Opportunistic Fresh Produce Commercialization under
Two-Market Disintegration

by

Hector M. Flores

A Thesis Presented in Partial Fulfillment
of the Requirements for the Degree
Master of Science

Approved July 2011 by the
Graduate Supervisory Committee:

Rene Villalobos, Chair
George Runger
Arnold Maltz

ARIZONA STATE UNIVERSITY

August 2011

ABSTRACT

This thesis develops a low-investment marketing strategy that allows low-to-mid level farmers extend their commercialization reach by strategically sending containers of fresh produce items to secondary markets that present temporary arbitrage opportunities. The methodology aims at identifying time windows of opportunity in which the price differential between two markets create an arbitrage opportunity for a transaction; a transaction involves buying a fresh produce item at a base market, and then shipping and selling it at secondary market price.

A decision-making tool is developed that gauges the individual arbitrage opportunities and determines the specific price differential (or threshold level) that is most beneficial to the farmer under particular market conditions. For this purpose, two approaches are developed; a pragmatic approach that uses historic price information of the products in order to find the optimal price differential that maximizes earnings, and a theoretical one, which optimizes an expected profit model of the shipments to identify this optimal threshold.

This thesis also develops risk management strategies that further reduce profit variability during a particular two-market transaction. In this case, financial engineering concepts are used to determine a shipment configuration strategy that minimizes the overall variability of the profits. For this, a Markowitz model is developed to determine the weight assignation of each component for a particular shipment.

i

Based on the results of the analysis, it is deemed possible to formulate a shipment policy that not only increases the farmer's commercialization reach, but also produces profitable operations. In general, the observed rates of return under a pragmatic and theoretical approach hovered between 0.072 and 0.616 within important two-market structures. Secondly, it is demonstrated that the level of return and risk can be manipulated by varying the strictness of the shipping policy to meet the overall objectives of the decision-maker. Finally, it was found that one can minimize the risk of a particular two-market transaction by strategically grouping the product shipments.

ACKNOWLEDGMENTS

TABLE OF CONTENTS

iv

LIST OF TABLES

Table   Page

LIST OF FIGURES

Figure   Page

1. PROBLEM DEFINITION

1.1   Introduction

One of the primary issues facing agricultural producers in today's market environment is the risk associated with their production. Due to the nature of the industry, production is dependent on a variety of external factors that are often outside of the farmers' immediate control, such as the risks associated with an under productive harvest, a catastrophic event, climate variability, etc. Furthermore, fresh produce items are often considered a perishable commodity, which allows the buyer to select from a broad variety of production sources giving the farmers little or no leverage when it comes to negotiating their prices. Consequently, the underlying situation is one in which the farmers assume most of the risk associated with production variability and in turn receive a reduced margin over the final profits.

To counteract their poor bargaining position, many farmers have sought to integrate vertically along the value chain of their products in order to curtail their direct competition, as well as increase their share of the total profit margins. Examples of this can be seen in several of today's global fresh produce industries. Such is the case of European farmers who have asserted themselves as global leaders by coming together through centralized, agricultural cooperatives. These cooperatives use 'logistic platforms' to engage in value-added practices that allow them to differentiate their products and gain greater ownership of their profit margins. In this manner, the farmers assume a greater role in their product's

added value, increase their bargaining leverage, and reduce the risk associated with price variability.

Figure 1-1 exemplifies a common integration often undertaken by farmers within the value chain of the product; from a traditional role to a more complex operational system. The traditional role of the farmer, as it shown in the diagram, is mainly associated with those activities related to production. This leaves the rest of the operations within the value chain, and most importantly its benefits, in the hands of its other members. On the other hand, as one can observe, a vertical integration allows the farmer to assume a greater role of the value-added and distribution activities. Similar to the integration observed within European markets, the implementation of a logistics platform allows the farmer to centralize and control the majority of the value-added activities of the product, in order to offer better service and get closer to the end consumer.
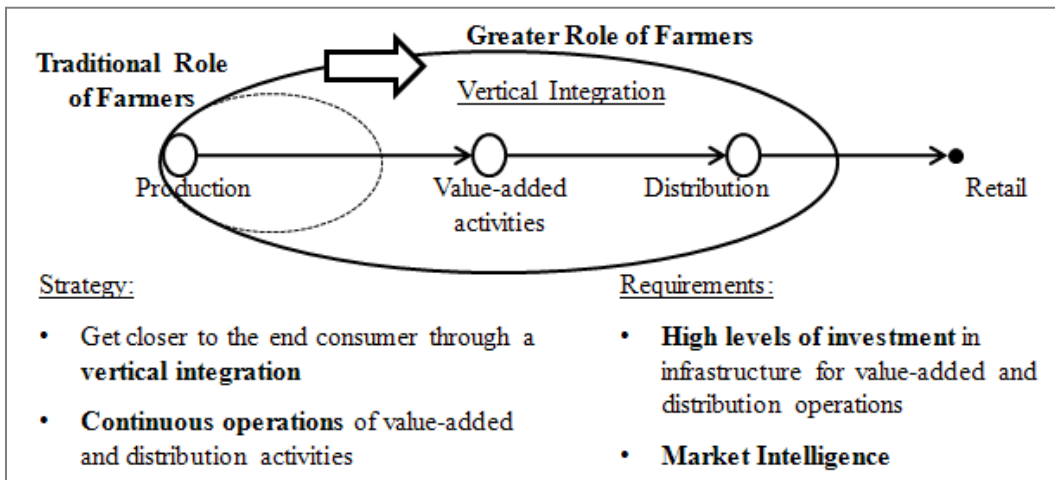
Figure 1-1 – Vertical Integration along the Product Value Chain

The vertical integration within the product value chain is a gradual process. In the case of European farmers, the initial market conditions allowed an

easier transition into more complex operations, in which the traditional structure of the farmers' operations was the commercialization of fresh produce items through basic cooperative auctions. Through these cooperatives, the independent farmer had an organized marketplace through which they could sell their products to wholesalers and retailers in simple auctions. The price of the product depended mainly on the conditions of supply and demand factors.

External factors to the European fresh produce industry slowly transformed the complexity of the fresh market environment. The consumer became more aware and concerned on the quality and standard of the food products they consume, specifically, fresh produce. This translated to higher quality standards for food retailers, their suppliers, and more importantly, farmers. In this case, the latter was left with the responsibility of updating its operational and organizational tools in order to improve the quality and reliability of the products and meet the demands of the end consumer. Among the investments required were infrastructure improvements, quality control programs, and an integration of value-added practices; in general, operations made possible by the implementation of centralized platforms.

The magnitudes of these investments were often too large for individual farmers to handle and thus many had to resort to alliances and partnerships. In the case of the European farmers, the formation of cooperatives became the best solution. Individual farmers formed alliances and created their own cooperatives, which they used as platforms to launch more complex operations. Through these platforms, the farmers were able to offer added value to the products, such as

repackaging, processing, maturation, etc. and have differentiated their products from the typical commodity product. Furthermore, they were able to coordinate logistic operations on a grander scale and more importantly, the farmer-owned cooperative structure gave them a greater amount of leverage to their negotiating position.

Today, private cooperative structures have allowed the European farmers to become primary players in the global fresh produce industry, as it provides a way for farmers to innovate and maintain high quality and service standards. They have also become role models for farmer organizations in other regions of the world in their hopes to develop sophisticated supply chains relative to those of the Europeans.

Nonetheless, the downside of such strategy is that low-to-mid level independent farmers are often unable to extend their commercialization reach due to their lack of capital for infrastructure investment. Thus, an important question to ask is if supply chain integration is the only way to circumvent the high levels of investment required to extend the reach of farmer.

This thesis argues that there exist alternative strategies that can still be economically beneficial for farmers without having to resort to high initial investment practices. As explained earlier, the fresh produce industry is often dependant on a multitude of external factors that add variability to the market. In the case of European farmers, their way to attack this problem is through the integration of their product's value chain. The proposed alternative is to consider the use of the underlying characteristics of the fresh produce markets the

advantage farmer. In this case, he/she can extend his/her primary role as a producer to become a speculator/investor and as a result profit from the price variability within fresh produce markets.

Before continuing, it is important to explain the general context in which this thesis is presented. The general concept is similar to the European case, but on a domestic level. The main focus is the development of alternative strategies in commercializing fresh produce without heavy investment requirements.

1.2   Problem Background: Commercialization of Mexican Farmers

In the Unites States, the fresh produce industry has not yet reached the level of maturity as the one observed in Europe. While the general tendency of the market place is towards more complex and dynamic structures, the domestic conditions still allow independent farmers to compete effectively. Nonetheless, the transition towards vertically integrated supply chains has been a model to follow for many farmer organizations, including those Mexican farmers aiming to extend their commercialization within the US.

Since the enactment of the North American Free Trade Agreement (NAFTA) in 1994, the commercial boundaries between the United States and Mexico liberated many of the commercial hurdles that had been present in earlier decades. This allowed the increase of agricultural exportations into the US, which led to a dramatic increase of the Mexican presence in the domestic fresh produce industry. All these developments have created a new playing field for all the parties involved in the industry.

Mexican farmers are not immune to the changes that are beginning to transform the US fresh produce industry. One clear example are the farmers from the state of Sinaloa, who have seen their profits reduced due to these changes and will most likely be obligated to restructure their US commercialization strategies in order to maintain competiveness within the industry. These farmers will have to transition from their traditional role, into more developed operational systems.

1.3   Traditional Role of Farmers from the State of Sinaloa

In previous studies led by a research group from Arizona State University, an analysis was performed on the situation of Mexican farmers from the state of Sinaloa (Villalobos et. al., 2007). For these farmers, the traditional distribution strategy is the commercialization of fresh produce at the US-Mexico international border of Nogales, AZ. Every winter season, the farmers take their products to this border and sell the products at Free-On-Board (FOB) price, which means that the price already accounts for transportation and the buyer gains full ownership of the product. At this point, brokerage and wholesaler companies buy the product based on their demand and make a profit from any value added to the products, which might include warehousing, repackaging, distribution, etc.

With this situation, both parties in the negotiation are benefitted. The revenues from such a simple operation are very lucrative for farmers from Sinaloa. However, this operation has a major disadvantage; the farmers are at the mercy of the variability that might occur in the US market. Furthermore, as the complexity of the fresh produce industry increases, so does the number of players

that play a role in the final sell of the product. The combination of all these factors leads to decreasing profit margins and lost economic opportunities.

After identifying this problem, the Mexican farmers have now entered the task of searching for ways to counteract their tough bargaining position; reminiscent of European farmers of the past decades. One of the recommendations from this study was the vertical integration of their value chain through the implementation of a 'logistics platform,' as a way to launch the commercialization operations of their products. This platform would serve as a strategically located hub closer to the final client, at which different logistics and value-added activities could be performed. In this manner, these farmers would be able to have greater ownership of their product, increase the profit margin over their products and increase the leverage in their negotiating endeavors.

But again, this scenario presents a problem; only those farmer entities with the sufficient investment capital have the ability to effectively compete and maintain innovative European-like practices within their supply chain structure. As a result, the limitation set by the investment requirements acts as an entrance barrier for those individual farmer entities incapable of generating the kind of necessary capital that would allow them to integrate vertically within their value chain. Thus, less financially-able entities have to find alternative ways to forego the capital requirements of such operations but still profit from the economic potential that is present in the produce industry.

Thus, it is important to assess and develop a methodology for those entities that may not be so inclined to integrate along their chain but still want to

profit from the opportunities present in the industry. The purpose of this thesis is to develop an alternative methodology that is less dependent on financial investment, but can still extend the commercialization reach of low-to-mid level farmers. The next section describes the market characteristics of the fresh produce industry that can be exploited for the development of an alternative marketing strategy.

1.4 Financial-Based Tools Application to Fresh Produce Markets

The main argument behind the methodology of this thesis is that the dynamic characteristics of the fresh produce industry create a prime environment of economic opportunity for those individuals that do not meet the minimum requirements to compete at full scale. As mentioned earlier, one potential strategy is to use the variability present within fresh produce markets as an advantage for the farmer. In this case, the idea is for the farmer to transition from a grower to a market speculator; role in which the farmer can concentrate his/her efforts in identifying ways to benefit from potential economic opportunities created by the fluctuation of the fresh produce markets.

There are several characteristics of the fresh produce industry that create the economic opportunities present within the two-market structure. Among these characteristics are the following:

- Integrated Markets
  - Variability at the supply source is indirectly transferred to the consumer market
  - Constant fluctuation of product prices at wholesale markets

- Homogenous Products

  - Under certain market scenarios, fresh produce can be categorized as undifferentiated, commodity products (without considering those that have received additional value-added activities)

  - Allows for easier ownership of acquired products

- Similar behavior characteristics as financial instruments

- Perishable characteristics of the products

  - Potential to add value to the products through transportation, warehousing, and distribution operations

  - Products cannot be stored, which opens market opportunities where supply is limited

The question then becomes how one uses the dynamic market characteristics of the fresh produce industry in order to develop a generalized method of operation that allows an entity to identify and take advantage of momentary economic arbitrage opportunities. The purpose is to have positive expectation of return on minimal levels of investment. Furthermore, the farmer would want to limit his/her risk exposure and only speculate on the market whenever the market conditions are favorable.

1.5 Problem Description

This thesis proposes an operational structure that allows the farmer to take advantage of the market price variability of fresh produce items; in his/her efforts to increase the commercialization reach of the products (Figure 1-2). In this case,

the farmer uses the price variability within two markets in order to identify particular time instances in which their price differential creates an opportunity for a transaction. This transaction is the process of moving a single product from the base to the secondary market under favorable conditions. If the transaction results in a positive profit, then that particular arbitrage opportunity is said to have been correctly identified and captured. If the transaction results in a negative profit, then the opportunity was incorrectly identified.
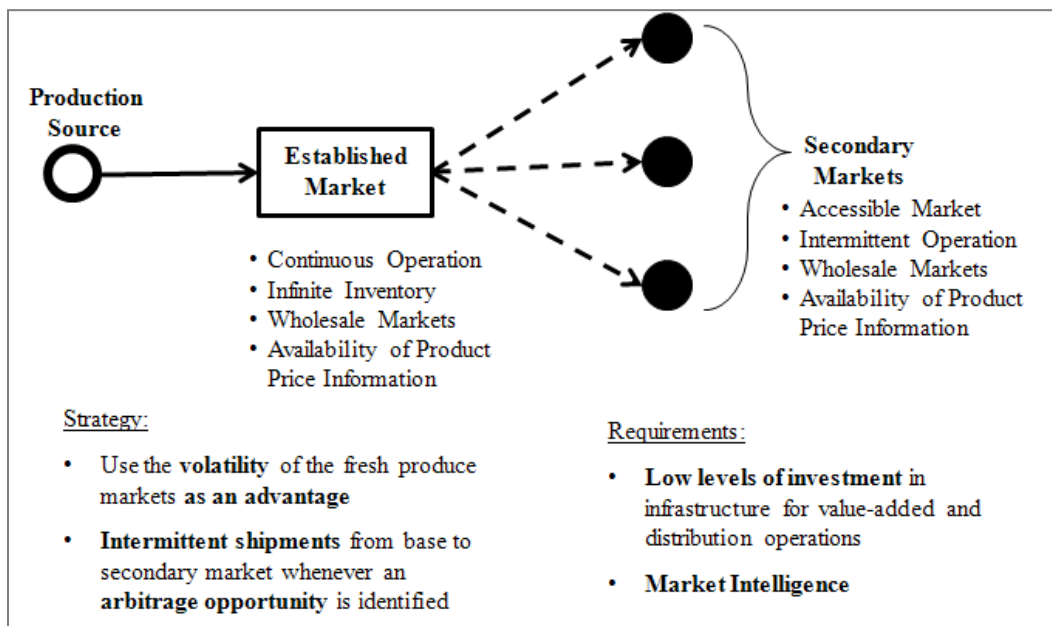
Figure 1-2 - Proposed Operational Structure

One of the main assumptions of this operational structure is that the base market is assumed to have continuous operations. On the other hand, operations within the secondary markets are performed intermittently, only when an arbitrage opportunity is identified. Furthermore, as it is discussed later in this study, the inventory at the base market is assumed infinite. Therefore, whenever an opportunity is identified within a particular secondary market, the product is

assumed to be on-hand and ready for shipment at the base market. Lastly, the operational structure considers wholesale market level prices for analyzing the daily fluctuations.

1.6 Study Objectives

The main premise of this thesis is to develop a methodology that correctly identifies and estimates the profitability of particular arbitrage opportunities within wholesale, two-market structures. These opportunities are in the form of favorable two-market differentials for a single product that optimize the long-run profitability of sending single container shipments. In this case, the first objective is to determine the specific shipment criteria from base to secondary market that optimizes the long – term expected profits of one-time shipments for a single product.

The second aspect of the operation is to develop risk reducing policies that limits the risk exposure of the shipment whenever an opportunity is identified. For this purpose, the market characteristics of each particular product are used in order to hedge the risk for a negative profit for any single fresh produce item. Financial engineering concepts are applied in order to determine an optimal shipment configuration that limits the overall risk of a one-time product shipment.

The development of these two policies should provide the farmer with important decision-making tools when operating within the proposed two-market structure. These tools should result in satisfactory levels of return for the operation, as well as risk exposure limitations of the decision-maker. The details

of these two main objectives are explained in-depth in later sections when detailing the methodology of this thesis.

1.7   Benefits of the Proposed Methodology

The overall objective of the proposed two-market structure is to allow the design of an operational strategy that not only increases the presence of the products within any given secondary market, but also creates positive revenues for the farmer. The investment requirements for such strategy are relatively small compared to those of a vertical integration. The methodology development and analysis is based on the price relationships of two-market systems for the purpose of identifying potential arbitrage opportunities.

Among the benefits of the proposed methodology is (1) the development of the commercialization expansion strategies based on reduced levels of investments, (2) the incorporation of statistical and financial engineering tools that offer a unique perspective of fresh produce market analysis in estimating the actual investment benefits of operating within a two-market system, and (3) a potential intermediate step in more complex operational objectives.

Overall, the methodology promotes the development of farmers looking for ways to effectively compete in fresh produce markets.

2. LITERATURE REVIEW

This chapter provides a general review of current literature on the topic of this thesis. The review is divided in the following sections that pertain accordingly to the proposed methodology. (1) First, the review presents the findings in current literature with regards to spatially integrated markets. The papers address the issue of price transmission and correlation among markets that aid in identifying potential opportunities of arbitrage. (2) The second set of papers in the in this review is indirectly related to decision-making analysis in the procurement of perishable items through two-market systems. These papers focus on ordering policies under uncertain demand.

2.1 Research Related to Spatial Market Integration

The transmission of prices of homogenous commodities within spatially separated locations is a concept that has received a high amount of attention in current literature. Padilla-Bernal, Thilmany, and Loureiro (2003) argue that in efficiently integrated market, the commodity price of one market should equal the price of the other (after adjusting for transaction cost) so that that the law of one price holds. Whenever this criterion is not upheld, an arbitrage opportunity could arise in the form of the price differential between the two markets. In this paper, the authors use parity bounds model (PBM) as a measurement of the level of integration of two markets for the same commodity product.

Baulch (1997) uses the parity bounds model mentioned above as a way to test for the integration of food markets. The author argues that conventional methods for analyzing co-movement of food prices rely solely on price data and

fail to capture the true transfer costs of the products. In his analysis, the author develops a model based on different trade regimes dependent on the market price differences and transfer costs, which dictate the possibility of trade. The author uses a market indicator as an extrapolation tool for the transfer cost. The model does not adjust for price lags of two markets, thus it uses price information of less frequency.

Econometric modeling is an alternative approach that has also been applied to two-market systems to analyze price relationships of commodity products. Several papers use co-integration testing as a way to analyze price relationships of commodities. According to Liang, Feuz, and Taylor (1997), co-integration is a statistical framework to test for short-run and long-run or steady-state equilibrium relationships among several non-stationary series; concept that was first introduced by Granger (1969). The concept was refined in Engle and Granger (1987). The formal definition co-integration of time series variables, as summarized by Liang, Feuz, and Taylor (1997), is as follows:

Time series x_{1t} and x_{2t} are said to be co-integrated of order d, b, where d$\geq$ b $\geq$ 0 written as

$x_{1t}, x_{2t} \sim$ CI (d,b) if

1. Both series are integrated of order d,

2. There exists a linear combination of these variables say $\alpha_1 \times x_{1t} + \alpha_2 \times$ x_{2t} , which is integrated of order (d-b).

Vector [α_1 , α_2] is said to be the co-integrating vector. In the situation where there is long-run relationship between two (or more) non-stationary variables (all

integrated of the same order), the deviations from this long-run path are stationary if the variables are co-integrated.

Furthermore, Liang, Feuz, and Taylor (1997) adapt four testing procedures suggested by Engle and Granger (1987) for co-integration, in which the null hypothesis for each test means "no co-integration" of the markets.

1. The Co-integration Regression Durbin Watson:

   $$y_t = \widehat{\alpha}\, x_t + c + \widehat{e}_t$$

   Test Statistic: $$DW = \frac{\sum_{t=2}^{T}(\widehat{e}_t - \widehat{e}_{t-1})^2}{\sum_{t=1}^{T}(\widehat{e}_t)^2}$$

   y_t and x_t are two price series. Reject if the DW is significantly different than zero.

2. Dickey Fuller Regression:

   $$\Delta\widehat{e}_t = -\phi\widehat{e}_{t-1} + \widehat{\varepsilon}_t$$

   where $\widehat{e}_t$ is previously defined and Δ is the first difference.

   Test Statistic : τ_ϕ (the t statistic for φ)

   Statistic tests whether the autoregressive parameter for the estimated residuals from the co-integrating regression (φ) is significantly different from one. If there is a unit root, then the two series are not co-integrated. Null hypothesis is rejected for values of φ which are significantly different from zero.

3. Restricted VAR:

   $$\Delta y_t = \hat{\beta}_1\widehat{e}_{t-1} + \varepsilon_{1t}\,, \qquad \Delta x_t = \hat{\beta}_2\widehat{e}_{t-1} + \hat{\gamma}\Delta y_t + \varepsilon_{2t}$$

Test Statistic: $\tau_{\hat\beta_1}^2 + \tau_{\hat\beta_2}^2$ (two sum of t-statistics for

$\hat\beta_1$ and $\hat\beta_2$)

This test involves the estimation of a vector error correction mechanism

for the co-integrating regression, which is based on the joint significance

of the error correction coefficients ($\hat\beta_1$ and $\hat\beta_2$). The null hypothesis is

rejected if the values for $\hat\beta_1$ and $\hat\beta_2$ are jointly different from zero.

4. Unrestricted VAR:

$$\Delta y_t = \hat\beta_1 y_{t-1} + \hat\beta_2 x_{t-1} + \hat c_1 + \varepsilon_{1t}$$

$$\Delta x_t = \hat\beta_3 y_{t-1} + \hat\beta_4 x_{t-1} + \gamma \Delta y_t + \hat c_2 + \varepsilon_{2t}$$

Test statistic: $2[F_1 + F_2]$

where F_1 is the F statistic for testing $\hat\beta_1$ and $\hat\beta_2$ both equal to zero in

first equation, and F_2 is the F statistic for testing $\hat\beta_3$ and $\hat\beta_4$ both equal to

zero in the latter equation.

This last test procedure utilizes an autoregression vector which is

not constrained on satisfying the co-integration constraints if parameters

$\hat\beta_1$ and $\hat\beta_2$ and $\hat\beta_3$ and $\hat\beta_4$ from both equations are jointly, significantly

different from zero. Failure to reject the null hypothesis indicates the

lack of statistically significant relationship between current changes and

past values of the economic variables, which also implies a general

failure of co-integration between variables.

