Actuarial Review for Price Volatility Factor Methodology

SUMARIA SYSTEMS, INC.

August 8, 2014

George Duffield
Sumaria Systems, Inc.
618-632-8450
gduffield@sumaria.com

Authored by:Barry K. Goodwin, Ardian Harri, Rodrick M. Rejesus, Keith H. Coble, and Thomas O. Knight

Sumaria

Table of Contents

EXECUTIVE SUMMARY 3
1. INTRODUCTION AND BACKGROUND 5
2. LITERATURE REVIEW: IMPLIED VOLATILITY IN CROP INSURANCE RATING 7
2.A. BLACK-SCHOLES MODEL, VOLATILITY SMILES, AND BIAS 7
2.B. IMPLIED VOLATILITY IN AGRICULTURAL MARKETS 13
2.C. IMPLIED VOLATILITY AND CROP INSURANCE 16
2.D. SUMMARY 18
3. DATA, ASSUMPTIONS, AND ADEQUACY OF RMA'S EXISTING METHOD FOR ESTABLISHING PRICE VOLATILITY FACTORS FOR COMBO PRODUCTS 19
4. REVIEW OF PROCEDURES FOR DETERMINING AND USING IMPLIED VOLATILITIES 23
4.A. USE OF THE AVERAGE IMPLIED VOLATILITY FROM THE FINAL FIVE DAYS OF PRICE DISCOVERY 23
4.B. EXAMINATION OF THE MECHANISM TO SIMULATE PRICE VOLATILITY DERIVED FROM OPTION PRICES 23
5. REVIEW THE USE OF THE PRICE VOLATILITY FACTOR WITHIN THE REVENUE RATE SIMULATION MODEL UNDERLYING COMBO 27
6. COMPARE AND CONTRAST TWO OR MORE ALTERNATIVE METHODS FOR CALCULATING IMPLIED VOLATILITY TO THE METHOD USED BY RMA 37
6.A. EMPIRICAL ANALYSIS OF PRICE VOLATILITY 37
6.B THE BLACK-SCHOLES MODEL AND ITS ALTERNATIVES 39
6.C. TRADING VOLUME ISSUES 40
6.D. SUMMARY AND CONCLUSIONS 45
7. SUMMARY AND RECOMMENDATIONS 67
7.A. CONTINUE TO USE THE BLACK SCHOLES FORMULA FOR PRICE VOLATILITY ESTIMATION 67
7.B. CONTINUE TO UTILIZE A PUBLICLY AVAILABLE AND EXTERNAL MEASURE OF THE MARKET PRICE VOLATILITY 67
7.C. CONTINUE TO USE THE UNDERLYING FUTURES PRICE AS A FORECAST OF FUTURE REALIZED PRICE 68
7.D. AVOID USING THINLY TRADED OPTIONS PRICES IN COMPUTING THE IMPLIED PRICE VOLATILITY 68
7.E. REVIEW AND UPDATE PRICE CORRELATIONS 68
7.F. WE RECOMMEND A REVISED FORMULA FOR PRICE VARIABILITY IN RATE SIMULATION 69
8. REFERENCES 70
APPENDIX A 76

Executive Summary

USDA/RMA tasked Sumaria Systems, Inc. research team to conduct an actuarial review of the price volatility methodology used in the development of premium rates for crop revenue insurance programs. Because revenue insurance protects against price risk, accurate estimates of the price risk component is fundamental to actuarially fair rates.

The price volatility factors used by the Risk Management Agency (RMA) are currently based on the average implied volatilities for close-to-the-money option contract puts and calls during the last five trading days of the Projected Price-monitoring period for the given commodity (as determined by Barchart.com). These data observations have the merit that they are from economic transactions with real financial outcomes. Given these transaction prices and the known characteristics of the contract one can infer the price volatility implied by the contract. Various techniques have been used to derive the implied price volatility, but the Black-Scholes model (BSM) formula dominates in applied use.

We begin our analysis by reviewing the literature related to forecasting volatility in financial and commodity markets. Overall, it seems that the BSM model is still considered the "cornerstone" option pricing model due to its ease of use and simplicity, and that it can effectively be used for calculating implied volatility (as a forecast of future volatility). However, the literature also recognizes that implied volatility from the BSM has shortcomings and it is sometimes inconsistent with price/volatility behavior observed in the market. This is the reason numerous studies have developed alternative option pricing models and model-free approaches to estimate implied volatility. Nevertheless, there is still mixed evidence with regards to the BSM's biasedness, predictive accuracy, and whether or not the BSM is better than ARCH- or GARCH-type forecasts (or alternative implied volatility calculation approaches). For agricultural commodities, the evidence is also mixed – some studies show that for a particular commodity implied volatility is biased while others do not. However, most agricultural commodity studies indicate that implied volatility forecasts tend to encompass information embodied in backward-looking time-series models and, hence, have better forecasting performance.

We obtained detailed data that included options prices and data on volume of trades at various strike. This analysis compared the BSM approach to several other alternative methods. This analysis examined volatility estimates and predictive accuracy for several crops for which

RMA offers revenue insurance. Ultimately, the results demonstrate that the BSM approach performs well as a predictor of future volatility. Further, the important role that the BSM plays in markets and the fact that it is transparent and is obtained from external sources offers important advantages from a public policy point of view. The differences between the BSM and other measures of volatility are modest and are likely to be minor relative to the uncertainty associated with other important rating factors.

We also conducted a review of the mathematical calculations used to translate the price volatility factor within the revenue rate simulation. We show that there is a mathematical inaccuracy in the transformation used. However, our empirical simulations suggest the magnitude of the error is not large.

Ultimately we make the following recommendations to RMA

- Continue to use the Black-Scholes formula for price volatility estimation
- Continue to utilize a publicly available and external source of market price volatility
- Continue to use the underlying futures price as a forecast of future realized price
- Avoid using thinly traded options prices in computing the implied price volatility
- Review and update Price/Yield correlations used in rating
- Revise the formula for price variability in rate simulation to make it mathematically accurate

1. Introduction and background

Price volatility factors are necessary to the development of premium rates for crop insurance programs offering revenue coverage. Because revenue insurance protects against price risk, accurate estimates of the price risk component is fundamental to actuarially fair rates. Ultimately, yield risk, price risk, and the correlation between price and yield must be estimated for these products. Based on 2013 Summary of Business data, revenue products that use these price volatility factors comprise 80.6% of all RMA premiums and 74.3% of liability for the program. It is important to note that RP and RP-HPE products use the same futures price volatility for all crop policies with the same sales closing date. Thus, billions of dollars of premium are affected by a single parameter estimate. This is in contrast to the yield risk component of revenue rates where the parameters are driven by local data and parameter estimates have local implications.

The price volatility factors used by the Risk Management Agency (RMA) are currently based on the average implied volatilities for close-to-the-money option contract puts and calls during the last five trading days of the Projected Price-monitoring period for the given commodity (as determined by Barchart.com). Futures options provide price risk protection in an exchange-traded market where traders take positions to hedge or speculate on the price of options contracts. Options contracts tied to underlying futures markets are defined with specific quantity, strike price, time period, and delivery points. The one negotiated aspect of the contract is the price (premium) for the option contract. What is observed in the option market is agreed upon prices between buyers and sellers for contracts with specific attributes. These data observations have the merit that they are from economic transactions with real financial outcomes. This adds credibility to the estimates. Further, many market participants may be in these markets, thus the equilibrium price reflects the information and beliefs of many firms.

Given these transaction prices and the known characteristics of the contract one can infer the price risk volatility implied from the transaction price of the contract. Various techniques have been used to derive the implied price volatility, but the Black-Scholes formula dominates in applied use.

Once a price volatility estimate is obtained from the futures market, that parameter feeds into a revenue simulation that incorporates yield deviations consistent with the underlying yield insurance rates and allows for correlation to exist between price and yield. Once the parameters of

the stochastic simulation are specified, a rate simulation is conducted which models the expected indemnity that would occur given revenue protection (RP), revenue protection with harvest price exclusion (RP-HPE), and yield protection (YP). The difference in premium rates are then reported as the rate adjustment required if RP or RP-HPE are selected rather than YP coverage.

Given this context, the focus of this review is the data and methodology used to estimate the implied volatility and then incorporate it into the RP and RP-HPE rating. The remainder of this report is organized as follows:

- Review existing literature regarding implied volatility determination
- Analyze the data and assumptions and assess the adequacy of RMA's existing method for establishing price volatility factors for COMBO products and use of implied volatilities in establishing premium rates
- Compare and contrast alternative methods for calculating implied volatility to the method used by RMA
- Review the use of the price volatility factor within the revenue rate simulation model for COMBO products, including an evaluation of the underlying price/yield correlations assumed and the interacting effect price volatility and price/yield correlation have on revenue rates
- Summary of findings and recommendations

.

2. Literature review: Implied volatility in crop insurance rating

In this section, we discuss the literature related to the calculation of implied volatility from options markets with particular focus on issues that may affect use of implied volatility in rating revenue insurance products. As discussed in the introduction, price volatility factors that are used in rating crop revenue insurance are based on average implied volatilities for close-to-the-money puts and calls during the last five trading days of the price discovery period for the insured commodity. The daily implied volatilities used in the calculation of the factor are derived based on the well-known Black-Scholes model (BSM) for options pricing and are taken from the Barchart.com website.

2.a. Black-Scholes Model, volatility smiles, and bias

Forecasting volatility of commodity prices is critically important in rating crop revenue insurance because of the need to capture the price risk covered in this type of policy. Future volatility is commonly estimated either by using a backward- or a forward-looking approach.

Backward-looking methods develop volatility forecasts using time-series statistical/econometric methods, like calculating the standard deviation of an asset's return and ARCH- (or GARCH)-type models. The increased popularity of backward-looking techniques that use historical data has generally been traced to the introduction of and subsequent advances in ARCH and/or GARCH time-series models (Engle, 1982; Bollerslev, 1986). Most empirical studies, primarily of financial markets, tend to confirm that these time-series models provide good predictions of short-term volatility (Anderson and Bollerslev, 1998; Poon and Granger, 2003). However, several studies have shown that ARCH and GARCH models do not perform as well for longer-term volatility predictions since forecasts from these models revert to the unconditional mean. Day and Lewis (1993) and Holt and Moschini (1992) find that ARCH- and GARCH-type models provide poor predictions of long-term volatility of crude oil futures and real hog prices, respectively. Christoffersen and Diebold (2000) show that if the interest is in volatility forecasts for intermediate and long-term horizons (i.e., beyond 10 to 20 days), ARCH- and GARCH-based models may have poor predictive power.

Forward-looking methods to estimate future volatility are typically based on a particular option pricing model and the estimate from this type of method is called the implied volatility. The most common (and considered the "cornerstone") option pricing formula used for calculating implied volatilities is the Black-Scholes model (BSM) (Black and Scholes, 1973; Black, 1976; Merton, 1973). With BSM, implied volatility is computed by inverting the BSM option pricing formula such that the current market price is equal to the calculated option price for given values of the other variables in the model (e.g., strike price, time to maturity, risk free interest rate). In this framework, implied volatility at time *t* represents a forecast of variability and is interpreted as the market's expectation of volatility over the option's maturity (from *t* to maturity at *T*).

According to theory, markets are efficient with respect to widely available information so that if implied volatility is the market expectation of future volatility it should be an unbiased and well-informed estimate that incorporates all of the information that can be obtained from observed past price behavior, as well as all other public information. In addition, the market will have access to other historical information, from returns in other markets, past news events, and so forth, as well as knowledge and expectations about current market conditions and anticipated future events (e.g., Federal Reserve policy, national and international economic and financial conditions, etc.). In other words, the volatility parameter implied by an option's current market price in an efficient market should accurately reflect all relevant past and future information (i.e., which is why it is a "forward-looking" estimate). In that case, once implied volatility is known, any volatility estimate based on past prices alone should be redundant. This is the reason why implied volatility is generally considered by both academics and practitioners to be superior to alternative volatility forecasts (Figlewski, 1997).

Even with the popularity of the forward-looking implied volatility from BSM among practitioners, concerns over the predictive accuracy of this approach have appeared over time. These concerns typically arise from questions about the validity of some of the inherent assumptions embedded within the BSM formula. For example, based on the BSM, volatility should be constant across moneyness (or strike prices) and time to maturity of the option. However, in numerous empirical studies, implied volatilities show different non-constant patterns across moneyness and time to maturity.

A well-known pattern is the "volatility smile", where implied volatility is non-constant across strike prices (or moneyness). In particular, the volatility smile refers to a phenomenon where the implied volatilities of at-the-

money options tend to be lower than those of in-the-money or out-of-the-money options (i.e., exhibiting a U- or smile-shaped pattern where implied volatilities become progressively higher as an option moves in-the-money (at lower strike prices) or out-of the money (at higher strike prices)). Foreign currency options exhibit a symmetric smile shape pattern over different strike prices (Hull, 2009). Another pattern is the volatility "skew" where implied volatility is downward sloping as strike price increases, which has been observed in post 1987 S&P futures options (see Rubinstein, 1994; Hull, 2009). A volatility "sneer" is also possible, which is the reverse of the smile pattern (i.e., in-the-money and out-of-the-money options have lower implied volatility than at-the-money). Regardless of the pattern of implied volatilities across moneyness, these non-constant patterns are still typically referred to as the volatility smile phenomenon. Similar to volatility smile, the term used to describe non-constant implied volatility over the option's time to maturity is typically called volatility term structure (see Hull, 2009).

One common explanation for the observed volatility smile is violation of the log-normality assumption inherent in BSM. For example, Hull (2009) has shown that foreign currency options have fatter tails than a log-normal distribution and this may have caused the volatility smile observed in this market. In addition to fatter tails, there are other observed features of options prices not accounted for in BSM that have been examined in the literature and pointed out as possible factors that cause inaccuracy in the implied volatility estimates. Some of the observed features that have been explored include stochastic (i.e. time varying) volatilities, discrete price jumps, measurement errors, and market microstructure features (i.e., non-frictionless markets, transactions costs, volatility risk premium).

Given that these observed features are not directly accounted for in BSM, one strand of the literature has focused on developing alternative option pricing models that relax the assumptions of the BSM and, consequently, account for or explain the observed implied volatility patterns. Alternative pricing models that have less stringent assumptions are expected to produce better implied volatility forecasts. For example, Hull and White (1987) and Heston (1993) developed stochastic volatility models that relax the constant volatility assumption in the BSM. Bates (1996) developed alternative pricing models that allow for jump processes and stochastic volatility. Bakshi, Cao, and Chen (1997) considered a comprehensive pricing framework that can accommodate stochastic volatility, stochastic interest rates, and jumps. Several other alternative models extend the BSM to account for trading/transactions costs and other market friction elements (see Leland, 1985; Boyle and Vorst, 1992). Previous studies have

shown that some of these alternative pricing models can successfully explain the volatility smile phenomenon (Dumas, Fleming and Whaley, 1998).

Another body of literature that grew as a response to the limitations of the BSM is the estimation of model-free implied volatility (MFIV) measures (Britten-Jones and Neuberger, 2000). MFIV incorporates option prices that span the full spectrum of exercise prices, but it does not depend on a particular pricing model and it has been shown to be robust to any underlying data generating process (Jiang and Tian, 2000; Carr and Wu, 2009). However, empirical research on the relative performance of MFIV has been limited and the existing evidence as to whether MFIV provides better volatility forecasts than BSM (or time series measures) has been mixed (Jiang and Tian, 2005; Andersen and Bondarenko, 2007; Taylor et al., 2010; Tsiarias, 2010; Wang and Fausti, 2011; Cheng and Fung, 2012).

Even with the growing number of alternative pricing models and model-free implied volatility approaches, the BSM still remains the cornerstone pricing model used by practitioners to estimate implied volatility. Barr (2009) argues that the BSM's ease of use, speed, and simplicity make it more attractive to practitioners, even though the inherent assumptions in the model are not consistent with observed features (like the volatility smiles). Barr (2009) also points out that most alternative pricing models require the use of Monte Carlo techniques wherein the parameters are still commonly calibrated based on implied volatility estimates from the BSM.

Since the BSM remains the predominant pricing model used in practice, another large strand of literature focuses on the bias and/or informational content of implied volatility estimates derived from the BSM. Studies in this area tend to focus on the ability of implied volatilities to predict future realized volatility. For example, since the BSM was developed to price European options on futures contracts (Black, 1976), there is concern that its use in pricing American type options generates upward bias in the implied volatility estimates. This potential bias has been found to be small for short-term options that are at-the-money (Whaley, 1986; Shastri and Tandon, 1986). Moreover, studies examining implied volatility estimation procedures that utilize weighting schemes (i.e., calculating implied volatility as the average implied volatility across strike prices) suggest that implied volatilities taken from nearest at-the-money options provide the most accurate volatility estimates (Beckers, 1981; Mayhew, 1995). At- or near-the-money options tend to contain the most information regarding future volatility because they are usually the most traded option (i.e., highest volume) and produce the largest vega (i.e., the rate of change in the options price due to changes in the volatility) (Mayhew, 1995). In

addition, Jorion (1995) indicates that the averaging of implied volatilities from both puts and calls can help reduce measurement errors, which has been noted as a possible source of volatility smiles.

In looking at the bias and informational content of implied volatilities calculated from the BSM, there is mixed evidence as to whether these volatilities predict future realized volatility well. Figlewski (1997) argues that there is ample evidence that implied volatility is a biased estimate of future volatility and does not impound all information provided by alternative forecasts (typically from backward-looking time series models). For example, studies by Day and Lewis (1992), Lamoureux and Lastrapes (1993), and Canina and Figlewski (1993), all of which studied either options on individual stocks or S&P 100 options, generally find that implied volatility is a poor forecast of the subsequent realized volatility over the remaining life of the option. Canina and Figlewski (1993) is the most extreme, suggesting that implied volatility forecasts are biased and have no statistically significant predictive power to forecast realized volatility. In contrast, Christensen and Prabhala (1998) find that implied volatility is a good predictor of realized volatility and subsumes information content of historical volatility for monthly, non-overlapping S&P 100 index options data. Using data from 35 futures options markets from eight exchanges, Szakmary et al. (2003) also found that, in most markets, implied volatility is a good predictor of realized volatility and that backward-looking time-series models contain no information that is not already embedded in the implied volatility forecast. On the other hand, another set of studies like Jorion (1995) and Fleming (1998) indicate that implied volatilities are biased but still have predictive power (i.e., outperforming backward-looking volatility measures). Chan, Cheng, and Fung (2010) also find that implied volatility forecasts outperform time-series forecasts, although the informational content of implied volatility depends on the realized volatility measure it is being compared against.

More recent studies on implied volatility tend to accept that the BSM implied volatility is a biased measure of future volatility but, in general, conclude that it tends to outperform backward-looking historical forecasts from time-series models. Therefore, the focus of these recent studies is to find procedures that can correct the bias in the BSM implied volatility estimates. Barr (2009) examined data from 26 options on commodity futures markets (encompassing agricultural commodities, soft commodities, livestock, precious metals, and energy) and revealed that, for 19 of the 26 markets examined, implied volatility estimated from at-the-money options is an upward biased estimator of realized volatility. For out-of-the-money and in-the-money options, implied volatility was found

to be an upward biased estimator in *all* markets. Barr (2009) also examined the possible sources of this positive bias and found that the bias is roughly equivalent to the transactions costs of option writers (e.g. commission charges). Hence, Barr (2009) suggests that people who want to use implied volatility as a forecast of realized volatility should first subtract the average bias (e.g., the average transaction fee for a round trip option purchase) from the actual option price before solving for the implied volatility.

Another study by Wu and Guan (2011) argues that a significant portion of the bias in implied volatility is not accounting for the volatility risk premium. Hence, they provide an adjustment to implied volatility that accounts for the volatility risk premium and indicate that this approach has better predictive power than backward-looking procedures when applied to corn futures. On the other hand, Xu (2012) suggests that the main source of bias in implied volatility forecasts is measurement error, and proposes an alternative implied volatility estimator that first nonparametrically smooths the option price function before inverting to get an implied volatility estimate.

Recent studies have also examined the information content and forecasting performance of MFIV measures vis-à-vis the implied volatility estimates from the BSM. The evidence is quite mixed. Jiang and Tian (2005), using S&P 500 index option data, strongly find that volatility forecasts from MFIV outperform both the volatility estimates from the BSM and time-series approaches. On the other hand, Andersen and Bondarenko (2007), using data from futures options, find that implied volatility from the BSM is a more informative measure of future volatility than MFIV. Cheng and Fung (2012), as well as Taylor et al. (2010), find similar results as Andersen and Bondarenko (2007). Tsiarias (2010), in contrast, did not find a clear winner between the BSM and MFIV. But note that Andersen and Bondarenko (2007), as well as Tsiarias (2010), find that a "model-free" corridor implied volatility (CIV) measure (i.e., an MFIV measure that truncates the tails of the return distribution) performs about the same as an implied volatility from the BSM and outperforms full MFIV estimates. These studies also suggest that the width of the corridor plays a crucial role in the forecasting performance of the CIV measure.

2.b. Implied volatility in agricultural markets

Most of the studies reviewed above are "general" studies that encompass financial (i.e., equity markets, stocks) and/or commodity markets. In this sub-section, we specifically review implied volatility studies that empirically focus on agricultural commodity markets.

One of the earliest studies of implied volatility in agricultural markets was by Wilson and Fung (1990) who investigated the informational content of implied volatility for corn, soybeans, and wheat futures. Results from Wilson and Fung (1990) found mixed results, implied volatility in corn and soybean markets correlated well with realized volatility but not in the wheat market. Simon (2002) also examined the predictive accuracy of implied volatilities (vis-à-vis a seasonal GARCH model) in the corn, soybean, and wheat futures markets. Using the Black (1976) model to calculate implied volatility over a 4 week horizon, Simon (2002) found that implied volatility estimates for soybeans and wheat were unbiased, and encompassed the forecasts from the seasonal GARCH models. However, for corn, the implied volatility estimate was biased; although it still encompassed the information from the GARCH model. Using daily futures contracts data for cocoa, coffee, and sugar, Giot (2003) investigated whether lagged implied volatility forecasts have superior informational content as compared to GARCH procedures. Results from Giot (2003) indicated that lagged implied volatility reflects all available information in the cocoa market, but that GARCH estimates in the coffee and sugar markets marginally improve the information content from the lagged implied volatility estimates.

As mentioned in the previous sub-section, the study by Szakmary et al. (2003) comprehensively examined predictive accuracy of implied volatility forecasts (for up to 70 trading days) in 35 futures options markets, which encompassed equity, interest rate, currency, energy, metals, agriculture, and livestock markets. For all 13 agricultural and livestock commodities examined, Szakmary et al. (2003) found that implied volatility forecasts are biased and, except for sugar, have more explanatory power than historical volatility estimates. In addition, GARCH forecasts in most agricultural markets do not add additional information beyond what is already embodied in the implied volatility estimates (i.e., the exceptions are in the soybean meal, sugar, feeder cattle, live cattle, and lean hog markets). In contrast to the results of Szakmary et al. (2003) for live cattle markets, Manfredo and Sanders (2004) found that implied volatility estimates still encompass all information provided by a time-series alternative (i.e., GARCH) even though they are biased and inefficient.

Using daily live and feeder cattle data from 1984 to 2009, Brittain, Garcia, and Irwin (2011) found results similar to Manfredo and Sanders (2004) – that is, in live and feeder cattle markets implied volatilities were upwardly biased and inefficient in both markets, but implied volatility forecasts still encompass GARCH forecasts in both markets. Manfredo, Leuthold, and Irwin (2001) also examined implied volatility performance in fed cattle, feeder cattle, and corn markets but focused on its ability to forecast cash price volatility rather than the realized volatility of their futures prices. Their main finding was that no single method of volatility forecasting (i.e., implied volatility, time series, or a composite approach) provided superior accuracy across alternative data sets and horizons. Although composite forecasting methods that combine implied volatility and time-series forecasts tend to provide improved volatility forecasts for almost all horizons examined.

Results from most of the studies reviewed above led Garcia and Leuthold (2004, p. 252) to conclude that "implied volatilities provide reasonable forecasts of nearby price variability." They also note that implied volatilities are often biased, but nevertheless appear to embody information in the market. Garcia and Leuthold (2004, p. 259-260) then went on to suggest that further research "is warranted to determine thoroughly the characteristics and magnitude of the bias, its sources, and its economic implications for decision-makers. It also seems useful to explore forecasting of volatility for distant horizons, as decision-makers, particularly in agricultural markets, need this information."

Another recent area examined in the literature is with regards to the forecasting performance of implied forward volatility (rather than implied volatility per se). Egelkraut, Garcia, and Sherrick (2007) define implied forward volatility as the volatility forecast generated from two options with consecutive maturities, and represent the expected average volatility for the non-overlapping future time interval between their expiration dates. Given the ability of implied forward volatility to forecast particular intervals within a corn crop's growing and non-growing seasons, Egelkraut, Garcia, and Sherrick (2007) developed a flexible procedure to calculate the term structure of implied forward volatility (i.e., the changing pattern of implied volatilities over some period; see Ferris, Guo, and Su, 2003) and compare its performance with historical forecast measures (i.e., three-year moving average of past realized volatility and a year-lagged realized volatility). Using corn futures market data, results of their analysis suggest that implied forward volatilities anticipate realized volatility well over various time horizons. When forecasting for nearby (i.e. short-term) intervals, the implied forward volatilities provide unbiased forecasts and are superior to

forecasts based on historical volatilities. For more distant intervals, early year corn options predict the direction and magnitude of future volatility changes as well as or better than the alternative historical volatility forecasts.

