PSERC

Integrated Financial and Operational Risk Management in Restructured Electricity Markets

Final Project Report

Power Systems Engineering Research Center

Empowering Minds to Engineer
the Future Electric Energy System
Since 1996

Integrated Financial and Operational Risk Management in Restructured Electricity Markets

Final Project Report

Research Team Faculty

Shijie Deng, Project Leader
Sakis Meliopoulos
Georgia Institute of Technology

Shmuel Oren
University of California at Berkeley

Research Team Students

Jieyun Zhou and Li Xu, Georgia Institute of Technology
Yumi Oum and Yongheon Lee, University of California at Berkeley

PSERC Publication 09-13

October 2009

Information about this project

For information about this project contact:

Shijie Deng, Ph.D.
Georgia Institute of Technology
School of Industrial and Systems Engineering
Atlanta, GA 30332
Tel: 404-894-6519
Fax: 404-894-2301
Email: deng@isye.gatech.edu

Power Systems Engineering Research Center

This is a project report from the Power Systems Engineering Research Center (PSERC). PSERC is a multi-university Center conducting research on challenges facing a restructuring electric power industry and educating the next generation of power engineers. More information about PSERC can be found at the Center's website: http://www.pserc.org.

For additional information, contact:

Power Systems Engineering Research Center
Arizona State University
577 Engineering Research Center
Box 878606
Tempe, AZ 85287-8606
Phone: 480-965-1643
Fax: 480-965-0745

Notice Concerning Copyright Material

Acknowledgements

This is the final report for the Power Systems Engineering Research Center (PSERC) research project entitled “Integrated Financial and Operational Risk Management in Restructured Electricity Markets.” (PSERC project M-17). The project began June 2007 and was completed in June 2009. We express our appreciation for the support provided by PSERC’s industry members.

The authors thank all PSERC members for their technical advice on the project, especially Art Altman (EPRI), Hung-Po Chao (ISO-New England), Mark Sanford (GE Energy), and Todd Strauss (PG&E) who were our industry advisors.

Executive Summary

In the restructured electric power industries, how to manage the extremely high price volatility in the electricity wholesale markets has been a crucial factor to the smooth and viable business operations of all parties, including independent power producers, system operators and load serving entities and the likes. Compounded with the price risk, quantity or volumetric risk that arises from demand uncertainty due to weather conditions and load migration, presents major challenges and opportunities for the above mentioned market participants. The financial exposures to these two sources of risk that could result in severe financial losses are amplified by the positive correlation between load and price, which prevails in electricity markets. Therefore, managing these risks is essential to the financial success of participants in the electricity industry.

This project investigates the integration of financial and operational risk management mechanisms to facilitate market operations and enhance market efficiency in the restructured electricity industry. Financial and operational hedging strategies utilizing existing standard and prospective instruments have been studied. This work has developed methods for pricing such instruments and assessing their effectiveness.

I. Electricity Price Curve Modeling and Forecasting

We established a novel non-parametric approach for the modeling and analysis of electricity price curves by applying the manifold learning methodology—locally linear embedding (LLE). The prediction method is based on manifold learning, and reconstruction is employed to make short-term and medium-term price forecasts. Our method not only performs accurately in forecasting one-day-ahead prices, but also has a great advantage in predicting one-week-ahead and one-month-ahead prices over other methods. The forecast accuracy is demonstrated by numerical results using historical price data taken from the Eastern U.S. electric power markets.

II. An Equilibrium Pricing Model for Weather Derivatives in a Multi-commodity Setting

We developed an equilibrium-pricing model for weather derivatives in a multi-commodity setting. The model is constructed in the context of a stylized economy where market participants optimize their hedging portfolios, which include weather derivatives that are issued in a fixed quantity by a financial underwriter. The demand of weather derivatives resulting from hedging activities of buyers and the supply by the underwriters are combined in an equilibrium-pricing model under the assumption that all participants maximize some risk-averse utility function. We analyzed the gains due to the inclusion of weather derivatives in hedging portfolios and examined the components of that gain attributable to risk hedging and to risk sharing.

III. Hedging Quantity Risks with Standard Power Options

We analyzed the quantity risk in the electricity market, and explored several ways of managing it. The research also addressed the price and quantity risk hedging problem of a load serving entity (LSE), which provides electricity service at a regulated price in electricity markets. Exploiting the correlation between consumption volume and spot price of electricity, we derived an optimal zero-cost hedging function characterized by

the payoff as a function of spot price. How such a hedging strategy can be implemented through a portfolio of forward contracts and call and put options was also illustrated.

IV. Optimal Static Hedging of Volumetric Risk

We developed a static hedging strategy for an LSE or a marketer whose objective is to maximize a mean-variance utility function over net profit, subject to a self-financing constraint. Since quantity risk is non-tradable, the hedge consists of a portfolio of price-based financial energy instruments, including a bond, a forward contract and a spectrum of European call and put options with various strike prices. The optimal hedging strategy, which varies in contract timing, is jointly optimized with respect to contracting time and the portfolio mix under specific price and quantity dynamics, and the assumption that the hedging portfolio, which matures at the time of physical energy delivery, is purchased at a single point in time. Explicit analytical results are derived for the special case where price and quantity have a joint bivariate lognormal distribution.

V. VaR Constrained Hedging of Fixed Price Load-Following Obligations

We developed a self-financed hedging portfolio consisting of a risk free bond, a forward contract and a spectrum of call and put options with different strike prices. A popular portfolio design criterion is the maximization of expected hedged profits subject to a Value-at-risk (VaR) constraint. Unfortunately, that criterion is difficult to implement directly due to the complicated form of the VaR constraint. We show, however, that under plausible distributional assumptions, the optimal VaR constrained portfolio is on the efficient Mean-Variance frontier. Hence, we proposed an approximation method that restricts the search for the optimal VaR constrained portfolio to that efficient frontier. The proposed approach is particularly attractive when the Mean-Variance efficient frontier can be represented analytically, as is the case, when the load and logarithm of price follow a bivariate normal distribution. We illustrate the results with a numerical example.

Potential uses of the developed analytical tools

In order to show the practical usage of the model discussed in this project, we have developed a graphic **User Interface** for industry members to investigate the hedging performance of the optimal portfolios suggested by our model. We implemented the model developed in Oum, Oren, Deng 2006 as an illustration. Our intention is that, with real market data inputted and utility functions specified by the industry users, the interface could provide the corresponding payoff functions, the positions of forward contracts and options, and the performance of hedging the price and volumetric risks.

Future work

On the side of hedging with financial instruments, a credit limit constraint, which limits the amount of money that can be borrowed to construct the portfolio, needs to be considered in future extension of our work. A dynamic hedging strategy rather than the static approach is likely to improve the hedging performance and should also be considered. On the other side, we would like to incorporate a broad range of demand-side management programs into the analytic framework and investigate the impact of these programs in hedging the price and volumetric risks. The valuation and role of other tools, for example, “out-of -money” power plant should also be explored.

Table of Contents

List of Tables

List of Figures

1. Introduction

In the restructured electric power industries, how to manage the extremely high price volatility in the electricity wholesale markets has been a crucial factor to the smooth and viable business operations of all parties, including independent power producers, system operators and load serving entities and the likes. Compounded with the price risk, quantity, or volumetric, risk that arises from demand uncertainty due to weather conditions and load migration, presents major challenges and opportunities for the above mentioned market participants. Non-storability of electricity, the steep rise in the supply function, limited demand response and demand fluctuation which are largely driven by weather conditions are among the major factors contributing to high price volatility in restructured wholesale electricity markets. On the other hand, vertical unbundling of the generation and distribution sectors has removed some of the natural hedging that previously existed in the vertically integrated industry and exposed both generation investments and consumers to the spot price risk. This price risk must now be managed through financial hedging and long term contracting. In addition, resources such as operating reserves and planning reserves that were deployed through command and control are now procured through market and economic incentives, which creates operational risk exposures threatening system reliability. Managing such operational risks requires "real options" that come with physical assets and operating protocols.

Overall an efficient electricity industry requires efficient and reliable operation as well as competitive and liquid markets for trading and risk allocation. Unfortunately, while market design efforts in the US and abroad over the last decade have focused on the development of market mechanisms that ensure spot market efficiency, little has been done toward understanding and facilitating efficient markets for trading and allocating risk in the electricity supply chain. Exposure to price risk of the three major utilities in California during the electricity crisis in 2000 and 2001 led them to bankruptcy or near bankruptcy. More recently during the ice storms in Texas in February 2003, a retail energy provider went into bankruptcy after incurring a devastating loss attributed to high price spikes.

Significant economic risks in a restructured electricity market do not come from price fluctuation alone. Quantity, or volumetric, risk that arises from demand uncertainty due to weather conditions and load migration, also presents major challenges and opportunities for market participants. For example, an LSE who has purchased a forward contract in order to serve its native load (at fixed regulated retail prices) may find that the demand realization will be less than expected, requiring the LSE to resell the residual electricity in the spot market at lower prices than the purchasing cost. Likewise during a hot day the LSE may become short in supply quantity and have to meet the extra demand through purchases in the wholesale spot market at prices that may exceed its regulated retail rate. Such exposure that could cause extreme financial losses is amplified by the positive correlation between load and price, which prevails in electricity markets. A similar situation is faced by energy merchants who bear quantity risk by entering into annual "Default Service Contracts" that obligate them to serve a fixed percentage of electricity loads at a fixed price per MWh. Such contracts are auctioned off annually in New Jersey by the local utility. The holder of the contract must then decide how to manage its risk through a combination of physical generation capacity and financial hedging.

Our project addressed the problem of integrating financial and operational hedging of price and volumetric risks. We examined the design and evaluation of hedging mechanisms utilizing both standard electricity financial instruments such as forward contracts and options and non-standard derivatives such as weather derivatives. Obviously, many industries, including the energy industry, are directly or indirectly exposed to weather risk. Although catastrophic events such as storms and hurricanes cause serious damage to most industries, even less extreme weather conditions can significantly affect the revenue of weather-sensitive industries. In terms of the energy industry, it is exposed to weather risk because the energy demand is highly dependent on weather conditions. Unexpected weather changes will affect energy demand and sudden demand increases result in spot price spikes. Thus, the price, volumetric and weather risks are all correlated. Weather derivatives provide an effective way to mitigate financial losses due to weather. They are financial instruments providing predetermined compensation in proportion to the deviation of the average temperature over a fixed time interval from a fixed norm (typically 65°F), which can be viewed as weather insurance. Such insurance can offset losses associated with extreme weather conditions such as those resulting from excess load during high price periods. Weather derivatives are particularly attractive as a supplement to operational hedges such as tolling contracts or distributed generation facilities since they are relatively liquid and they enable risk diversification and risk sharing across multiple commodities whose consumption and pricing are correlated with weather. In order to better understand the hedging effectiveness of weather derivatives, first of all, we need to understand the major factors driving the electricity price dynamics and identify the linkage between these factors to the weather factor. In Chapter 2, we propose a manifold-based dimension reduction to identify the non-linear mapping between fundamental supply-demand factors and the dynamics of electricity price curves. Local Linear Embedding is demonstrated to be an efficient method for extracting the intrinsic low-dimensional structure of electricity price curves. Using price data taken from the New York ISO, we found that there exists a low-dimensional manifold representation of the day-ahead price curve in the NY Power Pool, and specifically, the dimension of the manifold is around 4. The interpretation of each dimension in the low-dimensional space is attributed to the mean, standard deviation and skewness of the price curve, which are all strongly correlated with weather data such as temperature. The cluster analysis was performed to confirm that these identified factors capture the electricity price curve dynamics very well in the sense that the cluster pattern based on the first 3 factors match the pattern based on the original price curves.

On the weather derivative pricing end, because the weather derivatives market is incomplete, there is no effective pricing model. In Chapter 3, we formulate an equilibrium-pricing model in a multi-commodity setting that is driven by demand of weather derivatives, which is derived from hedging and risk diversification activities in weather sensitive industries. Specifically, the model is constructed in the context of a stylized economy where risk-averse agents optimize their hedging portfolios, which include weather derivatives that are issued in a fixed quantity by a financial underwriter. The supply and demand resulting from hedging activities and the supply by the underwriter are combined in an equilibrium-pricing model under the assumption that all agents maximize some utility functions.

As a part of our analysis, we measure the risk hedging and sharing effects of the weather derivatives, both of which contribute to increasing the expected utility of agents who trade these

hedging instruments. To price the weather derivatives, we assume that there are buyers and an issuer in a closed and frictionless endowment economy and all of them are utility maximizers. By solving the utility maximization problems of the market participants, we determine the optimal demand and supply functions for weather derivatives and obtain their equilibrium prices by invoking a market clearing condition. In the multi-commodity economy, the weather derivative has two effects: the risk hedging effect and the risk sharing effect. While in a single-commodity economy, there is only a risk hedging effect since there is no counter-party to share risk. We measure these effects in terms of certain equivalent differences among various cases. Under the mean-variance utility function, we were able to derive closed form expressions for equilibrium prices and the measurement of the risk hedging and sharing effects. Such expressions will be useful in future empirical work that will attempt to calibrate the model parameters to market data. Numerical examples employing Monte-Carlo simulations show that the equilibrium price tends to increase as the correlation between temperature and demand increases due to the high demand for the weather derivative. In addition, the numerical examples verify that weather derivative improves hedging and risk diversification capability, especially in situations where commodity derivatives are not available.

In addition to weather derivatives, forward contracts and derivatives such as call and put options have become common tools for mitigating price and quantity risks in electricity markets. An electricity forward contract obligates a party to buy and the other party to sell a specified quantity on a given delivery date in the future at a predetermined fixed price. At the delivery date if the market price is higher than the contracted forward price, then the buyer will benefit, conversely, if the market price is lower than the forward price, then the buyer will suffer. Put or call options are used for hedging either downside or upside price risks alone. The buyer of an electricity put (call) option pays a premium for the right to sell (buy) electricity at a specified price, called strike price, at a specific time in the future. LSEs would use call options to avoid the risk of high electricity purchasing cost and still enjoy the benefit of low electricity spot prices. While it is relatively simple for a power market participant who has obligations in delivering power to hedge price risk for a given quantity, it is more difficult to hedge price risk when the quantity demanded is uncertain and correlated with the price. The price-demand correlation is evident in electricity market and should be considered in solving hedging problems. For example, the correlation coefficient between price and load in Northern California from April 1998 to March 2000 was about 0.5. Our research for this PSerc project has exploited this correlation in developing hedging strategies that address both price and quantity risks.

One front is on constructing the optimal static hedge of the volumetric risk of loads through a portfolio of a risk-free bond and a set of forwards/standard options contracts written on electricity traded in competitive markets, which is discussed in Chapter 4. This work is based on the earlier work of this PSerc project (Oum, Oren and Deng 2006). We obtained the optimal hedging strategy that used electricity derivatives to hedge price and volumetric risks by maximizing the expected utility of the hedged profit. When such a portfolio is held by an LSE, the call options with strikes being below the spot price will be exercised so that the amount of the options being exercised is procured at the strike prices. Using this strategy, the LSE can set an increasing price limit on incremental load by paying the premiums for the options. This strategy is shown to be quite effective in managing quantity risk and it was also suggested in the market

design literature such as Chao and Wilson 2004, Oren 2005, and Willems 2006 as means to achieve resource adequacy, mitigate market power, and reduce spot price volatility.

We extended the single-period setting in Oum, Oren and Deng 2006 by allowing contract procurement to take place anywhere between the decision time at the onset of the period and the exercise time at the end of the period (when delivery occurs) as long as the entire hedging portfolio is procured at a single point in time. Within this framework we co-optimize the mix and procurement time of the hedging portfolio. We first solve for the optimal payoff of a general static hedging function given the procurement time, and then find a replicating portfolio that consists of forward, European calls and puts which yields the optimal payoff. Prices of the forwards and options contracts that are included in our hedging portfolio change as the time approaches delivery time, reflecting the changing expectations in the market. Thus, the mix of the optimal hedging portfolio also changes with the hedging time.

Our result shows that hedging too late can increase risk sharply. Optimizing such timing decisions requires solving an integrated problem of selecting the optimal hedging portfolio and time. For mean-variance expected utility, we solved for the optimal hedging time, under classical assumption regarding the stochastic processes governing forward price and load-estimate. Through numerical examples, we showed that generally there is a critical time beyond which the uncertainty in profit increases sharply while the uncertainty remains relatively constant before this critical time. Sensitivity analysis results indicate that the optimal hedging time gets closer to the delivery period if the positive correlation between the forward price and load-estimate is higher, and if the load-estimate volatility is higher. It is also observed that delaying the hedging time past the optimum time can be very risky, while the earlier hedging makes little difference as compared with hedging at the optimal time. This suggests that in practice one should err by hedging early rather than taking the chance of being too late.

The other front is to exploit the inherent positive correlation between wholesale electricity price and demand volume to develop a hedging strategy which maximizes the expected profit subject to a value-at-risk (VaR) constraint. The model is proposed in Chapter 5. A VaR constraint on a portfolio limits the lowest level below which the portfolio value wouldn't fall during a specified time period with 95% confidence. Specifically, we developed a hedging strategy for the LSE's retail positions (which is in fact a short position on unknown volume of electricity) using electricity standard derivatives such as forwards, calls, and puts. However, VaR constrained problems are generally very hard to solve analytically unless the value or profit under consideration is normally distributed. In our case, the profit depends on the product of the two correlated variables. Moreover, our hedging strategy is characterized by a nonlinear function of a random variable. We addressed this difficulty by limiting our search to feasible VaR-constrained self-financed hedging portfolios on the mean-variance efficient frontier. We provide theoretical justification to such an approximation and derive, an analytic representation of hedging portfolios on the mean-variance efficient frontier as function of the risk aversion factor. The computation of an approximate solution to the VaR constrained problem on the mean variance efficient frontier is facilitated by the fact that it corresponds to the smallest risk-aversion factor whose associated VaR meets the constraint limit.

2. Modeling and Forecasting the Electricity Price Curve

2.1. Introduction

In the competitive electricity wholesale markets, market participants, including power generators and merchants alike, strive to maximize their profits through prudent trading and effective risk management against adverse price movements. A key to the success of market participants is to model the electricity price dynamics well and capture their characteristics realistically. Researches on modeling electricity price processes focus on the aspect of derivative pricing (e.g., Johnson and Barz 1999, Deng 2000, Lucia and Schwartz 2002) and on forecasting spot or short- term electricity prices, especially the day-ahead prices (e.g., Davison et al. 2002, Nogales et al. 2002, Contreras et al. 2003, Conejo et al. 2005, etc.). While spot price modeling is important, successful trading and risk management operations in electricity markets also require knowledge on an electricity price curve consisting of prices of electricity delivered at a sequence of future times instead of only at the spot. Audet et al. 2004 proposes a parametric forward price curve model for the Nordic market, which does not model the movements of the expected future level of a forward curve. Lora et al. 2006 employs a weighted average of nearest neighbors approach to model and forecast the day-ahead price curve. These works offer little insight on understanding the main drivers of the price curve dynamics. We proposed a novel nonparametric approach for modeling electricity price curves. Analysis on the intrinsic dimension of an electricity price curve is offered, which sheds light on identifying major factors governing the price curve dynamics. The forecast accuracy of our model compares favorably against that of the ARX and ARIMA model in one-day-ahead price predictions. In addition, our model has a great advantage on the predictions in a longer horizon from days to weeks over other models.

In general, the task of analytically modeling the dynamics of such a price curve is daunting, because the curve is a high-dimensional subject. Each price point on the curve essentially represents one dimension of uncertainty. To reduce the dimension of modeling a price curve and identify the major random factors influencing the curve dynamics, Principle Component Analysis (PCA) is proposed and has been widely applied in the real-world data analysis for industrial practices. As PCA is mainly suited for extracting the linear factors of a data set, it does not appear to perform well in fitting electricity price curves with a linear factor model in a low-dimensional space. A natural extension to the PCA approach is to consider the manifold learning methods, which are designed to analyze intrinsic nonlinear structures and features of high-dimensional price curves in the low-dimensional space. After obtaining the low-dimensional manifold representation of price curves, price forecasts are made by first predicting each dimension coordinate of the manifold and then utilizing a reconstruction method to map the forecasts back to the original price space. The conceptual flowchart of our modeling approach is illustrated by Figure 2-1. Our major contribution is to establish an effective approach for modeling energy forward price curves, and set up the entire framework in Figure 2-1. The other major contribution is to identify the nonlinear intrinsic low-dimensional structure of price curves. The resulting analysis reveals the primary drivers of the price curve dynamics and facilitates accurate price forecasts. This work also enables the application of standard times series models such as Holt-Winters in the forecast step from box 1 to box 2.

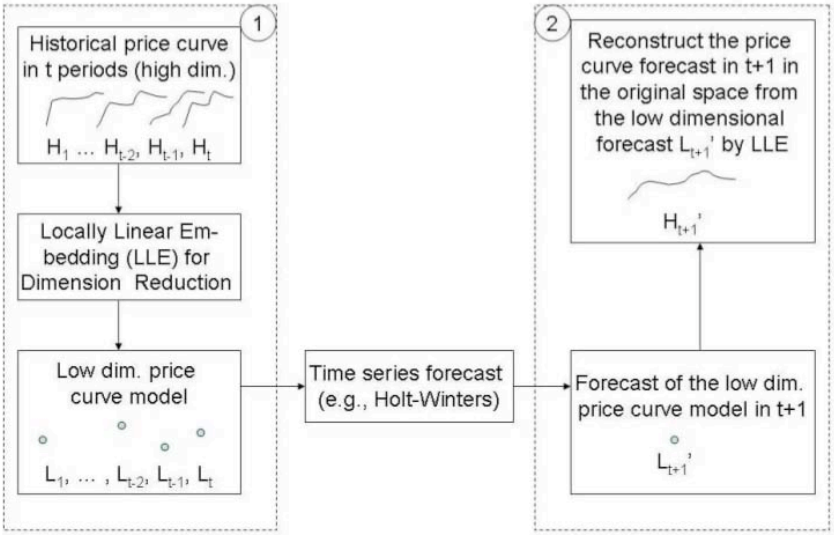


Figure 2-1: The conceptual flow chart of the model

2.2. Manifold Learning Algorithm

2.2.1. Introduction to Manifold Learning

Manifold learning is a new and promising nonparametric dimension reduction approach. Many high-dimensional data sets that are encountered in real-world applications can be modeled as sets of points lying close to a low-dimensional manifold. Given a set of data points $x_1, x_2, \cdots, x_N \in R^D$, we can assume that they are sampled from a manifold with noise, i.e.,

$$x_i = f(y_i) + \varepsilon_i, i = 1, \cdots, N \tag{2.1}$$

where $y_i \in R^d, d << D$, and ε_i are noises. Integer d is also called the intrinsic dimension. The manifold based methodology offers a way to find the embedded low-dimensional feature vectors y_i from the high-dimensional data points x_i.

Many nonparametric methods were created for nonlinear manifold learning, including multidimensional scaling (MDS), locally linear embedding (LLE), Isomap, Laplacian eigenmaps, Hessian eigenmaps, local tangent space alignment (LTSA), and diffusion maps.

Among various manifold based methods, we find that locally linear embedding (LLE) works well in modeling electricity curves. Moreover, LLE and LLE-reconstruction are fast and easy to implement. In the next two subsections, we introduce the algorithms of LLE and LLE reconstruction, respectively.

2.2.2. Locally Linear Embedding (LLE)

Given a set of data points $x_1, x_2, \cdots, x_N \in R^D$ in the high-dimensional space, we are looking for the embedded low-dimensional feature vectors $y_1, y_2, \cdots, y_N \in R^d$. LLE is a nonparametric method that works as follows:

1. Identify the k nearest neighbors based on Euclidean distance for each data point $x_i, 1 \le i \le N$. Let N_i denote the set of the indices of the k nearest neighbors of x_i.
2. Find the optimal local convex combination of the k nearest neighbors to represent each data point x_i. That is, the following objective function (2.2) is minimized and the weights w_{ij} of the convex combinations are calculated.

$$E(w) = \sum_{i=1}^{N} \| x_i - \sum_{j \in N_i} w_{ij} x_j \|^2$$

(2.2) where $\| \cdot \|$ is the l_2 norm and $\sum_{j \in N_i} w_{ij} = 1$.

The weight w_{ij} indicates the contribution of the jth data point to the representation of the ith data point. The optimal weights can be solved as a constrained least square problem, which is finally converted into a problem of solving a linear system of equation.

3. Find the low-dimensional feature vectors $y_i, 1 \le i \le N$, which have the optimal local convex representations with weights w_{ij} obtained from the last step. That is, y_i's are computed by minimizing the following objective function:

$$\Phi(y) = \sum_{i=1}^{N} \| y_i - \sum_{j \in N_i} w_{ij} y_j \|^2 \qquad (2.3)$$

It can be shown that solving the above minimization problem (2.3) is equivalent to solving an eigenvector problem with a sparse $N \times N$ matrix. The d eigenvectors associated with the d smallest nonzero eigenvalues of the matrix comprise the d-dimensional coordinates of y_i's. Thus, the coordinates of y_i's are orthogonal.

LLE does not impose any probabilistic model on the data; however, it implicitly assumes the convexity of the manifold. It can be seen later that this assumption is satisfied by the electricity price data.

2.2.3. LLE Reconstruction

Given a new feature vector in the embedded low-dimensional space, the reconstruction method is used to find its counterpart in the high-dimensional space based on the calibration data set. Reconstruction accuracy is critical for the application of manifold learning in the prediction. There are a limited number of reconstruction methods in the literature. For a specfic linear manifold, the reconstruction can be easily made by PCA. For a nonlinear manifold, LLE reconstruction is derived in the similar manner as LLE. Among all the reconstruction methods, LLE reconstruction has the best performance for the electricity data. This is an important reason for us to choose LLE and LLE reconstruction.

Suppose low-dimensional feature vectors $y_1, y_2, \cdots, y_N$ have been obtained through LLE in the previous subsection. Denote the new low-dimensional feature vector as y_0. LLE reconstruction is applied to find the approximation $\hat{x}_0$ of the original data point x_0 in the high-dimensional space based on $x_1, x_2, \cdots, x_N$ and $y_1, y_2, \cdots, y_N$. There are three steps for LLE reconstruction:

1. Identify the k nearest neighbors of the new feature vector y_0 among $y_1, y_2, \cdots, y_N$. Let N_0 denote the set of the indices of the k nearest neighbors of y_0.
2. The weights of the local optimal convex combination w_j are calculated by minimizing
$$E(w) = \| y_0 - \sum_{j \in N_0} w_j y_j \|^2, \qquad (2.4)$$
subject to the sum-to-one constraint, $\sum_{j \in N_0} w_j = 1$.
3. Data point $\hat{x}_0$ is reconstructed by $\hat{x}_0 = \sum_{j \in N_0} w_j x_j$.

Remark: Solving optimization problems (2.2) and (2.4) is equivalent to solving a linear system of equations. When there are more neighbors than the high dimension or the low dimension, i.e., $k > D$ or $k < d$, the coefficient matrix associated with the system of linear equations is singular, which means that the solution is not unique. This issue is solved by adding an identity matrix multiplied with a small constant to the coefficient matrix (see Saul and Roweis 2003). We adopt this approach here.

Suppose $x_0^{(j)}, 1 \le j \le D$, is the jth component of vector x_0. The reconstruction error (RE) of x_0 is defined as

$$RE(x_0) = \frac{1}{D} \sum_{j=1}^{D} \frac{| x_0^{(j)} - \hat{x}_0^{(j)} |}{x_0^{(j)}} \qquad (2.5)$$

The reconstruction error of the entire calibration data set (TRE) is defined as

$$TRE = \frac{1}{N \times D} \sum_{i=1}^{N} \sum_{j=1}^{D} \frac{| x_i^{(j)} - \hat{x}_i^{(j)} |}{x_i^{(j)}} \qquad (2.6)$$

by regarding each y_i as a new feature vector y_0.

2.3. Electricity Price Curve Modeling with Manifold Learning

The data of the day-ahead market locational-based marginal prices (LBMPs) and integrated real-time actual load of electricity in the Capital Zone of the New York Independent System Operator (NYISO) are collected and predicted in this project. The data are available online (www.nyiso.com/public/market_data/pricing_data.jsp). In this section, two years (731 days) of price data from Feb 6, 2003 to Feb 5, 2005 are used as an illustration of modeling the electricity price curves by manifold based methodology. Figure 2-2(a) plots the hourly day-ahead LBMPs during this period, where the electricity prices are treated as a univariate time series with 24 × 731 hourly prices. Figure 2-2(b), 2-2(c) and 2-2(d) illustrate the mean, standard deviation and skewness of 24 hourly log prices in each day after outlier processing.

2.3.1. Preprocessing

1) Log Transform: The logarithmic (log) transforms of the electricity prices are taken before the manifold learning. There are several advantages to deal with the log prices. First, the electricity prices are well known to have the non-constant variance, and log transform can make the prices less volatile. The log transform also enhances the efficiency of manifold learning, by making the embedded manifold more uniformly distributed in the low-dimensional space and the reconstruction error of the entire calibration data set (TRE) reduced. Moreover, the log transform has the interpretation of the returns to someone holding the asset.

2) Outlier Processing: Outliers in this paper are defined as the electricity price spikes that are extremely different from the prices in the neighborhood. To deal with the outliers, we replace the prices in the day with outliers by the average of the prices in the days right before and right after. We remove the outliers because the embedded low-dimensional manifold is supposed to extract the primary features of the entire data set, rather than the individual and local features such as extreme price spikes. The efficiency of manifold learning is improved after outlier processing. Moreover, outliers, which represent rarely occurring phenomena in the past, often have very small probability to occur in the near future, so the processing of outliers does not severely affect the prediction of the near-term regular prices.

In the illustrated data set, only one extreme spike is identified on the right of Figure 2-2(a), which belongs to Jan 24, 2005. In the low-dimensional manifold, the days of outliers can also be detected by the points that stand far away from the other points. Figure 2-3 shows that the point corresponding to Jan24, 2005 lies out of the main cloud of the points on the embedded three-dimensional manifold. Thus, we regard Jan24, 2005 as a day with outliers. Figure 2-4 shows that the low-dimensional manifold after removing the outliers is more uniformly distributed.

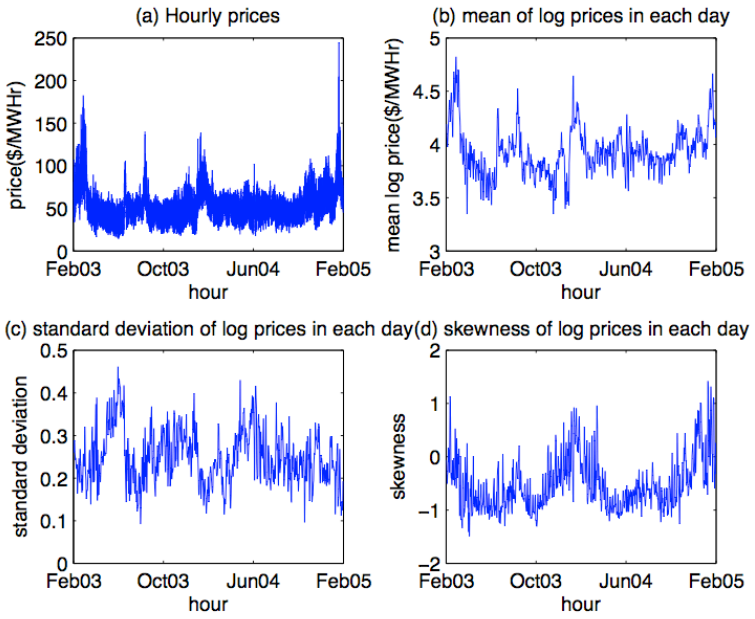


Figure 2-2: Day-ahead LBMPs from Feb 6, 2003 to Feb 5,2005 in the Capital Zone of NYISO

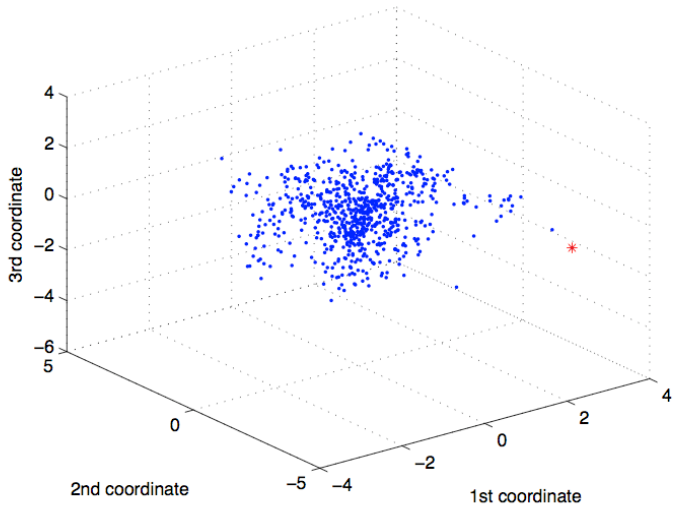


Figure 2-3: Embedded three-dimensional manifold without any outlier preprocessing (but with log transform and LLP smoothing). "*" indicates the day with outliers -- Jan 24, 2005

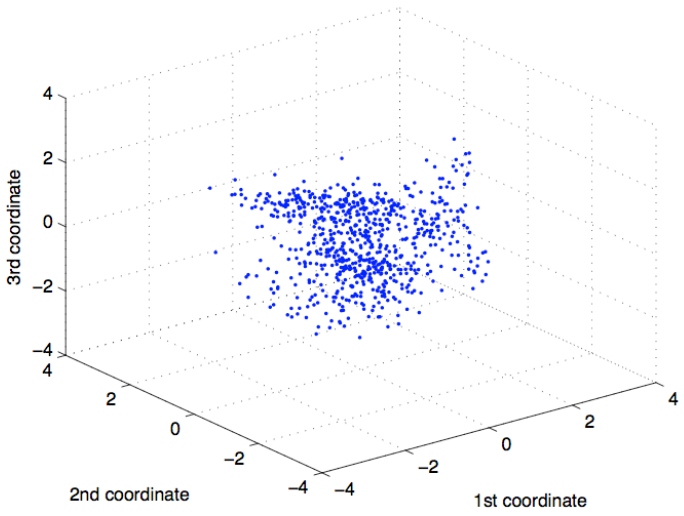


Figure 2-4: Embedded three-dimensional manifold after log transform, outlier preprocessing and LLP smoothing

3) LLP Smoothing: The noise in (2.1) can contaminate the learning of the embedded manifold and the estimation of the intrinsic dimension. Therefore, locally linear projection (LLP) (Huo and Chen 2002, Huo 2003) is recommended to smooth the manifold and reduce the noise. The description of the algorithm is given as follows:

ALGORITHM: LLP

For each observation $x_i, i = 1,2,\cdots,N$,

1. Find the k-nearest neighbors of x_i. The neighbors are denoted by $\tilde{x}_1, \tilde{x}_2, \cdots \tilde{x}_k$
2. Use PCA or SVD to identify the linear subspace that contains most of the information in the vectors $\tilde{x}_1, \tilde{x}_2, \cdots \tilde{x}_k$. Suppose the linear subspace is A. Let k_0 denote the assumed dimension of the embedded manifold. Then subspace A_i can be viewed as a linear subspace spanned by the singular vectors associated with the largest k_0 singular values.
3. Project x_i into the linear subspace A_i and let $\breve{x}_i, i = 1,2,\cdots,N$, denote the projected points.

