value of a responsive supply chain. No references could be found in literature on research work linking improvements in KPIs to improvements in sales or on measuring the value of a responsive supply network. This area of research appears to be unexplored territory.

2.3 Analytical Framework to Measure Responsiveness

Developing an analytical framework to link improvements in KPIs with improvements in sales requires a framework that can identify relationships between KPIs and sales. Besides identifying relationships between KPIs and sales, the framework should also measure and explain the causality in the relationships. Search for similar frameworks in the field of supply chain research led to a research paper by Narasimhan & Jayaram (1998). In their paper, the authors examined the relationship among a set of performance measures such as sourcing decisions, manufacturing goals, customer responsiveness, and manufacturing performance. The authors used a framework that included a combination of hypothesis testing, regression analysis and causal modeling to examine the relationship among various performance measures. The approach used by Narasimhan & Jayaram led to the development of a similar analytical framework used in this research work for LargeCo.

The analytical framework for this research uses econometric and causal models. The building blocks for the econometric model were developed using multiple reference texts on econometrics. In particular publications of Kennedy (2003), Kahane (2007) and Anderson et al., (2007) led to the development of the two stage econometric model that included correlation analysis and hypothesis testing using regression analysis. Refining the econometric model to include deseasonalization and standardization were based on the works of Sims (1974) and Feinstein (1996). Addition of instrumental and dummy variables to enhance the econometric model was based on notes and publications of Ejrnæs & Kongsted (2004) and Suits (1975).

Schwab (2005) provides reasoning for using causal models to analyze findings from econometric model. McConnell et al., (2004) provide the reasoning for going beyond just an econometric model by arguing that correlation does not explain causality. Development of causal diagrams was based on lecture notes and publications on System Dynamics by Sterman (2000).

Incorporation of all these elements led to the development of the analytical framework which became the foundation of this research work.

3 Methods

The focus of this research is to determine if improvement in KPIs and improvement in sales have a relationship. If a relationship is identified then the relationship needs to be explained. This is a two-stage process; stage1 determines if a relationship exists between any of the KPIs and sales, and stage 2 explains the relationship if it exists. An analytical framework was developed to manage this two-stage process. The framework uses two models – an econometric model and a causal model. The econometric model helps determine if a relationship exists between any of the KPIs and sales (stage 1). The causal model helps explain the relationship if one exists (stage 2).

The analytical framework in made up of 5 steps. These are:

1. Define Hypothesis
2. Identify Variables
3. Remove Effects
4. Assess Relationship
5. Finalize Inferences

Figure 3.1 provides an illustrated representation of the analytical framework with a summary of each of these 5 steps. These steps are explained following the illustration.

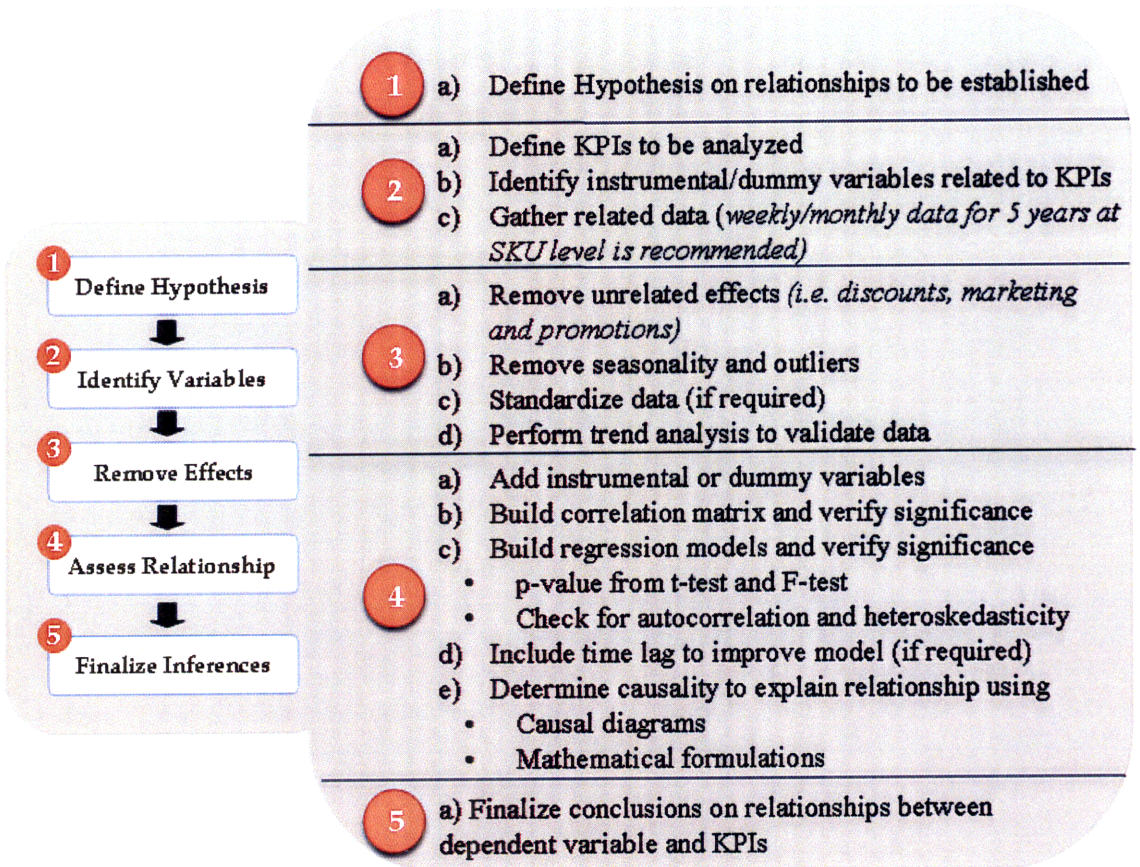

Figure 3.1 Analytical Framework

3.1 Analytical Framework

3.1.1 Define Hypothesis

The foundation of the econometric model is the hypothesis that is to be tested. Defining the hypothesis involves the following steps:

a) Stating the hypothesis to be tested

b) Listing the assumptions in testing the hypothesis

c) Identifying the test statistic to be used

d) Defining the criterion to accept or reject the hypothesis

a) Stating the hypothesis to be tested

In econometrics, the hypothesis to be tested is defined as null hypothesis – H_0. Hypothesis testing results in acceptance of the hypothesis or rejection of the hypothesis. When the null hypothesis is rejected, the alternative hypothesis H_1 is accepted.

The research question for this research work is 'Are improvements in KPIs related to improvements in sales?". Linear regression can be used to formulate the response to this question. Linear regression helps define a linear relationship between a dependent variable (y) and an independent variable (x) as shown below (Levin & Rubin, 1996):

$$y = f(x, \beta_1) \text{ (equation 3.1)}$$

$$y = a + x * \beta_1 + e \text{ (eq. 3.2)}$$

Given that LargeCo's interest is in determining if improvements in KPIs drive improvements in sales, the relationship between each KPI and sales is defined as linear regression with sales as the dependent variable and each KPI as independent variable. Therefore

$$y = f(KPI_n, \beta_1) \text{ (eq. 3.3)}$$

$$y = a + KPI_n * \beta_1 + e \text{ (eq. 3.4)}$$

where y = sales and KPI_n varies from KPI_1 to KPI_5. KPI_1 is Days of Inventory, KPI_2 is Fill Rate, KPI_3 is Logistics Cost, KPI_4 is Supply Chain Cycle Time and KPI_5 is Time Between Production Runs. a and β_1 are coefficients in the linear regression function symbolizing the y intercept and slope of the line. e is the error term or residual that explains the difference between the predicted value of y and the actual value of y. The linear regression when simplified to its generic form for any KPI will read as

$$Y = f(KPI, \beta_1) \text{ (eq. 3.5)}$$

$$y = a + KPI * \beta_1 + e \text{ (eq. 3.6)}$$

Data for y and each KPI is used to estimate the values of a and β_1 by adopting the Ordinary Least Squares (OLS) method as the estimation method. While many pairs of values can be generated for a and β_1 that fit eq. 3.6, the OLS method helps estimate the values of a and β_1 in such a way that the value of *e* is minimized. If a relationship exists between improvements in any of the KPIs and improvements in sales then β_1 will have a non-zero value $(\beta_1 \neq 0)$. If $\beta_1 = 0$ eq. 3.6 is reduced to y = a. This implies that y is independent of KPI and therefore improvements in sales do not have a relationship with improvements in KPI. Therefore the null hypothesis and the alternative hypothesis can be stated as:

