ABSTRACT

MENG, SHU. Three Essays on Agricultural Commodity Market Linkages. (Under the direction of Barry Goodwin).

The three essays of this thesis consider topics on agricultural economics. The first essay investigates the validity of the law of one price among Oriented Strand Board (OSB) markets using copula models, the second essay investigates the law of one price among plywood markets using stochastic differential equations (SDE), and the third essay investigates the market linkages between agricultural commodity markets and financial markets.

In the first chapter, we use different types of multivariate copulas to model the dependence of the spatial prices in the US and Canadian oriented strand board (OSB) markets. In this case, the joint probability distribution of spatial prices is estimated so that we can test the "Law of One Price," or spatial price parity. To model the joint probability distribution, we use copula models, which are becoming widespread methods. This paper uses a new flexible and tractable copula model, *the factor Durante generator (FDG) copula*, which allows for tail asymmetry to study the price linkages among markets. With interest in the OSB market, we apply this copula model to 6 important markets for OSB that are spatially separated across the US using 1996-2021 price data. After comparing the Mean Absolute Error (MAE) and tail dependence coefficients of FDG, Gaussian, t and Vine Copulas based on the data, we conclude that Vine copulas are better at modeling the association of markets, while the FDG Frechet copula and t copula are better at capturing tail dependence.

In the second chapter, we propose an alternative approach to analyzing the law of one price. We consider a continuous time threshold model using SDE to describe the changes in prices across markets with the "neutral band" resulting from transactions costs. This model is applied to weekly prices for lumber products. After utilizing a generalized impulse

response function to evaluate the dynamic paths of adjustment to shocks, it confirms converging responses consistent with integration of markets.

In the third chapter, we apply a cointegration model and a Threshold Vector Error Correction Model (TVECM) using the basis (i.e., the the error correction term) as the forcing variable to examine the relationship between cash and futures prices in corn, soybean, wheat, cotton and lean hogs markets before and during COVID. Our results support the cointegration relationships both in pre-COVID and COVID periods in those markets, and confirm the asymmetric price adjustment responding to positive and negative shocks. We do observe different results in pre-COVID and COVID periods and the reasons could be the changes of inventory levels, discruption of supply chain and decrease in foods demand caused by COVID.

© Copyright 2022 by Shu Meng

All Rights Reserved

Three Essays on Agricultural Commodity Market Linkages

by
Shu Meng

A dissertation submitted to the Graduate Faculty of
North Carolina State University
in partial fulfillment of the
requirements for the Degree of
Doctor of Philosophy

Economics

Raleigh, North Carolina
2022

APPROVED BY:

Walter Thurman	Denis Pelletier
Nicholas Piggott	Barry Goodwin Chair of Advisory Committee

DEDICATION

To my parents.

BIOGRAPHY

Shu Meng was born and raised in Beijing, China. She received her Bachelor of Science in Statistics from Beijing Normal University, then she started her journey pursuing higher education in Economics.

ACKNOWLEDGEMENTS

First and foremost, I would like to express my deepest gratitude to my advisor, Dr. Barry Goodwin, for his insightful advice, and kind support. In addition to his guidance and assistance on research, he has always been caring and understanding. Barry has also been my role model who inspires me to be a better person.

I extend my gratitude to the members of my dissertation committee. I would like to express my appreciation to Dr. Walter Thurman for his valuable comments and encouragement. My gratitude also goes to Dr. Denis Pelletier and Dr. Nick Piggott for their professional feedback and suggestions.

Last but not least, I want to express my utmost gratitude to my parents for their unconditional support and endless love, without which I would never have been able to make this accomplishment.

TABLE OF CONTENTS

List of Tables . . . . . **vii**

List of Figures . . . . . **ix**

Chapter 1 Modeling Spatial Price Parity Using High-Ordered Models of Dependence **1**
1.1 Introduction 1
1.2 Empirical Framework 5
1.2.1 Copula 5
1.2.2 The FDG Copula 6
1.2.3 Vine Copula 6
1.2.4 Other Copulas 7
1.2.5 Marginals 8
1.3 Empirical Application 9
1.3.1 Data 10
1.3.2 Marginals 15
1.3.3 Copula Results 15
1.3.4 Pre-COVID Results 28
1.4 Conclusion 33

Chapter 2 Modeling Spatial Price Linkages Using a Threshold SDE Model **44**
2.1 Introduction 44
2.2 Empirical Framework 48
2.2.1 Brownian Motion 49
2.2.2 Ornstein-Uhlenbeck Model 50
2.2.3 The Threshold SDE Model 50
2.3 Empirical Application 52
2.3.1 Data 52
2.3.2 Thresholds Selection 53
2.3.3 Results 55
2.3.4 Impulse Response Analysis 61
2.4 Conclusion 62

Chapter 3 An Analysis of the Relationship Between Agricultural Commodity Markets and Financial Markets: A Regime Switching Approach **65**
3.1 Introduction 66
3.2 Empirical Framework 72
3.2.1 Engle–Granger two-step cointegration model 72
3.2.2 Threshold Vector Error Correction Model 73

3.3 Empirical Application 75
3.3.1 Data 75
3.3.2 Cointegration Test 83
3.3.3 Threshold VECM 83
3.3.4 Impulse Response Analysis 86
3.4 Results 87
3.4.1 Cointegration Test 87
3.4.2 Threshold VECM 96
3.4.3 Impulse Response Analysis 101
3.5 Conclusion 103

References **105**

APPENDIX **109**
Appendix A Map of OSB Market Zones 110

LIST OF TABLES

Table 1.1 Statistics of Six Markets 13
Table 1.2 ADF Test Result 14
Table 1.3 Goodness of Fit for Marginals 16
Table 1.4 Parameter Estimates for Marginals 19
Table 1.5 ρ: Correlation of Six Markets 19
Table 1.6 λ^U: Upper Tail Dependence of Six Markets 20
Table 1.7 λ^L: Lower Tail Dependence of Six Markets 21
Table 1.8 Bivariate Copula Selection 22
Table 1.9 Multivariate Copulas Comparison 23
Table 1.10 λ: Upper and Lower Tail Dependence Estimated from FDG.frechet Copula 24
Table 1.11 λ: Upper and Lower Tail Dependence Estimated from t Copula 24
Table 1.12 Decomposition of Vine Copulas 25
Table 1.13 λ: Upper and Lower Tail Dependence Estimated from D-Vine Copula 26
Table 1.14 λ: Upper and Lower Tail Dependence Estimated from C-Vine Copula 27
Table 1.15 Statistics of Six Markets before COVID 28
Table 1.16 ADF test result before COVID 28
Table 1.17 Goodness of Fit for Marginals before COVID 30
Table 1.18 Parameter Estimates for Marginals before COVID 31
Table 1.19 λ^U: Upper Tail Dependence of Six Markets before COVID 31
Table 1.20 λ^L: Lower Tail Dependence of Six Markets before COVID 31
Table 1.21 Bivariate Copula Selection before COVID 32
Table 1.22 Multivariate Copulas Comparison before COVID 33

Table 2.1 Statistics of Four Markets: Prices 54
Table 2.2 Statistics of Six Pairs of Log Price Differentials 54
Table 2.3 Parameters of West-Southeast 57
Table 2.4 Thresholds of West-Southeast 57
Table 2.5 Comparison Between Real Data and Simulated Data of West-Southeast 58
Table 2.6 Parameters of West-Southwest 58
Table 2.7 Thresholds of West-Southwest 58
Table 2.8 Comparison Between Real Data and Simulated Data of West-Southwest 58
Table 2.9 Parameters of West-Southcentral 58
Table 2.10 Thresholds of West-Southcentral 59
Table 2.11 Comparison Between Real Data and Simulated Data of West-Southcentral 59
Table 2.12 Parameters of Southeast-Southcentral 59
Table 2.13 Thresholds of Southeast-Southcentral 59