Egelkraut and Garcia (2006) build on their other study by investigating the performance of implied forward volatility over a wider array of agricultural commodities (e.g., corn, soybeans, soybean meal, wheat, and hogs). In general, Egelkraut and Garcia (2006) find that the implied forward volatility dominates forecasts based on historical volatility information, but that predictive accuracy is affected by the commodity's characteristics. Due to the fairly well-established volatility patterns in corn and soybean markets, the implied forward volatilities in these markets were found to be unbiased and efficient. For soybean meal, wheat, and hogs, volatility is less predictable, and the implied forward volatilities in these markets are biased.

The topic of volatility smiles has also been examined in agricultural markets. Guo and Su (2004) examined corn futures options data from 1991 to 2000 and found evidence of the presence of a volatility smile in this market. That is, as one moves farther away-from-the-money, implied volatility increases monotonically. Guo and Su (2004) also indicate that implied volatility in the corn market decreases as time to maturity increases.

Barr (2009) examined options for 19 agricultural and livestock commodities (e.g., corn, cotton, spring wheat, oats, rice, soybeans, soybean meal, wheat no. 2, barley, flaxseed, lumber, cocoa, milk, orange juice, coffee, white sugar, raw sugar, feeder cattle, and live cattle) and investigated the existence of bias and volatility smiles in these markets. In these agriculture-related markets, Barr (2009) indicates that implied volatility from at-the-money options is an upward biased estimator in 14 out of the 19 markets (i.e., the non-biased markets are cotton, wheat no. 2, oats, cocoa, and orange juice). For out-of-the-money and in-the-money options, implied volatility is an upward biased measure in all agriculture-related markets. Moreover, Barr (2009) finds that some degree of volatility smile is observed in all agriculture-related markets. The exceptions are in cotton, barley, feeder cattle, and live cattle where volatility skew is more evident (i.e., downward sloping across strikes).

Recent studies by Wang, Fausti, and Qasmi (2011) and Wu and Guan (2011) developed new implied volatility calculation procedures and compared the performance of these estimators to traditional volatility

measures. Wang, Fausti, and Qasmi (2011) developed a new implied volatility measure based on what they call a "model free variance swap approach" that is akin to the model-free VIX volatility measure for the S&P 500 index (i.e., the same concept as the MFIV). Comparing this new model-free measure to the traditional implied volatility measure derived from Black (1976) and a GARCH model, Wang, Fausti, and Qasmi (2011) conclude that their new measure provides better forecasts of realized corn futures volatility in the sense that it encompasses more information and generates less forecasting error than the other alternatives. They also find that implied volatilities from their approach and the Black (1976) model tend to have an upward bias (relative to realized volatility measures) and they are time-varying.

Wu and Guan (2011) also developed a new approach to calculating an implied volatility measure. As mentioned in the previous sub-section, they provide an adjustment to implied volatility that accounts for the volatility risk premium and indicate that this approach has better predictive power than backward-looking procedures (e.g., three-year moving average of realized volatilities and one-year lagged realized volatility) when applied to corn futures.

2.c. Implied volatility and crop insurance

The role of implied volatility in the premium rate calculations for revenue coverage under the Common Crop Insurance Policy (i.e., COMBO Policy) is discussed in detail in RMA (2009). Essentially, an implied volatility estimate is used to characterize the variability of the price distribution employed in the simulation that determines the revenue add-on of the premium rate (i.e., the rate added to the yield protection policy premium to account for price risk in the revenue coverage). Based on its importance in the rating process, Bulut, Schnapp, and Collins (2011) carefully assessed how it is currently used in crop insurance rating and evaluated whether its use in the rate-making process makes sense. They identified four major issues.

First, Bulut, Schnapp, and Collins (2011) point out that implied volatility from the BSM are assumed to be constant and point out that observed price volatility tends to vary over time. They suggest considering GARCH type models to account for time-varying volatilities. Second, they indicate that the RMA approach of only averaging implied volatilities over the last five days of the discovery period ignores other implied volatility information available prior to this date (i.e., for example, implied volatility estimates in the month prior to the last five days of the discovery period).

Related to this issue, Bulut, Schnapp, and Collins (2011) also discuss how this procedure for calculating the volatility factor adversely affects the ability of insurance agents to provide accurate quotes to customers in a timely manner.

The third issue identified in Bulut, Schnapp, and Collins (2011) is that the revenue protection policy is essentially a yield-adjusted Asian (YAA) put option (as described in Barnaby, 2011) and the payoff depends on the average of futures prices in the harvest price discovery period. They point out that this is inconsistent with options traded on the Chicago Board of Trade (CBOT) which have payoff that depends on the price at the time of sale (i.e., spot price) and is the type of option used in determining implied volatility. Lastly, Bulut, Schnapp, and Collins (2011) noted that the sensitivity (elasticity) of premiums with respect to changes in volatility needs to be investigated further and their preliminary results suggest that in volatility ranges below (45%), which is where volatilities range from 2006 to 2011, premium rates tend to be very sensitive to changes in implied volatility. This is consistent with Barnaby (2013a, 2013b) who points out that the implied volatility has a major impact on premiums and it is likely the main factor that drives revenue insurance premiums, rather than the price level.

Two studies by Bozic et al. (2012a, 2012b) examined the role of implied volatility in Livestock Gross Margin Insurance for dairy cattle (LGM-Dairy). Note that the COMBO rating methodology is partly adapted from LGM rating methods (as well as previous revenue policies – Revenue Assurance (RA) and Crop Revenue Coverage (CRC)) (See RMA, 2009). In Bozic et al. (2012a), a parametric bootstrap procedure is developed to test whether implied volatilities for Class III milk, corn, and soybean meal futures are unbiased. Bozic et al. (2012a) found that implied volatilities for corn and soybean meal are unbiased predictors of end-of-term volatility, but implied volatility for Class III milk is biased downward. When accounting for the bias is Class III milk futures in LGM-Dairy rating, Bozic et al. (2012a) revealed through simulations that LGM-Dairy premiums will likely increase from 3% to 21%. With these estimates, they conclude that implied volatility biases in LGM-Dairy rating do not produce excessive premiums.

In a related paper, Bozic et al. (2012b) explored how volatility smiles (and skews) in Class III milk, corn, and soybean meal futures prices affect LGM-Dairy premium rates. In particular, since skewness and kurtosis of price distributions likely cause the volatility smiles and skews, Bozic et al. (2012b) investigated how changes in skewness and kurtosis (i.e., to better reflect the observed volatility smiles and skews in the data) influence the

premium rates in LGM-Dairy. In simulation models where additional skewness and kurtosis were added (i.e., above those for log-normality) for corn and soybean meal prices, Bozic et al. (2012b) found no effect of financial importance in the LGM-Dairy rates (i.e., changes were less than 4%). However, in scenarios where they only altered the skewness and kurtosis of corn price distributions by 50% above log-normal (i.e., assuming milk and soybean meal prices are known with certainty), they found premium rate increases up to 30%. Bozic et al (2012b) suggest that the "basket" nature of LGM-Dairy (i.e., with multiple price risks) may have tempered the effects of volatility smiles in the individual price distributions.

2.d. Summary

There is a rich literature related to forecasting volatility in financial and commodity markets. Overall, it seems that the BSM model is still considered the "cornerstone" option pricing model due to its ease of use and simplicity, and that it can effectively be used for calculating implied volatility (as a forecast of future volatility).

However, the literature also recognizes that implied volatility from the BSM has shortcomings and it is sometimes inconsistent with price/volatility behavior observed in the market. This is the reason numerous studies have developed alternative option pricing models and model-free approaches to estimate implied volatility. Nevertheless, there is still mixed evidence with regards to the BSMs biasedness, predictive accuracy, and whether or not the BSM is better than ARCH- or GARCH-type forecasts (or alternative implied volatility calculation approaches). For agricultural commodities, the evidence is also mixed – some studies show that for a particular commodity implied volatility is biased while others do not. However, most agricultural commodity studies indicate that implied volatility forecasts tend to encompass information embodied in backward-looking time-series models and, hence, have better forecasting performance.

Only a few studies have examined the role of implied volatility in crop insurance. And results from these studies seem to indicate that premium rates are sensitive to implied volatility estimates (except for one scenario with LGM-Dairy). But most of these studies suggest that further research needs to be undertaken to fully understand how sensitive premium rates are to implied volatility changes and whether there are other approaches that can improve implied volatility calculation procedures in crop insurance.

3. Data, assumptions, and adequacy of RMA's existing method for establishing price volatility factors for COMBO products

The COMBO rating method evolved from the rating of two previous insurance products – Revenue Assurance (RA) and Crop Revenue Coverage (CRC), both of which were introduced in the 1990s. The COMBO rating procedure probably remains closest to and follows from RA rating in the sense that rates are derived from parametric yield distributions calibrated to the loss cost based APH rates. COMBO rating also follows the RA approach in assuming that the price distribution is log-normal and its second moment can be computed based on an options-based volatility measure. The parameters of the yield and price distributions, together with an assumed yield-price correlation, are then used in a simulation procedure to calculate a revenue rate at various coverage levels. Given these simulations, a "revenue load" is then calculated by taking the difference between the simulated revenue rate and a corresponding simulated yield rate (for yield insurance coverage). This revenue load becomes an additive factor that is charged to an insured who chooses revenue coverage under the COMBO policy.

The particular futures contract and time periods used for price discovery and price volatility discovery depends on sales closing date. For example, the Projected Price for Missouri soybeans with a sales closing of March 15 is determined using the harvest year's CBOT November soybean futures contract. Daily settlement prices for the month of February on the November soybean futures contract are averaged and this average February settlement price serves as the Projected Price for determining the amount of insurance coverage. For the harvest price, the November futures contract's daily settlement prices are averaged in October, which is the harvest price discovery month for Missouri soybeans. RMA derives a measure of price volatility based on observed option contract prices for up to four in-the-money strike prices (2 put and 2 call options) by using the BSM.

The Black-Scholes option pricing model was developed under the aforementioned assumptions, including that asset prices are log-normally distributed. Premiums of European call and put options are expressed as

$$V_C = \delta_t \int_0^\infty \max(0, F_t - S)\varphi(\theta, F_t)dF_t \text{ , and}$$

$$V_p = \delta_t \int_0^{\infty} \max(0, \mathrm{S} - F_t)\varphi(\theta, F_t)dF_t \ ,$$

where V_C and V_P represent the value of calls and puts, δ_t is the relevant discount factor and $\varphi(\theta, F_t)$ represents the probability density function having parameters θ underlying the distribution of the futures price F_t with strike price *S*. The standard Black-Scholes specification assumes that $\varphi(.)$ is log-normal with a mean given by F_t and a variance that is represented by a transformation of the volatility parameter.[1] Thus, for any combination of futures and options quotes at a particular strike price, these expressions can be inverted to obtain a unique measure of the volatility. As noted, this volatility should be the same across all strike prices and option types traded at the same time if the assumptions underlying the pricing model are correct. Again, as noted earlier, empirical experience has shown that these volatilities may increase significantly as the distance between the strike price *S* and the futures price F_t increases—the aforementioned smile and smirk feature.

It has been widely demonstrated that the call option valuation equation can be inverted to solve for the implied volatilities using the following decompositions:

$$V_c = e^{-rt}(F_t N(d_1) - SN(d_2)), \text{ where}$$

$$d_1 = (\ln\left(\frac{F_t}{S}\right) + .5\sigma^2 t) / \sigma\sqrt{t}$$

$$d_2 = d_1 - \sigma\sqrt{t} \ ,$$

and *N(.)* is the standard normal cdf, *t* is the term of the option (year or fraction of year before expiration), and *r* is a constant risk free interest rate. An analogous expression exists for pricing put calls in terms of the same variables. Note that this assumes a log-normal distribution on prices and no-arbitrage conditions (which implies F_t is the mean of expected prices). This can be calibrated over a range of concurrent options quotes to obtain a measure of the implied volatility that uses information across all of the strike prices on a given day. This may allow for a more flexible and robust measure of the volatility that uses all available options while maintaining the assumption that the futures price is an unbiased expectation of the future spot price.

[1] As Black (page 174, 1976) notes, "the mean of the possible spot prices at time *t** (at expiration of the futures contract) will be the current futures price.

A drawback of the Black-Scholes formula for estimating implied volatility is a lack of a closed form solution. Typically a Black-Scholes implied volatility is determined through an iterative process that equates the market observed option premium to the imbedded variables which are underlying futures price, strike price, time to expiration and interest rates. Implied volatilities are provided by numerous financial reporting services which use the BSM or some other computational methods to estimate the implied volatility. RMA has for some years obtained data from barchart.com. An example of the Barchart data is given in Table 3.1.

Table 3.1 Example Barchart Data

Symbol	Date	Open	High	Low	Settle	Volume	Open Interest	Implied Volatility
CZ12	2/23/2012	5.625	5.63	5.555	5.5875	39201	269587	28.2
CZ12	2/24/2012	5.54	5.58	5.5	5.58	39629	269726	27.7
CZ12	2/27/2012	5.5425	5.58	5.5425	5.57	34730	270173	28.2
CZ12	2/28/2012	5.575	5.635	5.5725	5.635	34734	269342	27.7
CZ12	2/29/2012	5.645	5.695	5.645	5.685	37299	271305	28.1

The closing implied volatility for the contract for a particular crop/location is defined in the Commodity Exchange Price Provisions (CEPP) of the Common Crop Insurance Policy Basic Provisions (11-BR). One observation is obtained each trading day. As indicated earlier, the RMA volatility factor for a given crop is based on the average of the time-adjusted volatility factors for the last five days of the Projected Price discovery period. However, the implied volatility must be adjusted to take into account any differences between the expiration of the options contract and the time period RMA uses to establish the harvest price.

The exchange prices used are:

- corn, barley and grain sorghum use the corn CBOT contract
- soybeans (CBOT)
- rice (CBOT)
- wheat (CBOT, MGE, or KCBT depending on the location and type of wheat)
- cotton (ICE)

- canola/rapeseed (ICE)
- oil-type sunflower use Soybean oil futures (CBOT).

The steps for determining the volatility are then:

Step 1. Determine the Projected Price and Harvest Price monitoring periods from the CEPP.

Step 2. For each of the last five days of the Projected Price discovery period determine the number of days from that date until the midpoint of the Harvest Price discovery period (the 16th day of the Harvest Price discovery month), and divide by 365.

Step 3. Determine the square root of the value obtained in step 2.

Step 4. Multiply the value in step 3 by the implied volatility for the contract for the day.

Step 5. Calculate the simple average of the five values in step 4 and round to 2 decimals.

4. Review of procedures for determining and using implied volatilities

4.a. Use of the average implied volatility from the final five days of price discovery

RMA gives equal weights to the implied volatilities from the last five days of the price discovery period. By choosing to average five days RMA obtains some temporal smoothing of the volatility measure. Clearly, one could posit both longer and shorter time periods. A longer period would provide more smoothing and more days. However, market efficiency would suggest that the final day of the period would contain all relevant information available in previous days. Thus, one might also argue for using only the implied volatility for the final day of the price discovery period. Intuitively, some averaging of days avoids some anomalous market event influencing the implied volatility. For example there might be an extreme event leading to a reduced trading volume on a single day.

We recommend RMA's method continue to ignore the information on the futures contract prior to the last five days of the price discovery period. But also we are hesitant to recommend going to single day because little additional effort is required to collect four additional days of data. The cost of doing so seems minor relative to the problems that might arise from low trading volume in the options market on a single trading day.

4.b. Examination of the mechanism to simulate price volatility derived from option prices

The issue addressed in this section is related to the derivation of the premium rate for revenue coverage under the COMBO Policy and more specifically with the price simulation.

The goal of the RMA rating simulation is to obtain log-normally distributed prices. To achieve this:

1. Start with standard normal random draws p_i. So, $p_i \sim N(0,1)$
2. Transform the standard normal draws to a normal distribution with mean μ_p and standard deviation σ_p : $p_i^N = \mu_p + p_i\sigma_p$. So, $p_i^N \sim N(\mu_p, \sigma_p^2)$

3. Transform p_i^N using exponentiation to arrive at log-normal prices: $\tilde{p} = \exp(p_i^N)$. $\tilde{p}$ are then used in simulations in combinations with yield draws to derive the premium rate for revenue coverage.

The question is what are the values of μ_p and σ_p to be used in the above steps?

The "Cost Estimator Detailed Worksheet" for the "Cost Estimator" available at http://www.rma.usda.gov/ftp/Publications/M13_Handbook/2014/approved/P11_1_PLAN_01_02_03_PREMIUM_CALCULATION.PDF uses the following formulas:

$$\sigma_p = \ln(Volatility^2 + 1)$$
$$\mu_p = \ln(\bar{p}) - \frac{1}{2}\sigma_p.$$

In the "Cost Estimator" document, μ_p is referred to as "LnMean", σ_p is referred to as "LnVariance", and *Volatility* is referred to as "Price Volatility Factor". The document uses the symbol σ_p that, by convention, is used to denote the standard deviation, to instead denote the variance. We recommend that the symbol σ_p, in the context that is being used, be replaced with σ_p^2. The calculation of the *Volatility* ("Price Volatility Factor") is described earlier in this report and can be found at: http://www.rma.usda.gov/pubs/2011/volatilitymethodology.pdf.

This *Volatility* is simply the implied volatility obtained from the options market and adjusted for the length of the maturity period.

So the two yet undetermined parameters to use in the simulation of prices are μ_p and σ_p.
Next we provide the correct formulas for μ_p and σ_p and then provide an example. The correct formulas are:

$$\sigma_p = Volatility$$
$$\mu_p = \ln(\bar{p}) - \frac{1}{2}\sigma_p^2.$$

1. The formula for σ_p is simply: σ_p = *Volatility*.

2. The formula for μ_p is: $\mu_p = \ln(\bar{p}) - \frac{1}{2}\sigma_p^2$, where $\bar{p}$ is the expected price obtained from futures markets.

The formula for σ_p in the document is $\sigma_p = \ln(v^2 + 1)$, clearly different from the formula above.
The formula for μ_p is $\mu_p = \ln(\bar{p}) - \frac{1}{2}\sigma_p$, which leaves out the quadratic exponent for σ_p in the formula above. Following our recommendation to replace σ_p with σ_p^2 the two formulas would be the same.

<u>An example</u>

The goal is to simulate prices from a log-normal distribution $p \sim LN(m_p, s_p^2)$. We know that $\ln(p) \sim N(\mu_p, \sigma_p^2)$. We observe m_p from the futures markets, let's say \$5.00/bu. We also observe σ_p from the options markets, let's say 0.4 (this after the adjustments as described in http://www.rma.usda.gov/pubs/2011/volatilitymethodology.pdf).

From here we can calculate $\mu_p = \ln(p) - \frac{1}{2}\sigma_p^2$ = 1.5294. Now we can start the simulations. To achieve this:

1. Start with standard normal random draws z_i. So, $z_i \sim N(0, 1)$.
2. Transform the standard normal draws to a normal distribution with mean μ_p and standard deviation σ_p:
   $\ln(p) = \mu_p + z_i\sigma_p = 1.5294 + z_i * 0.4$. So, $\ln(p) \sim N(\mu_p, \sigma_p^2)$.
3. Transform $\ln(p)$ using exponentiation to arrive at log-normal prices: $p = \exp(lnp)$. These prices are then used in simulations in combination with yield draws to derive the premium rate for revenue coverage.

One of the issues with RMA's approach is the unnecessary transformation used to obtain σ_p^2. The RMA documents propose using

$\sqrt{\ln(0.4 \wedge 2 + 1)} = \sqrt{0.1482} = 0.3853$, clearly different from 0.4.

Further, using 0.3853 for σ_p^2 instead of 0.4 in the calculation of μ_p leads to μ_p=1.5352 > 1.5294. So, the effect is to increase the mean of the

distribution and reduce its variance. The combined effect on premium rates is addressed in the next section.

5. Review the use of the price volatility factor within the revenue rate simulation model underlying COMBO

The COMBO rating process has four key components: (1) calculating price-yield correlations, (2) estimation of the mean and standard deviations (i.e. the parameters) of the yield and price distributions, (3) generation of potentially correlated yield and price draws, and (4) simulating indemnities which allows calculating revenue premium rates.

According to the report by Coble et al. (2010), the yield-price correlations used in COMBO rating are calculated from historical yield and price data. Our understanding is that National Agriculture Statistics Service (NASS) county yield data are detrended using a linear trend and the yield deviates are calculated as the percentage deviation from trend. The futures price deviates are calculated as the percent change in price from the planting time expected price to the harvest price. Once the price and yield deviates have been calculated, the county-level, yield-price correlations are derived and then state-level, yield-price correlations are computed by taking the weighted average of the county-level correlations (i.e. weighted by production). The state-level correlations are then adjusted downward to more accurately reflect the yield-price correlation at the individual level. One limitation of the procedure used to calculate the yield-price correlations is that it imposes a constant yield-price correlation for all producers in the state. The current sets of correlations used by RMA are shown in figure 5.1. For, cotton rice, and canola, the rate simulations assume independence of price and yield risk based on the analysis of historical price and yield data. Negative correlations are found in some corn, soybean, and wheat producing states. The most extreme correlation estimate for corn occurs in the major Corn Belt states of the Midwest.

All else equal, as negative correlations increase in absolute value, RP and RP-HPE rates decrease. A couple of issues relate to the use of correlations in the rating procedure. First, how much spatial variability exists in the relations between price and yield? Figure 5.1 shows that correlation is held fixed for all counties in a state. However, Coble et al. (2010) discussed the potential for price-yield correlation to vary within a state due to differing production practices and spatial location. An illustration of this issue is found in a report by Lubben and Jansen (2010). A second issue is whether price yield relationships remain stable across time. For example do price-yield relationships change functionally when world

stocks vary or when widespread droughts occur? Some methods to model random variable relationships are more flexible than the Iman-Conover procedure (Zhu, Ghosh, and Goodwin, 2008). However, because we only observe one price-yield combination per year, the added flexibility of these procedures may result in spurious relationships being found in small samples. Our suggestion is that RMA consider updating price-yield correlations both spatially and using newer data. Further we suggest RMA follow the recent developments related to applying copulas for modeling revenue. However, we make no stronger recommendation at this time.

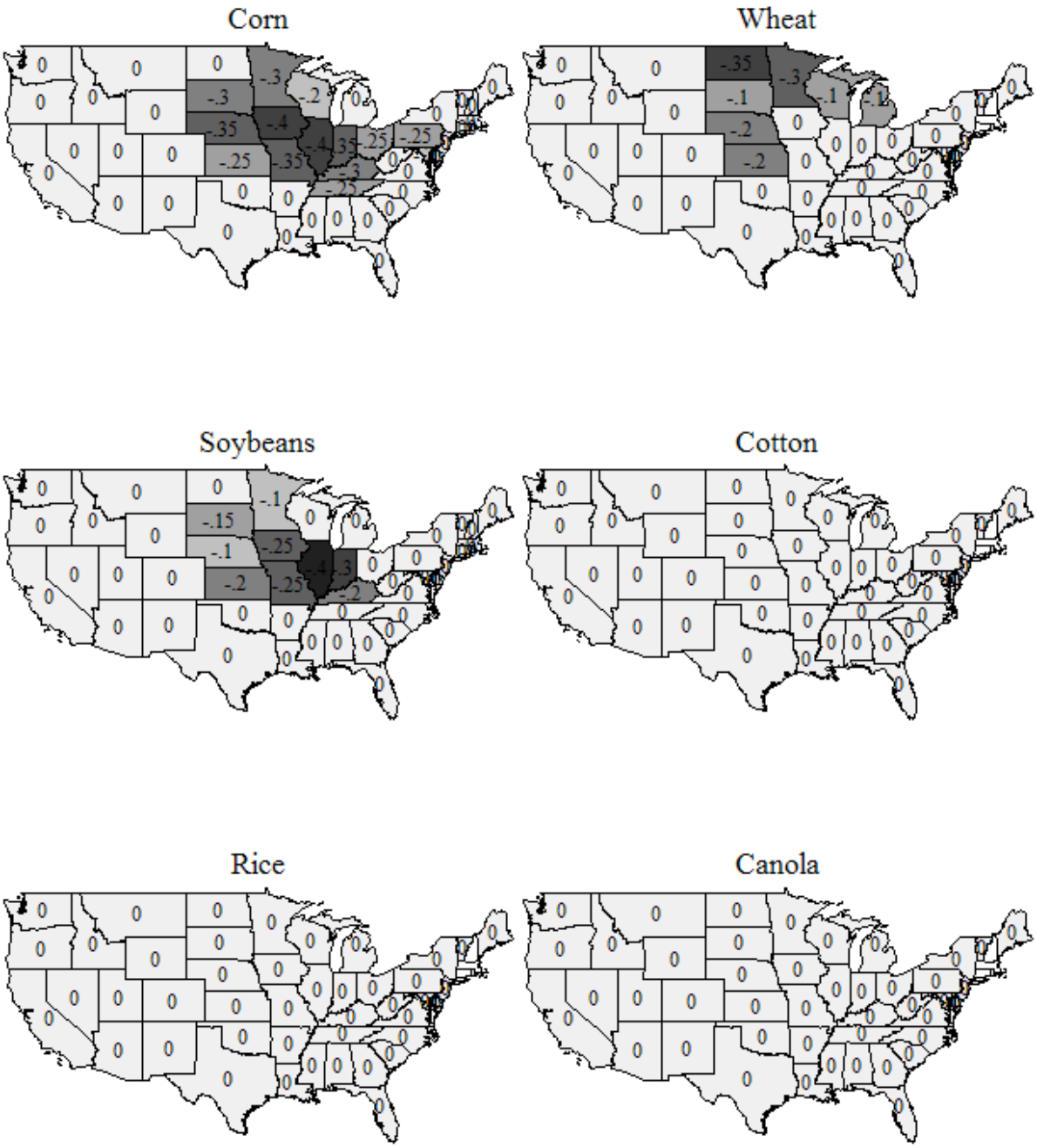


Figure 5.1 Current Correlations Used by RMA for Rating RP and RP-HPE

As mentioned above, the parameters of the price distribution (i.e. mean and standard deviation) are calculated primarily using the BSM volatility

measure and assuming that prices are log-normally distributed. The yield distribution in the COMBO rating method is assumed to follow a censored normal distribution. Coble et al. (2010) summarize the procedure to derive the parameters (i.e. the mean and standard deviation) of the censored normal yield distribution corresponding to the APH yield insurance rate at the 65% coverage level as follows:

(1) Normalize the APH yield to a value of 100 (i.e. μ=100)
(2) Select a target APH rate (i.e. the target rate)
(3) Find μ and σ that ensure that the following two equations hold:

$$100 = \frac{1}{5000}\sum_{i=1}^{5000}\max(y_i, 0)$$

$$\text{Target Rate} = \frac{\frac{1}{5000}\sum_{i=1}^{5000}\max(0, 65 - \max(y_i, 0))}{65},$$

where $y_i = z_i \cdot \sigma + \mu$ and z_i is the standard normal deviate. This involves solving two equations with two unknowns. Numerically the values for μ and σ may be approximated by iterating on μ and σ values until the right-hand side value is sufficiently close to the left-hand side of the equations.