Blank and Schimiesing (1988) apply causality and path analysis

(predecessor to co-integration analysis) to test for spatial relationships between

two markets. The concept of causality is derived from Granger (1969), in which the author uses testing criteria to determine the direction of causality of two related variables. Granger's causality test does not measure the relative strength of the relationships; neither can it distinguish between relationships that are real and those that are spurious (Ziemer and Collin 1984).

On the other hand, path analysis, developed a while back by Wright (1921), measures the relative strengths of relationships in a model developed from theory or some unique earlier hypothesis to be tested. In their study, Blank and Schimiesing (1988) combines both methods (causality and path analysis) to provide additional capabilities and counter act the limiting assumptions of the other in what the authors call the Causality and Path (CP) Method.

McNew (1996) argues that spatial price models should consider non-linear price responses, as opposed to traditional linear models. In the paper's empirical study (corn prices along the Mississippi River), the no-arbitrage condition suggests a piecewise linear relationship model between spatial prices, in which the slope of the functions serves as a binary indicator of integration between two locations. Serra et al. (2006) apply non-parametric techniques considering the non-linear nature of spatial price behavior. It assumes that there exist different trade regimes, and depending on the regime at any point in time, a particular dynamic pricing behavior will occur. Furthermore, the author argues that all costs related to spatial arbitrage and trade (transportation and freight charges, risk premium, refrigeration costs, etc.) may result in non-linear price adjustments.

It is important to note that although it is implicit that the methodology developed in the literature may be generalized and applied to almost any type of market, the specific effect of a product's perishable characteristics in market integration is a topic that has not yet been addressed. This effect is often generalized under the transaction cost of the products (Serra et al. 2006).

2.2   Studies Related to Two-Market Ordering Policies

Another aspect of the operations within the proposed commercialization strategy is the ordering policy that the farmer would have to implement. In this case, the farmer would design a system in which he would be able to effectively supply the demands of two markets; one being under arbitrage conditions. The farmer would need to have sufficient inventory in the base market to meet the intermittent spikes in demand. Nonetheless, his choices are limited once an order has been made due the perishable characteristics of fresh products. Once the shelf life for the product has been exceeded, its value would decrease until it reaches a zero value.

The single-period situation described above is similar to that of the newsvendor problem; a popular topic in literature. In this case, a decision has to be made on the optimal inventory policy under probabilistic demand and limited shelf life of the product. Khouja (1999) reviews different extensions to the news-vendor problem through a variety of topics. Among the extensions explored by this review pertaining to the topic of this thesis are the following: different objectives and utility functions, random yields, and constrained multi-product.

18

One of the topics that will be addressed by the proposed methodology is the perception of risk associated with the decision-making entity in the operations. Eeckhoudt, Gollier, and Schlesinger (1995) examine the effects of risk and risk aversion under single-period circumstances. In this paper, the authors attempt to consider the effects of two types of increases in risk in the optimal newspaper order: (i) the addition of an independent risk to the newsboy's background wealth (other risk besides demand variability), and (ii) an increase in the riskiness of newspaper demand. Choi, Li, and Yan (2008) address the risk associated with the newsvendor problem through minimum-variance approach, in which they explore different order policies under the three possible decision-making attitudes (seeker, neutral, and averse).

Other papers related to two-market ordering policies in the subject include Rudi, Kapur, and Pyke F. (2001), Lin, Dan (2007), and Burnetas and Smith (2000). The approach of these papers is related to inventory and ordering policies for the purpose of creating pricing strategies and increasing business value under a two-location situation. Although, the studies do not directly relate fully to the topic of this study, they do provide possible alternate extensions.

2.3 Thesis Contribution to Current Literature

The concepts addressed in this study have been studied extensively in literature. The application of co-integration testing to analyze the behavior of commodity markets is a subject that has been given much attention. The main purpose of the application has been to assess the level of integration of two-location trade of commodity products. Among the common approaches that have

been used in this type of assessment is an econometric-based analysis to capture market correlations and propensity for trade.

Other applications for market integration might be a possible extension to current research in the subject. The application of co-integration analysis to commodity trade markets from a financial standpoint. From what has been found in current literature, there has been little attention given to assess the effect of the level of market integration on the opportunity of arbitrage within commodity markets, especially with perishable items such as fresh produce. Furthermore, stemming from the purpose of this study, an application of market integration structured around financial concepts has not been done.

Studies related to inventory and ordering policies under two-market system is a complex subject that has been given relatively high amount of attention in the past. Based on the demand structure of the study (uncertain and perishable), literature regarding single-stage inventory, similar to a newsvendor problem, was reviewed in order to get a sense of current advances in the area. Applications under a risk prone environment are a common theme for single-stage inventory policies. However, literature is relatively limited when applied to portfolio mean-value theory.

The present study will attempt to link several aspects of a commercialization expansion that would limit the level of investment requirements. This paper will expand current advances in the field by combining concepts of co-integration for commodity market assessments through the use of financial engineering tools. Furthermore, this study will aim at address inventory

policies that are based not on demand, but on the probabilistic nature of the

arbitrage opportunities present within the two-market structures.

3. PROBLEM FORMULATION

3.1   Introduction

This thesis develops a commercialization strategy based on historical fresh produce market prices that aims to expand the commercialization reach of region-based farmers.   The proposed strategy involves reaching secondary markets through intelligence based operations that require minimum level of investments. The overall objective of the operations is to maintain an attractive balance between the levels of potential returns and associated risks.

The basic premise of this thesis is the development of a methodology that permits an established farmer with basic local operations to expand his/her commercialization reach into secondary markets, by way of financial engineering tools and statistical analysis.   Specifically, the purpose is to provide the farmer with the necessary decision-making tools that will help him/her identify and take advantage of momentary arbitrage opportunities for product placement in secondary markets.   Finally, this thesis will address additional aspects of the operation, such as risk limitation policies, that allow effective and profitable operations.

As explained in chapter 1, this study will seek to answer the following set of questions:

- Under what market conditions should the farmer engage in trade within the two-market system?

- What combination of products would optimize the balance between a potential return and a loss?

Each question addresses different aspects of the operation that need to be addressed throughout the planning process of the commercialization expansion. By way of this thesis, a methodology is proposed for each of these operational aspects.

3.2   Base Operations

In chapter 1 (Figure 1-2), a brief description of a potential structure of operation is presented that allows the farmer to engage in risk controlled practices within secondary markets.   This operation involves the price relationships between two consumer markets that are geographically distant. Specifically, this thesis addresses the fluctuations in the market price differentials caused by price transmission inefficiencies and geographical separation. It is believed that the fluctuation in the price differentials can be used as a starting block for a commercialization expansion that is not only effective, but also profitable.

The basic commercialization operation addressed is based on the price relationships that exist between two consumer markets. The two-market structure consists of a base market that is supplied on a continuous basis, and a secondary one which acts as complementary to the first. The secondary market is supplied intermittently depending on the availability of the arbitrage opportunities created by the fluctuations in the price differentials. The main objective is to develop a decision-making policy aimed at identifying and capturing these opportunities.

The main assumption of this study is that long-run average price differentials between markets will not be large enough to allow continuous trade operations over extended periods of time. In this case, trade is limited to the

availability of scarce arbitrage opportunities. For purposes of this thesis, an arbitrage opportunity is defined as the economic opportunity present when the difference between the selling price at a base and secondary markets is large enough to allow a profitable transaction. The transaction involves the acquisition of the product at the base market and its transportation to the secondary market.

One must note that there are obvious risks involved with this type of operation, such a wrongly identifying an opportunity or actually losing money on a particular transaction. Thus, one should design a methodology that correctly monitors and assesses these opportunities; a difficult task given the complexity of the markets and the non-stationary characteristics of the price movements. In order to accomplish this, careful consideration must be given to the assessment of the magnitudes and durations of these price differentials

Among the decision variables of the operation is the timing of product shipments and the configuration of potential commercialization baskets. The following sections will develop a methodology will address each of the different aspects of this basic operations. In general, the proposed strategy will be based on the operation dynamics addressed in this section.

3.3   Methodology Development

The commercialization consists of two operational strategies. The first will consider a basic approach of distribution. The farmer will sell his products directly from his production source to a base market. These operations will be assumed to be continuous enough to have available products for commercialization in other markets. The second operational strategy will be to

engage in two-market transactions whenever the price differential provides a potential arbitrage opportunity. This second type of operation will be the basis for the methodology developed in this thesis.

In this section, the proposed methodology of this thesis is presented. This methodology includes the general assumptions that are made with regards to the markets and products, the development of a shipment strategy that meet the operational objectives, both on short and long-run terms, and the development of a product combination strategy that limits the risk exposure of the shipments.

3.3.1 General Assumptions

Given the complexity of the fresh produce market environment, some assumptions facilitate the development of the methodology. This allows the application of financial tools and statistical analysis to fresh produce markets. Among the assumptions considered are the following:

- High transaction liquidity (Easily tradable and acquirable products)

- Easily accessible markets

- Price characteristics similar to that of financial instruments

- Infinite product availability at the base market

Furthermore, it is assumed that the demand of the target market will always be enough to cover the size the full load of the shipment. Therefore, the amount and frequency of the shipments will only be dependent on the price conditions of the two-market system and associated transaction costs.

Finally, one assumes that the decision-maker will have access to sufficient price information that will allow daily monitoring of price fluctuations at both the

base and secondary markets during a prolonged period of time. This permits the analysis of long-term behavior and trends of the price differentials. The methodology developed through this thesis was not tested under scenarios with limited price accessibility.

3.3.2 General Methodology of the Study

The scope of this study is defined by the decisions faced when operating in the proposed two-market structure operations (Figure 3-1). The decision-maker must determine the optimal market conditions (price differential) within the two-market system that maximizes his/her profits per product over a defined period of time. Along the same lines, he must also determine the combination of products that would not only improve the overall returns of his operations, but also reduce risk exposure for a collective portfolio of products. Each of these aspects is addressed through this thesis.

- Under what market conditions should the farmer engage in trade within the two-market system?

Ordering Policy
- Determine shipment criteria for two-market trade that maximizes the potential for a return

- What combination of products would optimize the balance between a potential return and a loss?

Product Basket
- Define level of profitability and risk per commercialized product
- Determine optimal product mix for two-market trade

Figure 3-1 - General Methodology of the Study

The following sections describe each phase of the operation in further detail. These sections are divided in the following manner: Section 3.4 addresses the shipment policy of the operation and Section 3.6 develops the tools for constructing adequate product baskets.

3.4 Shipment Policy for Individual Products

The first step in the methodology is to develop an individual shipment policy for each commercialized product within the two-market system. This shipment policy considers current and historic market conditions to determine when the farmer should engage in a two-market transaction. Based on the historic prices and market conditions, a shipment policy is developed that attempts to maximize the expected, long-term profits over a defined period of time. The

shipment policy also incorporates additional components, such as the risk exposure per strategy and the frequency of the market opportunities in order to make more informed decisions.

The second aspect of the shipment policy is the development of a decision-making tool for estimating short-term probabilities of a positive or negative profit given particular market conditions. This tool uses the probabilistic nature of the opportunities to estimate the probability of a gain or loss given a particular price differential and/or transportation lead time. It assists in the decision-making process when faced with immediate market conditions and may be used for short-term projections of profit.

3.4.1 Potential Decision Outcomes

In general, the decision to commit to a transaction is made based on future expectations of the selling price. This expectation depends on the conditions of the two-market structure and the relative likelihood that an opportunity will be present and maintained over an estimated transaction period. These decisions have three possible outcomes that may result in a potential profit or loss. The following are the possible scenarios:

1. Price differential is not greater than a reference price or threshold value, which under current conditions does not trigger a transaction since the inventory at hand, if any, can be sold at existing prices at base market.

    a. No gains or losses from the secondary market

2. Price difference is above a threshold value that triggers a transaction

     a. Price difference remains or increases during transportation lead time, thus

     b. A positive return on the investment is expected.

3. Price difference under current conditions triggers a transaction

     a. Price difference falls below the trigger threshold during transportation lead time, thus

     b. A negative return on the investment is expected.

All three scenarios are based on the value of the "reference" price differential, or threshold, which is an important aspect of the operation. Furthermore, determining its optimal value is one of the main objectives of this thesis. The following sections address this problem by incorporating these basic scenarios into the development of an adequate shipment policy

3.4.2 Distribution Fit for Market Price Differentials

The first step in developing an optimal shipment strategy is to design a way through which the decision-maker can summarize the behavior and tendencies of the price differentials. For this purpose, it was determined that the best way to achieve this was through statistical distribution fits on the historic observations, which would mean certain assumptions with regards to the price differentials. In this case, it is assumed that the movement of the prices (and differentials) at both the base and the secondary markets are randomly distributed.

The statistical fits on the distribution of the histogram observations is made according to a chi-square goodness of fit test. For application purposes, traditional software programs, such as Matlab or StatFit, are capable of generating

29

the results of these tests and may be used in the decision analysis process when selecting a distribution to fit. Although, a good statistical distribution is important, a perfect selection is not crucial for an effective application of the methodology. The next few sections explain the use of the distributional fits on the price differentials for the development of the methodology.

3.4.3 Determining Long-Term Shipment Policy

The second part of the methodology is to use the information of the market price differentials in order to optimize expected long-term profits. It is hypothesized that a specific threshold value, or price differential (K), will result in a long-term, maximum profit and/or rate of returns. For the purpose of identifying the optimal value of K, two approaches are used; (1) a pragmatic approach which only considers the opportunities within the two-market structure in an iterative manner, and (2) a theoretical one, which is based on the theoretic value of the optimal price differential. Both approaches are implemented and considered in the development of the shipment strategy. Additional components of the evaluation process, such as the calculation of the shipment frequency and the value-at-risk analysis, are explained in APPENDIX A.

The objective of the pragmatic process is to determine the specific long-term condition, or particular price differential (K), that maximize the expected rates of return or profits by iteratively applying different K values. This value, referred to as the threshold level, represents the minimum price differential that triggers a shipment. Once the price differential crosses this value, the decision is made to send a single, truckload shipment of the product in question.

30

An iterative decision process flow is developed in order to pragmatically find the optimal K value (Figure 3-2). Starting from a threshold level of zero, the market price differential, minus transaction cost, is assessed at the initial value, t_0. If this differential at this time is higher than a defined threshold, then the decision to ship is made. The actual profit observed from the shipment is the lagged price differential, or the product price when it finally reaches the secondary market; accounting for the transportation time.

The process iterates for every instance of time in the data set. The price differentials are collected and analyzed at the end of each data set. Once the process has passed through the set, a new value of K is assessed. At the end of each iteration, the threshold value is increased by a user-defined stepsize. The threshold starts at zero and ends at some value at which the total number of observations over the defined period of time drops below 20 observations (enough to create a relatively reliable distributional fit). After a full iteration has been performed, the profits (actual) over the defined period of time are collected. A histogram is created under each threshold value in order to analyze the frequency distribution of the observations.
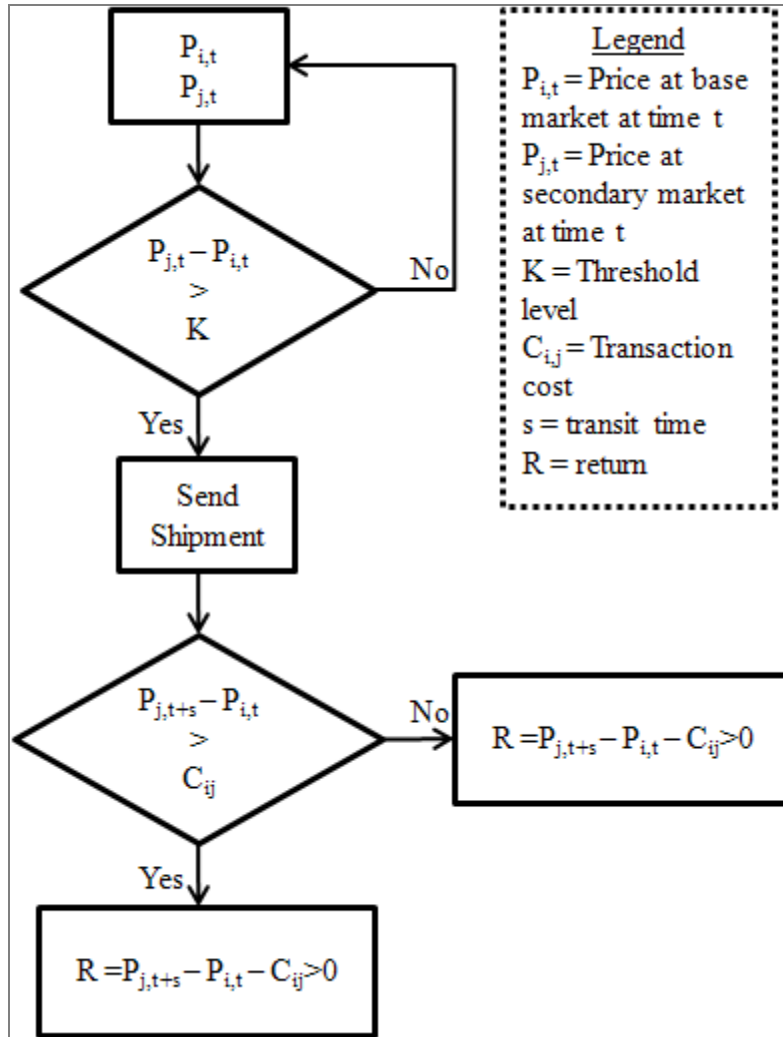
Figure 3-2 – Decision Process Flow

In order to determine the expected profit per threshold value, as well as the standard deviation of the observations, a statistical distribution is fitted on each histogram created. Since the objective is to directly compare among the different values of the threshold, one needs a statistical distribution that adequately explains the behavior of each histogram. Furthermore, since the operation also considers the possibility of potential losses, a statistical distribution that can handle negative values is needed. As it was mentioned before, a perfect selection

32

for this part is not crucial for an effective implementation of the methodology; however, some statistical knowledge on the part of the user is needed in order to choose an adequate distribution fit.

Once a statistical distribution is chosen for the collection of histograms, the parameters (mean and standard deviation) of each fit are used to determine the optimal value of K. However, in order to determine the value that optimizes the long-term expected profit, one must also consider additional components of the operation, such as the number of times an opportunity is presented for each value of K (frequency) and its associated risk. The estimation of these additional components is addressed in the next section.

The second approach in determining an optimal value of K is to develop an optimization model in a more theoretical manner. This approach is based on a theoretic value that would optimize the total profits of the operation. The model contains the different components of the shipment policy, such as the transaction cost between markets and the acquisition cost of the products. Furthermore, the theoretic optimal value is based on the distribution of the price differentials of the two-market structure. The following parameters form a part in the optimization model that is developed:

$P_{i,t}$ = Product price per pound at market i at time t

C_{ij} = Transaction cost between market i to market j

s = Transaction time from market i to market j

The first two parameters account for the product prices at different markets and the transaction cost of taking the product from the base to the

secondary market, respectively. The transaction cost is assumed to be fixed throughout the operational period for each two-market structure. The transaction time is depicted by variable s, and it accounts for the time it takes to transport the product between markets.

All these parameters are combined to form the random variables in the expected profit model. The ultimate purpose is to develop a function in terms of the price differential with and without lag. This allows the representation of the profits in terms of the statistical behavior of the price differentials. The following are the random variables that form part of the price differential function:

$$x = P_{i,t} - P_{j,t} \qquad\qquad \textbf{Eqn. 3-1}$$

$$y = P_{i,t+s} - P_{j,t} \qquad\qquad \textbf{Eqn. 3-2}$$

Variable y depicts the price differential between markets at present time; the differential with no lag in time. On the other hand, the x variable depicts the lagged price differential, considering transaction time. The final function of the differentials is represented by both of these variables. One must note that these variables are correlated if the actual profits captured (lagged differential) are dependant or conditioned on the current price differential that dictates the decision to ship. A scatter plot and correlation analysis is performed for this combination of random variables based on empirical data.

The expected profit model in terms of these variables is represented by the following:

34

Maximize:

$$E[P(k)] = \int_K^\infty \int_{-\infty}^\infty \left(P_{j,t+s} - P_{i,t} - C_{ij}\right) * f(x,y)\,dy\,dx \qquad \text{Eqn. 3-3}$$

$$= \int_K^\infty \int_{-\infty}^\infty \left(y - C_{ij}\right) * f(x,y)\,dy\,dx \qquad \text{Eqn. 3-4}$$

From this model, one can observe that the final profit of the shipments is just the difference between the product price at the secondary market, j, at time t + s and the price at base market, i, at time t. This difference also accounts for the transaction cost, C_{ij}. This value is then multiplied by the joint probability distribution function of the differentials. The integrals on this joint pdf state that a shipment is triggered based on current conditions (K →infinity over x variable) and that the opportunity is seized, regardless if this is a positive or negative profit once it arrives at the secondary market (- → + infinity over the y variable). Finally, the market price differential in the model is substituted by the y variable, since these two are equivalent.

In this case, the important component in the model is the non-lagged price differential that triggers a shipment, which is ultimately the decision factor. This value dictates the expected profit model:

Decision Factor:

K = non-lagged price differential

In order to find the optimal value of the K, the expected profit model is differentiated with respect to this decision factor. Then, the resulting equation is set to zero in order to find optimality. The following represents the differentiated

35

version of the previous model. With some model manipulation, the finalized

decision-making tool is depicted in terms of the expected value of the y variable,

or the lagged differential, conditioned on the value of *x*, in this case, the desired

threshold level, *K*.

$$\frac{\partial\, E[P(K)]}{\partial K} = -\int_{-\infty}^{\infty} y * \frac{f_{XY}(K,y)}{f_X(K)}\, dy + C_{ij} \int_{-\infty}^{\infty} \frac{f_{XY}(K,y)}{f_X(K)}\, dy \qquad \textbf{Eqn. 3-5}$$

$$= -E[y \mid x = K] + C_{ij} = 0 \qquad \textbf{Eqn. 3-6}$$

One should note that the conditional expectation on the non-lagged

differential in reality would be represented by the inequality x > K. This

complicates the evaluation process of the optimal threshold. For simplification

purpose, it is assumed that the threshold obtained through the optimization model

is only an approximation; close enough to the actual optimal value.

One could represent the expected value of the x and y random variables in

terms of a probability distribution function that fits the behavior of the market

conditions. For this purpose, a bivariate distribution is assumed to represent this

joint probability function in order to attain an estimate for this expected value.

Based on the resultant model, the value of the threshold is solved for and used as

the main decision-making criteria. One should keep in mind that the objective of

this theoretical model is to optimize the long-term profits of the operation. In the

case in which the objective is to optimize other aspects of the operation, then one

could restructure the optimization model and solve for the threshold value.

3.4.4 Short-Term Shipment Policies

The methodology described in section 3.4.3 identifies an optimal threshold level applicable for long-run operations. This allows the decision maker to develop long-term policies within the two-market structure. However, the decision-maker may also want to know about short term probabilities of a profit based on current market conditions. This section develops a methodology for estimating the probabilities of a negative or positive profit to perform just this.

Overall, positive or negative outcomes at any given time are dependent on the probability that the market price differential triggers a shipment and that the market price differential lasts past the transportation lead time (in this thesis, referred to as duration of an opportunity). For purposes of this thesis, the duration of the positive differentials are also assumed to be randomly distributed, as it was previously done for the market price differentials. As such, one is able to fit a statistical distribution on the histogram observations.