Table 2.14 Comparison Between Real Data and Simulated Data of Southeast-Southcentral . . . 59
Table 2.15 Parameters of Southwest-Southcentral . . . 60
Table 2.16 Thresholds of Southwest-Southcentral . . . 60
Table 2.17 Comparison Between Real Data and Simulated Data of Southwest-Southcentral . . . 60
Table 2.18 Parameters of Southwest-Southeast . . . 60
Table 2.19 Thresholds of Southwest-Southeast . . . 60
Table 2.20 Comparison Between Real Data and Simulated Data of Southwest-Southeast . . . 61

Table 3.1 Chow Test Results . . . 77
Table 3.2 Statistics of Five Markets: Prices . . . 77
Table 3.3 Statistics of Five Markets: Log Prices . . . 78
Table 3.4 ADF Test for Prices Series: p value . . . 84
Table 3.5 Cointegration Test Results . . . 84
Table 3.6 Cointegration Estimates . . . 84
Table 3.7 F Test for Threshold Effect . . . 87
Table 3.8 Estimates of Corn Market before COVID . . . 87
Table 3.9 Estimates of Corn Market during COVID . . . 93
Table 3.10 Estimates of Soybean Market before COVID . . . 94
Table 3.11 Estimates of Soybean Market during COVID . . . 95
Table 3.12 Estimates of Wheat Market before COVID . . . 95
Table 3.13 Estimates of Wheat Market during COVID . . . 96
Table 3.14 Estimates of Cotton Market before COVID . . . 97
Table 3.15 Estimates of Cotton Market during COVID . . . 97
Table 3.16 Estimates of Lean Hogs Market before COVID . . . 98
Table 3.17 Estimates of Lean Hogs Market during COVID . . . 98

LIST OF FIGURES

Figure 1.1 Prices 12
Figure 1.2 First Difference of Log Prices 14
Figure 1.3 Marginals: x1 x2 x3 17
Figure 1.4 Marginals: x4 x5 x6 18
Figure 1.5 Scatters of Six Markets 20
Figure 1.6 FDG.Frechet Copula 35
Figure 1.7 t Copula 35
Figure 1.8 C-Vine Copula 36
Figure 1.9 D-Vine Copula 36
Figure 1.10 β Estimated from t Copula 37
Figure 1.11 β Estimated from FDG.frechet Copula 37
Figure 1.12 β Estimated from D-Vine Copula 38
Figure 1.13 β Estimated from C-Vine Copula 38
Figure 1.14 Prices before COVID 39
Figure 1.15 First Difference of Log Prices before COVID 39
Figure 1.16 Marginals before COVID: $x1_{pc}$ $x2_{pc}$ $x3_{pc}$ 40
Figure 1.17 Marginals before COVID: $x4_{pc}$ $x5_{pc}$ $x6_{pc}$ 41
Figure 1.18 β Estimated from FDG.frechet Copula 42
Figure 1.19 β Estimated from D-Vine Copula 42
Figure 1.20 β Estimated from C-Vine Copula 43

Figure 2.1 Plywood Prices 53
Figure 2.2 GIRFs of Six Pairs of Markets 63

Figure 3.1 Corn Prices 78
Figure 3.2 Soybean Prices 79
Figure 3.3 Wheat Prices 80
Figure 3.4 Cotton Prices 81
Figure 3.5 Lean Hogs Prices 82
Figure 3.6 Responses in Corn Market: Before COVID 88
Figure 3.7 Responses in Corn Market: During COVID 88
Figure 3.8 Responses in Soybean Market: Before COVID 89
Figure 3.9 Responses in Soybean Market: During COVID 89
Figure 3.10 Responses in Wheat Market: Before COVID 90
Figure 3.11 Responses in Wheat Market: During COVID 90
Figure 3.12 Responses in Cotton Market: Before COVID 91
Figure 3.13 Responses in Cotton Market: During COVID 91
Figure 3.14 Responses in Lean hogs Market: Before COVID 92
Figure 3.15 Responses in Lean hogs Market: During COVID 92

Figure A.1 Map of OSB Markets in US and Canada 111

CHAPTER

1

MODELING SPATIAL PRICE PARITY USING HIGH-ORDERED MODELS OF DEPENDENCE

1.1 Introduction

The single market does not stand alone in an economy, and the price in one specific market is always influenced by the prices in other related markets. Therefore, it is important to investigate the price linkages among markets, or the degree of market integration. The "Law

of One Price"(LOP) implies that prices for homogeneous products at different geographic locations in freely functioning markets should differ by no more than transport and transactions costs. There has been considerable research on the empirical validity of the LOP and numerous measures have been developed to test it, that is, to analyze the degree of market integration. Early studies used the correlation coefficient as an index of market integration and used linear regression-based tests to study spatial market integration (Fackler and Goodwin (2001)). Recent research has turned to nonlinear models in order to better capture the nonlinearities inherent in spatial price linkages and empirical analysis provides strong evidence that copula models better capture the nonlinearities (Goodwin et al. (2018)).

Copulas are used to model the dependence between random variables and are used by researchers who are interested in an empirical representation of the joint probability distribution of the spatial prices and in examination of their dependence. While a joint distribution implicitly describes the marginal behavior of individual variables and their dependency structure, copula models can isolate the dependency structure (Aas et al. (2009)). Sklar (1959) has established methods modeling dependence among a set of random variables using copula models. Copulas allow a parametric or semi-parametric representation of a joint probability distribution function in terms of a single (or composite) function $C(\cdot)$ and the individual marginal distributions $F(\cdot)$ of individual random variables. Copulas extend the nonlinear time-series models that have recently been applied to models of spatial price parity. The models attempt to represent unobservable transactions costs, which result in nonlinear price linkages, by considering a variety of copula models. A recent example includes the work of Goodwin et al. (2018).

An abundance of bivariate copula specifications have been proposed in the existing literature. This includes Archimedean copulas, symmetric elliptical copulas, and mixtures. Much less is known about copulas that are appropriate for higher-ordered ($n > 2$) models,

where n is the number of arguments of $C(\cdot)$. Recent research has developed vines (Aas et al. (2009)), hierarchical copulas (Joe (1997)), and factor copulas (Krupskii and Joe (2013)) to empirically model such higher-ordered copulas. Each of these methods, however, imposes a strong set of parametric assumptions that serve to limit the flexibility of copulas. Methods for pursuing more flexible higher-ordered copula models are the focus of ongoing research. Recent research has identified methods that may provide a more flexible approach to modeling these higher-ordered problems. Mazo et al. (2015) recently introduced a copula model that may offer such improved flexibility—the factor Durante generator (FDG) copula. This copula adopts simplifying assumptions that are much more flexible, while still maintaining tractability and flexibility and allowing for tail asymmetry. In the case of FDG copulas, a single-factor copula model is constructed in terms of appropriate linking copulas. These FDG linking copulas are from the Durante class of bivariate copulas, which consist of copulas of the form $C(u,v) = min(u,v)f(max(u,v))$ where u is a common factor, v is a uniform transformed random variable, and f is a differentiable and monotonically increasing function such that $f(1) = 1$. The use of a Durante linking function makes the FDG copula nonparametric.