(4) Transform the parameters in step (3) for APH yields other than 100 using the following formulas:

$$\tilde{\mu}_y = \frac{APH \times \mu}{100}$$

$$\tilde{\sigma}_y = \frac{\sigma}{\mu} \times \tilde{\mu}_y .$$

Once these yield and price parameters are calculated, the correlated yield and price draws can be obtained. The RMA procedure takes 500 draws calculated from the inverse normal of Babcock's Nearly Uniform Sequence (a variant of the rectangular rule also called quasi-random sequences) that ensures that even with a lower number of draws, on average, these draws can still be consistent with a uniform distribution.

After the 500 yield draws are determined, the Iman and Conover (IC) (1982) method is used to impose state-level, yield-price correlations. This approach allows the censored normal yield distribution and the log-normal price distribution to be combined into a joint distribution with a

specified correlation. The IC procedure involves simulating independent variables and then re-sorting them using information derived from the correlation matrix. This process is computationally intensive, but in a two-random-variable context this is not a particularly critical concern. The procedure has gained wide usage due to its intuitive appeal and relative simplicity. Anderson, Harri, and Coble (2011) note that IC tends to have a slight downward bias in the absolute value of the modeled correlation such that modeled correlations are nearer to zero than the specified correlation. In the context of crop insurance where there is some negative correlation of price and yield for some crops, this bias of IC simulated correlations would slightly inflate premium rates.

Using the outcomes from the 500 correlated draws, yield and revenue rates (e.g., the Harvest Price Revenue Rate (HP Rate) and the Harvest Price Exclusion Option Revenue Rate (RP-HPE Rate)) can then be calculated using the following formulas (i.e. this process is also called a quasi-Monte Carlo simulation):

$$\text{YP Rate} = \sum_{i=1}^{500} \frac{\max(0, C \cdot Y - \max(0, y_i \cdot \tilde{\sigma}_y + \tilde{\mu}_y))}{500 \cdot Y \cdot C}$$

$$\text{RP Rate} = \sum_{i=1}^{500} \frac{\max(0, C \cdot Y \cdot \min(2 \cdot P, \max(P, \tilde{p})) - \max(0, (y_i \cdot \tilde{\sigma}_y + \tilde{\mu}_y) \cdot \min(2 \cdot P, \tilde{p})))}{500 \cdot Y \cdot C \cdot P}$$

$$\text{RP-HPE Rate} = \sum_{i=1}^{500} \frac{\max(0, C \cdot Y \cdot P - \max(0, (y_i \cdot \tilde{\sigma}_y + \tilde{\mu}_y) \cdot \min(2 \cdot P, \tilde{p})))}{500 \cdot Y \cdot C \cdot P}$$

where *C* is the coverage level, *Y* is the APH yield, *P* is the planting time price, y_i is the yield draw, and $\tilde{p}$ is the log-normally distributed harvest time price draw (calculated based on the parameters of the log-normal price distribution).

The "revenue load" can then be calculated as follows:

RP COMBO Revenue Load = RP Rate – Yield Rate
RP-HPE COMBO Revenue Load = RP-HPE Rate – Yield Rate.

The resulting COMBO premium rates are then derived using the following formulas:

RP COMBO Premium Rate = APH Base Premium Rate + RP COMBO Revenue Load

RP-HPE COMBO Premium Rate =

APH Base Premium Rate + RP-HPE COMBO Revenue Load.

To investigate the implications of the procedure used to calculate the price volatility and incorporate it into RP and RP-HPE, we take data developed as in the previous section and investigate the effect on revenue rates. Figures 5.2 through 5.7 show plots of Iman and Conover simulated rates given a range of implied volatilities from 0.15 to 0.5. The mean price, mean yield and yield variability are all held constant while the price volatility is varied on the X-axis. Representative corn, soybean and wheat scenarios are modeled with the parameters reported in Table 5.1

Table 5.1. Attributes of Rate Simulation Scenarios

Attribute	Corn	Soybeans	Wheat
E(price)	$ 5.00	$ 12.00	$ 6.50
E(Yield)	180	50	35
Standard Deviation (Yield)	30	10	10
E(price) x E(Yield)	$ 900.00	$ 600.00	$ 227.50
Correlation (Price Yield)	-0.4	-0.3	0

Separate figures are reported for RP and for RP-HPE for each crop. The current (RMA) rate reflects the current RMA rate procedure. The other, labeled (ALT) reflects the alternative calculation that we believe makes the correct transformation of the implied volatility. In both cases premium rates rise as price volatility increases, as expected. Ultimately, the alternative rate calculation results in higher rates for all price volatility scenarios. However, the difference in rates approaches about one percentage point at very high volatilities. At a more typical volatility of 25 percent the difference in rates is typically slightly less than 0.1 percentage point as shown in Table 5.2.

Table 5.2 Difference of Rates Under Current and Alternative Price Volatility Calculations Given Price Volatility of 0.25

Crop	RP-HPE Difference	RP- rate Difference
Corn	0.090%	0.097%
Soybeans	0.086%	0.097%
Wheat	0.084%	0.099%

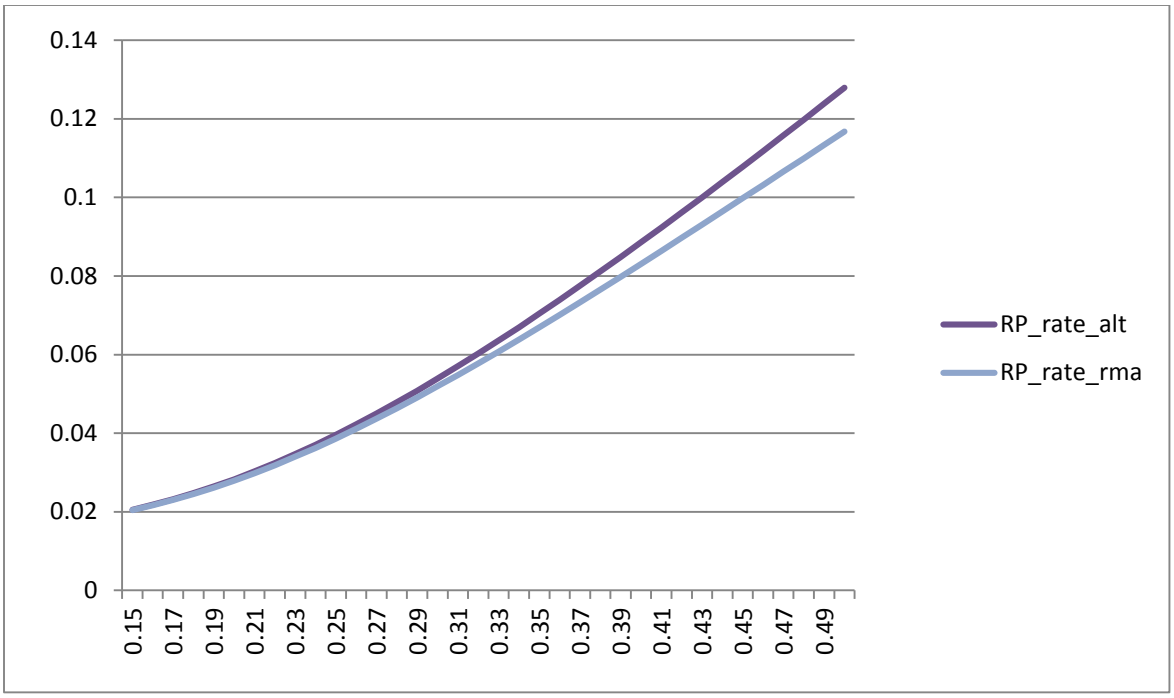


Figure 5.2 Examination of the Current and Alternative Rates for Corn RP Given Price Volatility of 0.15 to 0.50 (Mean Price = $5, E(Yield)=180, Yield Std.Dev.= 30,PY Correlation = -0.40)

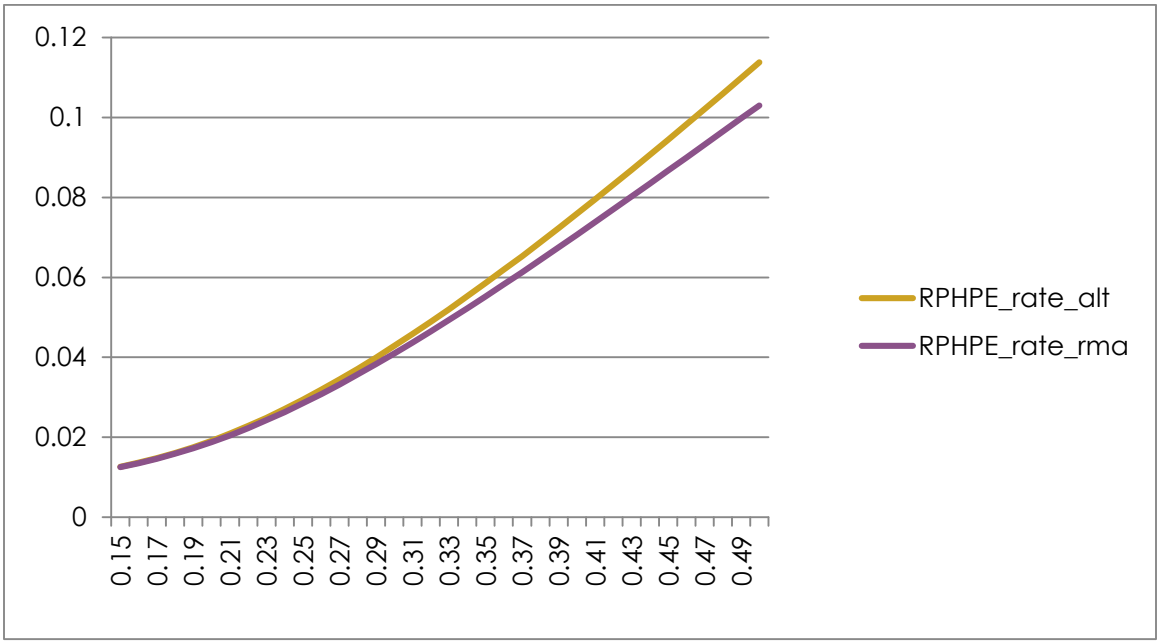


Figure 5.3 Examination of the Current and Alternative Rates for Corn RP-HPE Given Price Volatility of 0.15 to 0.50 (Mean Price = $5, E(Yield)=180, Yield Std.Dev.= 30,PY Correlation = -0.40)

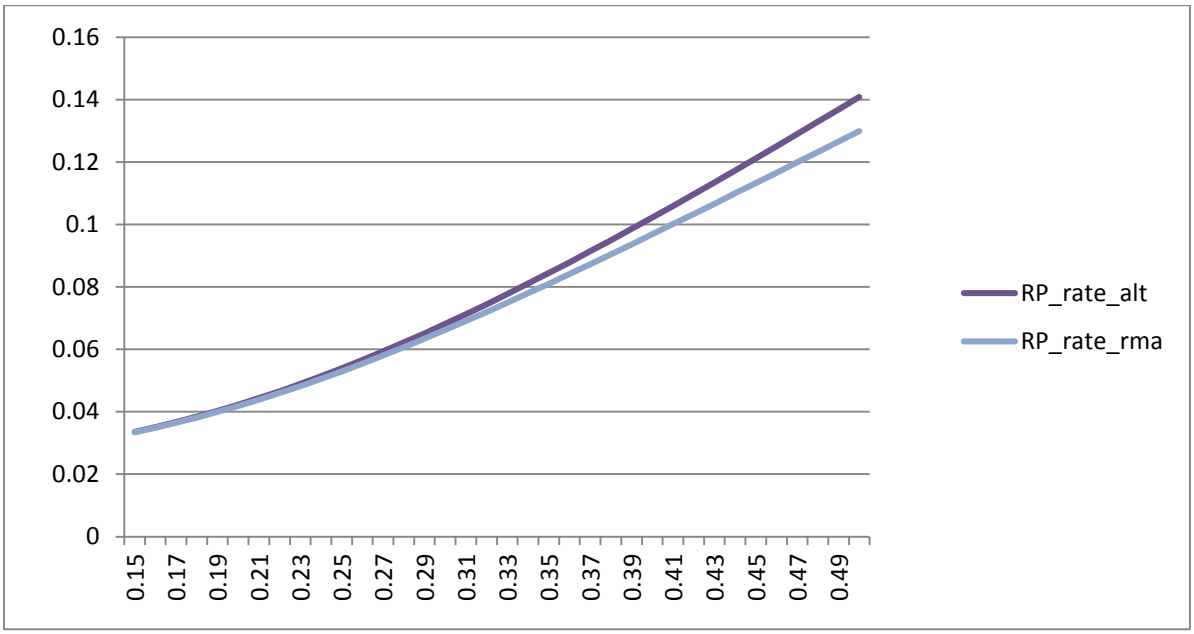


Figure 5.4 Examination of the Current and Alternative rates for Soybean RP Given Price Volatility of 0.15 to 0.50 (Mean Price = $12, E(Yield)=50, Yield Std.Dev.= 10,PY Correlation = -.30)

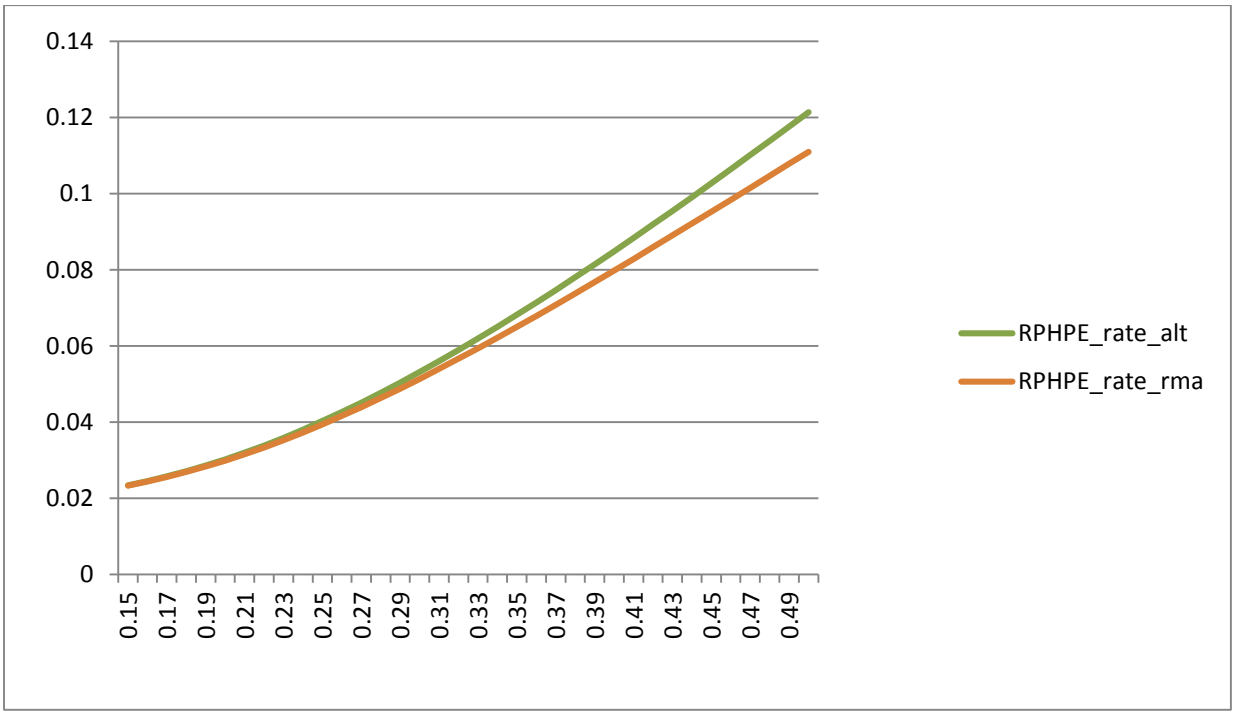


Figure 5.5 Examination of the Current and Alternative Rates for Soybean RP-HPE Given Price Volatility of 0.15 to 0.50 (Mean Price = $12, E(Yield)=50, Yield Std.Dev.= 10,PY Correlation = -0.30)

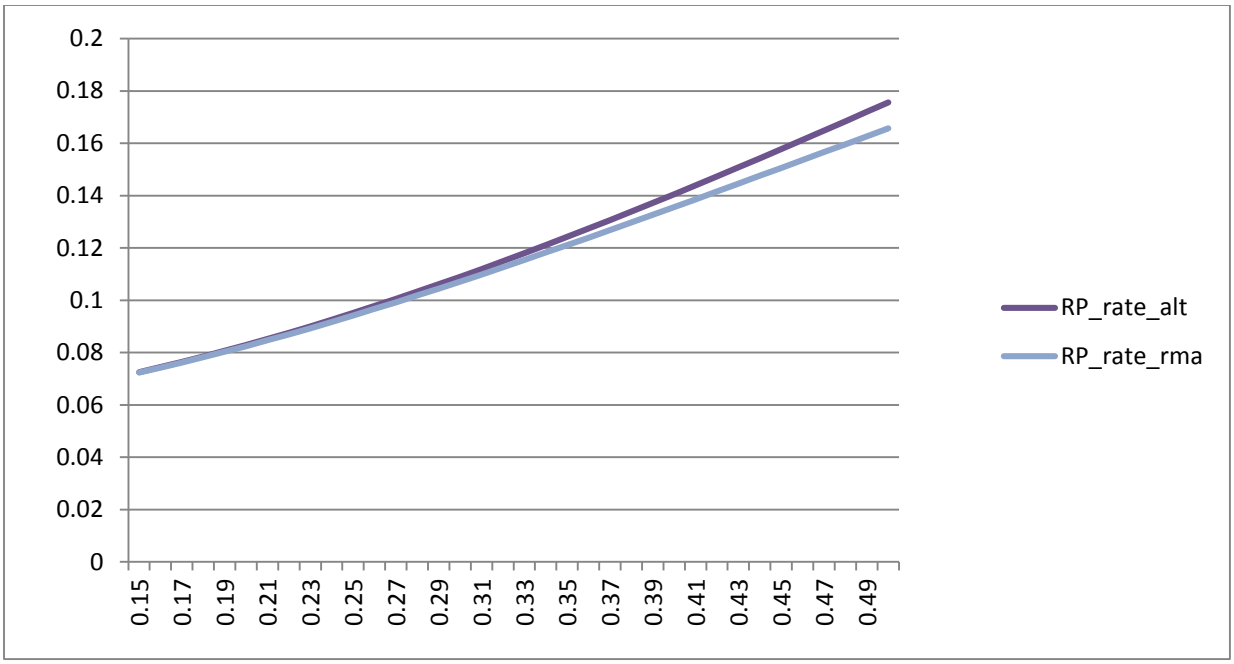


Figure 5.6 Examination of the Current and Alternative Rates for Wheat RP Given rice Volatility of 0.15 to 0.50 (Mean Price = $6.50, E(Yield)=35, Yield Std.Dev.= 10, PY Correlation = -0.0)

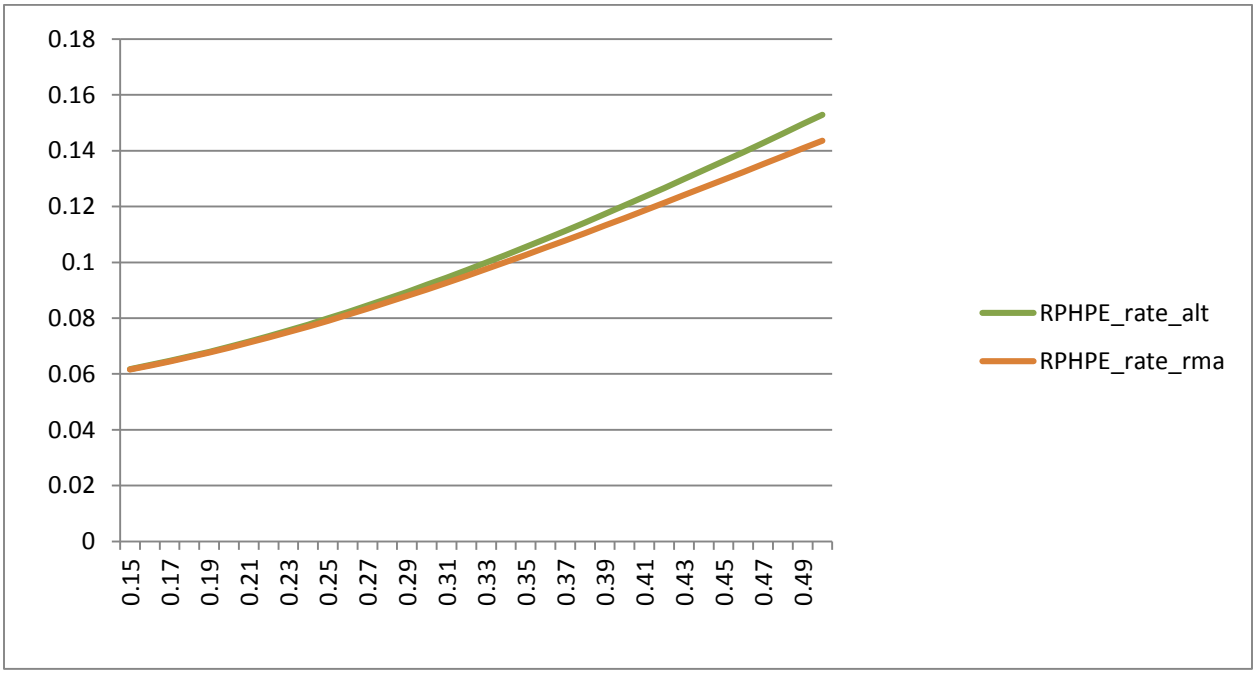


Figure 5.7 Examination of the Current and Alternative Rates for Wheat RP-HPE Given Price Volatility of 0.15 to 0.50 (Mean Price = $6.50, E(Yield)=35, Yield Std.Dev.= 10,PY Correlation = -0.0)

6. Compare and contrast two or more alternative methods for calculating implied volatility to the method used by RMA

6.a. Empirical analysis of price volatility

Current rating and contract design practices in the federal crop insurance program utilize published measures of the Black-Scholes implied volatility. The Black-Scholes option pricing model was introduced by Black and Scholes in 1973 and has become the dominant mechanism for communicating volatility concepts in the world of finance. Warren Buffett (pages 20-21, 2009) states that

"The Black-Scholes formula has approached the status of holy writ in finance, and we use it when valuing our equity put options for financial statement purposes … the formula represents conventional wisdom and any substitute that I might offer would engender extreme skepticism."[2]

In spite of its dominance as a market-based measure of price volatility, the formula is subject to a significant number of strong assumptions and criticisms. These assumptions include a log-normal distribution of asset prices, no transactions costs, no riskless arbitrage, risk neutrality, continuous trading, and a constant, real discount rate. Such assumptions are, in practice, often violated and thus the Black-Scholes model is subject to a number of limitations that may lead to biases or errors in the measurement of volatility. However, any alternative pricing model is likely to be subject to the same or similar assumptions and limitations. In nearly every alternative, a central assumption involves the rationality and efficiency of markets—the "no-arbitrage" assumption. A central tenet in most economic theories is that efficient markets will act to eliminate riskless arbitrage opportunities. Put differently, the no-arbitrage assumption plays a central role in most if not all option pricing models.