Given the characteristics of the opportunities, the expected probabilities are composed of a two-dimensional interaction between the two-market price differentials and the durations of positive differentials. Figure 3-3 is an example of a contour plot of conditional probabilities as a function of both the probability that the two-market price differential exceeds the threshold and that the duration of the opportunity lasts longer than the lead time period. The user-defined factor is the threshold value that triggers a shipment, or the current market conditions. The probability function of both the duration and the price differentials is used to generate these conditional probabilities based on different values of the threshold.
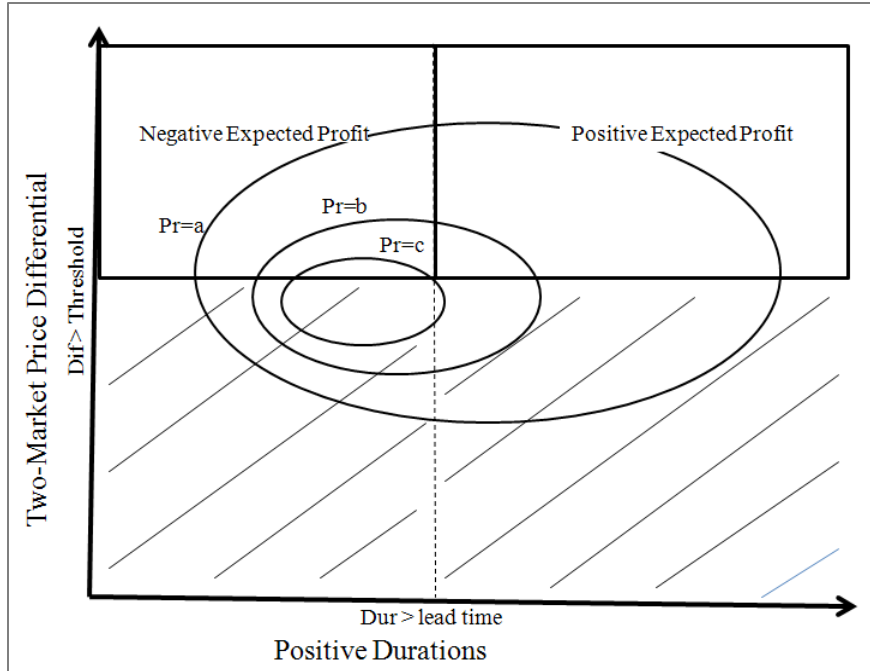
Figure 3-3 – Two-Market Structure Conditional Probability

The situation presented by the figure above can be thought as a two dimensional problem; one depending on the probability that an opportunity is present and another, on the probability of a duration given a particular threshold. In this case, one determines the probability of a certain time duration conditioned on the magnitude of an observed price difference. Similar to the methodology developed in the previous section, historic information of price movement over the transaction period are used to generate a representational contour plot of the conditional probabilities.

A mathematical representation of these probabilities is developed. The following parameters are used in these probability representations:

D_{ijt} = Price difference between market i and market j at time t

s = Transaction time from market i to market j

38

d_{ij} = Duration of positive difference between market i and market j

The first parameter represents the price differential within the two-market structure at any given time t, while the s parameter details the transaction time of moving the product from the base to the secondary market. The last parameter details the duration (in days) of a positive price differential between the two markets.

As it has been mentioned before, the probability of a positive expected profit is based on the probability that the duration outlasts the transaction time conditioned on a particular market price differential. One must note that the distribution of the price differentials would most likely be assumed continuous, and thus, a probability value at a single point is zero. In this case, the conditional probabilities are based on intervals of the K value. The K value is dictated by the market conditions. The following is a representation of a positive profit probability.

$$P[Profit_{positive}] = P(d_{ij} > s | D_{ij,t} = K)P(D_{ij,t} = K) \qquad \textbf{Eqn. 3-7}$$

On the other hand, a loss on a particular shipment would happen in the case in which the decision maker sends a product shipment based on current market conditions and the opportunity at the secondary market ceases to exist when the product arrives to its destination. In this case, a loss is conditionally dependent on the probabilities that the price differential is above a certain threshold level (trigger shipment) and that the duration of the opportunity is actually less than the transportation lead time.

$$P[Profit_{negative}] = P(\overline{d_{ij}} > s | D_{ij,t} = K)P(D_{ij,t} = K)) \qquad \textbf{Eqn. 3-8}$$

Based on the results of the application of this methodology, then one is able to estimate the probability for a profit or loss at any time given present market conditions. One could even think of repeating the same process for several periods and form binomial lattice structures. Based on these binomial lattice structures, one is able to form long term probabilistic projections of the profits and losses in the two-market operations. However, given time constraints, this application is presented as potential topic for future research.

3.4.5 Conclusion

Once a threshold limit has been determined; one which maximizes the expected returns for a particular product, the same process is repeated for a defined collection of fresh products. Thus, each product will have its own critical threshold that maximizes the returns over a determined period of time. Also, the additional components of their profitability are used to determine the product configurations of the shipments. The next few sections detail the application of these concepts.

3.5 Inventory Policy for Commercialization Products

An inventory policy that guarantees product availability whenever an opportunity is present is an important aspect of the operation. The policy must link the probabilistic nature of the price differentials and respective arbitrage opportunities to actual levels of inventory. This involves the development of an inventory model that would most likely be in terms of the expected costs and/or profits of the operations. Among the costs that should be considered in the

40

development of a potential inventory policy model are those of carrying inventory, the devaluation of the perishable products as a function of time, and the cost of not capitalizing on an opportunity.

The development of this model is within the scope of this study. However, the development of an inventory policy that depends not only on the probabilistic nature of the market price opportunities, but also on the intermittency of the arbitrage opportunities, requires a more in-depth analysis and focus that surpassed the time constraints of this thesis. Nonetheless, it is promising topic for future research (section 6.3.5).

For the purposes of this thesis, one assumes that there exists enough product availability to send one-time product shipments from the base to the secondary market. In this case, one does not need to constraint the one-time shipment quantities by the availability of the products.

3.6 Configuration of Commercialization Baskets

Once a shipment policy has been selected for an individual product, the next step is to develop a configuration strategy that uses the individual price characteristics of the products to limit the risk exposure of any particular one-time shipment. Specifically, an optimal configuration is created based on the variability and covariance of the products' market prices, in such a way that the potential for a loss on a shipment is minimized. As it was mentioned in Chapter 1, the behavior of the profits from the operational strategy is assumed similar to that of instruments, in common financial markets; estimated rates of return, standard deviations, and their respective covariance are used.

APPENDIX B explains the background of the mean-value portfolio theory and is relation to the methodology developed in this thesis. The following section details the process of estimating the necessary components to set-up the Markowitz model. Finally, it explains the method used to solve the problem and find a solution.

3.6.1 Estimating Rates of Return for Individual Products

For the application of the Mean-Variance portfolio theory and Markowitz problem approach, it is necessary to estimate the basic components of the product market price characteristics. These components are in the form of the estimated profits, associated with each particular shipment strategy. The estimation of these components facilitates the application of these concepts.

The first step is to translate the profits for each shipment policy (per product) into financial-based terms, such as returns and rates of return. The following details how the two-market price differentials are translated into these two terms:

Return: $\quad \dfrac{amount\ received}{amount\ invested} = \dfrac{P_{j,t+s}}{P_{i,t} + C_{ij}}$ **Eqn. 3-9**

Rate of Return: $\quad \dfrac{amount\ received - amount\ invested}{amount\ invested}$ **Eqn. 3-10**

Where:

S = transaction time from market i to j

$P_{j,t+s}$ = the price at the secondary market at time t +s

$P_{i,t}$ = the price at the base market at time t

C_{12} = the transaction cost to move the product from market i to market j.

42

These values are collected for each time a shipment is made to the secondary market under a particular shipment strategy. Secondly, once the values of return and rate of return have been collected for every shipment, the next step is to estimate the overall average rate of return and its standard deviation. To do this, one could assume that the rates of return follow a lognormal distribution given that common financial instruments tend to behave in this manner. However, in order to add accuracy to the estimates, individual theoretical statistical distributions are fitted on the histogram observations of the rates of return using a Chi-Square Goodness-of-Fit Test. The parameters of the best fit are used to estimate the average rate of return and standard deviation.

Finally, it is argued that the expected long-run rate of return for any shipment composed of n products the following:

Rate of return (Portfolio): $r_p = w_1 r_1 + w_2 r_2 + \cdots + w_n r_n$ **Eqn. 3-11**

Where:

r_p = average rate of return of the collection of products

$w_1, w_2, \ldots, w_n$ = weight allocation for product i

$r_1, r_2, \ldots, r_n$ = average rate of return for product i

Furthermore, the expected overall standard deviation of the rates of return for an n product shipment is as follows:

$$\sigma_p = \sum_{i,j=1}^{n} w_i w_j \sigma_{ij}$$ **Eqn. 3-12**

Where:

σ_p = overall standard deviation of the collection of products

$w_1, w_2, \ldots, w_n$ = weight allocation for product i

σ_{ij} = covariance of the return of product i with product j

The objective at this point is to develop a configuration policy of this collection of products that limits the profit variability of any one-time shipment.

3.6.2   Markowitz Model Application

The first step in determining the optimal product configuration is to set-up the Markowitz model. For this, the parameters in the model are estimated and used. These parameters are:

$\overline{r_p}$=Average rate of return of the collection of products in portfolio

$\overline{r_i}$=Average rate of return per product i

σ_{ij}=Covariance between product i and product j

The decision maker must decide on the weight amount assigned to each product on a one-time shipment. In this case, the allocation is in terms of the portion of the total shipment invested in any one particular product. This weight assignation thus becomes the decision variable in the objective function. This assignation is described by the following:

Decision Variable:

W_i= Allocation weight per product i

The objective function described by the Markowitz model is based on the weight assignation given to each product in the portfolio (or shipment) and the overall variance created by the variability of the individual products. The objective then becomes on minimizing the overall variance of the portfolio by strategically arranging the weights assigned to each individual products. In this

case, σ_{ij} represents the covariance between the price differentials in the product basket. The optimization model is represented by the following:

Minimize:

$$\frac{1}{2} \sum_{i,j=1}^{n} W_i W_j \sigma_{ij}$$

Eqn. 3-13

This model has two main constraints, the average rate of return of the products and the total weight of the portfolio itself. These constraints are represented by the following:

Average Rate of Return: $\qquad \sum_{i=1}^{n} W_i \bar{r}_i = \bar{r_p} \qquad \forall i \in I$

Investment Limitation: $\qquad \sum_{i=1}^{n} W_i = 1 \qquad \forall i \in I$

The first constraint restricts the model to a defined overall rate of return. This constraint equals the weighted individual rate of return per product to an overall portfolio rate of return value. This value is user-defined, and thus, can take on any value. For purposes of this thesis, this value is assumed to be the overall portfolio rate of return. The second constraint forces the sum of the individual weights to equal one.

3.6.3 Method for Solving Markowitz Model

In order to solve this model, a Lagrangian relaxation technique is used. The equation above is transformed to include the restrictions as part of the objective function. In this case, one has to determine the relaxation parameters that optimizes the function. The following represents the transformed model:

$$\frac{1}{2}\sum_{i,j=1}^{n} W_i W_j \sigma_{ij} - \lambda\left(\sum_{i=1}^{n} W_i \bar{r}_i - \bar{r}_p\right) - \mu\left(\sum_{i=1}^{n} W_i - 1\right) \qquad \textbf{Eqn. 3-14}$$

In order to find the value of the individual values of weight and Lagrangian relaxation parameters, matrix manipulation is used. The following explains the process of transforming this model into matrix form.

$$A = \begin{bmatrix} \sigma_{11} & \cdots & \sigma_{1n} & -\bar{r}_1 & -1 \\ \vdots & \ddots & \vdots & \vdots & \vdots \\ \sigma_{n1} & \cdots & \sigma_{nn} & -\bar{r}_n & -1 \\ \bar{r}_1 & \cdots & \bar{r}_n & 0 & 0 \\ 1 & \cdots & 1 & 0 & 0 \end{bmatrix} \qquad \textbf{Eqn. 3-15}$$

Matrix A is organized in a form so that the solution of the Markowitz model can be solved by matrix manipulation. The upper-left hand quadrant represents the covariance matrix of the price differentials. In the lower-left hand quadrant are the values of the individual rates of return, as well as a row of 1's representing the total weight. In the upper-right hand quadrant the negative values of the lower-left hand quadrant is located. Lastly, the lower-right hand quadrant is composed of a 2x2 zeros matrix.

$$X = \begin{bmatrix} w_1 \\ \vdots \\ w_n \\ \lambda \\ \mu \end{bmatrix} \qquad \textbf{Eqn. 3-16}$$

Matrix X above represents the variables being solved for in the final equation. The variables include the values of the individual weight assignations, as well as the values of the Lagrangian relaxation parameters.

$$Y = \begin{bmatrix} 0 \\ \vdots \\ 0 \\ \bar{r_p} \\ \mu \end{bmatrix}$$

Eqn. 3-17

Finally, matrix Y represents the target solution of the Markowitz model in a manner in which the individual components of matrix X are satisfied. The combination of these three components composes the overall matrix form of the Markowitz model solution.

$A\,X = Y$ **Eqn. 3-18**

$X = A^{-1}Y$ **Eqn. 3-19**

As it is detailed above, the final solution of the Markowitz is equal to the inverse of matrix A multiplied by the target matrix Y. This product details the weight assignation given to each product in the portfolio (or shipment). Furthermore, it details the values given to the relaxation parameters of the constraints.

3.6.4 Summary of Chapter 3

The purpose of this chapter was to explain the methodology followed in designing the two-market strategy. The following two chapters present the results of the analysis performed through the methodology and the potential benefits of the proposed strategy. Finally, Chapter 6 provides conclusions and recommendations for future research on specific opportunities areas of this thesis.

4.   DEVELOPMENT OF SHIPMENT STRATEGIES

4.1   Introduction

As it was mentioned in earlier sections, the objective of this thesis is to develop a low investment expansion strategy for mid-level farmers who wish to extend their commercialization reach within secondary markets. The strategy consists of taking advantage of two-market structure conditions, specifically arbitrage opportunities. The methodology and analysis is applied to a case study involving established farmers in a similar situation who wish to extend their marketing reach into other secondary markets.

The analysis of this thesis is divided among the different aspects of the decision-making process. The first step is to determine the appropriate product shipment policy that would allow the farmer to enter a secondary market through the exploitation of two-market structure conditions. The proposed shipment polity addresses a strategy for both long and short term conditions of the market. For purposes of this thesis, one assumes product availability at the base market, whenever an opportunity is present. Lastly, the farmer needs to limit his risk exposure within the two-market structure by hedging his/her one-time shipments through strategic product groupings. The case study analyzed focuses on addressing first and last aspect of the operation, in order to validate the methodology developed in chapter 3.

4.2   General Assumptions

The present thesis is based on several assumptions with regards to the case study and the application of the proposed methodology.   These assumptions

48

consider different components of the operations, such as the type of markets, the potential product basket, and the mode of transportation. The definition of these operational aspects facilitates the application of the methodology and its validation.

4.2.1   Terminal Markets

The main premise of this thesis relies on the conditions of two-market structure. Thus, it is important to define the type of markets used to perform the analysis. For this thesis' case study, the markets are selected based in part to the availability of price information from the United States Department of Agriculture (Agricultural Marketing Service) databases. The prices reported by the database are on wholesale, terminal market level at important consumption points the US. The prices are reported on a weekday basis throughout the entire year (except holidays and weekends).

The selection process for determining which two-market structure to assess is not crucial for the purpose of this thesis, since the methodology developed should be applicable to any two-market structure. However, only those markets with sufficient price information, enough to provide daily price reports throughout the whole year were considered. This allows an effective implementation of the analysis since its basis is on continuous monitoring of price fluctuations.

Even though, the two-market structure selection is not important for the application of the methodology, the analysis reflects the problem context through which this thesis is developed. Thus, a geographical representation of the actual

49

Mexican farmer's position is desirable. For this purpose, the terminal market at Dallas was selected as the base market for the operations given its proximity to Mexico and its convenient geographical position with respect to other domestic markets within the US.

The list of potential secondary markets was chosen based on relative market importance and demand potential. Also, these secondary markets represent strategic, attractive positions for direct access into even more distant regions. Lastly, price information for these markets is readily available throughout the analysis period. Based on these basic filtering characteristics, the selected markets were the following (Figure 4-1):

- Atlanta, GA

- Boston, MA

- Chicago, IL

- Washington, D.C. (District of Columbia)

- New York, NY

Figure 4-1 – Base and Secondary Markets

For purposes of simplifying the application and results of the proposed methodology, the Dallas-Boston, two-market structures is used for explaining the analysis process presented in this chapter. This structure represents the largest geographical separation. Reference to these results will be given accordingly within the context of this thesis.

4.2.2 Potential Products

Another basic component of the operation is the product selection for the commercialization basket. For this case study, the composition of the commercialization basket is based on the problem context through which this thesis was developed. The application of the methodology is focused on expanding the commercialization capabilities of Mexican farmers of the state of Sinaloa. Thus, the list of products will attempt to replicate their current market situation.

The background behind the selected list of products is based on a previous study regarding the implementation of logistics platform for the distribution of fresh produce in the US (Sanchez 2007). In this study, the author determines a particular US target market on which to expand, as well as the planning process to implement the platform for local distribution. Furthermore, the author uses a basket composed of representative products commercialized by farmers from the state of Sinaloa, MX.

Given that the methodology developed through this thesis derives from a similar problem context, the list of products analyzed in this thesis is similar. One should reiterate that the methodology developed through this thesis is robust enough to handle different kind of perishable products within any given two-market structure. The following list of products will be used for the application of the proposed methodology:

- Tomato (Plum Type)

- Cucumber

- Eggplant

- Squash

- Bell Pepper

An analysis is performed for each these products based on their historic market price information. However, given the relative importance of the tomato (plum type) as the leading commercialization item of Mexican sells in the US, this item will be used for presenting the results of the application process of the

52

methodology. A summary of the results for remaining items will be given when developing basket configurations during the last phase of the methodology.

4.2.3 Mode of Transportation

The mode of transportation used for moving products from the base to the secondary market is an important aspect of the methodology. Nonetheless, the operation is independent of the specific route used. One of the factors that may affect the behavior of the price differential is the product transportation lead time and the transaction cost (transportation, intelligence, etc.) depending on the mode of transportation used.

These factors may affect the amount of variation observed for the market price differentials observed over a defined period of time. For instance, if one decides to use air transportation to move the products, the lead time is reduced, but the transaction cost is increased. This in turn influences the frequency of the shipments, as well as the variations of the differentials during the lead time. Variation in the mode of transportation is fixed for this analysis. The mode of transportation considered for this thesis is truck.

Table 4-1 presents the estimated lead times for each of two-market structure analyzed in the study. For example, the lead time for transporting a product from Dallas to Boston is approximately 3 days. This means that if one observes a favorable two-market condition today and decides to send a shipment, the final selling price is made on the third day (after today). As the lead time is increased, so does the potential for price fluctuations at the secondary market, which in turn may cause an increased risk of loss for one-time shipments.

Table 4-1 – Transportation Lead Times per Mode of Transportation

Origin	Destination	Lead Time by Truck (days)
Dallas	Atlanta	2
Dallas	Chicago	2
Dallas	Boston	3
Dallas	DC	3
Dallas	New York	3

Table 4-2 presents the transportation costs for the mode of transportation considered for this operation. These costs are associated with the transaction costs of transporting the product from the base market to the secondary. The quoted transportation costs are translated to dollars per pound. This information was attained based on refrigerated container quotes reported by transportation companies operating within the US.

Table 4-2 – Transportation Costs of Mode of Transportation

Origin	Destination	Transportation Cost (Truck) ($/lb)
Dallas	Atlanta	0.01248
Dallas	Chicago	0.01845
Dallas	Boston	0.05516
Dallas	DC	0.02983
Dallas	New York	0.05589

The lead time and cost information is used as a key component in the decision analysis process. As mentioned earlier, the lead time limits the amount of variability that can be expected after the product is initially shipped and before it is sold at the additional market price. Lastly, the transportation cost defines the necessary price differential that triggers a shipment.

4.2.4 Analysis Time Period

In order to fully capture the long-term behavior of the two-market structure, a long enough timeline should be analyzed. In the case of this analysis, an arbitrary 10-year period of information was considered. A daily, product price database was created for each of the products within the different terminal markets for the 10-year period, ranging from January 2000 to December 2009. The whole extent of the analysis was performed on information derived from this database during this time period.

An important aspect to consider when analyzing data that extends over such a long period of time is the fact that the information is subject to time-dependent factors. Factors such as inflation could hinder the accuracy of the results. Thus, one of the ways that the database was corrected was through the use of the Consumer Price Index (CPI); a time-dependent index that measures the changes in the price level of consumer goods. Yearly averages of the CPI were collected for the 10-year period from the U.S. Bureau of Labor Statistics (http://www.bls.gov/cpi/, 2011). Table 4-3 presents the yearly CPI average values, as well as the relative values using the year 2000 as base.

Table 4-3 – Consumer Price Index 2000-2009

Year	Consumer Price Index	CPI (Year 2000=1.0)
2000	155.4	1.000
2001	160.9	1.035
2002	163.1	1.050
2003	167.0	1.075
2004	173.6	1.117
2005	177.0	1.139
2006	180.2	1.160
2007	187.4	1.206
2008	198.7	1.279
2009	199.4	1.283

The yearly CPI relative values were used to correct yearly product prices by multiplying them by the quoted terminal market prices. The result then accounted for yearly increases in price due to economic inflation within the consumer good markets. In the case of this study, this allows direct comparisons between years, as well as the development of a single shipment policy that is relevant for the entire 10-year span.

4.3 Market Price Analysis

The long-run average prices for those markets considered suggest that continuous shipment operations in these structures are not profitable (Table 4-4). As one can observe from the table below, the average long term prices at the base market in general tend to be close, if not, higher than at the secondary. Consequently, the price differentials between the structures are not large enough to allow continuous profitable transactions. In order to capture the opportunities that are present within the market price differentials, one has to search for specific

opportunity windows in which one might be able to observe from gains from engaging in a two-market trade.

