Extreme Correlation Between Daily Basis Returns of Local Corn Markets in North Carolina

Yixing Zheng, Department of Agricultural and Applied Economics, Virginia Tech, yixing19@vt.edu;
Ford Ramsey, Department of Agricultural and Applied Economics, Virginia Tech, aframsey@vt.edu

Selected Paper prepared for presentation at the 2022 Agricultural & Applied Economics Association Annual Meeting, Anaheim, CA; July 31-August 2

Extreme Correlation Between Daily Basis Returns of Local Corn Markets in North Carolina *†

Abstract

This paper examines the correlation of corn basis returns among six local market pairs in North Carolina during extreme events, employing a formal statistical approach based on extreme value theory without assuming normality. By focusing on basis returns, this study provides a unique perspective on local market dynamics, independent of national aggregate market movements. Results indicate that correlations among the six market pairs strengthen during crises compared to booming periods. An asymmetric dependence between the right and left tails suggests that extreme events influence local price-setting behavior in distinct ways. Among the examined pairs, Candor/Cofield exhibits the strongest extreme correlation. These findings offer valuable implications for market participants, including farmers and processors, by enhancing risk management and supply chain resilience. The high correlation observed in certain markets highlights a reliance on shared infrastructure, underscoring the need for improved coordination among grain storage operators, logistics companies, and transportation firms. Policymakers may also leverage these insights to design targeted stabilization measures. Additionally, the non-normal and asymmetric tail distributions challenge traditional models that assume normal residuals. This paper contributes to a deeper understanding of price dependencies and risks in regional agricultural markets, equipping stakeholders with practical tools to navigate market uncertainties and make informed decisions.

Keywords: Extreme Correlation; Corn Basis; Tail risk; Extreme Value Theory

JEL Codes: G11; G13; Q14; Q19

*Note: Preliminary draft. Do not cite.

1 Introduction

The measurement and identification of dependencies between different markets are crucial components in various financial applications, including portfolio management, risk assessment, pricing, and hedging. Understanding the co-movements between markets is particularly important for developing effective diversification strategies. When markets are less correlated, investors have better opportunities to diversify their portfolios by selecting assets from these markets. Conversely, a high correlation among markets might suggest diminishing returns from diversification efforts. Agricultural commodities are one of the most important asset classes and are often included in investors' portfolios due to their diversification benefits and the ability to hedge against inflation (Erb and Harvey, 2006).

Agricultural commodities are produced over a wide spatial area and are therefore expensive to transport in relation to their entire value. These features result in a complex set of spatial price linkages that are frequently investigated to reveal insight into market performance and market relationships. The literature on price linkages has been well-established and there is a wide range of issues relating to price transmission, regional price linkages and spatial market integration, market performance, and efficiency (Guney, Goodwin, and Riquelme, 2019; Hussain and Li, 2021; Sephton, 2003).

Many early empirical studies use simple price correlation analysis to evaluate market integration and market dependencies (Mohendru, 1937; Stigler and Sherwin, 1985). These studies usually assume normal distributions of prices such that a simple constant linear correlation was computed using the whole price series between different markets. Under the normality assumption, the tails and dependencies of returns are symmetric which can be captured by linear correlation. However, it is known that agricultural commodity price return distributions may exhibit skewness, asymmetries, and fat tails which are non-normal features (e.g., Geman, 2005). Constant correlation measures the overall strength of the association, but may not reflect the true dependence relationships. This shortcoming is particularly evident when extreme behaviors are to be modeled or inferred from data. This limitation highlights the importance of asymptotic dependence, which refers to the behavior of extreme values in statistical distributions and their relationships. In extreme events, the relationship between variables may be more pronounced or behave differently than what is suggested by a constant correlation measure. Recognizing and accounting for asymptotic dependence is essential for a more accurate and comprehensive analysis of market integration and dependencies, particularly in scenarios involving extreme market behaviors.

Agricultural commodity prices are highly volatile, and agricultural marketing is inherently risky. For instance, corn prices decreased by a staggering 55% in the second half of 2008 due to the financial crisis. There was a significant increase in corn prices in June and July 2012 because of a severe drought in the United States. The COVID-19 pandemic also heavily affected corn prices in 2019. More recently, there has been a sharp increase in corn prices due to fertilizer shortages related to the Russia-Ukraine conflict. These extreme events such as droughts, pests, disease outbreaks, and geopolitical tensions have a low probability of occurrence but can result in high-impact losses with major implications for producer profitability.

In light of these challenges, Extreme value theory (EVT) is well-suited to test whether the correlation between commodities markets changes during extreme events. Extreme value theory is a technique that studies the behavior of extreme observations. It concentrates directly on the tails of the distribution rather than the whole distribution and provides asymptotic results that hold for a wide range of parametric distributions other than the normal distribution. Therefore, it can better describe extremely volatile markets compared to standard approaches that assume normal distributions. EVT also has the advantage of capturing the asymptotic tail distribution with very few parameters and does not depend on the exact cumulative distribution function of returns. Moreover, the left and right tails of the distribution can be tackled independently so that the skewed distribution can be modeled appropriately. EVT could potentially perform better than other approaches in terms of prediction of unexpected changes in extremes.

This paper investigates the extreme correlation of corn basis returns among six local market pairs in North Carolina using a robust statistical framework based on Extreme Value Theory, which does not assume normality. By focusing on basis returns, our study provides a distinctive lens to analyze local market dynamics independently of the national aggregate market's movements. This focus is critical because extreme tail dependence highlights how markets behave under volatile conditions, providing valuable insights for various stakeholders. Understanding these extreme co-movements offers practical implications for farmers, agribusinesses, and other market participants. For instance, farmers might mitigate risks by diversifying their sales to less correlated markets during periods of high price volatility. Similarly, buyers such as exporters, ethanol plants, or feedlots could utilize this information to optimize their sourcing strategies, negotiate better pricing, or anticipate regional price shifts. Grain storage operators, logistics firms, and transportation companies could also leverage these findings to improve inventory coordination, resource allocation, and infrastructure investment strategies. Moreover, the analysis of asymmetric dependencies sheds light on how local price-setting behaviors may vary during extreme events, providing market analysts and decision-makers with actionable insights into market microstructure.

Furthermore, extreme values are often dependent because meteorological events occur on a spatial scale that includes multiple locations. Due to this dependence, information from one location may be useful in estimating the extreme value behavior at another location. Quantifying the dependence across locations is informative, as extreme events occurring over different regions can have more serious implications for risk assessment and market stability than localized extreme events. Harmonization across financial markets is a major issue for risk analysts, particularly when risk or investment is spread across various commodities within a single portfolio. To understand the overall level of risk entailed by a specific portfolio, questions about the extent of dependence between extremes of a number of series become unavoidable. Examining dependence structures sheds light on the degree of integration among regional markets. Strong co-movements might indicate high integration, whereas weak dependencies suggest market segmentation. This insight is essential for understanding regional market stability and resilience to shocks.

This paper is organized as follows. Section 1 provides a background on the importance of understanding dependencies between markets and introduces the application of Extreme Value Theory (EVT) in assessing extreme correlations in local agricultural markets. Section 2 reviews relevant literature on EVT applications in financial and agricultural markets, highlighting the gap this study addresses. Section 3 details the Modeling of Extreme Values, covering the marginal distribution of tails and dependence structures as well as the Estimation Procedure. Section 4 describes the Data used for the study, and Section 5 presents the Results of extreme correlation analysis across six local corn market pairs in North Carolina. Section 6 discusses Model Checking to validate model robustness, and Section 7 provides Conclusion and Discussion, exploring the implications.

2 Literature Review

Extreme Value Theory (EVT) has been widely applied across disciplines, including climatology, hydrology, and biology, with notable use in financial markets due to the significant economic impacts of extreme events. Early applications of EVT in finance primarily focused on testing distributional hypotheses, as highlighted in the early works (e.g., Koedijk, Schafgans, and De Vries, 1990; Lux, 2000; Vilasuso and Katz, 2000). More recently, EVT methods have been used to estimate risk measures such as Value-at-Risk (VaR) and Expected Shortfall (ES), highlighting their growing relevance in risk management.

One key area of EVT application in finance is the modeling of dependence and extreme correlations across markets. This research can be divided into two main categories. The first category examines dependencies between countries within the same financial sector. For instance, Longin and Solnik (2001) used EVT and threshold exceedance methods to analyze equity returns, finding that extreme correlations between the US and other G5 countries are more related to market trends than to volatility and that these correlations display asymmetric dependence. Similarly, Bekiros and Georgoutsos (2008) studied extreme return correlations between seven Asia-Pacific stock markets and the US, concluding that US investors may benefit from diversification by including Asian assets. However, questions remain about the reliability of using correlation coefficients to measure extreme dependence. For example, Poon, Rockinger, and Tawn (2004) used bivariate EVT to analyze the upper tail dependence of daily returns on G5 stock indices, finding that only 13 out of 84 country pairs exhibited non-zero coefficients, challenging the assumption of asymptotic dependence in most cases. Likewise, Zeevi and Mashal (2002) state that the co-movement of G5 national equity markets does not follow the "correlation-based" Gaussian dependence structure. Furthermore, Christoffersen et al. (2012) compared the tail dependencies between the equity markets in emerging countries and developed countries using a dynamic asymmetric copula model with EVT. They find the tail dependence increased over time for 16 developed countries and 17 emerging countries but the increase is relatively low in emerging countries.

The second category of research focuses on tail dependencies across different financial activities. For example, Hussain and Li (2021) applied EVT and a dynamic copula approach to examine the dependence structure between oil futures and other commodity futures in China, revealing significant right-tail dependence, particularly with oil and other commodities. Additionally, Sukcharoen and Leatham (2016) combined GARCH and EVT methods to explore the daily stock price indices for ten sectors in the US, showing that sector diversification benefits have diminished over time.

For agricultural commodities, EVT has primarily been applied to estimate commodity price risk through measures such as VaR and ES. For instance, Fretheim and Kristiansen (2015) estimated the tail shape parameters of future price return distributions for commodities like corn, wheat, and soybeans, finding no systematic increase in tail risk. Similarly, van Oordt, Stork, and de Vries (2021) utilized EVT to estimate VaR levels in agricultural commodities, using back-testing methods to validate accuracy. Bayesian approaches to EVT have also been explored. Park and Maples (2018) applied a Bayesian hierarchical model to analyze serial dependence in extreme corn futures price changes, estimating conditional VaR to evaluate model performance. Singvejsakul, Chaiboonsri, and Sriboonchitta (2021) used Bayesian inference and Newton-optimal methods to identify extreme maximum points in future prices for US commodities, including cocoa, coffee, and corn.

Despite the extensive research on EVT in financial markets, there is limited empirical work on local-level co-movements and dependencies across agricultural commodity markets during extreme events. This study aims to address this gap by examining extreme correlations in local corn markets in North Carolina, thereby contributing to the understanding of dependence structures within regional agricultural markets.