Null Hypothesis: $\beta_1 = 0$ (eq.3.7)

Alternative Hypothesis: $\beta_1 \neq 0$ (eq. 3.8)

The null and alternative hypotheses can also be stated as (Kahane, 2007):

Null Hypothesis: The impact of improvements in KPI on Sales is not significantly different from zero

Alternative Hypothesis: The impact of improvements in KPI on Sales is significantly different from zero

b) Listing the assumptions in testing the hypothesis

Some of the key assumptions in formulating and testing the hypothesis are (Kennedy, 2003):

i) The regression coefficients a and β_1 should be linear, i.e., they should be first order coefficients

ii) There is no independent variable other than KPI that affects sales. Violation of this assumption leads to a poor definition of relationship between a KPI and sales. This assumption is discussed further in the section on 'Identify Variables'

iii) The expected value of the error term, e is zero. The error terms are uncorrelated (no autocorrelation) and variance of the error terms is constant (homoscedasticity). These assumptions imply that estimates of a and β_1 will be unbiased and consistent

c) Identifying the test statistic to be used

In a simple linear regression, the t-test is used to determine if the regression coefficient β_1 is significantly different from zero. In a multiple regression model given multiple regression coefficients β_1 to β_n, the F-test is used to verify if all of the regression coefficients are significantly different from zero (Mooney & Swift, 1999). Given the simple linear regression relationship between KPI and sales, the econometric model uses the t-test. The t-test is the ratio of estimated regression coefficient to its standard error (Allen, 2004). The t-ratio helps compute the probability which is then used to accept or reject the null hypothesis (Kahane, 2007).

d) Defining the criterion to accept or reject the hypothesis

The t-ratio helps determine whether a relationship identified between any KPI and sales is significant or a chance correlation (Ejrnæs & Kongsted, 2004). Using the t-ratio and the standard t-distribution table, a probability known as p-value is computed. p-value ranges from 0 to 1 and signifies the probability that the coefficient is significantly different from zero. For example, a p-value of 0.1 indicates that the null hypothesis can be rejected at 10% (=0.1) significance level or at 90% (=1.0 - 0.1) confidence level[1] (Kahane, 2007). In applied econometric models, p-value of 0.05 or less (5% significance level or 95% confidence level) is used to reject the null hypothesis. The econometric model for this research uses a p-value of 0.05.

Defining the hypothesis leads to the next step of finalizing variables and collecting data for these variables so that the hypothesis can be tested.

3.1.2 Identify Variables

LargeCo uses five different KPIs – Days of Inventory, Fill Rate, Logistics Cost, Supply Chain Cycle Time and Time Between Production Runs to track the responsiveness of its supply chain. The econometric model in the analytical framework is designed to test the relationship between a dependent variable and a set of independent variables. Sales is the dependent variable. Each of the five KPIs is to be tested independently for its relationship with sales. So in effect the hypothesis defined in the previous section can be re-written as

Null Hypothesis-1: The impact of improvements in Days of Inventory on Sales is not significantly different from zero

Null Hypothesis-2: The impact of improvements in Fill Rate on Sales is not significantly different from zero

[1] confidence level = 100 – significance level

Null Hypothesis-3: The impact of improvements in Logistics Cost on Sales is not significantly different from zero

Null Hypothesis-4: The impact of improvements in Supply Chain Cycle Time on Sales is not significantly different from zero

Null Hypothesis-5: The impact of improvements in Time Between Production Run on Sales is not significantly different from zero

It is likely that sales is affected not just by endogenous variables such as KPI but also by exogenous economic variables that may be hard to measure. An alternative to finding hard-to-measure variables to add to the linear regression would be the adoption of the instrumental variable method (Ejrnæs & Kongsted, 2004). Instrumental variables are variables that affect the dependent variable only through an independent variable. In the case of LargeCo, no data was available to identify suitable instrumental variables.

Other than instrumental variables, dummy variables are another set of independent variables that can strengthen a linear regression model. A dummy variable can be used to model events that are not usually measured using a numerical scale (Suits, 1957). For example, if sales is affected by an exogenous event such as a natural disaster, then the occurrence or absence of a natural disaster can be modeled using a dummy variable to assess its impact on the dependent variable, sales. The occurrence of an event would be represented as 1 and the absence of the event would be represented as 0. In the case of LargeCo, no data was available to identify suitable dummy variables. Therefore in the absence of suitable instrumental and dummy variables, the econometric model used each KPI as the only independent variable.

Finalization of variables in the econometric model leads to collection of data for the dependent and independent variables. Data should be collected in a disaggregated form so that meaningful conclusions can be drawn from hypothesis testing. Aggregation of data may lead to misleading or spurious conclusions (Brook & Arnold, 1985; Garrett, 2003). For hypothesis testing linked to KPI and sales, data obtained at the lowest level of time and product dimensions would provide the lowest level of disaggregation. Lowest suitable level of time data for KPI and sales would be weekly data. Lowest level of product data would be data at the level of Stock Keeping Units (SKUs). At LargeCo data was available in monthly time buckets and for groups of SKUs aggregated as a product line.

3.1.3 Remove Effects

Collected data needs to be reviewed so that it can be cleaned for unrelated effects. Sales data, for example, needs to be cleansed of effects such as discounts, marketing and promotions. If sales data with these unrelated effects is used to assess relationship with a KPI, it could lead to misleading conclusions. This is because effects other than that of the KPI are present in sales but the econometric model is designed to assess only the effect of the KPI. If effects of discounts, marketing and promotions cannot be removed from sales then these effects should be modeled as additional independent variables or as instrumental variables. LargeCo uses a proxy measure of sales called SNE. This is a measure of sales with the effects of discounts, marketing and promotions removed.

Besides unrelated effects, effects of seasonality need to be removed from sales data. Seasonality can affect the econometric model resulting in misleading conclusions on the effect of a KPI on sales. Data can either be deseasonalized using smoothing techniques (Sims, 1974) or seasonality can be modeled as a dummy independent variable (Harvey & Proietti, 2005). LargeCo's sales data did not show effects of seasonality and was not deseasonalized.

Regression coefficient β_1 is sensitive to the scale of the independent variable (Feinstein, 1996). For example if the value of an independent variable varies either between two large extremes or two narrow extremes, it is bound to affect the regression coefficient disproportionately. To avoid this effect, the independent variable should be standardized. Standardization is done by subtracting the mean of the independent variable from each observation and then dividing it by its standard deviation. For example, if each observation of independent variable is labeled x_i, the mean of all observations of that independent variable is μ and the standard deviation is σ then the standardized value z_i of each observation is calculated as

$$z_i = \frac{x_i - \mu}{\sigma} \quad \text{(eq. 3.9)}$$

Standardization using this approach allows for rescaling the independent variable to draw meaningful inferences from the analysis using the econometric model.

Finally trend analysis of data helps determine if a linear trend is present. If trend plots of the dependent variable and independent variable do not reflect a straight line best-fit then the relationship between the two variables may not be linear (Keppel & Zedeck, 1989). For LargeCo, trend analysis of sales and each of the KPI confirmed a linear relationship between the two sets of variables.

The review and validation of data leads to the next step of assessing the relationship between the two sets of variables.

3.1.4 Assess Relationship

The analytical framework uses a two-step process to assess the relationship between each KPI measure and sales. The econometric model is the first step in the process and it helps determine if a relationship exists between the two variables. The causal model is then used to explain the relationship.

3.1.4.1 Econometric Model

Prior to performing any econometric analysis, any instrumental or dummy variables identified in step 3.1.2 should be incorporated in the econometric model. For LargeCo no instrumental or dummy variables were identified. Therefore the hypothesis to be tested is limited to assessing the relationship between two variables - KPI and sales, alone.