Table 4-4 – Long-Term Average Prices

	Dallas	Boston	Atlanta	Chicago	DC	NYC
Tomato	$0.70	$0.76	$0.70	$0.71	$0.72	$0.66
Squash	$0.58	$0.46	$0.49	$0.50	$0.53	$0.46
Eggplant	$0.94	$0.86	$0.57	$0.83	$0.55	$0.77
Cucumber	$0.39	$0.37	$0.33	$0.39	$0.31	$0.36
Bell Pepper	$1.07	$0.67	$0.99	$0.97	$1.01	$0.84

For this, one needs to dwell a bit deeper into the market price differentials, in order to identify the specific opportunity windows that indicate potential arbitrage in a two-market transaction. Figure 4-2 presents an in-depth glimpse of the market price differentials for an arbitrary year of 2005 within the Dallas-Boston market structure. The values observed in this graph account for the non-lagged differentials between these markets. This differential is the price of the product (per pound) at the secondary market minus the price at the base market and the cost of transaction, during the same day.
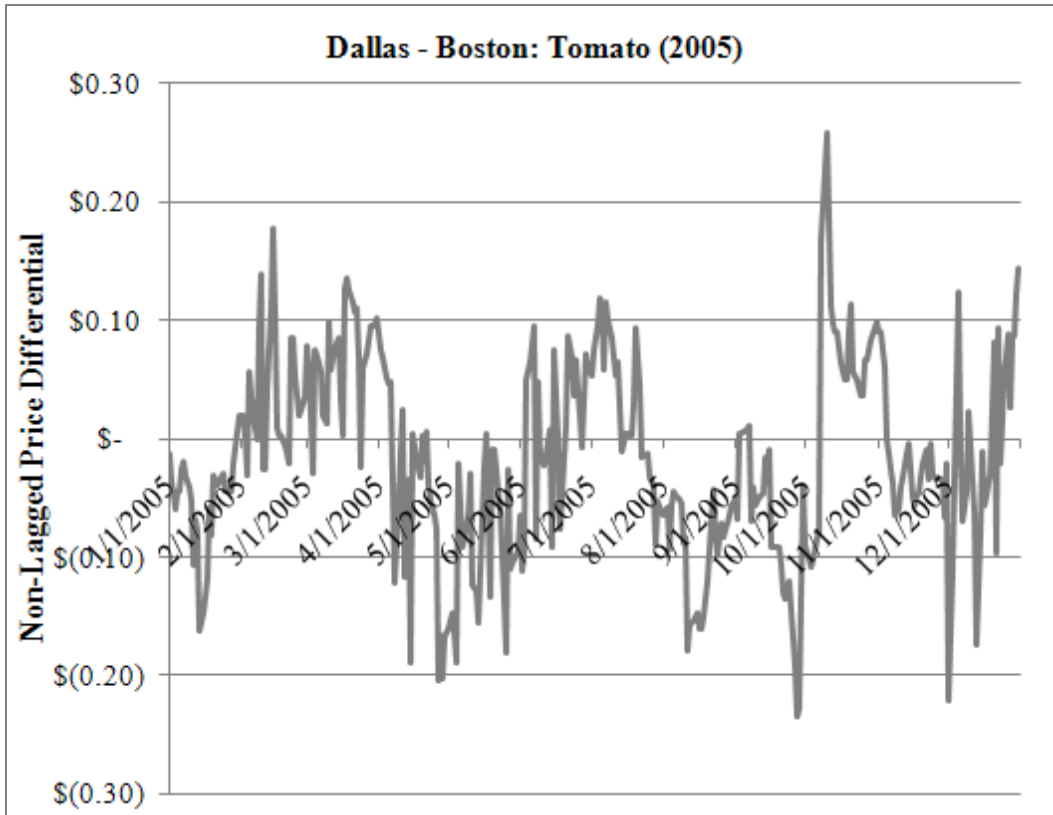
Figure 4-2 – Non-Lagged Price Differentials: Dallas – Boston

As one can observe from this figure, on the average, the differentials appear to be negative, and thus, normally, one would not venture to have continuous operations in the Boston market. However, from a visual inspection of these points, there appears to be particular time windows in which the differential per pound is relatively high. These time instances could potentially become opportunity windows for the decision-maker if a shipment is made. For this, it is necessary to observe the lagged price differentials, which represent the actual profits that would be made if the product is acquired at the base market and sold at the secondary 3 days later. Figure 4-3 presents the lagged price differentials, accounting for the transaction cost, between Boston and Dallas.

58

This value is the price of the product (per pound) at the secondary minus the price

at the base market and the transaction cost, 3 days earlier.
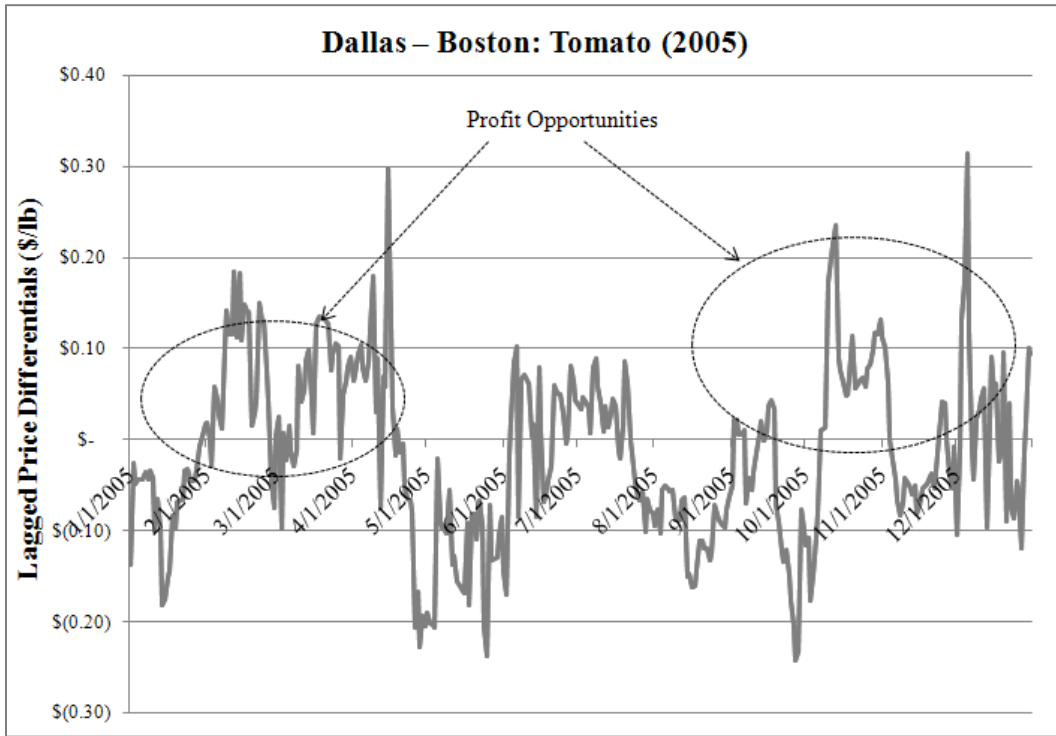
Figure 4-3 – Lagged Price Differentials: Dallas - Boston

In the graph above, each point above the zero marker represents a

potential arbitrage opportunity for the decision maker. In the figure, the encircled

points at the left side of the figure represent opportunities that may be a bit riskier

given the high variability in the differentials. On the other hand, those points

encircled on the right would mean more stable profit opportunities. Whatever the

case may be, one can conclude based on these differentials that there are windows

of time throughout the year, which may indicate a potential profit opportunity and

that a lucrative operational strategy may be developed.

The focus of the analysis then becomes on generalizing the attributes of specific price differentials and identifying those that most likely translate into a positive transaction. Thus, one must determine the characteristics of particular price differential values, in terms of expected profit, associated risk, frequency and duration, in order to decide on which type of opportunity is the most beneficial to the decision-maker given specific objectives of the operation. The application of the methodology developed in this thesis attempts to summarize and identify the behavior of particular price differentials that are inductive to positive profits. The results provide the user with a decision-analysis tool to pick and choose on the most beneficial opportunities and send product shipments when it is deemed profitable. The next few sections detail the application process of the methodology.

4.4 Long-Term Threshold Values for Shipment Policy

The first phase of the methodology involves determining an appropriate shipment policy that maximizes long-term expected profits. This policy is in terms of identifying the specific price differential that increment the overall profits per shipment. The methodology involves two main approaches; one which involves a pragmatic process of attaining the estimated profits under several price differentials and another which attempts to follow a theoretic approach to determining threshold value that optimizes the profits. The following sections present the application process of the methodology developed in section 3.4.

4.4.1 Distribution Fit for Two-Market Price Differentials

The first step in the methodology is to summarize the behavior of the price differentials throughout the 10-year period in an iterative manner. For demonstration purposes, the methodology is applied to the price information of tomato (plum type) within the Dallas-Boston, two-market structure. Based on this structure, one assumes a transportation lag of 3 days (truck mode) and a total transportation cost of $0.05516/lb shipped. One should note that the methodology is also applied to the other products in the commercialization basket within the other two-market structures. The results of this analysis are used in the following sections to develop appropriate shipment grouping strategies.

The summary of the price behavior is attained through the application of the iterative decision process flow developed in the methodology and represented by Figure 3-2. The decision flow is applied to the price information data set for a range of price differential values. As it was explained in the methodology, a summary of the profit results (lagged price differentials) is collected for each threshold value, K. These summaries are analyzed individually in order to determine their respective expected profit, frequency, rates of return and associated risks.

The summary process starts by creating a histogram of the profit results based on different K values. For demonstration purposes, Figure 4-4 presents the histogram of historic lagged price differentials for the Dallas-Boston, two-market structure under a threshold value of $K>0.15$. In this figure, the x-axis represents different interval bins of the price differentials, while the y-axis represents the

proportion of the entire data set population contained within each bin. Since these price differentials already account for the transportation lead time and transaction cost of the shipment, these observations are the actual profits in dollars per pound of tomato shipped. The iterative process is repeated for various values of *K* as described in section 3.4.3.
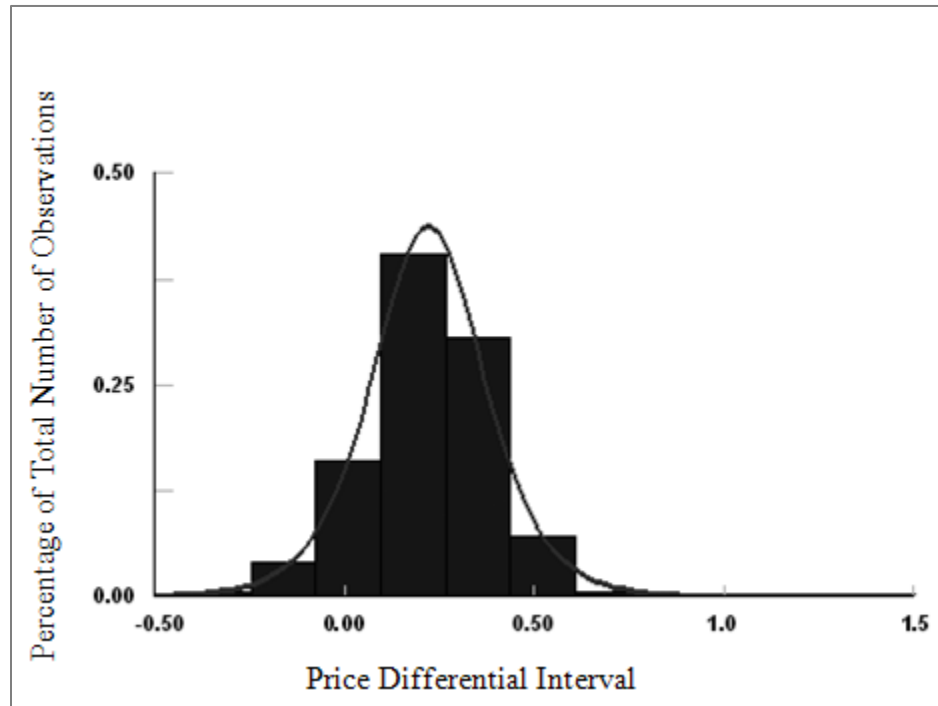
Figure 4-4 – Histogram of Price Differentials

For the development of the different histograms, a stepsize of 0.05 was used. This would mean that the iterative process starts a threshold value of $K > 0$ and is iteratively increased by 0.05.   If one wishes to attain higher resolution in the results, a smaller stepsize may be used. However, in this case, the application of the methodology is performed for demonstration purposes, in order to observe and present the general tendencies of the historic profits under different threshold values.

Figure 4-5 presents the behavior of these market price differentials for the Dallas-Boston, two-market structure under values of *K* that range from 0.05 to 0.40. As it was described earlier, the histograms of these observations are the lagged price differentials and represent the actual historical profits. In this figure, each of the histograms is similar to Figure 4-4. The x axis represents the interval bin for the differentials, while the y axis represents the percentage of the total observations corresponding to each bin. On the top right corner of each histogram is the threshold value that is used for the iterative process.
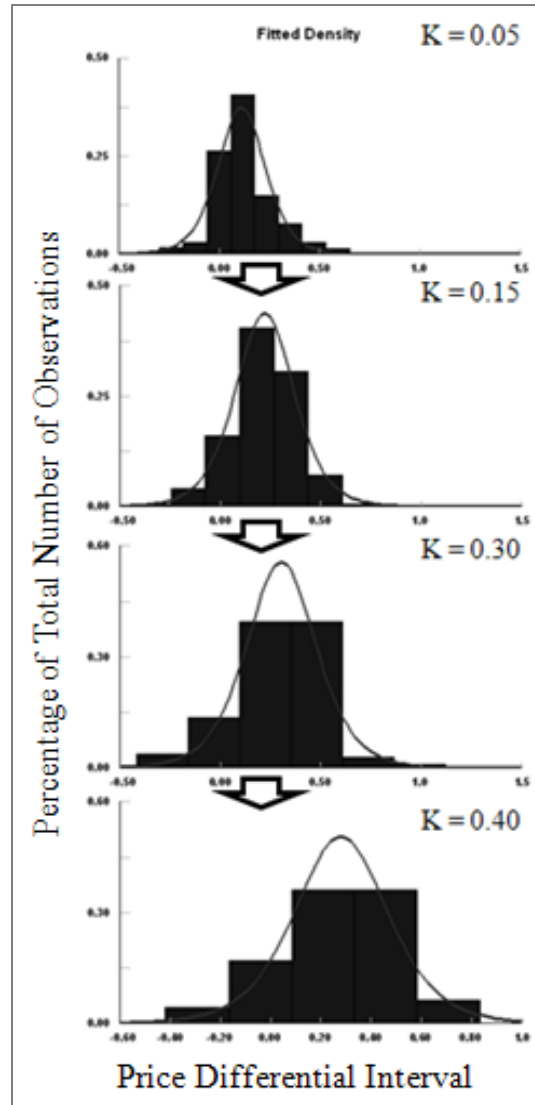
Figure 4-5 – Histogram of Price Differential under Various Values of K

From the figure above, one can observe the general tendencies of the histogram distributions as the threshold value is increased. From the general distribution, one can detect that the bulk of the observations has a tendency to move rightward. This indicates that as the value of K is increased, the expected profit per shipment also gets larger. Furthermore, one can observe that the variance of the profits, or the width/spread of the observations, tends to increase

as the threshold value gets larger, which might indicate higher variability. However, since the center point of the observations is far to the right, this variability might mean little in terms of potential losses.

This series of histograms provide a general idea of the effect of the shipment threshold criterion. However, in order to attain an estimate of its actual effect, the next step is to summarize the behavior of the histogram observations through theoretical distributions. Furthermore, since the purpose of the statistical fits is to compare the results among the different shipment strategies, a single type of theoretic distribution was selected to represent the histogram observations per threshold value. This allows easier comparisons among the different strategies. Lastly, as it was mentioned in earlier sections, one will assume the price differentials to be continuous random variables.

Using a Pearson's chi-square goodness of fit test, it was determined that the logistics distribution function best fitted the collection of histograms per value of K. The underlying analysis of this goodness of fit test establishes whether or not an observed frequency distribution differs from a theoretical distribution. In this case, the distribution of the histogram is compared against different theoretical distributions (e.g. normal, lognormal, etc.) using a p-value of 5%. StatFit software was used to simplify the application of the goodness of fit test, and to compare among the different distributions. The results for this goodness of fit test can be found in APPENDIX C.

Figure 4-6 presents a logistics distribution function that is fitted on the histogram described in Figure 4-4. In this figure, the x-axis represents the lagged

price differentials, while the y-axis represents the corresponding distribution function values. As one can observe from the figure, these observations assume a logistics distribution fit with a mean of 0.22 and a standard deviation of 0.0976. These values are attained from the StatFit software, which generates and fits the parametric values that best fits the distribution of the histogram observations.
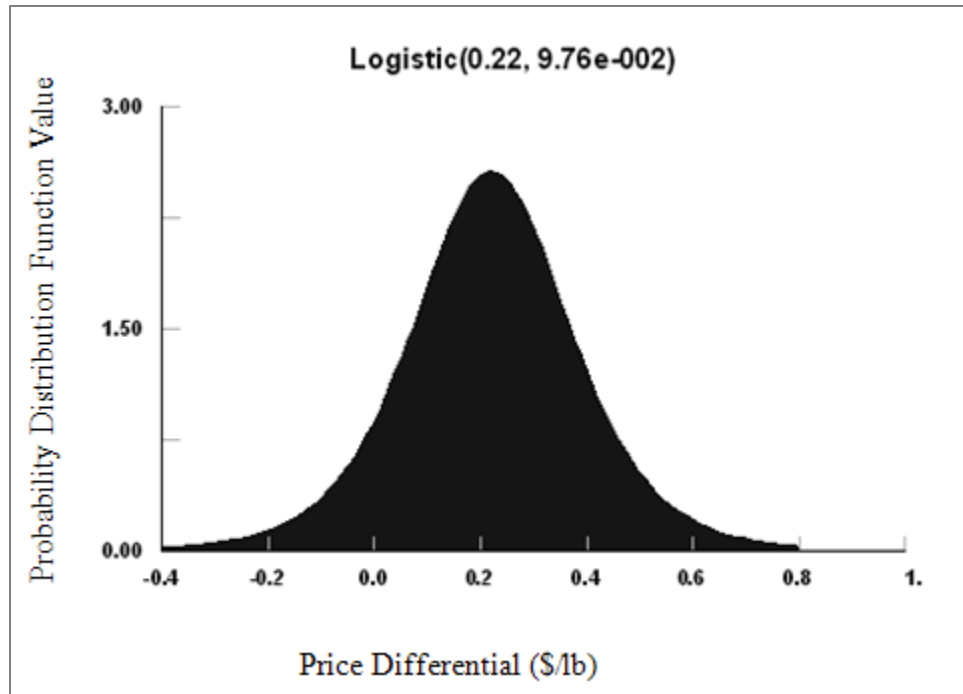
Figure 4-6 – Distribution Fit of Market Price Differentials

The parametric values of the fitted distribution function allows the user to gather information regarding the expected mean profit ($/lb), variance, and probabilities, for various threshold values. In this case, the mean of the fitted logistics distribution function is assumed to be the expected profit per pound of the individual shipments, while the standard deviation parameter is assumed to be the variance of the nominal profits. Furthermore, one is able to use these

parameters to estimate probabilistic values from their respective probability distribution function.

A visual representation of these distribution functions per value of K is presented in Figure 4-7. These distribution functions are created based on the parameters of the mean and standard deviation of the histogram observations. As one can observe from the different functions, the tendency of the peak point is to move rightward in the positive direction. At the same time, the standard deviation, or the width of the curve, starts to increase for larger values of K, just as it was noted for the histograms in Figure 4-5. Again, this might indicate higher volatility, but as it was pointed out, the potential for losses is less given that the central point is farther away from zero.
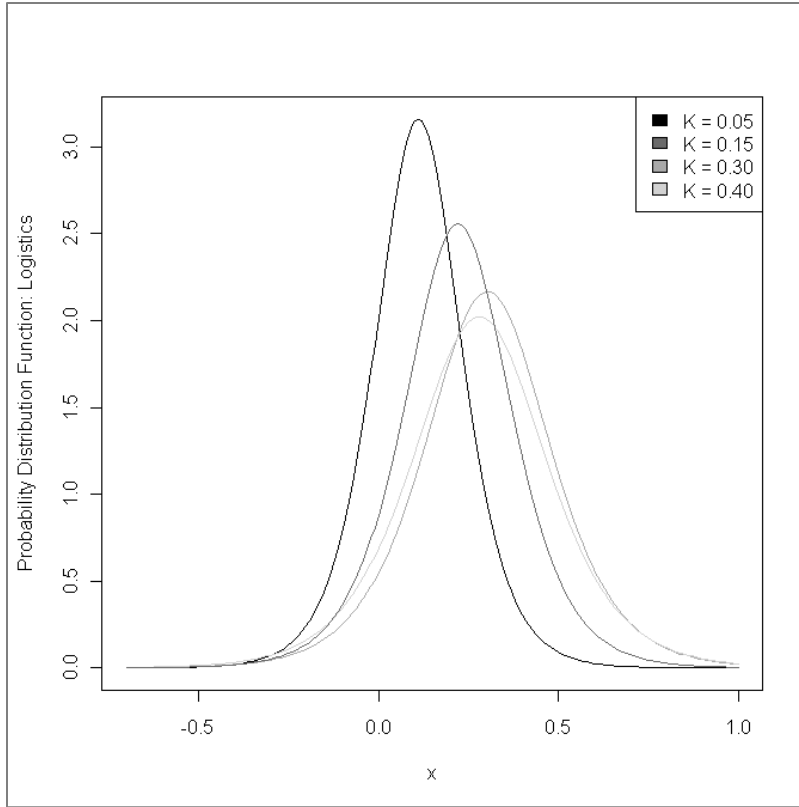
Figure 4-7 – Logistic Distribution for Different Values of K

Based on the parameters of these distribution functions, one is able to estimate an expected profit and standard deviation under different values of K. These values are summarized in Table 4-5 for an iterative stepsize of 0.05 of the decision process flow. Furthermore, Figure 4-8 presents a graphical representation of these values. In this graph, the values of the expected mean and the standard deviation are fixed on the primary and secondary vertical axis, respectively.

Table 4-5 – Logistics Distribution Expected μ and σ

K	μ	σ
0.05	0.110	0.079
0.1	0.168	0.093
0.15	0.220	0.098
0.2	0.252	0.101
0.25	0.281	0.106
0.3	0.303	0.115
0.35	0.309	0.121
0.4	0.280	0.124

As one can observe from the table above, the estimated values of profit and standard deviation increase, as the threshold value gets larger, as it has been noted graphically with the histogram observations and the logistic distribution fits. Also, one should point out that as the price threshold value is increased, logically, the number of observations that meet this criterion decreases significantly. Given this tendency, it was observed that for K values of 0.40 and higher, the number of observed points (<20) is not enough to perform reliable distribution fits, and thus were not included as part of the analysis process.
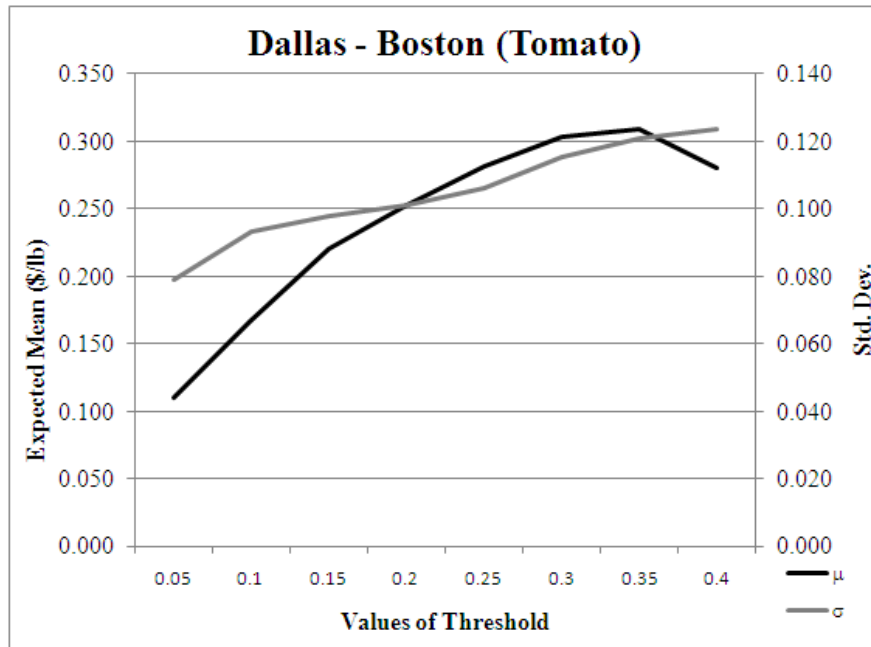
Figure 4-8 – Expected μ and σ

The information generated by the distributions will allow us to develop a methodology for determining the optimum value of the threshold based on the particular objective of the decision-maker. However, additional information of the historic profits will be needed to better compare among the different shipment policies. Among this information is the frequency of the opportunities and the associated risk per value of K. The following section addresses this issue and determines the optimal value of the threshold following the steps detailed in the methodology in section 3.4.3.

4.4.2 Additional Components of the Operations

The previous section details the effect of varying the price differential threshold on the estimation of the profits' expected mean and standard deviation. Based on a general analysis, it was determined that as the threshold value gets larger, both the mean and standard deviation of the statistical distribution fits also

70

increase.   These values contribute an important part in defining the general characteristics of the profits per threshold value. However, as it was mentioned in the methodology, one needs to assess other aspects of the shipment strategy that further aids the decision making process.

As it was noted earlier, stricter shipment policies reduce the number of times that an opportunity is identified; the number of one-time shipments over a defined period of time is reduced. Table 4-6 details the frequency of the historical shipments under each particular value of *K*. This information ranges over the 10-year period of operations considered. As one can observe from this table, the frequency of the shipments decreases, as the threshold value is increased.