After denoising, the efficiency of manifold learning is enhanced, and the reconstruction error (TRE) of the entire calibration data set is reduced. For the illustrated data set with the intrinsic dimension being four, the TRE is 3.89% after LLP smoothing, compared to 4.41% without LLP smoothing. The choice of the two parameters in LLP, the dimension of the linear space and the number of the nearest neighbors, will be discussed in detail in subsection 2.3.4.

2.3.2. Manifold Learning by LLE

Each price curve with 24 hourly prices in a day is considered as an observation, so the dimension of the high-dimensional space D is 24. The intrinsic dimension d is set to be four. The

number of the nearest neighbors k for LLP smoothing, LLE, and LLE reconstruction is selected to be a common number 23 for all the numerical studies. The details of the parameter selections are discussed in subsection 2.3.4. Due to the ease of visualization in a three-dimensional space, all the low-dimensional manifolds are plotted with the intrinsic dimension being three. We apply LLE to the denoised data $\bar{x}_i, i = 1,2,\cdots,N$, which are obtained after LLP smoothing. Figure 2-4 provides the plot of the embedded three-dimensional manifold. As the low-dimensional manifold is nearly convex and uniformly distributed, LLE is an appropriate manifold based method. Figure 2-5 plots the time series of each coordinates of the feature vectors in the embedded four-dimensional manifold.

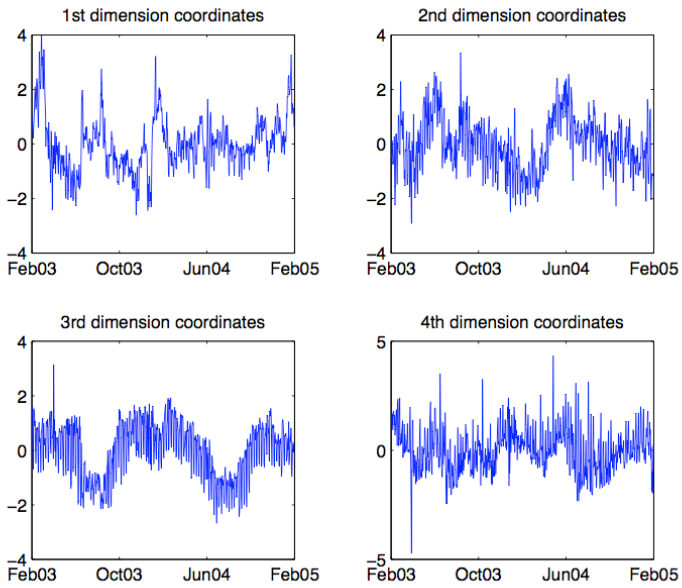


Figure 2-5: Coordinates of the embedded 4-dim manifold

Table 2-1 shows the TRE of different reconstruction methods. LLE reconstruction has the minimum reconstruction error among all the methods. LTSA reconstruction has a very large TRE, because it is an extrapolation-like method, and the reconstruction of some of the price curves has very large errors. Therefore, LLE and LLE reconstruction are selected to model the electricity price dynamics.

Table 2-1: The TRE of different reconstruction methods

Reconstruction method	TRE(%)
LLE and LLE reconstruction	3.89
PCA and PCA reconstruction	4.26
LTSA and LTSA reconstruction	4.55×10^6
LLE and nonparametric regression reconstruction	4.77

2.3.3. Analysis of Major Factors of Electricity Price Curve Dynamics with Low-Dimensional Feature Vectors

1) Interpretation of Each Dimension in the Low-Dimensional Space: There are some interesting interpretations for the first three coordinates of the feature vectors in the low-dimensional space. For each price curve, we can calculate the mean, standard deviation, range, skewness and kurtosis of the 24 hourly log prices. The sequence of each coordinates of the low-dimensional feature vectors comprises a time series. The correlation between each time series and mean log prices (standard deviation, range, skewness and kurtosis) is calculated. Table 2-2 shows the one of the four-dimensional coordinates, which has the maximum absolute correlation with mean log prices (standard deviation, range, skewness and kurtosis), and the corresponding correlation coefficients. The comparison between Figure 2-2 and Figure 2-5 gives more intuition about the correlations. It is found that the first coordinates have a very high correlation coefficient 0.9964 with the mean log prices within each day, and the second coordinates are highly correlated with the standard deviation of the log prices in a day with a correlation coefficient 0.7073. This also means that the second coordinates contain some other information besides standard deviation, and Table 2-2 demonstrates that the second coordinates are also correlated, but not significantly, with range and skewness. The third coordinates show both weekly and yearly seasonality in Figure 2-5. Weekly seasonality is well known for electricity prices. Yearly seasonality may be caused by the shape change of the price curves over the year. The shape of price curves is often unimodal in the summer and bimodal in the winter.

Table 2-2: The one of the four - dimensional coordinates which has the maximum absolute correlation coefficient with the mean (standard deviation, range, skewness and kurtosis) of log Prices in a day in embedded four-dimensional space.

	Mean	Standard Deviation	Range	Skewness	Kurtosis
Coordinate	1st	2nd	2nd	2nd	3rd
Correlation Coefficient	0.9964	0.7073	0.5141	-0.5646	0.2611

2) Cluster Analysis: The yearly seasonality of the electricity price curves can be clearly demonstrated by the cluster analysis of low-dimensional feature vectors.

Cluster analysis (also known as data segmentation, see Hastie et al. 2001) groups or segments a collection of objects into subsets (i.e., clusters), such that those within each cluster are more closely related to each other than those assigned to different clusters.

The K-means clustering algorithm is one of the mostly used iterative clustering methods. Assume that there are K clusters. The algorithm begins with a guess of the K cluster centers. Then, the algorithm iterates between the following two steps until convergence. The first step is to identify the closest cluster center for each data point based on some distance metric. The second step is to replace each cluster center with the coordinate-wise average of all the data points that are the closest to it.

For the electricity price data, we apply K-means clustering with Euclidean distance to the low-dimensional feature vectors that are obtained from manifold learning. The number of clusters is set to be three, as the yearly seasonality can be clearly illustrated with three clusters. The coordinate-wise average of price curves in each cluster is plotted in Figure 2-6. The distribution of clusters is illustrated in the first graph of Figure 2-7, where x axis is the date of the price curves, and y axis is the corresponding clusters. The two graphs show that the first cluster represents the price curves from the summer, which are featured with unimodal shape, and the second cluster represents the ones from the winter, which are characterized with bimodal shape. The price curves in the third cluster reveal the transition from unimodal shape to bimodal shape. The average price curves in the 3 clusters closely resemble the typical load shapes observed in summer, winter, and rest-of-year, respectively.

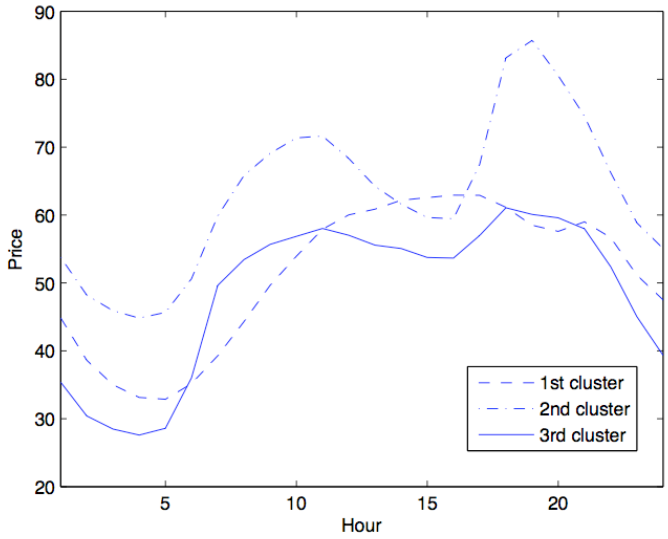


Figure 2-6: The coordinate-wise average of the actual price curves in each cluster, where clustering is based on low-dimensional feature vectors.

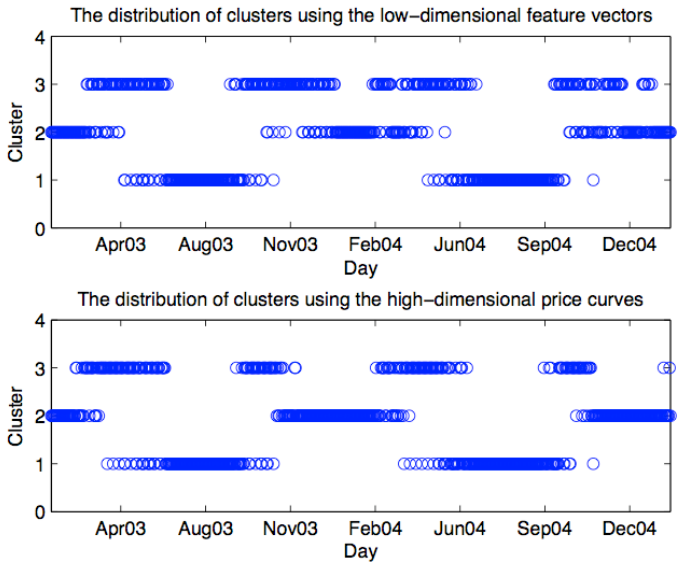


Figure 2-7: Distribution of clusters.

The second graph of Figure 2-7 shows the distribution of clusters by applying K-means clustering with correlation distance to the high-dimensional price curves. The two graphs in Figure 2-7 have the similar patterns, which gives a good illustration that low-dimensional feature vectors capture the major factors of the price curve dynamics.2.3.4. Parameter Setting and Sensitivity Analysis

1) Intrinsic Dimension: Intrinsic dimension d is an important parameter of manifold learning. Levian and Bickel 2005 and Verveer and Duin 1995 provide several approaches of estimating the intrinsic dimension. In Levian and Bickel 2005, the maximum likelihood estimator of the intrinsic dimension is established. In Verveer and Duin 1995, the intrinsic dimension is estimated based on a nearest neighbor algorithm. Without LLP smoothing, the two methods show that the intrinsic dimension is some value between 4 and 5. Thus, it is reasonable to set the dimension of the linear space as 4 in LLP smoothing. After LLP smoothing, the intrinsic dimension is reduced to a value between 3 and 4. The numerical experiments indicate that LLP smoothing cannot only denoise, but also improve the efficiency of estimating the intrinsic dimension.

Another empirical way of estimating the intrinsic dimension is to analyze the sensitivity of the TRE to the different values of the intrinsic dimension. Figure 2-8 shows that the TRE is a decreasing function of the intrinsic dimension with an increasing slope. The slope of the curve in the figure has a dramatic change when the intrinsic dimension is around four. Therefore, we choose the intrinsic dimension as four here.

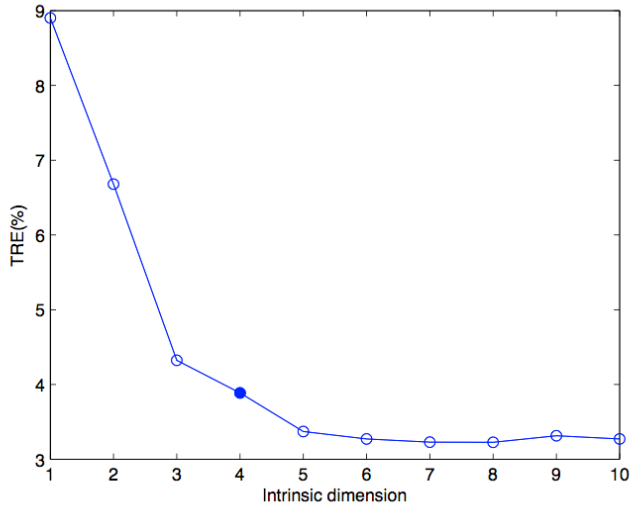


Figure 2-8: The sensitivity of TRE to the intrinsic dimension
(data length=731 days, number of the nearest neighbors=23).

2) The Number of the Nearest Neighbors: The plot of the TRE against the number of the nearest neighbors is used to select the appropriate number of the nearest neighbors. Figure 2-9 indicates the TRE first falls steeply when the number of the nearest neighbors is small, and then remains steady when the number of the nearest neighbors is greater than 22. We set the number of the nearest neighbors to be 23 for all the numerical studies. This is only one of the many choices as the construction error is not sensitive to the number of the nearest neighbors within a range.

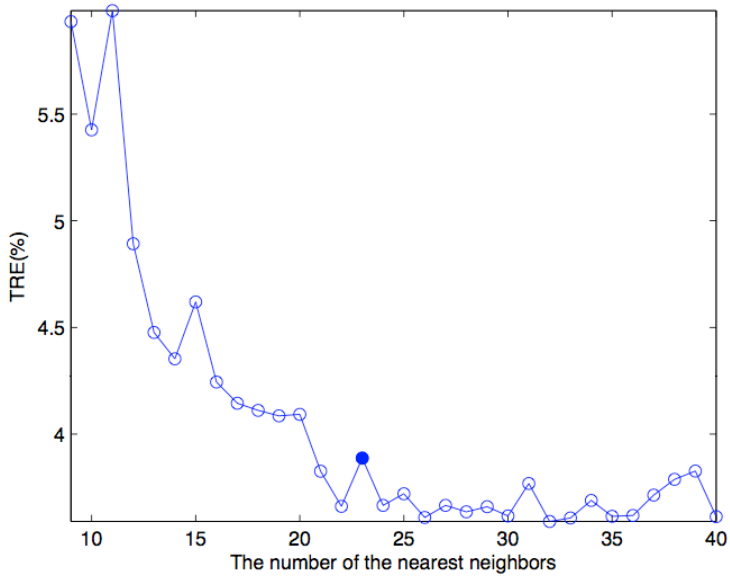


Figure 2-9: The sensitivity of TRE to the number of the nearest neighbors
(data length=731days, intrinsic dimension=4).

3) The Length of the Calibration Data: The plot of the TRE against the length of the calibration data in Figure 2-10 illustrates that the TRE is not very sensitive to the data length. Two years of data are applied to the manifold learning, and it helps to study whether there is yearly seasonality.

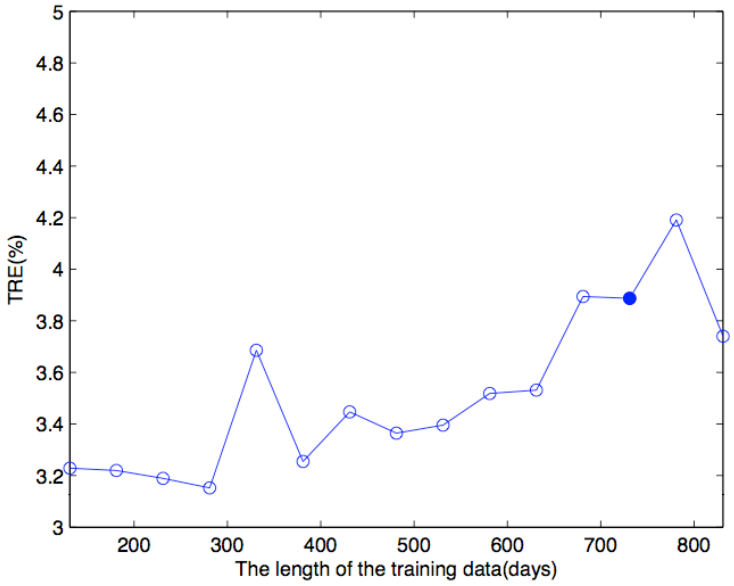


Figure 2-10: The sensitivity of TRE to the length of the calibration data (intrinsic dimension=4, number of the nearest neighbors=23).

2.4. Prediction of Electricity Price Curve

The prediction of future electricity price curves is an important issue in the electricity price market, because accurate predictions enable market participants to increase their profit by trading energy and hedge the potential risk successfully. However, it is difficult to make accurate predictions for the electricity prices due to their multiple seasonalities—daily and weekly seasonality. Unique features of the electricity price data often results in complicated models to forecast future electricity prices, which are often over fitting and fail to make accurate predictions in a longer horizon. Our method converts the hourly electricity price time series with multiple seasonalities into several time series with only weekly seasonality by manifold learning. After conversion, each data point in the new time series represents a day rather than an hour. The simplification of the new time series makes the longer horizon prediction easier and more accurate. Therefore, our method has an advantage in the longer horizon prediction over many other prediction methods.

A large amount of existing forecasting methods focus on one-day-ahead price predictions, i.e., the horizon of prediction is one day (24hours). Misiorek et al. 2006 and Conejo et al. 2005 give a good review on many prediction methods, and make a comparison on their performance. Here, we compare our prediction methods with three models —ARIMA, ARX and the naive method. The ARIMA model in Contreras et al. 2003 and the naïve method are pure time series

methods. The ARX model (also called dynamic regression model) includes the explanatory variable, load, and is suggested to be the best model in Conejo et al. 2005 and one of the best models in Misiorek et al. 2006. The longer horizon prediction has not drawn much attention so far. However, it also plays an important role in biding strategy and risk management. Our numerical results show that our prediction methods not only generate competent results in forecasting one-day-ahead price curves, but also produce more accurate predictions for one-week-ahead and one-month-ahead price curves, compared to ARX, ARIMA and the naive method. Moreover, as the new time series generated by manifold learning are simple, it is very easy to identify the time series models or utilize some nonparametric forecasting techniques. Our prediction methods also allow larger size of data for model calibration and incorporate more past information, but the size of the calibration data for ARIMA and ARX is often restricted to be several months.

2.4.1. Prediction Method

In our prediction method, we first make the prediction in the low-dimensional space, and then reconstruct the predicted price curves in the high-dimensional space from the low-dimensional prediction. There are three steps in detail:

1) Learn the low-dimensional manifold of electricity price curves with LLE. The sequence of each coordinates of the low-dimensional feature vectors comprises a time series.

2) Predict each time series in the low-dimensional space via univariate time series forecasting. Three prediction methods are applied: the Holt-Winters algorithm (HW), the structural model (STR) in Brockwell 2003 and the seasonal decomposition of time series by loess (STL) in Cleveland et al. 1990. Each data point in the time series represents one day, so for the one-week-ahead (one-day-ahead or one-month-ahead) price curve predictions, seven (one or 28) data points are forecasted for each time series.

3) Reconstruct the predicted price curves in the high-dimensional space from the predictions in low- dimensional space with LLE reconstruction.

The first and third step have been described in the previous sections. In the second step, we make the univariate time series forecasting for each coordinates of the feature vectors rather than making the multivariate time series forecasting for all the time series in the low-dimensional space, because the coordinates are orthogonal to each other.

There are a variety of methods of univariate time series forecasting, among which Holt-Winters algorithm, structural model and STL are selected. Both the Holt-Winters algorithm and structural model are pure time series prediction methods (models), and do not require any model identification as in ARIMA. The STL method can involve the explanatory variable in the prediction. All the prediction methods can be easily and fast implemented in statistical software R. The following is some brief description of the three prediction methods.

1) Holt-Winters Algorithm (HW): In Holt-Winters filtering, seasonals and trends are computed by exponentially weighted moving averages. In our numerical experiments, Holt-Winters algorithm is executed with starting period equal to 7 days and 14 days respectively. This

choice is due to the weekly effect of the electricity prices.

2) Structural Models (STR): Structural time series model is a (linear Gaussian) state-space model for (univariate) time series based on a decomposition of the series into a number of components—trend, seasonal and noise.

3) Seasonal Decomposition of Time Series by Loess (STL): The STL method can involve explanatory variables in the prediction. As the effect of temperature is usually embodied in electricity loads, only load is utilized as an exploratory variable. We first learn the manifold with the intrinsic dimension four for both prices and loads, and then decompose each time series in the low-dimensional space of price and load curves into seasonal, trend and irregular components using loess. Let $P_{i,t}$ and $Z_{i,t}$ denote the trend 2 of the ith coordinates of the feature vectors for prices and loads at time t. Then, we regress $P_{i,t}$ on $Z_{i,t}$ and the lagged $P_{i,t}$ with the lag three. As the relationship between prices and loads are dynamic, the history data we applied to train the model are 70 days. The model is written as:

$$P_{i,t} = \beta_0 + \beta_1 Z_{i,t} + \beta_2 P_{i,t-1} + \beta_3 P_{i,t-2} + \beta_4 P_{i,t-3} + \varepsilon_t$$

2.4.2. The Definition of Weekly Average Prediction Error

To assess the predictive accuracy of our methodology, three weekly average prediction errors are defined for one-day- ahead, one-week-ahead and one-month-ahead price predictions, respectively.

1) Weekly Average One-Day-Ahead Prediction Error: For the ith day of a certain week, $i = 1,\cdots,7$, the calibration data are set to be the two-year data right before this day, and then one-day-ahead predictions are made, i.e., the horizon of the prediction is one day. The predictions are denoted as $\hat{x}_{(i,1)}$, which is a 24-dimensional vector. The one-day-ahead prediction error for the ith day is defined as

$$WPE_d^{(i)} = \frac{1}{24} \frac{\| x_{(i,1)} - \hat{x}_{(i,1)} \|_1}{\bar{x}_{(i)}^d}$$

where $\bar{x}_{(i)}^d$ is the average of the actual electricity prices on the ith day. $\|\cdot\|_1$ is the L_1 norm of a vector, which is the sum of the absolute values of all the components in the vector.

The weekly average one-day-ahead prediction error is defined as

$$WPE_d = \frac{1}{7} \sum_{i=1}^{7} WPE_d^{(i)}$$

2) Weekly Average One-Week-Ahead Prediction Error: For the ith day of a certain week, $i = 1,\cdots,7$, the calibration data are set to be the two-year data right before this day, and then one-week-ahead predictions are made, i.e., the horizon of the prediction is one week. The jth-day-ahead predictions are denoted as $\hat{x}_{(i,j)}$, $j = 1,\cdots,7$. The one-week-ahead prediction error for the ith day is defined as

$$WPE_w^{(i)} = \frac{1}{7 \times 24} \sum_{i=1}^{7} \frac{\| x_{(i,j)} - \hat{x}_{(i,j)} \|_1}{\bar{x}_{(i)}^w}$$

where $\bar{x}_{(i)}^w$ is the average of the actual electricity prices of the one-week-ahead predictions.

The weekly average one-week-ahead prediction error is defined as

$$WPE_w = \frac{1}{7} \sum_{i=1}^{7} WPE_w^{(i)}$$

3) Weekly Average One-Month-Ahead Prediction Error:

For the ith day of a certain week, $i = 1, \cdots, 7$, the calibration data are set to be the two-year data right before this day, and then one-month-ahead (28-days-ahead) predictions are made, i.e., the horizon of the prediction is one month. The jth-day-ahead predictions are denoted as $\hat{x}_{(i,j)}$ $j = 1, \cdots, 28$. The one-month-ahead prediction error for the ith day is defined as

$$WPE_m^{(i)} = \frac{1}{28 \times 24} \sum_{j=1}^{28} \frac{\| x_{(i,j)} - \hat{x}_{(i,j)} \|_1}{\bar{x}_{(i)}^m}$$

where $\bar{x}_{(i)}^m$ is the average of the actual electricity prices of the one-month-ahead predictions.

The weekly average one-month-ahead prediction error is defined as

$$WPE_m = \frac{1}{7} \sum_{i=1}^{7} WPE_m^{(i)}$$

We define σ_d, σ_w and σ_m as the standard deviations of $WPE_d^{(i)}$, $WPE_w^{(i)}$ and $WPE_m^{(i)}$, respectively.

2.4.3. Prediction of Electricity Price Curves

Our numerical experiments are based on 12 weeks from February 2005 to January 2006, which consist of the second week of each month. Three weekly average prediction errors as defined above are calculated for each week, respectively. For each data set, the same parameter values taken from the previous section are used. The number of the nearest neighbors and the intrinsic dimension are set to be 23 and 4, respectively. Only one day, Jan 24, 2005, is identified with outliers. As we only have the forecasts of loads for six future days from the NYISO website, the weekly average one-week-ahead prediction error for STL and ARX is actually the weekly average six-days-ahead prediction error.

Table 2-3 and 2-4 provides the weekly average one-day-ahead prediction errors for the 12 weeks and their standard deviations. Our prediction methods—Holt-Winters, structural model and STL—are compared with ARX, ARIMA and the naive method. The naive predictions of a certain week are given by the actual prices of the previous week. Holt-Winters and structural model outperform all the other methods. It seems that involving the exploratory variable does not necessarily improve the prediction accuracy. STL performs slightly worse than Holt-Winters and structural model, and ARX also has less accuracy than ARIMA. This is not consistent with the results in Misiorek et al. 2006 and Conejo et al. 2005, where ARX has better performance than ARIMA. A potential cause is that the predictions of loads are not precise, or the correlation

between loads and prices is not high enough in NYPP.

Table 2-3: Comparison of $WPE_d(\%)$ of one - day- ahead predictions for 12 weeks.

date	HW7*	HW14	STR	STL	ARIMA	ARX	Naive
02/06/05 ~ 02/12/05	7.14	6.97	7.34	7.07	7.94	6.58	15.99
03/06/05 ~ 03/12/05	6.29	5.82	5.48	6.03	5.66	8.02	9.81
04/03/05 ~ 04/09/05	6.54	7.11	6.59	7.73	6.10	6.25	12.38
05/08/05 ~ 05/14/05	6.45	6.00	6.22	7.39	7.47	6.17	5.89
06/05/05 ~ 06/11/05	9.87	9.38	9.86	9.01	9.85	11.95	31.46
07/03/05 ~ 07/09/05	7.79	8.46	7.55	7.07	7.38	6.06	17.90
08/07/05 ~ 08/13/05	5.16	5.17	5.42	8.56	6.05	7.03	13.09
09/04/05 ~ 09/10/05	6.98	8.14	7.55	7.75	7.20	5.58	14.55
10/02/05 ~ 10/08/05	6.15	6.08	6.45	6.63	6.37	6.36	9.64
11/06/05 ~ 11/12/05	6.71	6.65	6.11	6.30	5.91	5.41	18.78
12/04/05 ~ 12/10/05	8.66	8.84	8.96	7.95	8.47	12.75	26.95
01/08/06 ~ 01/14/06	8.63	8.72	8.26	9.28	10.49	8.17	15.61
mean	7.20	7.28	7.15	7.56	7.41	7.53	16.00

*HW7 and HW14 stand for Holt-Winter algorithm with starting period equal to 7 days and 14 days respectively.

Table 2-4: Comparison of $\sigma_d(\%)$ of one - day- ahead predictions for 12 weeks.

date	HW7	HW14	STR	STL	ARIMA	ARX	Naive
02/06/05 ~ 02/12/05	3.03	3.34	3.10	2.21	4.58	2.28	4.66
03/06/05 ~ 03/12/05	1.89	1.96	1.81	2.21	2.16	3.15	5.67
04/03/05 ~ 04/09/05	2.24	2.88	2.34	2.76	1.83	3.13	3.33
05/08/05 ~ 05/14/05	3.86	3.36	3.47	3.34	4.07	2.54	1.40
06/05/05 ~ 06/11/05	3.31	3.36	4.03	2.53	4.31	6.13	5.62
07/03/05 ~ 07/09/05	3.62	4.68	3.73	4.01	3.37	2.15	8.79
08/07/05 ~ 08/13/05	1.65	1.53	2.28	3.59	3.01	2.63	4.37
09/04/05 ~ 09/10/05	3.68	3.59	3.09	3.83	4.48	2.16	8.38
10/02/05 ~ 10/08/05	2.27	2.02	2.26	3.36	2.18	3.06	4.45
11/06/05 ~ 11/12/05	2.57	2.41	2.60	2.57	2.36	3.08	7.06
12/04/05 ~ 12/10/05	4.14	4.27	3.69	2.36	2.80	6.62	10.55
01/08/06 ~ 01/14/06	4.61	4.42	3.73	3.93	5.79	4.49	4.71
mean	3.07	3.15	3.01	3.06	3.41	3.45	5.75

In Table 2-5 and 2-6, the weekly average one-week-ahead prediction errors for the 12 weeks and their standard deviations are presented. All of our prediction methods outperform ARX, ARIMA, and the naive method. The ARIMA model acts even worse than the naive method for one-week-ahead predictions. Since the ARIMA model is a very complicated model with multiple

seasonalities, it is often overfitting and makes the longer horizon predictions less accurate. The ARX model is a little simpler and given more information by the load forecasts, so it performs better than ARIMA. However, both ARX and ARIMA need to predict 168 data points for one-week-ahead predictions, while our prediction methods only need to predict seven data points for each time series. Therefore, our prediction methods have a great advantage in the longer horizon predictions. Among Holt-Winters, structural model and STL, STL has slightly worse performance than other two, and structural model is the most accurate.

Table 2-5: Comparison of $WPE_w(\%)$of one-week-ahead predictions for 12 weeks

date	HW7	HW14	STR	STL	ARIMA	ARX	Naive
02/06/05 ~ 02/12/05	8.31	8.33	7.65	8.28	14.57	9.00	12.19
03/06/05 ~ 03/12/05	10.85	10.59	10.19	12.57	13.28	15.30	12.33
04/03/05 ~ 04/09/05	11.03	10.88	10.53	11.74	10.68	11.46	14.59
05/08/05 ~ 05/14/05	7.55	7.15	7.36	6.89	14.21	6.16	6.71
06/05/05 ~ 06/11/05	15.58	15.23	13.88	11.39	21.64	19.49	26.68
07/03/05 ~ 07/09/05	10.85	10.04	10.41	9.60	14.63	7.03	14.44
08/07/05 ~ 08/13/05	6.82	7.09	6.42	12.87	9.49	9.14	10.28
09/04/05 ~ 09/10/05	7.46	7.95	7.31	8.52	10.36	6.86	13.17
10/02/05 ~ 10/08/05	9.78	9.60	10.11	8.97	11.84	8.30	11.57
11/06/05 ~ 11/12/05	9.45	9.44	8.99	9.42	11.24	9.00	15.18
12/04/05 ~ 12/10/05	12.94	13.09	13.30	11.55	21.78	22.92	23.94
01/08/06 ~ 01/14/06	13.95	14.08	13.28	15.00	26.01	16.37	11.52
mean	10.38	10.29	9.95	10.57	14.98	11.75	14.38

Table 2-6: Comparison of $\sigma_w(\%)$of one-week-ahead predictions for 12 weeks

date	HW7	HW14	STR	STL	ARIMA	ARX	Naive
02/06/05 ~ 02/12/05	1.32	1.66	1.14	1.85	7.06	2.62	2.35
03/06/05 ~ 03/12/05	3.80	3.73	3.86	3.88	7.06	2.22	1.18
04/03/05 ~ 04/09/05	3.12	3.48	3.41	2.90	3.50	4.18	2.47
05/08/05 ~ 05/14/05	3.45	3.13	3.50	1.99	6.28	1.82	0.88
06/05/05 ~ 06/11/05	5.91	6.10	6.46	4.64	10.30	8.18	4.75
07/03/05 ~ 07/09/05	3.10	3.44	3.43	1.94	5.51	0.55	1.50
08/07/05 ~ 08/13/05	1.02	0.95	1.40	3.10	3.77	2.60	1.48
09/04/05 ~ 09/10/05	3.19	2.96	2.33	2.75	4.88	1.17	2.28
10/02/05 ~ 10/08/05	2.06	1.73	2.01	4.32	5.50	3.11	1.34
11/06/05 ~ 11/12/05	2.17	2.42	2.30	2.40	3.92	2.14	3.05
12/04/05 ~ 12/10/05	4.88	4.41	5.17	4.53	13.77	4.47	4.65
01/08/06 ~ 01/14/06	3.25	3.37	2.15	6.15	11.60	3.70	2.66
mean	3.11	3.12	3.10	3.37	6.93	3.06	2.38

The proposed method can be applied to forecast prices in a longer horizon than one week, e.g., two weeks or even one month. As there are only a few methods associated with one-month-ahead price predictions, we apply three naive methods to compare with. The first naive method takes the last month prices in the calibration data set as the predictions. The second method repeats the last week prices four times, and the third one replicates the prices of last two weeks twice, respectively, as the predictions. Table 2-7 and 2-8 provide the weekly average prediction errors of the one-month-ahead price predictions for the 12 weeks and their standard deviations. The notations—naive1, naive2 and naive3—stand for the three naive methods. From the comparison, the proposed methods outperform all the naive methods. We notice that the total stand deviation of the structural model is larger than that of the naive methods, and it is mainly due to an inaccurate prediction for one day in week five. Thus, Holt-Winters algorithm has the best performance among all the methods for one-month-ahead price predictions.