We use the FDG copula to model the dependence of the spatial price relationships in the US and Canadian oriented strand board (OSB) markets. The North America region (especially the US and Canada) continues to be the world's top producer, consumer, and exporter of forest products. Since the passage of NAFTA, international transactions between U.S. and Canada became higher than that in the past. The OSB market is the largest sector of the overall panel wood products market in North America. Oriented strand board accounts for the most wood products used in structural panel applications in new construction. For example, OSB accounted for roughly two-thirds of wood panels used in wall sheathing in 2019 according to Builder Practices Survey data compiled by the National Association of

Home Builders (NAHB). During the first six months of 2021, manufacturers of OSB saw a huge price surge. Summertime new home construction reached its highest levels in over a decade, leading homebuilders to make large purchases of paneling. As of July 2021, per the National Association of Home Builders (NAHB), the price of OSB has seen a 510% increase since January 2020 and even a 650% to 662% increase for certain OSB products, like 3/8" OSB sheathing. Causes for the dramatic price increase cover a number of contributing factors, largely related to supply and demand of OSB. The pandemic and COVID-19 have also had an impact on the OSB market. At the initial outbreak in 2020, mills cut down on production while also unloading inventory due to concerns about an imminent housing crash. However, the opposite is what happened. Low interest rates in the US, due to a recession, helped spur a boom in housing while people in the US also took up Do It Yourself (DIY) projects and proceeded to buy up materials. Healthcare-related construction also saw an increase during the pandemic, elevating demand for OSB. In addition, in 2021, OSB production was very delayed because the resin, the glue, came from China and deliveries were constrained, which caused a supply problem.

In this paper, we consider 6 important markets for OSB that are spatially separated across North America. These markets are of particular importance in light of recent trade disputes involving wood materials trade with Canada and large fluctuations in the demand for construction materials. The data set consists of OSB for the Western Canada market, Eastern Canada market, North Central, Mid-Atlantic, Southwest and Southeast US markets. The six markets yield 15 pairs of bivariate price comparisons. Using standard notation, this defines copulas of the form $F(p_1, \dots p_6) = C(F(p_1), \dots, F(p_6))$. Then we compare the Mean Absolute Error (MAE) of Spearman's ρ and tail dependence coefficient λ of FDG Copula, as well as Gaussian, t and Vine Copulas based on the data, and discuss the advantages and disadvantages of different types of copulas.

1.2 Empirical Framework

It has become more popular to use copulas to model high ordered dependence in recent research, such as vines (Aas et al.(2009)), hierarchical copulas (Joe(1997)), and factor copulas (Krupskii and Joe(2013)). However, each of these methods has a strong set of parametric assumptions that limits their flexibility. To solve the problem, Mazo et al. (2015) recently introduced the factor Durante generator (FDG) copulas, that maintain tractability and flexibility while allowing for tail asymmetry.

1.2.1 Copula

The word copula is a Latin noun that means "a link, tie, bond," which is a way of studying scale-free measures of dependence in statistics. Sklar's theorem is the building block of the theory of copulas (Sklar (1959)), which is stated as follows:
Let F be a d-dimensional distribution function with univariate margins $F_1, F_2, \ldots, F_d$. Then there exists a copula C such that for all $(x_1, x_2, \ldots, x_d) \in \mathbb{R}^d$,

$$F(x_1, x_2, \ldots x_d) = C(F_1(x_1), F_2(x_2), \ldots, F_d(x_d)) \tag{1.1}$$

Copula models are well suited to considering tail dependence, which quantifies the degree of dependence in the joint tail of a multivariate distribution function. For a bivariate distribution, the coefficient of lower tail dependence λ_L and upper tail dependence λ_U are defined by

$$\lambda^L := \lim_{u \to 0} \frac{C(u,u)}{u} = \lim_{u \to 0} P(X_1 \leq F_1^{-1}(u) | X_2 \leq F_2^{-1}(u)) = \lim_{u \to 0} P(U_1 \leq u | U_2 \leq u) \tag{1.2}$$

$$\lambda^U := \lim_{u \to 1} \frac{1 - 2u + C(u,u)}{1-u} = \lim_{u \to 1} P(U_1 > u | U_2 > u) \tag{1.3}$$

And Spearman's ρ is defined as

$$\rho = 12 \int_0^1 \int_0^1 C(u_1, u_2) d\,u_1 d\,u_2 - 3 \tag{1.4}$$

1.2.2 The FDG Copula

The construction of the FDG Copula is:

$$\begin{aligned} C(u_1, \ldots, u_d) = u_{(1)}[(\prod_{j=2}^{d} u_{(j)}) \int_{u_{(d)}}^{1} \prod_{j=1}^{d} f_j^{'}(x) d\,x \\ + f_{(1)}(u_{(2)})(\prod_{j=2}^{d} f_{(j)}(u_{(j)})) \\ + \sum_{k=3}^{d} (\prod_{j=2}^{k-1} u_{(j)})(\prod_{j=k}^{d} f_{(j)}(u_{(j)})) \int_{u_{(k-1)}}^{u_{(k)}} \prod_{j=1}^{k-1} f_j^{'}(x) d\,x] \end{aligned} \tag{1.5}$$

Following Mazo et al. (2015), C is a one-factor copula, in which the linking copula belongs to the Durante class with given generator f_i, and $u_{(i)} := u_{\sigma(i)}, f_{(i)} := f_{\sigma(i)}$ and σ is the permutation of $(1, \ldots, d)$ such that $u_{\sigma(1)} \leq \cdots \leq u_{\sigma(d)}$.

1.2.3 Vine Copula

We concentrate on two special cases of regular vines, the C-vine and the D-vine.

The d-dimensional density corresponding to a C-vine is given by

$$\prod_{k=1}^{d} f(x_k) \prod_{j=1}^{d-1} \prod_{i=1}^{d-j} c_{j,j+i|1,\ldots,j-1}(F(x_j|x_1, \ldots, x_{j-1}), F(x_{j+i}|x_1, \ldots, x_{j-1})) \tag{1.6}$$

The d-dimensional density corresponding to a D-vine is given by

$$\prod_{k=1}^{d} f(x_k) \prod_{j=1}^{d-1} \prod_{i=1}^{d-j} c_{i,i+j|i+1,\ldots,i+j-1}(F(x_i|x_{i+1},\ldots,x_{i+j-1}), F(x_{i+j}|x_{i+1},\ldots,x_{i+j-1})) \tag{1.7}$$