The academic literature surveyed in chapter 2 of this report has documented specific aspects of the Black Scholes model and various circumstances of markets that may lead to violations of the assumptions underlying the volatility. One of the most commonly perceived shortcomings of the Black Scholes model involves departures from the log-

[2] Warren E. Buffett, Berkshire-Hathaway letter to shareholders, February 27, 2009, available at http://www.berkshirehathaway.com/2008ar/2008ar.pdf.

normality assumption in the tails of the distribution. Any two-parameter probability distribution can be calibrated to a single futures option quote. This involves the inherent assumption that the futures price represents an unbiased prediction of the future spot price. Using the price of the underlying futures contract and any concomitant options price, the Black-Scholes option pricing model can be solved for the implied volatility. When considered across a range of concurrent options with different strike prices, if the underlying assumptions of the option pricing model are valid, identical measures of the volatility should be derived. In practice, the volatilities often tend to not be constant but rather increase as one moves into the tails of the distribution. Such rising patterns of volatility are commonly referred to as "smiles" and "smirks" because of the patterns they exhibit when plotted across different strike prices.

The presence of such volatility patterns suggests that one or more of the Black-Scholes assumptions is violated. The strong assumption that asset prices are log-normally distributed is often pointed to as a reason for departures from the constant volatility implied by Black-Scholes. A number of different alternative option pricing models have been developed in an attempt to address these shortcomings. We consider several variants of a prominent alternative that has been widely applied in the literature (see, for example, Egelkraut, Garcia, and Sherrick, 2007). A common approach is to calibrate an entire range of concurrent strike prices to a constant volatility (see, for example, Fackler and King, 1989, and Sherrick, Garcia, and Tirupattur, 1996). This is typically done using a log-normal or gamma distribution though any valid probability distribution function can be used in this manner.

Before proceeding to a consideration of alternative option pricing models and their potential utility in the federal crop insurance program, it is important to emphasize several practical features a measure of price volatility derived for use in pricing a publicly subsidized revenue insurance plan should have. We believe that the Black-Scholes formulation, despite its potential shortcomings, has several important advantages in terms of these practical features.

We believe that transparency is of paramount importance in establishing the terms of coverage and premium rates in the Federal Crop Insurance Program. Important parameters such as the price volatility used in rating revenue coverage should be publicly available and easy to reference. Volatility measures derived from a method or model that does not have this feature of transparency may be perceived as coming from a "black-box" that is difficult to justify, communicate, understand, and replicate. We also believe that there are important advantages to the use of a common and widely accepted measure of price volatility that is

generated and reported from well-established and trusted industry sources rather than from the RMA itself. Criticisms will always be present with any rating method, but having an external source for rating parameters that is widely utilized in the financial sector may be preferable to a model-based measure constructed internally by RMA. The Black-Scholes volatility has these valuable attributes and we are unaware of any prominent alternative that would have similar advantages.

Of course, such advantages must be weighed against any potential gains in accuracy that may be derived from adopting an alternative to the Black-Scholes model. The overarching goal of this segment of our analysis is to consider how different alternative measures of market price volatility may be. We also consider a type of "out-of-sample" evaluation of the forecasting performance of alternative volatility measures by comparing the model-based projected volatilities to an empirical measure of the actual realized volatility. This approach, though illuminating, is constrained by the fact that the actual realized volatility is unobservable and must be represented by a proxy measure. We utilize the cumulative sum of daily returns over the life of the contract as a measure of the realized volatility and compare alternative model-based measures of volatility to this empirical measure.

In the empirical analyses which follow, we focus on the options contracts that are relevant to pricing Revenue Protection (RP) crop insurance for corn, soybeans, wheat, and rice. The specific contracts and periods of price discovery are as follows: for corn, the December contract quoted in February; for soybeans, the November contract quoted in February; for rice, the November contract quoted from January 15 – February 14; and for wheat, the July contract quoted from August 15 – September 14 in the previous calendar year. Realized prices are given by the average closing price over the month prior to the contract closing.

6.b The Black-Scholes Model and its alternatives

The Black-Scholes option pricing model was reviewed in section 3. There are many alternatives to this standard specification. We focus on variations of a model proposed by Egelkraut, Garcia, and Sherrick (2007). Their model essentially calibrates a range of put and call options being traded concurrently to determine parameters of the probability distribution of prices. Specifically, they choose parameters of $\varphi(.)$ to minimize the following objective function

$$\sum_{i=1}^{k}[V_{C,i} - \delta_t \int_0^{\infty} \max(0, F_t - S_i)\varphi(\theta, F_t)dF_t]^2 + \sum_{i=1}^{k}[V_{p,j} - \delta_t \int_0^{\infty} \max(0, F_t - S_j)\varphi(\theta, F_t)dF_t]^2$$

This specification differs from that above in that it estimates a mean and variance of the distribution of prices using the range of puts and calls and does not impose the restriction that the mean is equal to the futures price. We consider both versions of the extended log-normal option pricing model.

A third approach to pricing options and deriving implied volatilities makes use of a completely model free method. Such model-free approaches have recently gained prominence and a model-free volatility index (the VIX) for the S&P 500 stock index is now traded on the Chicago Board Options Exchange (CBOE). The VIX is usually derived for much shorter trading periods than those that we consider here (i.e., those pertinent to pricing revenue insurance over the planting to harvest period) and thus is not likely to have a great deal of relevance to our specific objectives. However, it does provide an interesting basis for comparison. The VIX type index is given by

$$\sigma^2 = \frac{2}{t}\sum_i \frac{\Delta S_i}{S_i^2} e^{rt} Q(S_i)$$

where $Q(S_i)$ is the value of the option at strike S_i and ΔS_i is the average of the difference between the two adjoining strikes.

6.c. Trading volume issues

Market-based measures of important parameters such as the volatility have the important advantage of being based upon the collective wisdom and assessment of a group of profit-motivated traders. Profit seeking should eliminate any biases that the market recognizes. However, many futures and options contracts are thinly traded. This is particularly the case for many of the put and call options that are used in pricing revenue insurance. The informational content of contracts that are not traded is questionable. The exchange has methods that involve the

judgment of experts and market participants and a consideration of previous settlement prices.[3]

Figure 6.1 illustrates the annual average daily proportions of contracts (over the relevant price discovery periods) that had no trades. The proportions are high and even approach 100% for some years and commodities. Rice options are particularly thinly traded. In the empirical analysis that follows, we attempt to control for the potential biases that may result from including option quotes for contracts that were not traded. We do this in several ways. Our baseline analysis excludes any day/contract combinations that had zero trades. We also exclude any day that did not have at least 4 different options (puts and calls) that had positive trade values. Of course, we must have at least two different options in order to calibrate a two-parameter distribution. Requiring 4 options to be traded makes the comparison to Barchart methods as consistent as possible, since four contracts (the two nearest to the money puts and two nearest calls) are used to derive the volatility measure.

A related issue pertains to the fact that trading volumes tend to be lower for far into or out of the money contracts. Figure 6.2 compares the in/out degree of "moneyness" (measured as the ratio of the strike price to the underlying futures price minus one, such that a value of zero represents an at the money option) to the total volume of options contracts traded each day. The figure illustrates two important points. First, a great many contracts experience days with no trades. This is particularly true for rice and, to an extent, for wheat. Second, those contracts close to the money tend to be traded in much greater volumes than contracts that are far in or out of the money. This has important implications for methods that use a range of strike prices to determine market volatility. To the extent that one is concerned about the informational content of an option that is not traded, incorporating such options in the determination of a volatility measure may result in significant biases. Put differently, a non-traded option may not reflect the assessment of market participants and thus may not be relevant in the determination of a market-based volatility measure.

To address this potential limitation of options with no or low volumes, we consider alternative volatility measures that weight options by total daily trading volumes. This naturally gives greater weight to options trading near the money since their volumes tend to be higher. We expect that this will result in volatility measures that are more similar to the

[3] Exchange settlement procedures are detailed by the CME at http://www.cmegroup.com/market-data/files/cme-group-settlement-procedures.pdf

conventional Black-Scholes (Barchart) volatility measures that are averages of the nearest to the money puts and calls.

We used numerical optimization routines to calculate several different volatility measures.[4] The specific measures of volatility that were considered are:

1. A standard Black-Scholes implied volatility calculated as the average of the volatilities from the two nearest the money puts and two nearest the money calls.
2. A log-normal volatility calibrated across the range of strike prices each day. The mean of the log-normal distribution was constrained to be equal to the underlying futures price.
3. A log-normal volatility calibrated as above but with each option weighted by total trading volume.
4. A log-normal volatility calculated using the calibration methods of Egelkraut, Garcia, and Sherrick (2007). The calibration estimates two parameters—corresponding to a mean price and the volatility of prices. This allows for differences between the mean price and the underlying futures price.
5. A volume-weighted version of the Egelkraut, Garcia, and Sherrick (2007) volatility.

A measure of the realized price volatility was calculated by summing the squared values of daily returns between the price discovery period and the month immediately preceding the expiration of the contract. This method has been widely applied in empirical comparisons of alternative measures of volatility (see, for example, Anderson and Bondarenko (2007)). The specific measure of the realized volatility is given by

.

$$\sqrt{\sum \ln\left(\frac{p_t}{p_{t-1}}\right)^2}\ .$$

Table 6.1 presents a summary of the differences in forecasted and realized volatilities. In this evaluation, contracts with zero trades and days without at least four different strikes are omitted. Although this type of evaluation is standard, it should be noted that the measure of the realized volatility may be suspect since it is an empirical representation of an unobservable variable. The results in the table lead to two clear

[4] The SAS code used to derive the alternative volatility measures in presented in an appendix.

conclusions. First, differences in volatility measures across the different methods are small. The minimal differences according to each evaluation criterion are highlighted. A second important conclusion is that the standard Black Scholes estimate appears to come closest to representing the realized volatilities in the majority of cases, though a log-normal calibrated across all traded option contracts produces nearly identical measures and is preferred according to these criteria. Rice is omitted due to the very low number of traded contracts on most days. As expected, the model-free volatility measure does not represent realized volatility very well across the significant term of these contracts (8-10 months). Allowing the mean to deviate from the underlying futures price does not appear to improve the volatility measures, though again the differences are modest. Our interpretation of these results is that, with the exception of the VIX-type volatility, any volatility measure would appear to be appropriate since the differences are slight. Further to this point, we would note that there are many other aspects of the rating procedures (measurement of price/yield correlation, measurement of yield risk, catastrophic loading, etc.) that are likely to produce much more variation and potential for measurement error than the choice of a volatility measure.

A number of different reasons for biases to exist in volatility measures have been offered in the literature. These reasons include risk premiums, transactions costs, and asymmetric information. Our results suggest that such biases, to the extent they exist, are minor and their presence does not result in alternative volatility measures having superior predictive power relative to the conventional Black Scholes methods.

One of the key advantages of the methods of Egelkraut, Garcia, and Sherrick (2007) is that the mean is not restricted to equal the underlying futures price, though it certainly nests this restriction as a special case. This additional parameter must necessarily produce a tighter fit in the calibration exercise. However, this added flexibility does not necessarily translate into superior out-of-sample forecasting performance, such as that represented in the comparisons to realized volatilities.

We would argue that the underlying futures price may convey a significant amount of information about the expectations of future prices. The fact is that futures contracts tend to be much more highly traded than is the case for options contracts. To examine this point, we considered the ratio of total daily trading volume on the underlying futures contract to the total volume of options traded each day. Figure 6.3 presents the annual averages of the daily ratios of total volume on options to total volume on the underlying contract. The results

demonstrate that the volume on futures is many times that of the corresponding options (typically less than 1%). Thus, the futures price may convey a great deal of information about traders' expectations about future market conditions that is not captured in a specification that allows the mean of prices to depart from the futures price.

We also compared the alternative mean (expected) prices generated by the model of Egelkraut, Garcia, and Sherrick (2007) to realized spot prices (measured by the simple average of the futures price in the month immediately preceding expiration of the contract). Table 6.2 contains the results of this comparison along with an analogous comparison of the futures price to the realized price. In every case, the underlying futures price appears to offer a more accurate forecast of the realized price than what is generated by the model-based alternatives. This can be taken to suggest that the models that omit the important information inherent in the very significant trading of futures contracts do not capture biases impacting the mean of the distribution of expected prices. We also compared the differences in the modeled mean prices to the concomitant futures prices. Table 6.3 presents a summary of this evaluation. Although specific statistical testing of the significance of the bias is complicated by the non-independence of daily quotes, the mean differences (which represent average biases) are very small when compared to the standard deviations of the differences.

We repeated the volatility forecast evaluation using a larger sample that included options contracts that were not traded on the day of the quote. We also lowered the required number of strikes from four to two. The results, which now include rice, are presented in Table 6.4. The results are fully consistent with those presented in Table 6.1. Because the samples differ, a direct comparison of forecasting performance is not available in comparing the results. However, the various metrics of forecasting performance are slightly smaller when the larger sample that includes non-traded contracts is used. This may suggest that additional information is conveyed in the settlement prices of contracts with no volume, though the differences are very slight.

A central question to the analysis involves the extent to which the alternative volatility measures produce different estimates. Figure 6.4 plots the conventional Black-Scholes estimates, calculated as described above, against each of the alternative volatility measures. The figures include a 45-degree line representing equality of the measures. As one might expect, the log-normal estimates that restrict the mean of prices to equal the futures price are very close to the conventional Black-Scholes estimates. Differences between the conventional and the Egelkraut, Garcia, and Sherrick (2007) model-based estimates are larger but are still

relatively modest. The Black-Scholes volatilities tend to be slightly higher than the model-based estimates.

Finally, we considered the degree of "smile" and "smirk" behavior that may be inherent in the Black-Scholes volatility estimates. Figure 6.5 plots the volatilities across alternative strikes against the ratio of the strike to futures price. The volatilities are normalized by dividing by the Black-Scholes estimate at the money, thereby allowing all volatilities to be compared in a single chart. The smiles and smirks so often noted in the literature are apparent and deviations from the at-the-money volatilities tend to be higher when one moves far in or out of the money. The aforementioned limited volumes on strikes far from the futures price tends to minimize the impacts of these deviations in the forecast comparisons presented above.

6.d. Summary and conclusions

Current revenue insurance rating methods utilize a conventional Black-Scholes based implied volatility that is taken from publicly accessible sources external to the RMA (i.e., Barchart). This measure has significant practical advantages in terms of transparency and in keeping with common market practices. The Black-Scholes implied volatility is, as Warren Buffet stated, a "holy writ" in finance that represents "conventional wisdom" against alternatives that would engender "skepticism." This is not to say that the Black-Scholes model is not without faults. As we have shown, significant differences in volatility appear across the range of strikes, suggesting that the log-normality assumption and other maintained hypotheses of the Black-Scholes model may not always be correct. Further to this point, measures of risk in the tails are particularly important in rating revenue insurance contracts. However, such limitations of any model must be weighed against its advantages and disadvantages relative to viable alternatives.

Our evaluation suggests that careful consideration should be given to trading volumes when using market quotes to derive important parameters such as market volatilities. In the case of options markets, trade volumes are frequently very thin. The extent to which a contract that is quoted, but not traded, actually reflects market conditions is questionable. Using options contracts close to the money addresses this problem since such contracts are more fluidly traded. However, it bears emphasizing that the potential for measurement errors associated with the risks of extreme price movements may arise when near the money contracts are heavily weighted. Rice bears special mention in that the trade volumes are very low and are frequently zero. In lieu of a viable

alternative, we do not recommend adoption of any other approach to measuring volatility. However, we do recommend that trading volumes be carefully monitored to ensure that rating parameters use a volatility measure that reflects actual market conditions.

We find that differences between the Black-Scholes methodology currently in use and various alternatives are modest. The Black-Scholes methodology generally performs better than or equivalent to alternative methods for the contracts considered here (i.e., those contracts relevant to pricing revenue coverage) in terms of forecasting realized price variability. The differences that do exist are so small as to be modest when compared to measurement issues associated with other important parameters, such as yield/price correlations. Alternative approaches to representing the distribution of prices generally do not provide superior forecasts of the mean of prices.

A summary of the recommendations emerging from our empirical analysis is as follows:

1. We recommend that RMA continue to utilize a publicly available and external measure of the market price volatility. The current use of Barchart as a source seems entirely reasonable and we see no reason to recommend consideration of any alternatives. In light of the prominence of the Black-Scholes volatility, we believe it is likely that any publicly available alternatives are likely to use the same methodology as that applied by Barchart.
2. We do not recommend adoption of any alternative that requires extensive analytical calculations by RMA or its contractors. We considered several alternatives and did not find improvements in accuracy that would merit such a change. Transparency is a paramount concern in rating and the use of a complex "black-box" would diminish such transparency and potentially lead to disputes over the basis for rating parameters. The overwhelming dominance of the Black-Scholes methodology as a mechanism for communicating volatility concepts in financial markets provides considerable justification for its use in rating revenue insurance. These advantages would be lost with any other currently available alternatives.
3. We recommend that RMA carefully consider trade volumes when utilizing options and futures quotes. Again, the differences associated with using volume-weighting or other methods do not merit adopting a method that is not external to RMA and that is not publicly available. However, contracts and/or days with no trades may need to be excluded in determining rating parameters.

4. We recommend that RMA continue to use the underlying futures price as an unbiased forecast of the future realized price. Models that allow for differences were not supported by our empirical analysis.

Table 6.1. Comparisons of Alternative Volatility Measures to Realized Volatility

(Excluding Zero Volume Options and Days with Fewer Than 4 Strikes)

Method	All (n=1,251)	Corn (n=539)	Soybeans (n=554)	Wheat (n=158)
	Mean Absolute Error			
Standard Black Scholes (Barcharts)	0.0479	0.0418	0.0389	0.1031
Log-Normal (Restricted Mean)	0.0487	0.0417	0.0410	0.1021
Weighted Log-Normal (Restricted Mean)	0.0496	0.0417	0.0429	0.1028
Egelkraut, Garcia, and Sherrick (Unweighted)	0.0528	0.0467	0.0408	0.1191
Egelkraut, Garcia, and Sherrick (Weighted)	0.0532	0.0475	0.0410	0.1191
Model-Free Volatility	0.1488	0.0456	0.2366	0.2012
	Mean Squared Error			
Standard Black Scholes (Barcharts)	0.0046	0.0033	0.0028	0.0162
Log-Normal (Restricted Mean)	0.0046	0.0033	0.0029	0.0159
Weighted Log-Normal (Restricted Mean)	0.0047	0.0033	0.0030	0.0159
Egelkraut, Garcia, and Sherrick (Unweighted)	0.0059	0.0045	0.0035	0.0202
Egelkraut, Garcia, and Sherrick (Weighted)	0.0060	0.0045	0.0034	0.0204
Model-Free Volatility	0.1656	0.0040	0.3514	0.0585
	Root Mean Squared Error			
Standard Black Scholes (Barcharts)	0.0680	0.0575	0.0528	0.1271
Log-Normal (Restricted Mean)	0.0681	0.0571	0.0541	0.1261
Weighted Log-Normal (Restricted Mean)	0.0685	0.0571	0.0551	0.1262
Egelkraut, Garcia, and Sherrick (Unweighted)	0.0769	0.0667	0.0588	0.1422
Egelkraut, Garcia, and Sherrick (Weighted)	0.0772	0.0674	0.0586	0.1427
Model-Free Volatility	0.4069	0.0631	0.5928	0.2420

Table 6.2. Comparisons of Alternative Measures of Expected Prices to Realized Prices

Method	All	Corn	Soybeans	Wheat
	Mean Absolute Error			
Egelkraut, Garcia, and Sherrick (Unweighted)	86.77	56.39	104.36	130.94
Egelkraut, Garcia, and Sherrick (Weighted)	87.22	56.73	105.01	131.12
Underlying Futures	86.17	56.11	103.73	129.35
	Mean Squared Error			
Egelkraut, Garcia, and Sherrick (Unweighted)	14,257.48	5,483.73	19,718.07	25,616.73
Egelkraut, Garcia, and Sherrick (Weighted)	14,349.51	5,514.22	19,863.98	25,730.85
Underlying Futures	14,170.18	5,465.15	19,642.17	25,240.37
	Root Mean Squared Error			
Egelkraut, Garcia, and Sherrick (Unweighted)	119.40	74.05	140.42	160.05
Egelkraut, Garcia, and Sherrick (Weighted)	119.79	74.26	140.94	160.41
Underlying Futures	119.04	73.93	140.15	158.87

Table 6.3. Summary of Bias between Projected Mean Prices and Realized Prices (Pti-Pt)

Statistic	All	Corn	Soybeans	Wheat
Difference in Modeled Mean Price and Futures (Unweighted)				
Mean	1.8499	0.6999	2.8296	2.3381
Std. Dev.	2.7776	1.0383	2.8194	4.6437
Min	-33.1037	-3.2119	-33.1037	-21.7052
Max	15.0282	3.7038	12.3198	15.0282
Difference in Modeled Mean Price and Futures (Weighted)				
Mean	3.0130	1.5021	4.4947	2.9721
Std. Dev.	4.1530	1.4700	4.6257	6.2419
Min	-27.0069	-4.4174	-27.0069	-18.5018
Max	52.2107	7.6112	21.8932	52.2107
Difference in Modeled Mean Price and Realized Price (Unweighted)				
Mean	8.3479	11.0994	5.3670	9.4699
Std. Dev.	119.1600	73.2840	140.4450	160.3070
Min	-317.4970	-182.6150	-317.4970	-273.1200
Max	509.6920	192.9090	509.6920	389.8480
Difference in Modeled Mean Price and Realized Price (Weighted)				
Mean	9.5174	11.9017	7.0321	10.1288
Std. Dev.	119.4590	73.3660	140.8910	160.6250
Min	-315.0300	-182.6230	-315.0300	-273.0580
Max	522.7580	192.8500	522.7580	392.7960
Difference in Futures and Realized Price				
Mean	6.4950	10.3995	2.5374	7.0816
Std. Dev.	118.9090	73.2600	140.2540	159.2460
Min	-320.8590	-183.1070	-320.8590	-275.5600
Max	504.9780	191.6840	504.9780	386.3410

Table 6.4. Comparisons of Alternative Volatility Measures to Realized Volatility

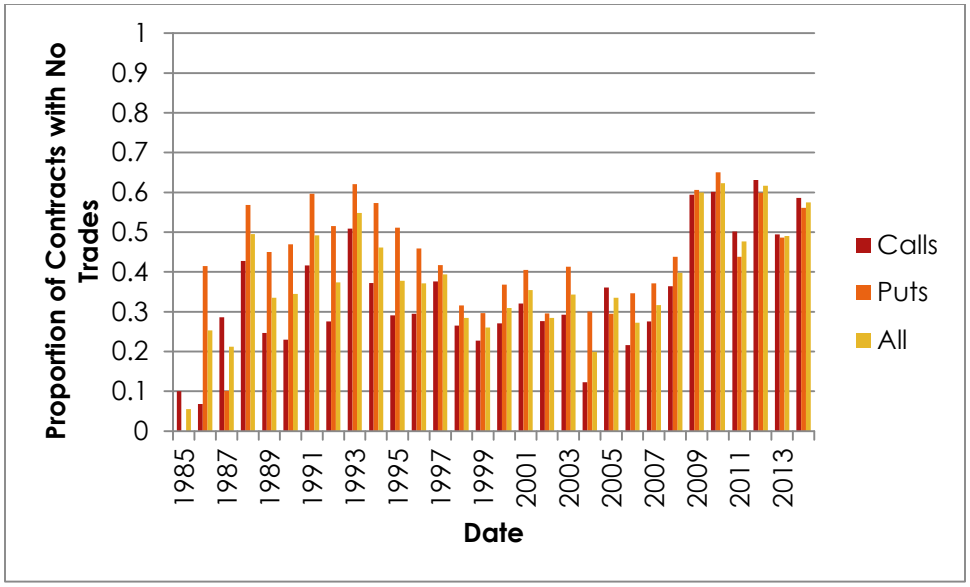
(Excluding Days with Fewer Than 2 Strikes—Includes Contracts with No Trades)