Table 2-7: Comparison of $WPE_m(\%)$ of one-month-ahead predictions for 12 weeks

date	HW7	HW14	STR	Naive1	Naive2	Naive3
02/06/05 ~ 02/12/05	8.87	9.32	9.42	29.63	12.07	25.72
03/06/05 ~ 03/12/05	12.52	12.07	12.16	17.17	13.71	14.39
04/03/05 ~ 04/09/05	17.62	17.28	16.51	14.96	15.96	14.20
05/08/05 ~ 05/14/05	12.00	11.81	11.93	13.62	11.52	12.04
06/05/05 ~ 06/11/05	16.67	16.71	25.48	24.37	23.93	27.55
07/03/05 ~ 07/09/05	15.88	15.58	16.72	16.99	13.63	13.91
08/07/05 ~ 08/13/05	10.70	10.69	11.36	12.96	12.84	15.21
09/04/05 ~ 09/10/05	14.53	14.21	13.74	19.48	15.91	18.95
10/02/05 ~ 10/08/05	18.22	18.40	19.96	19.70	17.10	19.67
11/06/05 ~ 11/12/05	13.98	13.79	14.58	31.93	18.35	28.49
12/04/05 ~ 12/10/05	18.99	18.87	19.48	30.00	26.17	27.64
01/08/06 ~ 01/14/06	13.08	13.11	12.04	31.80	14.27	14.99
mean	14.42	14.32	15.28	21.88	16.29	19.40

Table 2-8: Comparison of $\sigma_m(\%)$ of one-month-ahead predictions for 12 weeks

date	HW7	HW14	STR	Naive1	Naive2	Naive3
02/06/05 ~ 02/12/05	1.13	1.55	2.63	0.18	2.79	6.82
03/06/05 ~ 03/12/05	3.64	3.33	4.11	0.47	0.46	1.01
04/03/05 ~ 04/09/05	3.02	2.06	3.49	0.46	2.25	1.08
05/08/05 ~ 05/14/05	1.22	1.18	1.22	0.40	1.38	0.65
06/05/05 ~ 06/11/05	4.59	4.28	25.41	0.36	5.96	0.62
07/03/05 ~ 07/09/05	3.95	4.50	3.58	1.15	2.90	0.89
08/07/05 ~ 08/13/05	0.35	0.36	1.03	0.62	1.24	0.58
09/04/05 ~ 09/10/05	3.22	3.82	2.82	0.92	0.50	0.46
10/02/05 ~ 10/08/05	4.20	4.01	6.80	0.80	2.77	1.95
11/06/05 ~ 11/12/05	1.09	1.21	1.62	1.51	3.47	3.77
12/04/05 ~ 12/10/05	2.40	2.21	3.29	0.54	4.33	1.82
01/08/06 ~ 01/14/06	6.04	6.25	3.44	4.16	1.83	1.76
mean	2.90	2.90	4.95	0.96	2.49	1.78

In summary, our prediction methods without an exploratory variable—Holt-Winters and structural model—outperform all of ARX, ARIMA and the naive method in both one-day-ahead and one-week-ahead predictions. STL is competent with ARX and ARIMA in one-day-ahead predictions, and performs better in one-week-ahead predictions. Our prediction methods have a great advantage in the longer horizon predictions spanning days to weeks.

3. An Equilibrium Pricing Model for Weather Derivatives

3.1. Overview of Weather Derivatives Market

Early trading of weather based instruments among energy companies started as over-the-counter (OTC) trades which means that each contract is individually negotiated. OTC trades are still used for weather derivatives for local cities which are not listed in exchanges. In September 1999, the first electronic market place for standardized weather derivatives was launched by the Chicago Mercantile Exchange (CME) with the aim of increasing liquidity, market integrity, and accessibility. This market experienced phenomenal growth and currently Cooling Degree Day (CDD) and Heating Degree Days (HDD) futures and options for 19 cities in the US, 9 cities in Europe 6 cities in Canada and 2 cities in Japan, are being traded on the CME. These include New York, Chicago, Philadelphia, London, Paris and Berlin. Other types of contracts based on frost days and snowfall are also traded on the CME. The weather derivatives markets are expanding rapidly as diverse industries seek to manage their exposure to weather risks. The notional value of CME weather products in 2004 was $2.2 billion, and grew ten-fold to $22 billion through September 2005, with volume surpassing 630,000 contracts traded[1]. In 2006 the value of traded weather instruments rose to 45 billion.

The most commonly traded weather indices are monthly or seasonal HDD/CDD strips. The calculation of CDD/HDD is based on the average daily temperature on a day i, which is defined as the average of the maximum and minimum temperature during that day, i.e.,

$$T_i = \frac{T_i^{\max} + T_i^{\min}}{2} \qquad (3.1)$$

From here on in referring to temperature, we mean daily average temperature. Daily CDD/HDD are defined as:

$$\text{daily CDD}_i = \max(T_i - 65^\circ F, 0)$$

$$\text{daily HDD}_i = \max(65^\circ F - T_i, 0)$$

Monthly or seasonal CDD/HDD can be defined by summing up daily CDD/HDD over the month or season. Seasonal strips bundle two or more consecutive months into a single contract. The HDD index can be interpreted as a measurement of the coldness during the contract periods relative to the industry standard 65°F at which people are supposed to feel comfortable. Similarly CDD is a measure of heat over the contract period relative to the 65°F norm.

The CME offers weather futures and options which are the same as financial futures and options except for the underlying basis. Weather options are agreements to buy or sell the value of the CDD/HDD index over the contract periods or alternatively can be interpreted as bets on the value of the CDD/HDD. Weather options give the owners the right, but not the obligation, to

[1] CME (2005). An introduction to CME weather products. www.cme.com/weather.

buy or sell at a specified strike level the specified weather index. A tradable weather derivative contract specifies six attributes: the contract type, the contract period, the underlying index, the contract city where the official temperature will be measured, the strike level, and the tick size (i.e., payoff in dollars per index unit). On the CME, for instance, the value of a degree day index, called a tick size, is $20. The contract period should be specified as a calendar month or seasonal strip from November to March for the HDD and May to September for the CDD.

In this chapter, we derive an equilibrium pricing model for weather derivatives and measure risk hedging and sharing gains that accrue to the market participants due to the inclusion of such instruments in their volumetric hedging portfolios. First, we will derive the optimal portfolio choices from the expected utility maximization problems of market participants. Using derived optimal demand, we calculate an equilibrium price for the weather derivative by applying the market clearing condition requiring that the aggregate demand be equal to the aggregate supply. The number of the weather derivative supplied will be decided based on the issuer's single period expected utility maximization problem, however, for some industries it may make sense to take short positions which effectively increase the supply of the shorted instruments and will affect their prices. Clearly, the primary role of weather derivatives is to hedge weather risk. In a single-commodity economy, the risk hedging gain is the only gain possible. To measure the risk hedging effect, we use the certain equivalent difference of maximized utility between two cases, with and without weather derivatives in a single-commodity economy. In a multi-commodity setting, weather derivatives also provide a mechanism for risk sharing. Any two agents share risk if they employ state-contingent transfers to increase the expected utility of both by reducing their risk. Such risk sharing is possible due to the diversity in exposure to weather risk and different risk preferences among industries participating in weather derivatives markets. We measure the risk sharing effect in terms of certain equivalent difference of maximized utility between a single and multi-commodity economy. We note that the risk sharing effect measured by the above method includes not only the risk sharing effect but also a price effect, since in the multi-commodity economy higher demand for weather derivatives due to more market participants makes the equilibrium price higher and paying more can reduce the maximized utility level of buyers. To correct such distortion, we adjust the risk aversion coefficient for the issuer so as to equalize the equilibrium price in a single and multi-commodity economy.

3.2. Pricing Model for Weather Derivatives

3.2.1. Assumptions and Notation

We assume that our economy is a frictionless endowment economy in a single-period planning horizon. It is implicit in the assumption of an endowment economy that the issuer of the weather derivative (underwriter) will supply a fixed number of derivatives at time 0 which will subsequently be traded in the market. Hence the price of the weather derivative is determined by the initial number of instruments issued and by the market demand. This aspect distinguishes our model from an actuarial based approach where an insurer issue as many contracts as demanded at a price determined by the issuer based on a stochastic model of temperature risk. In addition to the issuer who is typically a financial entity, our economy consists of weather-sensitive industries whose output is a commodity for which there is a liquid derivatives market (e.g., electricity, gas, wheat). We also assume that there are weather sensitive industries with no liquid derivative market for their output (e.g., tourism, sky resorts). The economy is closed in the sense

that all the supply and demand for weather derivatives comes from the parties described above. We further assume that none of the market participants is involved in speculative trades of commodity derivatives other than the commodity specific to their industry, thus, all hedging activities by parties involve derivatives of the commodity they produce (or consume) if available, and weather derivatives. We assume that all market participants are expected utility maximizers. We further assume that retail prices for all commodities are stable while wholesale prices and demand quantities are volatile and correlated with weather. This is definitely true in the energy industry, which is the primary focus of this part of the project. In the electric power industry, for instance, electric utilities have an obligation to serve all their customers' load at fixed regulated retail prices, while they procure the power in a competitive wholesale market where spot prices are highly volatile. Thus, the buyers' profit function is given by (retail price - wholesale spot price) times demand. From the profit function, we can see that each company faces not only spot price risk but also volumetric risk. We assume that the spot price, the demand, and temperature are all correlated.

In the stylized economy described above, there exist three types of financial assets; a risk free bond, a plain-vanilla weather call option with a strike K, and commodity derivatives that include forward contracts and European call and put options for which the underlying asset is the commodity spot price. All the financial instruments mature at time 1, at which point the physical commodity is delivered and paid for. Each agent can trade financial assets at time 0 to hedge its net revenue risks so as to maximize the expected utility at maturity. In other words, each trading party is faced with the problem of maximizing the expected utility of terminal wealth subject to a budget constraint at time 0. The issuer decides on the number of the weather call options supplied into the market so as to maximize her expected utility of terminal wealth at time 1.

Under the multi-commodity economy, the weather derivatives create two social welfare enhancing effects, a risk hedging effect and a risk sharing effect. When considering a single-commodity case, only the hedging effect is relevant and it can be measured by the certain equivalent difference of the maximized utility with and without weather derivatives. The risk sharing effect reflects possible diversification of weather risk across industries with different weather dependence (e.g. some industries may benefit from high temperature while others may be adversely affected). Such risk sharing effect can be measured by the certain equivalent difference of the maximized utility between the multi-commodity and a single-commodity economy. In Section 3.3 we will provide a general form of an equilibrium pricing model and numerical examples illustrating the results of our analysis.

Denote a probability space triplet by (Ω, F, P). Also let Q denote a risk-neutral probability measure.

Notation

- $i \in \{0,1,\ldots,m\}$: indices 0 to m-1 represent buyers where u buyers have a liquid derivatives market and v buyers don't and index m represents the issuer of weather derivatives.
- $U_i(\cdot)$: The utility function of type i industry where $U_i : R \to R$ is smooth, increasing and strictly concave on R and has a continuous derivative $U_i'(\cdot)$ on R.
- $\Pi_{i,n}$: The profit function of the type i at time n

- P_i^R : The unit retail price of type i industry
- P_i : The unit spot price of type i at terminal time
- $P_i^{'} = P_i^R - P_i$: The marginal profit of selling a type i commodity
- D_i: The random demand for type i at terminal time
- $B_n = (1+r)^n B_0$: The riskless bond price at time n, where r is the interest rate and B_0 = 1
- W_n: The weather derivative price
- $W_{i,n}^1$: The weather derivative price at time n in a type i industry economy
- $I_i(D_i, P_i)$: The income function of type i industry at terminal time
- $x_{i,n}(P_i)$: The portfolio payoff consisting of a risk-free bond, forwards, and call and put options with various strikes in a type i industry
- $\alpha_{i,n}$: A portfolio position of type i industry for the weather derivative at time n
- $J(\Pi_{i,n})$: The maximized expected utility of type i industry at time n
- $J_{wd}^1(\Pi_{i,n})$: The maximized expected utility of type i industry at time n in a single-commodity economy with the weather derivative and commodity derivatives
- $J_{wn}^1(\Pi_{i,n})$: The maximized expected utility of type i industry at time n in a single-commodity economy with the weather derivative and without commodity derivatives
- $J_{nd}^1(\Pi_{i,n})$: The maximized expected utility of type i industry at time n in a single-commodity economy without the weather derivative and with commodity derivatives
- $J_{nn}^1(\Pi_{i,n})$: The maximized expected utility of type i industry at time n in a single-commodity economy without the weather derivative and commodity derivatives
- $HE_{i,n}$: The hedging effect for the type i industry at time n
- $RS_{i,n}$: The risk sharing effect for the type i industry at time n

3.2.2. Multi-Commodity Economy

3.2.2.1. Utility Maximization Problem of Buyers with a Liquid Derivatives Market

We consider the utility maximization problem of buyers that have a liquid commodity derivatives market. For example, electricity and natural gas industry have liquid futures and options markets of which underlying asset is the spot price of electricity or natural gas. The buyer's profit function at time 1 is

$$\Pi_{i,1} = I_i(D_i, P_i) + x_{i,1}(P_i) + \alpha_{i,1} W_1 \tag{3.2}$$

where $x_{i,1}(P_i)$ represents the optimal payoff of the commodity portfolio which is the function of the commodity price P_i. The corresponding utility maximization problem of buyers at time 0 is

$$\begin{aligned} &\max_{\{x_{i,1}, \alpha_{i,1}\}} E_0[U_i(\Pi_{i,1})] \\ &s.t. \qquad E^{\mathbb{Q}}\left[\frac{x_{i,1}(P_i)}{B_1}\right] + \alpha_{i,1} W_0 = 0, \quad \forall i = 1, 2, \ldots, u \end{aligned} \tag{3.3}$$

where the constraint means that it costs zero to construct the portfolio with a commodity derivatives payoff $x_{i,1}(P_i)$ and weather derivatives payoff $\alpha_{i,1}W_0$. The expected discounted portfolio values under a risk-neutral probability measure Q is the price of the commodity derivatives' portfolio. Note that $x_{i,1}(P_i)$ is a decision variable. For each realization p of the random price P_i we will find the optimal payoff function $x(p)$ by solving the above maximization problem (3.3).

Carr and Madan (2001) show that any twice continuously differentiable function, $f(S)$, of the terminal stock price S can be replicated by a unique initial position of $f(S_0) - f'(S_0)S_0$ unit discount bonds, $f'(S_0)$ shares, and $f''(K)dK$ out-of-the-money options of all strikes K:

$$f(S) = [f(S_0) - f'(S_0)S_0] + f'(S_0)S + \int_0^{S_0} f''(K)(K-S)^+ dK + \int_{S_0}^{\infty} f''(K)(S-K)^+ dK \tag{3.4}$$

Using the result (3.4), the optimal payoff function $x(P_i)$ can be rewritten as:

$$\begin{aligned} x(P_i) = x(F)\cdot 1 - x'(F)(P_i - F) + \int_0^F x''(K)(K-P_i)^+ dK \\ + \int_F^{\infty} x''(K)(P_i - K)^+ dK \end{aligned} \tag{3.5}$$

where F denotes the forward price at time 0. Because 1, $(P_i - F)$, $(K-P_i)^+$ and $(P_i - K)^+$ represent the payoff of the bond, forward contracts, and European put and call options respectively, the equation (3.5) shows that the optimal payoff $x(P_i)$ can be replicated by $x(F)$ units of the risk-free bonds, $x'(F)$ units of forwards, $x''(K)dK$ units of European put options with strike K for all K < F, and $x''(K)dK$ units of European call options with strike K for all K > F . Since in reality there are no derivatives with continuous strikes, we need to approximate the replication by existing derivatives. Oum et al. (2006) suggest one possible way to replicate the optimal payoff by approximating the option positions $x''(K)dK$ to the mean of two available strike prices. Determining the best discrete approximation to our continuous optimal portfolio is out of the scope of this work and will not be elaborated any further.

Back to our constrained maximization problem (3.3), the corresponding Lagrangian function is

$$\begin{aligned} L(x_{i,1}(P_i), \alpha_{i,1}, \lambda_i) = E_0[U_i(\Pi_{i,1})] - \lambda_i(E_0^{\mathbb{Q}}[\frac{x_{i,1}(P_i)}{B_1}] + \alpha_{i,1}W_0) \\ = \int_{-\infty}^{\infty} E_0[U_i(\Pi_{i,1})|P_i = p] f_i(p)dp - \lambda_i(\frac{1}{B_1}\int_{-\infty}^{\infty} x_{i,1}(p)g_i(p)dp + \alpha_{i,1}W_0) \end{aligned} \tag{3.6}$$

where $f_i(p)$ is a marginal probability density function of commodity spot price P_i under the real probability measure P and $g_i(p)$ is a risk-neutral probability density function of P_i. If the commodity market is incomplete there may exist infinitely many risk-neutral probability

measures. The ratio $\frac{g_i(p)}{f_i(p)}$ is a Radon-Nicodym derivatives for the type i commodity and satisfies $E\left[\frac{g_i(p)}{f_i(p)}\right]=1$. One of the decision variables, $x_{i,1}(P_i)$, is a function of P_i and we need the Euler equation for the functional derivatives. Taking partial derivative with respect to $x_{i,1}(P_i)$, $\alpha_{i,1}$ and λ_i gives us the first order necessary conditions as

$$\frac{\partial L}{\partial x_{i,1}(p)} = E_0[U_i'(\Pi_{i,1})\frac{\partial \Pi_{i,1}}{\partial x_{i,1}(p)}|p]f_i(p) - \lambda_i \frac{g_i(p)}{B_1} = 0 \tag{3.7}$$

$$\frac{\partial L}{\partial \alpha_{i,1}} = E_0[U_i'(\Pi_{i,1})\frac{\partial \Pi_{i,1}}{\partial \alpha_{i,1}}] - \lambda_i W_0 = 0 \tag{3.8}$$

$$E_0^{\mathbb{Q}}[\frac{x_{i,1}(P)}{B_1}] + \alpha_{i,1} W_0 = 0 \quad \forall i = 1,2\ldots,u \tag{3.9}$$

Note that the first order conditions are sufficient for optimality because the utility function is assumed to satisfy $U'(\cdot) > 0$ and $U''(\cdot) < 0$. Moreover under the assumption $E_0[|U_i(\cdot)|] < \infty$ the partial derivative and the expectation operator are interchangeable. By solving the above three equations with three unknowns we can determine the optimal portfolio choices, $x_{i,1}(P_i)$ and $\alpha_{i,1}^*$ i.e., the structure of the optimal derivatives portfolio payoff of type i industry, and the quantity of weather derivatives that should be purchased at time 0 in order to maximize the expected utility. The optimal $\alpha_{i,1}^*$ is a function of an equilibrium price W_0 and will be used as a demand function for the weather derivative.

3.2.2.2. Utility Maximization Problem of Buyers without a Liquid Derivatives Market

If a type i industry does not have a liquid derivatives market, a risk-free bond and the weather derivative are the only available financial assets for hedging volumetric and price risk. Then the profit function is of the form

$$\Pi_{j,1} = I_j(D_j, P_j) + \alpha_{j,1} W_1 + \beta_{j,1} B_1 \tag{3.10}$$

where $\beta_{j,1}$ is the amount of money invested in a risk-free bond. The corresponding utility maximization problem of buyers at time 0 is;

$$\max_{\{\alpha_{j,1},\beta_{j,1}\}} E_0[U_j(\Pi_{j,1})]$$
$$s.t. \qquad \alpha_{j,1} W_0 + \beta_{j,1} B_0 = 0, \quad j = 1,2,\ldots,v \tag{3.11}$$

where the constraint implies that the portfolio with $\alpha_{j,1}$ and $\beta_{j,1}$ has zero cost at time zero. The constraint can be rewritten as $\beta_{j,1} = -\frac{\alpha_{j,1} W_0}{B_0}$ and can be substituted into the profit function (3.10). Then we have the unconstrained maximization problem:

$$\max_{\{\alpha_{j,1}\}} E_0[U_j(\Pi_{j,1})] \tag{3.12}$$

where $\Pi_{j,1} = I_j(D_j, P_j) + \alpha_{j,1}(W_1 - W_0 B_1)$. Since there is only one decision variable $\alpha_{j,1}$, the optimality condition will be

$$\frac{\partial E_0[U_j(\Pi_{j,1})]}{\partial \alpha_{j,1}} = 0 \tag{3.13}$$

From the above optimality condition (3.13), we can find the optimal quantity of weather derivatives in the case of no commodity derivatives market and the derived quantity will be regarded as the demand function for weather derivatives with an argument W_0.

3.2.2.3. Utility Maximization Problem of an Issuer

Now we consider the issuer's or under-writer's problem. The issuer is assumed to be a purely financial firm that specializes in weather derivatives and balances its budget by trading a risk-free bond, but it takes no positions in any commodity derivatives. At time 0 the underwriter will issue the weather derivatives and receive a price W_0. At time 1 the issuer will pay the realized payoff W_1 for the issued weather derivatives. Consequently the issuer's profit function at time 1 is

$$\Pi_{m,1} = \alpha_{m,1}(W_0 * B_1 - W_1) \tag{3.14}$$

The corresponding problem of the issuer is as follows.

$$\max_{\{\alpha_{m,1}\}} E_0[U_m(\Pi_{m,1})] \tag{3.15}$$

The first order condition is

$$\frac{\partial E_0[U_m(\Pi_{m,1})]}{\partial \alpha_{m,1}} = 0 \tag{3.16}$$

The issuer can determine the quantity of weather derivatives that will be supplied in this economy. Only this number of the weather derivatives, $\alpha^*_{m,1}$, will prevail in the market.

3.2.2.4. Equilibrium Price

We next derive an equilibrium pricing formula by applying the market clearing condition. The market clearing condition means the aggregate demand should be equal to the aggregate supply and can be graphically interpreted as the intersection of the demand and supply curve. The market clearing condition is given by

$$\sum_{i=1}^{u} \alpha_{i,1}^{*}(W_0) + \sum_{j=1}^{v} \alpha_{j,1}^{*}(W_0) = \alpha_{m,1}^{*}(W_0) \tag{3.17}$$

From the above equation we can derive the equilibrium price for the weather derivative. Then, the optimal choices can be expressed as real numbers and the maximized expected utility of the buyer i at time 1 with a liquid derivatives market, denoted by $J_{wd}(\Pi_{i,1})$, is

$$J_{wd}(\Pi_{i,1}) = E[U_i(I_i(D_i, P_i + x_{i,1}^{*}(P_i) + \alpha_{i,1}^{*} W_1)] \quad \forall i = 1, 2, \ldots, u \tag{3.18}$$

The certain equivalent, denoted by CE_{wd}, will be

$$CE_{wd} = U_i^{-1}(J_{wd}(\Pi_{i,1})) \tag{3.19}$$

If there is no liquid derivatives market the maximized utility and its certain equivalents are written as

$$J_{wn}(\Pi_{j,1}) = E[U_j(I_j(D_j, P_j) + \alpha_{j,1}^{*} W_1 + \beta_{j,1}^{*} B_1)] \quad \forall j = 1, 2, \ldots, v \tag{3.20}$$

$$CE_{wn} = U_j^{-1}(J_{wn}(\Pi_{j,1})) \tag{3.21}$$

3.2.3. Single Commodity Economy

3.2.3.1. Utility Maximization Problem of Buyers with a Liquid Derivatives Market

In the case that there is a weather derivative in the type i commodity economy the corresponding problems of the buyer and the issuer are exactly the same as the case of the multi-commodity economy case resulting in the same optimality conditions (3.7) and (3.16) for the buyer and the issuer respectively. The only difference between a single and multi-commodity economy is the market clearing condition since a single-commodity economy has one buyer and one issuer. Instead of the aggregated demand, we use a single buyer's demand function in the market clearing condition.

$$\alpha_{i,1}^{*}(W_{i,0}^{1}) = \alpha_{m,1}^{*}(W_{i,0}^{1}) \quad \forall i = 1, 2, \ldots, u \tag{3.22}$$

This condition gives us the equilibrium price for a single-commodity economy. The maximized expected utility of the buyer with the weather derivatives in a single-commodity economy is of the form

$$J_{wd}^{1}(\Pi_{i,1}) = E_0[U_i(I(D_i, P_i) + x_{i,1}^{*}(P_i) + \alpha_{i,1}^{*} W_{i,1}^{1})] \tag{3.23}$$

And the corresponding certain equivalent is $CE_{wd}^{1} = U_i^{-1}(J_{wd}^{1}(\Pi_{i,1}))$.

If the weather derivative is not available and only a risk-free bond and type i commodity derivatives are traded, the profit function of the buyers will be changed. The buyers will hedge risk only via the type i commodity derivatives and a risk-free bond. The issuer does not have any role in this case. The buyer's profit function is of the form

$$\Pi_{i,1} = I(D_i, P_i) + x_{i,1}(P_i) \tag{3.24}$$

The corresponding maximization problem is

$$\max_{\{x_{i,1}(P_i)\}} E_0[U_i(\Pi_{i,1})]$$

$$s.t. \qquad E^{\mathbb{Q}}[x_{i,1}(P_i)] = 0 \quad \forall i = 1, 2, \ldots, u \tag{3.25}$$

The above problem (3.25) is solved by Oum et al. (2006) under the CARA and the mean-variance utility function. The optimality condition can be obtained by defining the Lagrangian function and taking derivatives with respect to $x_{i,1}(P)$ and λ_i.

$$\frac{\partial L}{\partial x_{i,1}(P)} = E_0[U_i'(\Pi_{i,1}) \frac{\partial \Pi_{i,1}}{\partial x_{i,1}(P)} |p] f_i(p) - \lambda_i g_i(p) = 0 \tag{3.26}$$

$$E_0^{\mathbb{Q}}[x_{i,1}(P)] = 0 \quad i = 1, 2 \ldots, u \tag{3.27}$$

After solving the above optimality conditions we can find the maximized utility and the certain equivalent as

$$J_{nd}^1(\Pi_{i,1}) = E_0[U_i(I_i(D_i, P_i) + x_{i,1}^*(P_i))] \tag{3.28}$$

$$CE_{nd}^1 = U_i^{-1}(J_{nd}^1(\Pi_{i,1})) \tag{3.29}$$

3.2.3.2. Utility Maximization Problem without a Liquid Derivatives Market

If weather derivatives are available but there is no market for type j commodity derivatives, the optimality condition is the same as (3.13). After applying the market clearing condition with a single, not aggregated, demand and supply we can find a new equilibrium price in this single-commodity economy and the maximized utility has the form of

$$J_{wn}^1(\Pi_{j,1}) = E_0[U_j(I_j(D_j, P_j) + \alpha_{j,1}^* W_1^1 + \beta_{j,1}^* B_1)] \quad \forall j = 1, 2, \ldots, v \tag{3.30}$$

The corresponding certain equivalent is

$$CE_{wn} = U_j^{-1}(J_{wn}^1(\Pi_{j,1})) \tag{3.31}$$

If there is no weather derivatives and liquid derivatives market, the buyers are exposed to all the risk. Then the buyer's profit function is $\Pi_{j,1} = I_j(D_j, P_j)$ and the expected utility and the certain equivalent are written as

$$J_{nn}^1(\Pi_{j,1}) = E_0[U_j(\Pi_{j,1})] \quad \forall j = 1, 2, \ldots, v \tag{3.32}$$

$$CE_{nn}^1 = U_j^{-1}(J_{nn}^1(\Pi_{j,1})) \tag{3.33}$$

3.2.4. Hedging and Risk Effects

In this section, we evaluate the hedging and the risk sharing effect. As mentioned before one role of the weather derivative is to hedge the volumetric risks of the buyers. The hedging effect can be measured by subtracting the certain equivalent without the weather derivative from the

certain equivalent with the weather derivative in a single-commodity economy. Thus, with a liquid derivatives market the hedging effect is given by

$$HE_{i,1} = CE^1_{wd} - CE^1_{nd} \quad \forall i = 1, \ldots, u \tag{3.34}$$

Without a liquid derivatives market the hedging effect is given by

$$HE_{j,1} = CE^1_{wn} - CE^1_{nn} \quad \forall j = 1, \ldots, v \tag{3.35}$$

The risk sharing effect exists only in the multi-commodity economy with weather derivatives. By holding weather derivatives buyers can increase their expected utilities. The difference between the certain equivalent in a single and multi-commodity economy will represent the risk sharing effect. With a liquid commodity derivatives market the risk sharing effect is

$$RS_{i,1} = CE_{wd} - CE^1_{wd} \quad \forall i = 1, \ldots, u \tag{3.36}$$

Without a liquid commodity derivatives market the risk sharing effect is

$$RS_{j,1} = CE_{wn} - CE^1_{wn} \quad \forall j = 1, \ldots, v \tag{3.37}$$

3.3. Mean-Variance Utility Case

3.3.1. Multi-Commodity Economy

Under the mean variance utility the Lagrangian function corresponding to problem (3.3) is given by

$$\begin{aligned} L(x_{i,1}(P_i), \alpha_{i,1}) &= \mu_{I_i} + \mu_{x_{i,1}} + \alpha_{i,1}\mu_{W_1} - \frac{\nu_i}{2}(\sigma^2_{I_i} + \sigma^2_{x_{i,1}} + \alpha^2_{i,1}\sigma^2_{W_1} \\ &+2(\sigma_{I_i x_{i,1}} + \alpha_{i,1}\sigma_{x_{i,1}W_1} + \alpha_{i,1}\sigma_{W_1 I_i})) - \lambda_i(E^{\mathbb{Q}}[\frac{x_{i,1}(P_i)}{B1}] + \alpha_{i,1}W_0) \end{aligned} \tag{3.38}$$

If we take the point-wise partial derivative with respect to $x^*_{i,1}(p)$ and the partial derivatives with respect to $\alpha_{i,1}$and λ_i we have the following optimality conditions.

$$\begin{aligned} \frac{\partial L}{\partial x_{i,1}(p)} &= f_i(p)(1 - \nu_i(x_{i,1}(p) - E[x_{i,1}(P_i)] + E[I_i|p] - \mu_{I_i} \\ &\quad +\alpha_{i,1}(E[W_1|p] - \mu_{W_1}))) - \lambda_i \frac{g_i(p)}{B_1} \end{aligned} \tag{3.39}$$

$$\frac{\partial L}{\partial \alpha_{i,1}} = \mu_{W_1} - \nu_i(\sigma^2_{W_1}\alpha_{i,1} + \sigma_{x_{i,1}W_1} + \sigma_{W_1 I_i}) - \lambda_i W_0 = 0 \tag{3.40}$$

$$E^{\mathbb{Q}}_0[\frac{x_{i,1}(P)}{B_1}] + \alpha_{i,1}W_0 = 0 \quad \forall i = 1, 2, \ldots, u \tag{3.41}$$

By solving the above three equations we obtain the optimal $x^*_{i,1}(P_i)$ as

$$x_{i,1}^*(P_i) = \frac{1}{\nu_i}\left(\frac{E^{\mathbb{Q}}[\frac{g_i(P_i)}{f_i(P_i)}] - \frac{g_i(P_i)}{f_i(P_i)}}{E[\frac{g_i(P_i)}{f_i(P_i)}]}\right) - (E[I_i|P_i] - E^{\mathbb{Q}}[E[I_i|P_i]])$$
$$+E^{\mathbb{Q}}[x_{i,1}^*(P_i)](1 + \frac{1}{W_0 B_1}(E[W_1|P_i] - E^{\mathbb{Q}}[E[W_1|P_i]])) \quad (3.42)$$

where

$$E^{\mathbb{Q}}[X_{i,1}(P_i)] = \frac{B_1 W_0}{E[E[W_1|P_i]W_1] - \mu_{W_1}^2 - \sigma_{W_1}^2}$$
$$\left(\frac{1}{\nu_i}\left(\frac{E[\frac{g_i(P_i)}{f_i(P_i)}W_1]}{E[\frac{g_i(P_i)}{f_i(P_i)}]} - B_1 W_0\right) - (\mu_{I_i}\mu_{W_1} - E[E[I_i|P_i]W_1]) - \sigma_{W_1 I_i}\right) \quad (3.43)$$

Finally, we can find the optimal quantity of the weather derivative from the optimality condition (3.41) as

$$\alpha_{i,1}^* = \frac{\frac{1}{\nu_i}\left(\frac{E[\frac{g_i(P_i)}{f_i(P_i)}W_1]}{E[\frac{g_i(P_i)}{f_i(P_i)}]} - B_1 W_0\right) - (\mu_{I_i}\mu_{W_1} - E[E[I_i|P_i]W_1]) - \sigma_{W_1 I_i}}{\mu_{W_1}^2 + \sigma_{W_1}^2 - E[E[W_1|P_i]W_1]} \quad (3.44)$$

If there is no commodity derivatives market in the type j commodity we have the following optimal choice from the equation (3.13).

$$\alpha_{j,1}^* = \frac{\mu_{W_1} - B_1 W_0 - \nu_j \sigma_{I_j W_1}}{\nu_j \sigma_{W_1}^2} \qquad \forall j = 1, \ldots, v \quad (3.45)$$

$\alpha_{j,1}^*$ is linear in W_0.