Table 4-6 – Shipment Frequency per Value of K

K	μ	σ	Frequency
0.05	0.110	0.079	930
0.1	0.168	0.093	515
0.15	0.220	0.098	324
0.2	0.252	0.101	223
0.25	0.281	0.106	154
0.3	0.303	0.115	105
0.35	0.309	0.121	74
0.4	0.280	0.124	48

A higher frequency on the shipments would suggest that lower K values represent the best strategy for the decision maker, since in the long run this would translate into higher total earnings over the entire operational period. One can observe from Table 4-7 this general trend; as the threshold level is increased, the total net earnings (10-year-span) are decreased. In this case, the total earnings are equal to the product between the expected profit per pound shipped, the frequency

of these shipments, and the assumed container capacity. One main assumption is that for each opportunity, a single shipment is made with a capacity of 40, 000lbs. based on the general characteristics of refrigerated containers.

Table 4-7 – Net Earnings per Value of K (Dallas – Boston)

K	μ ($/lbs)	Net Earnings (Thousands $)
0.05	0.110	5,132.40
0.10	0.168	4,588.80
0.15	0.220	2,848.00
0.20	0.252	891.60
0.25	0.281	1,682.00
0.30	0.303	1,222.40
0.35	0.309	878.80
0.40	0.280	500.00

The table above suggests that relaxing the shipment criterion may translate into higher earnings for the operation. However, an important component of the decision analysis process is the risk associated with each shipment strategy. Especially, if one wants to reduce the potential for a loss whenever a shipment has been made. For this purpose, a Value-at-Risk (VaR) analysis was used through the application of two methods, in order to estimate the potential maximum loss that can be encountered at any time during an opportunity period.

The first approach for calculating the VaR was through the use of a historic method. In this case, the returns per threshold value are collected, and the 5th percentile of those return values is calculated; this value represents the 5% VaR under that particular shipment strategy. The main advantage of this method is that it does not rely on a statistical assumption of the returns distribution.

The second approach is to estimate the VaR through the use of a variance-covariance method. This involves assuming a normal distribution on the set of returns and using its parameter values (mean and standard deviation) to estimate the 5% VaR. The main advantage of this method is its easiness to use given that it only needs two parameters. However, its main disadvantage is the normality assumption, especially for the highly skewed distribution of the returns, which result in a higher tendency for imprecise estimates.

Table 4-8 summarizes the VaR per pound of tomato by using both the historical and variance-covariance methods. As one can observe from the table below, the general tendency of the VaR is to reduce in risk as the threshold is increased. There is a bit of discrepancy in the estimated values of both methods, due in part to the normality assumption that was discussed earlier on the variance-covariance method. Given that the historical method does not heavily depend on the distribution of the returns, its estimates of the VaR are assumed to be more in accordance to the actual associated risk of each strategy. Thus, these values are selected for continuation in the decision analysis process.

Table 4-8 – VaR per Value of K

K	μ	VaR (Hist)	VaR (Var-Cov)
0.05	0.110	0.005	0.174
0.10	0.168	-0.018	0.136
0.15	0.220	-0.022	0.099
0.20	0.252	-0.097	0.076
0.25	0.281	-0.097	0.044
0.30	0.303	-0.091	-0.016
0.35	0.309	-0.023	-0.039
0.40	0.280	-0.024	-0.121

For these estimates of VaR, one can observe that as the threshold level is increased, the associated risk of the strategy tends to be less. In fact, the VaR move from positive to negative estimates for higher threshold levels. As it is explained in APPENDIX A, a negative VaR indicates a higher potential for a return during the next period. For example, under a $K > 0.25$ strategy, one can say with 95% confidence that the return will not be less than 0.097 over the next period. Thus, based on these VaR estimates, the associated risk is less from mid-range values of the threshold. Figure 4-9 provides a visual representation of this behavior.
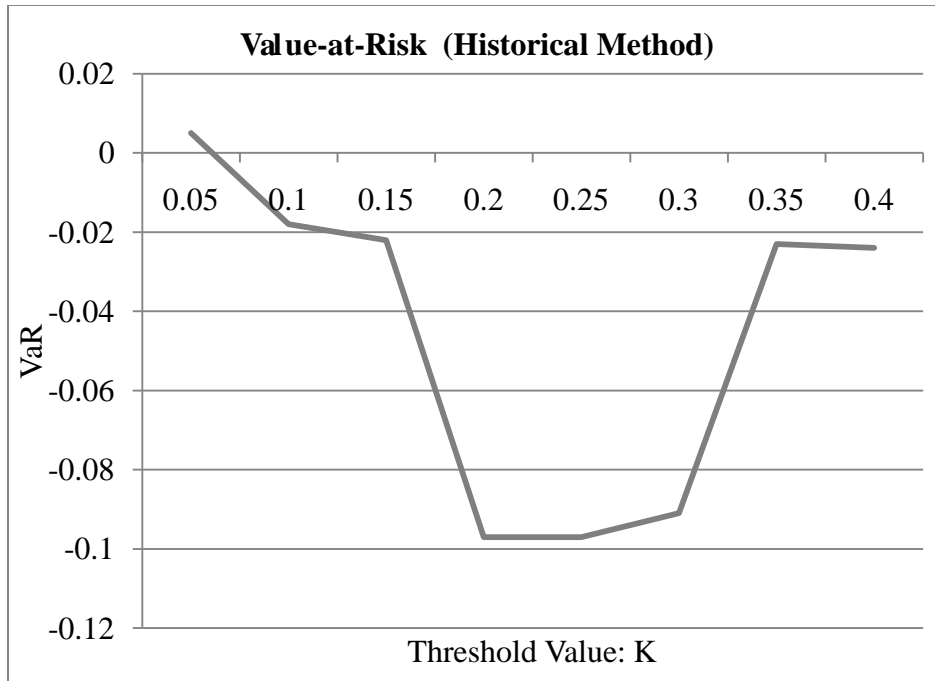
Figure 4-9 - VaR per Threshold Level

The frequency, total earnings and VaR are all additional components that can facilitate the decision-making process. The next section details a shipment strategy that incorporates all these different components into meeting particular objectives of the operation. This will be done in both a pragmatic and theoretical manner.

4.4.3 Optimal Threshold Value: Pragmatic Approach

The objective of this part of the operation is to determine a shipment strategy that meets the ultimate objective of the decision maker. As it was demonstrated in the previous section, one must decide on whether to maximize the total earnings of the operations over the entire period of time, or if to maximize the expected profit per shipment and reduce the associated risk. In

order to determine these values, two approaches were proposed in the methodology section (pragmatic and theoretical).

Under the pragmatic approach, the methodology that was used to analyze the general behavior in the previous section is applied to determining optimal threshold levels. In this case, the resolution on the results is increased by reducing the resolution of the iterative stepsize. This stepsize was reduced from 0.05 to a value of 0.001, in order to attain more precise results. Lastly, an analysis was performed for different operational objectives, such as maximizing total earnings or rates of return.

In the representation of the threshold, the transaction cost from Dallas to Boston is added to the K value. In this case, the summation represents the threshold without considering the transaction cost between Dallas-Boston. However, the analysis or the shipment decision process is not affected in any way. It only changes the reference point.

Table 4-9 – Optimal K per Objective

Objective	Total Earnings (thousand \$)	E[Profit] (\$/lb)	Value-at-Risk	Rate of Return	$K + C_{ij}$
Max Earnings	\$5,555.20	0.075	-0.018	0.122	0.060
Max Profits	\$943.68	0.318	-0.024	0.406	0.400
Min VaR	\$1,853.28	0.278	-0.097	0.345	0.289
Max ROR	\$943.68	0.318	-0.024	0.406	0.400

Table 4-9 presents the optimal value of the threshold based on the underlying objective of the decision-maker. In this table, the total earnings column represents the overall net profits of the shipment, while the second column is the expected profit per pound of tomato shipped. For example, in the

case in which the objective is to maximize the total earnings of the operational period, a small threshold level is needed, which in turn increases the frequency of the shipments. On the other hand, if the objective is to maximize the expected profit per individual shipment, then one would opt to increase the threshold value. This would also hold true if the decision-maker wishes to minimize the VaR of the shipments.

4.4.4   Optimal Threshold Value: Theoretical Approach

The second approach is the development of an expected profit model with the purpose of identifying a theoretic threshold level that optimizes the long-term operational profits. As it was described in the methodology, an expected profit model is developed and differentiated with respect to the threshold value, K. Also, the expected profit model is based on a particular bivariate distribution that represents the joint probability function with respect to the random variables, x and y. It was mentioned that one would assume a distribution in order to estimate the values of this joint p.d.f.

Figure 4-10 details a scatter plot of lagged versus non-lagged differentials in order to observe the general relationship between these two random variables. For the Dallas-Boston market structure, the correlation of tomato prices for lagged and non-lagged is approximately 0.857. High levels of correlation is to be expected given the dependency of both variables. As one can observe in this figure, the relationship between these two appear to be fairly linear.

77
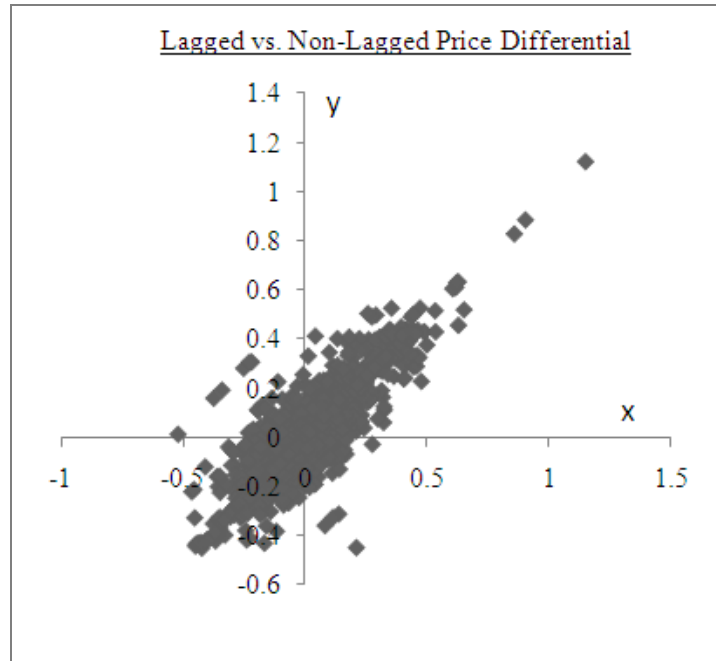
Figure 4-10 – Scatter Plot: x vs. y

For simplification purposes and in order to test the developed theoretical approach, a bivariate normal distribution is assumed for the representation of the general behavior of x and y. This assumed distribution represents the joint probabilty function that is used to evaluate the differentiated expected profit model (Eqn. 3-6). Figure 4-11 is a visual representation of the bivariate normal distribution for the laggged and non-lagged price differentials in the Dallas-Boston market structure for tomato. As one observe, the function follows the general behavior of the scatter plot shown in the previous figure.
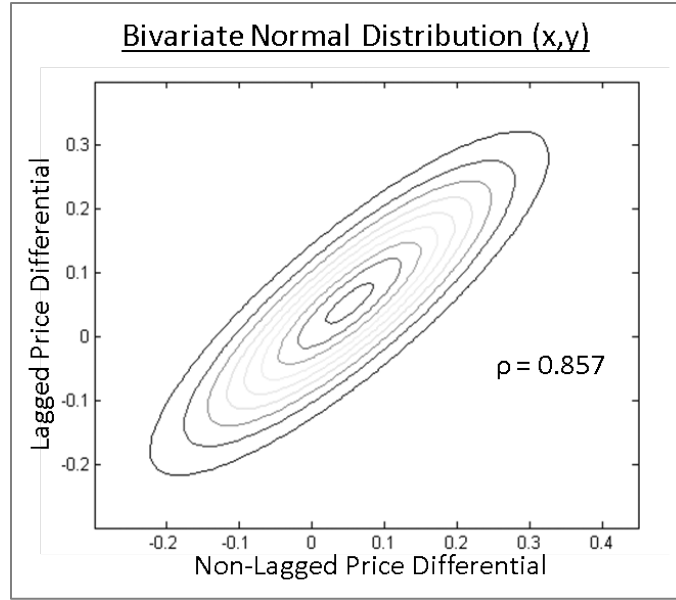
Figure 4-11 – Assumed Bivariate Distribution for x and y

After assuming joint probabilty function of the x and y, it is much easier to solve for the threshold value that optimizes the expected profit model. Equation 4-1 presents the conditional expectation of a bivariate normal distribution plus the transaction cost, which represents the simplified version of the optimized expected profit model of Equation 3-6.

$$-E[y \mid x = K] + C_{ij} = \mu_y + \rho\sigma_y \frac{(x - \mu_x)}{\sigma_x} + C_{ij} = 0 \qquad \textbf{Eqn. 4-1}$$

If one evaluates the model, assuming x is conditioned on the value of K, then one can solve for the optimal threshold. As it was explained earlier in the methodology, this solution is an approximation to optimality, since x is assumed equal to K, when in fact, it should be represented by an inequality $(x > K)$. The additional parameters of the expected profit model are estimated based on a normal distribution of emprircal data on x and y ($\sigma_x = 0.1277, \sigma_y = 0.1301$,

$\mu_x = 0.0521$, $\mu_y = 0.0518$, $C_{ij} = 0.0552$, and $\rho = 0.8576$). Equation 4-3 is

the optimized model in terms of the threshold.

$$-E[y \mid x = K] + C_{ij} = \mu_y + \rho\sigma_y \frac{(K - \mu_x)}{\sigma_x} + C_{ij} = 0 \qquad \textbf{Eqn. 4-2}$$

$$K = \mu_x + \frac{(C_{ij} - \mu_y)}{\rho}\left(\frac{\sigma_x}{\sigma_y}\right) \qquad \textbf{Eqn. 4-3}$$

$$K = 0.1227 + \frac{(0.0552 - 0.0518)}{0.8576}\left(\frac{0.1277}{0.1301}\right) \qquad \textbf{Eqn. 4-4}$$

$$K = 0.05605$$

After substituting the indvidual values for each component, one arrives at

the solution K=0.05605. Again, one must note that in the decision process one

uses the inequality K>0.05605 as the optimal shipmen criteria. Figure 4-12

represents the total historical profits as the value of the threshold (presented in the

x-axis) is increased in a pragmatic manner. The total historical profits are shown

on the primary vertical axis, while the average profits per shipment are shown on

the secondary. As one can note, the total profits during the operational period is

increased as the value of the threshold, K, is increased, and it hits a local maxim at

some point in this region.
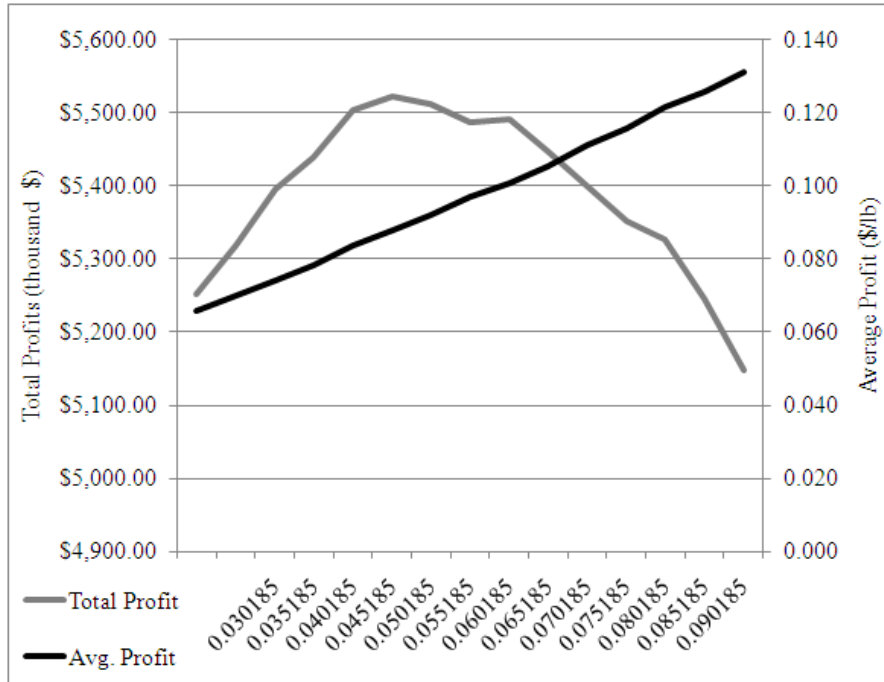
Figure 4-12 – Total and Average Profits vs. Threshold Value

Table 4-10 presents the actual values that are represented by the figure above. In this table, one can see more clearly that the local maxima observed actually falls within the interval $0.0502 < K < 0.0602$, which matches the results obtained from the theoretical approach. This means that the optimal price differential (without considering transaction cost) that maximizes the long-term profit models is above approximately $K > 0.06505$.

Table 4-10 – Total and Average Profit per Threshold Value

Threshold + Cij	Total Profit	Avg. Profit
0.0452	$ 5395.584	0.0784
0.0502	$ 5438.908	0.0836
0.0552	$ 5502.928	0.0880
0.0602	$ 5522.660	0.0921
0.0652	$ 5512.480	0.0972
0.0702	$ 5487.104	0.1008
0.0752	$ 5490.544	0.1055
0.0802	$ 5448.264	0.1113

Overall, one can conclude that the results obtained from the theoretical approach tested the strength of the theoretical profit model. Based on these results, one can develop a shipment strategy that optimizes the long–term expected profits during an operational period. In the case in which one wants to optimize other aspects, such as the rate of return per shipment, then one needs to restructure the optimization model in a way in which a threshold is obtained for that specific purpose.

4.4.5   Two-Market Structure: Short Term Probabilities

Another aspect of the methodology is the development of a decision-making tool designed for estimating short-term probabilities of negative and positive profits based on particular market conditions. It was detailed in section 3.4.4 that that the chance of a positive profit is based on the probability that the market conditions will trigger a shipment and that the duration of the opportunity will outlast the transaction time. The ultimate purpose is to develop a probability grid in terms of both the duration of the opportunities and the non-lagged price

differential. Based on its values, the decision-maker is able to make intelligent short-term shipment decisions based on the conditions of the market price differentials.

4.4.5.1 Probability Distribution Fits on Duration

In order to develop a probability grid that is in function of both the differential and duration variables, one must assume particular distributions for each. Earlier sections detailed the process of fitting statistical distributions on the differential observations. In this case, a similar process will be used to fit a statistical distribution on the histogram observations of the durations. The result should be a short-term decision making tool that is able to estimate discrete probability value of profits based on particular values of the differentials.

The first step in this process is to fit a theoretic distribution on the histogram observations of the duration, which are conditioned on particular differential values. One must note that since the differential is assumed to be a continuous random variable, one cannot evaluate the probability for a single point of the differential. Thus, intervals of the differentials are used to capture these probabilities.

For the Dallas-Boston market structure, an interval width of 0.02 is used for the price differentials, while the total number of intervals range from K= 0 up to 0.40. Then, the durations of the opportunities within each interval are recorded and placed on a histogram. Finally, a statistical distribution that best fits the characteristics of the histogram observations is selected. The following presents a more in-detail summary of the analysis process.

Figure 4-13 presents the histogram observations for durations conditioned on a price differential interval of $0.00 < K < 0.02$. In this histogram, the x-values correspond to price duration interval bins, while the y-values is the corresponding proportion of the whole data set. As one can observe, the proportion of the data set contained in higher duration values reduces for higher duration values. A theoretic distribution will be fitted on these histogram observations.
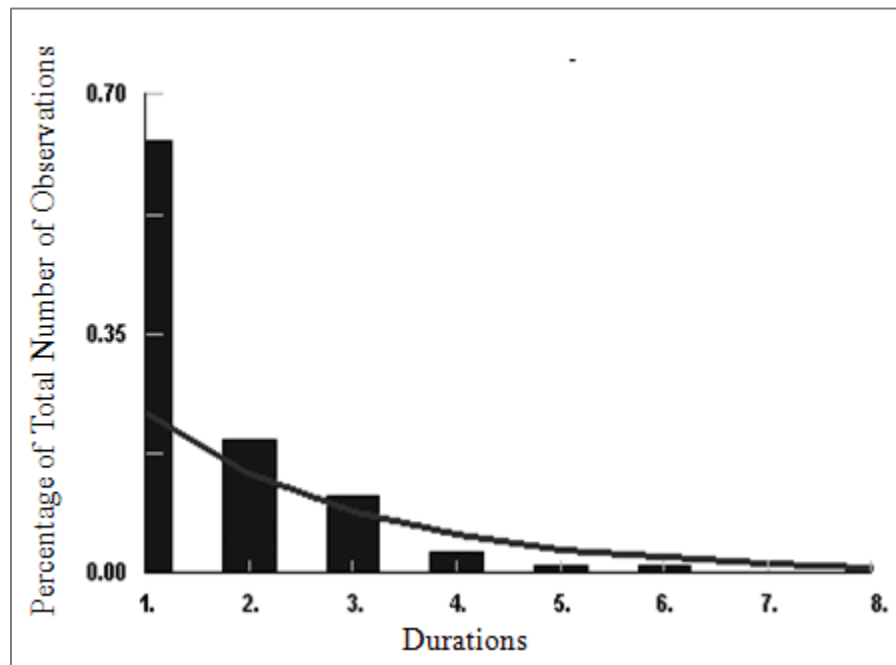
Figure 4-13 – Histogram of Positive Differential Durations

A Chi-square goodness of fit test was used to determine the statistical distribution that best fits the histogram observations of the duration. StatFit software is used to produce the results of this test (APPENDIX D). It was found that a geometric function would be the best choice for summarizing its behavior. Figure 4-14 presents the statistical distribution fit of the previous histogram

observations. In this case, the x-axis represents the durations, while the y-axis represents the corresponding probability mass function (pmf) values.
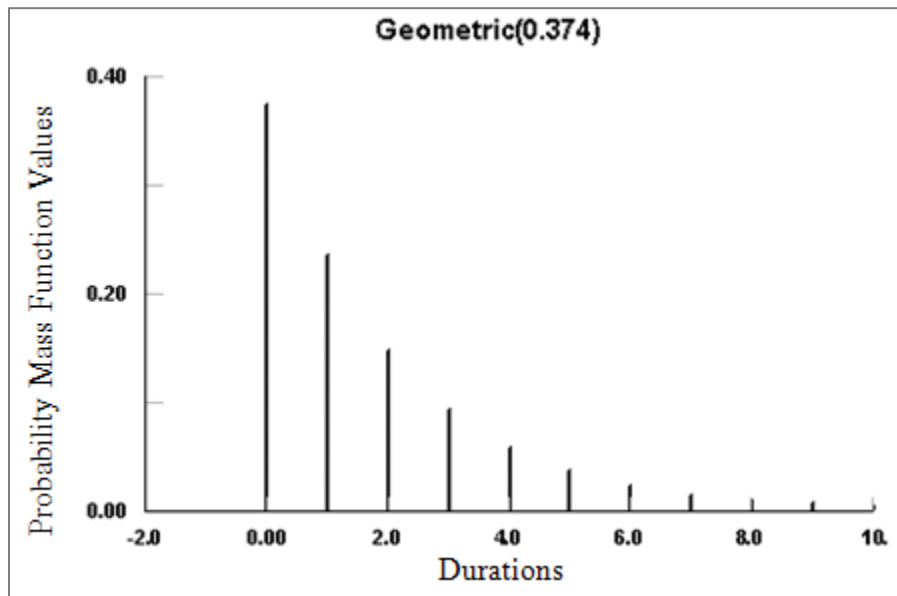
Figure 4-14 – Distribution Fit of Differential Durations

This process is repeated for the rest of the price differential intervals. A geometric function is assumed for each of these intervals, since it fits the distribution of each histogram. Now, the next step is to use these probability mass functions to estimate the probabilities of profit and loss conditioned on particular values of the market price differential. The following section will detail this process.