Method	All (n=2,072)	Corn (n=634)	Rice (n=222)	Soybeans (n=651)	Wheat (n=565)
	Mean Absolute Error				
Standard Black Scholes (Barcharts)	0.0518	0.0388	0.0462	0.0358	0.0882
Log-Normal (Restricted Mean)	0.0518	0.0382	0.0455	0.0377	0.0873
Weighted Log-Normal (Restricted Mean)	0.0529	0.0373	0.0570	0.0398	0.0915
Egelkraut, Garcia, and Sherrick (Unweighted)	0.0565	0.0451	0.0509	0.0363	0.0965
Egelkraut, Garcia, and Sherrick (Weighted)	0.0522	0.0453	0.0476	0.0326	0.0903
Model-Free Volatility	0.1322	0.0429	0.7833	0.0512	0.0566
	Mean Squared Error				
Standard Black Scholes (Barcharts)	0.0056	0.0029	0.0044	0.0024	0.0131
Log-Normal (Restricted Mean)	0.0055	0.0028	0.0042	0.0026	0.0128
Weighted Log-Normal (Restricted Mean)	0.0055	0.0028	0.0058	0.0027	0.0132
Egelkraut, Garcia, and Sherrick (Unweighted)	0.0068	0.0040	0.0054	0.0030	0.0154
Egelkraut, Garcia, and Sherrick (Weighted)	0.0060	0.0038	0.0049	0.0024	0.0144
Model-Free Volatility	0.1446	0.0035	1.2442	0.0046	0.0063
	Root Mean Squared Error				
Standard Black Scholes (Barcharts)	0.0748	0.0536	0.0663	0.0491	0.1143

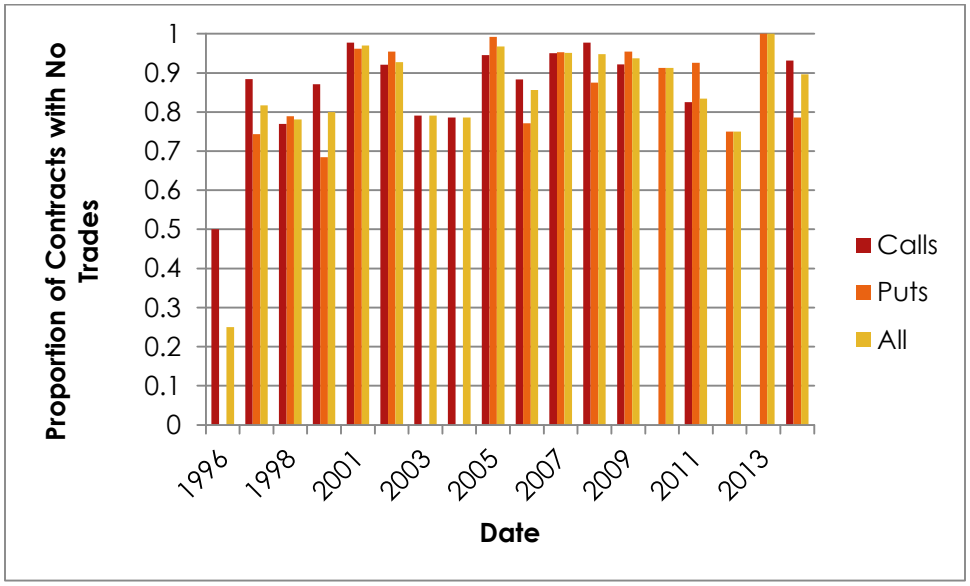
Log-Normal (Restricted Mean)	0.0743	0.0532	0.0648	0.0506	0.1130
Weighted Log-Normal (Restricted Mean)	0.0743	0.0529	0.0763	0.0515	0.1148
Egelkraut, Garcia, and Sherrick (Unweighted)	0.0827	0.0630	0.0738	0.0544	0.1242
Egelkraut, Garcia, and Sherrick (Weighted)	0.0776	0.0616	0.0699	0.0489	0.1201
Model-Free Volatility	0.3802	0.0591	1.1154	0.0676	0.0792

Figure 6.1. Average Proportion of Options Contracts with No Daily Trading Volume

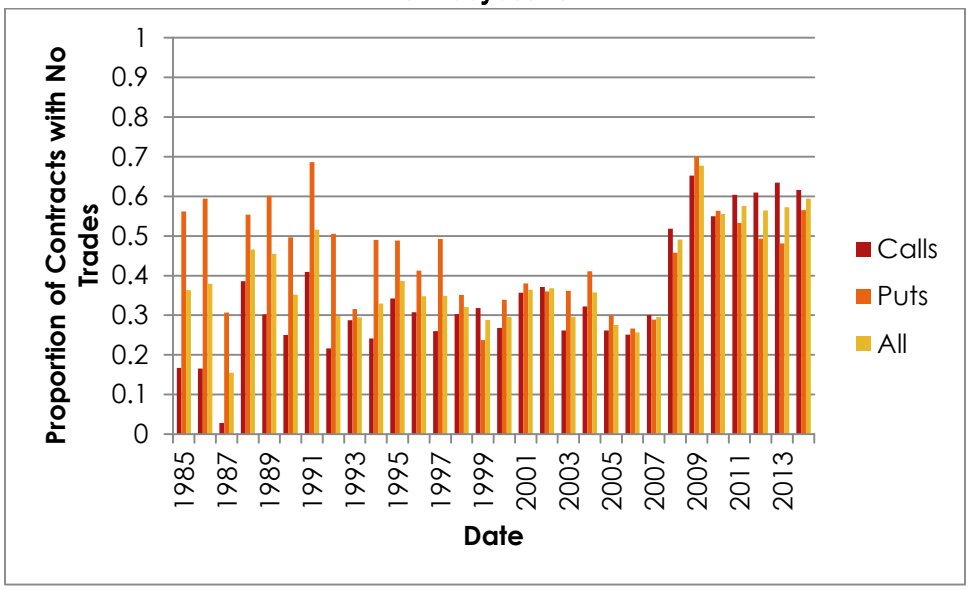
A. Corn



B. Rice



C. Soybeans



D. Wheat

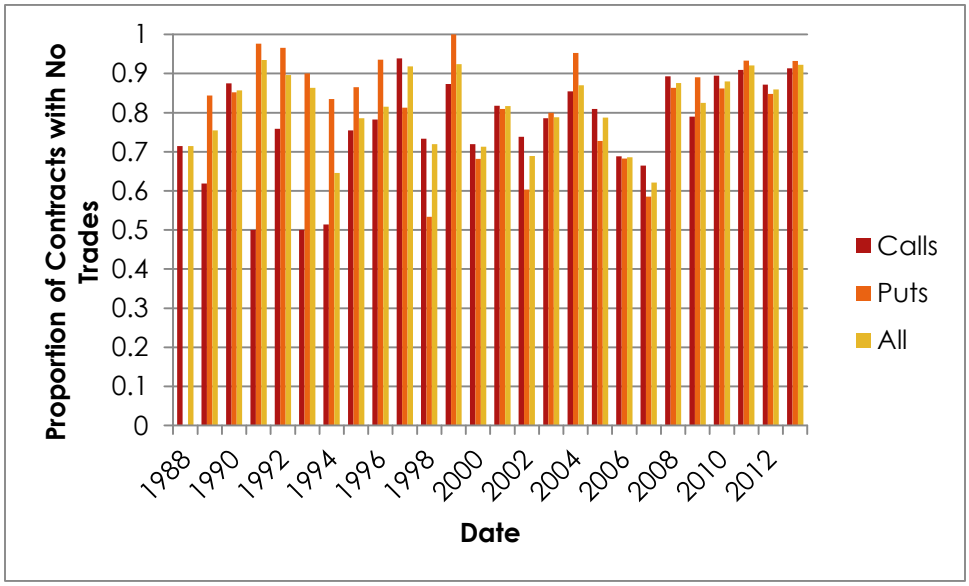
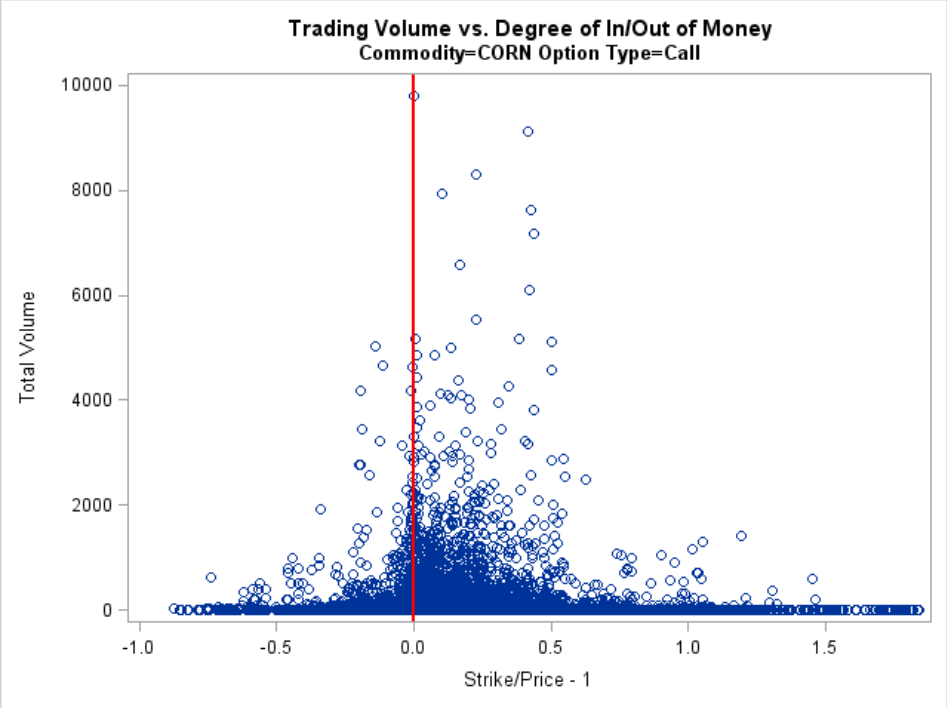
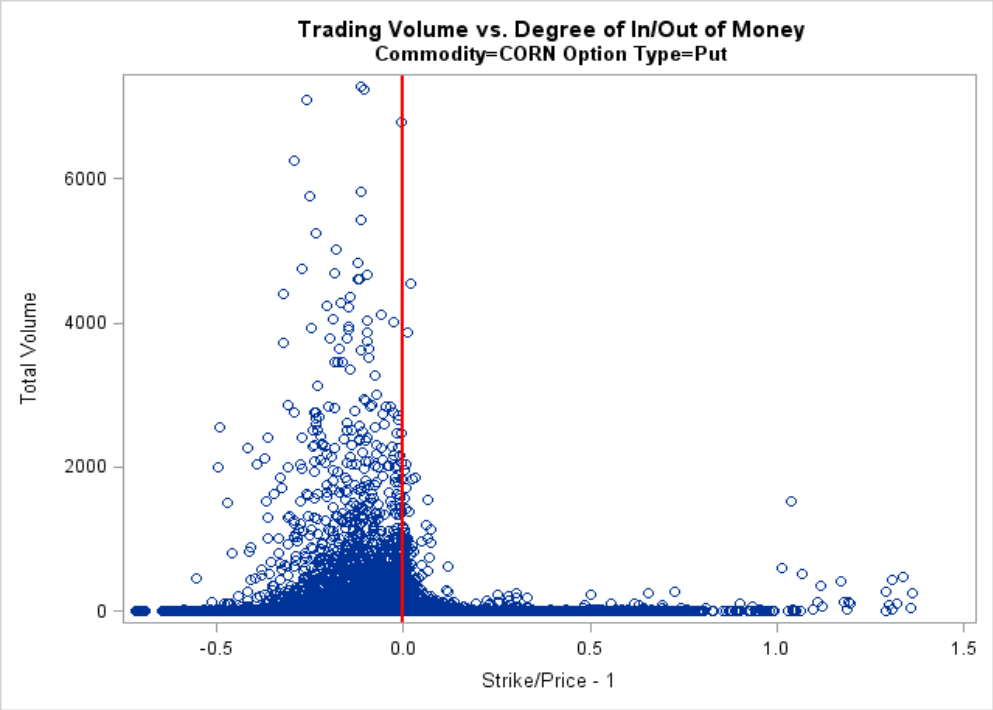


Figure 6.2. Daily Options Trading Volumes and Relative "Moneyness" (Strike/Futures – 1)

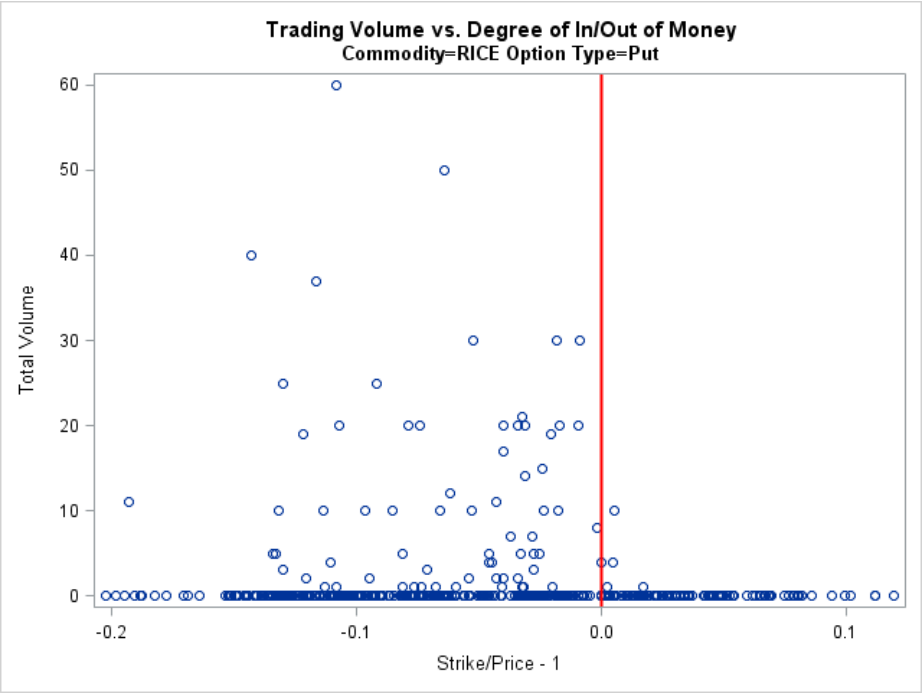
A. Corn Call Options



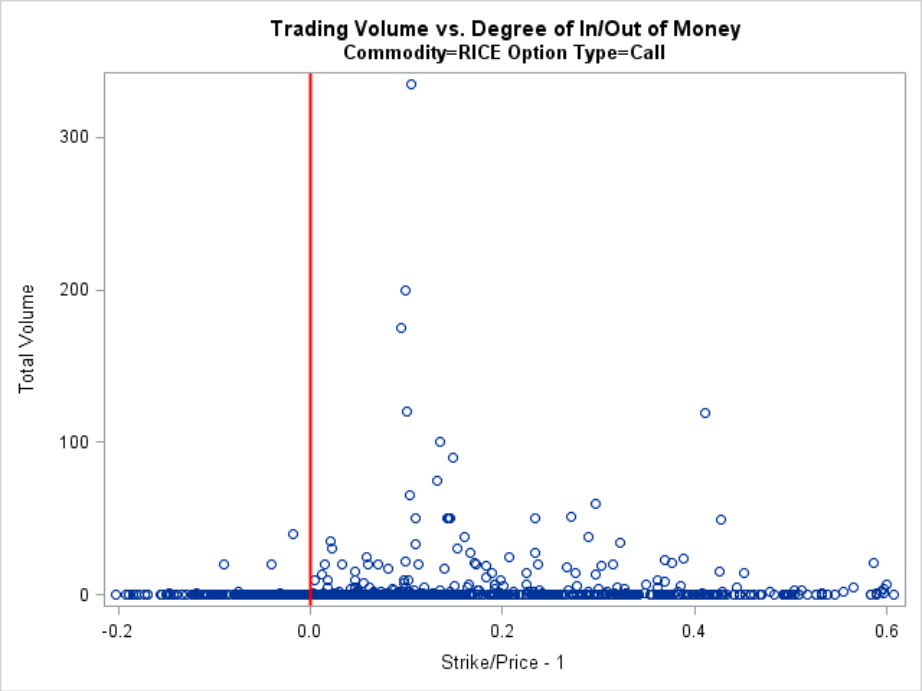
B. Corn Put Options



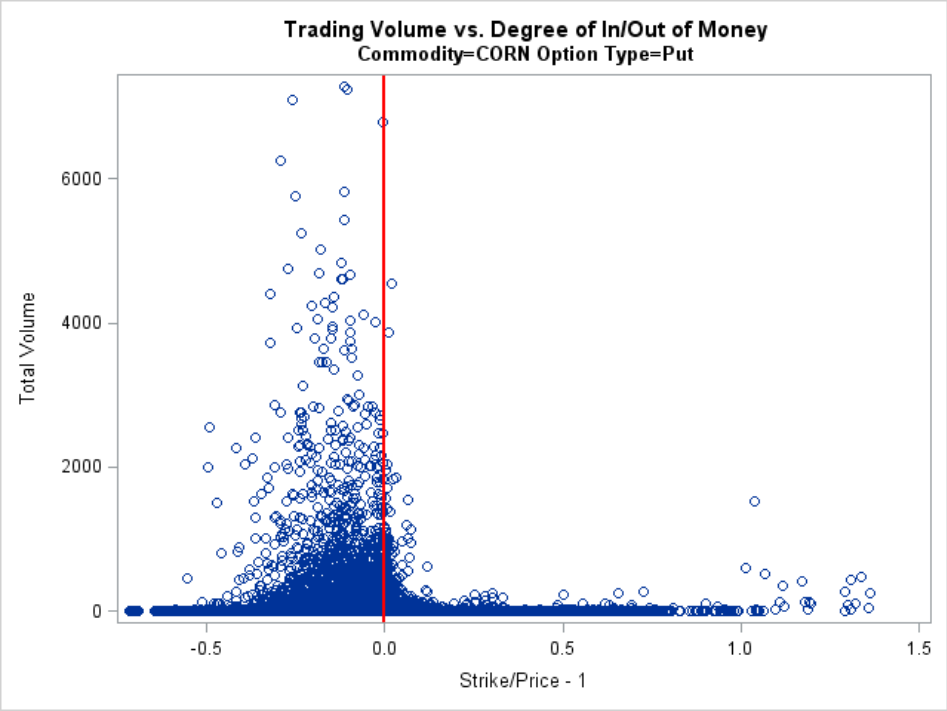
Rice Put Options



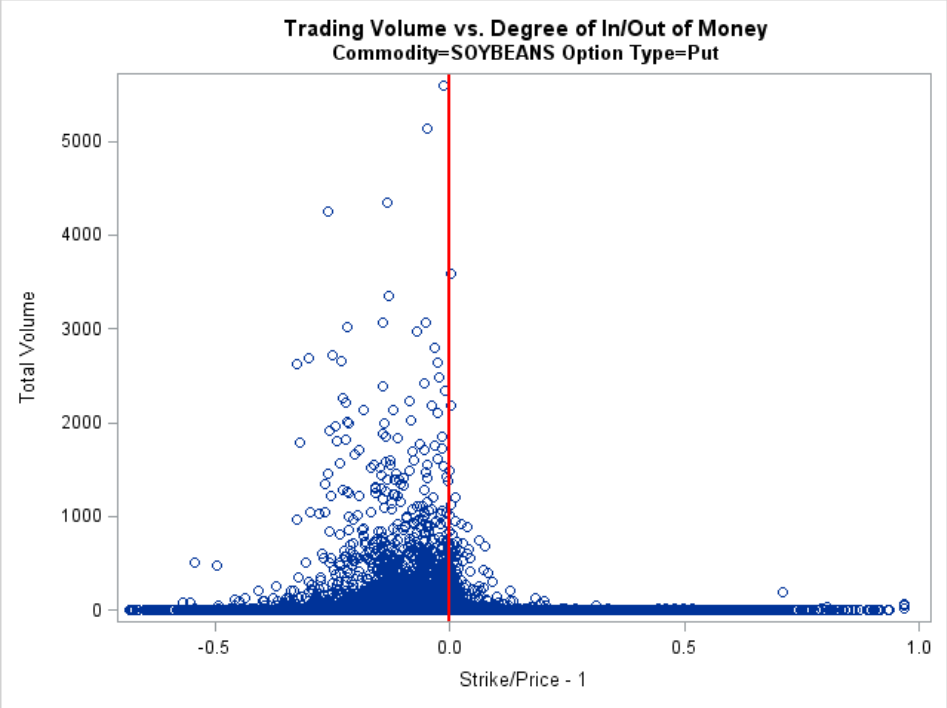
C. Rice Call Options



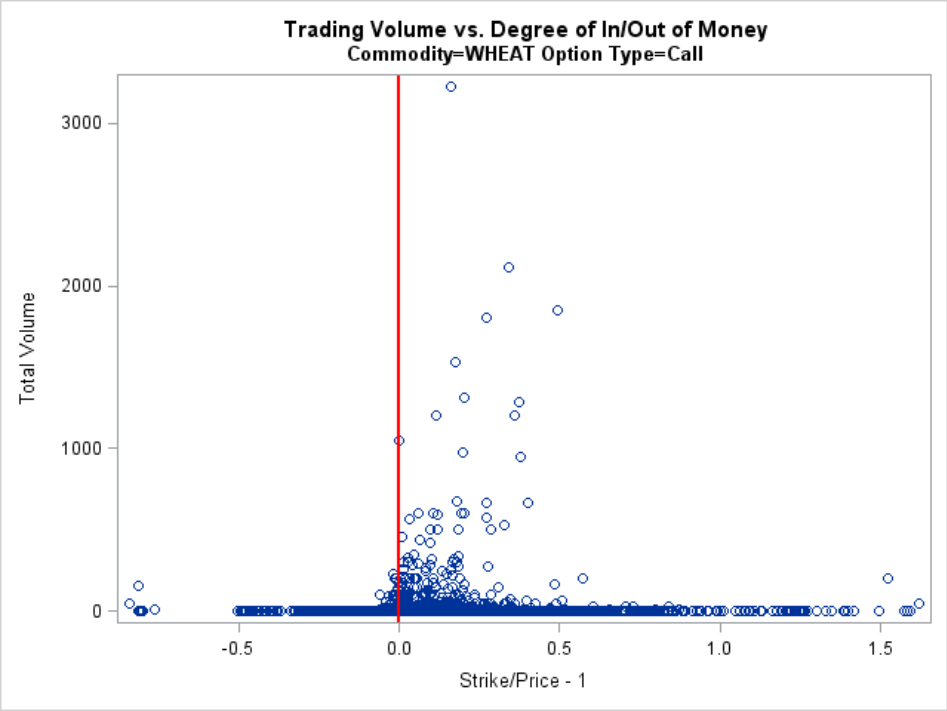
D. Soybean Call Options



E. Soybean Put Options



F. Wheat Call Options



G. Wheat Put Options

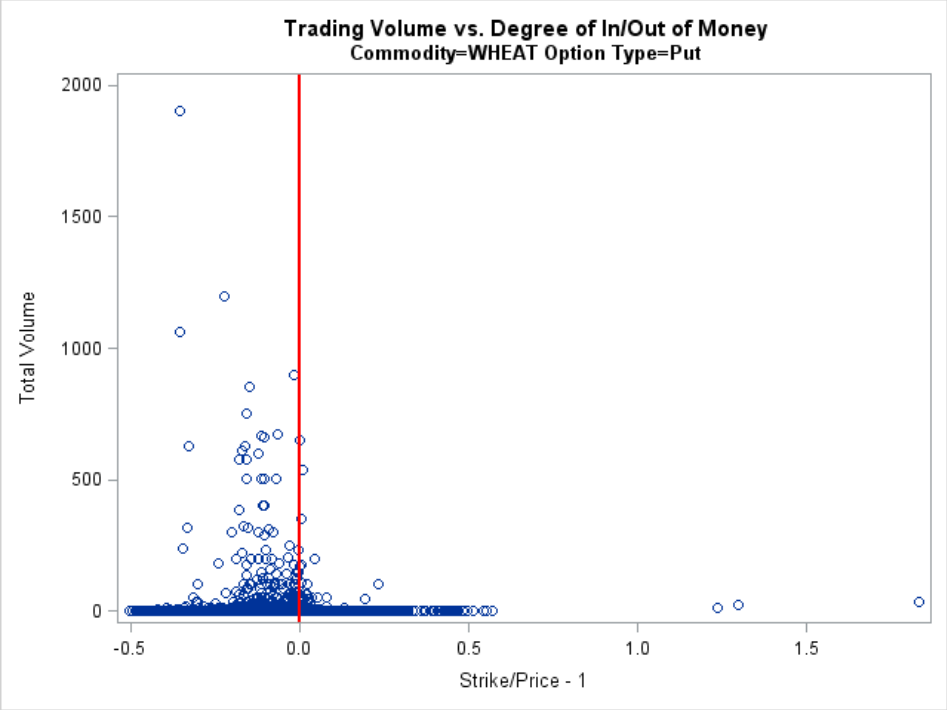
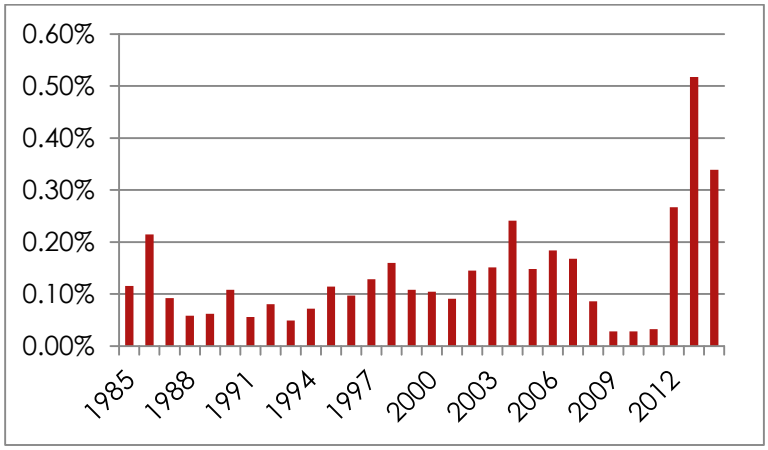
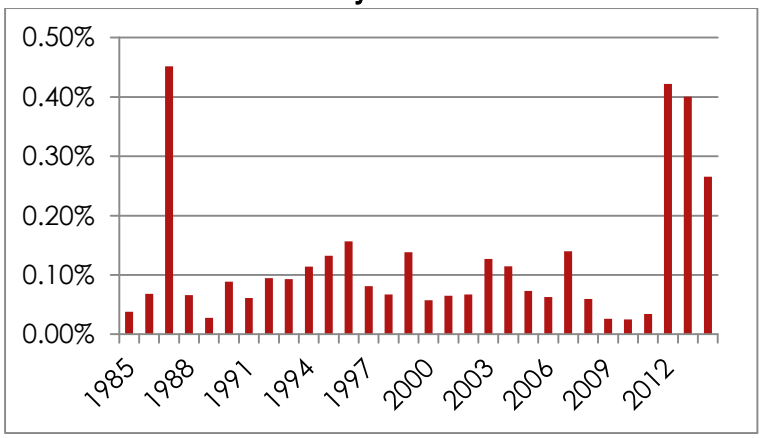