Next the issuer's problem (3.15) under the mean-variance case becomes;

$$\max_{\alpha_{m,1}} \alpha_{m,1}(B_1 W_0 - \mu_{W_1}) - \frac{\nu_m}{2}\sigma_{W_1}^2 \alpha_{m,1}^2 \quad (3.46)$$

From the optimality condition (3.16), the number of the weather derivative supplied in this economy will be

$$\alpha_{m,1}^* = \frac{B_1 W_0 - \mu_{W_1}}{\nu_m \sigma_{W_1}^2} \quad (3.47)$$

Here the supply function for weather derivative is also linear in W_0. Therefore, in the mean-variance utility function, the demand and supply of the weather derivative are all linear. The intersection of the aggregate demand and the supply function will clear the weather derivative market. In other words, the equilibrium price of the weather derivative can be calculated from the following market clearing condition.

$$\sum_{i=1}^{u} \frac{\frac{1}{\nu_i}\left(\frac{E[\frac{g_i(P_i)}{f_i(P_i)}W_1]}{E[\frac{g_i(P_i)}{f_i(P_i)}]} - B_1W_0\right) - (\mu_{I_i}\mu_{W_1} - E[E[I_i|P_i]W_1]) - \sigma_{W_1 I_i}}{\mu_{W_1}^2 + \sigma_{W_1}^2 - E[E[W_1|P_i]W_1]}$$
$$+\sum_{j=1}^{v} \frac{\mu_{W_1} - B_1W_0 - \nu_j\sigma_{I_jW_1}}{\nu_j\sigma_{W_1}^2} = \frac{B_1W_0 - \mu_{W_1}}{\nu_m\sigma_{W_1}^2} \tag{3.48}$$

From the above equation, the equilibrium price of the weather derivative can be obtained if we specify a probability density function of the spot price Pi for the commodity i under P and Q and the income structure of type i industry. As a result, the maximized utility of type i buyer is

$$J_{wd}(\Pi_{i,1}) = U_i(I_i(D_i, P_i) + x_{i,1}^*(P_i) + \alpha_{i,1}^* W_1) \qquad \forall i = 1, \ldots, u \tag{3.49}$$

The maximized utility of type j buyer without a liquid commodity derivatives market is

$$J_{wn}(\Pi_{j,1}) = U_j(I_j(D_j, P_j) + \alpha_{j,1}^* W_1 + \beta_{j,1}^* B_1) \qquad \forall j = 1, \ldots, v \tag{3.50}$$

3.3.2. Single-Commodity Economy

In a single-commodity economy with weather derivatives and a liquid commodity derivatives market, there is one buyer and one issuer. The revised market clearing condition under the mean-variance utility is

$$\frac{\frac{1}{\nu_i}\left(\frac{E[\frac{g_i(P_i)}{f_i(P_i)}W_1]}{E[\frac{g_i(P_i)}{f_i(P_i)}]} - B_1W_0\right) - (\mu_{I_i}\mu_{W_1} - E[E[I_i|P_i]W_1]) - \sigma_{W_1 I_i}}{\mu_{W_1}^2 + \sigma_{W_1}^2 - E[E[W_1|P_i]W_1]}$$
$$= \frac{B_1W_0 - \mu_{W_1}}{\nu_m\sigma_{W_1}^2} \tag{3.51}$$

The maximized utility in this case will be

$$J_{wd}^1(\Pi_{i,1}) = U_i(I(D_i, P_i) + x_{i,1}^*(P_i) + \alpha_{i,1}^* W_1) \quad \forall i = 1, 2, \ldots, u \tag{3.52}$$

If there is a weather derivative but no commodity derivatives market, the market clearing condition is

$$\frac{\mu_{W_1} - B_1W_0 - \nu_j\sigma_{I_jW_1}}{\nu_j\sigma_{W_1}^2} = \frac{B_1W_0 - \mu_{W_1}}{\nu_m\sigma_{W_1}^2} \tag{3.53}$$

And the maximized utility is given by

$$J_{wn}^1(\Pi_{j,1}) = U_j(I(D_j, P_j) + \alpha_{j,1}^* W_1 + \beta_{j,1}^* B_1) \quad \forall j = 1, 2, \ldots, v \tag{3.54}$$

In a single-commodity economy without weather derivatives and with a liquid derivatives market, the optimal solution to the type j buyers is

$$x_{i,1}^*(p) = \frac{1}{\nu_i}\left(E^{\mathbb{Q}}\left[\frac{g_i(p)}{f_i(p)}\right] - \frac{g_i(p)}{f_i(p)}\right) + E^{\mathbb{Q}}[E[I_{i,1}|p]] - E[I_{i,1}|p] \tag{3.55}$$

Then the buyer's maximized utility is given by

$$J_{nd}^{1}(\Pi_{i,1}) = U_i(I_i(D_i, P_i) + x_{i,1}^{*}(P_i)) \quad \forall i = 1, 2, \ldots, u \tag{3.56}$$

If there are no weather derivatives and liquid commodity derivatives market, the buyer's maximized utility is the form of

$$J_{nn}^{1}(\Pi_{j,1}) = U_j(I_j(D_j, P_j)) \quad \forall j = 1, 2, \ldots, v \tag{3.57}$$

Now we can measure the hedging effect (3.34) and (3.35) and the risk sharing effect (3.36) and (3.37) under the mean-variance preference.

3.4. Numerical Example

In this section we illustrate the equilibrium pricing model and the risk hedging and sharing effects with a numerical example based on the mean variance utility function. We apply our pricing model to a plain-vanilla weather call option with a strike of 85°F, which may be in the money during a hot summer day. However, this example can be extended to CDD/HDD indexed call or put options if we specify the probability density functions of CDD/HDD indices during the contract period. We assume that there are 5 market participants, an issuer and the four types of the buyers. Each buyer's commodity demand and spot price are positively or negatively correlated with temperature and they may have a liquid derivatives market. For convenience we label the four buyers as buyer 1 trough buyer 4. Buyer 1, which may be an electricity distribution company, faces positive correlation among demand, spot price, and temperature and it can trade commodity derivatives in the liquid derivatives market. Demand faced by buyer 2 is negatively correlated with temperature and there are tradable commodity derivatives. Buyer 3 faces a positive correlation between demand and temperature but does not have a commodity derivatives market. Buyer 4 faces a negative correlation between demand and temperature and has no derivatives market to trade in. All the buyer types have the same form of the income function which reflects selling at a fixed retail price and buying in a volatile wholesale price (a typical situation for energy utilities in the restructured electricity or gas markets), i.e. $I_i(D_i, P_i) = (P_i^R - P_i)D_i$. We also assume that P = Q in each commodity market. This assumption has been justified in the Nordic electricity market by Audet et al. (2004).

We define temperature, demand, and the spot price as

$$T = \mu_T + \sigma_T Z \tag{3.58}$$

$$D_i = e^{a_i Z + b_i Z_{i,1} + c_i} \tag{3.59}$$

$$P_i = e^{d_i Z + e_i Z_{i,1} + f_i Z_{i,2} + g_i}, \forall i = 1,2,3,4 \tag{3.60}$$

where a_i, b_i, c_i, d_i, e_i, f_i, and g_i are constant and Z, $Z_{i,1}$, and $Z_{i,2}$ are independent standard normal random variables. We then have the mean vector of demand and the spot price as

$$(\mu_{D_i}, \mu_{P_i}) = (e^{c_i + 0.5(a_i^2 + b_i^2)}, e^{g_i + 0.5(d_i^2 + e_i^2 + f_i^2)}) \tag{3.61}$$

and the covariance matrix Σ as shown in Table 3-1. Because we have 7 parameters and 7 equations from the mean vector and the covariance matrix for each i, we can determine the

parameters a_i , b_i , . . . , g_i so that random variables T , D_i , and P_i have the specified correlations.

Table 3-1: Covariance matrix

	T	D_i	P_i
T	σ_T^2	$a_1\sigma_T e^{0.5(a_i^2+b_i^2)+c_i}$	$d_i\sigma_T e^{0.5(d_i^2+e_i^2+f_i^2)+g_i}$
D_i	·	$e^{2c_i+a_i^2+b_i^2}(e^{a_i^2+b_i^2}-1)$	$e^{c_i+g_i+0.5(a_i^2+b_i^2+d_i^2+e_i^2+f_i^2)}(e^{a_i d_i+b_i e_i}-1)$
P_i	·	·	$e^{2g_i+d_i^2+e_i^2+f_i^2}(e^{d_i^2+e_i^2+f_i^2}-1)$

In this example, we vary the correlation coefficient corresponding to buyer 1, denoted by ρ with temperature. All other parameters are fixed. In addition, temperature T is assumed to be normally distributed as N(80,25). The risk aversion coefficients of all buyers are equal to 0.1. The issuer's risk aversion coefficient is assumed to be 0.01. Table 3-2 shows the correlation coefficient corresponding to the four buyers among temperature, demand, and spot price. Other market parameters for buyer 1 which represent an electricity distribution company are estimated from Energy Information Administration web sites[2]. Demand and spot price are 9.474 × 10^6MWh and $99.47/MWh. Finally, we assume that the variances of demand and spot price of buyers are 3 and 2; 2 and 2; 2 and 2;and 2 and 2.

Table 3-2: Correlation Coefficient of the Buyers

(a) Buyer 1

	T	D_1	P_1
T	1	vary	0.3
D_1	vary	1	0.7
P_1	0.3	0.7	1

(b) Buyer 2

	T	D_2	P_2
T	1	-0.7	-0.6
D_2	-0.7	1	0.7
P_2	-0.6	0.7	1

(c) Buyer 3

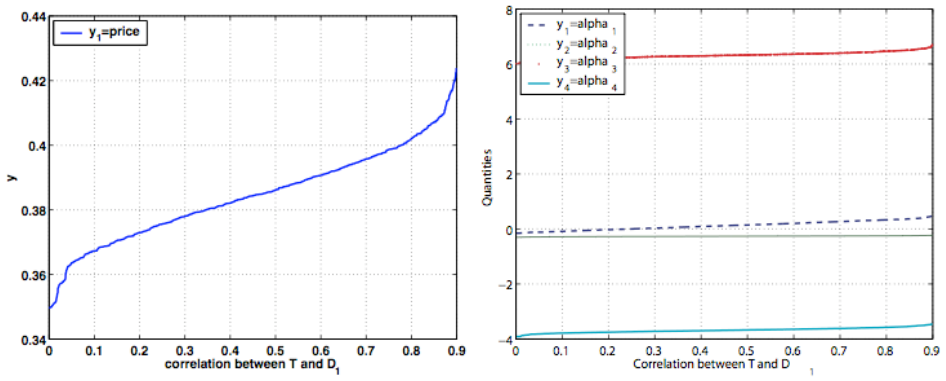
	T	D_2	P_2
T	1	0.8	0.8
D_2	0.8	1	0.7
P_2	0.8	0.7	1

(d) Buyer 4

	T	D_2	P_2
T	1	-0.8	-0.8
D_2	-0.8	1	0.7
P_2	-0.8	0.7	1

Figure 3-1 shows the equilibrium price and the optimal quantities of weather call options of the four buyers. As the correlation between T and D_1 increases, the price and the optimal quantities also increase. This means that all buyers want to buy more weather call options to effectively hedge the volumetric risk. High correlation and the increased demand $\alpha_{1,1}^*$ cause the higher equilibrium price.

[2] www.eia.doe.gov



(a) Equilibrium Price (b) Optimal Choices

Figure 3-1: Equilibrium Price and Choices

Figure 3-2 shows the linear aggregated demand and supply curves under the mean-variance utility function with correlation 0.8.

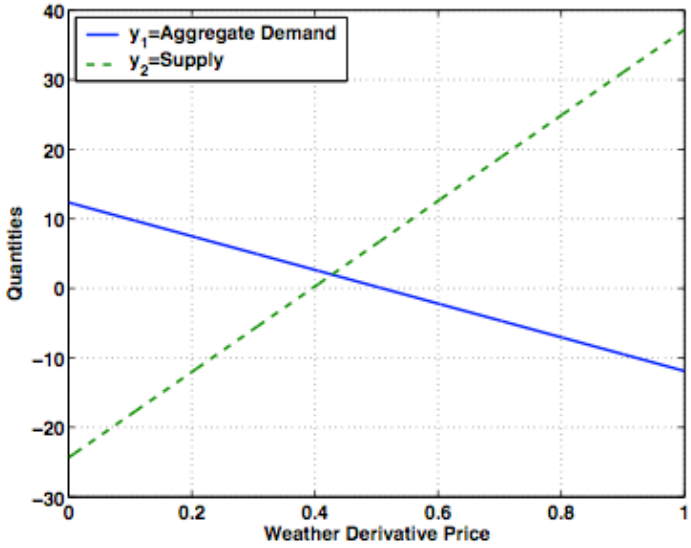
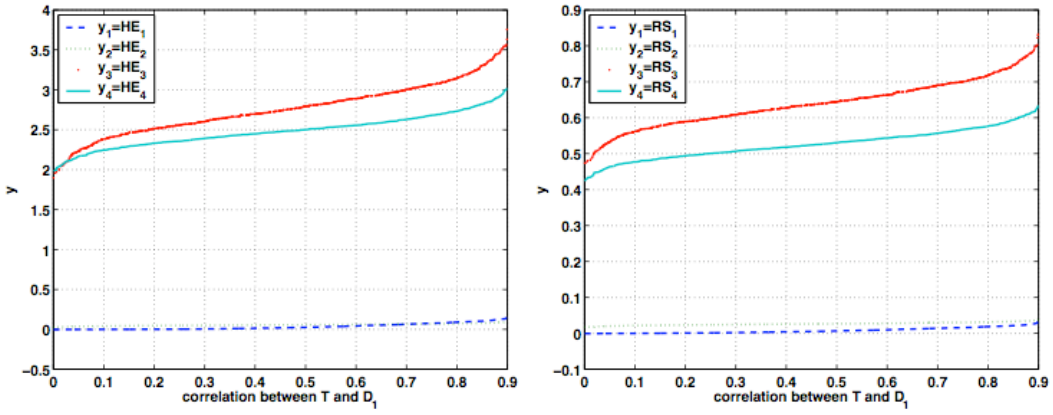


Figure 3-2: Supply and Demand Curve

The hedging and the risk sharing effects corresponding to each buyer are shown in Figure 3-3. Figure 3-3 demonstrates how buyers can increase their maximum utility by employing weather call options under high correlation between T and D_1. The risk hedging and sharing effects for buyer 1 and 2 are relatively small but more significant for buyers 3 and 4. This result is due to the existence of commodity derivatives. Because buyers 1 and 2 can construct portfolios consisting of commodity derivatives with continuous strike prices and the weather call options are only one part of their hedging portfolios, the impact of the weather call option for buyers 1 and 2 is much smaller than the corresponding impact for buyers 3 and 4 that must rely only on weather call options in order to hedge risk. Importantly, if the strike prices of commodity options are discrete with only a few number of the strike prices, which is more realistic, the risk hedging and sharing effects attributable to weather derivatives are expected to be higher.

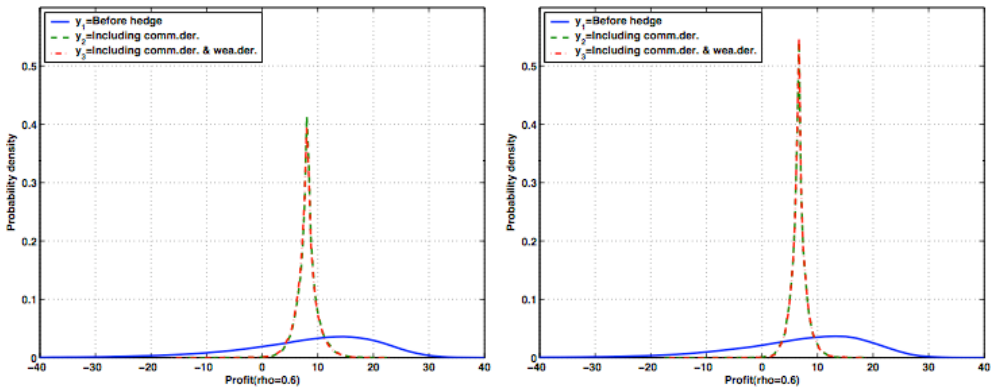


(a) Risk Hedging Effect (b) Risk Sharing Effect

Figure 3-3: Hedging and Risk Sharing Effects

The utility improvement shown by Figure 3-3 is caused by reducing the variance of demand and spot price via the weather call option and other commodity derivatives (if available) and sharing the volumetric risk with other market participants. This risk (or variance) reduction can be shown by the probability density functions (p.d.f.) of the profit function Πi before and after hedging risk.

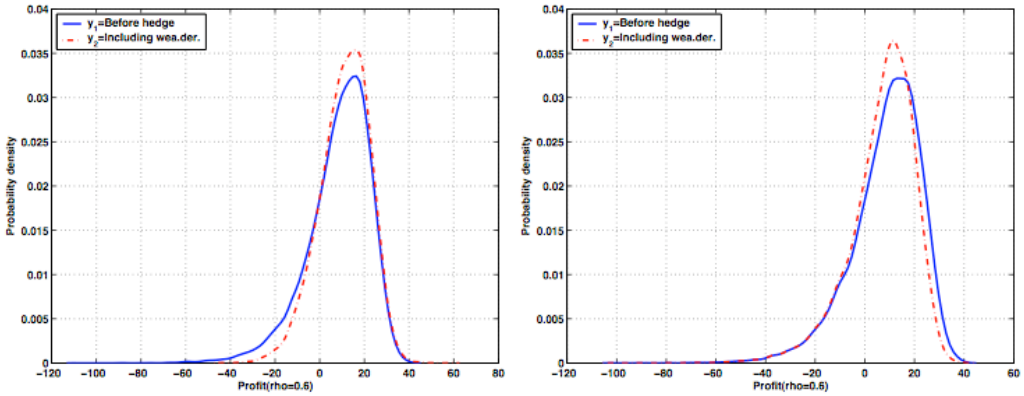
Figure 3-4 illustrates the p.d.f. of the buyers 1 and 2's profit functions for three cases; exposure to all risk, after hedging with commodity derivatives only, and after hedging with commodity derivatives and the weather call option. The p.d.f. before hedging is widely spread, which means that the buyer is exposed to high net revenue risk, but after including commodity derivatives, the risk is greatly reduced. However, the p.d.f. of the profit function after including commodity derivatives and the weather call option is very similar to the p.d.f. of the profit function with commodity derivatives only.



(a) Buyer 1 (b) Buyer 2

Figure 3-4: P.D.F. of Buyer 1 and 2's Profit Function (ρ1=0.6)

Figure 3-5 shows the p.d.f. of the profit function for buyers 3 and 4 before and after hedging. By employing the weather call options, buyers 3 and 4 can reduce the variance of the profit function, which improves their mean-variance utility values.



(a) Buyer 3 (b) Buyer 4

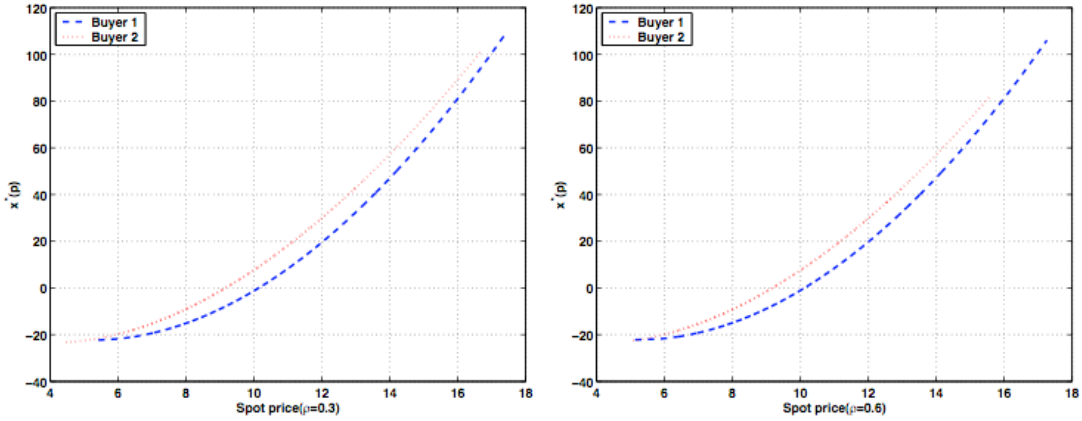
Figure 3-5: P.D.F. of Buyer 3 and 4's Profit Function (ρ1=0.6)

Table 3-3 shows the variance of the profit function when the buyer1's correlation to temperature is 0.6.

Table 3-3: Variance of the Profit Function

	Buyer 1	Buyer 2	Buyer 3	Buyer 4
Before Hedging	170.0013	164.4937	192.6060	197.4552
Including Comm.Der.	3.9816	2.5077		
Including Wea.Der.			126.3753	175.2812
Including Comm.Der. and Wea. Der.	3.9410	2.4158		

Finally, Figure 3-6 illustrates the optimal payoff of the portfolio with commodity derivatives for buyers 1 and 2 when the correlation ρ between temperature T and demand D_1 is 0.3 and 0.6. Because we change the correlation ρ but not the correlation between temperature and spot price, the resulting graphs look very similar.



(a) Optimal $x^*(P)(\rho = 0.3)$ (b) Optimal $x^*(P)(\rho = 0.6)$

Figure 3-6: Optimal Payoff $x^*(P)$ of the Commodity Derivatives Portfolio

4. Static Hedging of Volumetric Risk

4.1. Optimal Static Hedging in a Single-period Setting

4.1.1. Obtaining the Optimal Hedge Payoff Function

Consider an LSE who is obligated to serve an uncertain electricity demand q at the fixed price r. Assume that the LSE procures electricity that it needs in order to serve its customers, from the wholesale market at spot price p. To protect against price risk, the LSE can enter into forward contracts to fix the buying price at the forward price F. First, the number of forward contracts to be purchased needs to be determined. Suppose that the LSE decides to purchase an amount $\overline{q}$ of forwards; then, the actual demand would be $\overline{q}+\Delta q$. Then, the profit that is at risk is $\Delta q \cdot (r - p)$. The LSE would want to protect against the situation where either spot price p is higher than r and $\Delta q > 0$, or p is less than r and $\Delta q < 0$. The LSE's strategy could be buying call options with strike prices which are higher than r and exercised when $\Delta q > 0$ and buying put options with strike prices less than r and exercised when $\Delta q < 0$. Of course, prices of the call/put options are not negligible.

Suppose that a hedging portfolio consisting of electricity derivatives is constructed at time 0 whose payoff at time 1, x(p), is a function of the spot price p at time 1, is received at time 1. The hedging portfolio may also include money market accounts, allowing the LSE to finance hedging instruments through loans payable at time 1. Let y(p, q) be the LSE's profit from serving the customers' demand q at the fixed retail rate r at time 1. Then, the hedged profit Y (p, q, x(p)) - total profit including the net payoffs of the hedging portfolio - is given by

$$Y(p,q,x(p)) = y(p,q) + x(p) \tag{4.1}$$

where

$$y(p,q) = (r - p)q$$

The LSE's risk preference is characterized by a concave utility function U defined over the total profit Y (·) at time 1. LSE's beliefs on the realization of spot price p and load q are characterized by a joint probability function f (p, q) for positive p and q, which is defined on the probability measure P. On the other hand, let Q be a risk-neutral probability measure based on which the hedging instruments are priced, and g(p) be the probability density function of p under Q. Because the electricity market is incomplete, there may exist infinitely many risk-neutral probability measures. In our work, it is assumed that a specific measure, Q, was picked according to some criteria.

Then, the formulation of the optimal static hedging problem is as follows:

$$\begin{aligned} &\max_{x(p)} \; E\left[U[Y(p,q,x(p))]\right] \\ &s.t. \; E^{Q}[x(p)] = 0 \end{aligned} \tag{4.2}$$

Where $E[\cdot]$ and $EQ[\cdot]$ denote expectations under the probability measure P and Q, respectively; Y is a function that defines the profit; x(p) is the outcome of the hedge portfolio; U is a concave utility function which can give us a unique solution to the maximization problem. It is important to note that E(x(p)) = 0; which implies that the cost of setting up the hedge portfolio is zero.

We require the manufacturing cost of the portfolio to be zero under a constant risk-free rate. This zero-cost constraint implies that purchasing derivative contracts may be financed from selling other derivative contracts or from the money market accounts. In other words, under the assumption that there is no limits on the possible amount of instruments to be purchased and money to be borrowed, our model finds a portfolio from which the LSE obtains the maximum expected utility over total profit.

The goal here is to find a portfolio x(p) that maximizes the utility function. This is done by considering the Lagrangain function of the optimization as a function of x(p), and equating its partial derivative with x(p) with zero (since we are trying to maximize the problem).

CARA Utility (Proposition 4.1)

A CARA utility function has an exponential form:

$$U(Y) = -\tfrac{1}{a}e^{-aY}$$

where a is the coefficient of absolute risk aversion. With CARA utility, the optimal payoff function $x^*(p)$, which satisfies the above equation, is obtained as

$$\begin{aligned} x^*(p) =& \frac{1}{a}\Big(\ln \frac{f_p(p)}{g(p)} + \ln E\big[e^{-ay(p,q)}\big|p\big]\Big) \\ & - \frac{1}{a}\Big(E^Q\big[\ln \frac{f_p(p)}{g(p)}\big] + E^Q\big[\ln E\big[e^{-ay(p,q)}\big|p\big]\big]\Big) \end{aligned} \tag{4.3}$$

Mean Variance Approach (Proposition 4.2)

The mean-variance approach is to maximize a mean-variance objective function, which is linearly increasing in the mean and decreasing in the variance of the profit:

$$E[U(Y)] \;=\; E[Y] \;-\; \tfrac{1}{2}aVar(Y)$$

It follows from Var(Y) = $E[Y^2] - E[Y]^2$ that

$$U(Y) \equiv Y - \frac{1}{2}a(Y^2 - E[Y]^2)$$

for the mean-variance objective function in an expected utility form. Then, the optimal solution $x^*(p)$ that satisfies the above equation is obtained as

$$x^*(p) = \frac{1}{a}\left(1 - \frac{\frac{g(p)}{f_P(p)}}{E^Q[\frac{g(p)}{f_P(p)}]}\right) - E[y(p,q)|p] + E^Q[E[y(p,q)|p]]\frac{\frac{g(p)}{f_P(p)}}{E^Q[\frac{g(p)}{f_P(p)}]} \tag{4.4}$$

Bivariate Longnormal-normal distribution for price and load

Suppose the marginal distributions of *p* and *q* as follows:

$$\begin{array}{ll} \text{Under } P: & \log p \sim N(m_1, s^2), \quad q \sim N(m, u^2), \quad Corr(\log p, q) = \rho \\ \text{Under } Q: & \log p \sim N(m_2, s^2) \end{array}$$

Then, we can get the explicit functions for the optimal payoff. For the **CARA utility**, the optimal payoff function reduces to

$$x^*(p) = \frac{1}{a}(A_1(p) + A_2(p)) \tag{4.5}$$

where

$$\begin{aligned} A_1(p) &\equiv \ln\tfrac{f_p(p)}{g(p)} - E^Q\big[\ln\tfrac{f_p(p)}{g(p)}\big] = -\tfrac{m_2-m_1}{s^2}(\log p - m_2) \\ A_2(p) &\equiv \ln E[e^{-ay(p,q)}|p] - E^Q[\ln E[e^{-ay(p,q)}|p]] \\ &= -ar\rho\tfrac{u}{s}(\log p - E^Q[\log p]) + a(m - \rho\tfrac{u}{s}m_1)(p - E^Q[p]) \\ &\quad + a\rho\tfrac{u}{s}(p\log p - E^Q[p\log p]) \\ &\quad + \tfrac{1}{2}a^2(-2r(p - E^Q[p]) + p^2 - E^Q[p^2])u^2(1-\rho^2) \\ &= -ar\rho\tfrac{u}{s}(\log p - m_2) + a(m - \rho\tfrac{u}{s}m_1)(p - e^{m_2+\frac{1}{2}s^2}) \\ &\quad + a\rho\tfrac{u}{s}\left(p\log p - (m_2+s^2)e^{m_2+\frac{1}{2}s^2}\right) \\ &\quad + \tfrac{1}{2}a^2(-2r(p - e^{m_2+\frac{1}{2}s^2}) + p^2 - e^{2m_2+2s^2})u^2(1-\rho^2) \end{aligned}$$

and for the **mean-variance utility**, the optimal payoff function reduces to

$$x^*(p) = \frac{1}{a}(1 - B_1(p)) - B_2(p) + B_3B_1(p) \tag{4.6}$$

where

$$
\begin{aligned}
B_1(p) &\equiv \frac{\frac{g(p)}{f_P(p)}}{E^Q\left[\frac{g(p)}{f_P(p)}\right]} = \exp\left(\tfrac{m_2-m_1}{s^2}\log p + \tfrac{m_1^2-m_2^2}{2s^2} - \tfrac{(m_1-m_2)^2}{s^2}\right) \\
&= e^{-\frac{(m_1-m_2)(m_1-3m_2)}{2s^2}} p^{\frac{m_2-m_1}{s^2}} \\
B_2(p) &\equiv E[y(p,q)|p] = E[(r-p)q|p] = (r-p)\big(m + \rho\tfrac{u}{s}(\log p - m_1)\big) \\
B_3 &\equiv E^Q[E[y(p,q)|p]] \\
&= (r - E^Q[p])(m - \rho\tfrac{u}{s}m_1) + \rho\tfrac{u}{s}(rE^Q[\log p] - E^Q[p\log p]) \\
&= (r - e^{m_2+\frac{1}{2}s^2})(m - \rho\tfrac{u}{s}m_1) + \rho\tfrac{u}{s}(rm_2 - (m_2+s^2)e^{m_2+\frac{1}{2}s^2})
\end{aligned}
$$

Bivariate Lognormal distribution for price and load

Suppose the marginal distributions of *p* and *q*, on the other hand, follow bivariate lognormal distributions as follows:

$$
\begin{aligned}
&\text{Under } P: \quad \log p \sim N(m_1, s^2), \quad \log q \sim N(m_q, u_q^2), \quad Corr(\log p, \log q) = \phi \\
&\text{Under } Q: \quad \log p \sim N(m_2, s^2)
\end{aligned}
$$

Then, we can get the explicit functions for the optimal payoff for the **mean variance utility**:

$$
x^*(p) = \frac{1}{a}(1 - B_1(p)) - B_2'(p) + B_3' B_1(p) \tag{4.7}
$$

where

$$
B_2'(p) \equiv E[y(p,q)|p] = E[(r-p)q|p] = (r-p)e^{m_q + \phi\frac{u_q}{s}(\log p - m_1) + \frac{1}{2}u_q^2(1-\phi^2)}
$$

since $\log q|p \sim N(m_q + \phi\frac{u_q}{s}(\log p - m_1), u_q^2(1-\phi^2))$, and

$$
\begin{aligned}
B_3' \equiv E^Q[E[y(p,q)|p]] &= re^{m_q + \phi\frac{u_q}{s}(m_2-m_1) + \frac{1}{2}u_q^2(1-\phi^2) + \frac{1}{2}\phi^2\frac{u_q^2}{s^2}s^2} \\
&\quad - e^{m_2 + m_q + \phi\frac{u_q}{s}(m_2-m_1) + \frac{1}{2}u_q^2(1-\phi^2) + \frac{1}{2}(\phi\frac{u_q}{s}+1)s^2}
\end{aligned}
$$

4.1.2. Replicating the Optimal Payoff Function

In this analysis, we replicate a portfolio that replicates payoff x(p). It can be written in the following form for any arbitrary positive value of s.

$$
x(p) = [x(s) - x'(s)s] + x'(s)p + \int_0^s x''(K)(K-p)^+ dK + \int_s^\infty x''(K)(p-K)^+ dK
$$

Let F be the forward price for a delivery at time 1. Evaluating the equation at S=F and rearranging it gives

$$x(p) = x(F) \cdot 1 + x'(F)(p - F) + \int_0^F x''(K)(K - p)^+ dK + \int_F^\infty x''(K)(p - K)^+ dK.$$

Note that 1, $(p-F)$, $(K-p)^+$ and $(p-K)^+$ in the above expression represent payoffs at time 1 of a bond, forward contract, put option, and call option, respectively.

Therefore,
$x(F)$ units of bonds,
$x'(F)$ units of forward contracts,
$x''(K)dK$ units of put options with strike K for every $K < F$, and
$x''(K)dK$ units of call options with strike K for every $K > F$
gives the same payoff as $x(p)$.