Following Aas et al. (2009), for example, the general expression for the five-dimensional C-vine structure is

$$\begin{aligned} f(x_1,x_2,x_3,x_4,x_5) = {} & f(x_1)\cdot f(x_2)\cdot f(x_3)\cdot f(x_4)\cdot f(x_5) \\ & \cdot c_{12}(F(x_1),F(x_2))\cdot c_{13}(F(x_1),F(x_3))\cdot c_{14}(F(x_1),F(x_4))\cdot c_{15}(F(x_1),F(x_5)) \\ & \cdot c_{23|1}(F(x_2|x_1),F(x_3|x_1))\cdot c_{24|1}(F(x_2|x_1),F(x_4|x_1))\cdot c_{25|1}(F(x_2|x_1),F(x_5|x_1)) \\ & \cdot c_{34|12}(F(x_3|x_1,x_2),F(x_4|x_1,x_2))\cdot c_{35|12}(F(x_3|x_1,x_2),F(x_5|x_1,x_2)) \\ & \cdot c_{45|123}(F(x_4|x_1,x_2,x_3),F(x_5|x_1,x_2,x_3)) \end{aligned} \tag{1.8}$$

and the general expression for the D-vine structure is

$$\begin{aligned} f(x_1,x_2,x_3,x_4,x_5) = {} & f(x_1)\cdot f(x_2)\cdot f(x_3)\cdot f(x_4)\cdot f(x_5) \\ & \cdot c_{12}(F(x_1),F(x_2))\cdot c_{23}(F(x_2),F(x_3))\cdot c_{34}(F(x_3),F(x_4))\cdot c_{45}(F(x_4),F(x_5)) \\ & \cdot c_{13|2}(F(x_1|x_2),F(x_3|x_2))\cdot c_{24|3}(F(x_2|x_3),F(x_4|x_3))\cdot c_{35|4}(F(x_3|x_4),F(x_5|x_4)) \\ & \cdot c_{14|23}(F(x_1|x_2,x_3),F(x_4|x_2,x_3))\cdot c_{25|34}(F(x_2|x_3,x_4),F(x_5|x_3,x_4)) \\ & \cdot c_{15|234}(F(x_1|x_2,x_3,x_4),F(x_5|x_2,x_3,x_4)) \end{aligned} \tag{1.9}$$

1.2.4 Other Copulas

We also compare the results of FDG Copula with Gaussian Copula and t Copula. To be more specific, we estimate the Gaussian and t Copula out and simulate from the fitted Copulas,

then compare the correlations with that of original data.

For multivariate Copulas, we compare the estimated spatial dependence relationships to simple empirical measures involving Spearman's ρ and tail dependence λ. Then we compare the MAE_ρ (Mean Absolute Error of ρ), MAE_λ (Mean Absolute Error of λ) and ME_λ (Mean Error of λ) to choose the best Copula. The definitions are as follows:

$$MAE_\rho = \frac{1}{p}\sum_{i<j}|\rho_{empirical} - \rho_{estimated}| \tag{1.10}$$

$$MAE_{\lambda^L} = \frac{1}{p}\sum_{i<j}|\lambda^L_{empirical} - \lambda^L_{estimated}| \tag{1.11}$$

$$MAE_{\lambda^U} = \frac{1}{p}\sum_{i<j}|\lambda^U_{empirical} - \lambda^U_{estimated}| \tag{1.12}$$

$$ME_{\lambda^L} = \frac{1}{p}\sum_{i<j}(\lambda^L_{estimated} - \lambda^L_{empirical}) \tag{1.13}$$

$$ME_{\lambda^U} = \frac{1}{p}\sum_{i<j}(\lambda^U_{estimated} - \lambda^U_{empirical}) \tag{1.14}$$

We have 6 markets and 15 pairs here, so $p = 15$.

1.2.5 Marginals

We estimate the marginals and copulas in a two-step procedure. We use nonparametric representations of the marginal densities by considering the quantiles of the distributions and focus on a parametric specification of the copula model representing dependence. We use ranks of each individual random variable to determine nonparametric quantile

functions. It is common to start with plots of histograms to characterize the empirical distribution and assume the specific families of the marginals. Then distribution parameters θ are by default estimated by maximizing the likelihood function.

1.3 Empirical Application

For the empirical analysis, we consider six North American markets for oriented strand board (OSB). OSB is a widely used, versatile structural wood panel. The Structural Board Association (SBA) reports that OSB panel production in North America was 751 million square feet (on a 3/8-inch basis) in 1980, but that this number had grown to 25 billion square feet by as early as 2005. The SBA also reports that OSB production enjoyed a 60% market share among all panel products in North America by 2006. With the bulk of OSB produced in North America originating in the southern United States and Canada, Canada and the southern U.S. produced nearly 90% of all OSB produced in North America in 2009 and 2010 (Engineered Wood Product Association 2010). A growth in domestic housing markets (approximately 1991 through 2006), the sharp contraction in U.S. new home construction (2006-2008) and the corresponding economic downturn have brought about a number of significant shocks to this rapidly expanding industry. Construction market responses to large hurricanes such as Katrina in 2005 are another source of OSB market price volatility (Goodwin et al. (2011); Goodwin et al. (2018)).

Recently, North America dominated the Oriented Strand Board Market with a global share of more than 45.4%. OSB, traded at 1527 dollars per thousand board feet in June 2021, marked a 97% increase since the start of 2021, according to Bloomberg data. The rally in OSB is due in part to a storm in Texas earlier this year that caused a shortage in resin, a chemical substance needed to make the product.

The wood industry has been one of the biggest beneficiaries of the pandemic, as more people built and renovated their homes. According to the Engineered Wood Association (APA), new home construction is the primary end-user for the OSB industry in North America, accounting for more than 50% of the OSB consumption. When new home construction and home-improvement sales boomed earlier this year, inventory for lumber and OSB, among other commodities, were running low which could be one of the causes for the dramatic price increase of OSB.

We are interested in the spatial market linkages of OSB because it is widely traded in North America and we would like to know whether the degree of market integration would be varied when the price change is relatively high, just like what happened during the pandemic.

1.3.1 Data

The data set consists of six markets: (1) North Central US (a producing region for OSB that includes Minnesota, Michigan, and Wisconsin); (2) Southwest US (a producing zone for OSB that includes Texas, Louisiana, Arkansas, and Oklahoma); (3) Southeast US (a producing zone for OSB consisting of the states of Georgia, Alabama, Mississippi, South Carolina, and Tennessee); (4) Eastern Canada (a producing region for OSB encompassing plants in Ontario and Quebec); (5) Mid-Atlantic US (a producing zone for OSB that includes North Carolina, Virginia, and West Virginia); and (6) Western Canada (a producing region for OSB that includes B.C., Alberta, Manitoba, and Saskatchewan). You can see the corresponding states in the map in the Appendix, Figure A.1. Therefore we need to examine 15 pairs of price relationships. The price data are obtained from *Random Lengths* and for panels of 7/16th's inch thickness in US dollars per thousand square feet.

For the data collection process of *Random Lengths*, Editors on the staff contact sources who are directly involved in selling or buying the products. The number of sources contacted varies, depending on the product group. Some sources are contacted each week, some are contacted every two or three weeks, and some are contacted on an irregular basis. Sources of price and market information range from large corporations to small, single-mill operations, and from large wholesale/distribution firms to one-person wholesale offices. Calls are made to buyers as well as producers to get equal input from both sides of the trading table. The types of buyers contacted include office wholesalers, stocking distributors, co-ops, and buyers for treating operations and secondary manufacturing plants. Retail buying offices that are not buying directly from mills are not contacted; however, sources can include centralized buying offices for retail operations.

The data contains 1286 weekly observations, from September 6, 1996, to July 23, 2021. The price data are illustrated in Figure 1.1.