4.4.5.2 Conditional Probability Grid

Once a pmf has been fitted for the durations under the different K intervals, the next step is to estimate the conditional probability of a particular duration. For this, the probability mass function of the duration is evaluated under different duration values. For example, the probability mass function

conditioned on a price differential interval of $0.02 < K < 0.04$ is evaluated for a range of durations (0<duration<15). In this case, each estimate of the pmf corresponds to the probability of a particular duration conditioned on this price differential interval.

If this process is repeated for different values of the duration and differentials, the final result is a conditional probability grid based on these two variables. Figure 4-15 represents a surface plot of this conditional probability grid. As one can observe from the figure, as the price differential is increased, so does the probability for smaller durations. This means that for higher price differentials, it is more possible that the duration will not hold through the 3 day transportation period.
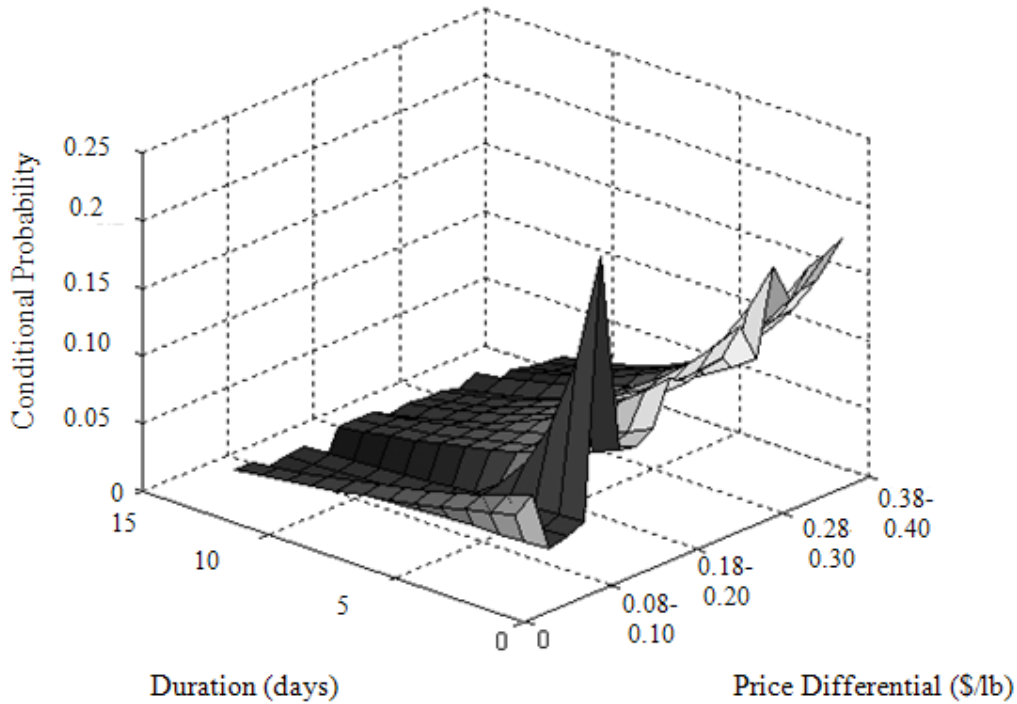
Figure 4-15 – Conditional Probability of Duration of Differential

In order to estimate the actual probability of a positive profit given a particular price differential, one simply needs to aggregate the probabilities that the duration falls beyond the three day period. On the other hand, if one wishes to estimate the probability that a price differential will result in a loss, then one should aggregate the probabilities of a duration less than the transaction period. Table 4-11 presents these cumulative probabilities per K interval. As one can observe, the probabilities of a negative profit increase as the price differential is greater; the probabilities of a positive profit are at their higher levels for mid-range values of the K-interval.

Table 4-11 – Probabilities of a Profit given K interval

K-Intervals	Positive	Negative
0.00 - 0.02	0.422	0.130
0.02 - 0.04	0.447	0.146
0.04 - 0.06	0.256	0.062
0.06 - 0.08	0.278	0.069
0.08 - 0.10	0.311	0.080
0.10 - 0.12	0.254	0.379
0.12 - 0.14	0.461	0.157
0.14 - 0.16	0.457	0.154
0.16 - 0.18	0.471	0.167
0.18 - 0.20	0.495	0.224
0.20 - 0.22	0.494	0.208
0.22 - 0.24	0.495	0.218
0.24 - 0.26	0.495	0.219
0.26 - 0.28	0.487	0.251
0.28 - 0.30	0.493	0.201
0.30 - 0.32	0.439	0.302
0.32 - 0.34	0.483	0.258
0.34 - 0.36	0.458	0.287
0.36 - 0.38	0.452	0.292
0.38 - 0.40	0.438	0.303

Figure 4-16 provides a visual representation of the table above. The probabilities of positive and negative profits have a tendency to converge as the interval of K is increased. However, price data availability for higher K intervals was limited, and therefore this cannot conclude with strong certainty. Since the objective of the operation is to increase the probability of a profit while minimizing that of a loss, one would want to perform a transaction whenever the differential between these two probabilities is higher. Based on these results, one would want to perform a transaction whenever the differential is approximately between 0.12 and 0.26.
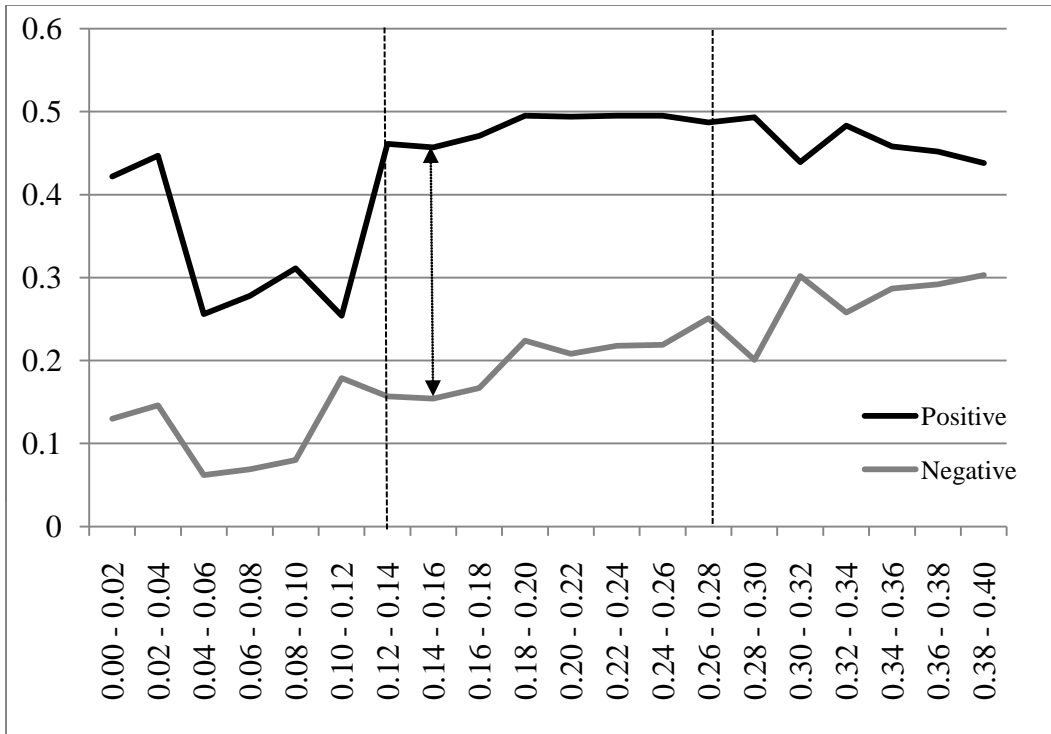
Figure 4-16 – Expected Profits per K Interval

One should note that the analysis so far only accounts for the profit probability once a particular price differential is identified. However, a decision-maker might also want to consider the probability of a profit at any instance in time. In this case, one must also consider the probability that a particular price differential is present, which greatly reduces the probability estimates of a profit. However, this is to be expected given that one is assuming that the opportunities for a transaction are relatively scarce. As one will observe in the following analysis, this scarcity becomes more obvious for stricter shipment policies.

In order to estimate the probabilities of profits at any moment in time, one needs to include the probability that the price differential of the markets will fall within a particular price interval. In essence, this converts the probabilities into estimates of the joint probability function with respect to the duration and the

89

non-lagged price differential (Figure 4-17). Nonetheless, the estimates of the joint probability function are generated in the same manner as before, only that it now includes the probability of the differential.

From a visual inspection of the grid below, one can conclude that the probability estimates are greatly reduced, given that it now includes the probability that the price differential will fall within each interval (the peak point observed in the middle of the graph is considered an outlier). Also, the grid suggests that the probabilities of encountering an opportunity within the highest differential interval are very small in the Dallas-Boston market structure for tomato. This would further confirm that the operations within the structure would be done intermittently; only when an opportunity is identified.
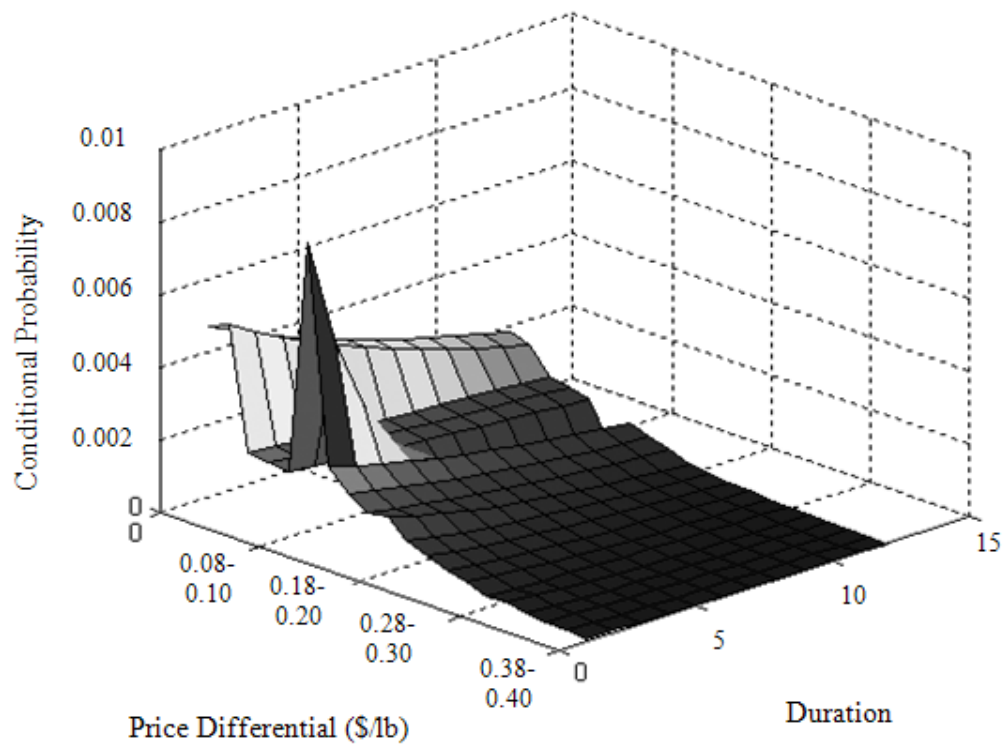
Figure 4-17 – Joint Probability Function Grid

The next step is to calculate the estimates of these probabilities, which is done in the same manner as before. In order to estimate the probabilities of a positive profit at any instance in time, the duration values after the transaction period are aggregated. On the other hand, for a negative profit probability, the duration values before the transaction period are aggregated. Table 4-12 presents the results of these calculations.

Table 4-12 – Probabilities of Profit at any Time

K-Intervals	Positive	Negative
0.00 - 0.02	0.032	0.010
0.02 - 0.04	0.031	0.010
0.04 - 0.06	0.016	0.004
0.06 - 0.08	0.015	0.004
0.08 - 0.10	0.014	0.004
0.10 - 0.12	0.009	0.014
0.12 - 0.14	0.013	0.004
0.14 - 0.16	0.010	0.003
0.16 - 0.18	0.008	0.003
0.18 - 0.20	0.006	0.003
0.20 - 0.22	0.005	0.002
0.22 - 0.24	0.004	0.002
0.24 - 0.26	0.003	0.001
0.26 - 0.28	0.002	0.001
0.28 - 0.30	0.001	0.001
0.30 - 0.32	0.001	0.001
0.32 - 0.34	0.001	0.000
0.34 - 0.36	0.001	0.000
0.36 - 0.38	0.000	0.000
0.38 - 0.40	0.000	0.000

Based on the results presented in the table above, one can conclude that the decision-maker cannot base his/her operations on the general probabilities of the market. He/she has to condition the operations based on the price differentials; a common theme throughout this thesis. Lastly, the summary of these probabilities can be used for making short-term projections of profit within the structures. A possible extension to the estimates that were generated is the application of a binomial lattice structure. However, given the resource limitations, this is regarded as a potential research opportunity.

4.4.6 Conclusions

This section presented the application of the methodology developed for identifying and optimizing long-run expected profits of the individual shipments, while also reducing the level of associated risk. For this purpose, two approaches were proposed and applied to the dataset of price information. It was found that there exist particular shipment strategies that do maximize the level of return per shipment. Furthermore, one could use the results of these approaches to identify shipment strategies that can meet other kind of objectives, such as that of maximizing the total earnings over the entire operational period or minimizing the VaR.

The second part of the shipment strategy involved the development of a methodology that would help analyze short-term market opportunities. In this case, the purpose was to estimate the probability of positive and negative profits based on particular conditions of the market. It was further concluded that the probabilities of a positive profit are greatly enhanced if the shipment decisions are conditioned on particular price differentials.

5. DEVELOPMENT OF COMMERCIALIZATION BASKETS

The last phase of the operations involves the development of a shipment configuration that limits the risk exposure of the decision-maker. In this case, the market price characteristics of each individual component (under a defined shipment strategy) are used to hedge the risk. It is assumed that one can apply mean-value portfolio theory to the collection of shipment components in such a way that one can manipulate the overall rate of return and variance. Ultimately, the objective is to determine an optimal shipment configuration that minimizes the variability of the returns for a particular component.

For demonstration purposes, it is assumed that one wants to limit the risk exposure of tomato shipments from Dallas to Boston. It is also assumed that the general objective of the decision-maker is to maximize his/her long-term profits of the shipments during a defined operational period. Thus, the market price information for the rest of the shipment components is limited to the opportunity time windows of tomato within the Dallas-Boston structure. The main objective is to reduce the variability of the rates of return for tomato shipments from Dallas to Boston.

For this purpose, two approaches are considered. One in which the profit variability of the shipment is reduced by using a product mix to the same market (Dallas-Boston). In this case, one sends a strategic combination of fresh produce items to Boston. Another option explored is one in which tomato shipments are sent to different secondary markets, in such a way that the variability of the returns is also reduced.

94

The following sections details the set-up of the Markowitz model, the method of solving the model, and a general interpretation of the results obtained for each approach.

5.1 Assumptions

As part of the configuration of the shipment container, one is not considering the product's physical characteristics and storage requirements, or the container capacity of approximately 40,000lbs. In practicality, the investor must allocate the individual product investments in such a way that the capacity of the container is not compromised, as well as that the environmental storage requirements are met. Further research may be needed in this area to combine the physical characteristics with cost aspects when determining investment allocations.

5.2 Shipment Configuration Policy: Product Mix

As it was mentioned earlier, the objective of the decision-maker is to maximize the long-term profits of the operation. For this purpose, the solution that was obtained previously from the theoretical approach is used as the optimal shipment criteria ($K>0.05605$). Under the first approach, it is assumed that whenever a tomato shipment is triggered, a product configuration mix can be created that minimizes the overall return variability (Figure 5-1).
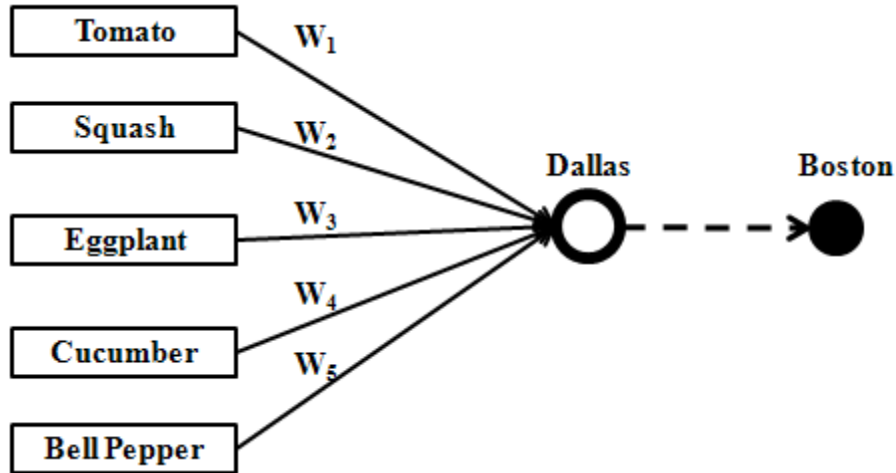
Figure 5-1 – Product Mix Configuration

In this process, the first step is to obtain the necessary components of the Markowitz model set-up. Then, one can solve the model by matrix manipulation, and arrive at an optimal solution. The solution obtained represents the weight assignation given to each component of a particular shipment. One should keep in mind that this weight assignation is not in terms of physical mass, but rather in terms of the investment, or cost, assigned to each individual component of the shipment.

5.2.1 Rates of Return and Covariance: Product Mix

Table 5-1 summarizes the average rates of return that is observed for the rest of the fresh produce items during the time windows of opportunity of tomato in Boston. As one can observe, the rates of return for the rest of the products are negative, which may be an alarming sign (as expected, the only positive rate of return is observed for tomato since the shipment strategy is based on this item). However, as it will be demonstrated shortly, the Markowitz model allows the user

96

to force the expected average rate of return of the total shipment to a desired target value.

Table 5-1 – Product Mix Configuration: Rates of Return

Product	Dallas - Boston
Tomato	0.114
Squash	-0.264
Eggplant	-0.078
Cucumber	-0.146
Bell Pepper	-0.357

Table 5-2 details the correlation matrix of the rates of return during the time windows of opportunity for tomato. As one can observe the correlation between the products is not extremely high. The highest level of correlation with tomato observed is with squash, which approximates 0.128, while the highest correlation is between cucumber and eggplant (0.237). As a result, these low correlation levels decrease the amount by which the variability of the returns can be reduced.

Table 5-2 – Product Mix Configuration: Correlation Matrix

	Tomato	Squash	Eggplant	Cucumber	Bell Pepper
Tomato	1.000	0.128	-0.060	0.044	0.069
Squash	0.128	1.000	0.062	0.178	-0.005
Eggplant	-0.060	0.062	1.000	0.237	-0.080
Cucumber	0.044	0.178	0.237	1.000	0.116
Bell Pepper	0.069	-0.005	-0.080	0.116	1.000

Table 5-3 presents the covariance matrix of the products within the Dallas-Boston market structure. Again, these values represent the covariance of the rates of return during moments of opportunity for tomato. One should note that the diagonal matrix represents to the individual variance of each product under this shipment policy.

Table 5-3 – Product Mix Configuration: Covariance Matrix

	Tomato	Squash	Eggplant	Cucumber	Bell Pepper
Tomato	0.035	0.003	-0.004	0.002	0.002
Squash	0.003	0.016	0.003	0.006	0.000
Eggplant	-0.004	0.003	0.143	0.024	-0.005
Cucumber	0.002	0.006	0.024	0.071	0.005
Bell Pepper	0.002	0.000	-0.005	0.005	0.031

The values obtained from the covariance matrix are used as part of the process for solving the Markowitz model.

5.2.2 Solution for the Markowitz Model: Product Mix for Dallas-Boston

The next step in determining the optimal product mix is to set up the Markowitz model using the different components; covariance matrix, rates of return, and target solution. In order to solve the model and arrive at a solution, the methodology developed in section 3.6.3 is used.

In reference to the application of the methodology, each component of the overall matrix model is presented. Matrix B represents the covariance matrix of the five product components (order is the same as that of Table 5-3), while Matrix C is composed of two rows; one is the average rate of return per product and the other a row of 1's. Lastly, Matrix D is the negative transpose of C, while E is a 2x2 zero matrix.

$$B = \begin{bmatrix} 0.035 & 0.003 & -0.004 & 0.002 & 0.002 \\ 0.003 & 0.016 & 0.003 & 0.006 & 0.000 \\ -0.004 & 0.003 & 0.143 & 0.024 & -0.005 \\ 0.002 & 0.006 & 0.024 & 0.070 & 0.005 \\ 0.002 & 0.001 & -0.005 & 0.005 & 0.031 \end{bmatrix}$$

Eqn. 5-1

$$C = \begin{bmatrix} 0.114 & -0.264 & -0.078 & -0.146 & -0.357 \\ 1 & 1 & 1 & 1 & 1 \end{bmatrix}$$

Eqn. 5-2

$$D = \begin{bmatrix} -0.114 & -1 \\ 0.264 & -1 \\ 0.078 & -1 \\ 0.145 & -1 \\ 0.357 & -1 \end{bmatrix}$$

Eqn. 5-3

Matrix A is a combination of these four matrix and is represented by the following.

$$A = \begin{bmatrix} B & D \\ C & E \end{bmatrix}$$

Eqn. 5-4

As one can observe, the upper-left hand quadrant represents the covariance matrix of the secondary markets, while the lower-left hand quadrant details their individual average rate of return. Lastly, in the upper-right hand quadrant are the transposed negative values of the lower-left hand quadrant, and the last quadrant is a simple 2x2 zero matrix.

$$Y = \begin{bmatrix} 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0.114 \\ 1 \end{bmatrix}$$

Eqn. 5-5

Matrix Y represents the target solution of the Markowitz model. In this case, the sixth row represents the target average rate of return for the tomato shipment and forces the solution of the weighted rates of return to produce this value. Given that the objective is to maintain the profits of a tomato shipment, while reducing the return variance, the target is set to equal the tomato's average rate of return.

Once the different components of the matrix model version are obtained, the next step is to solve for the solution matrix X. The following details the matrix manipulation process:

$$AX = Y \qquad \text{Eqn. 3-18}$$

$$X = A^{-1}Y \qquad \text{Eqn. 3-19}$$

In the solution matrix X, the first five rows represent the weight assignation given to each of the products in the shipment. The last two rows are the values of the Lagrangian relaxation parameters for each constraint, target rate of return and total weight assignation.

$$X = \begin{bmatrix} w_1 \\ w_2 \\ w_3 \\ w_4 \\ w_5 \\ \lambda \\ \mu \end{bmatrix} = \begin{bmatrix} 0.874 \\ 0.028 \\ 0.119 \\ 0.112 \\ -0.133 \\ 0.068 \\ 0.022 \end{bmatrix} \qquad \text{Eqn. 5-6}$$

The mathematical solution to the Markowitz model suggests that it is feasible to send a product mix that can maintain the average rate of return of tomato shipment to Boston, while also reduce the expected variance. Table 5-4 details the variance of the rates of return before and after configuring a shipment product mix.

Table 5-4 – Variance Before/After Product Mix (Dal-Bos)

	Variance without Mix (only tomato)	Variance with Mix (Shipment Configuration)
Total	0.0346	0.0297

As one can observe from the table above, the overall variance of a tomato shipment from the base to the secondary market is reduced by creating a product

mix.   In this case, one can expect a variance reduction in the rates of approximately 14%.

5.2.3   Interpretation of the Results: Product Mix

The mathematical solution presented previously suggests that it is mathematically feasible to reduce the overall variance of the rate of returns for a tomato shipment and still keep a satisfactory level of return.   However, it is important to detail the interpretation of the results under a meaningful, real-world application.