The above implies that unless the optimal payoff function is linear, the optimal strategy involves purchasing (or selling short) a spectrum of both call and put options with continuum of strike prices. This result proves that LSEs should purchase a portfolio of options to hedge price and quantity risk together. Even if prices go up with increasing loads, more call options with higher strike prices are exercised, having an effect of putting price caps on each incremental load.

4.1.3. An Example

In this section, we illustrate the method that we derived in the previous sections. We consider the on-peak hours of a single summer day as time 1. Parameters were approximately based on the California Power Exchange data of daily day-ahead average on-peak prices and 1% of the total daily on-peak loads from July to September, 1999. Specific parameter values are imposed as follows:

- Price is distributed lognormally with parameters m1 = 3.64 and s = 0.35 in both the real world and risk-neutral world: log p ~ N (3.64, 0.35^2) in P and Q. The expected value of the price p under this distribution is \$40.5/MWh.
- The fixed rate r = \$100/M W h is charged to the customers who are served by the LSE.
- For CARA utility, the risk aversion is a = 1.5.
- Load is either normally distributed with mean m = 300 and $u^2 = 30^2$, or lognormally distributed with parameter m = 5.77 and u = 0.09.

We would like to point out a significant correlation-effect in the profit distributions. Figure 4-1 shows that the profit distributions become quite different as the correlation between load and logarithm of price changes. Considering that the correlation coefficient of our data is 0.7, we observe that the correlation coefficient cannot be ignored in the analysis of profit. The optimal payoff functions for a CARA utility LSE are drawn in Figure 4-2 for various correlation coefficients between log p and q. Generally, low profit from high loads for very high spot prices and from low load for very low spot price is compensated with the cases where spot prices and loads are around the expected value. This can be seen from the graph where as the spot price

goes away from r, positive payoff is received from the optimal portfolio while the payoff is negative around r. We also note that larger payoff can be received when the correlation is smaller. This is because the variance of profit is bigger when the correlation is smaller as we can see from Figure 4-1. Therefore, even when the correlation is zero, the optimal payoff function is nonlinear. Figure 4-3 illustrates the numbers of contracts to be purchased in order to obtain the payoff x*(p) for an LSE with a CARA utility function. We see that the numbers of option contracts are very high relative to the mean volume. This is because we don't restrict the model with constraints such as credit limits. The zero-cost constraint (4.2) that we only included in our model enables borrowing as much money as needed to finance any number of derivative contracts.

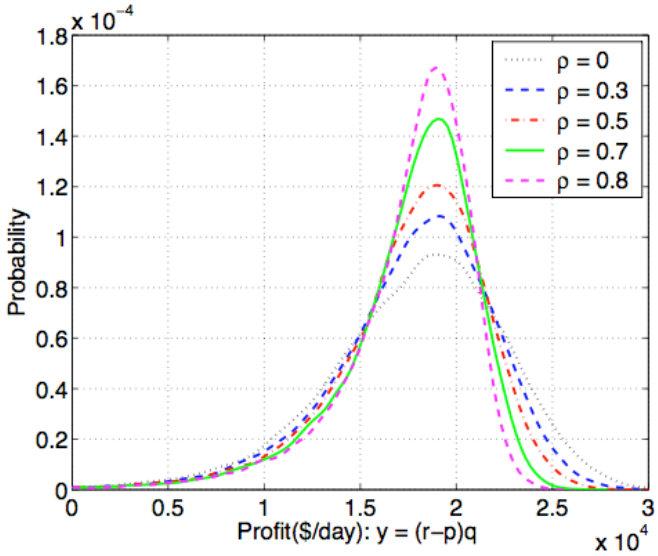


Figure 4-1: Profit distribution for various correlation coefficients.

Generated 50000 pairs of (p,q) from a bivariate normal distribution of (logp, q) with a various correlation ρ's, where logp ~N(3.64, 0.35^2) and q~N(300, 30^2), and plotted estimated probability density functions of the profit using normal kernel (assuming r=$100/MWh).

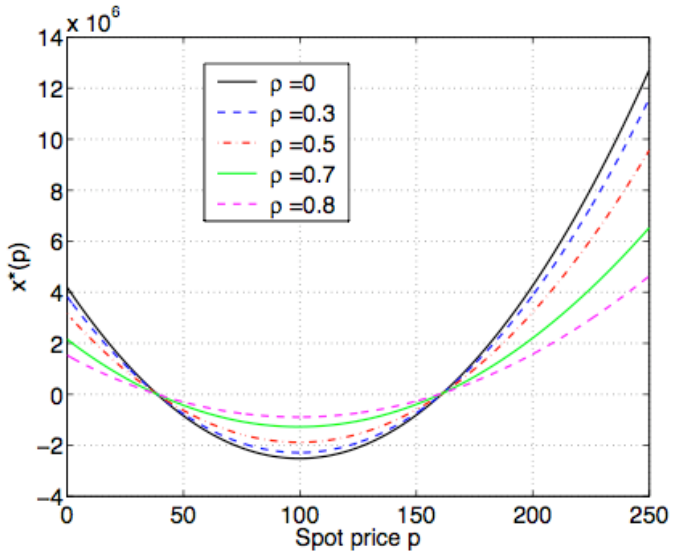
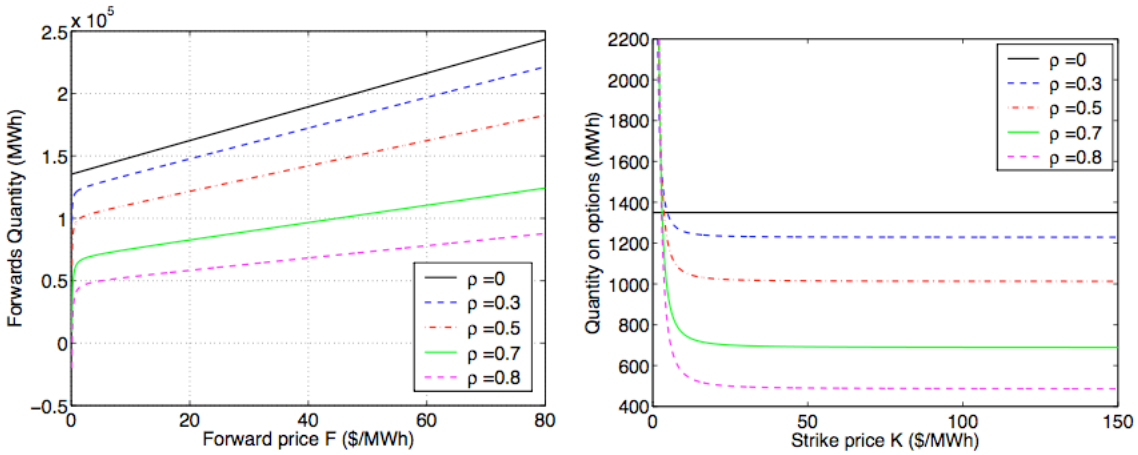


Figure 4-2: The optimal payoff function for an LSE with CARA utility

Price and load follow bivariate lognormal-normal distribution logp~N(3.64, 0.35^2), and q~N(300, 30^2) with correlation coefficient ρ



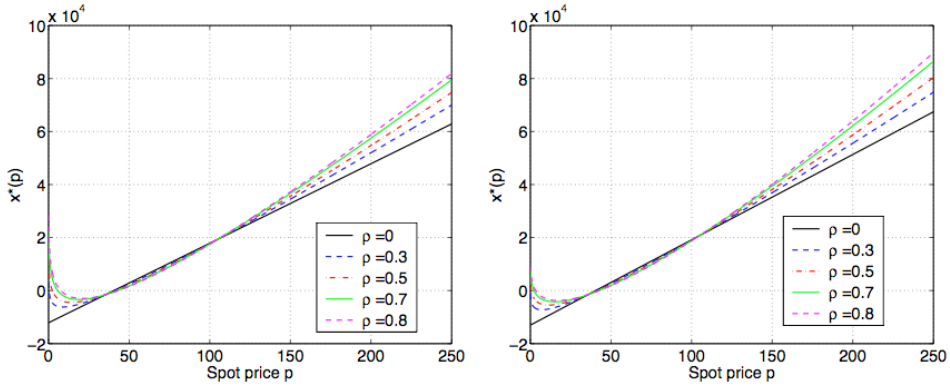
(a) $x'(F)$ (b) $x''(K)$

Figure 4-3: Optimal numbers of forward and options contracts for the LSE with CARA utility

The graphs show numbers of forward and option contracts to be purchased in order to replicate the optimal payoff x*(p) that is obtained for the LSE with CARA utility. In this example, the forward price is $40.5/MWh, thus, the optimal portfolio includes forward contracts for x′(40.5)MWh, put options on x″(K)*dK MWh for K<40.5 and call options on x″(K)*dK MWh for K>40.5

For an LSE with mean-variance utility, the optimal payoff functions are drawn in Figure 4-4.

They show the tendency of mean-variance utility to protect against high price and low quantity. For an illustration of the numbers of contracts to be purchased in order to obtain payoff x*(p), see Figure 4-5. Note that in our examples the number of option contracts to be purchased in the optimal portfolio is positive for any strike prices. This implies that we borrow money from the bank and purchase a portfolio of option contracts.



(a) Bivariate lognormal-normal (p, q) (b) Bivariate lognormal (p, q)

Figure 4-4: Optimal payoff functions for an LSE with mean-variance utility.

(a) corresponds to (logp,q) ~N(3.64,300, 0.35^2, 30^2,ρ), and (b) corresponds to (logp,logq) ~N(3.64,5.77, 0.35^2, 0.09^2,ρ).

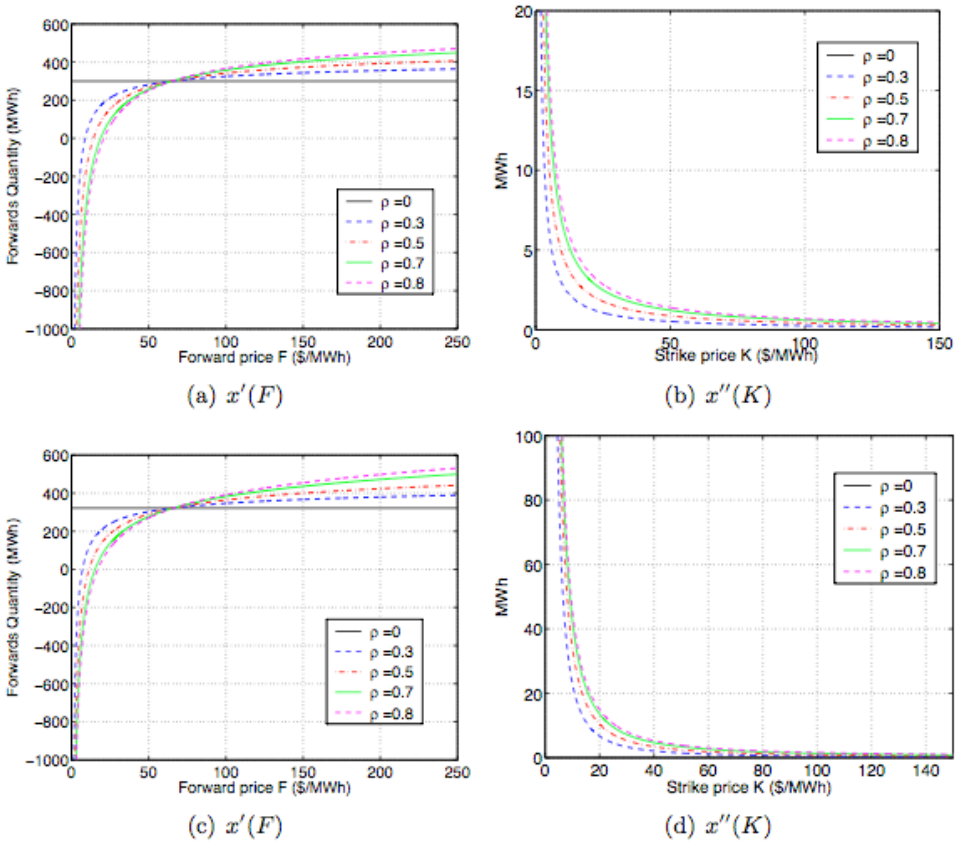


Figure 4-5: Optimal numbers of forward and options contracts for the LSE with mean-variance utility

The graphs show numbers of forward and option contracts to be purchased in order to replicate the optimal payoff x*(p) that is obtained for the LSE with mean-variance utility. In this example, the forward price is $40.5/MWh, hence, the optimal portfolio includes the forward contract for x′(40.5)MWh, put options on x′′(K)*dK MWh for K<40.5 and call options on x′′(K)*dK MWh for K>40.5. The upper panels (a) and (b) correspond to price and load following a bivariate lognormal-normal distribution, and the lower panels correspond to price and load following a bivariate lognormal distribution.

Figure 4-6 compares distribution changes between profit without hedging, profit after price hedge and profit after the optimal price and quantity hedge. The graph shows significant improvements in reducing risks when we hedge price and quantity risk together.

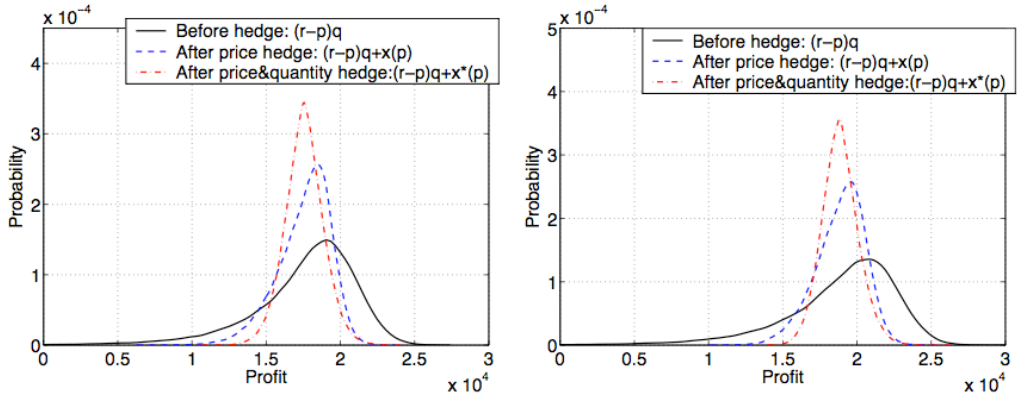


Figure 4-6: The comparison of profit distribution for an LSE with mean-variance utility

The comparison of profit distribution for an LSE with mean-variance utility for three cases: before hedge, after price hedge, and after price and quantity hedge, assuming the correlation coefficient between price and load to be 0.7.

In Figure 4-7 we explore the sensitivity of the optimal payoff function with respect to the divergence between the risk-neutral distribution and the assumed physical distribution of prices. Specifically we assume that the joint distribution for quantity and price under both measures P and Q are represented by a bivariate lognormal-normal density function with possible differences in the mean logarithmic price, which we vary. The results depend on the utility function used. For CARA utility, the overall payoff is higher than the optimal payoff if the expected price is higher than the market price, but the difference is not that significant with respect to the payoff changes for differing p. For the mean-variance case, however, the difference between payoffs for varying mean logarithmic price m2 is more noticeable.

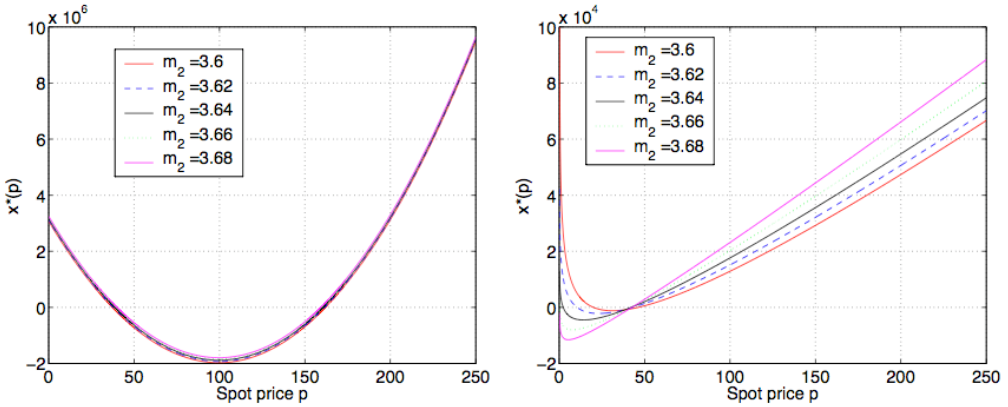
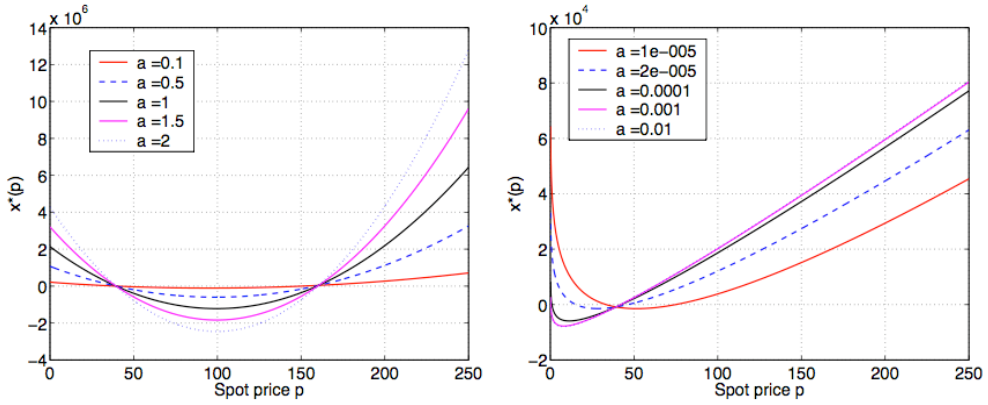


Figure 4-7: Sensitivity of the optimal payoff function

The graphs correspond to the case when price and load follow a bivariate lognormal-normal distribution with correlation coefficient 0.5. m2 represents the mean of logarithm of price under the risk-neutral probability measure with m2=3.64 corresponding to the case $P \equiv Q$.

Figure 4-8 shows how hedging strategies change with risk aversion. Figure 4-8(a) displays the optimal payoff functions for CARA utility with different levels of risk aversion. It shows the payoff function with high risk aversion is more sensitive to the unit change in spot price, indicating that a more risk averse LSE will enter into more active hedging. On the other hand, mean-variance utility shows different aspects. In Figure 4-8(b), as *a* gets close to 0.01, the optimal payoff doesn't change much; the mean-variance objective function gives more weight to variance as *a* gets bigger, so *a* won't affect the optimal payoff function above a certain level and the objective turns into minimizing variance. However, for smaller risk aversion, the mean-variance objective function puts more weights on the mean of profit; LSEs with low risk aversion will protect more against the lower spot price worrying that the expected profit is low from decreased load when spot price is low.



(a) Optimal payoff with CARA utility (b) Optimal payoff with mean-variance utility

Figure 4-8: Optimal payoffs with different risk aversion

Optimal payoffs for the case when price and load follow a bivariate lognormal-normal distributions: N(3.64,300, 0.35^2 , 30^2 ,0.5) under P while the log-price distribution is N(3.66,0.352) under Q.

4.1.4. Potential Use of Developed Tools

In order to show the practical use of the model discussed in the previous section, we have developed a graphic User Interface for industry members to investigate the hedging performance of the optimal portfolios suggested by our model. Our intention is that, with real market data inputted and utility functions specified by the industry users, the interface could provide the corresponding payoff functions, the positions of forward contracts and options, and the performance of hedging the price and volumetric risks.

To start with, we developed a simpler version to reproduce the results of the example in the last section. The screen shot is showed in Figure 4-9. Users need to provide the following parameters as inputs and specify their desired plots as outputs.

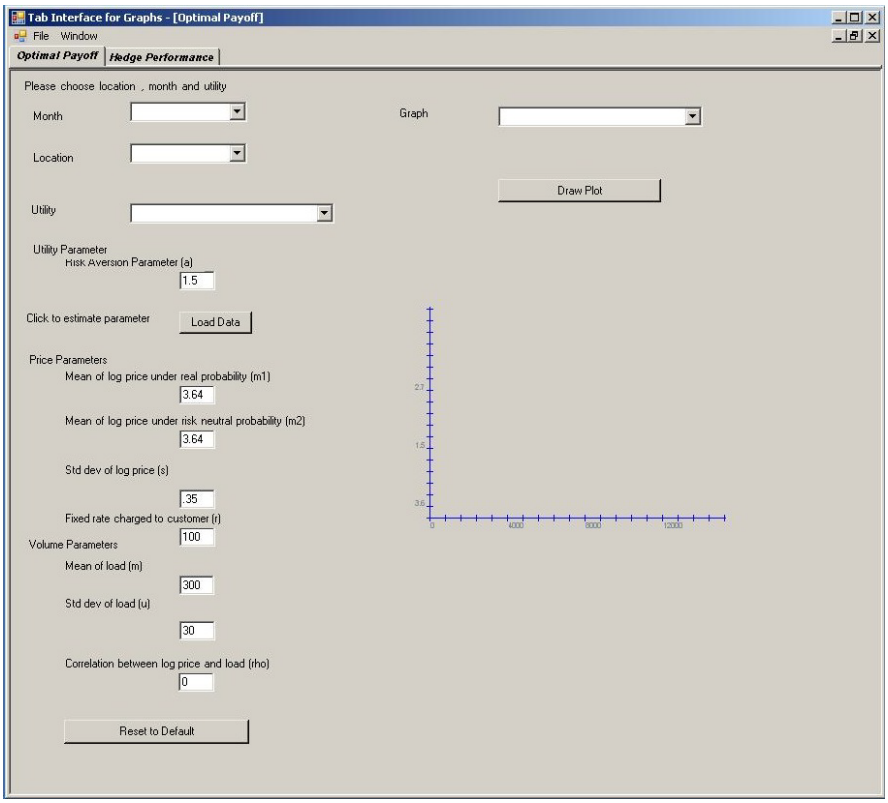


Figure 4-9: Screen Shot for the Basic User Interface

INPUTS:

- The utility function: CARA, Mean-Variance Bivariate Lognormal-normal and Mean-Variance Bivariate Lognormal.
- The risk aversion parameter, “a”.
- Fixed rate charged to customers, “r”.
- Price parameters: mean of log price under real probability “m1” and under risk neutral probability “m2”, standard deviation of log price “s”.
- Volume parameters: Mean of load “m” and standard deviation of load “u”.
- The correlation between log price and load.

OUTPUTS:

- Optimal Payoff Function
- Number of the Forward Contracts in optimal hedge portfolio
- Number of the Option Contracts in optimal hedge portfolio
- Profit distribution before hedge
- Profit distribution after just the price hedge
- Profit distribution after both price and quantity hedge

After providing the required information, users need to click “Draw Plot” to see the plots in the right bottom panel.

For illustration, we implement the model by using the same sets of parameters in the previous example.

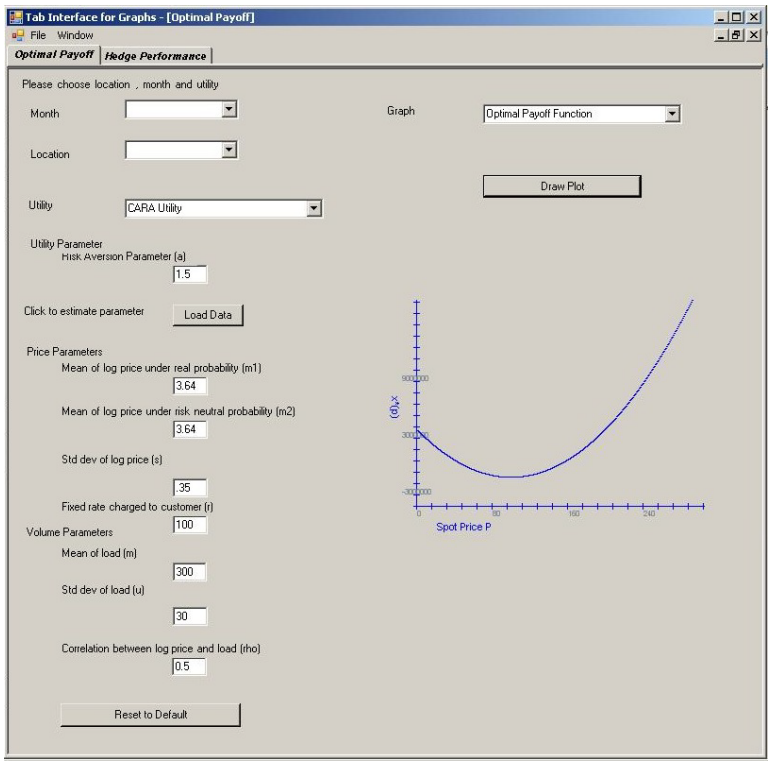


Figure 4-10: Screen Shot: Optimal Payoff Function under CARA Utility

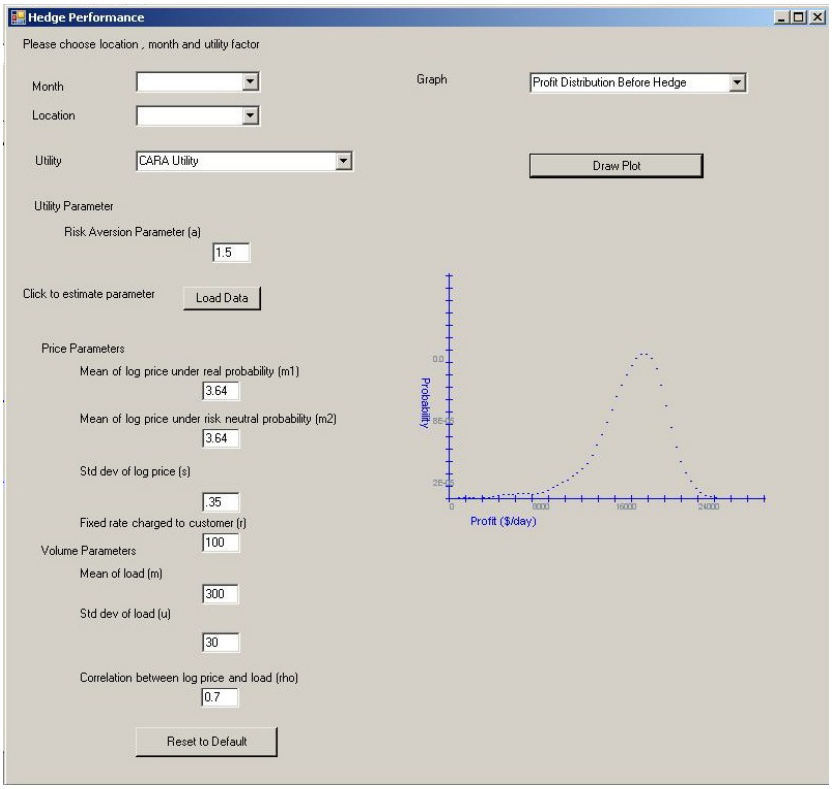


Figure 4-11: Screen Shot: Profit Distribution under CARA Utility

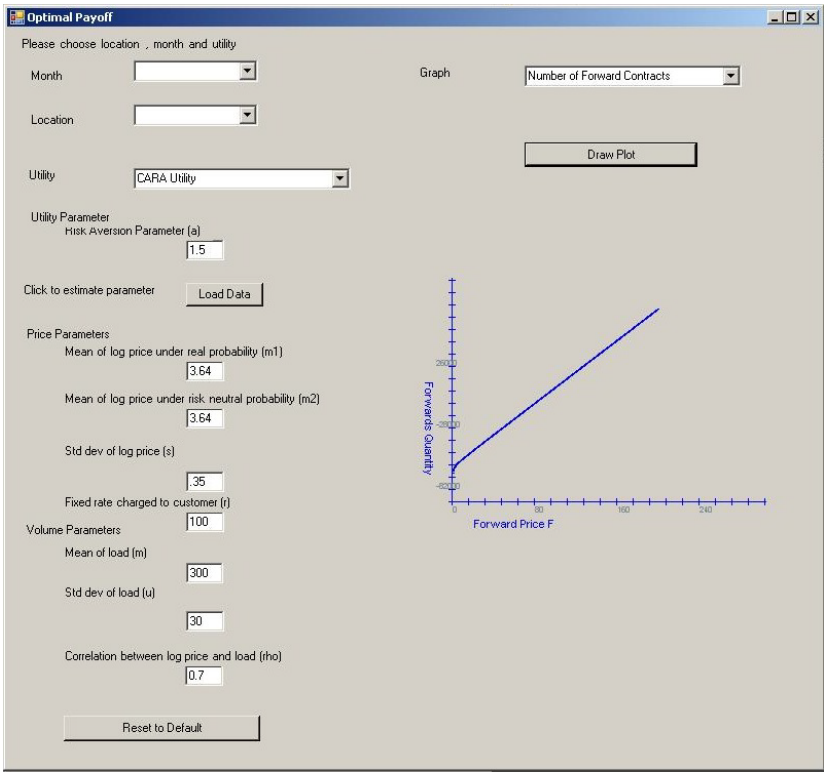


Figure 4-12: Screen Shot: Number of Forwards under CARA Utility

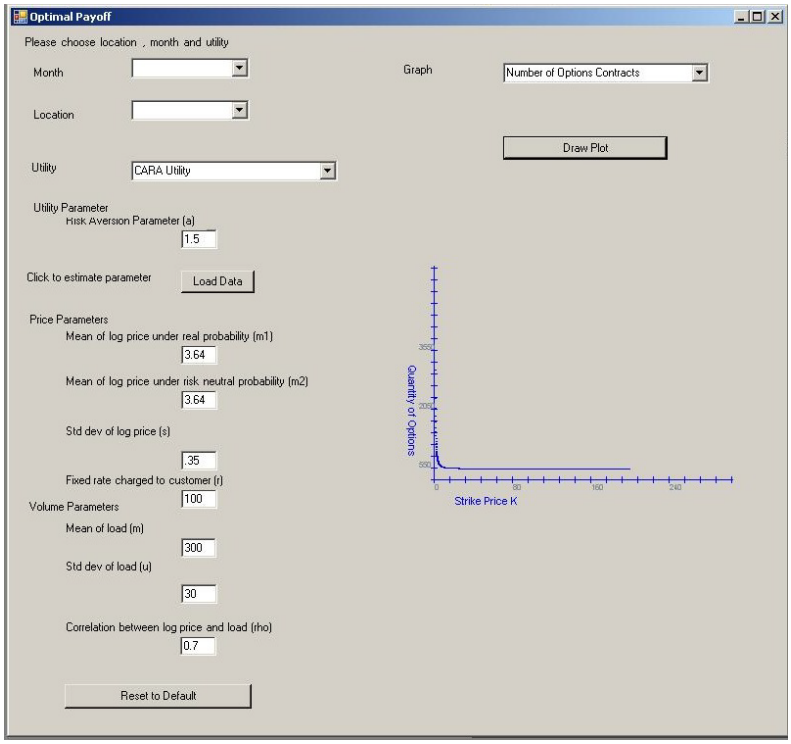


Figure 4-13: Screen Shot: Number of Options under CARA Utility

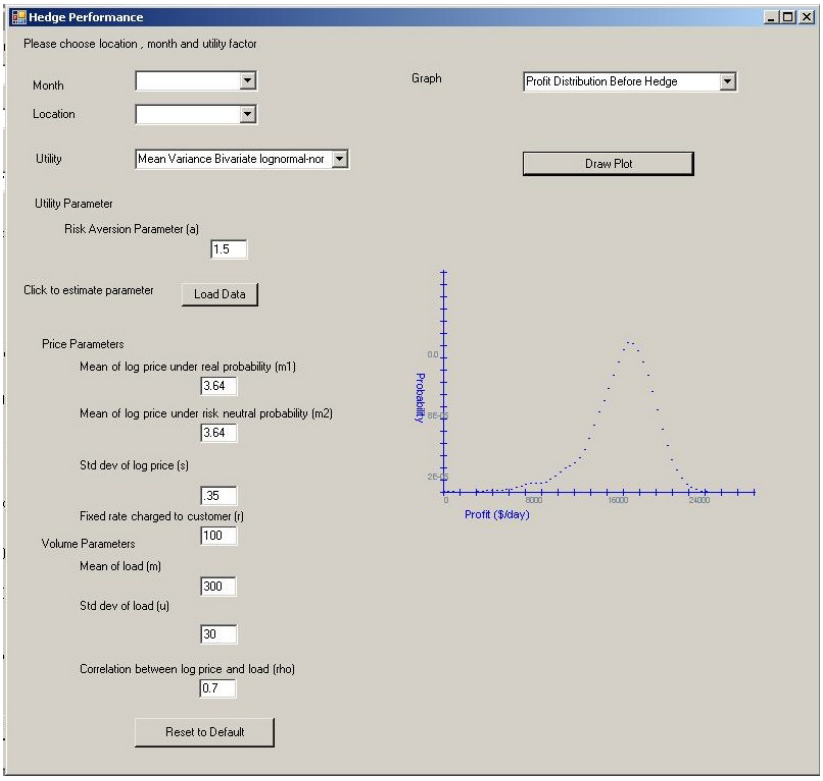


Figure 4-14: Screen Shot: Profit Distribution Before Hedging under Bivariate Lognormal-normal Utility