3 Modeling of Extreme Values

Extreme Value Theory (EVT) is distinguished by its focus on quantifying the stochastic behavior of processes at unusually large or small levels. Unlike traditional statistical approaches that concentrate on the central part of a distribution, EVT aims to estimate the probability of events that are more extreme than any previously observed. It is well-suited to model the extreme market risk inherent to agricultural commodity markets with fat tails.

Classical Extreme Value Theory has two primary models: the block maxima model based on generalized extreme value distribution (GEV), and the threshold model based on generalized Pareto distribution (GPD). The block maximum method involves dividing the data into non-overlapping blocks of equal size (e.g., months or years) and extracting the maximum value from each block. These maxima are then asymptotically approaches to the Generalized Extreme Value (GEV) distribution. A drawback of the block maxima (minima) method is that it leads to a waste of information as it only considers the extreme values within each block and ignores other potentially valuable data points. For example, if one block happens to contain more extreme events than another, it is wasteful.

To analyze extreme market events, however, we are not only interested in maxima or minima but also the behavior of large movements over or under some threshold. If an entire time series of observations is available, then better use is made of the data by avoiding altogether the procedure of blocking. The threshold method focuses on all observations that exceed a specified high threshold. The exceedances have a corresponding approximate distribution within the generalized Pareto Distribution (GPD). The GPD models a distribution of excess over (under) the threshold by considering the use of all maximum (minimum) log return values higher (lower) than the threshold rather than the collection of block maxima (minima) values within some periods. Efficiency is improved because all observations that are extreme in the sense of exceeding a high threshold can be used in model fitting. Therefore, the threshold exceedance model is employed in this paper.

EVT involves analyzing the tail of the distribution where such extreme events reside. EVT can also be extended to model the dependence structure between the tails of multiple distributions. Multivariate extreme value models enable the calculation of the probability of simultaneously extreme events. This requires an asymptotic approximation for the dependence at extreme levels and the extreme value behavior of each individual series. Therefore two modeling aspects are involved: the marginal distributions' tails and extreme observations' dependence structure.

3.1 Modeling Marginal Distribution of Tails

Let R represent a stationary series of returns on a corn basis, with F_R denoting the cumulative distribution function of R. The lower and upper endpoints of the associated density function are denoted by l and u, respectively.[1]

In the context of extreme value theory, we focus on the tails of the distribution, particularly through the threshold exceedance model in the paper. Here, the extreme returns are identified as those exceeding a specific threshold μ. Positive exceedances correspond to all observations of R that are greater than the threshold μ. Conversely, negative exceedances, which can be derived from positive exceedances by considering symmetry, correspond to observations of R that are less than μ. A return R is greater than μ with probability p, and less than μ with probability $1 - p$. The probability p is linked to the threshold μ and the distribution F_R by the relation $p = 1 - F_R(\mu)$.

When $x > \mu$, defining the right tail of the distribution, the cumulative distribution function for exceedances over μ, denoted by F_R^μ, is expressed as:

$$F_R^\mu(x) = \frac{F_R(x) - F_R(\mu)}{1 - F_R(\mu)} \quad \text{for } x > \mu$$

If the distribution of the whole F_R is known, then the distribution of exceedances can be determined. However, in most cases, the exact distribution of returns F_R is not well-defined, which makes the distribution of return exceedances uncertain. This is where the asymptotic behavior of exceedances becomes critical. As established by the works of Balkema and De Haan (1974) and Pickands III (1975), the Generalized Pareto Distribution (GPD) serves as the appropriate model for the distribution of exceedances as the threshold μ approaches the upper endpoint u of the distribution. In that sense, EVT can be thought of as a central limit theory for extremes. The GPD, G_R is expressed by:

[1] In the case of a normally distributed variable, $l = -\infty$ and $u = +\infty$

$$G_R(x) = 1 - \left(1 + \frac{\xi(x-\mu)}{\sigma}\right)^{-\frac{1}{\xi}}$$

where:

- ξ is the shape parameter, indicating the tail behavior:
  - If $\xi > 0$, the distribution has a power-declining tail (fat-tailed distribution).
  - If $\xi = 0$, the distribution has an exponentially declining tail (thin-tailed distribution).
  - If $\xi < 0$, the distribution has a finite endpoint, indicating no tail (finite distribution).
- σ is the scale parameter, which depends on the threshold μ and the underlying distribution F_R.
- μ is the threshold, and x represents the observations of returns.

In the context of this study, we model the exceedances of basis returns from four local corn markets, denoted as R_1, R_2, R_3, and R_4. Each market's threshold is denoted by μ_i, for i=1,2,3,4. The tail distribution of each basis return, represented by $F_{R_i}^{\mu_i}$ is modeled as:

$$F_{R_i}^{\mu_i}(x_i) = (1-p_i) + p_i \cdot G_{R_i}^{\mu_i}(x_i) = 1 - p_i\left(1 + \frac{\xi_i(x_i-\mu_i)}{\sigma_i}\right)^{-1/\xi_i}, \quad i = 1,2,3,4 \qquad (1)$$

which indicates that a basis return R_i either falls below the threshold μ_i(not the tail), with a probability of $1-p_i$, or it exceeds the threshold(tail) and follows Generalized Pareto Distribution (GPD) G_{R_i} with probability p_i. In other words, returns that do not surpass the threshold contribute to the model only by indicating that they are below the threshold, without providing their specific values.[2] For negative exceedances, where $R_i < \mu_i$ (a negative threshold), the same principle applies, with the roles of the tail and the threshold adjusted accordingly.

3.2 Modeling Dependence Structure

To model the tails of the basis series for the six pairs of corn markets, we employ a bivariate threshold exceedance method. The joint distribution of two basis exceedances can be expressed in terms of their respective marginal distributions and the dependence structure between them.

We begin by deriving univariate marginal distribution as the $F_{R_i}^{\mu_i}(x_i)$ which are asymptotic to generalized Pareto distributions for the tails of each basis return. The model F_R^u of the bivariate distribution of basis exceedances is given by

$$F_R^u(x_1, x_2) = exp(-D_l(-1/logF_{R_1}^{u_1}(x_1), -1/logF_{R_2}^{u_2}(x_2))) \tag{2}$$

where D_l represents the dependence function between the two basis exceedances.

Dependence Characterizations: Dependence structures can be characterized as independent, perfectly dependent, asymptotically independent, or asymptotically dependent. For positively related variables, large values occur together more often if asymptotically dependent, and less often if asymptotically independent. As variables approach their upper limits, the probability of both being high goes to zero for asymptotically independent variables but remains nonzero for asymptotically dependent ones. Consequently, extreme values occur simultaneously only for asymptotically dependent variables.

Asymptotic dependence between two variables R_1 and R_2 can be characterized by:

[2] When constructing the likelihood function, a return below the threshold is treated as censored at that level.

$$P(z) = \Pr\left(F_{R_1} > z \mid F_{R_2} > z\right)$$

where $P(z)$ is nonzero as z approaches the endpoint of F. Otherwise, the two variables R_1, and R_2 are asymptotically independent.

Two nonparametric measures can be used to examine the asymptotic dependence of two variables: χ and $\bar{\chi}$

$$\chi = \lim_{z\to\infty} \frac{\Pr(R_1 > z, R_2 > z)}{\Pr(R_2 > z)} > 0$$

where $\chi > 0$ indicates asymptotic dependence, meaning that extreme values of R_1 and R_2 are likely to occur together even as we move into the tails of the distribution. Conversely, if $\chi = 0$, the variables are asymptotically independent, which means that the probability of simultaneous extreme values approaches zero as we move further into the joint tail of the distribution. In fact, $\chi = 0$ indicates that the two variables become independent of each other in the limit as we move further into the joint tail of the distribution. However, at finite levels of extremes, strong dependence can still exist.

The degree of dependence in the tails can also be assessed using the alternative measure $\bar{\chi}$:

$$\bar{\chi} = \lim_{z\to\infty} \frac{2\log\Pr(R_2 > z)}{\log\Pr(R_2 > z, R_1 > z)} - 1$$

The measure χ lies within $0 \leq \chi \leq 1$, and $\bar{\chi}$ lies within $-1 \leq \bar{\chi} \leq 1$. The variables R_1 and R_2 are asymptotically independent when $\chi = 0$, with the degree of dependence characterized by $\bar{\chi}$. If $\bar{\chi} = 1$, the variables are asymptotically dependent, and the degree of dependence is given by χ.

3.2.1 Parametric Dependence Models

Parametric methods are employed to estimate the joint tail of the distribution, as they exist for both asymptotically dependent and asymptotically independent cases. Unlike nonparametric methods, which are typically restricted to asymptotic dependence, parametric models can address both scenarios. We consider a parametric model for each class to emphasize the differences between asymptotically dependent and asymptotically independent structures. Studies by Ledford and Tawn (1996) and Dupuis and Tawn (2001) suggest that when the model is appropriately chosen, the exact form of the dependence structure becomes relatively less crucial.

An Asymptotically Dependent Model: Logistic For cases where the components of the bivariate distribution of extreme returns are asymptotically dependent, the dependence function is often modeled using the logistic function proposed by Gumbel (1961). The dependence function for the logistic model is given by:

$$D_l(y_1, y_2, ..., y_q) = (y_1^{-1/\alpha} + y_2^{-1/\alpha} + ... + y_q^{-1/\alpha})^{\alpha} \tag{3}$$

where $y_i = -\frac{1}{\log F_{R_i}^{u_i}(x_i)}$ and $0 < \alpha < 1$. The parameter α determines the level of dependence between extreme basis returns, with larger values of α indicating weaker dependence. In the bivariate case ($q = 2$) in our study, this parameter α is related to the correlation coefficient of extremes ρ via the formula $\rho = 1 - \alpha^2$ (Tiago de Oliveira, 1973; Ledford and Tawn, 1996).

The bivariate distribution of basis exceedances is then given by:

$$\begin{aligned} F_R^u(x_1, x_2) &= \exp\left(-\left(y_1^{-1/\alpha} + y_2^{-1/\alpha}\right)^{\alpha}\right) \\ &= \exp\left(-\left(\left(-\frac{1}{\log F_{R_1}^{u_1}(x_1)}\right)^{-1/\alpha} + \left(-\frac{1}{\log F_{R_2}^{u_2}(x_2)}\right)^{-1/\alpha}\right)^{\alpha}\right). \end{aligned} \tag{4}$$

Given thresholds u_1 and u_2, the bivariate distribution of return exceedance is characterized by seven parameters $(p_1, p_2; \xi_1, \xi_2; \sigma_1, \sigma_2; \alpha)$.