The two econometric methods that can be applied on the variables are correlation analysis and regression analysis. Correlation analysis yields a correlation coefficient which measures the strength of a linear association between each KPI and sales (Anderson et al., 2007). The correlation coefficient is measured as

$$r_{xy} = \frac{s_{xy}}{s_x s_y} \quad \text{(eq. 3.10)}$$

where r_{xy} is the correlation coefficient, s_{xy} is the covariance, s_x is the standard deviation of x and s_y is the standard deviation of y for two sets of variables x and y. Covariance is calculated as

$$s_{xy} = \frac{\sum (x-\bar{x})(y-\bar{y})}{n-1} \quad \text{(eq. 3.11)}$$

Where $\bar{x}$ is the average of x and $\bar{y}$ is the average of y, n is the number of observations and the summation is across all observations of x and y.

The value of correlation coefficient varies from -1 to +1. A value of +1 indicates that KPI and sales are positively correlated implying that they move in the same direction; so improvement in KPI results in improvement in sales or deterioration in KPI causes deterioration in sales. A value of -1 indicates that KPI and sales are negatively correlated implying that improvement in KPI results in deterioration in sales. A value of zero or close to zero indicates that KPI and sales are not related.

While a correlation coefficient is a good measure of strength of relationship, additional measures are required to verify the relationship to ensure that the measure reflected by correlation coefficient is not a measure of chance. Regression analysis yields a coefficient of determination, known as r-squared, which also measures strength of relationship. Coefficient of determination is calculated as the square of correlation coefficient. Coefficient of determination is a useful measure of strength of relationship in the case of nonlinear relationships or for linear relationships with two or more independent variables (Anderson et al., 2007). Since the relationship between KPI and sales involves a single independent variable, the econometric model does not use the coefficient of determination as an additional measure.

While performing regression analysis in statistical packages such as SPSS or software such as Microsoft Excel, the package or software also provides another measure known as p-value. The p-value along with the t-ratio, as described in section 3.1.1, provides a measure to accept or reject the null hypothesis. This approach of using the t-ratio and p-value to accept or reject the null hypothesis is the t-test. If the relationship to be assessed involved more than one independent variable then the F-test, instead of the t-test, will be used. As stated in section 3.1.1, the t-test is used in this research to analyze the relationship between KPI and sales.

In general smaller values of p-value provide less support for the null hypothesis and the approach to using p-value is as follows (Anderson et al., 2007):

- p-value less than 0.01: Overwhelming evidence to reject null hypothesis H_0
- p-value between 0.01 and 0.05: Strong evidence to reject null hypothesis H_0
- p-value between 0.05 and 0.10: Weak evidence to reject null hypothesis H_0

- p-value greater than 0.10: Insufficient evidence to reject null hypothesis H_0

In each of these cases, rejection of null hypothesis implies acceptance of the alternative hypothesis H_1. In the case of LargeCo, p-value of 0.05 or lower is used to reject the null hypothesis H_0.

In situations where the econometric model does not indicate a relationship between KPI and sales, it is important to review if the effect of an improvement in KPI is felt in sales with a time lag. If so it may be necessary to incorporate a time lag in the analysis (Ostrom, 1990). A time lag can be incorporated by assessing the relationship between KPI from time period 1 with sales from time period (n+1). The value of n determines the lag in time period. Value of 1 implies a lag of 1 time period while a lag of 3 implies a lag of 3 time periods. It is important to note that larger the lag, fewer the observations to assess using the econometric model. Therefore the length of the lag in the case of LargeCo was limited to 4 time periods.

The econometric model helps determine if there is a strong correlation between improvements in KPI and improvements in sales. However it is important not to assume relationships based just on econometric measures. Correlation between two sets of variables does not always explain causality (McConnell et al., 2004). This drives the need for a causal model to support the econometric model.

3.1.4.2 Causal Model

Econometric models help determine if a relationship between two sets of variables is statistically significant. However the econometric models do not explain the causality in the relationship. Causal models are used to explain the causality in a relationship. The econometric models provide the data to support the causal models (Schwab, 2005). The analytical framework includes two types of causal models to explain relationships identified by the econometric model. The two causal models are causal diagrams and mathematical formulations.

Causal diagrams link dependent and independent variables to explain causality. Arrows are drawn from the independent variable ("cause") to the dependent variable ("effect") from left to right. Correlation of the causality is identified using + or – symbol on top of the arrow head. The + sign indicates positive correlation signifying relationship in the same direction while the – indicates negative correlation signifying relationship in the opposite direction. Examples of causal diagrams are shown below. The causal diagram below links inventory on hand (independent variable, x) and inventory costs (dependent variable, y). The + sign in the diagram indicates that as inventory on hand increases, inventory costs increase or as inventory on hand decreases, inventory costs decrease.

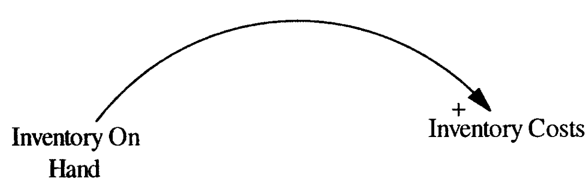

Figure 3.2 Causal Diagram: Example 1

The causal diagram below links stockout rate (independent variable, x) and sales (dependent variable, y). The – sign signifies that as stockout rate decreases, sales increase or vice versa.

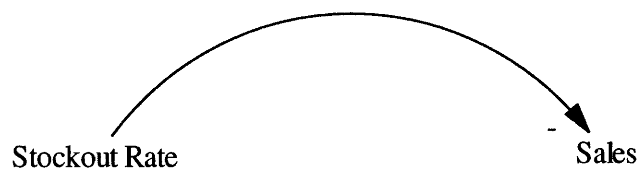

Figure 3.3 Causal Diagram: Example 2

Another approach to explain causality in relationships identified using econometric models is deriving mathematical formulations linking independent and dependent variables. For example the mathematical formulation below links inventory on hand and inventory costs.

InventoryCost = InventoryOnHand/2 * InventoryHoldingCost (eq. 3.12)

Identifying and explaining causality using causal diagrams and mathematical formulations is the final step in identifying and explaining relationships between dependent and independent variables.

3.1.5 Finalize Inferences

Using the econometric and causal models, inferences can be drawn from relationships defined and explained using the models. Inferences can be quantitative or qualitative in nature. Quantitative inferences, for example, can measure the impact of independent variable on dependent variable. Qualitative inferences, on the other hand, provide insights into the relationship between the two sets of variables. This is the final step in the analytical framework in assessing relationships between the identified variables.

3.2 Summary

The analytical framework helps identify and define relationships between variables by building a hypothesis connecting the variables, gathering and processing data, assessing relationships and drawing inferences. The next section provides details of analysis performed for LargeCo and inferences drawn from the analysis.

4 Analysis and Findings

This section covers details of application of the analytical framework using data provided by LargeCo to determine if improvements in KPI drive improvements in sales. Sales and KPI data for a period of 32 months for five product lines with varying levels of details were available. The analytical framework with its five step process was used to analyze the data and derive inferences.

4.1 Define Hypothesis

Five sets of null hypothesis were defined to assess the relationship between each of the five KPIs and sales. These are:

$$\text{Sales} = a_1 + DI * \beta_1 + e_1 \Rightarrow \beta_1 = 0 \text{ (eq. 4.1)}$$

$$\text{Sales} = a_2 + FR * \beta_2 + e_2 \Rightarrow \beta_2 = 0 \text{ (eq. 4.2)}$$

$$\text{Sales} = a_3 + LC * \beta_3 + e_3 \Rightarrow \beta_3 = 0 \text{ (eq. 4.3)}$$

$$\text{Sales} = a_4 + SCCT * \beta_4 + e_4 \Rightarrow \beta_4 = 0 \text{ (eq. 4.4)}$$

$$\text{Sales} = a_5 + TBPR * \beta_5 + e_5 \Rightarrow \beta_5 = 0 \text{ (eq. 4.5)}$$

In addition a sixth hypothesis was defined to verify if the five KPIs had a combined effect on sales. This is stated as:

$$\text{Sales} = a_6 + DI * \beta_6 + FR * \beta_7 + LC * \beta_8 + SCCT * \beta_9 + TBPR * \beta_{10} + e_6 \Rightarrow$$

$$\beta_6, \beta_7, \beta_8, \beta_9, \beta_{10} = 0 \text{ (eq. 4.6)}$$

The six hypotheses were tested using the analytical framework for each of the five product lines based on availability of data.