- P_1: PAAS = OSB Sheathing (North Central) 7/16" Prices Net f.o.b. Mill
- P_2: PACS = OSB Sheathing (Southwest) 7/16" Prices Net f.o.b. Mill
- P_3: PADS = OSB Sheathing (Southeast) 7/16" Prices Net f.o.b. Mill
- P_4: PAKU = OSB Sheathing (E Canada) 7/16" Prices Net f.o.b. Mill
- P_5: PALB = OSB Sheathing (Mid Atl) 7/16" Prices Net f.o.b. Mill
- P_6: PALI = OSB Sheathing (W Canada) 7/16" Prices Net f.o.b. Mill

Firstly, we need to prepared the six data series to be stationary, so we take the first difference of the log prices. That is,

- $x_1 = lnP_{1,t} - lnP_{1,t-1}$;

- $x_2 = lnP_{2,t} - lnP_{2,t-1}$;
- $x_3 = lnP_{3,t} - lnP_{3,t-1}$;
- $x_4 = lnP_{4,t} - lnP_{4,t-1}$;
- $x_5 = lnP_{5,t} - lnP_{5,t-1}$;
- $x_6 = lnP_{6,t} - lnP_{6,t-1}$;

And the ADF test results are presented in Table 1.2.

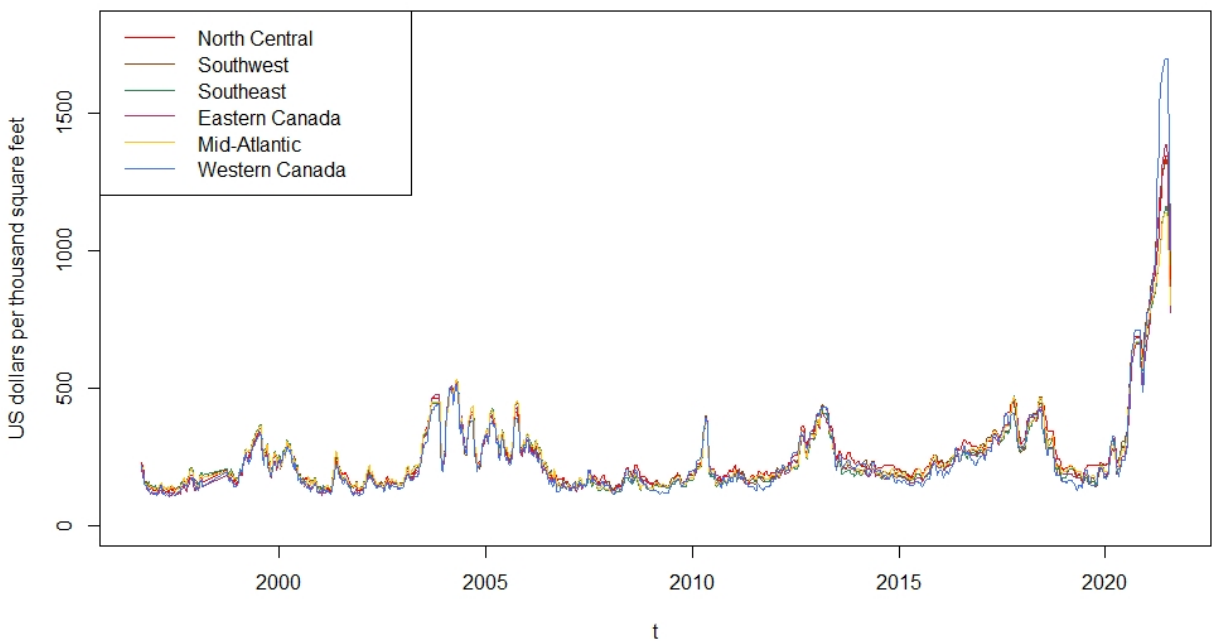


Figure 1.1: Prices

Table 1.1 shows the statistics of x_1, x_2, x_3, x_4, x_5, x_6 and from it we can see the average range of percentage change of prices among six markets is from -0.36 to 0.24, where Western

Table 1.1: Statistics of Six Markets

	min	mean	max
x1:North Central	-0.390	0.001	0.242
x2:Southwest	-0.336	0.001	0.228
x3:Southeast	-0.336	0.001	0.213
x4:Eastern Canada	-0.381	0.001	0.255
x5:Mid-Atlantic	-0.336	0.001	0.213
x6:Western Canada	-0.372	0.001	0.303

Canada market shows the highest postive price change and the North Central market shows the highest negative price change. Then x_1, x_2, x_3, x_4, x_5, x_6 pass the ADF test according to Table 1.2, so we are modeling $F(x_1, x_2, x_3, x_4, x_5, x_6)$ using $C(u_1, u_2, u_3, u_4, u_5, u_6)$, where $u_i = F(x_i), i = 1,2,3,4,5,6$. That is, we are modeling the relationship of percentage change in spatial prices in this paper. For example, if we obtain the result that x_1 and x_2 are highly dependent on each other, that means the percentage change of prices in market 1 is highly related to the percentage change of prices in market 2. That is, market 1 and market 2 are highly integrated. Also, if we regard the price relationship as equation (1.15), when $|\beta|$ is closer to 1, it means the degree of integration between market 1 and market 2 is higher. Therefore, using the copula models can give us nonlinear relationships between β and x_i, as presented in Figure 1.10-1.13 and Figure 1.18-1.20. That is, we could see the degree of market integration changes as the percentage change of price increases or decreases from those graphs.

$$x_2 = \beta x_1 + \epsilon \tag{1.15}$$

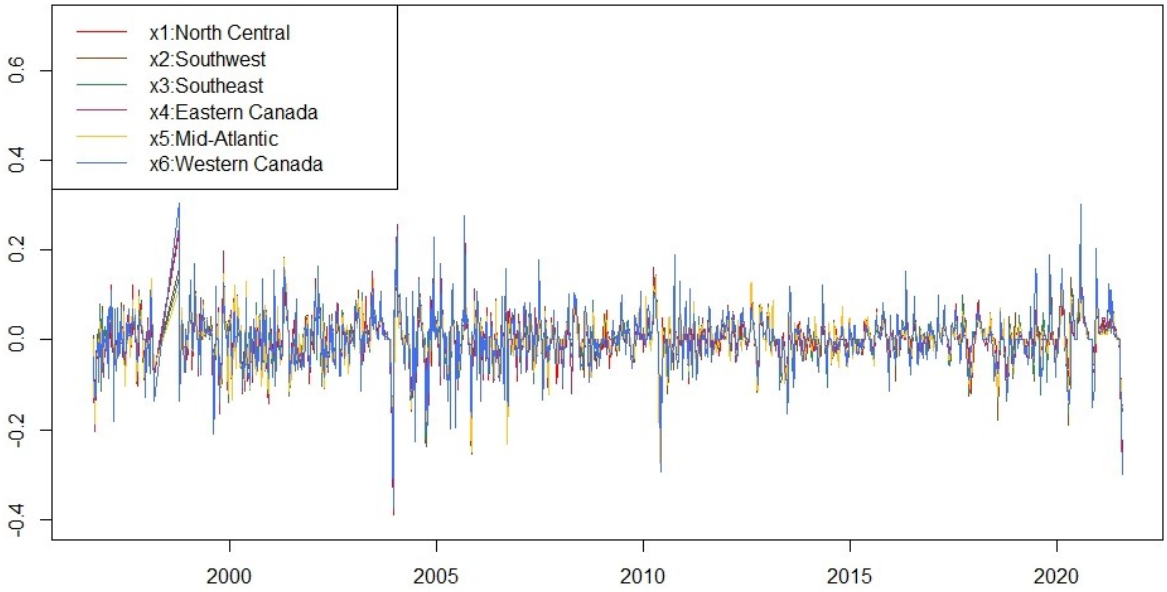


Figure 1.2: First Difference of Log Prices

Table 1.2: ADF Test Result

ADF test alternative hypothesis: stationary			
data	p value	data	p value
PAAS:North Central	0.03079	x1	<0.01
PACS:Southwest	0.5255	x2	<0.01
PADS:Southeast	0.4948	x3	<0.01
PAKU:Eastern Canada	<0.01	x4	<0.01
PALB:Mid-Atlantic	0.4292	x5	<0.01
PALI:Western Canada	<0.01	x6	<0.01

1.3.2 Marginals

We also need to estimate the marginals of each variables. Here we use logistic distributions for the six marginals according to the results in Table 1.3, Figure 1.3 and Figure 1.4. The parameter estimates are listed in Table 1.4.