An Asymptotically Independent Model: Gaussian For asymptotically independent series, the Gaussian joint-tail model proposed by Bortot, Coles, and Tawn (2000) is commonly employed. The joint distribution $F_R^u(x_1, x_2)$ of the two basis return exceedances is given by:

$$F_R^u(x_1, x_2) = \Phi_2\left(\Phi^{-1}\left(F_{R_1}^{u_1}(x_1)\right), \Phi^{-1}\left(F_{R_2}^{u_2}(x_2)\right); \rho\right) \tag{5}$$

$\Phi_2(\cdot, \cdot; \rho)$ represents the bivariate normal (Gaussian) distribution function, which describes the joint probability of the two standardized variables $\Phi^{-1}(F_{R_1}^{u_1}(x_1))$ and $\Phi^{-1}(F_{R_2}^{u_2}(x_2))$, with correlation ρ. The function Φ_2 incorporates the Gaussian dependence structure into the joint distribution. $\Phi^{-1}(\cdot)$ is the inverse of the standard normal cumulative distribution function (quantile function), which transforms the marginal distribution values into the standard normal scale.

ρ is the linear correlation coefficient between the two variables, representing the degree of dependence between the two markets. As you move into the extreme tails of the distribution (as x_1 and x_2 become large), ρ typically approaches zero, indicating that the variables are asymptotically independent, meaning there is weak or no dependence between the extreme events in the two markets.

3.3 Estimation Procedure

1. Test Corn Basis Return Stationarity, Trend, and Seasonality: Before applying Extreme Value Theory (EVT) to model the tails, it is essential to ensure that the series satisfies the stationarity assumption. If the series exhibits non-stationary behavior due to trends or seasonality, it should be removed before proceeding with tail modeling. To test whether the series is stationary, we apply standard statistical tests: the Augmented Dickey-Fuller (ADF) test, the Kwiatkowski-Phillips-Schmidt-Shin (KPSS) test, and the Ng-Perron test. These tests help detect whether the series contains unit roots, indicating non-stationarity. We also test the seasonality for the basis return. If a significant trend or seasonality is present, the series will be detrended or seasonally adjusted using techniques such as seasonal differencing or Seasonal Decomposition of Time Series (STL).

2. Threshold Selection: Choosing an appropriate threshold $u$is a critical step in EVT, as it determines what is considered an extreme event. There is a trade-off between bias and variance when choosing the threshold. Setting u too low will include too many observations as extremes, leading to biased estimates while setting it too high will reduce the number of extreme observations, increasing variance and reducing estimation efficiency. Following the approach of Longin and Solnik (2001), we consider multiple thresholds (e.g., 90th, 92nd, 95th, and 97th percentiles) for the left and right tails to estimate the tail index and to better capture the difference between the left tails and right tails. Additionally, bootstrap simulations are employed to select the optimal threshold that minimizes the asymptotic mean squared error (AMSE) criterion of the Hill estimator for each series. A Mean Residual Life (MRL) plot will also be generated to assess the choice of threshold visually. The MRL plot displays the average excess over different threshold values, aiding in identifying the point at which the plot stabilizes, indicating a suitable threshold for modeling.

3. Test for Data Clustering: In stationary series, extreme values may tend to cluster, which can complicate the application of EVT as it introduces dependence between extreme events for one variable. The extremal index, introduced by Smith and Weissman (1994), provides a means to quantify the degree of clustering. The extremal index θ is interpreted as the inverse of the mean cluster size and ranges from 0 to 1. If $\theta = 1$ is close to 1, it suggests that extreme values are largely independent. If $\theta = 1$is significantly less than 1, it suggests clustering, with lower values indicating stronger dependence among extreme events and larger mean cluster sizes. We begin by selecting a threshold to define extreme values and then calculate the extremal index. If clustering is detected, we apply a GARCH (Generalized Autoregressive Conditional Heteroskedasticity) model to capture volatility clustering in the data, and EVT is applied to the residuals.

4. Estimate Generalized Pareto Distribution (GPD) Parameters: Once the threshold has been selected, the parameters of the Generalized Pareto Distribution (GPD) are estimated using Maximum Likelihood Estimation (MLE). The GPD models the tail distribution of the basis return series, capturing the behavior of the exceedances above the selected threshold. The tail distributions for each of the four corn markets are then estimated using the fitted GPD parameters, providing insights into the extreme value behavior.

5. Plot χ and $\bar{\chi}$ to Inform the Choice of Dependence Function: The dependence structure between extreme events in two markets is characterized by the nonparametric measures χ and $\bar{\chi}$. These measures help determine whether the markets exhibit asymptotic dependence or independence. The asymptotic dependence between two series is characterized by $\bar{\chi} = 1$, which occurs if and only if $\chi = c > 0$. If $\bar{\chi} < 1$, the series are asymptotically independent, meaning $\chi = 0$. While much of the literature often assumes asymptotic dependence by default (e.g., Longin and Solnik, 2001), it is important to recognize that extreme events in one market do not necessarily coincide with those in another (e.g., Poon, Rockinger, and Tawn, 2004; Singh, Allen, and Powell, 2017; Ergen, 2014). To address this, Poon, Rockinger, and Tawn (2004) recommended first using nonparametric measures to test for dependence between two returns before fitting model parameters for the tails. To test for asymptotic dependence or independence, we follow the approach suggested by Poon, Rockinger, and Tawn (2004), which involves testing the null hypothesis $\bar{\chi} = 1$. If the test rejects this hypothesis, the series is considered asymptotically independent. we will initially employ Chi plots and Chi-bar plots to assess dependence between the markets, then fit the appropriate model parameters based on the test results.

6. Calculate Dependence Based on the Selected Dependence Function: Based on the dependence function informed by the χ and $\bar{\chi}$ plots, we then calculate the dependence between the six pairs of corn basis markets. If two markets are asymptotically dependent, the dependence for 6 pairs of corn basis markets will be modeled using the logistic function. 6 pairs corn basis markets

7. Model Checking and Goodness of Fit: To test the accuracy of the model, we perform several diagnostic checks:

- **Tail of Marginal**: It helps evaluate how well a fitted marginal distribution, such as a Generalized Pareto Distribution (GPD), fits the data in the tail (extreme values). It often plots the empirical tail against the fitted model tail, typically on a log-log scale, to visually assess the goodness of fit.

- **Contours of Bivariate Survival Function plot** : Visualizes the joint survival (or tail) probabilities of two variables, focusing on the tail behavior. Contour lines represent levels of constant survival probability, and the plot illustrates how likely both variables exceed large thresholds together. If contour lines are tightly clustered around the origin, it indicates stronger dependence at extreme values.

4 Data

Daily elevator bid price data were collected from four spatially distinct North Carolina corn markets—Candor, Cofield, Roaring River, and Statesville— covering the period from January 3, 2000, to November 4, 2015. The data are the same as Goodwin and Piggott (2001).[3] The observations were omitted on holidays for all four markets. The four markets

[3] We have updated data in this paper and we would like to thank Nick Piggott for providing his dataset

have about 10-15% missing values and missing values were interpolated using cubic splines, assuming a smooth price continuation.

The Candor, Cofield, Roaring River, and Statesville corn markets were selected based on their regional importance and data availability, reflecting each market's economic durability and agricultural relevance. Each market plays a critical role in North Carolina's diverse agriculture, especially in supporting the state's significant poultry and pork industries, which drive a high demand for corn. Located in North Carolina's northeastern coastal plain, Cofield is central to swine and grain production, with easy access to transport routes, facilitating both domestic and export activities. Moving inland, Candor and Statesville lie in the central region, where poultry production is concentrated. These markets serve as key distribution points, meeting the high feed demands of the local livestock sector. Roaring River, in the western part of the state, supports extensive poultry operations, with terrain and climate conditions suited to the area's agricultural needs. As North Carolina is a net importer of feed grains, these four markets play an essential role in the state's grain distribution network, primarily supplying feed mills to sustain livestock production.

In this paper, rather than studying local prices, we analyze the local basis return. Here the local basis is defined as the difference between a (logarithmic) local cash price and the corresponding price of a (logarithmic) nearby futures contract. The basis return is then calculated as $log(basis_t) - log(basis_{t-1})$. Basis is often used to reflect local market conditions which may be affected by factors such as transportation costs, local supply and demand, storage costs, handling costs, etc. Given the fact that these factors change by location, the basis also varies from one location to the next. Using basis is driven by two main considerations. First, the correlation of the spatial prices may be heavily related to the aggregate market factors such as inflation or population growth that affect all markets. Fackler and Goodwin (2001) indicate that such correlation may be spurious and may not accurately reflect the implications of spatially integrated markets. By analyzing the basis return series, we can investigate the correlation in local markets without considering the movement of the national aggregate market. Second, Extreme Value Theory (EVT) requires that data series be independent and identically distributed (i.i.d.). Stationarity tests (Table 1) confirm that all basis return series are stationary, whereas local prices are non-stationary. Additionally, seasonality plots reveal no seasonality in basis returns, further satisfying EVT requirements.

Table 1: Unit Root Test for Corn Log Basis Returns and Log Prices

	ADF	KPSS	Ng-Perron			
			MZa	MZt	MSB	MPT
	(1)	(2)	(3)	(4)	(5)	(6)
Log Basis Return						
Candor	-53.3879***	0.0142	-21.1893***	-3.2549***	0.1536***	1.1562***
Cofield	-49.2543***	0.0092	-711.2503***	-18.8580***	0.0265***	0.0344***
Roaring River	-52.793***	0.0107	-16.0706***	-2.8347**	0.1764**	1.5245**
Statesville	-53.0262***	0.0137	-7.0467**	-1.8771**	0.2663*	3.4768*
Log Price						
Candor	-2.2871	2.4816***	-11.4282	-2.3190	0.2029	8.3578
Cofield	-2.0707	2.4395***	-10.0743	-2.1548	0.2139	9.4715
Roaring River	-2.2189	2.3850***	-9.7696	-2.1263	0.2176	9.7125
Statesville	-2.1666	2.2564***	-11.1954	-2.2909	0.2046	8.5366

Log Basis Returns are tested using a constant for ADF, KPSS, and Ng-Perron models.
Log Prices are tested with both a constant and a trend;
Critical Values: ADF (Log Basis Return): -2.58 (1%), -1.95 (5%), -1.62 (10%).
KPSS (Log Basis Return): 0.347 (10%), 0.463 (5%), 0.574 (2.5%), 0.739 (1%).
Ng-Perron (Log Basis Return): 1% Level: MZa -13.8, MZt -2.58, MSB 0.174, MPT 1.78;
5% Level: MZa -8.1, MZt -1.98, MSB 0.233, MPT 3.17; 10% Level: MZa -5.7, MZt -1.62, MSB 0.275, MPT 4.45.
ADF (Log Prices): -3.96 (1%), -3.41 (5%), -3.12 (10%).
KPSS (Log Prices): 0.119 (10%), 0.146 (5%), 0.176 (2.5%), 0.216 (1%).
Ng-Perron (Log Prices): 1% Level: MZa -23.8, MZt -3.42, MSB 0.143, MPT 4.03;
5% Level: MZa -17.3, MZt -2.91, MSB 0.168, MPT 5.48; 10% Level: MZa -14.2, MZt -2.62, MSB 0.185, MPT 6.67

Table 2 shows the descriptive statistics and the clustering test results for all corn basis return series. We can see that the magnitude of the volatility is similar for all four corn markets in NC state. The average basis return series for Candor, Cofield, and Roaring River are around zero while it is negative for Statesville. All return distributions display positive skewness, suggesting a tendency for the returns to have longer tails on the right side of the distribution. This indicates that the distributions are not symmetrical. Most observations are clustered on the left (lower values), while fewer, higher values extend the tail to the right. Kurtosis of the four markets is larger than 3 which implies heavy tails for the series and sharper peaks compared to a normal distribution. Jarque-Bera tests further reject normality in all four basis return series. To assess clustering in extremes, which could undermine log-likelihood estimation by inducing dependence in observations, we calculated extremal indexes. The extremal indexes for the 4 corn basis returns are all around 0.9 at the 97th quantile threshold which is closer to 1. There is little evidence of clustering at thresholds selected at the 97th quantile thresholds.