4.2 Identify Variables

Varying level of detail was available for each of LargeCo's five product lines. The table below summarizes the information available for analysis in each of the five product lines.

Product Line	Observation Type	Number of Observations						Instrumental Variable	Dummy Variable
		Sales	FR	ITO	LC	SCCT	TBPR		
ProductLine1	Monthly	30	30	24	30	30	30	-	-
ProductLine2	Monthly	30	30	24	30	30	30	-	-
ProductLine3	Monthly	17	-	25	-	-	-	-	-
ProductLine4	Monthly	25	31	31	-	31	31	-	-
ProductLine5	Monthly	25	-	25	-	-	-	-	-

Table 4.1 Observation Statistics

4.3 Remove Effects

In this stage of analysis, sales data is stripped of effects, deseasonalized, normalized and analyzed for non-linear trends.

The data provided by LargeCo for sales was a proxy measure of sales called Sales Net of Effects. This measure of sales was devoid of the effects of discounts, marketing and promotions. The data was analyzed for seasonality and did not show any seasonal effects.

Fill Rate (FR) for the three product lines, for which information was available, varied within a very narrow range of 1 percentage point. As discussed in section 3.1.3, the scale of the independent variable has a significant impact on the regression co-efficient. If data for the independent variables falls within a narrow range or a very large range, its impact on the dependent variable as measured by linear regression is disproportionate. Therefore the data for Fill Rate was normalized using eq. 3.9 converting the values for Fill Rate to a scale of -3 to +3.

Trend analysis of the data for all product lines across KPIs and sales did not show any non-linear trend and therefore the relationship between KPIs and sales was considered to be linear.

4.4 Assess Relationship

In this stage of analysis, LargeCo's data is analyzed using econometric and causal models. The details of the analysis using these models and findings from the analysis are summarized below.

4.4.1 Econometric Model

4.4.1.1 Correlation Analysis

The econometric model uses correlation analysis and statistical tests for significance of relationship. The correlation coefficient helps measure the strength and direction of relationship. As stated in section 3.1.4.1, the value of correlation coefficient varies from -1 to +1 and is computed using eqn. 3.10. The correlation matrices for LargeCo's five product lines are tabulated below.

ProductLine1

The correlation matrix for ProductLine1 is as shown below.

	SNE	*TBPR*	*FR*	*DI*	*SCCT*	*LC*
SNE	1.00					
TBPR	-0.01	1.00				
FR	0.09	*-0.35*	1.00			
DI	**-0.49**	-0.14	*-0.22*	1.00		
SCCT	*-0.32*	*-0.28*	0.13	**0.69**	1.00	
LC	0.03	**-0.55**	*0.22*	0.08	*0.28*	1.00

Table 4.2 Correlation Matrix: ProductLine1

The correlation matrix provides the correlation coefficient for each pair of variables. The analytical framework relied on the constraint that only correlation coefficient values above 0.4 or below -0.4 were to be considered significant (shown in bold). Values between 0.2 to 0.4 or -0.2 to -0.4 were marked as noteworthy (shown in italics) and for further analysis. Correlation of a variable with itself is expected to be 1 as shown above and can be ignored.

From the correlation matrix it can be seen that a significant negative or inverse relationship exists between Days of Inventory and Sales. The sign of the coefficient signifies that as Days of Inventory decreases (increases) Sales increases (decreases). Other significant relationships identified by correlation analysis include an inverse relationship between Logistics Cost and Time Between Production Runs and a direct relationship between Supply Chain Cycle Time and Days of Inventory. Additional relationships found to be noteworthy are shown in italics.

ProductLine2

The correlation matrix for ProductLine2 is as shown below.

	SNE	*TBPR*	*FR*	*DI*	*SCCT*	*LC*
SNE	1.00					
TBPR	0.03	1.00				
FR	*0.24*	-0.10	1.00			
DI	**-0.42**	-0.02	**-0.57**	1.00		
SCCT	*0.23*	*-0.22*	-0.03	-0.17	1.00	
LC	-0.19	*-0.24*	0.14	*-0.25*	-0.05	1.00

Table 4.3 Correlation Matrix: ProductLine2

The correlation matrix shows that Days of Inventory has strong negative correlation with Sales indicating an inverse relationship between the two variables. This is similar to the finding for ProductLine1. In addition an inverse relationship appears to exist between Days of Inventory and Fill Rate.

As stated in section 3.4.1 correlation alone does not signify causality and therefore further analysis using t-test and causal diagrams will be required to confirm findings from the correlation analysis.

ProductLine3

Data only for Sales and Days of Inventory were available for ProductLine3. The correlation matrix for ProductLine3 is as shown below.

	SNE	*DI*
SNE	1	
DI	**-0.50**	1

Table 4.4 Correlation Matrix: ProductLine3

The analysis shows that the correlation coefficient for Sales and Days of Inventory at -0.50 appears to be significant and in line with the findings for ProductLine1 and ProductLine2.

ProductLine4

The correlation matrix for ProductLine4 is as shown below.

	SNE	*TBPR*	*FR*	*DI*	*SCCT*
SNE	1.00				
TBPR	*-0.30*	1.00			
FR	0.02	-0.11	1.00		
DI	0.01	*0.35*	*-0.34*	1.00	
SCCT	0.04	0.01	-0.02	**0.40**	1.00

Table 4.5 Correlation Matrix: ProductLine4

Correlation analysis for ProductLine4 does not show any significant correlation between sales and any of the five KPIs. Similar to the finding for ProductLine1 a direct relationship between Supply Chain Cycle Time and Days of Inventory appears to be significant. Figures in italics refer to relationships between variables that appear to be statistically noteworthy for further analysis.

ProductLine5

Similar to ProductLine3, for ProductLine5 data was available only for Sales and Days of Inventory. The correlation matrix for these two variables for this product line is as shown below.

	SNE	*DI*
SNE	1	
DI	*-0.29*	1

Table 4.6 Correlation Matrix: ProductLine5

The analysis shows that the correlation coefficient for Sales and Days of Inventory is statistically noteworthy (between -0.2 and -0.4). The inverse relationship between KPI and sales appears to be consistent with the findings for other product lines.

As discussed in section 3.1.4.1 a time lag was incorporated in the correlation analysis between KPIs and sales to determine if improvements in KPIs had a relationship with sales after a time lag. The rationale behind this approach is that for some KPIs or for some product lines it is likely that an improvement in the KPI begins to have an effect on sales after a time lag. To perform a time lag analysis, KPI data for periods 1 to n is compared with sales data for periods 1+m to n+m where m is the time lag. For example, if it is hypothesized that improvement in Supply Chain Cycle Time has an effect on Sales after a three month delay then data for Supply Chain Cycle Time for periods 1 to 27 will be compared with Sales data for periods 4 to 30 to determine if a relationship exists. The results of the time lag based correlation analysis are shown below only for those product lines and KPIs that showed a significant relationship.

KPI	**Product Line**	**Correlation with Sales**	**Time Lag**
SCCT	ProductLine1	-0.53	2 months
DI	ProductLine4	0.63	3 months
SCCT		0.50	3 months
TBPR		0.45	4 months

Table 4.7 Correlation Matrix with Time Lag

The analysis shows that Supply Chain Cycle Time has an inverse relationship with Sales when lagged for 2 months for ProductLine1. Days of Inventory, Supply Chain

Cycle Time and Time Between Production Runs appear to have a direct relationship with Sales when lagged. However the direct nature of relationship appears to indicate that for all these KPIs, sales decreases when these KPIs show an improvement (decrease). This is quite in contrast to generally accepted linkages in the supply chain function between sales and these KPIs. Therefore the results of the time lagged correlation analysis for ProductLine4 will be reviewed in detail in the next stage of analysis.

A summary of the relationships between KPIs and sales found to be significant or noteworthy is provided below. These relationships will be assessed further in the next stage of analysis.