To start estimating Copulas, it is necessary to transform $x_i's$ to $u_i's$, which are ready to be used to estimate the Copula models. First of all, we use the empirical CDFs of $x_i's$, that is, $u_i's$ to estimate the parameters of FDG Copulas, as well as Gaussian and t Copulas. Then we compare the empirical Spearman's ρ with different estimated ρ's and the empirical tail dependence coefficient λ with different estimated λ's. Lastly, we simulate data from the fitted Copulas and compare the correlation plots.

1.3.3 Copula Results

Figure 1.5, Table 1.5, Table 1.6 and Table 1.7 present the values of correlation ρ, tail dependence λ_U and λ_L, which are estimated using empirical Copulas (Schmidt and Stadtmüller (2006)). From Table 1.5 and Figure 1.5, ρ ranges from 0.78 to 0.95, which means the fifteen pairs of markets present relatively high dependence on each other. In particular, Southeast US market and Mid-Atlantic US market, and Southwest US market and Southeast US market are more strongly related to each other which may be caused by the geographical distance. From Table 1.6 and Table 1.7, North Central US market and Eastern Canada market, and Southeast US market and Mid-Atlantic US market have higher upper tail dependence, which means when the price increases a lot in one market, the price of other market has a high probability to increase as well. North Central US market and Eastern Canada market also share a relatively high lower tail dependence, which means when the price decreases a lot in one market, the price of the other market presumably would decrease correspondingly.

Table 1.3: Goodness of Fit for Marginals

	Normal			Logistic			Optimal
	AIC	BIC	lglk	AIC	BIC	lglk	
x1:North Central	-3932.605	-3922.316	1968.303	-4145.16	-4134.871	2074.58	Logistic
x2:Southwest	-3876.208	-3865.919	1940.104	-4057.647	-4047.358	2030.824	Logistic
x3:Southeast	-3819.305	-3809.016	1911.652	-3988.323	-3978.034	1996.161	Logistic
x4:Eastern Canada	-3894.26	-3883.971	1949.13	-4116.152	-4105.863	2060.076	Logistic
x5:Mid-Atlantic	-3847.203	-3836.914	1925.601	-4003.117	-3992.828	2003.559	Logistic
x6:Western Canada	-3555.244	-3544.956	1779.622	-3736.951	-3726.662	1870.475	Logistic