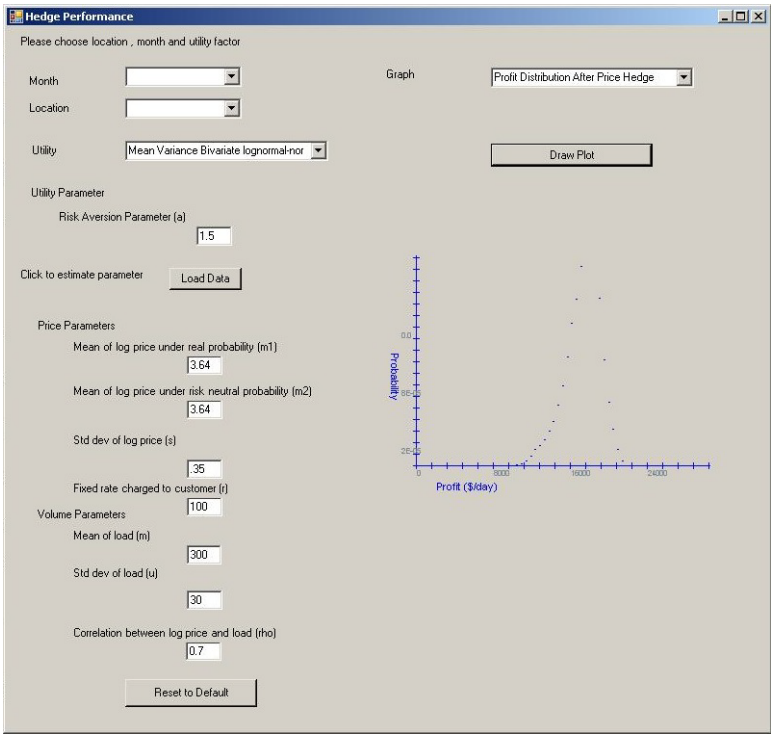


Figure 4-15: Screen Shot: Profit Distribution After Price Hedging under Bivariate Lognormal-normal Utility

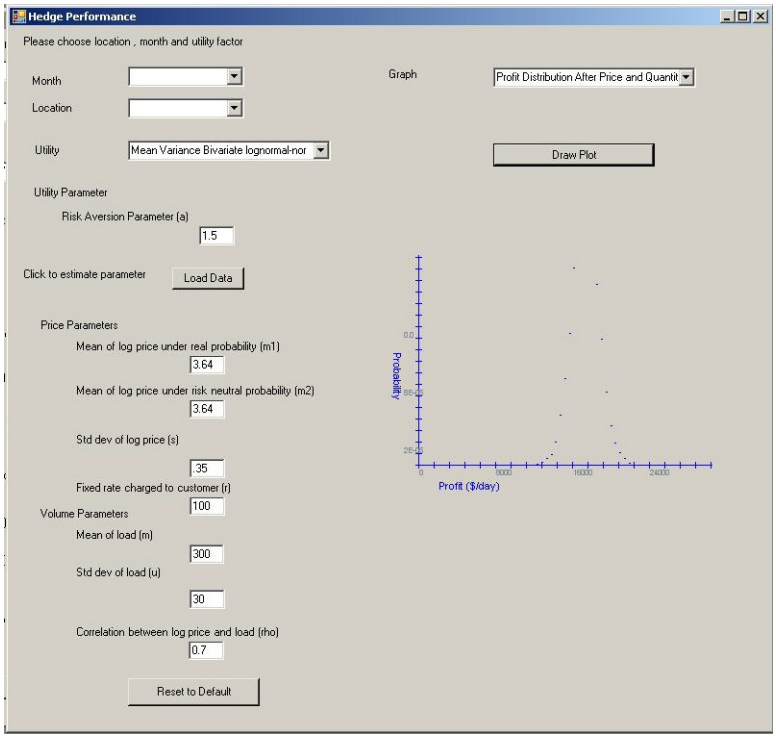


Figure 4-16: Screen Shot: Profit Distribution After Price and Quantity Hedging under Bivariate Lognormal-normal Utility

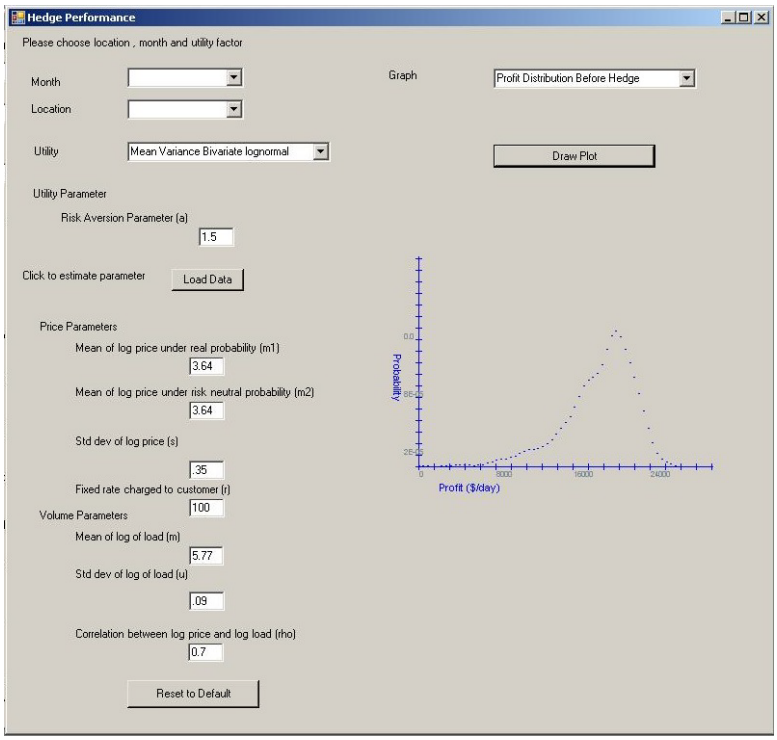


Figure 4-17: Screen Shot: Profit Distribution Before Hedging under Bivariate Normal Utility

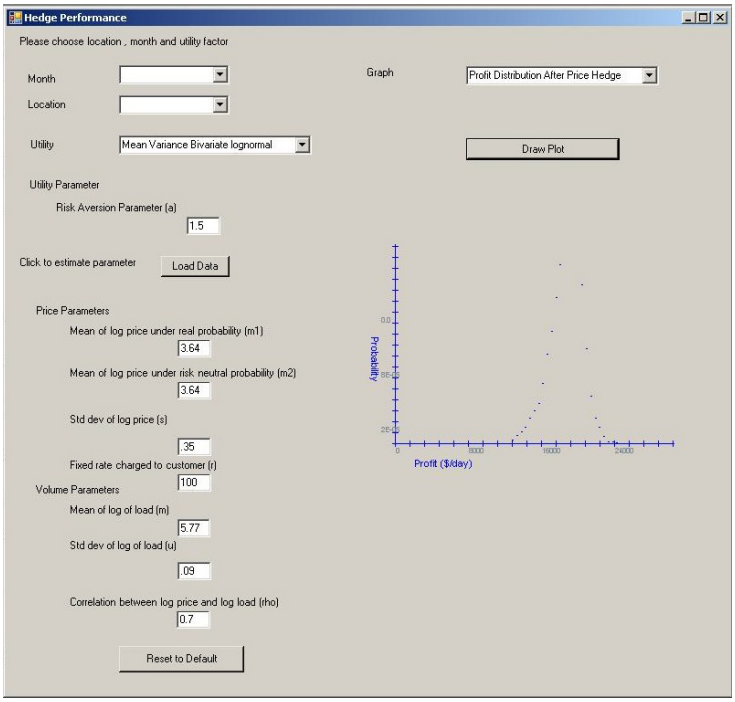


Figure 4-18: Screen Shot: Profit Distribution After Price Hedging under Bivariate Normal Utility

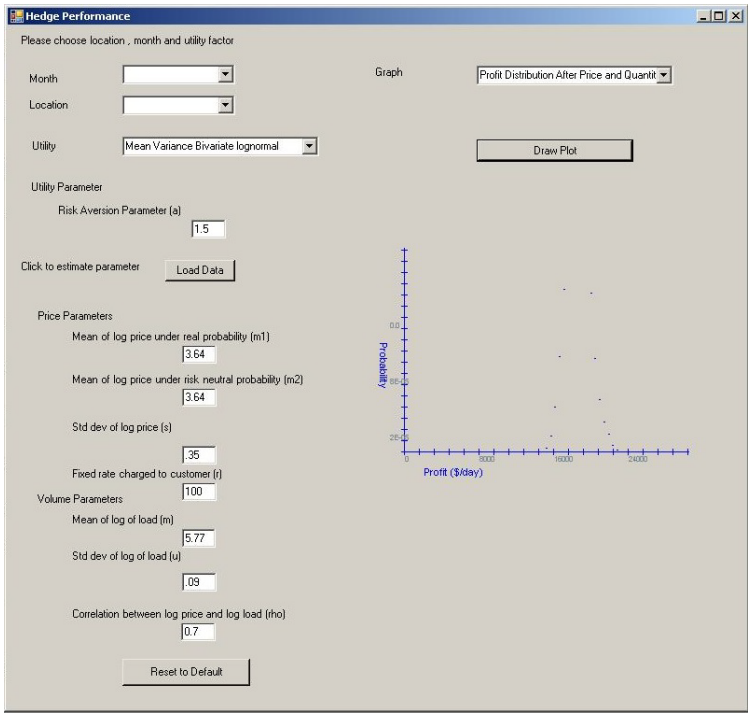


Figure 4-19: Screen Shot: Profit Distribution After Price and Quantity Hedging under Bivariate Normal Utility

The interpretation of these plots is given in Section 4.1.3. Now we improve our interface by using the parameters for price and load based on historical data rather than an editable text box for the user to fill in. The interface provides for the user to select the hour of the day and a selection of regions for which we can provide the parameters based on historical data. This is expandable to other regions (if desired by the user). One simply needs to add data into a new sheet in the workbook already provided and simply add the name of the sheet in the list of variables for the list box in its properties. Click on the 'Load data' button provides these parameters based on the historical data from the selected region and for the given hour selected. The screen shot of the extended interface is shown in Figure 4-20.

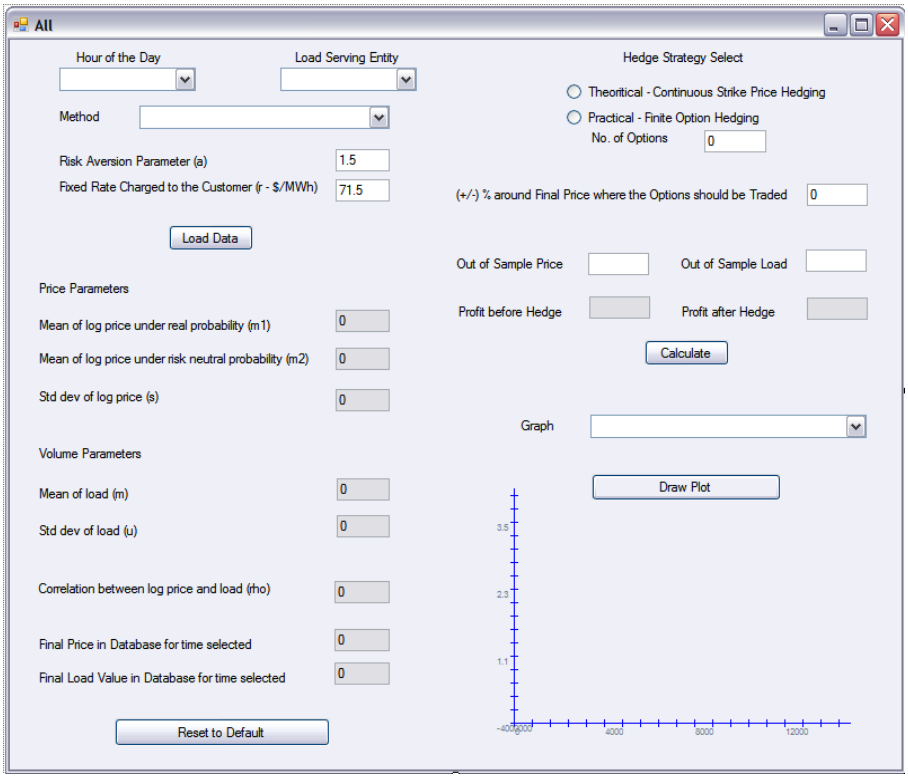


Figure 4-20: Screen Shot of the Extended Interface

Users have more options for inputs in this interface.

INPUTS:

- Hour of the day and Load Serving Region (specify these two information to locate the desired data set saved in the .xls file).
- Utility function.
- The risk aversion parameter, “a”.
- Fixed rate charged to customers, “r”.

Then you click “Load Data” to get the estimated parameters based on the chosen market data. We also provide for an option to select the theoretical optimal hedging or a practical hedging version (practical hedging is left for future development). For the practical hedging part the user can enter the number of call options as desired and the range in which he/she would like these options to have their strike prices. This simple feature allows the user to decide the strike prices approximately based on the liquidity is the OTM/ITM option market. The out of sample price/load point can be provided by the user or left to us, where we use the last point from the historical set as out of sample (we do not use the last point in parameter calculation).

The “Calculate” button will give the profit before and after hedge for this out of sample point based on either Theoretical/Practical method as selected. Options for outputs and the draw plot feature is same as the earlier interface.

For illustration, we have downloaded the data for our analysis from www.nyiso.com. Data has been collected for four LSE's: Long Island, NYC, North NY and West NY. Day Ahead Market and Load data have been downloaded for the period of Summer 2008 (June, July and August) on an hourly basis. Seasonality is a major issue in electricity prices and thus a particular period of the year is chosen.

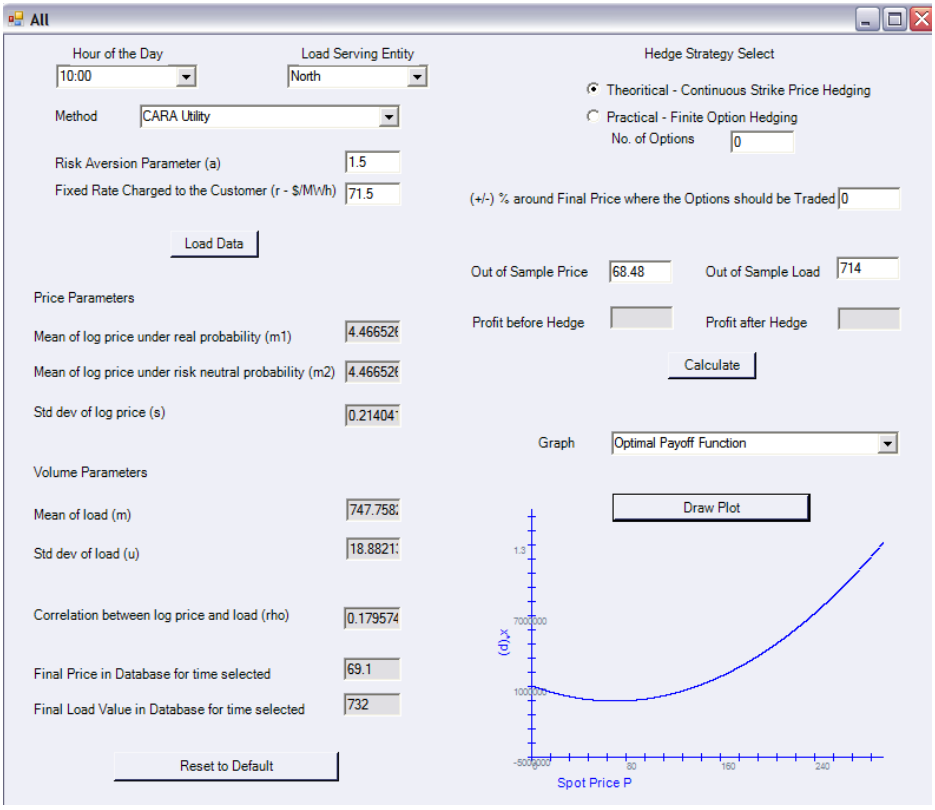


Figure 4-21: Screen Shot: Optimal Payoff Function under CARA Utility for 10 AM North NY

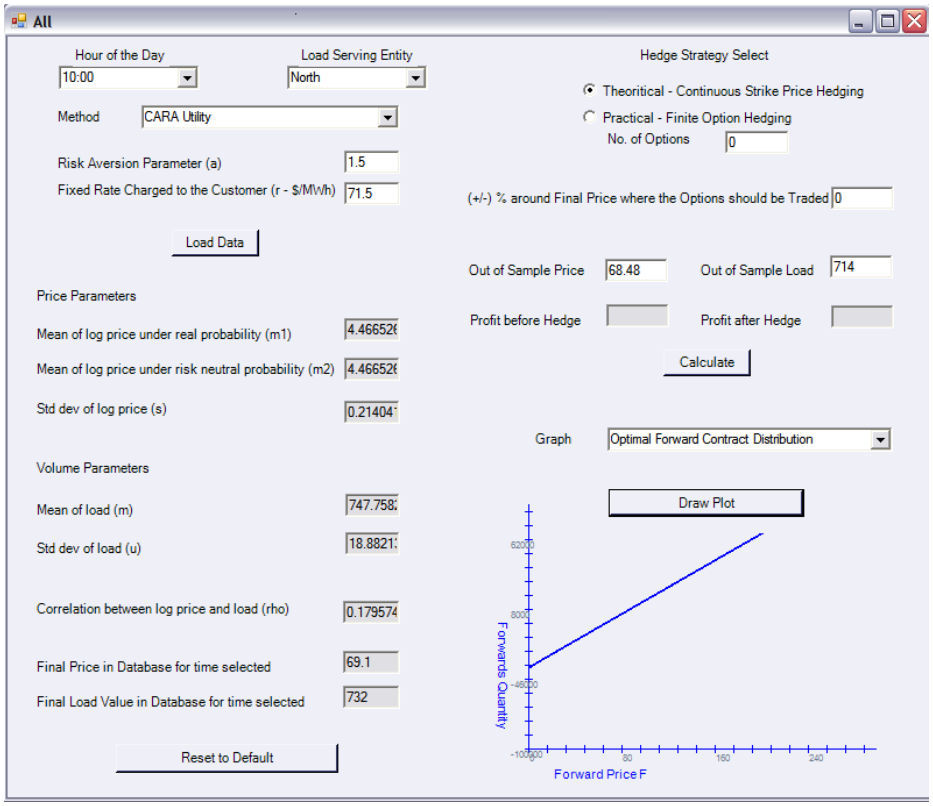


Figure 4-22: Screen Shot: Optimal Forward Contracts under CARA Utility for 10 AM North NY

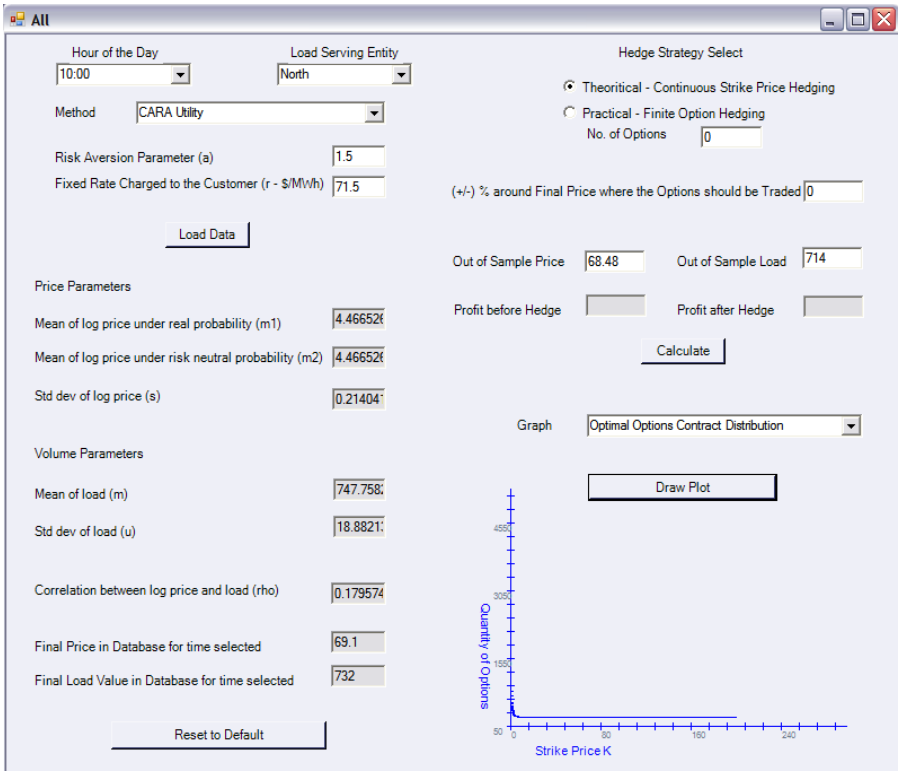


Figure 4-23: Screen Shot: Optimal Options under CARA Utility for 10 AM North NY

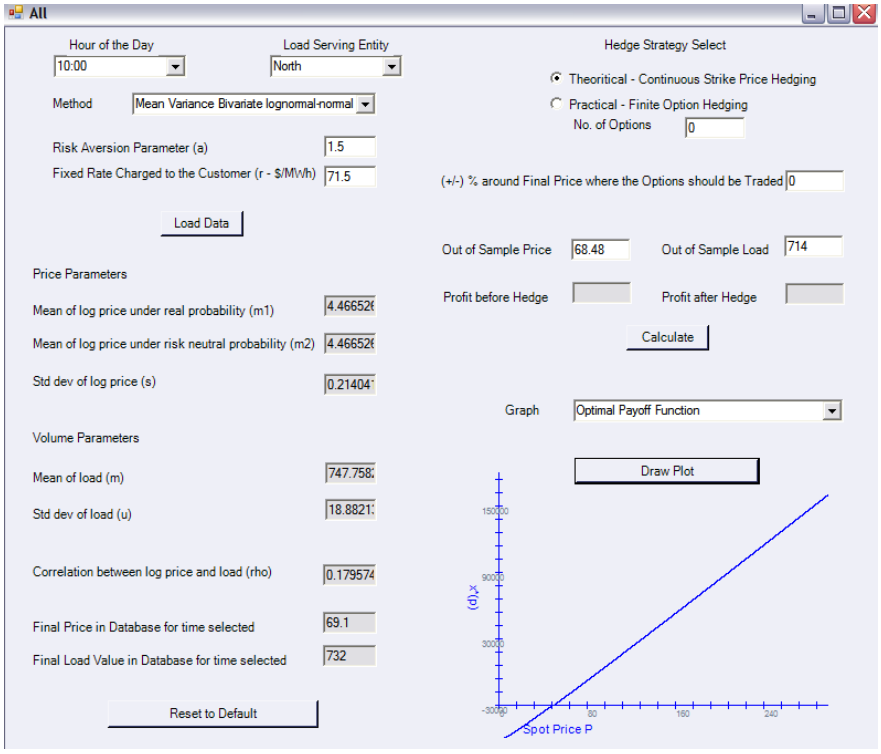


Figure 4-24: Screen Shot: Optimal Payoff Function under Bivariate Lognormal Normal Utility for 10 AM North NY

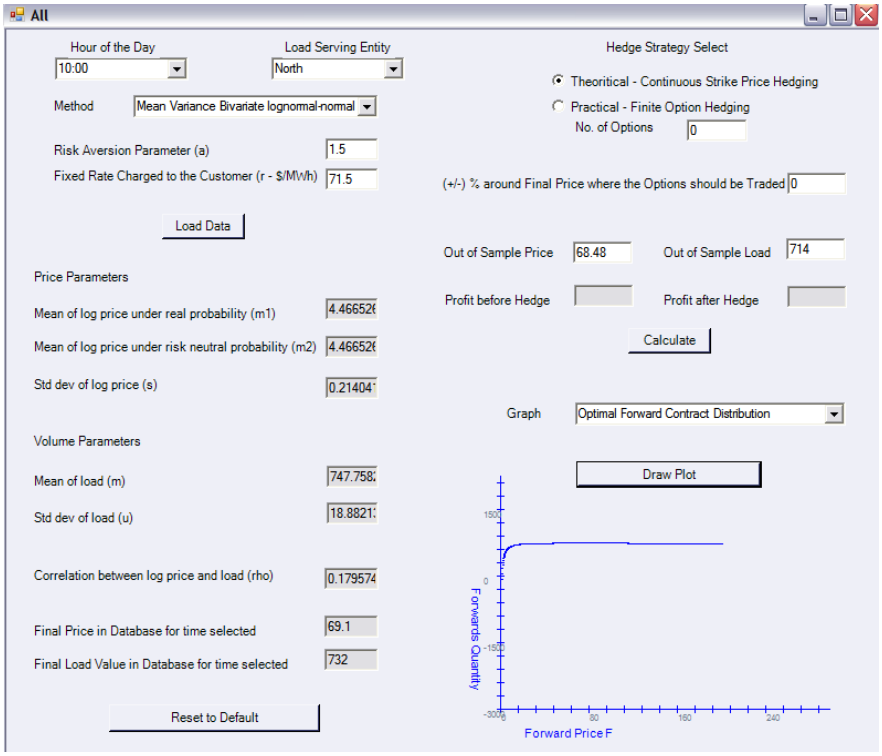


Figure 4-25: Screen Shot: Optimal Forward Contracts under Bivariate Lognormal Normal Utility for 10 AM North NY

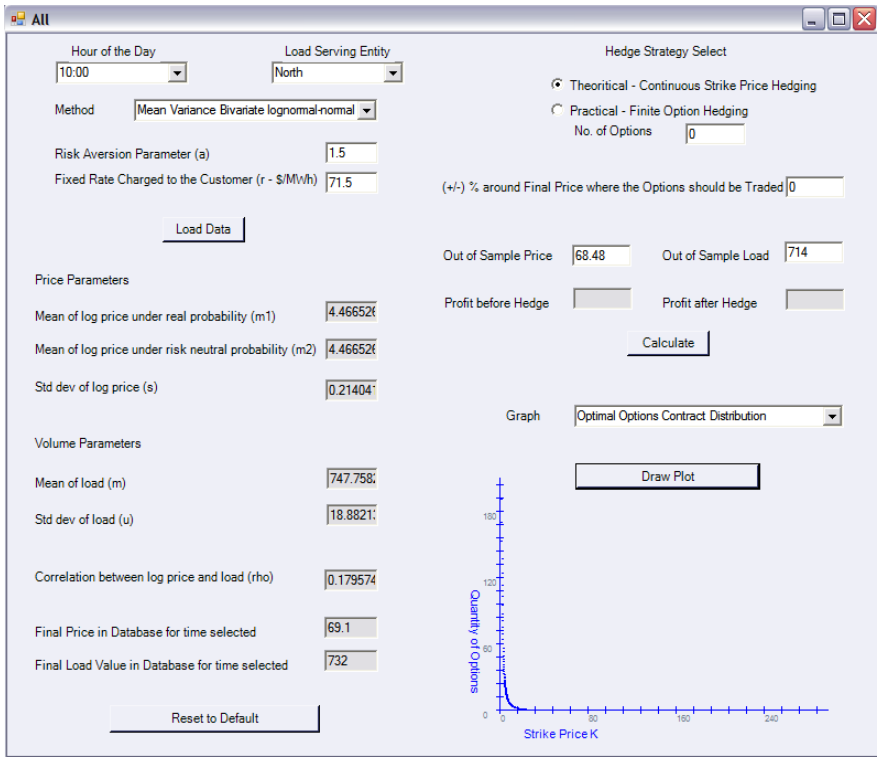


Figure 4-26: Screen Shot: Optimal Options under Bivariate Lognormal Normal Utility for 10 AM North NY

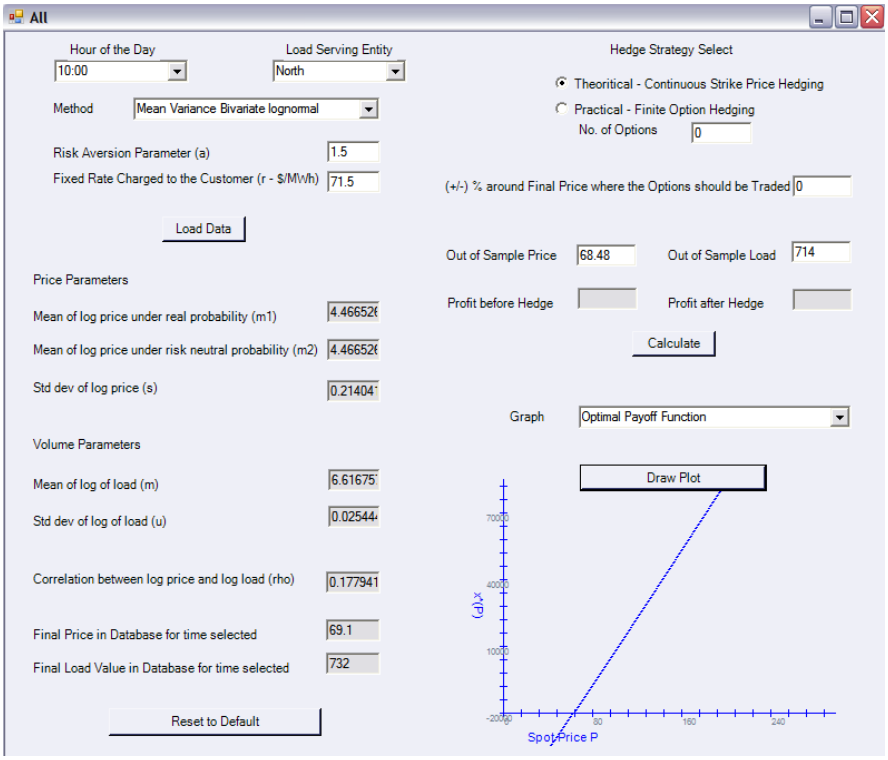


Figure 4-27: Screen Shot: Optimal Payoff Function under Bivariate Lognormal Utility for 10 AM North NY

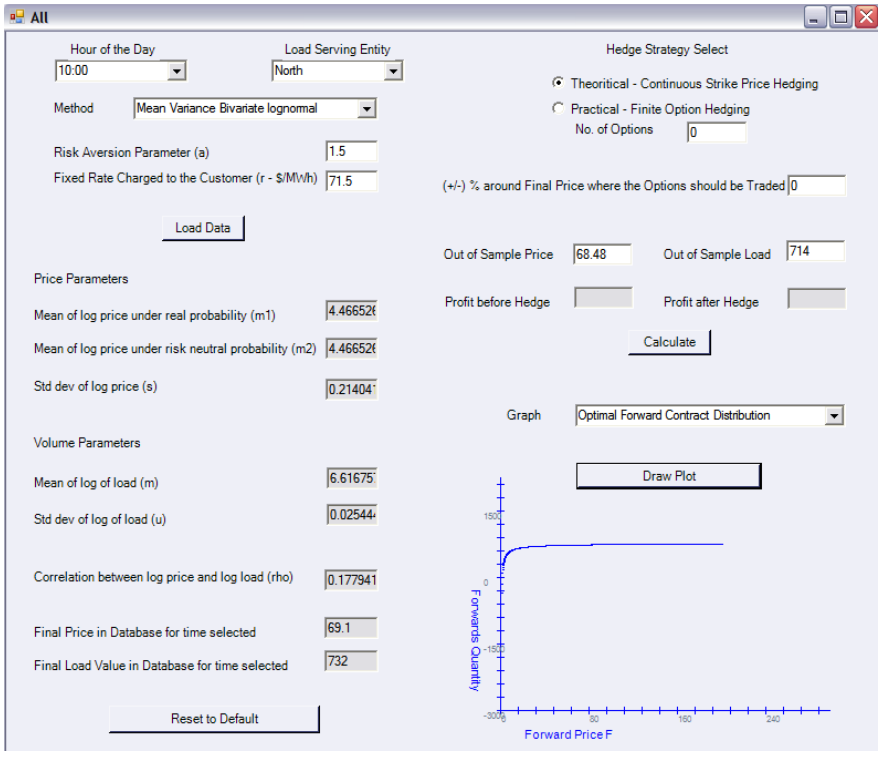


Figure 4-28: Screen Shot: Optimal Forward Contracts under Bivariate Lognormal Utility for 10 AM North NY

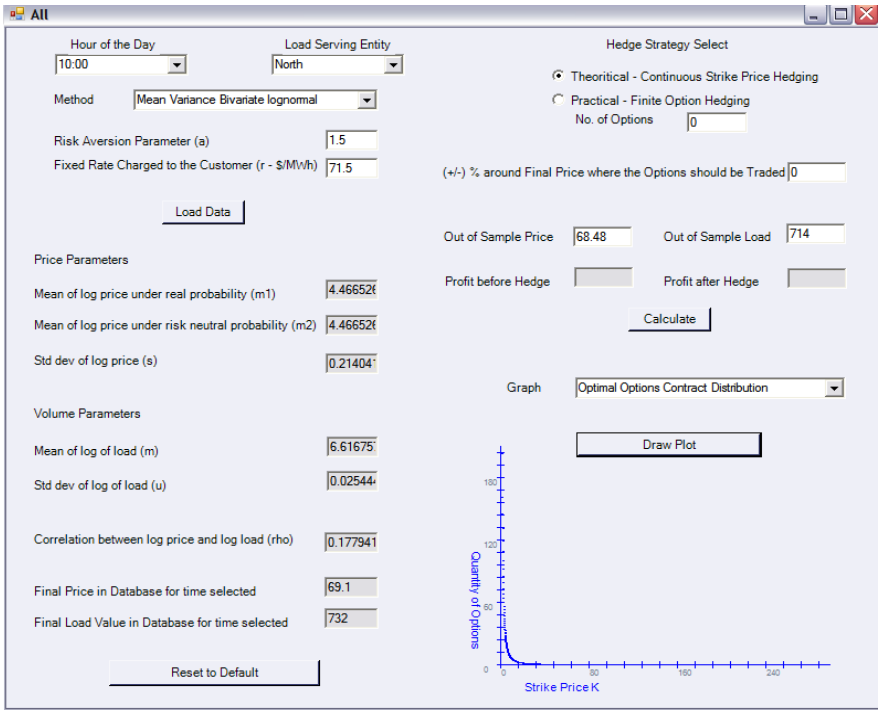


Figure 4-29: Screen Shot: Optimal Options under Bivariate Lognormal Utility for 10 AM North NY

4.2. Timing of a Static Hedge in a Continuous-time Setting

Let T be the delivery period and maturing date of the hedging instruments. We will assume that all the hedging instruments for the delivery period T are to be contracted at the same time τ. Contracting earlier reduces the risk by locking in the price of the contracts; while delaying the contracting enables more profitable hedging by exploiting more information that becomes available as we approach maturity. We will assume that the optimal hedging time is determined at time t_0 based on the information available at that time but the composition of the hedging portfolio is determined at hedging time based on the realized spot price. It should be noted that the optimal hedging time determined at t_0 might be no longer optimal at time $t_1 > t_0$ because more information on time T becomes available at time t_1, however we will assume that the contracting time is chosen irreversibly at time t_0.