Table 2: Descriptive Statistics and Extremal Index for Corn Basis Returns($/bushel)

Locations	Candor	Cofield	Roaring River	Statesville
Obs	4,116	4,116	4,116	4,116
Mean	1.20114 e-05	1.22553e-06	1.72228e-05	-4.38733e-06
Std. Dev.	0.01466	0.01219	0.01488	0.01442
Min	-0.1809	-0.14059	-0.19758	-0.26804
Max	0.25678	0.25585	0.3066	0.27414
Skewness	1.33345	1.44901	1.87542	1.947
Kurtosis	71.86203	65.66652	87.94430	110.2884
Jarque-Bera	814469***	674937***	1239877***	1976703***
Extremal Index	0.8871	0.9113	0.8548	0.8710

Jarque-Bera is the test for normality; *p<0.10, ** p<0.05, *** p<0.01;
Extremal Index is calculated at the 97th quantile threshold

Figure 1 illustrates the time series of log daily prices (top chart) and log basis returns (bottom chart) over the study period. While the log prices show some alignment across markets, suggesting a level of correlation, the basis returns exhibit distinct patterns of volatility. High-volatility periods in the basis returns do not always correspond clearly across markets, indicating potential variability in local market responses.

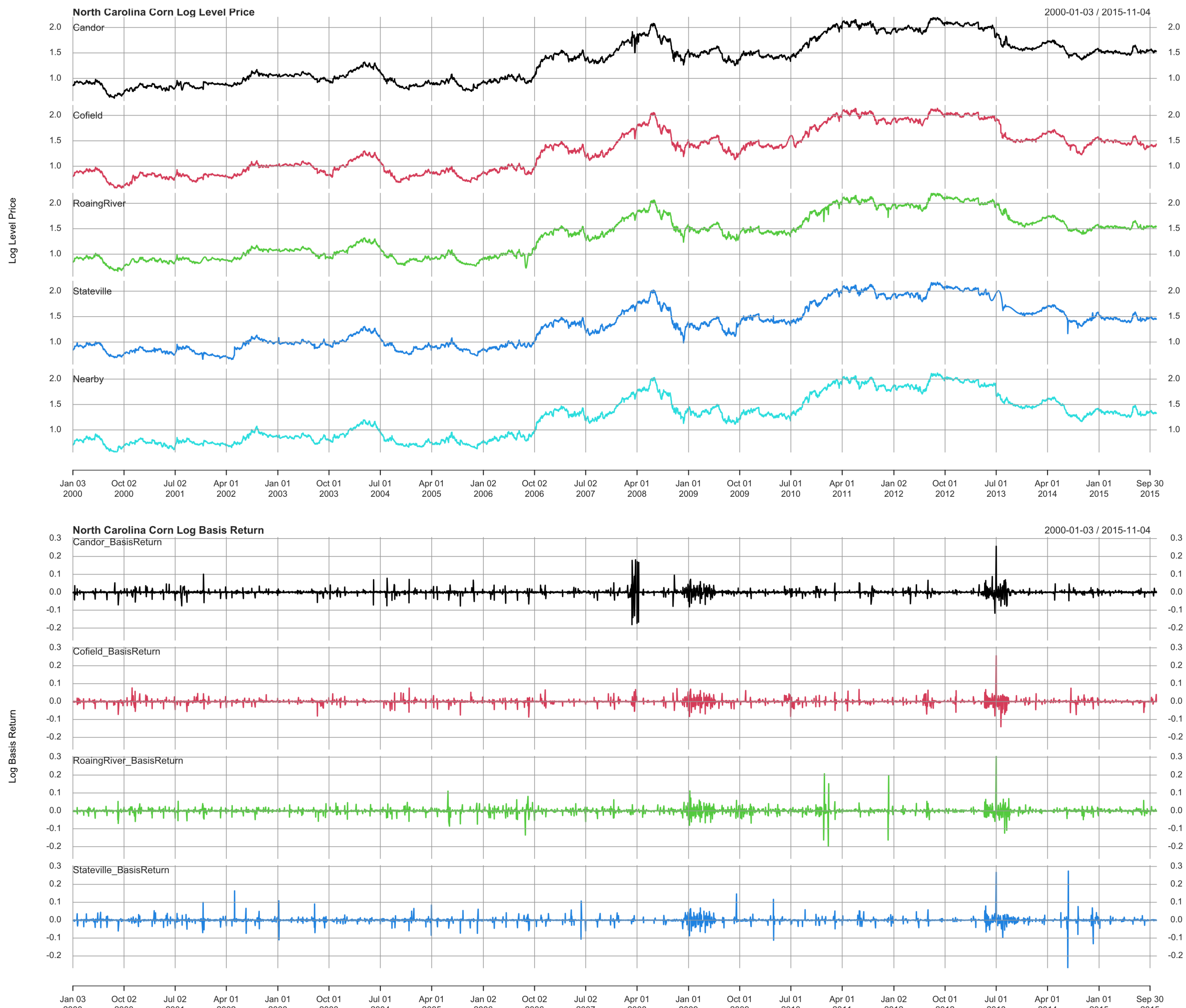


Figure 1: Corn log Level Price, Log Basis Return of Candor, Cofield, Roaring River, Statesville markets in NC

5 Results

Selecting a threshold above which the data can be used for statistical inference in the tail is one of the most fundamental problems in the field of extreme value analysis. Scarrott and MacDonald (2012) and Dey and Yan (2016) provide comprehensive reviews of threshold selection methods. A basic approach to selecting thresholds is through data visualization, with commonly used graphical diagnostics including the Zipf plot, Hill plot, QQ plot, and

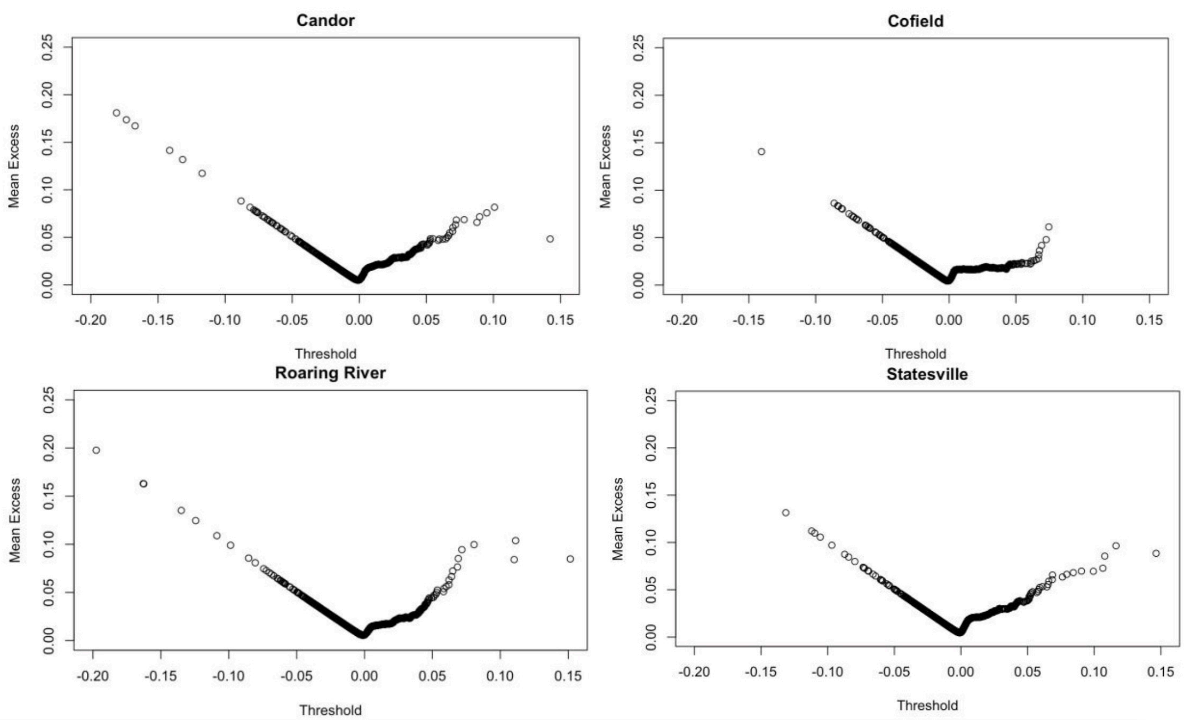


Figure 2: Mean Excess Function Plots for Candor, Cofield, Roaring River, and Statesville

mean excess plot. Figure 2 shows the mean excess function plots for the four markets. If the plot shows linear behavior above a threshold, it indicates that the Generalized Pareto Distribution(GPD) can effectively model the data beyond that threshold. An upward trend in the plot shows heavy-tailed behavior. In particular, a straight line with a positive gradient above the threshold is a sign of Pareto behavior in the tail. A downward trend shows thin-tailed behavior whereas a line with zero gradient shows an exponential tail. From Figure 2 we observe an upward trend above the threshold around zero for Candor, Roaring River, and Statesville, indicating heavy tails. In contrast, Cofield exhibits a relatively flat line beyond the threshold at zero, suggesting less extreme volatility. These observations indicate that the GPD may be suitable for modeling the tail distributions of the corn basis returns in all four markets, particularly beyond a threshold close to zero. However, a major drawback of the graphical method is it is subjectivity due to the necessarily personal interpretation of the plot. Additionally, manually selecting thresholds is time-consuming, particularly in high-dimensional analyses or when dealing with multiple datasets.

To capture extreme values more systematically, we also applied several predetermined threshold levels commonly used in the literature. We defined extremes based on thresholds at the ±3 percent, ±5 percent, ±8 percent, and ±10 percent of the total observations. Negative return exceedances (lower extremes) are values below the lowest 3%, 5%, 8%, and 10% of observations, capturing the most negative deviations in the distribution. Conversely, positive return exceedances (upper extremes) are values exceeding the highest 97%, 95%, 92%, and 90% of observations, capturing the most positive deviations. By setting separate thresholds for the upper and lower tails, we account for potential differences in extreme behavior across both tails. We also explored the optimal threshold selection approach, which aims to minimize the asymptotic mean squared error (AMSE) of the Hill estimator, balancing bias and variance. Following Caeiro and Gomes (2015), we used the optimal number of upper-order statistics that minimize the AMSE criterion. Additionally, we also applied a single bootstrap procedure to choose the optimal sample fraction to select the appropriate optimal threshold.