KPI	Product Line	Correlation with Sales	Time Lag
DI	ProductLine1	-0.49	None
SCCT		-0.53	2 months
DI	ProductLine2	-0.42	None
DI	ProductLine3	-0.50	None
DI	ProductLine4	0.63	3 months
SCCT		0.50	3 months
TBPR		0.45	4 months
DI	ProductLine5	-0.29	None

Table 4.8 Summary of Correlation Analysis

4.4.1.2 Regression Analysis

Given the findings from correlation analysis, it is necessary to confirm if the relationships identified by correlation analysis as significant or noteworthy are in fact statistically significant. Regression analysis with t-test helps achieve this objective. Software such as Microsoft Excel generates t-test statistics and p-values when performing the regression analysis. The results and findings from regression analysis for the five product lines are described below.

ProductLine1

The table below provides a summary of the statistics from the regression analysis for ProductLine1. It also includes the correlation coefficient values from the previous section to provide a summary of statistics for the analysis done using the econometric model.

KPI	Correlation	Adjusted R^2	t	p-value	Significance Level	Confidence Level	Lag
TBPR	-0.01	-0.04	-0.03	0.98	98%	2%	None
FR	0.09	-0.03	0.46	0.65	65%	35%	None
DI	**-0.49**	**0.20**	**-2.61**	**0.02**	**2%**	**98%**	None
SCCT	**-0.53**	**0.25**	**-3.20**	**0.00**	**0%**	**100%**	2 Months
LC	0.03	-0.04	0.14	0.89	89%	11%	None

Table 4.9 Regression Analysis: ProductLine1

The summary includes the correlation coefficient, adjusted r-squared value, the t test statistic value, p-value, the significance level, the confidence level and time lag. Adjusted r-squared is an adjusted value of the coefficient of determination, r-squared. It is calculated by adjusting for the number of independent variables in the linear regression function. Since the coefficient of determination can always be improved by adding additional independent variables, a true measure of the relationship is the adjusted r-squared (Ostle & Malone, 1988). It is always less than or equal to the value of r-squared. As stated in 3.1.4.1, given that the relationship between each KPI and sales is being assessed, there are only two variables – a dependent variable and an independent variable, involved in the linear regression. In such a scenario, the coefficient of determination does not provide any additional insight and therefore the econometric model does not use the r-squared or adjusted r-squared values for any conclusions. The **t** value is the t-test statistic as described in section 3.1.1. The p-value measures the probability with which the null hypothesis can be rejected. Significance level is the same as p-value but expressed in % terms. Confidence level is 1 - p-value expressed in % terms. As described in section 3.1.4.1, p-value of 0.05 or lower will be used to reject a null hypothesis. Rejecting a null hypothesis symbolizes accepting the hypothesis that the relationship between the corresponding KPI and sales is statistically significant. p-value greater than 0.05 will be used to accept the null hypothesis indicating that the relationship between the corresponding KPI and sales is not statistically significant.

Based on these stipulations, the findings show that Days of Inventory has a statistically significant relationship with Sales at 98% confidence level. The findings also show that Supply Chain Cycle Time has a statistically significant relationship with Sales with a 2-month time lag at 100% confidence level. The – sign for the correlation coefficient for both these KPIs signifies that there is an inverse relationship between these KPIs and sales implying that as these KPIs show improvement (decrease) sales is expected to increase.

An additional analysis was performed for ProductLine1 to determine if the cumulative impact of KPIs had a relationship with sales. For this analysis, Days of Inventory was ignored as it already appeared to have a relationship with Sales. Retaining it and adding more KPIs to the regression function will only increase the statistical significance of the result. Therefore the regression analysis for combined effect of KPIs was done with four KPIs excluding Days of Inventory. The results of the analysis are shown below.

KPI	**Adjusted R^2**	**t**	**p-value**	**Significance Level**	**Confidence Level**
	-0.01				
TBPR		-0.09	0.93	93%	7%
FR		0.52	0.61	61%	39%
SCCT		-1.84	0.08	8%	92%
LC		0.41	0.69	69%	31%

Table 4.10 Cumulative Regression Analysis: ProductLine1

In the cumulative regression analysis, relationship between sales and four KPIs together is being assessed. Given that the number of variables is now four, the adjusted r-squared carries significance. For a statistically significant relationship adjusted r-squared values should be closer to 1. Values closer to 0 indicate lack of statistically significant relationship. The regression analysis shows that with an adjusted r-squared value of -0.01 and p-values greater than 0.05, cumulatively KPIs other than Days of Inventory do not have a relationship with Sales for ProductLine1.

ProductLine2

The findings from regression analysis for ProductLine2 are presented below.

KPI	**Correlation**	**Adjusted R^2**	**t**	**p-value**	**Significance Level**	**Confidence Level**	**Lag**
TBPR	0.03	-0.03	0.16	0.87	87%	13%	None
FR	0.24	0.02	1.31	0.20	20%	80%	None
DI	**-0.42**	**0.14**	**-2.14**	**0.04**	**4%**	**96%**	**None**
SCCT	0.23	0.02	1.25	0.22	22%	78%	None
LC	-0.19	0.00	-1.00	0.33	33%	67%	None

Table 4.11 Regression Analysis: ProductLine2

From the analysis it can be seen that the only KPI with a statistically significant relationship to Sales is Days of Inventory. The correlation coefficient of -0.42 indicates that Days of Inventory has an inverse relationship with Sales signifying that as Days of Inventory decreases Sales increases. The p-value indicates that this relationship can be accepted at 96% confidence level. No other KPI, with or without a time lag, appears to have a statistically significant relationship with sales.

Regression analysis for the cumulative impact of KPIs (other than Days of Inventory) on sales is provided below.

KPI	Adjusted R^2	t	p-value	Significance Level	Confidence Level
	0.03				
TBPR		0.33	0.74	74%	26%
FR		1.51	0.14	14%	86%
SCCT		1.29	0.21	21%	79%
LC		-1.02	0.32	32%	68%

Table 4.12 Cumulative Regression Analysis: ProductLine2

The regression analysis shows that with an adjusted r-squared value of 0.03 and p-values greater than 0.10, cumulatively KPIs other than Days of Inventory do not have a relationship with Sales for ProductLine2.

ProductLine3

The output from regression analysis for ProductLine3 is presented below.

KPI	Correlation	Adjusted R^2	t	p-value	Significance Level	Confidence Level	Lag
DI	-0.50	0.20	-2.23	0.04	4%	96%	None

Table 4.13 Regression Analysis: ProductLine3

The output shows that Days of Inventory has a statistically significant relationship with Sales at 96% confidence level.. With no data available for other KPIs no analysis was done on cumulative effect of KPIs on sales.

ProductLine4

The output from regression analysis for ProductLine4 is presented below.

KPI	Correlation	Adjusted R^2	t	p-value	Significance Level	Confidence Level	Lag
TBPR	**0.45**	**0.16**	**2.17**	**0.04**	**4%**	**96%**	4 Months
FR	0.02	-0.04	0.10	0.92	92%	8%	None
DI	**0.63**	**0.36**	**3.58**	**0.00**	**0%**	**100%**	3 Months
SCCT	**0.50**	**0.23**	**2.60**	**0.01**	**1%**	**99%**	3 Months

Table 4.14 Regression Analysis: ProductLine4

The findings show that Time Between Production Runs, Days of Inventory and Supply Chain Cycle Time all have a statistically significant relationship with Sales at varying confidence levels, 96%, 100% and 99% respectively. These relationships are with a time lag, 4 months, 3 months and 3 months respectively. The + sign for the correlation coefficient for these three KPIs signifies that there is a direct relationship between these KPIs and sales implying that as these KPIs show improvement (decrease) sales is expected to decrease. This conclusion is quite in contrast to expected relationship between these KPIs and sales. Improvements in these KPIs should improve sales and not affect sales. Data for these KPIs were tested against themselves with the time lags listed above to determine if the statistical relationship of these KPIs with sales was driven by autocorrelation[2]. The correlation analysis did not reveal any autocorrelation effect. Therefore these findings will be reviewed in detail using a causal model to determine if these conclusions are meaningful even though they appear to be statistically significant.

Regression analysis for the cumulative impact of KPIs (other than Days of Inventory) on sales is provided below.