4.2.1. Mathematical Formulation

Let $\{p_t\}_{t\in[0,T)}$ be a process of forward price for delivery at time T and $\{q_t\}_{t\in[0,T)}$ be a process for load estimate for period T calculated at time t. Assume the forward price and load estimate processes evolve as the following Ito processes:

$$dp_t = p_t\left(\mu_p(t)dt + \sigma_p(t)dB_t^1\right) \tag{4.8}$$

$$dq_t = q_t\left(\mu_q(t)dt + \sigma_{pq}(t)dB_t^1 + \sigma_q(t)dB_t^2\right) \tag{4.9}$$

where B_t^1 and B_t^2 are independent Wiener processes. Then, p_T and q_T denote the spot price and demand at time T.

The solution of the following problem is the best hedging timing determined at time t_0 for a mean-variance optimizer:

$$\begin{aligned} &\max_{\tau \geq t_0} \; E_{t_0}[U((r - p_T)q_T + x_\tau(p_T))] \\ &\text{s.t.} \quad x_\tau(p_T) = \arg\max_{x(p_T)} E_\tau[U((r - p_T)q_T + x(p_T))] \text{ s.t. } E_\tau^Q[x(p_T)] = 0 \end{aligned} \tag{4.10}$$

$x_\tau(p_T)$ denotes the payoff from the optimal portfolio to be constructed at time τ. Thus, the formulation finds a time τ^*, hedging at that time maximizes the expected utility of the optimally hedged profit.

Throughout the section, it is assumed that the physical probability measure and risk-neutral probability measure are the same. It follows that from the zero-cost constraint, maximizing a mean-variance objective function is reduced to minimizing the variance of the hedged profit. The formula for $x_\tau(p_T)$ can be obtained from the results of the Section 4.1.1. Thus, the problem becomes a single-variable unconstrained optimization problem that can be easily solved numerically.

4.2.2. Finding the Optimal Payoff Function at Contracting Time

Proposition: Suppose $\{p_t\}_{t\in[0,T)}$ and $\{q_t\}_{t\in[0,T)}$ follow Ito processes given (4.8) and (4.9). Assuming P = Q, then $x_\tau(p_T)$ that solves (4.10) for a mean-variance utility function is obtained as follows:

$$x_\tau^*(p_T) = B_\tau(p_T - r)p_T^{A_\tau}p_\tau^{-A_\tau}q_\tau + rC_\tau q_\tau - D_\tau p_\tau q_\tau \tag{4.11}$$

where

$$\begin{aligned} A_\tau &= \frac{\int_\tau^T b_t d_t dt}{\int_\tau^T b_t^2 dt} \\ B_\tau &= \exp\left(\int_\tau^T c_t dt - A_\tau \int_\tau^T a_t dt + \tfrac{1}{2}\int_\tau^T (d_t^2 + e_t^2)dt - \tfrac{1}{2}\frac{(\int_\tau^T b_t d_t dt)^2}{\int_\tau^T b_t^2 dt}\right) \\ C_\tau &= e^{\int_\tau^T (c_t + \frac{1}{2}d_t^2 + \frac{1}{2}e_t^2)dt} \\ D_\tau &= e^{\int_\tau^T (a_t + c_t + \frac{1}{2}b_t^2 + \frac{1}{2}d_t^2 + \frac{1}{2}e_t^2 + b_t d_t)dt} \end{aligned}$$

Equation (4.11) is the payoff of the optimal portfolio to be constructed when hedging at time τ. One can see that the optimal portfolio incorporates the information of the forward price and load estimate available at the hedging time τ.

4.2.3. Determining the Optimal Hedging Time

With the assumption P = Q and the zero-cost constraint $E_\tau [x_\tau] = 0$, maximizing (4.9) reduces to minimizing

$$\Pi(\tau) \equiv Var((r - p_T)q_T + x_\tau^*(p_T)).$$

Given x_τ^* obtained in Section 4.2.2, the problem (4.9) is in fact an unconstrained optimization problem with a single decision variable in the interval [0, T]. Once Π is obtained as a function of τ, the problem is solvable numerically even though Π(τ) is neither convex or concave. This section is concluded with the calculation of Π(τ):

$$\Pi(\tau) = Var(x_\tau^*(p_T)) + 2cov((r - p_T)q_T, x_\tau^*(p_T)) + Var((r - p_T)q_T)$$

Where

$$x_\tau^*(p_T) = B_\tau p_T^{A_\tau+1} p_\tau^{-A_\tau} q_\tau - rB_\tau p_T^{A_\tau} p_\tau^{-A_\tau} q_\tau + rC_\tau q_\tau - D_\tau p_\tau q_\tau$$

Each term of Π(τ) is calculated as a function of τ as follows: (For notational convenience, subscript τ for A_τ, B_τ, C_τ and D_τ is omitted.)

$$\begin{aligned}
Var(x_\tau^*(p_T)) &= E[x_\tau^*(p_T)^2] \\
&= B^2E[p_T^{2A+2}p_\tau^{-2A}q_\tau^2] + r^2B^2E[p_T^{2A}p_\tau^{-2A}q_\tau^2] \\
&\quad +r^2C^2E[q_\tau^2] + D^2E[p_\tau^2q_\tau^2] - 2rB^2E[p_T^{2A+1}p_\tau^{-2A}q_\tau^2] \\
&\quad +2rBCE[p_T^{A+1}p_\tau^{-A}q_\tau^2] - 2BDE[p_T^{A+1}p_\tau^{-A+1}q_\tau^2] \\
&\quad -2r^2BCE[p_T^Ap_\tau^{-A}q_\tau^2] + 2rBDE[p_T^Ap_\tau^{-A+1}q_\tau^2] \\
&\quad -2rCDE[p_\tau q_\tau^2] \\
cov((r - p_T)q_T, x_\tau^*(p_T)) &= E[(r - p_T)q_Tx_\tau^*(p_T)] \\
&= rBE[p_T^{A+1}q_Tp_\tau^{-A}q_\tau] - r^2BE[p_T^Aq_Tp_\tau^{-A}q_\tau] \\
&\quad +r^2CE[q_Tq_\tau] - rDE[q_Tp_\tau q_\tau] \\
&\quad -BE[p_T^{A+2}q_Tp_\tau^{-A}q_\tau] + rBE[p_T^{A+1}q_Tp_\tau^{-A}q_\tau] \\
&\quad -rCE[p_Tq_Tq_\tau] + DE[p_Tq_Tp_\tau q_\tau] \\
Var((r - p_T)q_T) &= E[(rq_T - p_Tq_T)^2] - E[rq_T - p_Tq_T]^2 \\
&= r^2E[q_T^2] - 2rE[p_Tq_T^2] + E[p_T^2q_T^2] \\
&\quad -(rE[q_T] - E[p_Tq_T])^2
\end{aligned}$$

The expectation terms were calculated using

$$E[p_T^\alpha q_T^\beta p_\tau^\gamma q_\tau^\delta] = p_0^{\alpha+\gamma}q_0^{\beta+\delta}e^{\int_0^T(\alpha a_t+\beta c_t)dt}e^{\int_\tau^T(\frac{1}{2}(\alpha b_t+\beta d_t)^2+\frac{1}{2}\beta^2e_t^2)dt} \cdot e^{\int_0^\tau(\gamma a_t+\delta c_t+\frac{1}{2}((\alpha+\gamma)b_t+(\beta+\delta)d_t)^2+\frac{1}{2}(\beta+\delta)^2e_t^2)dt}$$

From

$$
\begin{aligned}
p_T^{\alpha} &= p_0^{\alpha} \exp\left(\int_0^T \alpha a_t dt + \int_0^T \alpha b_t dB_t^1\right) \\
q_T^{\beta} &= q_0^{\beta} \exp\left(\int_0^T \beta c_t dt + \int_0^T \beta d_t dB_t^1 + \int_0^T \beta e_t dB_t^2\right) \\
p_\tau^{\gamma} &= p_0^{\gamma} \exp\left(\int_0^\tau \gamma a_t dt + \int_0^\tau \gamma b_t dB_t^1\right) \\
q_\tau^{\delta} &= q_0^{\delta} \exp\left(\int_0^\tau \delta c_t dt + \int_0^\tau \delta d_t dB_t^1 + \int_0^\tau \delta e_t dB_t^2\right).
\end{aligned}
$$

4.2.4. An Example

We now illustrate the optimal hedging timing with a concrete example. The example assumes that the maturity of the portfolio is one year from now. Base values of the parameters are set according to the empirical estimates of Audet et al 2004, which were also used by Nasakkala and Keppo 2005. Specifically, set

$$\mu_p(t) = 0, \quad \sigma_p(t) = e^{-\psi(T-t)}\sigma$$

where σ is the spot volatility and ψ is a mean-reversion rate of the spot price process, i.e., a rate at which forward volatility is discounted from the spot volatility. We also set $\mu_q(t) = 0$. In addition, $\sigma_{pq}(t)$ and $\sigma_q(t)$ are assumed as constants, so as to have a constant load volatility and correlation:

$$\sigma_L = \sqrt{\sigma_{pq}^2 + \sigma_q^2}, \qquad \phi = \frac{\sigma_{pq}}{\sigma_L}.$$

The resulting process is then

$$
\begin{aligned}
\frac{dp_t}{p_t} &= e^{-\psi(T-t)}\sigma dB_t^1 \\
\frac{dq_t}{q_t} &= \phi\sigma_L dB_t^1 + \sqrt{1-\phi^2}\sigma_L dB_t^2.
\end{aligned}
$$

The forward price and load estimate for a month one year later is assumed to be 20Euro/MWh and 1000 MWh. The following table summarizes the base values of the parameters:

Parameter	T	r	p_0	q_0	ψ	σ	σ_L	ϕ
Value	1	40	20	1000	4.02	0.7	0.1	0.7

To study how the optimal hedging time is affected by various parameters, a sensitivity analysis of optimal hedging time with respect to parameter values is illustrated Figure 4-30.

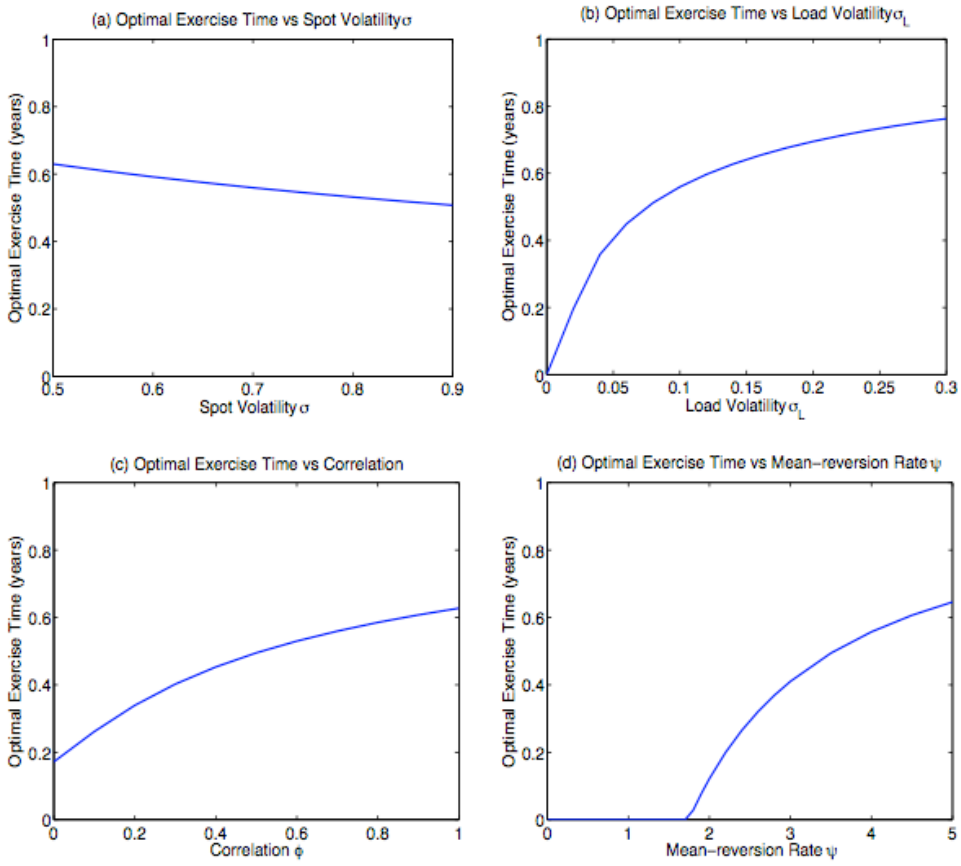


Figure 4-30: Optimal Hedging Time versus Other Parameter Values

Figure 4-30 (a) plots the optimal hedging time against the spot price volatility σ, and it shows that a higher spot volatility favors earlier hedging. Intuitively, a higher spot volatility increases uncertainties in the future information, which justifies locking in the price of hedging contracts earlier.

Figure 4-30 (b) plots the optimal hedging time against the load volatility σL. It shows that a higher volatility in the load estimate postpones the hedging time, confirming the intuition that the inaccuracy in the load estimate will delay the hedging so as to obtain more information.

Figure 4-30 (c) plots the optimal hedging time against the correlation between forward price and load estimate. It shows that a lower correlation makes earlier hedging more favorable which can be explained by the fact that high correlation reduces uncertainties in profit and thus delays hedging to take advantage of more information.

Figure 4-30(d) plots the optimal hedging time against the mean-reversion rate of spot price. The figure shows that increase in mean-reversion rate of the spot price postpones the hedging time, since higher mean-reversion rate of spot price decreases the volatility of forward prices, so it will not be as risky to postpone the hedging time.

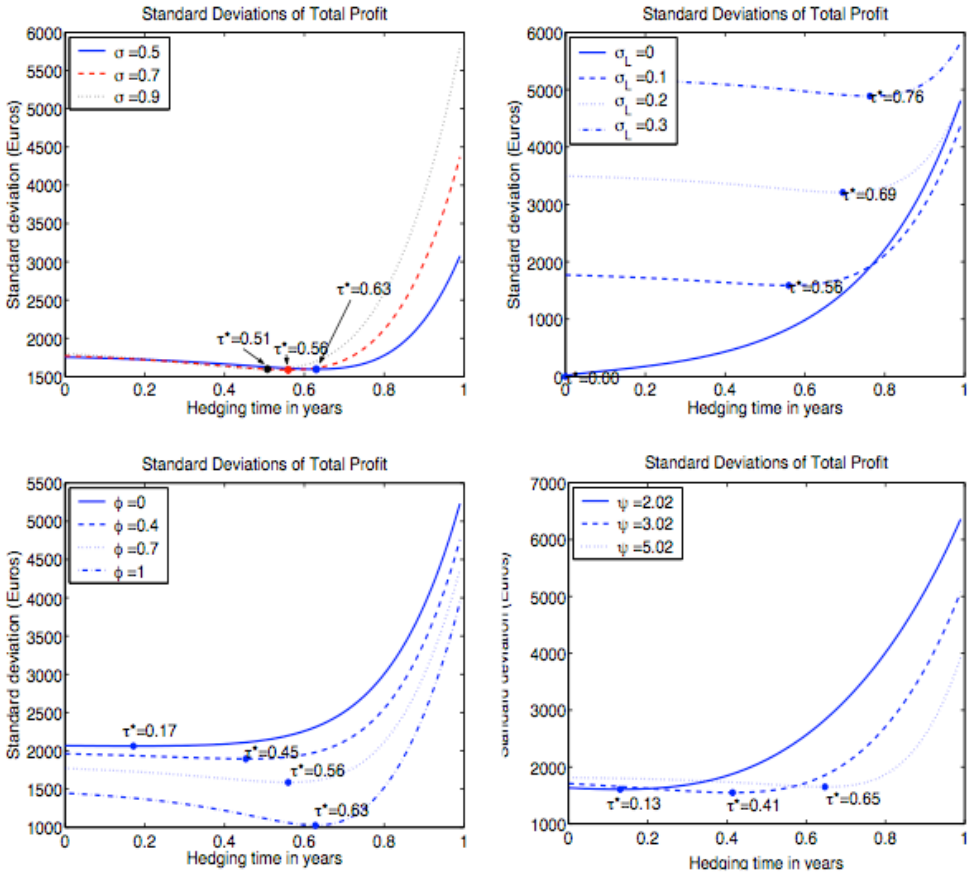


Figure 4-31: Standard Deviation of Hedged Profit Versus Hedging Times

Figure 4-31 the variance of the optimally hedged profit as function of the hedging time. We note that hedging at time 0 versus the optimal time τ make little difference in the variance of hedged profit inmost cases. However, the variance of profit increases rapidly if hedging is delayed beyond the optimal time.

Figure 4-31 also show the level of uncertainties changes with respect to the changes in σ, σL, φ, and ψ. The data displayed in the figure indicates that the profit uncertainty increases with the increases in spot and load volatility, and decreases with the mean reversion rate and correlation coefficient.

It is also note worthy that hedging at the optimal time may not make any differences in the variance of hedged profit even for different parameters such as volatility and mean-reversion rate of spot price. In other words, the increased uncertainty from higher volatility in the forward price can be overcome by the optimal choice of hedging time.

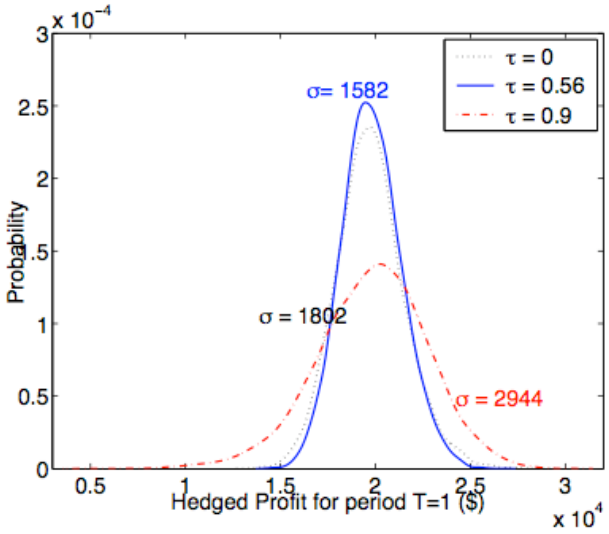
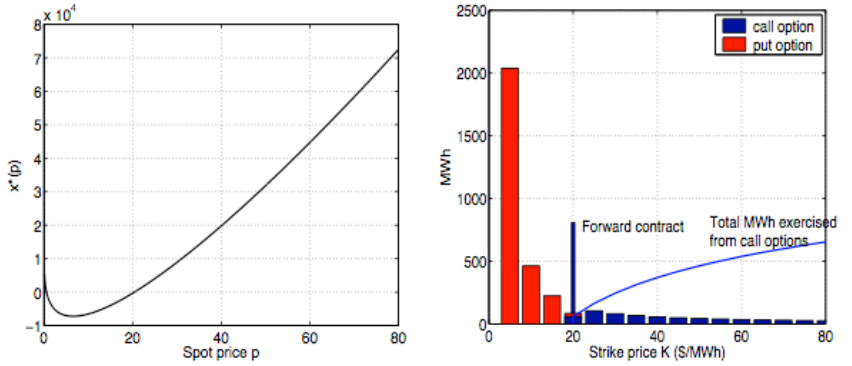


Figure 4-32: The distributions of profits when hedging at different times. The τ=0.56 is the distribution of the profits with hedging at the optimal hedging time, based on the information given at time 0.

Figure 4-32 compares the distributions of profits at delivery time when the hedging portfolio is purchased at time 0, at the optimal hedging time (0.56), and at time (0.9) close to delivery time. It also confirms that earlier hedging does not increase profit risk very much as compared to the optimal hedging time, but late hedging can have adverse consequences.

Finally, the optimal hedging strategy at time 0 under the base values of the parameters, is illustrated in Figure 4-33 which shows the optimal payoff function and its approximate replication (developed in Section 4.1.1) under the assumption that the hedging portfolio is constructed at time 0.



(a) The optimal payoff function

(b) The number of forward and options to purchase

Figure 4-33: The optimal payoff function and its replication when the hedging portfolio is constructed at time 0

5. VaR Constrained Static Hedging of Volumetric Risk

5.1. VaR-constrained Hedging Problem

We define VaR as a maximum possible loss at a $(1-\gamma)$ confidence level. In other words, VaR is the $(1-\gamma)$ percentile of the loss distribution. In this section, we present a model for the hedging portfolio subject to a VaR limit set by the risk manager for a specified horizon. This preset VaR level will reflect the risk tolerance of the risk manager.

Consider the LSE whose revenue is determined by a fixed retail price r and the uncertain demand q. Denoting uncertain wholesale electricity price per unit as p, the profit y(p, q) from retail sales at time 1 depends on the two random variable p and q. i.e.,

$$y(p,q) = (r-p)q$$

Let LSE's beliefs on the realization of spot price p and load q be characterized by a joint probability function f (p, q) for positive p and q, which is defined on the probability measure P.

Suppose the LSE hedges the profit through an exotic electricity option maturing at time 1. Let Y (x) be the hedged profit, then

$$Y(x) = y(p,q) + x(p) = (r-p)q + x(p)$$

where x(p) is a payoff function of the exotic option, which is contingent on the price of p.

With the VaR limit V_0, the VaR-constrained hedging problem is formulated as follows:

$$\begin{array}{ll} \max_{x(p)} & E[Y(x)] \\ s.t. & E^Q[x(p)] = 0 \\ & VaR_\gamma(Y(x)) \le V_0 \end{array} \tag{5.1}$$

for a random variable X , and with E[·] and EQ [·] denoting expectations under the probability measures P and Q, respectively. The formulation seeks the payoff function of a self-financing hedging portfolio at time 1, which maximizes the expected profit while requiring that a 1 $-\gamma$ percentile of the loss distribution does not exceed V_0.

The zero-cost constraint $E^Q[x(p)] = 0$ requires the manufacturing cost of the portfolio to be zero under a constant risk-free rate. This zero-cost constraint implies that purchasing derivative contracts may be financed from selling other derivative contracts or through money market accounts. In other words, under the assumption that there is no limit on the possible amount of instruments to be purchased and money to be borrowed, the model finds a portfolio from which the LSE obtains the maximum expected utility over total profit.

One might question the use of the optimal payoff function solved from the formulation (5.1). The optimal payoff function will eventually be used to derive the optimal quantities of forwards

and options at different strike prices of which the hedging portfolio consists. This approach of getting the payoff function first and then calculating the portfolio composition that replicates the payoff, not only makes the problem solvable but also provides valuable insights regarding the optimal hedging portfolio.

5.2. Optimal Payoff Function in the Mean-Variance Efficient Frontier

The VaR constraint in the formulation (5.1) cannot be written in a tractable form for optimization without very restrictive assumptions on the distribution of Y (x). If Y (x) is linear in the risk factors which are normally distributed, then it is possible to write VaR in a closed form. However, in the formulation (5.1), Y (x) has a multiplicative term of two risk factors and, moreover, a term of the unknown function x(p). Thus, a closed form of VaR(Y (x)) cannot be obtained in a form amenable to simple optimization.

The reason behind the normal distribution having been a common assumption when calculating VaR is the fact that the quantiles of the normal distribution (actually, VaR) can be expressed using mean and variance. Likewise, when VaR can be expressed using mean and variance - even in cases when a closed form of the VaR cannot be obtained - the VaR-constrained problem could be solved using the mean-variance framework.

Therefore, a key assumption throughout this section is that VaR(Y (x)) is solely determined by mean and variance of Y (x). In the following theorem adopted from Kleindorfer and Li 2005 we show that under such an assumption, monotonicity of the VaR in the mean and variance of the Y (x) corresponding to feasible hedging functions x(p) is sufficient to ensure that the mean-maximizing VaR-constrained solution to (5.1) lies on the efficient mean-variance frontier.

Theorem 1:

$$X(p) = \{x(p) : x(p) \text{ is a continuous function of } p \text{ such that } E^Q[x(p)] = 0\}.$$
$$\Psi = \{Y(x) : Y(x) = y(p,q) + x(p) \text{ where } x(p) \in X(p)\},$$
$$E = \{E[Y(x)] : Y(x) \in \Psi\}, \quad \Sigma = \{\sigma(Y(x)) : Y(x) \in \Psi\},$$

and $\Re$ *be a set of real numbers. Let's define* $VaR_\gamma(Y(x))$ *as* ν *such that*

$$P\{Y(x) \geq -\nu\} = 1 - \gamma.$$

Suppose now that there exists a continuous function $h : (E, \Sigma, \gamma) \rightarrow \Re$ *that satisfies*

$$VaR_{\gamma}(Y(x)) = h(\mu, \sigma, \gamma)$$

with $h(\mu, \sigma, \gamma)$ *which is increasing in* σ *and non-increasing in* μ *for* $\mu = E[Y(x)]$ *and* $\sigma^2 = V(Y(x))$.[6].

Then if $x^*(p)$ *solves the problem (1), then the following (a)* $\sim$ *(e) hold:*

(a) $x^*(p)$ *is on the efficient frontier of the (E-*VaR_{γ}*) plane*[7]*, on which any feasible* $x(p)$ *is mapped to a corresponding point* $(VaR_{\gamma}(Y(x)), E[Y(x)])$.

(b) $x^*(p)$ *is on the efficient frontier of the (E-V) plane*[8]*, on which any feasible* $x(p)$ *is mapped to a corresponding point* $(V(Y(x)), E[Y(x)])$.

(c) The variance on the efficient frontier in the (E-V) plane is non-decreasing in the mean.

(d) The variance on the efficient frontier in the (E-V) plane is a convex function of the mean.

(c) There exists $k > 0$ *such that* $x^*(p)$ *solves*

$$\max_{x(p) \in X(p)} E[Y(x)] - \frac{1}{2} k V(Y(x)). \qquad (3)$$

Theorem 2:

$$x^k(p) = \arg \max_{x(p) \in X(p)} E[Y(x)] - \frac{1}{2} k V(Y(x)).$$

Then $E[Y(x^k)]$ *and* $V(Y(x^k))$ *are monotonically non-increasing in* k.

5.3. The Optimal Payoff Function when the Demand and Log Price Follows Bivariate Normal Distribution

It is often assumed that the electricity demand and logarithm of price are normally distributed with some correlation. In Proposition 4.2 and equation (4.6), we show that under such assumption a closed form of $x_k(p)$can be obtained. We have also shown in the previous section that $E[Y(x_k(p))]$ and the variance $V[Y(x_k(p))]$ are non-increasing in k. We will now describe an approximation procedure that searches for an approximate solution to the VaR-constrained Expected-value-maximizing self-financed hedging function along the mean-variance efficient frontier. The justification for this approximation is motivated by the intuitively plausible properties of the VaR that make such an approximation exact. The approximation is also supported by the fact that the required properties are met by the Chebyshev upper bound on the VaR so that tightening the VaR constraints by replacing the VaR with its Chebyshev approximation will also produce results that lie on the mean-variance efficient frontier.

To obtain the approximate solution we characterized above, we start with $k = \varepsilon$. (ε is a small constant). Using the formula for $x_k(p)$ given in equation (4.6), we compute the corresponding VaR_γ (k) =VaR_γ (Y (x(p))) using a Monte-Carlo simulation such that

$$P\{(r-p)q + x^k(p) \geq -VaR_\gamma(k)\} = 1 - \gamma.$$

We then repeat the process incrementing k until VaRγ (k) ≤ V0 at which point we set $k^* = k$. The monotonicity of the mean in k coming from Theorem 2 guarantees that the first k at which the VaR constraint is satisfied will yield the largest expected value. We can replicate such exotic payoff functions by means of the formula suggested in Chapter 4.

5.4. An Example

In this section we demonstrate the computation of an approximate optimal VaR-constrained volumetric hedging problem using the method developed in the previous section. Consider a hypothetical LSE that charges a flat retail rate r = $120/MWh to its customers. The wholesale spot price p at which the LSE must purchase its power and the load q it is obligated to serve in any fixed time interval (typically 15 minutes), are distributed according to a bivariate distribution in quantity and log price:

$$\begin{array}{ll} \text{Under } P: & \log p \sim N(4, 0.7^2), \quad q \sim N(3000, 600^2), \quad Corr(\log p, q) = 0.8 \\ \text{Under } Q: & \log p \sim N(4.1, 0.7^2) \end{array}$$

Note that we assume here $P \neq Q$. Otherwise, the mean-variance problem has the same solution for all k. In such case, the VaR-constrained problem either has the same solution as the variance-minimizing problem, or is infeasible.

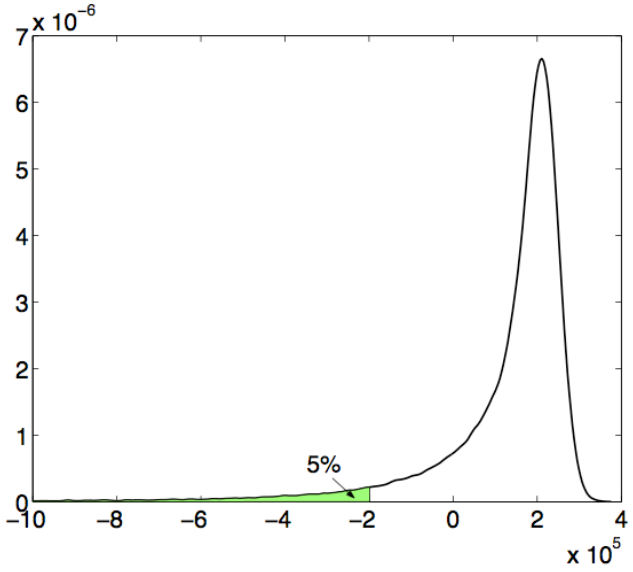


Figure 5-1: Distribution of the unhedged profit y(p,q)=(r −p)q

Figure 5-1 shows a distribution of unhedged profit,

$$y(p, q) = (120 - p)q.$$

95% VaR is also indicated in the figure, which is about $20, 000. The mean of the distribution is $127, 000. This implies that there is 5% chance that the LSE can take a loss of more than $20, 000. The VaR-constrained problem for the LSE which seeks a hedging strategy that maximizes the expected profit with at least $60, 000 profit with 95% probability is formulated as follows:

$$\begin{array}{ll} \max_{x(p)} & E[Y(x)] \\ s.t. & E^{Q}[x(p)] = 0 \\ & VaR_{\gamma}(Y(x)) \leq -60000 \end{array}$$

Where

$$Y(x) = (120 - p)q + x(p) \text{ and } Pr\{Y(x) \geq -VaR_{\gamma}(Y(x))\} = 0.95.$$

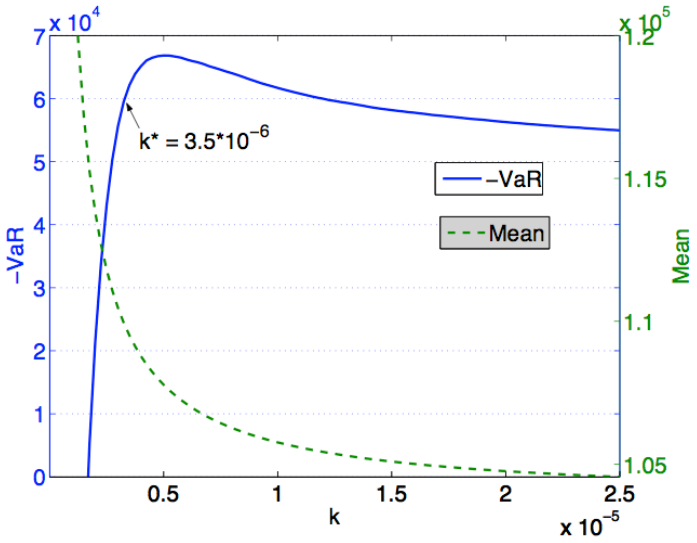
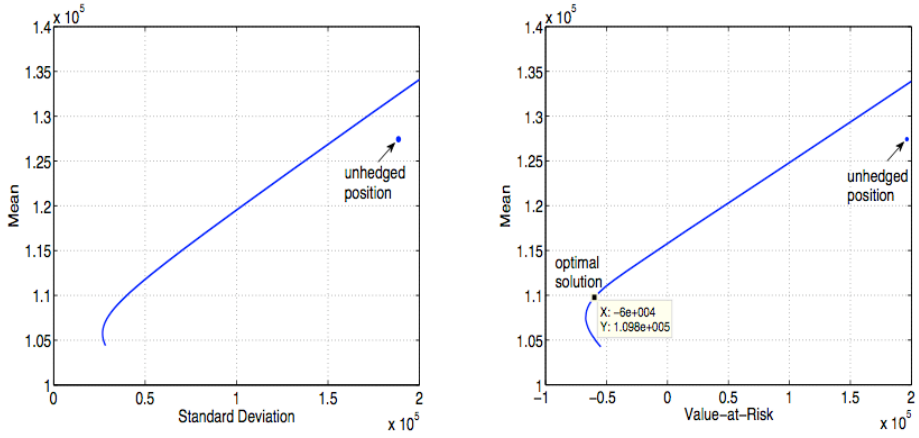


Figure 5-2: −VaR(k) in the left y-axis and E[Y(xk(p))] in the right y-axis. The optimal k^* is obtained as the first k that provides-VaR no less than the required level 60,000.