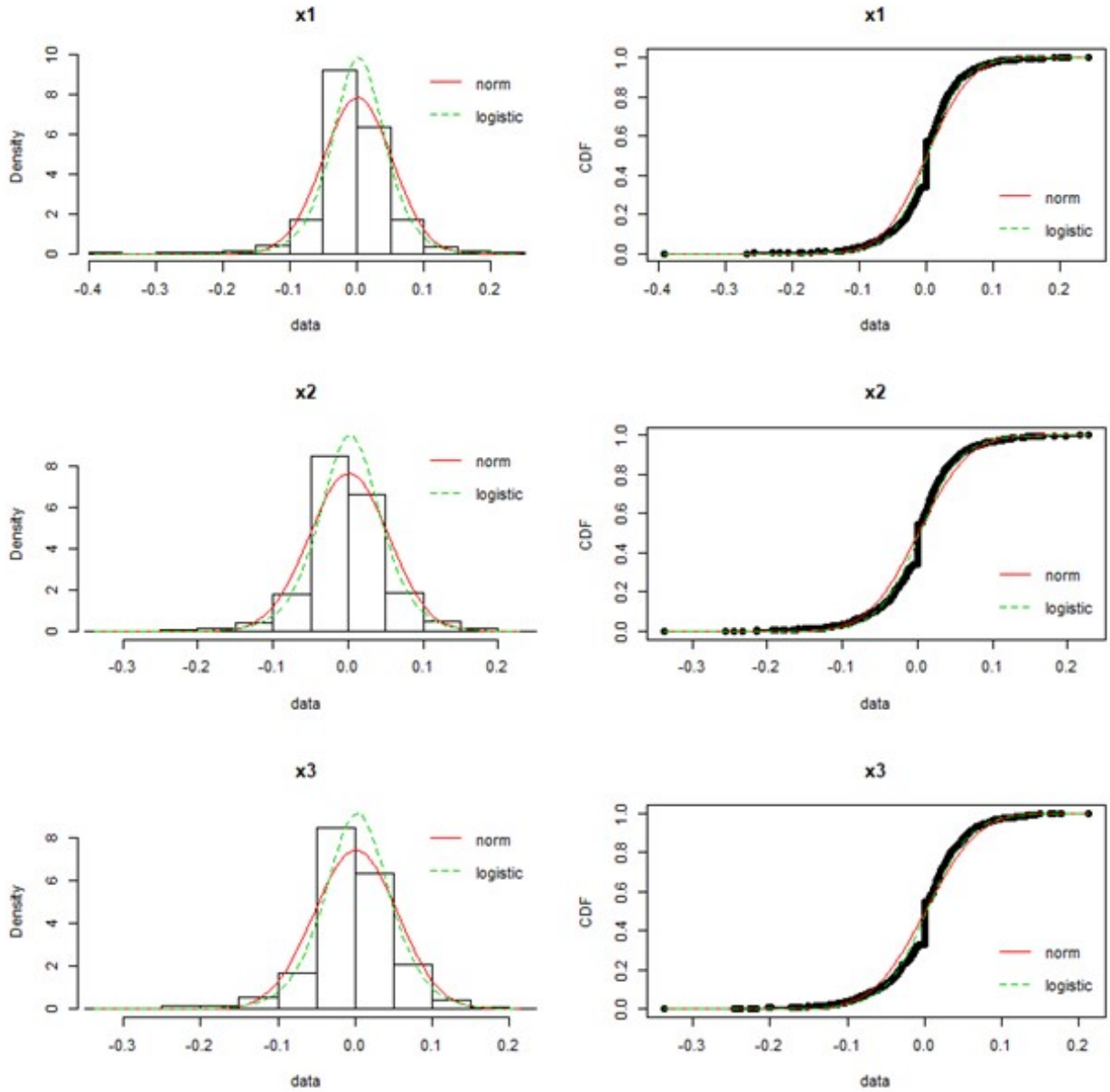


Figure 1.3: Marginals: x1 x2 x3

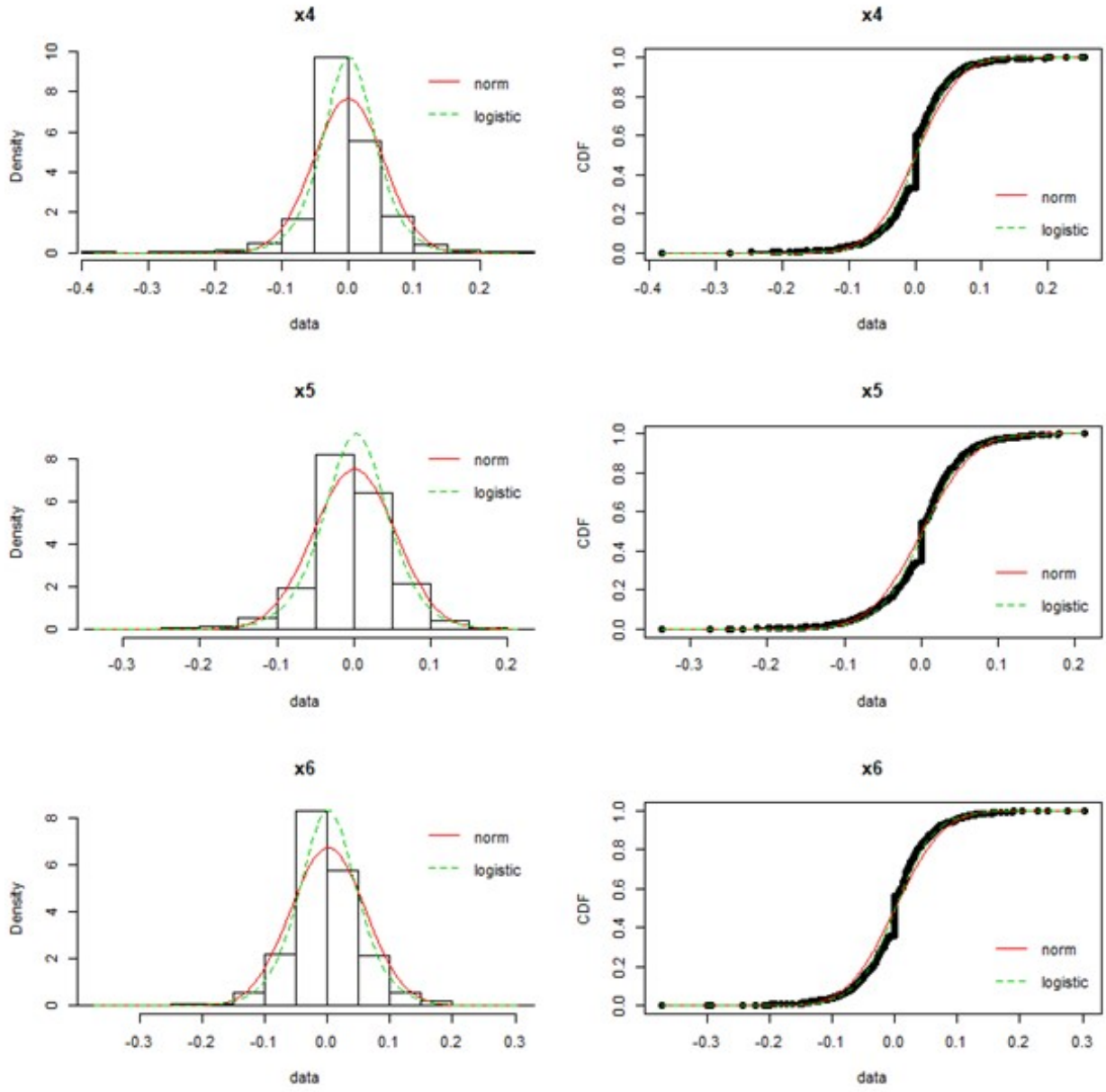


Figure 1.4: Marginals: x4 x5 x6

Table 1.4: Parameter Estimates for Marginals

	location		scale	
	estimate	*std error*	*estimate*	*std error*
x1:North Central	0.00234	0.001207	0.0254	0.0006116
x2:Southwest	0.00272	0.001254	0.0264	0.0006344
x3:Southeast	0.00309	0.001295	0.0272	0.0006524
x4:Eastern Canada	0.00186	0.001220	0.0257	0.00061793
x5:Mid-Atlantic	0.00283	0.001296	0.0271	0.0006490
x6:Western Canada	0.00172	0.001432	0.0301	0.0007199

Table 1.5: ρ: Correlation of Six Markets

	x1	x2	x3	x4	x5	x6
x1:North Central	1.00	0.83	0.83	0.89	0.85	0.83
x2:Southwest	0.83	1.00	0.91	0.82	0.90	0.80
x3:Southeast	0.83	0.91	1.00	0.84	0.95	0.78
x4:Eastern Canada	0.89	0.82	0.84	1.00	0.85	0.83
x5:Mid-Atlantic	0.85	0.90	0.95	0.85	1.00	0.79
x6:Western Canada	0.83	0.80	0.78	0.83	0.79	1.00

Bivariate Copula Results

Table 1.8 presents AIC and BIC criteria for the optimal bivariate copula specification and estimates results. All 15 pairs favor the t copula specification and all of them suggest significant tail dependence. Except for pairs of North Central US and Western Canada markets, Southwest US and Southeast US markets, and Southwest US and Mid-Atlantic US markets, estimated upper tail dependences in other pairs are lower than empirical upper tail dependences. Except for pairs of Southwest US and Southeast US markets, and Southeast US and Mid-Atlantic US markets, estimated lower tail dependences in other pairs are lower