Table 5-5 summarizes the weight assignation based on solution matrix X (Equation 5-8). As it was mentioned previously, the weight assignation is based on the investment, or cost, per fresh produce item and not on the physical characteristics of the products.   Based on these results, the highest investment assignation is given to tomato while the lowest is given to squash.

Table 5-5 – Container Weight Assignation per Product

Product	Weight Assignation
Tomato	0.874
Squash	0.028
Eggplant	0.119
Cucumber	0.112
Bell Pepper	-0.133

In the case of bell pepper, a negative investment assignation is given. Under a financial environment, this would mean acquiring a short position on that particular instrument of the portfolio for the decision-maker. In the application of this thesis, it has a similar interpretation. This means that the item is not owned

by the decision-maker but is borrowed from a broker in Boston. Thus, the transaction is made under current market conditions. Then, when the product arrives to the secondary market it is repaid (returned) to the broker. This represents a short position of the farmer with respect to bell pepper.

Table 5-6 – Reconfigured Weight Assignation per Product

Product	Optimal Weight Assignation	New Weight Assignation
Tomato	0.874	0.690
Squash	0.028	0.023
Eggplant	0.119	0.094
Cucumber	0.112	0.088
Bell Pepper	0.133	0.105
Total	**1.266**	**1.000**

Nonetheless, the product does take physical space within the shipment container. This is why in reality this weight assignation is a positive value as it is shown in Table 5-6. By doing this, however, the weight assignation restriction is violated and thus, one must reconfigure the shipment in order to meet it. The new assignation is given by far-right column, after considering the total weight restriction.

5.3   Shipment Configuration Policy: Market Configuration

The second approach explored uses the behavior of secondary markets as a way to reduce the variability of the rates of return. In this case, the approach of determining the optimal shipment configuration is very similar to that of before. However, in this case the shipment components are not necessarily the products, but rather the rest of the secondary markets. This means that whenever an

opportunity is identified for a tomato shipment in Boston, a shipment of the same product is also sent to the rest of the secondary markets (Figure 5-2).
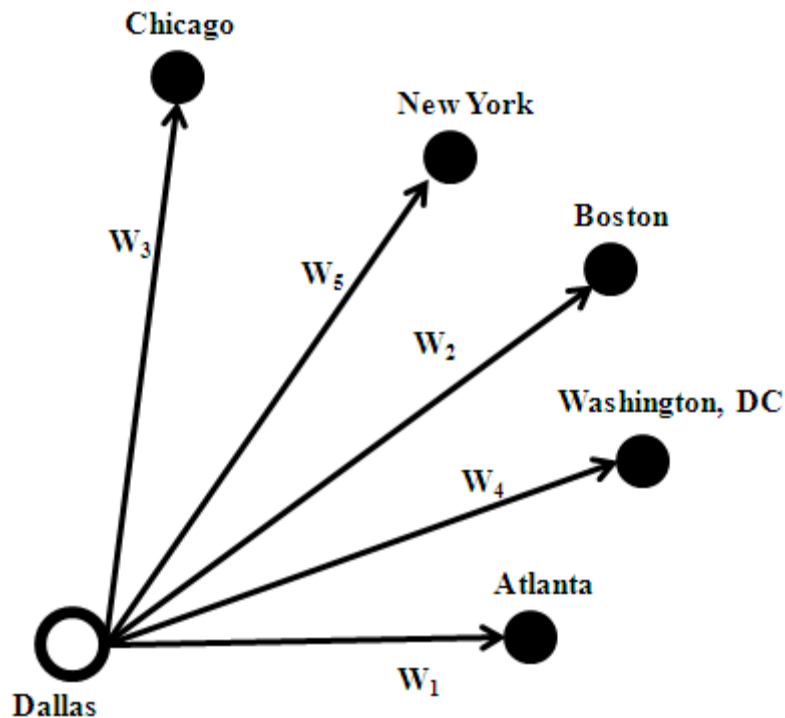
Figure 5-2 – Secondary Market Configuration

Again, the objective is to use the market characteristics within each secondary market in order to strategically invest in each for the sake of minimizing the variability observed in the returns.

5.3.1 Estimating Rates of Return: Market Configuration

For this approach, the objective is to hedge the risk of sending a tomato container to Boston, by also sending a shipment of the same product to the other secondary markets. The shipment policy used is that for tomato within the Dallas-Boston structure ($K>0.05605$), which optimizes the long-term profits for this item during the operational period. Thus, the information of rates of return

correlation and covariance is limited to the time windows of opportunity of this product.

Table 5-7 – Secondary Market Configuration: Average Rate of Return

Two-Market Structure	Rate of Return (Tomato)
Dal – Atl	0.060
Dal – Bos	0.114
Dal – Chi	0.045
Dal – DC	0.064
Dal – NY	-0.077

Table 5-7 presents the rates of return of tomato per secondary market under this shipment policy. As one can observe, the highest average rate is observed in Boston, while the lowest details a negative rate in New York. Furthermore, Table 5-8 summarizes the correlation levels observed between each secondary market, which are much higher than those observed in the product mix. In this case, the highest level of correlation with Boston is New York, which approximates 0.378. Overall, Atlanta and DC have the highest level of correlation with 0.571.

Table 5-8 – Secondary Market Configuration: Correlation Matrix

	Dal – Atl	Dal – Bos	Dal – Chi	Dal – DC	Dal – NY
Dal – Atl	1.000	0.340	0.359	0.571	0.503
Dal – Bos	0.340	1.000	0.209	0.260	0.378
Dal – Chi	0.359	0.209	1.000	0.241	0.350
Dal – DC	0.571	0.260	0.241	1.000	0.304
Dal – NY	0.503	0.378	0.350	0.304	1.000

Lastly, Table 5-9 summarizes the covariance matrix of the rates of return between each two-market structure. Again, the covariance is only during time

windows of opportunity in Boston. The diagonal matrix represents the individual

covariance per secondary market.

Table 5-9 – Secondary Market Configuration: Covariance Matrix

	Dal – Atl	Dal – Bos	Dal – Chi	Dal – DC	Dal – NY
Dal – Atl	0.024	0.010	0.012	0.013	0.016
Dal – Bos	0.010	0.035	0.008	0.007	0.014
Dal – Chi	0.012	0.008	0.043	0.007	0.014
Dal – DC	0.013	0.007	0.007	0.020	0.009
Dal – NY	0.016	0.014	0.014	0.009	0.039

The estimated values of the rates of return and the covariance are used for

the application of the Markowitz model in determining the optimal shipment

configuration. The purpose is to reduce the variance of the returns observed in

Boston.

5.3.2 Solution for the Markowitz Model: Market Configuration

Similar to the method used to find the optimal product mix, the Markowitz

model is solved in order to determine the optimal secondary market configuration.

In this case, the objective is to determine the investment weight assignation given

to each secondary market in a manner in which the rate of return variability of a

tomato shipment to Boston is reduced. The model is solved by simple matrix

manipulation using the components obtained in the previous section.

In reference to the application of section 3.6.3, the different components of

the Markowitz model is translated into matrix form. Matrix B represents the

covariance matrix of the five secondary markets, while Matrix C has the average

rates of return, as well as 1's row representing the total weight. Matrix D and E

are a transposed negative version of C and a 2x2 zero matrix, respectively.

$$B = \begin{bmatrix} 0.024 & 0.010 & 0.012 & 0.013 & 0.016 \\ 0.010 & 0.035 & 0.008 & 0.007 & 0.014 \\ 0.012 & 0.008 & 0.043 & 0.007 & 0.014 \\ 0.013 & 0.007 & 0.007 & 0.020 & 0.009 \\ 0.016 & 0.014 & 0.014 & 0.009 & 0.039 \end{bmatrix}$$ **Eqn. 5-7**

$$C = \begin{bmatrix} 0.060 & 0.114 & 0.045 & 0.064 & -0.767 \\ 1 & 1 & 1 & 1 & 1 \end{bmatrix}$$ **Eqn. 5-8**

$$D = \begin{bmatrix} -0.060 & -1 \\ -0.114 & -1 \\ -0.045 & -1 \\ -0.064 & -1 \\ 0.767 & -1 \end{bmatrix}$$ **Eqn. 5-9**

Matrix A summarizes these matrices into a form useful for solving the Markowitz model.

$$A = \begin{bmatrix} B & D \\ C & E \end{bmatrix}$$ **Eqn. 5-10**

In Matrix A, the upper-left hand quadrant represents the covariance matrix of the secondary markets, while the lower-left hand quadrant details their individual average rate of return. Lastly, in the upper-right hand quadrant are the negative values of the lower-left hand quadrant, and the last quadrant is a simple 2x2 zero matrix.

$$Y = \begin{bmatrix} 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0.114 \\ 1 \end{bmatrix}$$ **Eqn. 5-11**

Matrix Y represents the target solution of the Markowitz model. In this case, the sixth row represents the target average rate of return. Again, the target is

set as the average rate of return for the optimal tomato shipment policy within the Dallas – Boston market structure.

Simple matrix manipulation is used to arrive at an optimal solution. The following details the method of solving the model explained in section 3.6.3:

$$AX = Y \qquad\qquad \textbf{Eqn. 3-18}$$

$$X = A^{-1}Y \qquad\qquad \textbf{Eqn. 3-19}$$

Based on the matrix solution represented above, the solution is given by matrix X (Equation 5-14). In this matrix, the first five rows represent the weight assignation given to each secondary market, while the last two rows are the values of the Lagrangian relaxation parameters for the target average rate of return and total weight assignation.

$$X = \begin{bmatrix} w_1 \\ w_2 \\ w_3 \\ w_4 \\ w_5 \\ \lambda \\ \mu \end{bmatrix} = \begin{bmatrix} 0.250 \\ 0.387 \\ 0.166 \\ 0.443 \\ -0.246 \\ 0.057 \\ 0.010 \end{bmatrix} \qquad\qquad \textbf{Eqn. 5-12}$$

Based on the mathematical solution of the Markowitz model, it was found feasible to reduce the return variance for tomato shipments to the Boston market. In this case, the configuration is created based on different secondary markets rather than a product mix. Nonetheless, the results are still found to be satisfactory. Table 5-10 presents the variance of the rates of return before and after the development of a market configuration.

Table 5-10 – Variance Before/After Market Configuration

	Variance (only Dallas-Boston)	Variance (Market Configuration)
Total	0.0346	0.0168

As one can observe from the table above, the overall variance of the shipment in the Dallas-Boston structure is reduced by combining operations with other secondary markets. It is observed that the variance of the rates of return is reduced by 52%, even further than for the product mix.

5.3.3 Market Configuration: Interpretation of the Results

Again, it is important to determine the real-life meaning of the mathematical solution to the Markowitz model. In this case, it was found that one can reduce the variability of the rates of return by configuring the shipments based on different secondary markets for the same product. However, this has important implications in the real-life operational strategies.

Table 5-11 presents the investment weight assignation given to each secondary market. As it was mentioned earlier, this weight does not refer to the physical properties of the product, but to the investment, or cost, assigned to each secondary market. As one can observe, the highest assignation is given to Washington, DC., followed closely by Boston.

Table 5-11 – Container Weight Assignation per Secondary Market

Two-Market Structure	Weight Assignation
Dal – Atl	0.250
Dal – Bos	0.387
Dal – Chi	0.166
Dal – DC	0.443
Dal – NY	-0.246

In this case, New York is given a negative weight assignation, which means that one would short this component. In a real-world application, this has the same meaning as before, the product is borrowed from a broker at New York, and the product is delivered 3 days later. In this case, the transactions are made based on current market price conditions.

However, this product does have a physical presence and must be considered in the container shipment. Table 5-12 presents the weight assignation accounting a positive value for New York. Then, the weight assignation is reconfigured in order to meet the weight assignation restriction. The new investment configuration is represented by the far-right column of the table. Again, the highest assignations are given to Boston and the DC area.

Table 5-12 – Reconfigured Weight Assignation per Secondary Market

Secondary Market	Optimal Weight Assignation	New Weight Assignation
Dal – Atl	0.250	0.167
Dal – Bos	0.387	0.259
Dal – Chi	0.166	0.111
Dal – DC	0.443	0.297
Dal – NY	0.246	0.165
Total	**1.491**	**1.000**

The real-life application of this shipment strategy would probably not consider sending partial shipments to each of the secondary markets. In this case, a probable solution would be to borrow space within another entity's transportation fleet whenever a chance is presented. One could even think of reducing the number of secondary markets to less than five, in order to facilitate the operations. Whatever the practical solution may be, based on the results of this thesis, it is found plausible to reduce the variability of returns by configuring shipments to different secondary markets.

5.4   Final Observations

Based on the results obtained by the application of the methodology, one can conclude that the application of the Markowitz model approach can reduce the variability of the profits of a particular two-market operation. In this case, it was observed that if the main target is the Boston market, one can reduce the variability by either configuring an optimal product mix or by sending concurrent shipments to other secondary markets.

The implementation scope of this strategy could further be expanded to include other secondary markets and products. However, given the constraint on

the data availability and time, the analysis was limited to this set of products and markets. Future research is needed to expand on the results attained through this methodology in application for a more extensive list of products and secondary markets.

6. CONCLUSIONS AND FUTURE RESEARCH

This chapter highlights the most important analysis, results, conclusions, future research and recommendations made for this thesis in the modeling and development of the proposed commercialization strategy. A summary of each is given through the following sections.

6.1   Thesis Summary and Conclusions

The main premise of this thesis is based on the development of a new operational strategy that increases the commercialization reach of local, mid-level farmers into secondary markets. In this case, the farmer is constrained by the amount of capital investment that can be allocated for the start-up of a commercialization expansion. It was evident by a literature review in the subject that a commercialization strategy based on limited capital resources was needed. This thesis looked to expand on this issue by creating an intelligent, operational strategy that aims to extend the commercialization reach of mid-level farmers, while limiting the capital requirements and increasing the overall profits.

The main strategy was to take advantage of the variability that exists within the markets of the fresh produce industry in order to profit from the arbitrage opportunities that arise from the price differentials of two markets. It was found that in fact, one could profit from operating within a two-market structure, if the shipments are done in an intelligent and methodical manner. Furthermore, it was determined that by varying the shipment criterion, one could actually increase the expected profits and rates of return per shipment, while decreasing the potential for losses.

112

The strategy proposed by this thesis is robust enough to be applied to any two-market structure. The strategy was composed of two main operational phases; (1) a shipment strategy that attempts to increase the expected profits of the shipments while limiting the risk exposure, and (2) a product basket strategy that reduces the overall risk exposure of the shipment. The methodology developed combines a mixture of both pragmatic and theoretic approaches to develop decision-making tools. One must note, however, that subjective reasoning is needed on the part of the decision-maker to further improve on the selection of an appropriate operational strategy.

The results of the first phase of the methodology suggest that in fact one could benefit from engaging in operations within the two-market structure. Following a pragmatic approach, one can deduce that the total earnings of the operations in the long run increase under a loose shipment criterion. This would be in the case that a shipment is made every time a positive price differential is identified. On the other hand, once the shipment criterion becomes stricter the overall total earnings of the operations tend to drop. This in turn causes the rate of returns per shipment to increase and the risk potential for a loss to decrease.

A theoretic approach was also tested on the data set. In this case, the methodology involves the development of an expected profit model based on the distributions of the lagged and non-lagged price differential. This model was then differentiated with respect to the decision factor, K, in order to optimize the function. The main objective was to maximize the expected profits of the total

shipments. It was found that the results of the rates of return converged to similar values of that of the pragmatic approach.

The second phase of the operation involves the development of shipment groupings that could reduce the overall risk exposure of the decision-maker. For this, the characteristics of the price differential under each selected shipment strategy were used in order to hedge against the market changes that are frequent within the fresh produce industry. Financial engineering tools were used to achieve this, most specifically a Markowitz model approach.

The first step was to select a product for commercialization within a particular two-market structure (tomato within the Dallas-Boston structure). The objective was to reduce the overall variance of that shipment by defining an adequate shipment policy. For this purpose, two approaches were considered; one in which the variability of the returns was reduced by sending multiple products to same secondary market, and another in which the same product is sent to different secondary markets. It was determined that one could reduce the overall profit variability of a tomato shipment to Boston by implementing both strategies, creating a product mix or a market configuration. As it was concluded in a previous section, further research is needed to further test the adequacy of the strategy with a wider range of fresh products and secondary markets.

6.2   Thesis Contributions

The main contribution of this thesis is with respect to the development of an operational strategy for the commercialization expansion of mid-level farmers. The methodology developed could represent the first step in related studies,

especially for those entities who aim to increase their presence within secondary markets, but at the same time, also want to limit their risk exposure. The tools developed by the present study are meant to be utilized under practical circumstances and must be refined to fit the circumstances through which they will be used.

The case study through which the methodology was tested derives from a similar situation. This study aims to expand on the tools that were originally used for previous studies and apply them under different market conditions. This thesis addresses a subject area that may prove very beneficial, if one were to refine the tools and apply the methodology under a real-world scenario.

6.3   Recommendations for Future Research

The scope of the research topic addressed through this thesis covers different aspects of a specific operational strategy. This section will address five main areas that have high relevancy for additional studies and extensions in the future. These potential area of improvement include the development of a more robust optimization model that can incorporate additional components of the market characteristics when determining a shipment strategy, more accurate market price information including retail data, backhaul utilization of the two-market structure, more in-depth return predictions, and the development of an inventory policy model.

6.3.1   Additional Components of the Optimization Model

The methodology developed through this thesis focuses on determining levels of opportunities based on the basic characteristics of the markets.

115

However, additional components of the operations, such as varying transaction costs, the time dependency of the product value associated with the perishable characteristics and the seasonality factors in the product variations may help refine the results obtained.

The incorporation of variability in the transaction costs is a relevant topic in current research with respect to market integration of perishable items, and it is one of the main areas for improvement in future research. Given the limitation on reliable data for current and historic transaction costs, a time-weighted value based on the Consumer Price Index was used instead. Furthermore, the theoretic approach developed by this thesis is based on the acquisition and transaction cost of the items. It is strongly believed that more reliable data regarding this value would improve the accuracy of the results attained in the analysis.

Another component that could increase the resolution of the results is the addition of seasonality factors for identifying the opportunities, which would change the manner in which the price movements and trends are summarized. This would involve the addition of time-series analysis into the expected profit model and create a more complex identification process. Furthermore, given the large size of the database collected, data mining techniques could be used to identify additional relationships in these price movements.

Finally, the incorporation of the product value as a function of its perishable characteristics might be a helpful addition to the evaluation of the expected profit model. This would refine the decision-making process with regards to the perishable conditions of the individual products. However, this

would most definitely create a more complex optimization model and probably need different tools to apply them.

6.3.2 Retail Price Information and Behavior

Another area of opportunity for research is the application of the methodology developed into additional market environments. This includes the behavior of product prices higher up the supply chain and its effect on the potential returns of a more detailed operation. The methodology developed by this thesis would need to be modified to include additional components and assumptions. Consideration must be given to the increased number of echelons in the supply chain, as well as the liquidity of the products at the secondary markets.

6.3.3 Backhaul Utilization

An addition to the proposed operational strategy would be the utilization of the backhaul price relationships of the products. In this case, one could design a two-way operational system that can take advantage of the arbitrage opportunities that exist not only in the long haul of the shipments, but also reap the benefits of opportunities that might exist on the backhaul. This would mean a more complex environment and additional assumptions regarding the operation.

6.3.4 Binomial Lattice Application

An addition to the short-term projections of the available opportunities would be the incorporation of binomial lattice structures. The purpose of these structures would aid the decision-maker to predict long term fluctuations of the market prices, especially the probabilities for a profit and a loss. However, given that the operations are assumed to be intermittent, the application of these

structures is not as simple as they are under normal usage within a financial scope of a binomial lattice.
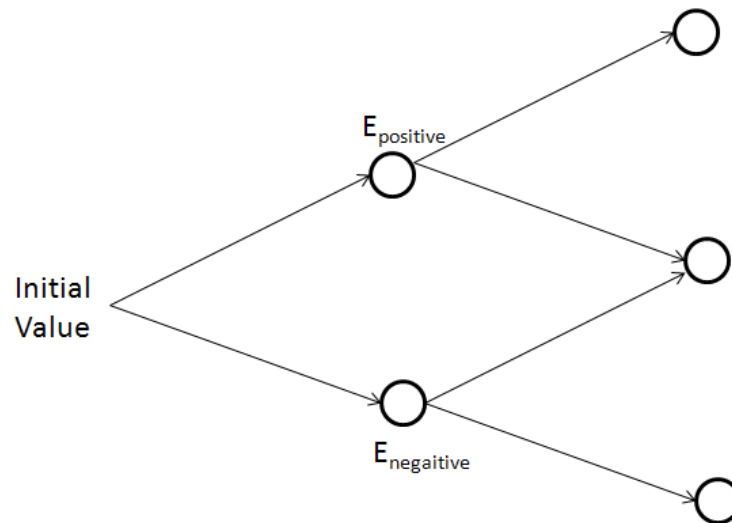
Figure 6-1 – Binomial Lattice

Figure 6-1 presents an example of a binomial lattice structure that can be created based on the projections of profit and loss. Each node represents a potential outcome of a series of prior events. Also, the lattice details the expectation of profit based on the probability of a particular event.

6.3.5 Theoretic Inventory Policy Model

Lastly, the development of an inventory policy that considers the probabilistic nature and intermittency of the opportunities is a promising topic for future research. In this case, one could use a newsvendor approach in determining the adequate inventory levels to maintain. However, the difference would be in that for the proposed strategy, the product demand is based on the availability of the opportunities. Thus, one has to develop a cost model in terms

of the probabilistic nature of the market price differentials, as well as determine a linkage between the opportunities and actual quantity levels of inventory.

Among the factors that can be considered for the development of this model is the time-dependent cost of inventory based on the perishable characteristics of the individual products. Additionally, the costs would also include the loss of a potential opportunity due to lack of product availability.

REFERENCES

Baulch, Bob. 1997. "Transfer Costs, Spatial Arbitrage, and Testing for Food Market Integration." *American Journal of Agricultural Economics* 79 (2) (May): 447-487.

Blank, Steven C., and Brian H. Schimiesing. 1988. "Modeling of Agricultural Markets and Prices Using Causality and Path Analysis." *North Central Journal of Agricultural Economics* 10 (1) (January): 35-48.

Burnetas, Apostolos N., and Craig E. Smith. 2000. "Adaptive Ordering and Pricing for Perishable Products." *INFORMS* 48 (3) (June): 436-443.

Choi, Tsang-Ming, Duan Li, and Houmin Yan. 2008. "Mean-Variance Analysis for the Newsvendor Problem." *Systems, Man and Cybernetics, Part A: Systems and Humans* 38 (5) (September): 1169-1180.

Eeckhoudt, Louis, Christian Gollier, and Harris Schlesinger. 1995. "The Risk-Averse (and Prudent) Newsboy." *Management Science* 41 (5) (May): 786-794.

Engle, Robert F., and C. W. J. Granger. 1987. "Co-Integration and Error Correction: Representatio, Estimation, and Testing." *Econometrica* 55 (2) (March): 251-276.

Granger, C.W.J. 1969. "Investigating Causal Relations by Econometric Models and Cross-Spectral Methods." *Econometrica* 37 (3) (July): 424-438.

Granger, C.W.J. 1981. "Some Properties of Time Series Data and their Use in Econometric Model Specification." *Journal of Econometrics*: 121-130.

Khouja, Moutaz. 1999. "The Single-Period (news-vendor) Problem: Literature Review and Suggestions for Future Research." *The International Journal of Management Science* 27 (5) (October): 537-553.