Figure. 6.3 Annual Average of Ratio of Volume of Options Traded to Volume of Futures Traded

A. Corn



B. Soybeans



C. Wheat

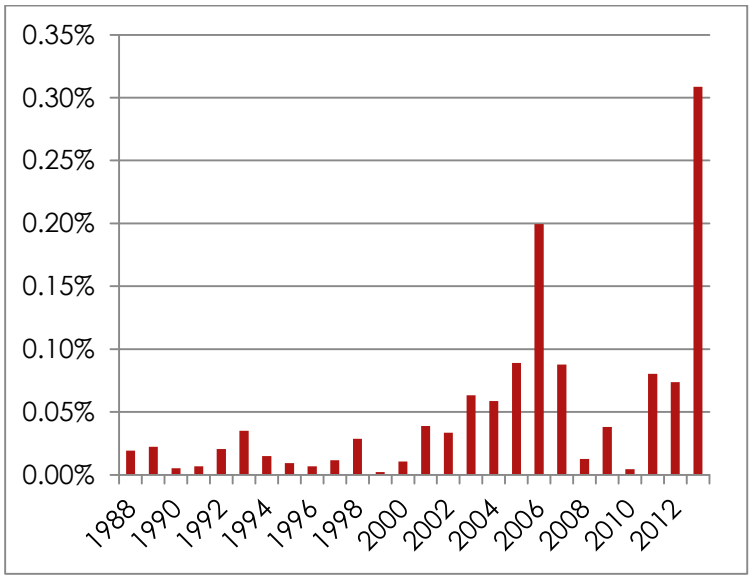
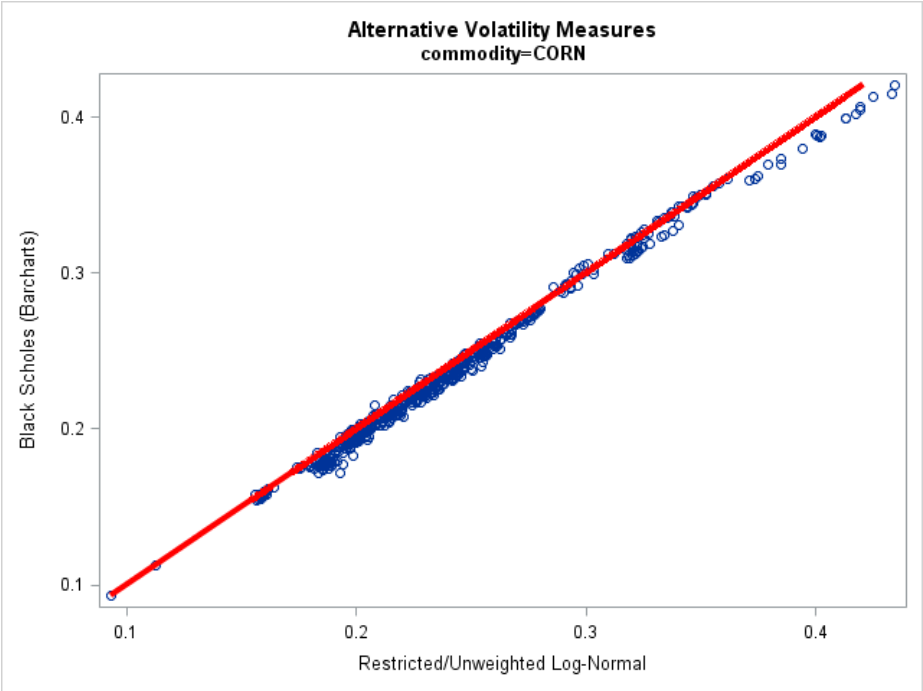
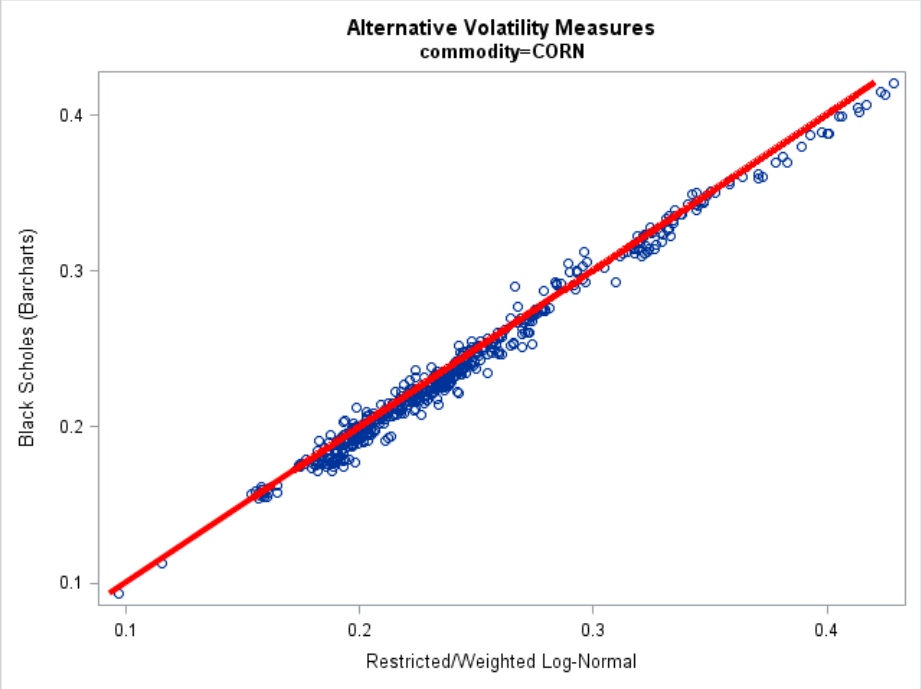


Figure 6.4. Comparisons of Alternative Volatility Measures to Conventional Black-Scholes for Corn

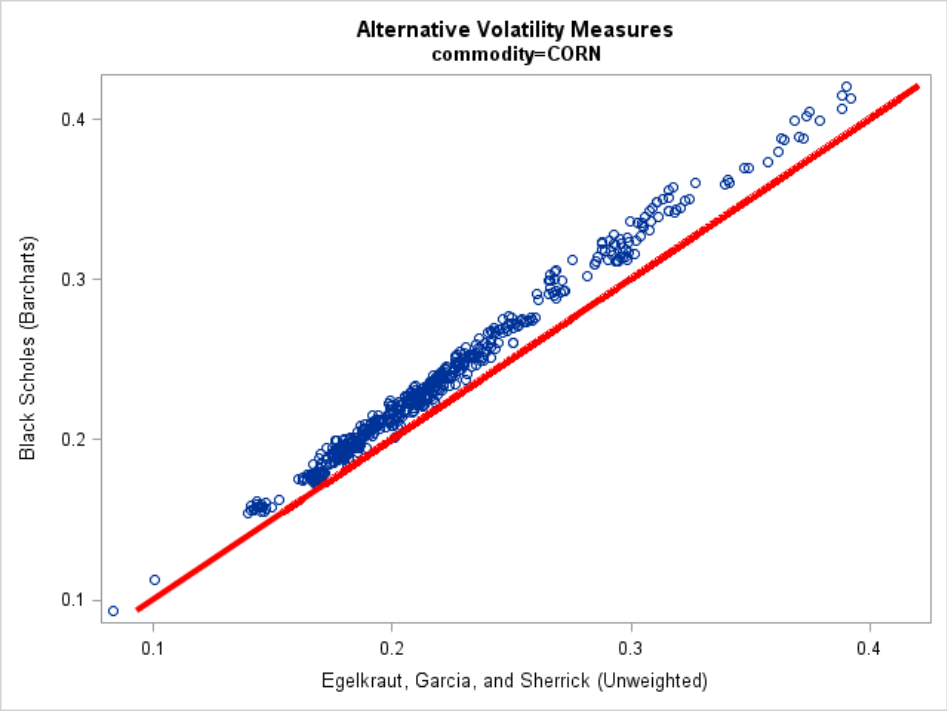
A. Restricted, Unweighted Log-Normal



B. Restricted, Weighted Log-Normal



C. Egelkraut, Garcia, and Sherrick (Unweighted)



D. Egelkraut, Garcia, and Sherrick (Weighted)

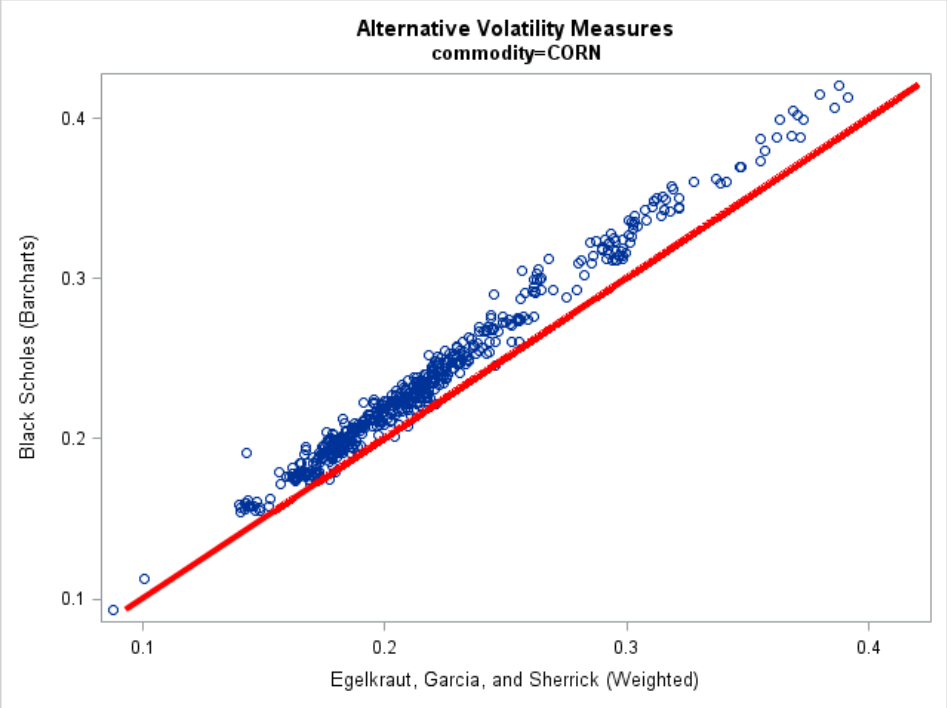
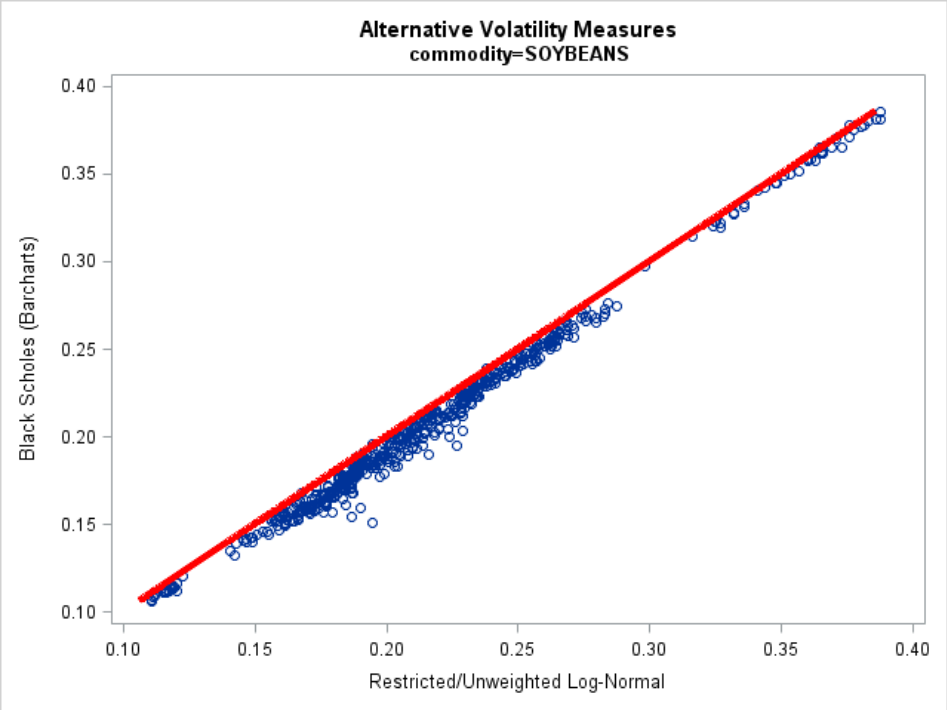
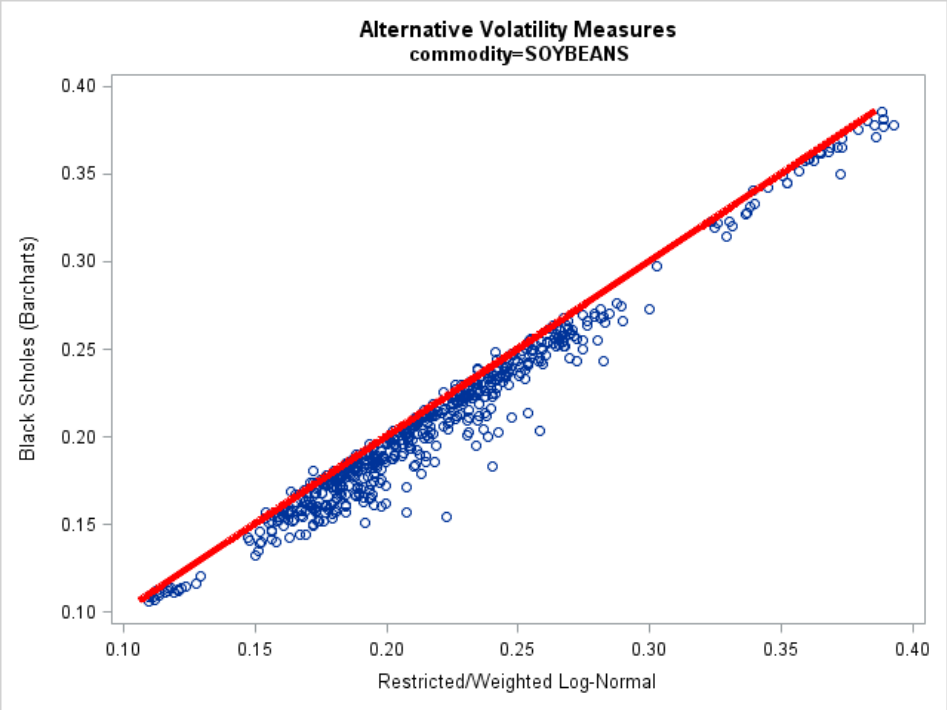


Figure 6.5. Comparisons of Alternative Volatility Measures to Conventional Black-Scholes for Soybeans

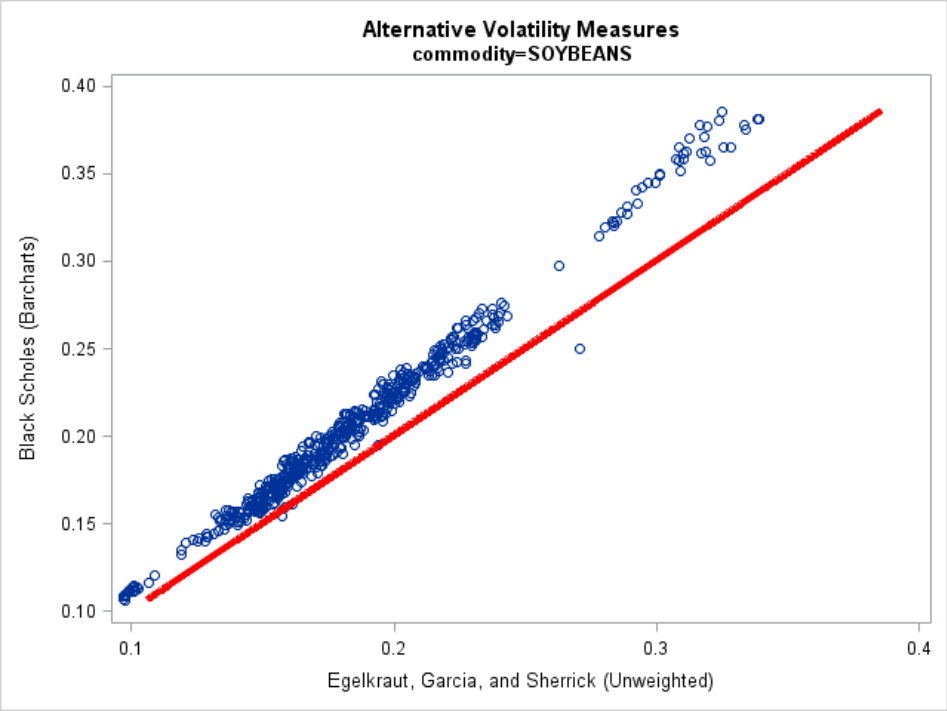
A. Restricted, Unweighted Log-Normal



B. Restricted, Weighted Log-Normal



C. Egelkraut, Garcia, and Sherrick (Unweighted)



D. Egelkraut, Garcia, and Sherrick (Weighted)

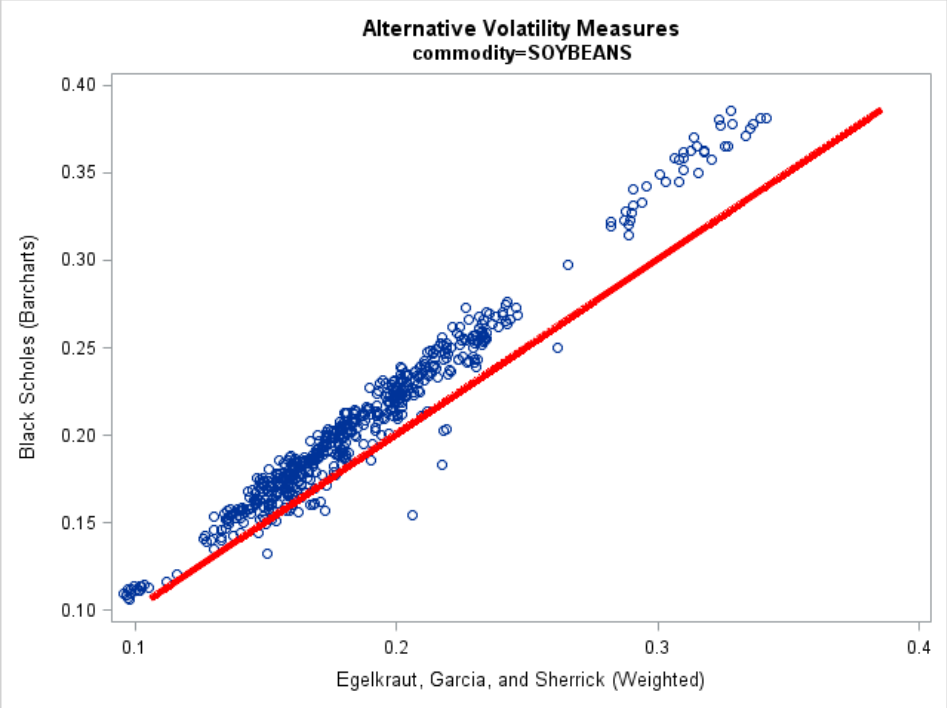
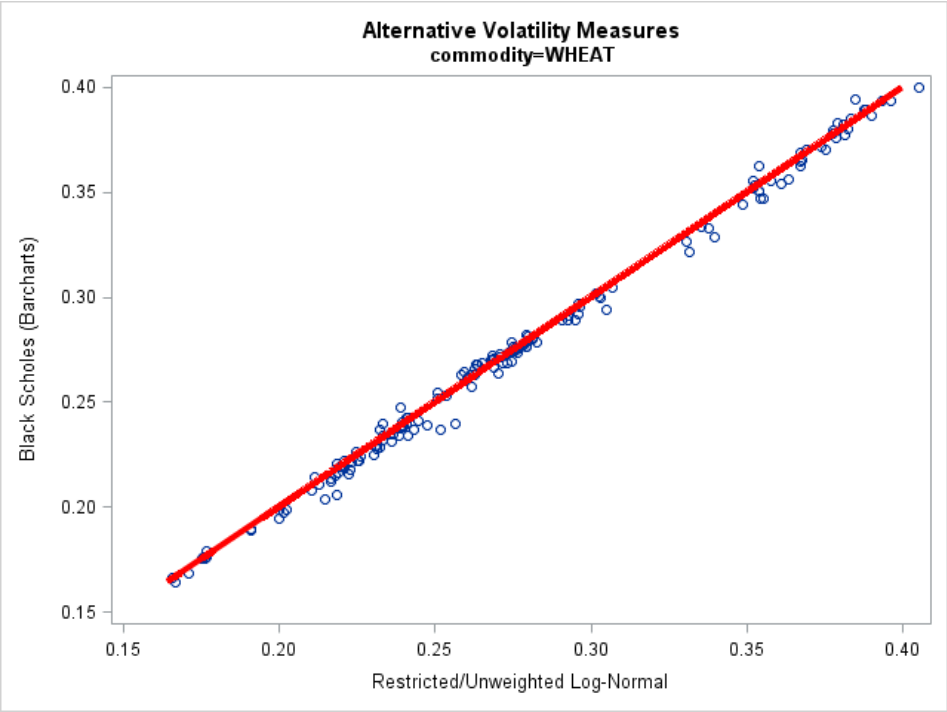
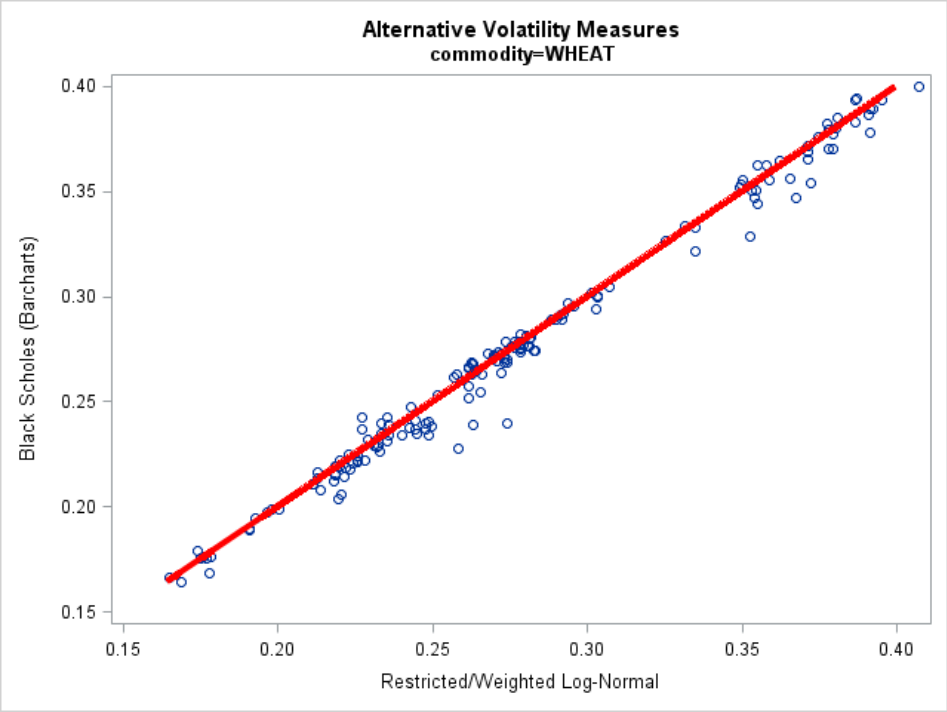


Figure 6.6. Comparisons of Alternative Volatility Measures to Conventional Black-Scholes for Wheat