Tables 3 to 9 and Figure 4 present the estimation results for bivariate distributions of corn basis return exceedances across six market pairs in North Carolina: Candor/Cofield, Candor/Roaring River, Candor/Statesville, Cofield/Roaring River, Cofield/Statesville, and Roaring River/Statesville for different predetermined threshold levels: ±3 percent, ±5 percent, ±8 percent, and ±10 and also the optimal threshold. All the estimates are statistically significant according to the standard errors. These tables also show the estimation of the tail distribution for each of the 4 markets. When we look at the shape estimates from Tables 3 to 8, they demonstrate the volatility across different markets as the mean excess plot suggested, noting that the choice of threshold influences the tail distribution. For the Candor, Roaring River, and Statesville markets, shape estimates above zero across all threshold

Table 3: Estimation of the Bivariate Distribution of Candor and Cofield Corn Basis Return Exceedances

$Threshold_{Candor}$	P_{Candor}	$Shape_{Candor}$	$Scale_{Candor}$	$Threshold_{Cofield}$	$P_{Cofield}$	$Shape_{Cofield}$	$Scale_{Cofield}$	$Dependence$
Positive Return Exceedances								
0.00513	10%	0.39017	0.01004	0.00422	10%	0.1269	0.01331	0.54245
		(0.07236)	(0.00084)			(0.05378)	(0.00095)	(0.01429)
0.00697	8%	0.30508	0.01252	0.00654	8%	0.074	0.01526	0.55737
		(0.06976)	(0.00108)			(0.04648)	(0.00109)	(0.01623)
0.01323	5%	0.29247	0.01477	0.01437	5%	0.10859	0.01461	0.56162
		(0.0835)	(0.00156)			(0.06226)	(0.00134)	(0.02008)
0.02181	3%	0.43946	0.01379	0.02246	3%	0.17241	0.01366	0.60642
		(0.13608)	(0.00216)			(0.0944)	(0.00174)	(0.02652)
0.01936	3.57143%	0.41628	0.01329	0.02019	3.47425%	0.15303	0.01391	0.59305
		(0.11877)	(0.00184)			(0.08369)	(0.00161)	(0.02436)
Negative Return Exceedances								
-0.00516	-10%	0.37409	0.01070	0.00388	-10%	0.17644	0.01334	0.47449
		(0.07474)	(0.00092)			(0.07266)	(0.00115)	(0.01394)
-0.00735	-8%	0.31690	0.01265	0.00607	-8%	0.03977	0.01696	0.47239
		(0.0799)	(0.00119)			(0.06373)	(0.00141)	(0.01531)
-0.01299	-5%	0.16292	0.01841	0.01326	-5%	-0.053	0.02023	0.49312
		(0.07718)	(0.00189)			(0.05907)	(0.00183)	(0.01977)
-0.02232	-3%	0.19396	0.01939	-0.02337	-3%	-0.04391	0.01942	0.50046
		(0.10739)	(0.00268)			(0.076)	(0.00227)	(0.02456)
-0.02014	-3.54713%	0.25278	0.01686	-0.02071	-3.52283%	-0.02770	0.01886	0.48784
		(0.10919)	(0.00226)			(0.07577)	(0.0021)	(0.02262)

All estimates are significant at the 1% level.
Standard errors are in parentheses.
Last row for each panel represent the estimation results for optimal threshold

levels indicate heavy-tailed behavior in both upper and lower tails. In contrast, the Cofield market shows a different pattern: the tail distribution is above zero for all positive return exceedances and negative return exceedances at the -10% and -8% threshold level, while there's no significant tail for lower tail exceedances at the -5%, and -3% threshold and the optimal threshold -3.52283%. This suggests higher volatility in the Candor, Roaring River, and Statesville markets, with pronounced extreme returns in both directions, compared to the Cofield market.

As discussed earlier, extreme events in one market do not always align with those in another market. Before estimating the extreme dependence, we use nonparametric measures—plot χ and $\bar{\chi}$—to test for tail asymptotic dependence and to inform the choice of dependence function before fitting model parameters. Figure 3 illustrates the χ and $\bar{\chi}$ plots for the bivariate corn basis returns of six market pairs, assessing tail dependence across quantiles.

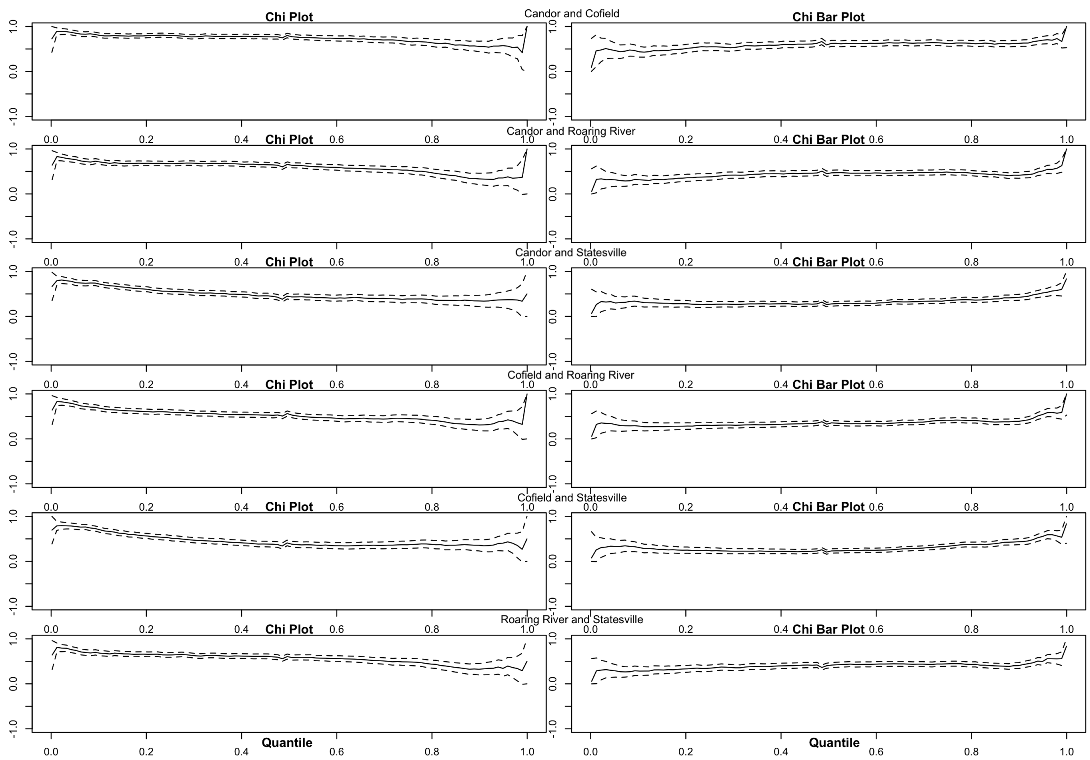


Figure 3: χ and $\bar{\chi}$ Plots for 6 market pairs of corn basis returns

Table 4: Estimation of the Bivariate Distribution of Candor and Roaring River Corn Basis Return Exceedances

$Threshold_{Candor}$	P_{Candor}	$Shape_{Candor}$	$Scale_{Candor}$	$Threshold_{RoaringRiver}$	$P_{RoaringRiver}$	$Shape_{RoaringRiver}$	$Scale_{RoaringRiver}$	$Dependence$
Positive Return Exceedances								
0.00513	10%	0.39017	0.01004	0.00759	10%	0.23888	0.01141	0.70844
		(0.0723)	(0.00084)			(0.05503)	(0.00082)	(0.01489)
0.00697	8%	0.30508	0.01252	0.01034	8%	0.24804	0.01178	0.72097
		(0.06976)	(0.00108)			(0.06066)	(0.00094)	(0.0167)
0.01323	5%	0.29247	0.01477	0.01698	5%	0.33375	0.01126	0.71743
		(0.0835)	(0.00156)			(0.08537)	(0.00119)	(0.02088)
0.02181	3%	0.43946	0.01379	0.02267	3%	0.3422	0.01352	0.73862
		(0.13608)	(0.00216)			(0.10849)	(0.00184)	(0.02648)
0.01936	3.57143%	0.41628	0.01329	0.02148	3.42566%	0.38173	0.01198	0.72958
		(0.11877)	(0.00184)			(0.11162)	(0.00161)	(0.02465)
Negative Return Exceedances								
-0.00516	-10%	0.37409	0.01070	-0.00650	-10%	0.21105	0.01359	0.60288
		(0.07474)	(0.00092)			(0.06241)	(0.00106)	(0.01482)
-0.00735	-8%	0.31690	0.01265	-0.00879	-8%	0.13706	0.01625	0.62577
		(0.0799)	(0.00119)			(0.05797)	(0.00128)	(0.01692)
-0.01299	-5%	0.16292	0.01841	-0.01734	-5%	0.16121	0.01634	0.62168
		(0.07718)	(0.00189)			(0.07239)	(0.00162)	(0.02097)
-0.02232	-3%	0.19396	0.01939	-0.02686	-3%	0.24839	0.01511	0.64301
		(0.10739)	(0.00268)			(0.10444)	(0.00203)	(0.02682)
-0.02014	-3.54713%	0.25278	0.01686	-0.02438	-3.40136%	0.21101	0.01586	0.63486
		(0.10919)	(0.00226)			(0.09314)	(0.00195)	(0.02506)

All estimates are significant at the 1% level.
Standard errors are in parentheses.

The x-axis represents the quantile q, and the y-axis represents the χ and $\bar{\chi}$ estimates, respectively. The left panels show Chi plots and the right panels show Chi-bar plots for market pairs. The solid lines represent the Chi and Chi-bar estimates, while the dashed lines indicate confidence intervals. The Chi function measures quantile-dependent asymptotic dependence. For perfectly dependent variables, $\chi(q) = 1$ across all quantiles q, while for asymptotically independent variables, $\chi(q) = 0$. The Chi-bar function is also quantile-dependent, providing insights into tail independence. If $\bar{\chi}(q) = 1$ at high quantiles, the variables are asymptotically dependent, but if $\bar{\chi}(q)$ approaches zero, the variables are asymptotically independent. For the six market pairs, all the $\bar{\chi}(q)$ values approach 1 as the quantile increases, and all the $\chi(q)$ plots are greater than 0 across the quantiles, indicating that the six bivariate corn basis returns between market pairs are asymptotically dependent. Therefore, we can use the logistic dependence function to assess the extreme correlation between markets.