Liang, Chyi-lyi, Dillon M. Feuz, and R. G. Taylor. 1997. Cointegration Tests of Spatial and Variety Price Linkages In Regional Dry Bean Markets. In Reno/Sparks, NV, July.

Lin, Dan. 2007. Exploration of Role of Market in Perishable Goods. The University of Texas at Austin, May.

McNew, Kevin. 1996. "Spatial Market Integration: Definition, Theory, and Evidence." *Agricultural and Resource Economi Review* 25 (1) (April).

Padilla-Bernal, Luz, Dawn D. Thilmany, and Maria L. Loureiro. 2003. "An Empirical Analysis of Market Integration an Efficiency for U.S. Fresh Tomatoes." *Journal of Agricultural and Resource Economics* 28 (3) (December): 435-450.

Rudi, Nils, Sandeep Kapur, and David Pyke F. 2001. "A Two-Location Inventory Model with Transshipment and Local Decision Making." *INFORMS* 47 (No. 12) (December): 1668-1680.

Sanchez, Octavio. 2007. Strategic Design of a Logisitics Platform for Fresh Produce. Tempe, AZ: Arizona State University, August.

Serra, Teresa, Barry K. Goodwin, Jose M. Gil, and Anthony Mancuso. 2006. "Non-Parametric Modelling of Spatial Price Relationships." *Journal of Agricultural Economics* 57 (3) (September): 501-522.

Wright, Sewall. 1921. "Correlation and Causation." *Journal of Agricultural Research* 20 (7): 557-582.

Ziemer, Rod F., and Glenn S. Collin. 1984. "Granger Causality and U.S. Crop and Livestock Prices." *Southern Journal of Agricultural Economics* 16 (July): 115-120.

APPENDIX A

ADDITIONAL COMPONENTS OF THE SHIPMENT STRATEGY

There exists additional components of any given shipment strategy that may influence the decision making process. Among these components are the frequency of shipments, total earnings and the relative risk of engaging in a two-market trade given particular market conditions. While the frequency of the shipments and total earnings are values that can be readily recorded, the associated risk with a particular shipment strategy involves more subjectivity. The following details the tools that are used in this thesis to estimate the levels of risk.

From a financial perspective, it is generally known that as the magnitude of the potential profits of an investment is increased, the underlying risk is also likely to increase. A common example is a traditional financial instrument, such as a bond, stocks, option, etc. In general, the advantage of investing in riskier financial instruments is that the potential payoffs of the investment are generally higher. However, the main disadvantage is that the general variability that creates the high payoffs in the first place can also result in higher level of losses.

It is believed that a similar trend occurs in the proposed strategy of this thesis. One would think that as the expected profits of the operation increase, the underlying risk will also increase, given the general tendency of financial securities. So the next question becomes how one quantifies this risk, in a way that the potential losses per shipment strategy can be analyzed and compared. For this, similar risk analysis tools that are commonly used within financial markets are applied for purposes of this thesis.

The risk analysis tool that is used to quantify the risk of the different shipment strategies is known as the Value-At-Risk analysis. The general purpose of this analysis is to estimate the amount by which the value of a particular financial portfolio can vary from one period to a next given a particular confidence level. In the context of this thesis, the value-at-risk is associated with the value by which the two-market price differential can drop during a period of opportunity.

There are various ways in which the VaR can be estimated. A basic approach is to use a historic method to quantify this risk. For this method, the first step is to order the historical returns of the shipments from worst to best. Assuming that history repeats itself, one determines the value of the returns that pertains to the 5th percentile of sorted set of returns. This value indicates that with 95% confidence, the returns on the shipment will be no larger than the value at the 5th percentile. The advantage of calculating the VaR through this method is that one does not need to assume a particular theoretic distribution to the returns.

Another approach is to use a procedure known as the Variance-Covariance Method. This method involves fewer steps, if one can summarize the behavior of the returns through a distribution fit. For this step, only the probability distribution parameters of the returns are needed to estimate the risk, assuming normally distributed profits. In this case, the only statistical parameters needed are the mean and standard deviation per distribution fit. However, this assumption is also its main disadvantage, since other statistical fits could better represent the distribution of the observations.

Conventionally, the VaR is represented as positive value, although it almost always represents a loss. In the case in which the VaR is negative, this would imply that the instrument has a high probability of making a profit in the next period of opportunity. For example, a one-day 5% VaR of negative $1 dollar implies that the product has 95% chance of making more than $1 dollar over the next period.

Overall, the frequency, total earnings and VAR are used to gauge the goodness of each particular strategy. These values are used not only to assess the adequacy of the pragmatic approach in the first part of this section, but will also be used to assess the results obtained from the more theoretic approach. As a whole, these variables will help identify the best strategy given the conditions of the market.

APPENDIX B

BACKGROUND TO MEAN-VALUE PORTFOLIO THEORY

The methodology behind determining an appropriate commercialization basket is based on the mean-variance portfolio theory. The theory is based on a common principle often observed within financial markets; as the potential return for any particular investment increases, so does the risk for a potential loss. In general, the average investor will seek to optimize the balance between the level of return and the potential for loss on a particular investment, based on his/her perception of risk. To accomplish this, a primary objective within financial applications is to design a portfolio composed of a variety of financial instruments that can limit the risk exposure of the investor, but still generate satisfactory levels of return. Next, a general background is presented in order to give the reader a more in-depth look on the dynamics of the mean-variance portfolio theory, and the method (Markowitz problem) used to solve it.
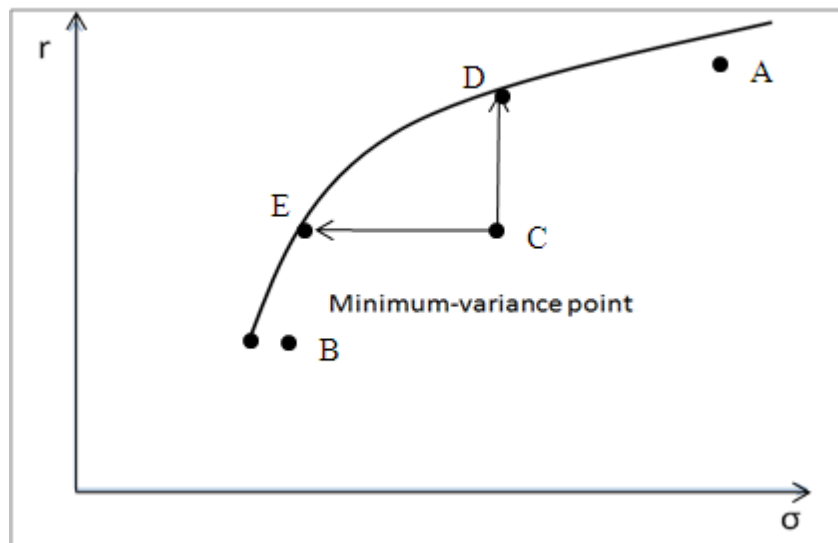

Figure 1 - Risk vs. Variability of General Product Basket

Figure 1 represents an x-space for potential returns and risk associated with any particular collection of products. The curve represents the feasible

frontier of return and loss for any portfolio collection. This curve acts as a feasibility restriction, and it means that one cannot expect extremely high returns on a low-risk financial product, or vice-versa. Also, within this space, point A represents a collection of products that for the most part are high-profit, high-risk investments. On the other hand, point B represents a portfolio composed of low-profit, low-risk products, while point C is a collection that fits right in the middle; probably, a combination of high and low risk products.

As one can observe from these points, the combination of return and risk per portfolio are not necessarily on the feasible frontier. One could re-configure the collection of products within a portfolio in order to move around this feasible space. This suggests that there is an opportunity for the investor to either increase the return fixating the risk (point C to D), or the other way around, which is to reduce the risk and fixate the returns (point C to E). Either way, there is an opportunity to improve the position of the investor, by getting closer to this "efficient" frontier. In general, an investor would seek to reduce the risk and fix the returns.

In finance, a common tool used to calculate the appropriate product configuration that allows the variance reduction detailed above is known as the Markowitz problem. This problem attempts to determine the investment weight assigned to each individual product in such a way that the overall risk of the portfolio is reduced. A similar approach is performed on operational strategy developed in this thesis.

For purposes of this thesis, one assumes that each individual shipment policy can be treated as a financial instrument. Thus, it is assumed that one can use the concepts of Mean-Variance theory and Markowitz problem to determine an optimal configuration of product shipment policies that can limit the risk exposure of general operations.

APPENDIX C

CHI-SQUARE GOODNESS-OF-FIT TEST ON LAGGED DIFFERENTIALS

Goodness of Fit Test: Lagged Differentials	
Market Structure:	Dallas – Boston
Theoretical Distribution:	Logistic
Level of Significance:	0.05
Threshold Level:	$K > 0.30$
Parameters:	
alpha:	0.303
beta:	0.115
Kolmogorov-Smirnoff	
data points :	104
ks stats :	0.0788
alpha :	0.05
ks stat [1224,0.05]:	0.131
p-value:	0.504
result	DO NOT REJECT
Anderson-Darling	
data points :	104
ad stats :	0.887
alpha :	0.05
ad stat [0.05]:	2.49
p-value:	0.422
result	DO NOT REJECT

APPENDIX D

CHI-SQUARE GOODNESS-OF-FIT TEST ON DIFFERENTIAL DURATIONS

Goodness of Fit Test: Durations	
Market Structure:	Dallas – Boston
Theoretical Distribution:	Geometric
Level of Significance:	0.05
Threshold Level:	$0<K<0.02$
Parameters:	
p:	0.374
Kolmogorov-Smirnoff	
data points :	98
ks stat :	0.0923
alpha :	0.05
ks stat [98,0.05]:	0..135
p-value:	0.352
result	DO NOT REJECT

Goodness of Fit Test: Rate of Return	
Market Structure:	Dallas – Boston
Theoretical Distribution:	Johnson SB
Level of Significance:	0.05
Parameters:	
minimum:	0.00337
lambda:	1.487
gamma:	1.278
delta:	0.706
Kolmogorov-Smirnoff	
data points :	1224
ks stats :	0.0361
alpha :	0.05
ks stat [1224,0.05]:	0.0387
p-value:	0.0799
result	DO NOT REJECT
Anderson-Darling	
data points :	1204
ad stats :	1.1
alpha :	0.05
ad stat [0.05]:	2.49
p-value:	0.308
result	DO NOT REJECT

APPENDIX E

MATLAB CODE FOR CALCULATING THRESHOLD VALUE PRAGMATIC

APROACH

```
%% This file follows a pragmatic approach to determining the optimal threshold
%% value
%% Secondary Market Index of the excel that is read
%% 1=Date, 2=Dallas, 3=Atlanta, 4=Boston,5=Chicago,6=Columbia, 7=NY
%% Prices that incorporate Inflation are located in 'Sheet2'

%% Reads the excel files and user-defined input
ln = @log;
ProductPrice=xlsread('2001-2009 Tomato Prices.xlsx','Sheet2');
city=4;

%% Determines the lag and transaction times based on the input
if city == 3;
   lag=2;
   TCost=0.01248;
else if city==4;
     lag=3;
     TCost=0.055185;
   else if city==5;
        lag=2;
        TCost=0.01845;
     else if city==6;
          lag=3;
          TCost=0.029825;
        else if city==7;
             lag=3;
             TCost=0.05589;
          end
        end
     end
   end
end

%% Instantiates arrays
nRows = zeros(size(ProductPrice));
DiffPNoLag = nRows(:,1);
DiffPLag = nRows(:,1);
Opportunity = zeros(size(nRows));

%% Price differences without lag under current market situation
for i=5: size(ProductPrice);
   DiffPNoLag(i)= ProductPrice(i,city)-ProductPrice(i,2)-TCost;
end

%% Price differences with transportation lag
```

```matlab
for i=5: size(ProductPrice)-5;
    DiffPLag(i)= ProductPrice(i,city)-ProductPrice(i-lag,2)-TCost;
end

%% Probability distribution of the non-lagged differentials (current)
PDTD=fitdist(DiffPNoLag,'logistic');
UTD=PDTD.Params(1,1);
STD=PDTD.Params(1,2);

%% Stepsize and range of threshold value
stepsize=0.01;
start=0;
interval=0.02;
F=40;

%% Instantiates arrays used in the iterative process
syms a b c;
CurrDiffWoutLag = zeros(size(DiffPLag));
ActualOpportunityWOpt=zeros(size(nRows,1),1);
CurrentOpportunity=zeros(size(nRows,1),1);
RateOfReturnWOpt=zeros(size(nRows,1),1);
ReturnsWOpt=zeros(size(nRows,1),1);
Data= zeros(size(ProductPrice,1),F);
DataTempActWOpt= zeros(size(ProductPrice,1),F);
DataTempCurWOpt= zeros(size(ProductPrice,1),F);
DataRORWOpt= zeros(size(ProductPrice,1),F);
DataTempReturnsWOpt= zeros(size(ProductPrice,1),F);

%% Start of iterative process
for j=1:F;
    %% Instantiates record keeping variables
    SumNetEarnings=0;
    SumRORWOpt=0;

    count=count+1;
    K=0+j*stepsize;
    K_Values(j)=K;

  for i=5:size(DiffPLag)-lag;
    %% Checks to see if threshold is met
    MktDiff = ProductPrice(i,city)-ProductPrice(i,2)-TCost;

    %% Records the results whenever a shipment is triggered
     if (MktDiff>=K);
        ActualOpportunityWOpt(i)=ProductPrice(i+lag,city)
```

```
         -ProductPrice(i,2)-TCost;
      CurrentOpportunity(i)=MktDiff;
      AmountReceived=ProductPrice(i+lag,city);
      AmountInvested=ProductPrice(i,2)+TCost;
      RateOfReturnWOpt(i)=(AmountReceived-
AmountInvested)/AmountInvested;
      ReturnsWOpt(i)=AmountReceived/AmountInvested-1;
      SumRORWOpt=SumRORWOpt+RateOfReturnWOpt(i);
      SumNetEarnings=SumNetEarnings+ActualOpportunityWOpt(i);
      count=count+1;
    end
  end

  ActOpp = ActualOpportunityWOpt(find(ActualOpportunityWOpt));
  CurOpp = CurrentOpportunity(find(CurrentOpportunity));
  RORWOpt = RateOfReturnWOpt(find(RateOfReturnWOpt));
  RWOpt = ReturnsWOpt(find(ReturnsWOpt));

  %% Estimates the VaR based on the 5th percentile of the returns
  ValueAtRisk_hist(j) = prctile(RWOpt,0.05);

  %% Records the information for rate of returns
  for i=1:size(RORWOpt)
    DataRORWOpt(i,j)=RORWOpt(i);
  end

  NumberOfOpportunitiesWOpt(j)=count;
  AverageRORWOpt(j)=SumRORWOpt/size(RORWOpt,1);
  TotalEarningsWOpt(j)=SumNetEarnings;

  %% Records the information for lagged differentials
  for r=1:size(ActOpp);
    DataTempActWOpt(r,j)=ActOpp(r);
  end

  %% Records the information for non-lagged differentials
  for r=1:size(CurOpp);
    DataTempCurWOpt(r,j)=CurOpp(r);
  end

  %% Records the information for returns
  for r=1:size(CurOpp);
    DataTempReturnsWOpt(r,j)=RWOpt(r);
  end
```

```matlab
  %% Fits a logistic distribution on the lagged differentials
  PDA = fitdist(ActOpp,'logistic');
  UA = PDA.Params(1,1);
  SA = PDA.Params(1,2);

  %% Fits a logistic distribution on the non-lagged differentials
  PDC = fitdist(CurOpp,'logistic');
  UC = PDC.Params(1,1);
  SC = PDC.Params(1,2);

  %% Fits a logistic distribution on the returns
  PDR = fitdist(RWOpt,'normal');
  UR = PDR.Params(1,1);
  SR = PDR.Params(1,2);

  %% Attains esimates of the mean and variance based on the
  %% distribution fits
  meanA(j)=UA;
  varA(j)=SA;

  %% Computes the VaR based on the fit on the lagged differentials
  %% Assumes a 5% confidence interval
  %% VaR is attained using a logistic and normal distribution fits
  %% for comparative purposes
  P_ConfidenceInt=0.05;

  ValueAtRisk_logist(j)=meanA(j)+varA(j)*ln(P_ConfidenceInt/
                      (1-P_ConfidenceInt));
  ValueAtRisk_norm(j) = UR - 1.65*SR;

  %% Resets the arrays for next iteration
  ActualOpportunityWOpt=zeros(size(nRows,1),1);
  CurrentOpportunity=zeros(size(nRows,1),1);
  RateOfReturnWOpt=zeros(size(nRows,1),1);
  ReturnsWOpt = zeros(size(nRows,1),1);

  count=0;
end

%% Transposes the arrays
ValueAtRisk_norm=ValueAtRisk_norm';
ValueAtRisk_logist=ValueAtRisk_logist';
AverageRORWOpt=AverageRORWOpt';
NumberOfOpportunitiesWOpt=NumberOfOpportunitiesWOpt';
```

```
%% Summarizes the results of the pragmatic approach
for r =1:size(ValueAtRisk_norm);
    ResultRORVAR(r,1)=AverageRORWOpt(r);
    ResultRORVAR(r,2)= ValueAtRisk_logist(r);
    ResultRORVAR(r,3)= ValueAtRisk_norm(r);
    ResultRORVAR(r,4)= ValueAtRisk_hist(r);
    ResultRORVAR(r,5)= NumberOfOpportunitiesWOpt(r);
    ResultRORVAR(r,6)= TotalEarningsWOpt(r);
    ResultRORVAR(r,7)= K_Values(r);
    ResultRORVAR(r,8)= meanA(r);
    ResultRORVAR(r,9)= varA(r);
end
% xlswrite('ActualOppK',ResultRORVAR);

%% Sorts the results array by the value of K
ResultRORVAR=sortrows(ResultRORVAR,-7);

%% END OF CODE
```

APPENDIX F

MATLAB CODE FOR CALCULATING DURATIONS

```
%% This code determines the duration of positive price differentials
%% based on specific k intervales
%% Secondary market index
%% 1=Date, 2=Dallas, 3=Atlanta, 4=Boston,5=Chicago,6=Columbia, 7=NY
%% Inflation prices is located in 'Sheet2'

%% Reads the product price excel and user-defined market
ProductPrice=xlsread('2001-2009 Tomato Prices.xlsx','Sheet2');
city=4;

if city == 3;
   lag=2;
   TCost=0.01248;
else if city==4;
     lag=3;
     TCost=0.055185;
   else if city==5;
        lag=2;
        TCost=0.01845;
      else if city==6;
           lag=3;
           TCost=0.029825;
         else if city==7;
              lag=3;
              TCost=0.05589;
           end
         end
      end
   end
end

%% Instantiates the arrays used
nRows = zeros(size(ProductPrice));
DiffPNoLag = nRows(:,1);
DiffPLag = nRows(:,1);
Opportunity = zeros(size(nRows,1),1);

%% Determines non-lagged price differentials
for i=5: size(ProductPrice);
   DiffPNoLag(i)= ProductPrice(i,city)-ProductPrice(i,2)-TCost;
end

%% Fits losgistic distribution on non-lagged price differentials
PDTD=fitdist(DiffPNoLag,'logistic');
UTD=PDTD.Params(1,1);
```

```
STD=PDTD.Params(1,2);

%% Instantiates variables use in probability distribution functions
syms c t v;

%% Sets the stepsize, range and interval size
s=0.01;
F=20;
interval=0.02;

%% Instantiates the arrays used in iterative process
ActualOpportunity=zeros(size(nRows,1),1);
Data= zeros(size(ProductPrice,1),F);
tempDuration=zeros(size(ActualOpportunity,1),1);
Duration=zeros(size(tempDuration,1),F);

low=0-interval;
high=0;

%% Begins the iterative process of collecting data points under
%% various K-intervals
for j=1:F;
    %% iterative process
    K=j*s;

    %% Adjusts lower and upper bound of intervals
    low = low+interval;
    high = high+interval;
    K_values(j)=K;

    %% Records the intervals
    intervals(j,1)=low;
    intervals(j,2)=high;

  for i=5:size(DiffPLag)-lag;
      %% Checks to see if non-lagged differentials falls within interval
      MktDiff = ProductPrice(i,city)-ProductPrice(i,2)-TCost;
      if ((low<=MktDiff)&&(MktDiff<high));
          %% Records the result
          ActualOpportunity(i)=ProductPrice(i+lag,city)-ProductPrice(i,2)-TCost;
      end

      MktDiff=0;
  end
```

```matlab
ActOpp = ActualOpportunity(find(ActualOpportunity));

%% Records the lagged price differentials per k interval
for i=1:size(ActOpp);
   DataTempAct(i,j)=ActOpp(i);
end

%% Binary variable to identify when an opportunity is identified
 for r=5: size(ActualOpportunity);
   if(ActualOpportunity(r)>0);
      Opportunity(r)=1;
   end
 end

 z = find(Opportunity==0);
 y = find(Opportunity==1);
 x=zeros(size(z));
 l=zeros(size(y));

 %% finds the time number of days between opportunites
 x(1)=z(1)-1;
 x(2:end)=diff(z)-1;
 %% finds the length of an opportunity
 l(2:end)= diff(y)-1;

 %% Records the durations of the opportunities
 durations=  x(find(x));
 for r=1:size(durations);
    Duration(r,j)=durations(r);
 end
ActOpp = ActualOpportunity(find(ActualOpportunity));

PDC = fitdist(ActOpp,'logistic');
U_Dif = PDC.Params(1,1);
S_Dif = PDC.Params(1,2);

%% Fits an exponential distribution on durations
PDDur = fitdist(durations,'exponential');
U_Dur(j) = PDDur.Params(1,1);

%% Fits a geometric distribution function on the durations
f = ezfit(durations,'(1-p)^x*p; p=0.5');

%% Calculates the mean of the durations
U_Dur(j) = (1-f.m(1,1))/f.m(1,1);
```

144

```matlab
    zd =U_Dur(j)*exp(-1*(U_Dur(j)*v));

    %% Probabilities of duration and particular differential values
    for r =1:12;
        ProbabilityDurAndDiff=(1-f.m(1,1))^r*f.m(1,1)*(int(yC,c,-10000,high)
                            -int(yC,c,-10000,low));

        %% uses exponential distribution fit
        %       ProbabilityDurGivenDiff=int(zd,v,0,r)-int(zd,v,0,r-1);

        ProbDurationAndDiffChartInputHlpr=ProbabilityDurAndDiff;

        %% Records the probability of a duration and a differential

ProbDurationAndDiffChart(r,j)=double(ProbDurationAndDiffChartInputHlpr);
    end

    %% Probabilities of duration conditioned on given differential
    for r =1:12;
        ProbabilityDurGivenDiff=(1-f.m(1,1))^r*f.m(1,1);

        %% uses exponential distribution fit
        %       ProbabilityDurGivenDiff=int(zd,v,0,r)-int(zd,v,0,r-1);

        %% Records the probability of a duration conditioned on a differential
        ProbDurationGivenDiffChartInputHlpr=ProbabilityDurGivenDiff;

ProbDurationGivenDiffChart(r,j)=double(ProbDurationGivenDiffChartInputHlpr)
;
    end

    %% Resets the arrays used in the iterativ process
    ActualOpportunity=zeros(size(nRows,1),1);
    Opportunity=zeros(size(Opportunity));
    durations=zeros(size(durations));
    l=zeros(size(l));
end

%% Plots surface of the duration conditioned on a differential
surf(ProbDurationGivenDiffChart);
rotate3d;
view([-56 48]);
xlabel('Price Differential ($/lbs)');
ylabel('Duration (days)');
zlabel('Conditional Probability')
```

```
%% Plots surface of the duration and a differential
surf(ProbDurationAndDiffChart);
rotate3d;
view([-56 48]);
xlabel('Price Differential ($/lbs)');
ylabel('Duration (days)');
zlabel('Probability')

%% END OF CODE
```