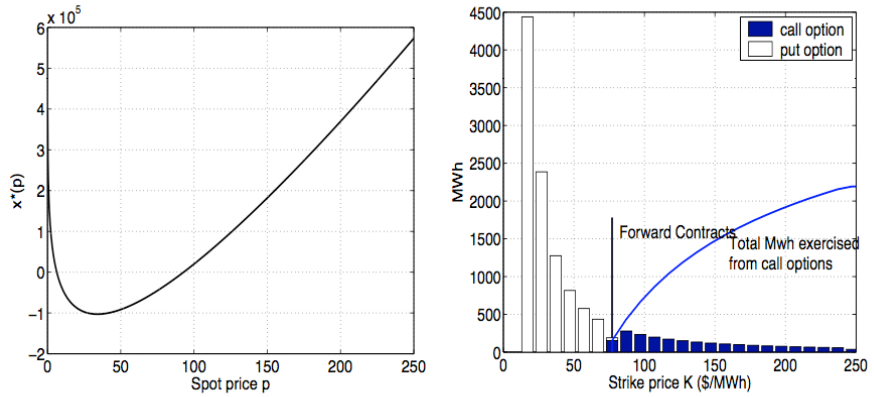
Motivated by Theorem 1 we restrict our search for solution to the VaR constrained problem to optimal solutions for the mean-variance problems for various risk-aversion levels k and for each such candidate solution we compute the corresponding VaR. The relationship between VaR and k is drawn in Figure 5-2 as an example. The figure also shows the mean of the hedged profit, E[Y (x_k)], on the right axis, which is non-increasing in k as proven in Theorem 2. Because of the monotonicity of the mean in k selecting the first value of k that meets the VaR constraint as k^* = 3.5×10^{-6} gives the largest mean value with −VaR $_{\gamma}$ (k) ≥ 60000 among all hedging portfolios that maximize a mean-variance criterion.



(a) Mean-variance frontier (b) Mean-VaR frontier

Figure 5-3: Mean-variance frontier and mean-VaR frontier

Figure 5-3 illustrates the mean-variance efficient frontier and the corresponding mean-VaR frontier for our example. Note that the mean-VaR frontier is the efficient mean-VaR frontier only if the distribution of hedged profit satisfies the monotonicity properties postulated in Theorem 1.



(a) The optimal payoff function (b) Replicating strategy

Figure 5-4: Hedging strategy for an LSE that maximizes the expected pay-off with VaR constraints of −$60,000. The underlying distributions of spot prices and load are logp ~N(4,0.72), q ~N(3000,6002), and Corr(logp,q)=0.8 (assuming r=$120/MWh)

The optimal mean-variance hedging strategy corresponding to k^* and hence, the approximation to the optimal mean-VaR hedging strategy, is shown in Figure 5-4. Figure 5-4(a) shows the payoff function $x^*(p) = x_k^*(p)$ obtained as an approximation for the VaR-constrained problem, and Figure 5-4(b) illustrates its replicating strategy consisting of forwards, calls, and puts, as described in Section 5.3.

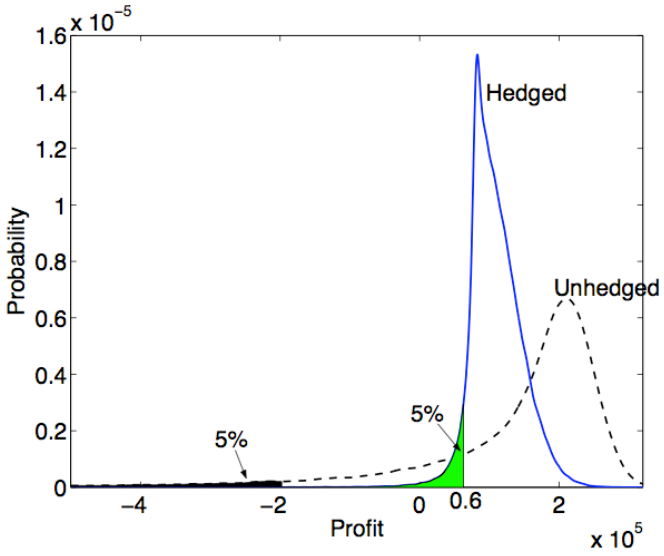


Figure 5-5: Profit distributions and VaRs before and after the optimal hedge

Figure 5-5 compares profit distributions before and after hedging. One can see that the hedge obtained as an approximate solution to the VaR-constrained problem reduces the left-tail of the profit distribution significantly.

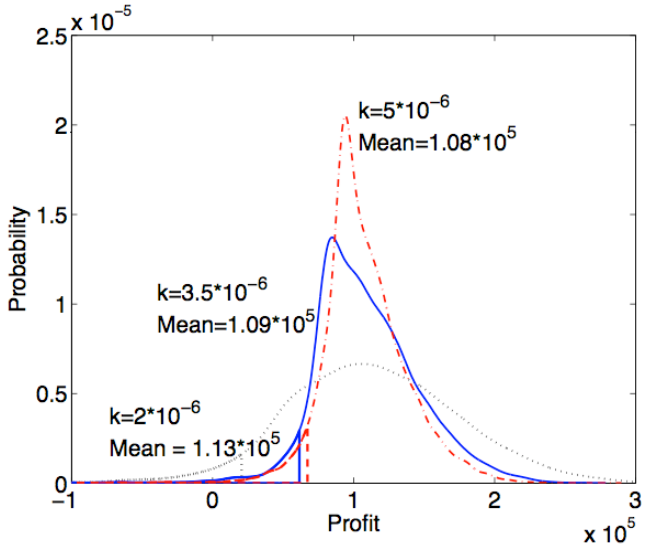


Figure 5-6: Profit distribution and its VaR for various levels of k

Figure 5-6 shows the profit distributions for different k. The corresponding VaR is represented as the vertical line from the distribution to the x-axis. $k = 3.5 \times 10^{-6}$ corresponds to profit after the optimal hedge. One can see that $k = 2 \times 10^{-6}$ gives the higher expected value, 1.13×10^5, than the optimal one, but it was rejected from the feasible hedge because its VaR level exceeds the required level of−\$60,000. The graph for $k = 5 \times 10^{-6}$ shows a case of VaR

satisfying the required level, but it was not chosen for the optimum since it provides a lower expected profit than the optimal one.

6. Conclusion

We apply manifold-based dimension reduction to electricity price curve modeling. LLE is demonstrated to be an efficient method for extracting the intrinsic low-dimensional structure of electricity price curves. Using price data taken from the NYISO, we find that there exists a low-dimensional manifold representation of the day-ahead price curve in NYPP, and specifically, the dimension of the manifold is around 4. The interpretation of each dimension and the cluster analysis in the low-dimensional space are given to analyze the main factors of the price curve dynamics. Numerical experiments show that our prediction performs well for the short-term prediction, and it also facilitates medium-term prediction, which is difficult, even infeasible for other methods.

We also propose an equilibrium pricing model in a multi-commodity setting that is driven by demand for weather derivatives which is derived from hedging and risk diversification activities in weather sensitive industries. As a part of our analysis, we measure the risk hedging and sharing effects of the weather derivative, both of which contribute to increasing the expected utility of risk averse agents that include these instruments in their hedging portfolios. To price the weather derivative we assume that there are buyers and an issuer in a closed and frictionless endowment economy and all of them are utility maximizers. By solving the utility maximization problems of the market participants we determine the optimal demand and supply functions for weather derivatives and obtain their equilibrium prices by invoking a market clearing condition. In the multi-commodity economy the weather derivative has two effects: the risk hedging effect and the risk sharing effect, while in a single-commodity economy there is only a risk hedging effect since there is no counter-party to share risk. We measure these effects in terms of certain equivalent differences among various cases.

Under the mean-variance utility function we were able to derive closed form expressions for equilibrium prices and the measurement of the risk hedging and sharing effects. Such expressions will be useful in future empirical work that will attempt to calibrate the model parameter to market data. Numerical examples employing Monte-Carlo simulations show that the equilibrium price tends to increase as the correlation between temperature and demand increase due to the high demand for the weather derivative. In addition, the numerical examples verify that weather derivative improves hedging and risk diversification capability, especially in situations where commodity derivatives are not available.

In addition, we developed a method of mitigating volumetric risk that load-serving entities (LSEs) and marketers of default service contract face in providing their customers' load following service at fixed or regulated prices while purchasing electricity or facing an opportunity cost at volatile wholesale prices. Exploiting the inherent positive correlation and multiplicative interaction between wholesale electricity spot price and demand volume, we developed a hedging strategy for the LSE's retail positions (which is in fact a short position on unknown volume of electricity) using electricity standard derivatives such as forwards, calls, and puts.

The optimal hedging strategy was determined based on expected utility maximization, which has been used in the hedging literature to deal with non-tradable risk. We derived an optimal payoff function that represents the payoff of the optimal costless exotic option as a function of

price. We then showed how the optimal exotic option can be replicated using a portfolio of forward contracts and European options. The examples demonstrated how call and put options can improve the hedging performance when quantity risk is present, compared to hedging with forward contracts alone. While at present the liquidity of electricity options is limited, the use of call options has been advocated by Oren 2005 and Chao and Wilson 2004 in the electricity market design literature as a tool for resource adequacy, market power mitigation, and spot volatility reduction. These authors advocated capacity payments in the form of option premiums that will incent capacity investment, and ensure electricity supply at a predetermined strike price. Our research contributes to better understanding of how options can be utilized in hedging the LSE's market risk, and hopefully increase their liquidity in the electricity market.

We also extended our framework by considering the optimal timing of a hedging portfolio as well as the co-optimization of the portfolio mix taking account of the timing. For mean-variance expected utility, we solved for the optimal hedging time, under classical assumption regarding the stochastic processes governing forward price and load-estimate. The example showed that generally there is a critical time beyond which the uncertainty in profit increases sharply while the uncertainty remains relatively constant before this critical time.

Sensitivity analysis results indicate that the optimal hedging time gets closer to the delivery period if the positive correlation between the forward price and load-estimate is higher, and if the load-estimate volatility is higher. It is also observed that delaying the hedging time past the optimum time can be very risky, while the earlier hedging makes little difference as compared with hedging at the optimal time. This suggests that in practice one should err by hedging early rather than taking the chance of being too late.

Finally, the hedging strategy is extended to maximize the expected profit under the VaR constraint, which limits the lowest level below which the hedged profit wouldn't fall with 95% confidence. However, VaR constrained problems are generally very hard to solve analytically unless the value of profit under consideration is normally distributed. In our case, the profit depends on the product of the two correlated variables. Moreover our hedging strategy is characterized by a nonlinear function of a random variable. We address this difficulty by limiting our search to feasible VaR-constrained self-financed hedging portfolios on the mean-variance efficient frontier. We provide theoretical justification to such an approximation and derive, an analytic representation of hedging portfolios on the mean-variance efficient frontier as function of the risk aversion factor. The computation of an approximate solution to the VaR-constrained problem on the mean variance efficient frontier is facilitated by the fact that it corresponds to the smallest risk-aversion factor whose associated VaR meets the constraint limit.

When one uses the mean-variance formulation, it is usually easy to solve the problem, but hard to decide what the appropriate risk-aversion factor is. The analysis in this section implies that one can use a VaR-constrained formulation as an alternative, which takes one of the mean-variance solutions but automatically chooses associated risk aversion at which the maximum mean is achieved while maintaining the required VaR level. The advantage of using the VaR-constrained formulation is that VaR is easier to interpret, and it is a widely used risk-measure in practice.

The model presented in Chapter 4 and 5 determined the best hedging portfolio assuming that an LSE has unlimited borrowing capability. In practice, credit limits can become an impeding factor in purchasing the optimal hedging portfolio. An LSE may not be able to borrow enough upfront money to finance the option contracts. Therefore, a credit limit constraint, which limits the amount of money that can be borrowed to construct the portfolio, needs to be considered in future extension of our model. A dynamic hedging strategy rather than the static approach adapted in this project is likely to improve the hedging performance and should be considered in future extension of this work.

7. Project Publications

[1] Oum Y, Oren SS, Deng SJ. Hedging quantity risks with standard power options in a competitive wholesale electricity market. Special Issue on Applications of Financial Engineering in Operations, Production, Services, Logistics, and Management. Naval Research Logistics 2006; 53: 679-712.

[2] Jie Chen, Shi-Jie Deng, and Xiaoming Huo, "Electricity Price Curve Modeling and Forecasting by Manifold Learning", IEEE Transactions on Power Systems, Vol 23, No. 3 (2008) pp 877-888.

[3] Yongheon Lee, Shmuel S. Oren, "An equilibrium pricing model for weather derivatives in a multi-commodity setting", Energy Economics, In Press (2009).

[4] S.J. Deng "Analysis on Cross-market Trading Strategy on Electricity with Manifold Learning and Logistic Smooth Transition Regression", working paper, Georgia Institute of Technology, January 2008.

[5] Yumi Oum, Shmuel S. Oren, "Optimal Static Hedging of Volumetric Risk in a Competitive Wholesale Electricity Market", working paper, UC Berkeley, September 2007.

[6] Oum Yumi and Shmuel Oren, "VaR Constrained Hedging of Fixed Price Load-Following Obligations in Competitive Electricity Markets", Journal of Risk and Decision Analysis, Vol 1, No.1 (2009) pp 43-56.

[7] S. J. Deng, L. Xu, "Mean-risk Efficient Portfolio Analysis of Demand Response and Supply Resources", Energy, Vol 34 (2009) pp 1523–1529.

8. References

[1] D.-H. Ahn, J. Boudoukh, M. Richardson, and R.F. Whitelaw. Optimal risk management using options. The Journal of Finance, 54:359–375, 1999.

[2] G.J. Alexander and A.M. Baptista. Economic implications of using a mean-var model for portfolio selection: A comparison with mean-variance analysis. Journal of Economic Dynamics & Control, 26:1159 – 1193, 2002.

[3] Ankirchner, Stefan, Peter Imkeller, Alexandre Popier. 2006. Optimal cross hedging of insurance derivatives. Working paper.

[4] N. Audet, P. Heiskanen, J. Keppo, and I.Vehvilainen. Modeling Electricity Forward Curve Dynamics in the Nordic Market, Modelling Prices in Competitive Electricity Markets. Wiley Series in Financial Economics, 2004.

[5] M. Belkin and P. Niyogi, "Laplacian eigenmaps for dimensionality reduction and data representation," Neural Computation, vol. 15, no. 6, pp. 1373–1396, June 2003.

[6] Bhattacharya K, Bollen M, Daalder J. Operation of Restructured Power System. Kluwer Academic Publishers, 2001.

[7] I. Borg and P. Groenen, Modern Multidimensional Scaling: Theory and Applications. New York: Springer-Verlag, 1997.

[8] Brockett, Patrick L., Mulong Wang. 2006. Portfolio effects and valuation of weather derivatives. The Financial Review 41.

[9] P. J. Brockwell, Introduction to Time Series and Forecasting, 2nd ed. Springer, 2003.

[10] G.W. Brown and K.B. Toft. How firms should hedge. The Review of Financial Studies, 15(4):1283–1324, 2002.

[11] Cao, Melanie, Jason Wei. 1999. Pricing weather derivative: an equilibrium approach. Working paper.

[12] Carr, Peter, Dilip Madan. 2001. Optimal positioning in derivative securities. Quantitative Finance 1 19–37.

[13] H-P. Chao and R. Wilson. Resource adequacy and market power mitigation via option contracts. In 2004 POWER Ninth Annual Research Conference, 2004.

[14] Chaumont, Sebastien, Peter Imkeller, Matthias Muller, Ulrich Horst. 2005. A simple model for trading climate risk. Quarterly Journal of Economic Research 74.

[15] R. B. Cleveland, W. S. Cleveland, J. McRae, and I. Terpenning, "STL: A seasonal-trend

decomposition procedure based on loess," Journal of Official Statistics, vol. 6, pp. 3–73, 1990.

[16] CME. 2005. An introduction to CME weather products. www.cme.com/weather.

[17] J. Contreras, R. Espınola, F. J. Nogales, and A. J. Conejo, "ARIMA models to predict next-day electricity prices," IEEE Transactions on Power Systems, vol. 18, no. 3, pp. 1014–1020, 2003.

[18] A. J. Conejo, J. Contreras, R. Espınola, and M. Plazas, "Forecasting electricity prices for a day-ahead pool-based electric energy market," International Journal of Forecasting, vol. 21, pp. 435–462, 2005.

[19] A. J. Conejo, M. A. Plazas, R. Espınola, and A. B. Molina, "Day-ahead electricity price forecasting using the wavelet transform and ARIMA models," IEEE Transactions on Power Systems, vol. 20, no. 2, pp. 1035–1042, 2005.

[20] J-P Danthine. Information, futures prices, and stabilizing speculation. Journal of Economic Theory, 17:79–98, 1978.

[21] M. Davison, L. Anderson, B. Marcus, and K. Anderson, "Development of a hybrid model for electricity spot prices," IEEE Transactions on Power Systems, vol. 17, no. 2, pp. 257–264, 2002.

[22] S. J. Deng, "Stochastic models of energy commodity prices and their applications: Mean-reversion with jumps and spikes," UCEI POWER Working Paper P-073, 2000.

[23] S. J. Deng and W. J. Jiang, "Levy process driven mean-reverting electricity price model: a marginal distribution analysis," Decision Support Systems, vol. 40, no. 3-4, pp. 483–494, 2005.

[24] Deng, Shijie, Shmuel. S. Oren. 2006. Electricity derivatives and risk management. Energy 31.

[25] D. L. Donoho and C. Grimes, "Hessian eigenmaps: new locally linear embedding techniques for high-dimensional data," Proceedings of the National Academy of Sciences, vol. 100, pp. 5591–5596, 2003.

[26] D. Duffie and T. Zariphopoulou. Optimal investment with undiversifiable income risk. Mathematical Finance, 3:135–148, 1993.

[27] Dutton, John A. 2002. Opportunities and priorities in a new era for weather and climate services. American Meteorological Society.

[28] A. Eydeland and K. Wolyniec. Energy and power risk management: new development in modeling, pricing, and hedging. John Willy & Sons, Inc., 2003.

[29] G. Feder, R.E. Just, and A. Schmitz. Futures markets and the theory of the firm under price uncertainty. The Quarterly Journal of Economics, March 1980.

[30] S.-E. Fleten, S.W. Wallace, and W.T. Ziemba. Hedging electricity portfolios via stochastic programming. Working paper, NTNU, 1999.

[31] A. M. Gonzalez, A. M. S. Roque, and J. G. Gonzalez, "Modeling and forecasting electricity prices with input/output hidden Markov models," IEEE Transactions on Power Systems, vol. 20, no. 2, pp. 13–24, 2005.

[32] J. Gussow. Power systems operations and trading in competitive energy markets. PhD thesis, HSG, 2001.

[33] Hamisultane, Helene. 2007. Extracting information from the market to price the weather derivatives. Icfai Journal of Derivatives Markets 4 17–46.

[34] T. Hastie, R. Tibshirani, and J. Friedman, The elements of statistical learning. Springer, 2001.

[35] H. He and H. Pages. Labor income, borrowing constraints, and equilibrium asset prices. Economic Theory, 3:663–696, 1993.

[36] F. Herzog. Optimal dynamic control of hydro-electric power production. Master's thesis, ETHZ, 2002.

[37] D.M. Holthausen. Hedging and the competitive firm under price uncertainty. American Economic Review, 69(5):989–995, 1979.

[38] X. Huo, X. Ni, and A. K. Smith, Mining of Enterprise Data. Springer, 2005, new york Ch. A survey of manifold-based learning methods, invited book chapter, to appear, also available at http://www2.isye.gatech.edu/statistics/papers/06-10.pdf.

[39] X. Huo and J. Chen, "Local linear projection (LLP)," in First IEEE Workshop on Genomic Signal Processing and Statistics (GENSIPS), Raleigh, NC, October 2002, http://www.gensips.gatech.edu/proceedings/.

[40] X. Huo, "A geodesic distance and local smoothing based clustering algorithm to utilize embedded geometric structures in high dimensional noisy data," in SIAM International Conference on Data Mining, Workshop on Clustering High Dimensional Data and its Applications, San Francisco, CA, May 2003.

[41] B. Johnson and G. Barz, Energy Modelling and the Management of Uncertainty. Risk Books, 1999, ch. Selecting Stochastic Processes for Modeling Electricity Prices, London.

[42] M.S. Kimball. Precautionary saving in the small and in the large. Econometrica, 58:53–73,

1990.

[43] M.S. Kimball. Standard risk aversion. Econometrica, 61:589–611, 1993.

[44] P. Kleindorfer and L. Li. Multi-period VaR-constrained portfolio optimization with applications to the electric power sector. Energy Journal, 26(1):1–26, 2005.

[45] C. Knittel and M. Roberts, "Empirical examination of deregulated electricity prices," Energy Economics, vol. 27, no. 5, pp. 791–817, 2005.

[46] S. Koekebakker and F. Ollmar. Forward curve dynamics in the nordic electricity market. Managerial Finance, 31:73–94, 2005.

[47] Kroll, Yoram, Haim Levy, Harry M. Markowitz. 1984. Mean-variance versus direct utility maximization. The Journal of Finance 39, No.1 47–61.

[48] J. B. Kruskal, "Multidimensal scaling by optimizing goodness of fit to a nonmetric hypothesis," Psychometrika, vol. 29, pp. 1–27, 1964.

[49] E. Levina and P. J. Bickel, "Maximum likelihood estimation of intrinsic dimension," in Advances in Neural Information Processing Systems 17 (NIPS2004). MIT Press, 2005.

[50] A. T. Lora, J. M. R. Santos, A. G. Exposito, J. L. M. Ramos, and J. C. R. Santos, "Electricity market price forecasting based on weighted nearest neighbors techniques," Working Paper, University of Sevilla, Spain, 2006.

[51] J. J. Lucia and E. S. Schwartz, "Electricity prices and power derivatives: Evidence from the nordic power exchange," Review of Derivatives Research, vol. 5, no. 1, pp. 5–50, 2002.

[52] Colin Loxley and David Salant. Default service auctions. Journal of Regulatory Economics, 26(2):201 – 229, 2004.

[53] Markowitz HM. Portfolio selection. Journal of Finance 1952; 7 (1): 77- 91.

[54] R.I. McKinnon. Future markets, buffer stocks, and income stability for primary producers. The Journal of Political Economy, 75(6):844–861, December 1967.

[55] A. Misiorek, S. Trueck, and R. Weron, "Point and interval forecasting of spot electricity prices: Linear vs. non-linear time series models," Studies in Nonlinear Dynamics and Econometrics, vol. 10, no. 3, 2006, article 2.

[56] T. D. Mount, Y. Ning, and X. Cai, "Predicting price spikes in electricity markets using a regime-switching model with time-varying parameters," Energy Economics, vol. 28, no. 1, pp. 62–80, 2006.

[57] G. Moschini and H Lapan. International Economic Review, (4), November 1995.

[58] B. Nadler, S. Lafon, R. R. Coifman, and I. G. Kevrekidis, “Diffusion maps, spectral clustering and reaction coordinates of dynamical systems,” Applied and Computational Harmonic Analysis: Special issue on Diffusion Maps and Wavelets, vol. 21, pp. 113–127, July 2006.

[59] E. Nasakkala and J. Keppo. Electricity load pattern hedging with static forward strategies. Managerial Finance, 31(6):115–136, 2005.

[60] F. J. Nogales, J. Contreras, A. J. Conejo, and R. Espınola, “Forecast next-day electricity prices by time series models,” IEEE Transactions on Power Systems, vol. 17, no. 2, pp. 342–348, 2002.

[61] Y. Oum, S. Oren, and S. Deng. Hedging quantity risks with standard power options in a competitive wholesale electricity market. Special Issue on Applications of Financial Engineering in Operations, Production, Services, Logistics, and Management, Naval Research Logistics, 53:697–712, 2006.

[62] S.S Oren. Generation adequacy via call option obligations: safe passage to the promised land. Electricity Journal, November 2005.

[63] Platen, Eckhard, Jason West. 2004. A fair pricing approach to weather derivatives. Asia-Pacific Financial Markets 11 23–53.

[64] Richards, Timothy J., Mark R. Manfredo, Dwight R. Sanders. 2004. Pricing weather derivatives. American Journal of Agricultural Economics 86.

[65] B. Ramsay and A. J. Wang, “A neural network based estimator for electricity spot-pricing with particular reference to weekend and public holidays,” Neurocomputing, no. 47-57, 1998.

[66] L. K. Saul and S. T. Roweis, “Think globally, fit locally: unsupervised learning of low dimensional manifolds,” Journal of Machine Learning Research, vol. 4, pp. 119–155, 2003.

[67] L. K. Saul and S. T. Roweis, “Nonlinear dimensionality reduction by locally linear embedding,” Science, vol. 290, pp. 2323–2326, 2000.

[68] B. R. Szkuta, L. A. Sanabria, and T. S. Dillon, “Electricity price short-term forecasting using artificial neural networks,” IEEE Transactions on Power Systems, vol. 14, no. 3, pp. 851–857, 1999.

[69] J. B. Tenenbaum, V. de Silva, and J. C. Langford, “A global geometric framework for nonlinear dimensionality reduction,” Science, vol. 290, pp. 2319–2323, 2000.

[70] G. Unger. Hedging Strategy and Electricity Contract Engineering. PhD thesis, The Swiss Federal Institute of Technology, Zurich, 2002.

[71] I. Vehvilainen and J. Keppo. Managing electricity market price risk. European Journal of Operations Research, 145:136–147, 2003.

[72] P. Verveer and R. Duin, “An evaluation of intrinsic dimensionality estimators,” IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 17, no. 1, pp. 81–86, 1995.

[73] B. Willems. Virtual divestitures, will they make a difference? In 2006 POWER Eleventh Annual Research Conference, 2006.

[74] C.K. Woo, R. Karimov, and I. Horowitz. Managing electricity procurement cost and risk by a local distribution company. Energy Policy, 32(5):635–645, 2004.

[75] Kit Pong Wong. Currency hedging with options and futures. European Economic Review, 47:833–839, 2003.

[76] M. Wagner, P. Skantze, and M. Ilic. Hedging optimization algorithms for deregulated electricity markets. Proceedings of the 12th Conference on Intelligent Systems Application to Power Systems, 2003.

[77] Mingxin Xu. Risk measure pricing and hedging in incomplete markets. Finance 0406004, EconWPA, June 2004. Available at http://ideas.repec.org/p/wpa/wuwpfi/0406004.html.

[78] L. Zhang, P. B. Luh, and K. Kasiviswanathan, “Energy clearing price prediction and confidence interval estimation with cascaded neural networks,” IEEE Transactions on Power Systems, vol. 18, no. 1, pp. 99–105, 2003.

[79] Z. Zhang and H. Zha, “Principal manifolds and nonlinear dimension reduction via tangent space alignment,” SIAM Journal of Scientific Computing, vol. 26, no. 1, pp. 313–338, 2004.

Appendix A: Optimal Payoff Function under CARA Utility

Proof of Proposition 4.1:

We see from the special property $U'(Y) = -aU(Y)$ of a CARA utility function that the following condition holds:

$$E[U(Y^*)|p] = -\frac{\lambda^*}{a}\frac{g(p)}{f_p(p)},$$

which implies that the utility which is expected at any price level p is proportional to $\frac{g(p)}{f_p(p)}$.

Then the optimal condition is reduced to

$$E\left[e^{-a(y(p,q)+x^*(p))}\middle|p\right] = \lambda^*\frac{g(p)}{f_p(p)}$$

for an LSE with a CARA utility function. Then,

$$\begin{aligned} x^*(p) &= \frac{1}{a}\ln\left(\frac{1}{\lambda^*}\frac{f_p(p)}{g(p)}E\left[e^{-ay(p,q)}\middle|p\right]\right) \\ &= \frac{1}{a}\left(-\ln\lambda^* + \ln\frac{f_p(p)}{g(p)} + \ln E\left[e^{-ay(p,q)}\middle|p\right]\right) \end{aligned} \tag{A.1}$$

The Lagrange multiplier λ^* in the equation should satisfy the zero-cost constraint, which is $\int_{-\infty}^{\infty} x^*(p)g(p)dp = 0$. That is,

$$\int_{-\infty}^{\infty}\frac{1}{a}\left(-\ln\lambda^* + \ln\frac{f_p(p)}{g(p)} + \ln E\left[e^{-ay(p,q)}\middle|p\right]\right)g(p)dp = 0. \tag{A.2}$$

Solving (A.2) for $\ln\lambda^*$ gives

$$\ln\lambda^* = \int_{-\infty}^{\infty}\left(\ln\frac{f_p(p)}{g(p)} + \ln E\left[e^{-ay(p,q)}\middle|p\right]\right)g(p)dp. \tag{A.3}$$

Substituting this into equation (A.2) gives the optimal solution. **QED**.

Appendix B: Optimal Payoff Function under Mean-Variance Utility

Proof of Proposition 4.2:

The Lagrangian function for the optimization problem (4.2) is given by

$$\begin{aligned} L(x(p)) &= E\big[U(Y(p,q,x(p)))\big] - \lambda E^{Q}[x(p)] \\ &= \int_{-\infty}^{\infty} E\big[U(Y)|p\big] f_p(p)dp - \lambda \int_{-\infty}^{\infty} x(p)g(p)dp \end{aligned}$$

with a Lagrange multiplier λ and the marginal density function $f_p(p)$ of p under P. Differentiating $L(x(p))$ with respect to x(·) results in

$$\frac{\partial L}{\partial x(p)} = E\Big[\frac{\partial Y}{\partial x}U'(Y)\Big|p\Big] f_p(p) - \lambda g(p) \tag{B.1}$$

by the Euler equation. Setting (B.1) to zero and substituting ∂ Y/∂ x = 1 yields the first order condition for the optimal solution $x^*(p)$ as follows:

$$E\big[U'(Y(p,q,x^*(p)))|p\big] = \lambda^* \frac{g(p)}{f_p(p)} \tag{B.2}$$

Here, the value of λ^* should be the one that satisfies the constraint $E^{Q}[x(p)] = 0$.

It follows from $Var(Y) = E[Y^2] - E[Y]^2$ that

$$U(Y) \equiv Y - \frac{1}{2}a(Y^2 - E[Y]^2).$$

From $U'(Y) = 1 - aY$, the optimal condition (B.2) is as follows:

$$E[1 - aY^*|p] = \lambda^* \frac{g(p)}{f_p(p)}.$$

Equivalently,

$$f_p(p) - aE[Y^*|p] f_p(p) = \lambda^* g(p). \tag{B.3}$$

Integrating both sides with respect to p from $-\infty$ to ∞, we obtain $\lambda^* = 1 - aE[Y^*]$. By substituting λ^* and $Y^* = y(p,q) + x^*(p)$ into (B.3) gives

$$f_p(p) - a\Big(E[y(p,q)|p] + x^*(p)\Big) f_p(p) = g(p) - a\Big(E[y(p,q)] + E[x^*(p)]\Big) g(p). \tag{B.4}$$

By rearranging, we obtain

$$x^*(p) = \frac{1}{a} - \frac{1}{a}\frac{g(p)}{f_p(p)} + \Big(E[y(p,q)] + E[x^*(p)]\Big)\frac{g(p)}{f_p(p)} - E[y(p,q)|p] \tag{B.5}$$

To cancel out $E[x^*(p)]$ in the right-hand side, we take the expectation under Q to the both sides to obtain

$$0 = \frac{1}{a} - \frac{1}{a}E^Q\left[\frac{g(p)}{f_p(p)}\right] + \Big(E[y(p,q)] + E[x^*(p)]\Big)E^Q\left[\frac{g(p)}{f_p(p)}\right] - E^Q[E[y(p,q)|p]], \tag{B.6}$$

and subtract Eq.(B.6)$\times\frac{g(p)/f_p(p)}{E^Q[g(p)/f_p(p)]}$ from Eq.(B.5). This gives the final formula for the optimal payoff function under mean-variance utility as

$$x^*(p) = \frac{1}{a}\left(1 - \frac{\frac{g(p)}{f_p(p)}}{E^Q[\frac{g(p)}{f_p(p)}]}\right) - E[y(p,q)|p] + E^Q[E[y(p,q)|p]]\frac{\frac{g(p)}{f_p(p)}}{E^Q[\frac{g(p)}{f_p(p)}]} \tag{B.7}$$

QED.