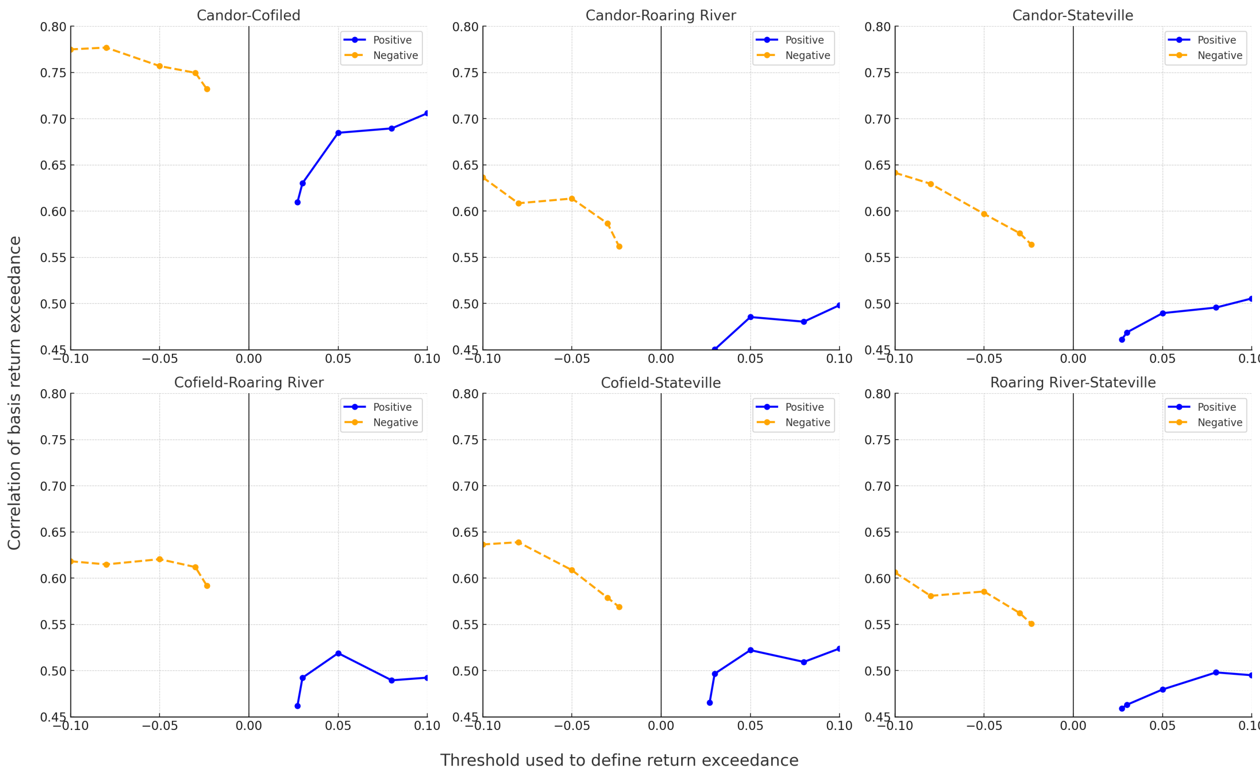


Figure 4: Corn Basis Return Extreme Correlation between 6 market pairs in North Carolina

Table 5: Estimation of the Bivariate Distribution of Candor and Statesville Corn Basis Return Exceedances

$Threshold_{Candor}$	P_{Candor}	$Shape_{Candor}$	$Scale_{Candor}$	$Threshold_{Statesville}$	$P_{Statesville}$	$Shape_{Statesville}$	$Scale_{Statesville}$	$Dependence$
Positive Return Exceedances								
0.00513	10%	0.39017	0.01004	0.00453	10%	0.50097	0.00858	0.70326
		(0.0723)	(0.00084)			(0.08668)	(0.00081)	(0.01488)
0.00697	8%	0.30508	0.01252	0.00577	8%	0.30859	0.01262	0.71025
		(0.06976)	(0.00108)			(0.07183)	(0.00111)	(0.01663)
0.01323	5%	0.29247	0.01477	0.01262	5%	0.30125	0.01446	0.71449
		(0.0835)	(0.00156)			(0.08516)	(0.00155)	(0.02081)
0.02181	3%	0.43946	0.01379	0.02078	3%	0.37851	0.01506	0.73037
		(0.13608)	(0.00216)			(0.12443)	(0.00223)	(0.02644)
0.01936	3.57143%	0.41628	0.01329	0.02214	2.69679%	0.3467	0.01658	0.73853
		(0.11877)	(0.00184)			(0.12418)	(0.00251)	(0.02626)
Negative Return Exceedances								
-0.00516	-10%	0.37409	0.01070	-0.00445	-10%	0.39464	0.01034	0.59880
		(0.07474)	(0.00092)			(0.08413)	(0.00097)	(0.01470)
-0.00735	-8%	0.31690	0.01265	-0.00605	-8%	0.21586	0.01448	0.60871
		(0.0799)	(0.00119)			(0.07297)	(0.00130)	(0.01642)
-0.01299	-5%	0.16292	0.01841	-0.01206	-5%	0.08638	0.01995	0.63488
		(0.07718)	(0.00189)			(0.06158)	(0.00184)	(0.02086)
-0.02232	-3%	0.19396	0.01939	-0.02353	-3%	0.14565	0.01825	0.65119
		(0.10739)	(0.00268)			(0.08576)	(0.00224)	(0.02623)
-0.02014	-3.54713%	0.25278	0.01686	-0.02501	2.81827%	0.15864	0.01792	0.66197
		(0.10919)	(0.00226)			(0.0907)	(0.00229)	(0.02606)

All estimates are significant at the 1% level.
Standard errors are in parentheses.

Table 9 presents the extreme correlation data for six pairs of corn basis return markets in North Carolina. Our findings reveal that correlations are influenced by both the magnitude and direction of threshold levels that define extreme returns. For negative return exceedances, we observe an increase in correlation across all market pairs. For example, the correlation between the Candor and Roaring River markets rises from 0.58654 at a -3% threshold to 0.63654 at a -10% threshold. On the other hand, for positive return exceedances, an increase in correlation is observed for the Candor and Cofield markets at higher thresholds, while correlations for other market pairs remain more stable during positive shocks.

Figure 4 further illustrates the extreme correlation of basis return exceedances across different threshold levels. The plot of the extreme correlation between 6 market pairs also shows that the correlation of basis return exceedances tends to increase with the threshold levels. Notably, the increase is more pronounced during negative return exceedances com-

Table 6: Estimation of the Bivariate Distribution of Cofield and Roaring River Corn Basis Return Exceedances

$Threshold_{Cofield}$	$P_{Cofield}$	$Shape_{Cofield}$	$Scale_{Cofield}$	$Threshold_{RoaringRiver}$	$P_{RoaringRiver}$	$Shape_{RoaringRiver}$	$Scale_{RoaringRiver}$	$Dependence$
Positive Return Exceedances								
0.00422	10%	0.1269	0.01331	0.00759	10%	0.23888	0.01141	0.71256
		(0.05378)	(0.00095)			(0.05503)	(0.00082)	(0.01516)
0.00654	8%	0.07400	0.01526	0.01034	8%	0.24804	0.01178	0.71463
		(0.04648)	(0.00109)			(0.06066)	(0.00094)	(0.01698)
0.01437	5%	0.10859	0.01461	0.01698	5%	0.33375	0.01126	0.69368
		(0.06226)	(0.00134)			(0.08537)	(0.00119)	(0.02101)
0.02246	3%	0.17241	0.01366	0.02267	3%	0.3422	0.01352	0.70600
		(0.09440)	(0.00174)			(0.10849)	(0.00184)	(0.02654)
0.02019	3.47425%	0.15303	0.01391	0.02148	3.42566%	0.38173	0.01198	0.69802
		(0.08369)	(0.00161)			(0.11162)	(0.00161)	(0.02482)
Negative Return Exceedances								
-0.00388	-10%	0.17644	0.01334	-0.00650	-10%	0.21105	0.01359	0.61785
		(0.07266)	(0.00115)			(0.06241)	(0.00106)	(0.01518)
-0.00607	-8%	0.03977	0.01696	-0.00879	-8%	0.13706	0.01625	0.62061
		(0.06373)	(0.00141)			(0.05797)	(0.00128)	(0.01694)
-0.01326	-5%	-0.05300	0.02023	-0.01734	-5%	0.16121	0.01634	0.61605
		(0.05907)	(0.00183)			(0.07239)	(0.00162)	(0.02112)
-0.02337	-3%	-0.04391	0.01942	-0.02686	-3%	0.24839	0.01511	0.62295
		(0.07600)	(0.00227)			(0.10444)	(0.00203)	(0.02664)
-0.02071	-3.52284%	-0.02770	0.01886	-0.02438	-3.40136%	0.21101	0.01586	0.62023
		(0.07577)	(0.0021)			(0.09314)	(0.00195)	(0.02500)

All estimates are significant at the 1% level.
Standard errors are in parentheses.

pared to positive ones for most market pairs. This reflects an asymmetry in the response of negative versus positive return exceedances across all market pairs, with stronger correlations in the lower (left) tail than in the upper (right) tail. This asymmetry implies that market pairs demonstrate stronger co-movement during negative shocks than positive shocks. When compared to linear correlations, these extreme correlations are more pronounced, suggesting that market co-movements intensify during periods of extreme volatility. Particularly, the Candor and Cofield market pair shows the strongest correlations in both severe downturns and upswings.

These findings suggest that farmers might consider diversifying their market options beyond Candor and Cofield during periods of significant price swings to mitigate risks. Similarly, buyers like exporters or processors that rely on corn as inputs should be aware of potential sourcing risks from these markets during extreme market conditions, as the elevated

Table 7: Estimation of the Bivariate Distribution of Cofield and Statesville Corn Basis Return Exceedances

$Threshold_{Cofield}$	$P_{Cofield}$	$Shape_{Cofield}$	$Scale_{Cofield}$	$Threshold_{Statesville}$	$P_{Statesville}$	$Shape_{Statesville}$	$Scale_{Statesville}$	$Dependence$
Positive Return Exceedances								
0.00422	10%	0.1269	0.01331	0.00453	10%	0.50097	0.00858	0.69001
		(0.05378)	(0.00095)			(0.08668)	(0.00081)	(0.01500)
0.00654	8%	0.07400	0.01526	0.00577	8%	0.30859	0.01262	0.70053
		(0.04648)	(0.00109)			(0.07183)	(0.00111)	(0.01686)
0.01437	5%	0.10859	0.01461	0.01262	5%	0.30125	0.01446	0.69134
		(0.06226)	(0.00134)			(0.08516)	(0.00155)	(0.02090)
0.02246	3%	0.17241	0.01366	0.02078	3%	0.37851	0.01506	0.70724
		(0.09440)	(0.00174)			(0.12443)	(0.00223)	(0.02627)
0.02019	3.47425%	0.15303	0.01391	0.00314	2.69679%	0.3467	0.01658	0.71626
		(0.08369)	(0.00161)			(0.12418)	(0.00251)	(0.02635)
Negative Return Exceedances								
-0.00388	-10%	0.17644	0.01334	-0.00445	-10%	0.39464	0.01034	0.60299
		(0.07266)	(0.00115)			(0.08413)	(0.00097)	(0.01489)
-0.00607	-8%	0.03977	0.01696	-0.00605	-8%	0.21586	0.01448	0.60095
		(0.06373)	(0.00141)			(0.07297)	(0.00130)	(0.01637)
-0.01326	-5%	-0.05300	0.02023	-0.01206	-5%	0.08638	0.01995	0.62555
		(0.05907)	(0.00183)			(0.06158)	(0.00184)	(0.02074)
-0.02337	-3%	-0.04391	0.01942	-0.02353	-3%	0.14565	0.01825	0.64899
		(0.07600)	(0.00227)			(0.08576)	(0.00224)	(0.02640)
-0.02071	-3.52284%	-0.0277	0.01886	-0.02501	-2.81827%	0.15864	0.01792	0.65271
		(0.07577)	(0.0021)			(0.0907)	(0.00229)	(0.02606)

All estimates are significant at the 1% level.
Standard errors are in parentheses.

correlations imply heightened volatility and increased co-movement between these markets during such events. Buyers like ethanol plants or feedlots may leverage these insights to negotiate pricing with suppliers, anticipating correlated movements in nearby markets. High market correlation might indicate reliance on shared transportation or processing infrastructure. Grain storage operators in highly correlated markets may need to coordinate inventory to manage risks. Logistics and transportation companies could better allocate transport resources strategies. Investments could be made to diversify infrastructure or increase capacity in these areas to mitigate risks.