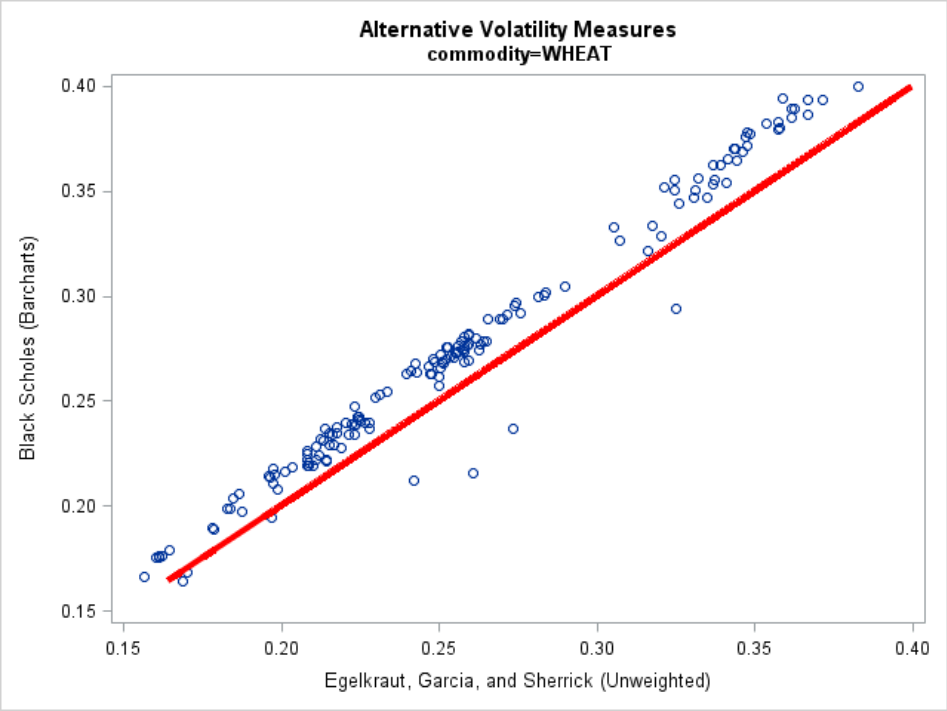
A. Restricted, Unweighted Log-Normal



B. Restricted, Weighted Log-Normal



C. Egelkraut, Garcia, and Sherrick (Unweighted)



D. Egelkraut, Garcia, and Sherrick (Weighted)

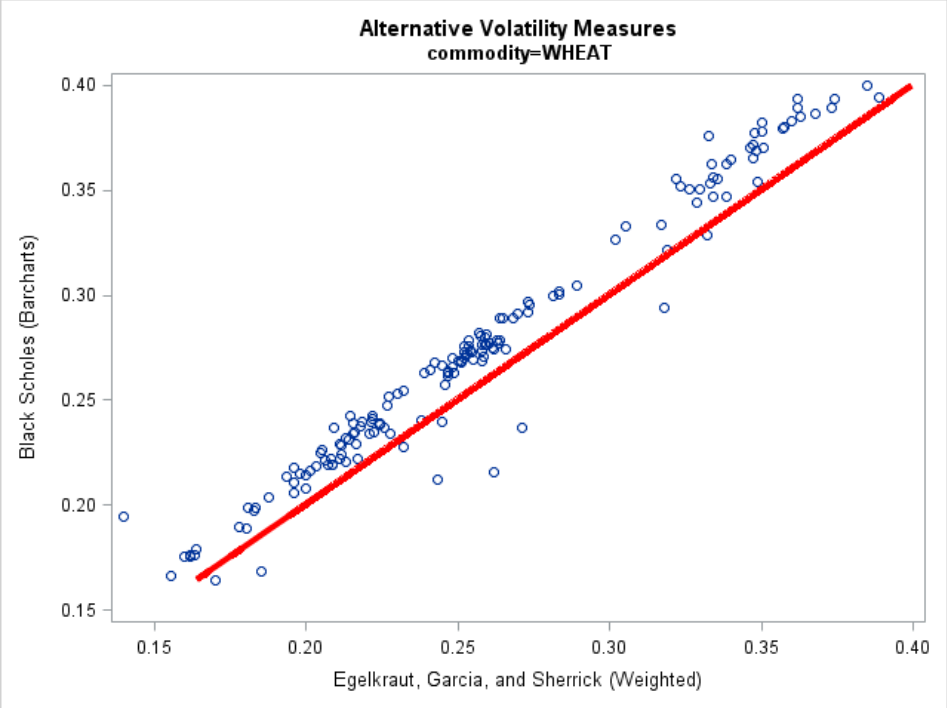
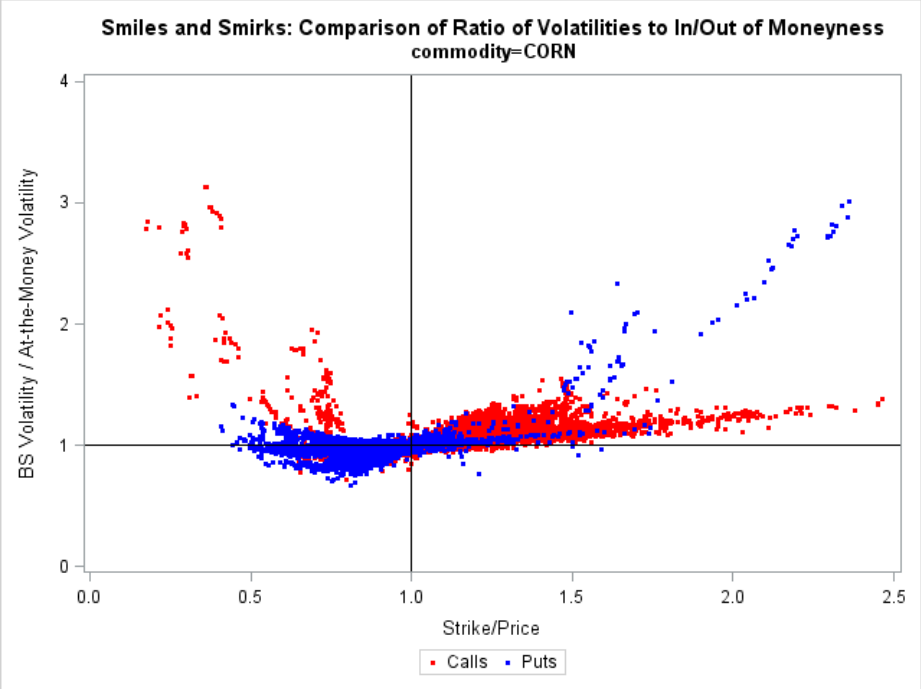
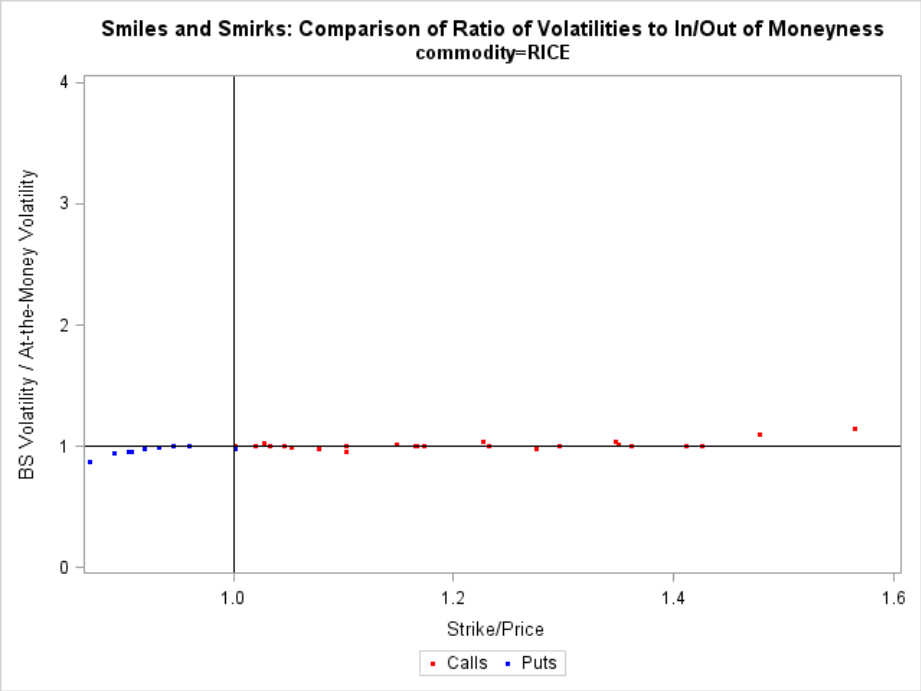


Figure. 6.7 Ratio of In/Out of Money Black Scholes Volatilities to Average Close to Money Volatilities ("Close to Money" is Average of Two Closest Puts and Two Closest Calls to Futures Price)

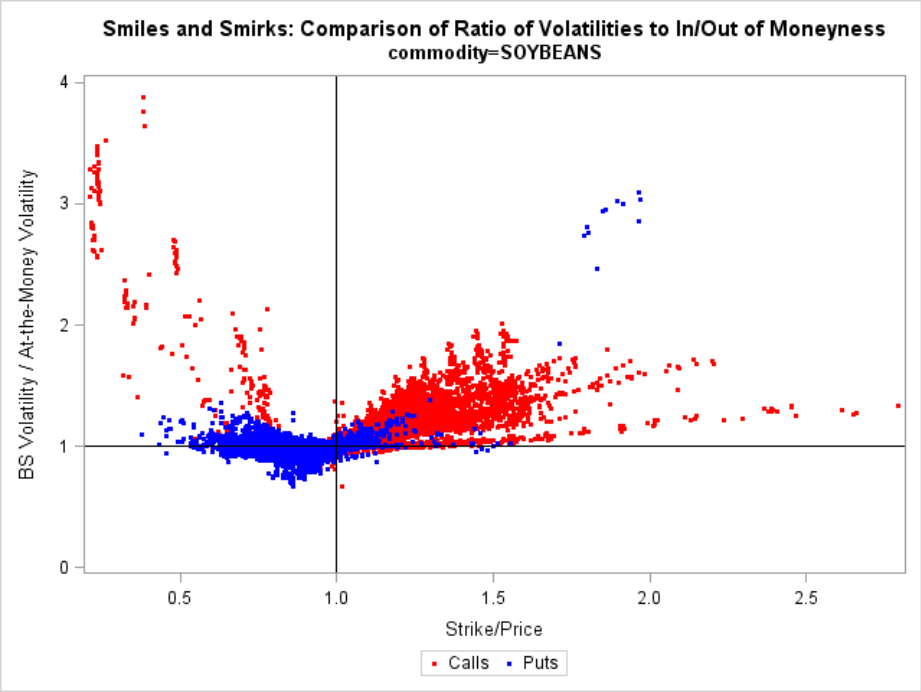
A. Corn



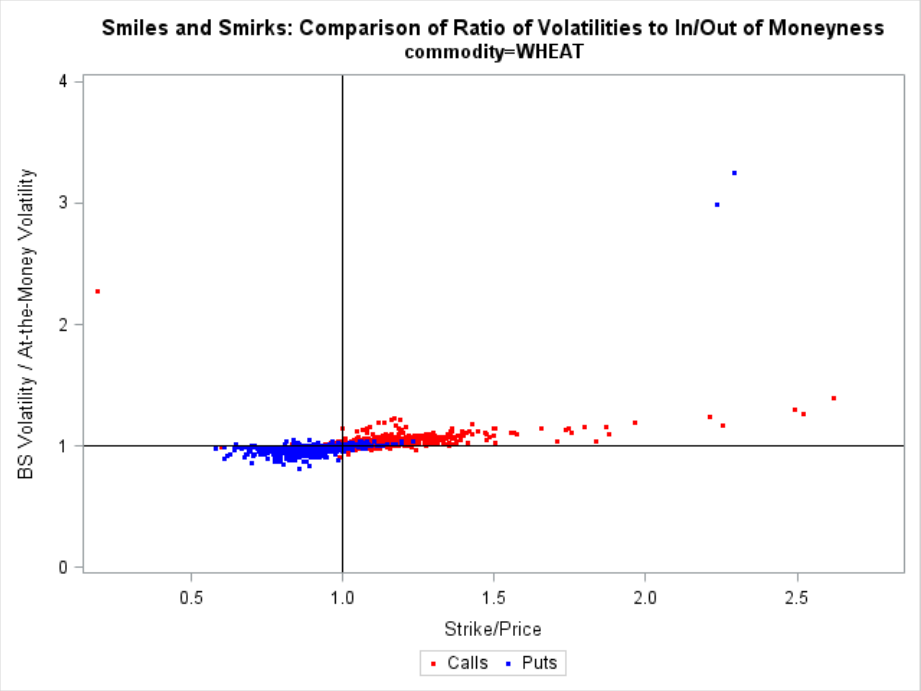
B. Rice



C. Soybeans



D. Wheat



7. Summary and recommendations

7.a. Continue to use the Black Scholes Formula for price volatility estimation

There is a vast literature related to forecasting volatility in financial and commodity markets. Overall, it seems that the BSM model is still considered the "cornerstone" option pricing model due to its ease of use and simplicity, and that it can effectively be used for calculating implied volatility (as a forecast of future volatility). However, the literature also recognizes that implied volatility from BSM has its shortcomings and it is sometimes inconsistent with price/volatility behavior observed in the market. This is the reason why numerous studies have developed alternative option pricing models and model-free approaches to estimate implied volatility. Nevertheless, there is still mixed evidence with regards to BSMs biasedness, predictive accuracy, and whether or not BSM is better than ARCH- or GARCH-type forecasts (or alternative implied volatility calculation approaches). Our analysis suggests BSM performs as well as alternatives for most crops. These results when combined with the widespread familiarity and use of BSM merits continued use. We considered several alternatives and did not find improvements in accuracy that would merit such a change. Transparency is a paramount concern in rating and the use of a complex "black-box" would diminish such transparency and potentially lead to disputes over the basis for rating parameters.

7.b. Continue to utilize a publicly available and external measure of the market price volatility

We recommend that RMA continue to utilize a publicly available and external measure of the market price volatility. The current use of Barchart as a source seems entirely reasonable and we see no reason to recommend consideration of any alternatives. In light of the prominence of the Black-Scholes volatility, we believe it is likely that any publicly available alternatives are likely to use the same methodology as that applied by Barchart.

Further, we recommend RMA's method continue to use the futures contract prior to last five days of the price discovery period (e.g. the last five trading days in February for March 15 sales closing dates). We

understand the argument to simply use the last possible trading day as it would reflect all available information. However there is risk of thin option market volume – especially for relatively thinly traded crops such as rice. Thus, we do not recommend going to single date because little additional effort is required to collect four additional days of data. The cost of doing so seems minor relative to the problems that might arise from a thin options market on a single trading day. Conversely we see little reason to extend the period to more than five days.

7.c. Continue to use the underlying futures price as a forecast of future realized price

We recommend that RMA continue to use the underlying futures price as an unbiased forecast of the future realized price. Models that allow for differences were not supported by our empirical analysis.

7.d. Avoid using thinly traded options prices in computing the implied price volatility

We recommend that RMA carefully consider trade volumes when utilizing options and futures quotes. Again, the differences associated with using volume-weighting or other methods do not merit adopting a method that is not external to RMA and that is not publicly available. However, contracts and/or days with no trades may need to be excluded in determining rating parameters.

7.e. Review and update price correlations

RMA should consider updating price yield correlations used in rating to include recent data. Further, we suggest considering more regionally specific correlation estimates. We also recommend that the issue of proper yield/price correlation relationship be monitored and updated as financial research on the topic progresses. This includes consideration of alternative distributional formulations and copula models that capture tail dependence and other important characteristics of revenue distributions. This line of research is nascent and rapidly developing. It is our opinion that no practical applications of such methods that would offer important benefits to the Federal Crop Insurance Program are currently apparent. However, the issue should be revisited as the literature advances.

7.f. We recommend a revised formula for price variability in rate simulation

The "Cost Estimator Detailed Worksheet" uses the following formulas:

$$\sigma_p = \ln(Volatility^2 + 1)$$
$$\mu_p = \ln(\overline{p}) - \frac{1}{2}\sigma_p.$$

μ_p is referred to as "LnMean", σ_p is referred to as "LnVariance", and *Volatility* is referred to as "Price Volatility Factor". This *Volatility* is simply the implied volatility obtained from the options market and adjusted for the length of the maturity period. We conclude that the mathematically correct formulas for μ_p and σ_p are:

$$\sigma_p = Volatility$$
$$\mu_p = \ln(\overline{p}) - \frac{1}{2}\sigma_p^2.$$

The alternative approach affects σ_p and indirectly affects μ_p. This modification is more consistent with the underlying BSM assumptions. Empirically, this change results in slightly higher rates that are more accurate. Empirically, the rate change is likely to be less than 0.1 percentage points.

8. References

Anderson, J.D., A. Harri and K.H. Coble, 2009 "Techniques for Multivariate Simulation from Mixed Marginal Distributions with Application to Whole Farm Revenue Simulation." *Journal of Agricultural and Resource Economics* 34(April):53-67.

Andersen, T. G. and O. Bondarenko. 2007. "Construction and interpretation of model-free implied volatility." NBER Working Paper 13449, NBER, Cambridge, MA.

Anderson, T.G. and T. Bollerslev. 1998. "Answering the Skeptics: Yes, Standard Volatility Models Do Provide Accurate Forecasts." *International Economic Review*. 39: 885-905.

Bakshi, G., C. Cao, and Z. Chen. 1997. "Empirical Performance of Alternative Pricing Models." *The J. of Finance*. 52:2003-2049.

Barnaby, G.A. 2011. "Farmers Who Cancel Their Revenue Coverage will only Save 2 Cents." Ag. Manager Publication, Kansas State Research and Extension, Kansas State University.

Barnaby, G.A. 2013a. "Did Some Wrong Assumptions Cause CBO to Overestimate Crop Insurance Costs." Ag. Manager Publcication, Kansas State Research and Extension, Kansas State University.

Barnaby, G.A. 2013b. "Professor, I am Not Buying Your Argument." Ag. Manager Publication, Kansas State Research and Extension, Kansas State University.

Barr, K.J. 2009. "The Implied Volatility Bias and Option Smile: Is there a Simple Explanation?" Unpublished PhD Dissertation, Iowa State University, Ames, IA.

Beckers, S. 1981. "Standard Deviations Implied in Option Prices as Predictors of Future Stock Price Variability." *J. of Banking and Finance*. 5: 363-382.

Black, F. 1976. "The Pricing of Commodity Contracts." *J. of Financial Economics*. 3: 167-179.

Black, F. and M. Scholes. 1973. "The Pricing of Options and Corporate Liabilities." *J. of Political Economy*. 81: 637-659.

Bollerslev, T. 1986. "Generalized Autoregressive Conditional Heteroskedasticity." *J. of Econometrics*. 31: 307-327.

Boyle, P.O. and T. Vorst. 1992. "Option Replication in Discrete Time with Transactions Costs." *The J. of Finance*. 47: 271-293.

Bozic, M., J. Newton, C.S. Thraen, and B.W. Gould. 2012a. "Parametric Bootstrap Tests for Futures Price and Implied Volatility Biases with Application to Rating Livestock Margin Insurance for Dairy Cattle." Staff Paper Series P12-9, Dept. of Applied Economics, University of Minnesota, St. Paul, MN.

Bozic, M., J. Newton, C.S. Thraen, and B.W. Gould. 2012b. "Do Volatility Smiles Matter for Pricing Asian Basket Options? The Case of Livestock Gross Margin Insurance for Dairy Cattle." Working Paper Number 12-02, Program on Dairy Markets and Policy, University of Wisconsin, Madison, WI.

Brittain, L., P. Garcia, and S.H. Irwin. 2011. "Live and Feeder Cattle Options Markets: Returns, Risk, and Volatility Forecasting." *J. of Ag. and Resource Econ*. 36(1): 28-47.

Britten-Jones, M. and A. Neuberger. 2010. "Option prices, implied price processes, and stochastic volatility." *Journal of Finance*. 55: 839–866.

Bulut, H., F. Schnapp, and K. Collins. 2011. "Volatility Factor in Concept and Practice." National Crop Insurance Services (NCIS) Special Reports, NCIS, Overland Park, KS.

Canina, L. and S. Figlewski. 1993. "The information content of implied volatility." *Rev. of Financial Studies*. 6: 659-681.

Carr, P. and L. Wu, L. 2009. "Variance risk premiums." *Review of Financial Studies*. 22: 1311–1341.

Chan, W.H., X. Cheng and J.K.W. Feng. 2010. "Forecasting Volatility: Roles of Sampling Frequency and Forecasting Horizon." *J. of Futures Markets*. 30(2): 1167-1191.

Cheng, X. and J.K.W. Fung. 2012. "The information content of model-free implied volatility." *J. of Futures Markets*. 32(8): 792-806

Christensen, B.J. and N.R. Prabhala. 1998. "The relation between implied and realized volatility." *J. of Financial Economics*. 50: 125-150.

Christofferson, P.F. and F.X. Diebold. 2000. "How relevant is volatility forecasting for financial risk management?" *Rev. Econ. Stat*. 82: 12-22.

Coble, K.H., T.O. Knight, B. J. Goodwin, M.F. Miller, and R. Rejesus, "A Comprehensive Review of APH and COMBO Rating, A report prepared for the USDA Risk Management Agency, March 17, 2010.

Day, T.E. and C.M. Lewis. 1992. "Stock Market Volatility and the Information Content of Stock Index Options." *J. of Econometrics*. 52: 267-287.

Day, T.E. and C.M. Lewis. 1993. "Forecasting Futures Market Volatility." *J. of Derivatives*. 1: 33-50.

Dumas, B., J. Fleming, and R. Whaley. 1998. "Implied Volatility Functions: Empirical Tests." *J. of Finance*. 53: 2059-2106.

Egelkraut, T.M. and P. Garcia. 2006. "Intermediate Volatility Forecasts Using Implied Forward Volatility: The Performance of Selected Agricultural Commodity Options." *J. of Ag. and Res. Econ*. 31(3): 508-528.

Egelkraut, T.M., P. Garcia, and B.J. Sherrick. 2007. "The Term Structure of Implied Forward Volatility: Recovery and Informational Content in the Corn Options Market." *American J. of Ag. Econ*. 89(1): 1-11.

Engle, R. 1982. "Autoregressive Conditional Heteroskedasticity with Estimates of the Variance of U.K. Inflation." *Econometrica*. 50: 987-1008.

Ferris, S., W. Guo, and T. Su. 2003. "Predicting Implied Volatility in Commodity Futures Options Markets." *Banking and Finance*. 1: 73-94.

Figlewski, S. 1997. "Forecasting Volatility." *Financial Markets, Institutions, and Instruments*. 6: 1-88.

Fleming, J. 1998. "The quality of market volatility forecasts implied by S&P index option prices." *J. of Empirical Finance*. 5: 317-345.

Garcia, P. and R.M. Leuthold. 2004. "A selected review of agricultural commodity futures and options markets." *European Rev. of Ag. Econ*. 31(3): 235-272.

Giot, P. 2003. "The information content of implied volatility in agricultural commodity markets." *J. of Futures Markets*. 23: 441-454.

Guo, W. and T. Su. 2004. "Factors in Implied Volatility in Corn Futures Options." Papers of the Nebraska Business and Economic Assoc. 16(1): 1-15.

Hull, J.C. 2009. *Options, Futures, and Other Derivatives*. Seventh Edition. Pearson Prentice Hall Pub., Upper Saddle, NJ.

Heston, S.L. 1993. "A Closed Form Solution for Options with Stochastic Volatility with Applications to Bond and Currency Options." *The Rev. of Financial Stud*. 6(2): 327-344.

Holt, M.T. and G. Moschini. 1992. "Alternative Measures of Risk in Commodity Supply Models: An Analysis of Sow Farrowing Decisions in the United States." *J. of Ag. and Res. Econ*. 17: 1-12.

Hull, J. and A. White. 1987. "The Pricing of Options on Assets with Stochastic Volatilities." *The J. of Finance*. 42: 281-300.

Iman, R.L. and W.J. Conover. 1982. "A distribution-free approach to inducing rank correlation among input variables." Communications in Statistics – Simulation and Computation. 11(3): 311-334.

Jiang, G. J. and Y.S. Tian. 2005. "Model free implied volatility and its information content." *Review of Financial Studies*. 18: 1305–1342.

Jorion, P. 1995. "Predicting Volatility in the Foregin Exchange Market." *The J. of Finance*. 50: 507-528.

Lamoureux, C.G. and W.D. Lastrapes. 1993. "Forecasting Stock-Return Variance: Toward and Understanding of Stochastic Implied Volatilities." *The Rev. of Fin. Stud*. 6: 293-326.

Leland, H.E. 1985. "Option Pricing and Replication with Transactions Costs." *The J. of Finance*. 40: 1283-1301.