KPI	Adjusted R^2	t	p-value	Significance Level	Confidence Level
	-0.04				
TBPR		-1.44	0.16	16%	84%
FR		-0.06	0.96	96%	4%
SCCT		0.19	0.85	85%	15%

Table 4.15 Cumulative Regression Analysis: ProductLine4

The regression analysis shows that with an adjusted r-squared value of -0.04 and p-values greater than 0.15, cumulatively KPIs other than Days of Inventory do not have a relationship with Sales for ProductLine4.

ProductLine5

The findings from regression analysis for ProductLine5 are presented below.

KPI	Correlation	Adjusted R^2	t	p-value	Significance Level	Confidence Level	Lag
DI	-0.29	0.05	-1.46	0.16	16%	84%	None

Table 4.16 Regression Analysis: ProductLine5

[2] Autocorrelation occurs when data for a variable has a correlation with itself when time-lagged

The output shows that Days of Inventory does not have a statistically significant relationship with Sales. With no data available for other KPIs no analysis was done on cumulative effect of KPIs on sales.

Starting with eight relationships (Table 4.8) for validation, the regression analysis helped verify which of these relationships are statistically significant. A summary of the relationships found to be statistically significant is provided below. These relationships will be assessed further in the next stage of analysis.

KPI	Product Line	Correlation	Adjusted R^2	t	p-value	Significance Level	Confidence Level	Lag
DI	ProductLine1	-0.49	0.20	-2.61	0.02	2%	98%	None
SCCT		-0.53	0.25	-3.20	0.00	0%	100%	None
DI	ProductLine2	-0.42	0.14	-2.14	0.04	4%	96%	None
DI	ProductLine3	-0.50	0.20	-2.23	0.04	4%	96%	None
DI	ProductLine4	0.63	0.36	3.58	0.00	0%	100%	3 Months
SCCT		0.50	0.23	2.60	0.01	1%	99%	3 Months
TBPR		0.45	0.16	2.17	0.04	4%	96%	4 Months

Table 4.17 Summary of Analysis using Econometric Model

The econometric model helped identify statistically significant relationships between KPIs and sales for the various product lines. These relationships will be validated using the causal model to determine if causality exists to support these relationships.

4.4.2 Causal Model

The causal model for this research work relies on two tools – mathematical formulations and causal diagrams, to explain causality and provide support to relationships identified as statistically significant using the econometric model.

4.4.2.1 Mathematical Formulations

This section covers developing mathematical formulations to link each of the KPIs with sales.

Causality between Sales and Days of Inventory

The mathematical formulation for calculating Days of Inventory is as shown below:

Days of Inventory =

$$\frac{365}{InventoryTurnover} = \frac{365}{\left(\frac{CostOfSales}{AverageInventory}\right)} = \frac{365 * AverageInventory}{CostOfSales} \quad \text{(eq. 4.7)}$$

From the formulation it can be seen that Days of Inventory is inversely related to Cost of Sales. Cost of Sales is directly related to Sales, i.e., as Sales increases Cost of Sales increases. From this it can be formulated that Days of Inventory is inversely related to Sales, i.e., as Sales increases Days of Inventory decreases.

$$Days\ of\ Inventory \propto \frac{1}{CostofSales} \propto \frac{1}{Sales} \Rightarrow Days\ of\ Inventory \propto \frac{1}{Sales} \quad \text{(eq. 4.8)}$$

From this formulation it can be seen that contrary to the original hypothesis that Sales was dependent on Days of Inventory, Sales drives Days of Inventory. The findings from the econometric model still hold good given that the econometric model helped determine the direction and strength of the relationship between these two variables. The causality shows that Sales drives Days of Inventory and therefore increase in Sales drives reduction in Days of Inventory.

Causality between Sales and Fill Rate

Sales of any organization selling products and services are defined by the portion of market demand for its products and services fulfilled by the organization. This % of met demand is knows as Fill Rate. Therefore the formulation is:

$$Sales = Demand * FillRate \Rightarrow Sales \propto FillRate \quad \text{(eq. 4.9)}$$

As can be seen from the formulation, Sales is directly related to Fill Rate and is dependent on Fill Rate. If Fill Rate declines Sales decline. The causality confirms that Fill Rate drives Sales and that increase in Fill Rate drives increase in Sales.

Causality between Sales and Logistics Cost

Logistics Cost is one of the sub-elements of the aggregate measure Cost of Sales. Other sub-elements of Cost of Sales include Material Cost, Labor Cost and other costs. Therefore the relationship between Logistics Cost and Cost of Sales can be expressed as

$$CostofSales \propto LogisticsCost \quad \text{(eq. 4.10)}$$

It is known that Cost of Sales is directly related to Sales, i.e., as Sales increases Cost of Sales increases. From this it can be formulated that Logistics Cost is directly related to Sales, i.e., as Sales increases Logistics Cost increases.

$$CostofSales \propto Sales \Rightarrow Sales \propto LogisticsCost \quad \text{(eq. 4.11)}$$

From these formulations it can be seen that contrary to the original hypothesis that Sales was dependent on Logistics Cost, Sales drives Logistics Cost and therefore increase in Sales drives increase in Logistics Costs.

Causality between Sales and Supply Chain Cycle Time

LargeCo defines Supply Chain Cycle Time as cumulative of Procurement Cycle Time, Production Cycle Time and Distribution Cycle Time. Procurement Cycle Time is the time taken to order and receive raw materials, Production Cycle Time is the time taken to manufacture products and Distribution Cycle Time is the time taken to deliver finished goods to customers. Therefore Supply Chain Cycle Time represents the aggregate cycle time across the supply chain.

Supply Chain Cycle Time is linked to capacity utilization of a production facility. As Supply Chain Cycle Time decreases capacity utilization increases thereby improving Fill Rate. Therefore Supply Chain Cycle Time is inversely related to Fill Rate. As Supply Chain Cycle Time decreases Fill Rate increases. Therefore the relationship between Fill Rate and Supply Chain Cycle Time can be expressed as

$$FillRate \propto \frac{1}{SupplyChainCycleTime} \quad \text{(eq. 4.12)}$$

It is known that Fill Rate is directly related to Sales (eq. 4.9), i.e., as Fill Rate increases Sales increases. From this it can be formulated that Sales is inversely related to Supply Chain Cycle Time, i.e., as Supply Chain Cycle Time decreases Sales increases.

$$Sales \propto FillRate \Rightarrow Sales \propto \frac{1}{SupplyChainCycleTime} \text{ (eq. 4.13)}$$

The formulation shows that Sales is inversely related to Supply Chain Cycle Time and dependent on it. The causality confirms that Supply Chain Cycle Time drives Sales and that decrease in Supply Chain Cycle Time drives increase in Sales.

Causality between Sales and Time Between Production Runs

LargeCo defines Time Between Production Runs as the duration (in days or weeks or months) between two production runs for a product. Time Between Production Runs is a sub-element of Production Cycle Time, the other sub-element being the actual Production Time. Since Production Cycle Time is a sub-element of Supply Chain Cycle Time, Time Between Production Runs is also a sub-element of Supply Chain Cycle Time. The formulation is as shown below:

$$TimeBetween\Pr oductionRuns \propto \Pr oductionCycleTime \propto SupplyChainCycleTime \text{ (eq. 4.14)}$$

This leads to the inference that Time Between Production Runs has the same relationship with Sales as Supply Chain Cycle Time. Therefore it can be shown that Sales is inversely related to Time Between Production Runs, i.e., as Time Between Production Runs decreases Sales increases.

$$Sales \propto \frac{1}{TimeBetween\Pr oductionRuns} \text{ (eq. 4.15)}$$

The formulation shows that Sales is inversely related to Time Between Production Runs and dependent on it. The causality confirms that Time Between Production Runs drives Sales and that decrease in Time Between Production Runs drives increase in Sales.

Summarizing the causal analysis using mathematical formulations, it can be seen that:

1. Sales drives Days of Inventory
2. Fill Rate drives Sales
3. Sales drives Logistics Cost
4. Supply Chain Cycle Time drives Sales
5. Time Between Production Runs drives Sales

Next the supply network of LargeCo will be analyzed using the causal diagram tool to verify the causality derived using the mathematical formulations.