Table 8: Estimation of the Bivariate Distribution of Roaring River and Statesville Corn Basis Return Exceedances

$Threshold_{RoaringR}$	$P_{RoaringR}$	$Shape_{RoaringR}$	$Scale_{RoaringR}$	$Threshold_{States}$	P_{States}	$Shape_{States}$	$Scale_{States}$	$Dependence$
Positive Return Exceedances								
0.00759	10%	0.23888	0.01141	0.00453	10%	0.50097	0.00858	0.71067
		(0.05503)	(0.00082)			(0.08668)	(0.00081)	(0.01502)
0.01034	8%	0.24804	0.01178	0.00577	8%	0.30859	0.01262	0.70855
		(0.06066)	(0.00094)			(0.07183)	(0.00111)	(0.01670)
0.01698	5%	0.33375	0.01126	0.01262	5%	0.30125	0.01446	0.72142
		(0.08537)	(0.00119)			(0.08516)	(0.00155)	(0.02097)
0.02267	3%	0.34220	0.01352	0.02078	3%	0.37851	0.01506	0.73022
		(0.10849)	(0.00184)			(0.12443)	(0.00223)	(0.02637)
0.02148	3.42566%	0.38173	0.01198	0.02214	2.69679%	0.3467	0.01658	0.73875
		(0.11162)	(0.00161)			(0.12418)	(0.00251)	(0.02646)
Negative Return Exceedances								
-0.00650	-10%	0.21105	0.01359	-0.00445	-10%	0.39464	0.01034	0.62750
		(0.06241)	(0.00106)			(0.08413)	(0.00097)	(0.01494)
-0.00879	-8%	0.13706	0.01625	-0.00605	-8%	0.21586	0.01448	0.64743
		(0.05797)	(0.00128)			(0.07297)	(0.00130)	(0.01695)
-0.01734	-5%	0.16121	0.01634	-0.01206	-5%	0.08638	0.01995	0.64378
		(0.07239)	(0.00162)			(0.06158)	(0.00184)	(0.02109)
-0.02686	-3%	0.24839	0.01511	-0.02353	-3%	0.14565	0.01825	0.66172
		(0.10444)	(0.00203)			(0.08576)	(0.00224)	(0.02689)
-0.02438	-3.40136%	0.21101	0.01586	-0.02501	-2.81827%	0.15864	0.01792	0.64282
		(0.09314)	(0.00195)			(0.0907)	(0.00229)	(0.02621)

All estimates are significant at the 1% level.
Standard errors are in parentheses.

6 Model Checking

All estimates in this study are significant at the 1% level, which provides initial support for the accuracy of our model. To further validate the fitness of the model, various diagnostic plots were employed to assess how well the model captures the extreme values in the data.

To assess the fit of the extreme value model in capturing the upper and lower tails of the marginal distributions, the tail of marginal plots was used to visualize tail behavior in each market. Figure 5 shows the upper tail plots (Panel A) and lower tail plots (Panel B), each plotted at the optimal threshold level to evaluate whether the Generalized Pareto Distribution (GPD) accurately represents the extreme tails observed in the corn basis returns for each market. In these plots, the x-axis represents extreme basis return values on a log scale, while the y-axis shows the exceedance probability, also on a log scale. Scatter points depict the empirical data, while a continuous line represents the fitted GPD model. A close

Table 9: Corn Basis Return Extreme Correlation between 6 market pairs in North Carolina

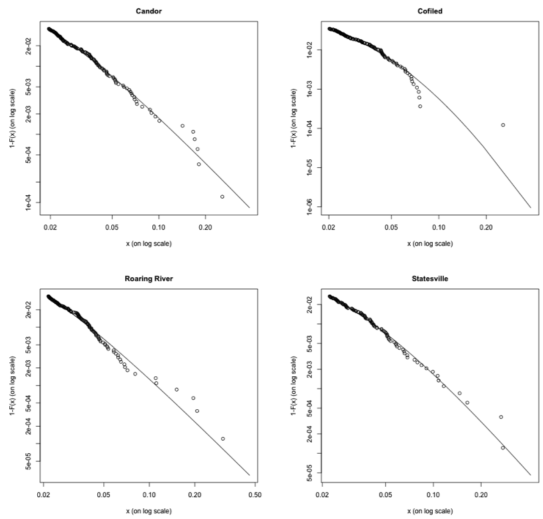
Threshold	Candor-Coefield	Candor-Roaring River	Candor-Statesville	Coefield-Roaring River	Coefield-Statesville	Roaring River-Statesville
Positive Return Exceedances						
10%	0.70575	0.49812	0.50543	0.49225	0.52388	0.49495
8%	0.68934	0.48020	0.49554	0.48930	0.50926	0.49795
5%	0.68458	0.48530	0.48951	0.51880	0.52205	0.47955
3%	0.63225	0.45444	0.46656	0.50157	0.49981	0.46278
Optimal	0.64829	0.46771	0.45458	0.51277	0.48697	0.45424
Negative Return Exceedances						
-10%	0.77486	0.63654	0.64144	0.61826	0.63640	0.60624
-8%	0.77685	0.60841	0.62948	0.61484	0.63885	0.58083
-5%	0.75684	0.61351	0.59692	0.62048	0.60869	0.58555
-3%	0.74954	0.58654	0.57596	0.61194	0.57881	0.56213
Optimal	0.76202	0.59696	0.56179	0.61532	0.57397	0.58678
Constant Correlation						
	0.58696	0.42010	0.34040	0.48160	0.42263	0.39240

All estimates are significant at the 1% level.

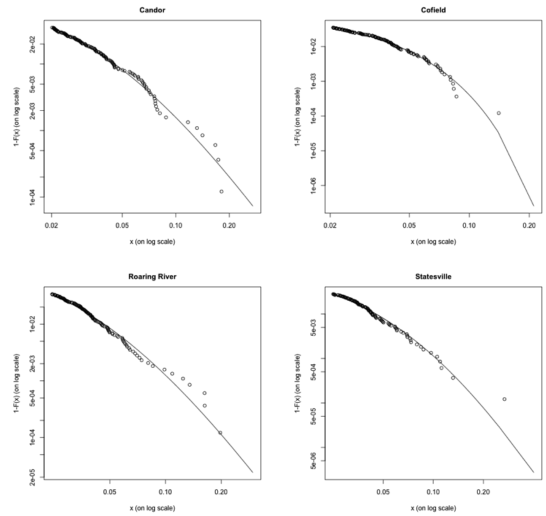
alignment between the empirical data and the fitted line in each plot suggests that the GPD model is well-suited to capture extreme behavior in the tails of the distributions. The shape and steepness of the fitted GPD line provide insights into the "fatness" of the tails. A slower decay in the line implies a heavy (fat) tail, where extreme values are more frequent, while a steeper decay signifies a thin tail, with rarer extreme events. For fat-tailed distributions, the data points in the tail region tend to spread out more broadly, whereas in thin-tailed distributions, the data points cluster closer together and exhibit a rapid decline.

For all markets–Candor, Cofield, Roaring River, and Statesville–the data points in both the upper and lower tails align well with the fitted GPD line, indicating a strong model fit. This alignment suggests that the GPD model reliably captures the tail behavior across these markets, accurately representing both the occurrence and magnitude of extreme events. However, the data points in Cofield's lower tail cluster closely and show a rapid drop-off, consistent with the steep slope of the GPD line. This observation is aligned with conclusions from our estimation, which suggests a thin tail distribution for Cofield's lower tail.

Figure 6 presents the Bivariate Survival Function Contour Plots for both the upper and lower tails of corn basis returns across various market pairs (Candor, Cofield, Roaring River, and Statesville). These contour plots are used to examine the model's fit in capturing the



(a) Upper Tail of Marginal Plots



(b) Lower Tail of Marginal Plots

Figure 5: Upper and Lower Tails of Marginal Plots of Corn Basis Return for Each Market

dependence structure of extreme basis returns between markets. Each contour line represents a joint survival probability level, indicating the likelihood that both markets in a given pair exceed a specific threshold [4] (upper tail) or fall below a threshold [5] (lower tail) simultaneously. The contour levels (e.g., 0.001, 0.002) indicate the joint survival probabilities, where smaller values further from the origin correspond to rarer joint events. Contours closer to the origin represent higher survival probabilities for simultaneous extreme values in both markets.

In the upper tail contour plots (Panel A), the contour lines for all market pairs are tightly packed and clustered near the bottom left corner, indicating a rapid decrease in survival probability as one moves away from the origin. This pattern reflects low probabilities for joint extreme events across both markets. Among these pairs, the Candor and Cofield pair shows relatively steep contour lines, suggesting a stronger level of joint dependence in the upper tail region. In the lower tail contour plots (Panel B), the contours are notably more sharply curved across all market pairs, indicating stronger dependence in the negative tails.

[4] The plot is generated using optimal upper threshold
[5] The plot is generated using optimal lower threshold

The Candor and Cofield pair, in particular, shows the steepest and sharpest contour lines, suggesting the strongest lower tail dependence among the pairs. This observation aligns with the results of our model estimation, which indicated a pronounced dependence in the lower tails for this pair.

Our model-checking results suggest that the GPD model provides a robust fit for capturing tail behavior across the four markets. The alignment of empirical data with the fitted GPD line in the marginal tail plots, combined with the distinct tail dependence patterns observed in the bivariate survival contour plots, supports the model's accuracy in representing both tails for each market and joint extremes.

7 Conclusion and Discussion

This paper aims to estimate local-level extreme co-movement in agricultural commodity markets, focusing on how extreme events influence price discovery and information flow. Understanding these dynamics is crucial for capturing market behavior during periods of significant volatility.

Our findings reveal an asymmetric correlation between the left and right tails for the 6 market pairs—stronger correlations in the lower tail than in the upper tail for 6 market pairs which indicates that extreme events impact price-setting behavior differently depending on whether the market is experiencing a negative or positive shock. The Candor/Cofield pair was found that have a stronger extreme correlation among 6 pairs.