Lubben, B. D., J.A. Jansen "Correlations Between Nebraska Crop Yields and Market Prices" December 2010. Department of Agricultural Economics, Institute of Agriculture and Natural Resources, University of Nebraska-Lincoln. http://agecon-cpanel.unl.edu/CorrelationsCropYieldsMarketPrices.pdf

Manfredo, M.R., R.M. Leuthold, and S.H. Irwin. 2001. "Forecasting Fed Cattle, Feeder Cattle, and Corn Cash Price Volatility: The Accuracy of Time Series, Implied Volatility, and Composite Approaches." *J. of Ag. and Applied Econ*. 33(3): 523-538.

Manfredo, M.R. and D.R. Sanders. 2004. "The Forecasting Performance of Implied Volatility From Live Cattle Options Contracts: Implications for Agribusiness Risk Management." *Agribusiness: An International Journal*. 20(2): 217-230.

Mayhew, S. 1995. "Implied Volatility." *Fin. Analysts Journal*. 54: 8-19.

Merton, R.C. 1973. "Theory of Rational Option Pricing." *Bell J. of Econ. and Mgt. Sci*. 41: 281-300.

Poon, S.H. and C.W.J. Granger. 2003. "Forecasting Volatility in Financial Markets: A Review." *J. of Econ. Lit*. 41:478-539.

Risk Management Agency (RMA). 2009. "RMA Revenue Rating An Analysis of the COMBO Rating Method." Unpublished Manuscript, RMA, Kansas City, MO.

Rubinstein, M. 1994. "Implied Binomial Trees." *The J. of Finance*. 49: 771-818.

Simon, D.P. 2002. "Implied Volatility Forecasts in the Grain Complex." *J. of Futures Markets*. 22: 959-981.

Shastri, K. and K. Tandon. 1986. "On the Use of European Models to Price American Options on Foreign Currency." *J. of Futures Markets*. 6: 93-108.

Szakmary, A., E. ORs, J.K. Kim, W.N. Davidson. 2003. "The predictive power of implied volatility: Evidence from 35 futures markets." *J. Banking and Finance*. 27: 2151-2175.

Taylor, S. J., P.K. Yadav, and Y. Zhang. 2010. "The information content of implied volatilities and model-free volatility expectations: Evidence from options written on individual stocks." *Journal of Banking and Finance*. 34: 871–881.

Tsiaras, L. 2010. "The forecast performance of competing implied volatility measures: the case of individual stocks." Working paper, Aarhus School of Business, Aarhus University, Aarhus, Denmark.

http://www.rma.usda.gov/pubs/2011/volatilitymethodology.pdf.

Wang, Z., S.W. Fausti, and B.A. Qasmi. 2011. "Variance Risk Premiums and Predictive Power of Alternative Forward Variances in the Corn Market." *J. of Futures Markets*. 32: 587-608.

Wilson, W.W. and H.G. Fung. 1990. "Information Content of Volatilities Implied by Option Premiums in Grain Futures Markets." *J. of Futures Markets*. 10: 13-27.

Whaley, R.E. 1986. "Valuation of American Futures Options: Theory and Empirical Tests." *The J. of Finance*. 41: 127-150.

Wu, F. and Z. Guan. 2011. "Forecasting Volatilities of Corn Futures at Distant Horizons." Selected Paper, 2010 Ag. and Applied Econ. Assoc. (AAEA) annual meetings, Denver, CO (July 25-27, 2010).

Xu, Z. 2012. "Essays on GMO effects on crop yields, the effects of pricing errors on implied volatilities and smoothing for seasonal time series with a long cycle." Unpublished PhD Dissertation, Iowa State University, Ames, IA.

Zhu, Y., S.K. Ghosh, and B.K. Goodwin. "Modeling Dependence in the Design of Whole Farm—A Copula-Based Model Approach," Presented at the 2008 AAEA Meetings.

Appendix A

SAS Code Used to Derive Volatilities

```
options linesize=max;

%let my_dir = c:\rma_pv\cmedata\;

libname my_dir "&my_dir";

proc datasets kill;


data options;format commodity $12.;  set my_dir.options2;

if Product_Code = 'CZO' then commodity="SOYBEANS";

else if Product_Code = 'LO' then commodity="CRUDE OIL";

else if Product_Code = 'PY' then commodity="CORN";

else if Product_Code = 'RRC' then commodity="RICE";

else if Product_Code = 'WZ' then commodity="WHEAT";

else commodity=" ";

trade_year=substr(compress(Trade_Date),1,4)/1;

trade_day=substr(compress(Trade_Date),7,2)/1;

trade_month=substr(compress(Trade_Date),5,2)/1;


if Product_Code = 'LO' then delete;                                        *
Dropping Crude Oil;


data options; set options;


proc sort; by commodity trade_year trade_month trade_day contract_year
contract_month ;


data fut; infile "c:\rma_pv\prices\prices2.csv" lrecl=1026 delimiter=',';

input  contract $ trade_month trade_day trade_year open h l p vol_ oi_;

proc sort; by trade_month trade_day trade_year;
```

```
data tbill; infile "c:\rma_pv\cmedata\tb2.csv" lrecl=1026 delimiter=',';
input  trade_month trade_day trade_year r;
proc sort; by trade_month trade_day trade_year;

data fut; merge fut tbill; by trade_month trade_day trade_year;


data fut;format commodity $12.; set fut;
crop = upcase(substr(contract,1,2));
cmonth=upcase(substr(contract,7,2));
Contract_year=substr(contract,3,4)/1;
if cmonth ='F' then contract_month=1;
if cmonth ='G' then contract_month=2;
if cmonth ='H' then contract_month=3;
if cmonth ='J' then contract_month=4;
if cmonth ='K' then contract_month=5;
if cmonth ='M' then contract_month=6;
if cmonth ='N' then contract_month=7;
if cmonth ='Q' then contract_month=8;
if cmonth ='U' then contract_month=9;
if cmonth ='V' then contract_month=10;
if cmonth ='X' then contract_month=11;
if cmonth ='Z' then contract_month=12;
if crop =  'S-' then commodity="SOYBEANS";
if crop =  'CL' then commodity="CRUDE OIL";
if crop =  'C-' then commodity="CORN";
if crop =  'RR' then commodity="RICE";
```

```
if crop =  'W-' then commodity="WHEAT";


proc sort data=fut; by commodity trade_year trade_month trade_day
contract_year contract_month ;

proc sort data=options; by commodity trade_year trade_month trade_day
contract_year contract_month ;


data all; merge options fut;  by commodity trade_year trade_month trade_day
contract_year contract_month ;

if strike_price <=0 then delete;

if p <= 0 then delete;


if trade_year <= 2007 then do;

if commodity="CORN" then strike_price=strike_price/10;

if commodity="SOYBEANS" then strike_price=strike_price/10;

if commodity="WHEAT" then strike_price=strike_price/10;

if commodity="RICE" then strike_price=strike_price/1000;

end;


if trade_year > 2008 and commodity="RICE" then do;

strike_price=strike_price/100; end;

ratio=Strike_Price/p;

if trade_year = 2008 then do;

scale=Strike_Price/p;

if p=. then scale=.;

if scale = . then sc=.;

else if scale < 5 then sc=1;

else if scale < 50 then sc=10;

else if scale < 500 then sc=100;

else if scale < 5000 then sc=1000;
```

```
else sc=.;
strike_price=strike_price/sc;
ratio=Strike_Price/p;
end;

rate=r/100;


term =(mdy(contract_month, 15, contract_year) - mdy(trade_month, trade_day,
trade_year))/365.25;
if put_call='C' then type="call";
else if put_call='P' then type="put";
else type=" ";
settlement=settlement/10;
*if total_volume<=0 then delete;
no_trade=(total_volume<=0);
if settlement<=0 then delete;
r_ = (1+rate/4)**4-1;

proc sort; by commodity trade_year trade_month trade_day contract_year
contract_month put_call strike_price;


data all; set all;
if commodity="RICE" then settlement=settlement/100;
strike_price_1=lag(strike_price);
settlement_1=lag(settlement);
if commodity=lag(commodity) and
trade_year=lag(trade_year) and
```

```
trade_month =lag(trade_month) and
trade_day =lag(trade_day) and
contract_year =lag(contract_year) and
contract_month =lag(contract_month) and
put_call=lag(put_call) then do;
        if put_call="P" then do;
                if strike_price>strike_price_1 and settlement<settlement_1 then z=1;
*delete;
        end;
        else if put_call="C" then do;
                if strike_price>strike_price_1 and settlement>settlement_1 then z=2;
*delete;
        end;
end;
t+1;


proc expand data=all out=all method = none;
 id t;
 convert z = z_lead1  / transformout=(lead 1);
run;

data all; set all;
keep commodity trade_month trade_year trade_day contract_month
contract_year type Open_Interest strike_price
settlement p p_ total_volume implied_volatility rate r__ r_
Settlement_Cabinet_Indicator term vol_ oi_ put_call z z_lead1 ;

proc freq; table Settlement_Cabinet_Indicator / missing; run;
```

```
data all; set all;
if commodity="CORN" then do;
        if contract_month ne 12 then delete;
        if trade_month ne 2 then delete;
        if trade_year ne contract_year then delete;
end;
if commodity="SOYBEANS" then do;
        if contract_month ne 11 then delete;
        if trade_month ne 2 then delete;
        if trade_year ne contract_year then delete;
end;
if commodity="WHEAT" then do;
        if contract_month ne 7 then delete;
        if trade_month ne 8 and trade_month ne 9 then delete;
        if trade_month = 8 and trade_day < 15 then delete;
        if trade_month = 9 and trade_day > 14 then delete;
        if trade_year ne (contract_year-1) then delete;
end;
if commodity="RICE" then do;
        if contract_month ne 11 then delete;
        if trade_month ne 1 and trade_month ne 2 then delete;
        if trade_month = 1 and trade_day < 15 then delete;
        if trade_month = 2 and trade_day > 14 then delete;
        if trade_year ne contract_year then delete;
end;
if commodity="RICE" then delete;
data volume; set all;  keep commodity trade_year trade_month trade_day
contract_year contract_month put_call vol_ total_volume strike_price p;
```

```
data volume; set volume;
io_mo=(strike_price/p-1);
data my_dir.volume; set volume; run;


data iv; set all; if z>0 then delete; if z_lead1 > 0 then delete;
rate=r_;
p_=exp(-rate*term)*p;
if commodity=" " then delete;

proc sort; by commodity trade_year trade_month trade_day contract_year
contract_month put_call;

data iv; set iv;
if commodity ne lag(commodity) or
trade_year ne lag(trade_year) or
trade_month ne lag(trade_month) or
trade_day ne lag(trade_day) or
contract_year ne lag(contract_year) or
contract_month ne lag(contract_month) then do; q+1;end;
run;


proc means noprint data=iv sum n;
var total_volume; by q;
output out=volsum sum=volsum n=no;

data iv; merge iv volsum; by q;
```

```
                    if volsum<=0 then delete;
        * Delete if no volume on any strikes;

                                        *if Settlement_Cabinet_Indicator="C"
then delete;* Delete CAB settlements (no trades);

                    if total_volume<=0 then delete;
        * Delete contracts with zero trades ;

proc means noprint data=iv n;

var total_volume; by q;

output out=volsum2 n=no2;


data iv; merge iv volsum2; by q;

        if no2 <=4 then delete;                                        *
Delete any day with less than 4 strikes;

data iv; set iv;

if q ne lag(q) then do; qq+1; end;

k_=strike_price/p;

mo=abs(k_-1);tt+1;


proc fcmp outlib=sasuser.funcs.options;

  function blkschc(strike_price, term, p_, rate, volty);

    return(blkshclprc(strike_price, term, p_, rate, volty));

  endsub;

    function bsvoltyc(settlement, strike_price,term, p_, rate);

          array opts[5] initial abconv relconv maxiter status

            (.3 .001 1.0e-6 500 -1);

  iv = solve("blkschc", opts, settlement, strike_price, term, p_, rate, .);

  return(iv);

  endsub;

run;
```

```
proc fcmp outlib=sasuser.funcs.options;
  function blkschp(strike_price, term, p_, rate, volty);
    return(blkshptprc(strike_price, term, p_, rate, volty));
  endsub;
    function bsvoltyp(settlement, strike_price,term, p_, rate);
          array opts[5] initial abconv relconv maxiter status
            (.3 .001 1.0e-6 500 -1);
  iv = solve("blkschp", opts, settlement, strike_price, term, p_, rate, .);
  return(iv);
  endsub;
run;


options cmplib=sasuser.funcs;
data iv; set iv;                                                    * Here
we set limits on implied volatility 1% < IV < 400% ;
   if put_call="C" then do;
          l_l = blkshclprc(strike_price, term, p_, rate, 0.01);
          u_l = blkshclprc(strike_price, term, p_, rate, 400);
          if settlement > l_l and settlement < u_l then iv1= bsvoltyc(settlement,
strike_price, term, p_, rate);
          else iv1=.;
          end;
else if put_call="P" then do;
   l_l = blkshptprc(strike_price, term, p_, rate, 0.01);
          u_l = blkshptprc(strike_price, term, p_, rate, 400);
          if settlement > l_l and settlement < u_l then iv1= bsvoltyp(settlement,
strike_price, term, p_, rate);
          else iv1=.;
          end;
else iv1=.;
```

```
run;

proc means; by commodity; var iv1;

data iv; set iv; q=qq;


data qq; set iv;
if q=lag(q) then delete;
keep q
commodity
trade_year
trade_month
trade_day
contract_year
contract_month ;

data iv_; set iv;
zz=(put_call="C");
if volsum>0 then w=total_volume/volsum;i_+1;


proc sort; by q type mo;

data iv_; set iv_;
if q ne lag(q) or type ne lag(type) then rz=1;
else do; rz+1;end;
```

```
proc means noprint mean data=iv_(where=(rz<=2));
var iv1; by q;
output out=conv_ mean=iv_1; run;

data iv_; merge iv_ conv_; by q; id=q;

proc freq data=iv; table q / missing; table commodity / missing;  table
q*commodity / missing; run;



data my_dir.iv__; set iv_; run;

/* Standard log-normal with constrained mean */

proc nlmixed data=iv_ maxit=500; by q;
parms s=0.30;
bounds s>0.01, s< 4;
d1=(log(p/strike_price)+.5*s**2*term)/(s*sqrt(term));
d2=d1-s*sqrt(term);
pf= (zz*(exp(-rate*term)*(p*cdf('NORMAL',d1,0,1)-
strike_price*cdf('NORMAL',d2,0,1))) +
    (1-zz)*(exp(-rate*term)*(strike_price*cdf('NORMAL',-d2,0,1)-p*cdf('NORMAL',-
d1,0,1))));
CL=-(sum((settlement-pf)**2));
model settlement ~ general(CL);
ods output ParameterEstimates=parm2;
run;

data parm2; set parm2; rename Estimate=s_2; rename StandardError=SE_2;
data parm2; set parm2; keep q s_2 SE_2;
```

```
/* Standard log-normal with constrained mean, weighted by volume */

proc nlmixed data=iv_ maxit=500;by q;
parms s=0.30;
bounds s>0.01, s< 4;
d1=(log(p/strike_price)+.5*s**2*term)/(s*sqrt(term));
d2=d1-s*sqrt(term);
pf= (zz*(exp(-rate*term)*(p*cdf('NORMAL',d1,0,1)-
strike_price*cdf('NORMAL',d2,0,1))) +
    (1-zz)*(exp(-rate*term)*(strike_price*cdf('NORMAL',-d2,0,1)-p*cdf('NORMAL',-
d1,0,1))));
CL=-(sum(w*((settlement-pf)**2)));
model settlement ~ general(CL);
ods output ParameterEstimates=parm3;
run;
data parm3; set parm3; rename Estimate=s_3; rename StandardError=SE_3;
data parm3; set parm3; keep q s_3 SE_3;

/*  Check of code to verify match (Compared to R Options Package Results) */
/*
proc nlp data=iv_ pcov phes maxit=500;
min CL;
parms s=0.30;
bounds s>0.01, s< 4;
d1=(log(p/strike_price)+.5*s**2*term)/(s*sqrt(term));
d2=d1-s*sqrt(term);
```

```
pf= (zz*(exp(-rate*term)*(p*cdf('NORMAL',d1,0,1)-
strike_price*cdf('NORMAL',d2,0,1))) +

    (1-zz)*(exp(-rate*term)*(strike_price*cdf('NORMAL',-d2,0,1)-p*cdf('NORMAL',-
d1,0,1))));

CL=sum((settlement-pf)**2);

run;*/



/*
                                                                    * This calculates IV at
single observation to check code (Compared to R Options Package Results);

proc nlp data=iv_(where=(i_=25)) pcov phes maxit=500;

min CL;

parms s=0.30;

bounds s>0.01, s< 4;

d1=(log(p/strike_price)+.5*s**2*term)/(s*sqrt(term));

d2=d1-s*sqrt(term);

pf= (zz*(exp(-rate*term)*(p*cdf('NORMAL',d1,0,1)-
strike_price*cdf('NORMAL',d2,0,1))) +

    (1-zz)*(exp(-rate*term)*(strike_price*cdf('NORMAL',-d2,0,1)-p*cdf('NORMAL',-
d1,0,1))));

CL=sum((settlement-pf)**2);

run;

*/


                                                                    * This is Sherrick et al.
calculation;


%macro bs;


%do q=1 %to 1289;
```

```
%put &q;

data iv0; set iv_; if q ne &q then delete;

proc iml;
use iv0;
read all var {id} into id;
read all var {settlement} into settlement;
read all var {p} into p;
read all var {p_} into p_;
read all var {strike_price} into strike;
read all var {term} into term;
read all var {rate} into rate;
read all var{implied_volatility} into iv0;
read all var {type} into type;
read all var {put_call} into put_call;
read all var {Settlement_Cabinet_Indicator} into cab;
read all var {total_volume} into vol;
read all var {open_interest} into oi;
read all var {iv1} into iv_;
read all var {iv_1} into iv_1;
read all var {w} into w;
close iv0;
zz=(put_call="C");

start bsss(x) global(_settlement, _strike, _term, _p, _rate, _zz, m, s, v, settlement,
strike, term, p, rate, zz, iv_1);

do i=1 to nrow(strike);
```

```
_p=p[i];
_strike=strike[i];
_settlement=settlement[i];
_term=term[i];
_p=p[i];
_rate=rate[i];
_zz=zz[i];


m=x[1]; v=x[2];
start ker1(y) global(_settlement, _strike, _term, _p, _rate, _zz, m, v, settlement,
strike, term, p, rate, zz, iv_1);

        pdf=pdf('LOGNORMAL',y,log(m**2/sqrt(v+m**2)),sqrt(log(v/m**2+1)))#(y-
_strike);
                return(pdf);
                finish ker1;
                start ker2(y) global(_settlement, _strike, _term, _p, _rate, _zz, m, v,
settlement, strike, term, p, rate, zz, iv_1);

        pdf=pdf('LOGNORMAL',y,log(m**2/sqrt(v+m**2)),sqrt(log(v/m**2+1)))#(_str
ike-y);
                return(pdf);
                finish ker2;


il2=0||_strike;
il1=_strike||.P;


if _zz=1 then do;z2=0;
```

```
call quad(z1,"ker1",il1) eps=1E-3 msg = "NO" cycles=1 peak=_strike;*print
_settlement, _strike, z1, _p;end;

else do;z1=0;

call quad(z2,"ker2",il2) eps=1E-3 msg = "NO" cycles=1 peak=_strike;*print
_settlement, _strike, z2, _p;end;


f=_zz#(_settlement - exp(-_rate#_term)#z1) + (1-_zz)#(_settlement - exp(-
_rate#_term)#z2);

if i=1 then res=f##2;

else res=res+f##2;*print res;

end;

return(res);

finish bsss;

m_=sum(p)/nrow(p);ivv=sum(iv_1)/nrow(iv_1); if ivv<=0.05 then ivv=0.3; print ivv;

v_=(ivv*m_)**2;print v_; print m_;

x=m_||v_;print x;

optn=j(1,11,.);optn[1]=0;optn[2]=1;con={0 0.01, . .};

call NLPNRR(rc,xres,"bsss",x,optn,con);m=xres[1];v=xres[2]; s=sqrt(log(v/m**2+1));

print s,m,v;

optimum=bsss(xres);print optimum;

CALL NLPFDD (crit, grad, hess, "bsss", xres) ; if min(eigval(hess))>0 then do;

cov=sqrt(diag(inv(hess))); print cov;end;

else do; cov={. ., . .};end;


sol1=&q||xres||cov[1,1]||cov[2,2]||s||m_||m||p[1,1]||optimum||rc;print
sol1;

varnames={"q" "b1_4" "b2_4" "se1_4" "se2_4" "s_4" "m__4act" "m_4" "p_4" "opt_4"
"rc_4"};

create sol1 from sol1 [colname=varNames];

append from sol1;
```

```
run;
proc append data=sol1 base=sol_1;


                                                    * This is Sherrick et al.
calculation weighted by volume;
proc iml;
use iv0;
read all var {id} into id;
read all var {settlement} into settlement;
read all var {p} into p;
read all var {p_} into p_;
read all var {strike_price} into strike;
read all var {term} into term;
read all var {rate} into rate;
read all var{implied_volatility} into iv0;
read all var {type} into type;
read all var {put_call} into put_call;
read all var {Settlement_Cabinet_Indicator} into cab;
read all var {total_volume} into vol;
read all var {open_interest} into oi;
read all var {iv1} into iv_;
read all var {w} into w;
read all var {iv_1} into iv_1;
close iv0;
zz=(put_call="C");

start bsss(x) global(_settlement, _strike, _term, _p, _rate, _zz, m, s, v, settlement,
strike, term, p, rate, zz, w, _w, iv_1);
```

```
do i=1 to nrow(strike);


_p=p[i];
_strike=strike[i];
_settlement=settlement[i];
_term=term[i];
_p=p[i];
_rate=rate[i];
_zz=zz[i];
_w=w[i];




m=x[1]; v=x[2];
start ker1(y) global(_settlement, _strike, _term, _p, _rate, _zz, m, v, settlement,
strike, term, p, rate, zz, w, _w, iv_1);

        pdf=pdf('LOGNORMAL',y,log(m**2/sqrt(v+m**2)),sqrt(log(v/m**2+1)))#(y-
_strike);
                return(pdf);
                finish ker1;
                start ker2(y) global(_settlement, _strike, _term, _p, _rate, _zz, m, v,
settlement, strike, term, p, rate, zz, w, _w, iv_1);

        pdf=pdf('LOGNORMAL',y,log(m**2/sqrt(v+m**2)),sqrt(log(v/m**2+1)))#(_str
ike-y);
                return(pdf);
                finish ker2;


il2=0||_strike;
il1=_strike||.P;
```

```
if _zz=1 then do;z2=0;

call quad(z1,"ker1",il1) eps=1E-3 msg = "NO" cycles=1 peak=_strike;*print
_settlement, _strike, z1, _p;end;

else do;z1=0;

call quad(z2,"ker2",il2) eps=1E-3 msg = "NO" cycles=1 peak=_strike;*print
_settlement, _strike, z2, _p;end;


f=_zz#(_settlement - exp(-_rate#_term)#z1) + (1-_zz)#(_settlement - exp(-
_rate#_term)#z2);

if i=1 then res=_w#f##2;

else res=res+_w#f##2;*print res;

end;

return(res);

finish bsss;

m_=sum(p)/nrow(p);ivv=sum(iv_1)/nrow(iv_1);  if ivv<=0.05 then ivv=0.3;  print ivv;

v_=(ivv*m_)**2;print v_; print m_;

x=m_||v_;print x;

optn=j(1,11,.);optn[1]=0;optn[2]=1;con={0 0.01, . .};

call NLPNRR(rc,xres,"bsss",x,optn,con);m=xres[1];v=xres[2]; s=sqrt(log(v/m**2+1));

print s,m,v;

optimum=bsss(xres);print optimum;

CALL NLPFDD (crit, grad, hess, "bsss", xres) ; print grad; print hess;

if min(eigval(hess))>0 then do;

cov=sqrt(diag(inv(hess))); print cov;end;

else do; cov={. ., . .};end;

sol2=&q||xres||cov[1,1]||cov[2,2]||s||m_||m||p[1,1]||optimum||rc;print
sol2;

varnames={"q" "b1_5" "b2_5" "se1_5" "se2_5" "s_5" "m__5act" "m_5" "p_5" "opt_5"
"rc_5"};
```

```
create sol2 from sol2 [colname=varNames];
append from sol2;
run;
proc append data=sol2 base=sol_2;

%end;

%mend;

%bs;
proc print data=qq;
proc print data=conv_;
proc print data=sol_1;
proc print data=sol_2;
proc print data=parm2;
proc print data=parm3;
proc print data=my_dir.model_free;
proc print data=volsum_;


data all; merge qq conv_ sol_1 sol_2 parm2 parm3 my_dir.model_free volsum; by
q; drop _:;
proc sort; by commodity contract_year;
data all; merge all my_dir.realized_iv;by commodity contract_year;
if commodity=" " then delete; if contract_year=. then delete; if iv_1=. then
delete;drop _:;

proc sort; by q;
```

```
proc print; var q commodity contract_year contract_month iv_1 s_2 s_3 s_4 s_5
mfvol realized_vol m_:;

run;

data my_dir.all_iv_1; set all;


run;
```