4.4.2.2 Causal Diagrams

As described in section 3.1.4.2 causal diagrams help explain causality for relationships identified using the econometric model. Causal models provide the data to support the econometric model. The illustration below shows a causal diagram for a supply network identifying the links between various elements of the supply network and how they drive costs and sales within the supply network. Of particular interest are the KPIs and how they are influenced by other factors in the network and how they drive costs or sales in the supply network. Details of the causal diagram follow the illustration.

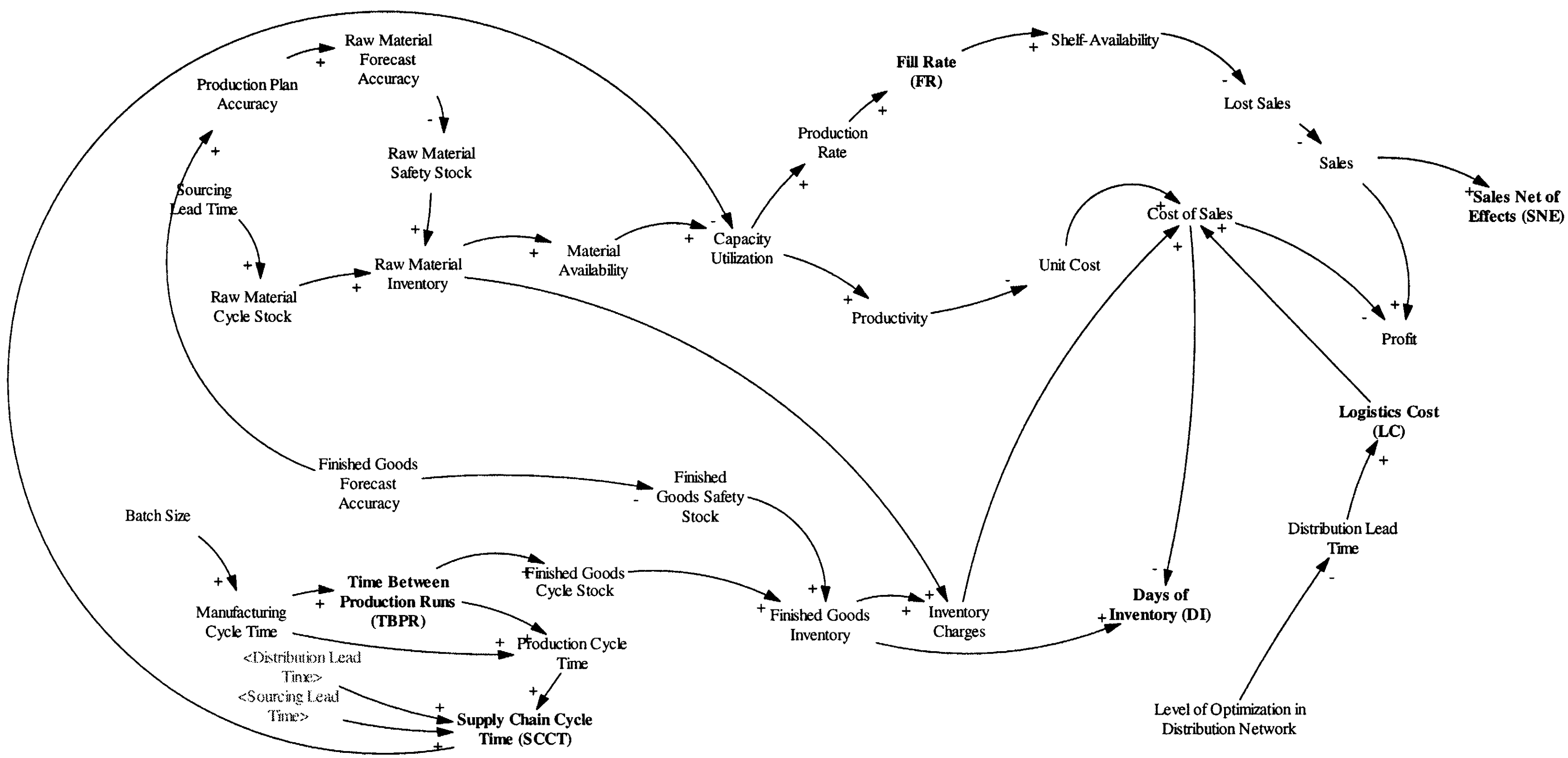

Figure 4.1 Causal Diagram of a Supply Network

The inferences from the causal diagram can be summarized as below:

1. Sales is not linked to Days of Inventory
2. Fill Rate is directly linked to Sales
3. Sales is not linked to Logistics Cost
4. Supply Chain Cycle Time is inversely linked to Sales through Fill Rate
5. Time Between Production Runs is inversely linked to Sales through Supply Chain Cycle Time

Using the combination of causal diagram and mathematical formulations relationships between the various KPIs and sales can be finalized as discussed in the next section.

4.5 Finalize Inferences

A summary of the relationships found to be significant based on econometric analysis and causality is listed below.

KPI	Product Line	Correlation	Adjusted R2	t	p-value	Significance Level	Confidence Level	Lag
DI	ProductLine1	-0.49	0.20	-2.61	0.02	2%	98%	None
SCCT		-0.53	0.25	-3.20	0.00	0%	100%	None
DI	ProductLine2	-0.42	0.14	-2.14	0.04	4%	96%	None
DI	ProductLine3	-0.50	0.20	-2.23	0.04	4%	96%	None

Table 4.18 Summary of Analysis using Analytical Framework

Given the summary above and the hypothesis defined for testing initially, it can be concluded based on econometric analysis and causality that:

$$\text{DI} = a_1 + Sales * \beta_1 + e_1 \text{ and } \beta_1 = 0 \quad \text{(eq. 4.16)}$$

$$\text{Sales} = a_4 + SCCT * \beta_4 + e_4 \text{ and } \beta_4 = 0 \quad \text{(eq. 4.17)}$$

The following hypothesis testing results are supported only by the causal model and not by the econometric model. The reasons for this behavior are discussed in the next chapter.

$$\text{Sales} = a_2 + FR * \beta_2 + e_2 \text{ and } \beta_2 = 0 \quad \text{(eq. 4.18)}$$

$$\text{LC} = a_3 + Sales * \beta_3 + e_3 \text{ and } \beta_3 = 0 \quad \text{(eq. 4.19)}$$

$$\text{Sales} = a_5 + TBPR * \beta_5 + e_5 \text{ and } \beta_5 = 0 \quad \text{(eq. 4.20)}$$

Regarding the sixth hypothesis defined to verify if the five KPIs had a combined effect on sales, it can be concluded that based on econometric analysis and causality:

$$\text{Sales} \neq a_6 + DI * \beta_6 + FR * \beta_7 + LC * \beta_8 + SCCT * \beta_9 + TBPR * \beta_{10} + e_6 \quad \text{(eq. 4.21)}$$

This is based on the finding that there is no combinatorial effect of the KPIs on sales.

5 Summary and Conclusions

In this section a summary of key findings and conclusions from the application of the analytical framework are presented.