These insights are particularly relevant for market participants and policymakers. Producers like farmers can use this information to make better decisions regarding risk management, structuring portfolios, and managing supply disruption. Buyers such as processors and exporters might design better supply chains that account for correlated basis risk to enhance resilience. High market correlation might indicate reliance on shared transportation or processing infrastructure. Grain storage operators and logistics and transportation companies in highly correlated markets may need to better coordinate to manage risks and allocate transport resources strategies. Investments could be made to diversify infrastructure or increase capacity in these areas to mitigate risks. Policymakers, in turn, can leverage these findings to design more effective revenue protection and insurance programs to mitigate the adverse impacts of extreme events. Targeted policies between highly correlated markets can stabilize interconnected regions.

Additionally, we found that the tail distributions for the four markets are non-normal and asymmetric. This implies that standard time series models, which often assume a normal distribution or symmetric distribution (e.g., student-t distribution) for residuals such as ARIMA, linear regression, and Vector Autoregression, may not be adequate for capturing extreme price movements. This challenges the validity of traditional modeling approaches that assume normal residuals, highlighting the need for models that better accommodate non-normal behavior in the tails. Incorporating such models will expand the econometric toolkit for understanding agricultural price series during extreme events.

The conventional dependence measure, Pearson correlation, calculates an average of deviations from the mean without distinguishing between large and small movements or positive and negative returns. It also assumes a linear relationship and a multivariate Gaussian distribution, which can significantly underestimate the risk from joint extreme events. In contrast, our approach focuses on extreme co-movement, offering a more accurate representation of market dependencies during periods of extreme volatility.

For future research, time-varying dependence models such as copulas could be valuable in capturing the dynamic nature of market co-movements. Extending the analysis to multivariate models capable of handling extreme co-movements across multiple markets could also provide a broader perspective. Furthermore, incorporating additional factors, such as weather events, policy changes, and global market disruptions into threshold exceedance analysis could help better identify the drivers behind extreme market behavior.

Overall, this paper contributes to a better understanding of price dependencies and risks in local-level regional agricultural markets, offering practical tools for stakeholders to make informed decisions in the face of market uncertainties.

References

Balkema, A.A., and L. De Haan. 1974. "Residual life time at great age." *The Annals of probability* 2:792–804.

Bekiros, S.D., and D.A. Georgoutsos. 2008. "The extreme-value dependence of Asia-Pacific equity markets." *Journal of Multinational Financial Management* 18:197–208.

Caeiro, F., and M.I. Gomes. 2015. "Threshold selection in extreme value analysis." *Extreme value modeling and risk analysis: Methods and applications*, pp. 69–87.

Christoffersen, P., V. Errunza, K. Jacobs, and H. Langlois. 2012. "Is the potential for international diversification disappearing? A dynamic copula approach." *The Review of financial studies* 25:3711–3751.

Dey, D.K., and J. Yan. 2016. *Extreme value modeling and risk analysis: methods and applications*. CRC Press.

Dupuis, D.J., and J.A. Tawn. 2001. "Effects of mis-specification in bivariate extreme value problems." *Extremes* 4:315–330.

Erb, C.B., and C.R. Harvey. 2006. "The strategic and tactical value of commodity futures." *Financial Analysts Journal* 62:69–97.

Ergen, I. 2014. "Tail dependence and diversification benefits in emerging market stocks: an extreme value theory approach." *Applied Economics* 46:2215–2227.

Fackler, P.L., and B.K. Goodwin. 2001. "Spatial price analysis." *Handbook of agricultural economics* 1:971–1024.

Fretheim, T., and G. Kristiansen. 2015. "Commodity market risk from 1995 to 2013: an extreme value theory approach." *Applied Economics* 47:2768–2782.

Geman, H. 2005. *Commodities and commodity derivatives: modeling and pricing for agriculturals, metals and energy*, vol. 302. John Wiley & Sons.

Goodwin, B.K., and N.E. Piggott. 2001. "Spatial market integration in the presence of threshold effects." *American Journal of Agricultural Economics* 83:302–317.

Guney, S., B.K. Goodwin, and A. Riquelme. 2019. "Semi-Parametric Generalized Additive Vector Autoregressive Models of Spatial Basis Dynamics." *American Journal of Agricultural Economics* 101:541–562.

Hussain, S.I., and S. Li. 2021. "Dependence structure between oil and other commodity futures in China based on extreme value theory and copulas." *The World Economy*, pp. .

Koedijk, K.G., M.M. Schafgans, and C.G. De Vries. 1990. "The tail index of exchange rate returns." *Journal of international economics* 29:93–108.

Ledford, A.W., and J.A. Tawn. 1996. "Statistics for near independence in multivariate extreme values." *Biometrika* 83:169–187.

Longin, F., and B. Solnik. 2001. "Extreme correlation of international equity markets." *The journal of finance* 56:649–676.

Lux, T. 2000. "On moment condition failure in German stock returns: an application of recent advances in extreme value statistics." *Empirical economics* 25:641–652.

Mohendru, I. 1937. “Some factors affecting the price of wheat in the Punjab.” *Board of Economic Inquiry Publication*, pp. .

Park, E., and J. Maples. 2018. “Serially Dependent Extreme Events in Agricultural Commodity Futures Markets.” Working paper.

Pickands III, J. 1975. “Statistical inference using extreme order statistics.” *the Annals of Statistics*, pp. 119–131.

Poon, S.H., M. Rockinger, and J. Tawn. 2004. “Extreme value dependence in financial markets: Diagnostics, models, and financial implications.” *The Review of Financial Studies* 17:581–610.

Rocco, M. 2014. “Extreme value theory in finance: A survey.” *Journal of Economic Surveys* 28:82–108.

Scarrott, C., and A. MacDonald. 2012. “A review of extreme value threshold estimation and uncertainty quantification.” *REVSTAT-Statistical journal* 10:33–60.

Sephton, P.S. 2003. “Spatial market arbitrage and threshold cointegration.” *American Journal of Agricultural Economics* 85:1041–1046.

Singh, A.K., D.E. Allen, and R.J. Powell. 2017. “Tail dependence analysis of stock markets using extreme value theory.” *Applied Economics* 49:4588–4599.

Singvejsakul, J., C. Chaiboonsri, and S. Sriboonchitta. 2021. “The Optimization of Bayesian Extreme Value: Empirical Evidence for the Agricultural Commodities in the US.” *Economies* 9:30.

Smith, R.L., and I. Weissman. 1994. “Estimating the extremal index.” *Journal of the Royal Statistical Society: Series B (Methodological)* 56:515–528.

Stigler, G.J., and R.A. Sherwin. 1985. "The extent of the market." *The Journal of Law and Economics* 28:555–585.

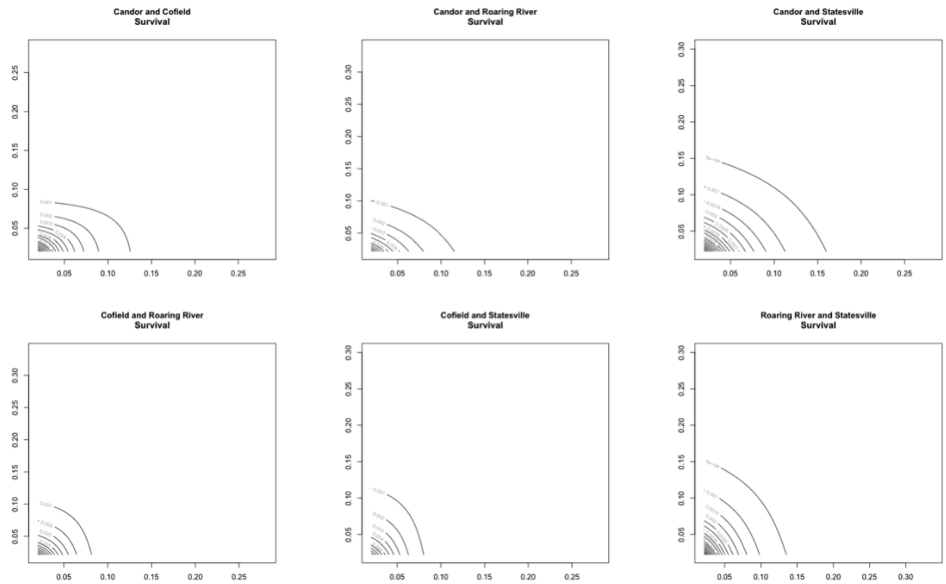
Sukcharoen, K., and D.J. Leatham. 2016. "Dependence and extreme correlation among US industry sectors." *Studies in Economics and Finance*, pp. .

Tiago de Oliveira, J. 1973. "Statistical Extremes-a survey." *Center of Applied Mathematics, Lisbon*, pp. .

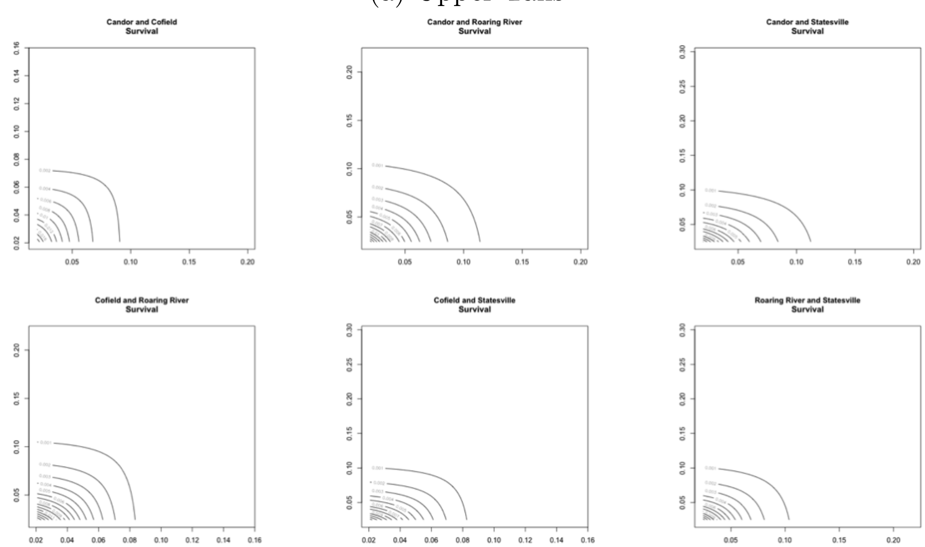
van Oordt, M.R., P.A. Stork, and C.G. de Vries. 2021. "On agricultural commoditiesâ extreme price risk." *Extremes* 24:531–563.

Vilasuso, J., and D. Katz. 2000. "Estimates of the likelihood of extreme returns in international stock markets." *Journal of Applied Statistics* 27:119–130.

Zeevi, A., and R. Mashal. 2002. "Beyond correlation: Extreme co-movements between financial assets." *Available at SSRN 317122*, pp. .



(a) Upper Tails



(b) Lower Tails

Figure 6: Bivariate survival function Plot for Upper and Lower Tails of Corn Basis Return for Each Market