5.1 Summary of Key Findings

Using the analytical framework and by applying the econometric and causal models, the following key findings emerge from the research work:

1. Using econometric and causal models **it is possible to establish relationships** between Key Performance Indicators and Sales
2. The initial hypothesis on Sales and Days of Inventory was that improvements in Sales were linked to improvements in Days of Inventory. The hypothesis also considered that improvements in Sales were driven by improvements in Days of Inventory. Econometric and causal models show that while Sales and Days of Inventory are linked, **improvement in Sales drives improvement in Days of Inventory**. Therefore LargeCo's **initiatives to improve Days of Inventory cannot be shown as improving Sales**
3. Causal models show that **improvement in Fill Rate has a relationship with improvements in Sales**. However this finding was not borne out by applying the econometric model to LargeCo's data linking Fill Rate and Sales. The inability of the econometric model to verify the relationship is possible due to a combination of factors
    a. Aggregated data for Fill Rate was used for analysis. The aggregation was at the level of product line and used monthly data instead of weekly data

b. LargeCo maintains very high Fill Rates and the improvement achieved in the Fill Rates through the initiatives have shown very marginal changes given the high Fill Rates

LargeCo can try assessing the relationship between Fill Rates and Sales at a SKU level using weekly data to determine if data disaggregation helps detect a relationship. If a relationship between Fill Rates and Sales is still not detected then it can be concluded that LargeCo **may not obtain noticeable benefits in Sales by trying to improve its Fill Rates** because current Fill Rates are quite high

4. The econometric model did not find a relationship between Sales and Logistics Cost. However the causal model shows that contrary to the initial hypothesis, **Sales drives Logistics Cost.** Therefore LargeCo's **initiatives to improve Logistics Costs cannot be shown as improving Sales**
5. The econometric model found a relationship between Supply Chain Cycle Time and Sales. The causal model confirmed that improvements in Supply Chain Cycle Time have a relationship with improvements in Sales. The econometric model found this relationship occurring with a lag. The relationship was not found for all product lines. This could be attributed to the aggregation of data at the level of product line and use of monthly data instead of weekly data. Nevertheless using the analytical framework it can be concluded that LargeCo's **initiatives to reduce Supply Chain Cycle Time will have an effect on Sales but with a time lag**
6. Causal analysis showed that Time Between Production Runs is linked to Sales in much the same way as Supply Chain Cycle Time is linked to Sales. Econometric

model could not find a relationship between Time Between Production Runs and Sales. This could again be attributed to the aggregation of data. However based on the causal analysis it can be concluded that **initiatives to reduce Time Between Production Runs will have an effect on Sales but with a time lag**

5.2 Summary of Key Influencing Factors

The research work helped identify a few key factors that influence the inferences that can be drawn from the analytical framework. These are:

1. Deseasonalization of data is critical to remove the influence of seasonality on the econometric model. If seasonality cannot be removed then the effect of seasonality should be modeled using dummy variables so that the effect of seasonality can be distinctly identified
2. Standardization of data helps scale data so that the econometric model can detect relationships between variables
3. Removal of effects of discounts, marketing and promotions from data is critical to derive useful inferences from the econometric model. If the effects of these factors cannot be removed they should be modeled using instrumental variables to detect their effects
4. Causal models are required to support inferences derived from econometric models. Deriving inferences on relationships using econometric models alone cannot be justified because correlation detected by econometric models is not necessarily causality

5. Data aggregation has a significant effect on the quality of inferences that can be drawn from the analysis. The recommended level of data aggregation is at the SKU level and with weekly data

5.3 Caveats

Listed below are a few key caveats to consider while using the inferences drawn from the research work.

1. Instrumental variables and dummy variables have not been considered in verifying relationship between Sales and KPIs. Addition of instrumental and dummy variables helps draw more meaningful conclusions from econometric models
2. Aggregation of data significantly affects inferences that can be drawn from the analysis
3. Only a small volume of KPI and Sales data (about 24 to 32 months) was available for analysis. Disaggregation of data or data for a longer time horizon will influence the analysis
4. The link between initiatives to improve responsiveness and improvements in KPIs is assumed as given
5. Trend analysis of data did not reveal non-linear trends between Sales and any of the KPIs. However additional analysis may be required to verify this conclusion

5.4 Future Research

The analytical framework developed through this research work provides LargeCo and others a framework to pursue further research on measuring the value of responsiveness. Listed below are some potential areas for additional research.

1. LargeCo can consider identifying and adding relevant instrumental and dummy variables to the econometric model to help obtain additional inferences
2. The analytical framework can be used without changes to
    a. Assess relationship between Sales and KPIs tested in this research for other product lines
    b. Assess relationships between Sales and KPIs other than those tested in this research

References

Allen, M. P. (2004). Understanding Regression Analysis (1st ed.). New York: Springer Science & Business.

Anderson, D. R., Sweeney, D. J., Williams, T. A., Freeman, J., & Shoesmith, E. (2007). Statistics for Business and Economics (1st ed.). London, UK: Cengage Learning.

Beamon, B. M. (1999). Measuring supply chain performance. International Journal of Operations And Production Management, 19(3), 275-292.

Benton, W. C., & Maloni, M. (2005). The influence of power driven buyer/seller relationships on supply chain satisfaction. Journal of Operations Management, 23(1), 1-22.

Brook, R. J., & Arnold, C. G. (1985). Applied Regression Analysis and Experimental Design (1st ed.). New York: CRC Press.

Chan, F. T. S., & Qi, H. J. (2003). An innovative performance measurement method for supply chain management. Supply Chain Management, 8(3), 209-223.

de Groote, X. (1994). Flexibility and marketing/manufacturing coordination. International Journal of Production Economics, 36(2), 153-167.

Ejrnæs, M., & Kongsted, H. C. (2004, April 20). Notes on Instrumental Variable Estimation [WWW Document]. URL http://www.econ.ku.dk/metrics/Okonometri1_F04/ive.pdf (visited 2009, February 27).

Feinstein, A, F. (1996). Multivariable Analysis: An Introduction (1st ed.). New Haven, CT: Yale University Press.

Fisher, M. L. (1997). What Is the Right Supply Chain for Your Product? Harvard Business Review, 75(2), 105-117.

Garrett, T. A. (2003). Aggregated versus disaggregated data in regression analysis: implications for inference. Economics Letters, 81(1), 61-65.

Harvey, A. C., & Proietti, T. (2005). Readings in Unobserved Components Models (1st ed.). New York: Oxford University Press.

Hausman, W. H. (2003). Supply Chain Performance Metrics. In T. P. Harrison, H. L. Lee & J. J. Neale (Eds.), The Practice of Supply Chain Management: Where Theory and Application Converge (pp. 61-73). New York: Springer Science & Business.

Hoover, W. E., Eloranta, E., Huttunen, K., & Holmstrom, J. (2001). Managing the Demand-Supply Chain (2nd ed.). New York: John Wiley and Sons.

Kahane, L. H. (2007). Regression Basics (2nd ed.). Thousand Oaks, CA: Sage Publications.

Kennedy, P. (2003). A Guide to Econometrics (5th ed.). Cambridge, MA: The MIT Press.

Keppel, G., & Zedeck, S. (1989). Data Analysis for Research Designs (1st ed.). New York: Macmillan.

Levin, R. I., & Rubin, D. S. (1996). Statistics for Management (6th ed.). Englewood Cliffs, NJ: Prentice-Hall.

McConnell, C. R., Brue, S. L., & Campbell, R. R. (2004). Microeconomics (16th ed.). New York: McGraw-Hill.

Mooney, D. D., & Swift, R. J. (1999). A Course in Mathematical Modeling (1st ed.). New York: Cambridge University Press.

Narasimhan, R., & Jayaram, J. (1998). Causal Linkages in Supply Chain Management: An Exploratory Study of North American Manufacturing Firms. Decision Sciences, 29(3), 579-606.

Ostle, B., & Malone, L. C. (1988). Statistics in Research: Basic Concepts and Techniques for Research Workers (4th ed.). Ames, IA: Iowa State University Press.

Ostrom Jr., C. W. (1990). Time Series Analysis: Regression Techniques (2nd ed.). Thousand Oaks, CA: Sage Publications.

Ramdas, K., & Spekman, R. E. (2000). Chain or Shackles: Understanding What Drives Supply-Chain Performance. Interfaces, 30(4), 3-21.

Randall, T. R., Morgan, R. M., & Morton, A. R. (2003). Efficient versus Responsive Supply Chain Choice: An Empirical Examination of Influential Factors. The Journal of Product Innovation Management, 20(6), 430-443.

Schwab, D. P. (2005). Research Methods for Organizational Studies (2nd ed.). New York: Routledge.

Sims, C. A. (1974). Seasonality in Regression. Journal of the American Statistical Association, 69(347), 618-626.

Skinner, W. (1974). The focused factory. Harvard Business Review, 52(3), 113-121.

Sterman, J. D. (2000). Business Dynamics: Systems Thinking and Modeling for a Complex World (1st ed.). New York: McGraw-Hill.

Suits, D. B. (1957). Use of Dummy Variables in Regression Equations. Journal of the American Statistical Association, 52(280), 548-551.

Wang, G., Huang, S. H., & Dismukes, J. P. (2004). Product-driven supply chain selection using integrated multi-criteria decision-making methodology. International Journal of Production Economics, 91(1), 1-15.