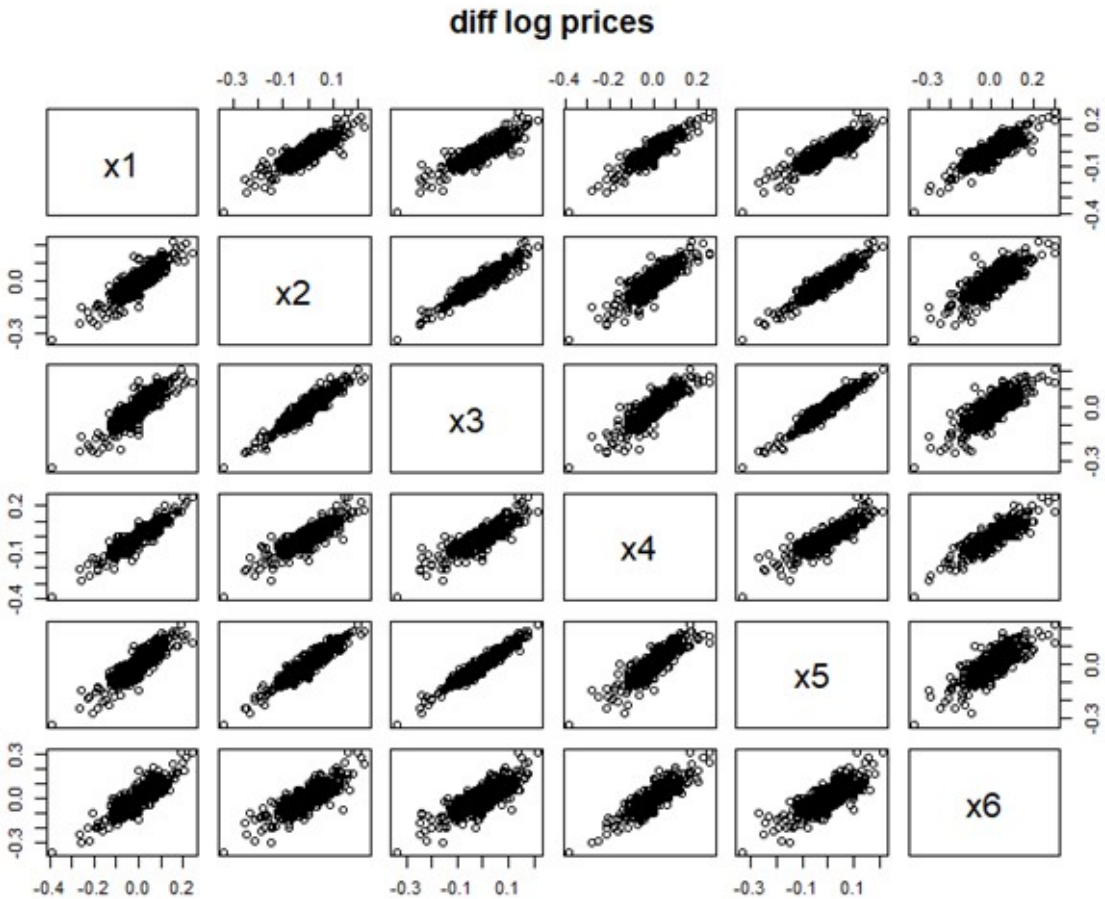


Figure 1.5: Scatters of Six Markets

Table 1.6: λ^U: Upper Tail Dependence of Six Markets

	x1	x2	x3	x4	x5	x6
x1:North Central	1.00	0.66	0.60	0.86	0.63	0.54
x2:Southwest	0.66	1.00	0.69	0.63	0.69	0.60
x3:Southeast	0.60	0.69	1.00	0.60	0.86	0.54
x4:Eastern Canada	0.86	0.63	0.60	1.00	0.69	0.51
x5:Mid-Atlantic	0.63	0.69	0.86	0.69	1.00	0.51
x6:Western Canada	0.54	0.60	0.54	0.51	0.51	1.00

Table 1.7: λ^L: Lower Tail Dependence of Six Markets

	x1	x2	x3	x4	x5	x6
x1:North Central	1.00	0.60	0.57	0.80	0.63	0.66
x2:Southwest	0.60	1.00	0.74	0.57	0.71	0.60
x3:Southeast	0.57	0.74	1.00	0.60	0.80	0.51
x4:Eastern Canada	0.80	0.57	0.60	1.00	0.66	0.63
x5:Mid-Atlantic	0.63	0.71	0.80	0.66	1.00	0.60
x6:Western Canada	0.66	0.60	0.51	0.63	0.60	1.00

than empirical lower tail dependences. Overall, the bivariate t copula could capture the tail dependence of the pairwise markets, but usually underestimates the tail dependence. In the case of t copula, $\hat{\theta}_1$ is the correlation parameter, and $\hat{\theta}_2$ is another parameter, the degrees of freedom. We could also conclude that t copula usually overestimates the correlation ρ of markets after comparing the $\hat{\theta}_1$ with the $\rho_{empirical}$ presented in Table 1.5.

Multivariate Copula Results

We compare Gaussian, t, Vine, fdg.frechet, fdg.cuadrasauge and fdg.sinus Copulas in this section. Table 1.9 presents the MAE_ρ, MAE_{λ^L}, MAE_{λ^U}, ME_{λ^L} and ME_{λ^U} of Gaussian, t, Vine and three families of FDG copulas. From the comparison of MAE_ρ's, C-Vine, D-Vine, FDG.frechet, FDG.cuadrasauge and t copulas are better than others. However, as the data set demonstrates obvious upper and lower tail dependences, it is also necessary to look at the MAE_λ's. FDG.frechet and t copula have smaller MAE_{λ^L}'s and MAE_{λ^U}'s than other copulas, which is also in concordance with theory. Theoretically, Gaussian copula have zero tail dependence, FDG.cuadrasauge copula has zero lower tail dependence, FDG.sinus copula has zero lower tail dependence, and only Vine, FDG.frechet and t copulas can capture both upper and lower tail dependences. Although t copula has lower MAE_λ's on both upper and

Table 1.8: Bivariate Copula Selection

markets	optimal copula	$\hat{\theta}_1$	$SE(\hat{\theta}_1)$	$\hat{\theta}_2$	$SE(\hat{\theta}_2)$	τ	λ_L	λ_U	ρ
(x1:North Central,x2:Southwest)	t	0.85	0.01	3.31	0.42	0.65	0.59	0.59	0.84
(x1:North Central,x3:Southeast)	t	0.85	0.01	3.58	0.5	0.65	0.57	0.57	0.84
(x1:North Central,x4:Eastern Canada)	t	0.9	0.01	2.13	0.22	0.72	0.72	0.72	0.89
(x1:North Central,x5:Mid-Atlantic)	t	0.86	0.01	4.39	0.68	0.66	0.55	0.55	0.85
(x1:North Central,x6:Western Canada)	t	0.85	0.01	2.6	0.31	0.65	0.62	0.62	0.84
(x2:Southwest,x3:Southeast)	t	0.92	0.01	2.02	0.22	0.75	0.75	0.75	0.91
(x2:Southwest,x4:Eastern Canada)	t	0.84	0.01	3.37	0.44	0.63	0.56	0.56	0.82
(x2:Southwest,x5:Mid-Atlantic)	t	0.91	0.01	2.59	0.32	0.73	0.71	0.71	0.90
(x2:Southwest,x6:Western Canada)	t	0.81	0.01	5.11	0.83	0.6	0.45	0.45	0.80
(x3:Southeast,x4:Eastern Canada)	t	0.84	0.01	4.05	0.6	0.64	0.54	0.54	0.83
(x3:Southeast,x5:Mid-Atlantic)	t	0.96	0.001	2	0.31	0.82	0.82	0.82	0.96
(x3:Southeast,x6:Western Canada)	t	0.79	0.01	5.36	0.95	0.58	0.42	0.42	0.78
(x4:Eastern Canada,x5:Mid-Atlantic)	t	0.86	0.01	3.77	0.53	0.66	0.57	0.57	0.85
(x4:Eastern Canada,x6:Western Canada)	t	0.83	0.01	4.4	0.72	0.63	0.51	0.51	0.82
(x5:Mid-Atlantic,x6:Western Canada)	t	0.8	0.01	5.07	0.85	0.59	0.44	0.44	0.78

Table 1.9: Multivariate Copulas Comparison

Copulas	MAE_ρ	$MAE_{\lambda L}$	$ME_{\lambda L}$	$MAE_{\lambda U}$	$ME_{\lambda U}$
fdg.frechet	0.0299	0.2014	0.2014	0.2089	0.2071
fdg.cuadrasauge	0.0299	0.6457	-0.6457	0.2413	0.2413
fdg.sinus	0.479	0.6457	-0.6457	0.36	0.36
Gaussian	0.0386	0.6457	-0.6457	0.64	-0.64
t	0.0293	0.0361	-0.0322	0.0467	-0.0265
CVine	0.0113	0.4128	-0.3492	0.4144	-0.3507
Dvine	0.0198	0.4662	-0.3642	0.4913	-0.3893

lower tail, it underestimates the tail dependence of the markets. In particular, for North Central US market and Eastern Canada market, the upper and lower tail dependence are around 0.8 while the estimates from t copula are around 0.6. FDG.frechet copula, on the contrary, overestimates the tail dependence. As the tail dependence coefficient describes the dependence between concurrent extremes at different locations, it is no problem to have overestimates so that we could take precausionary actions. Therefore, we could keep both t copula and FDG.frechet copula as our final models, which could give us a range of the estimated tail dependence. We also keep C-Vine and D-Vine copulas because they have the least MAE_ρ. Table 1.12 gives the specific decompositions of C-Vine and D-Vine copulas, and the Vine copulas are selected using AIC and Kendall's τ (Dissmann et al. (2013)).

In Figure 1.6, Figure 1.7, Figure 1.8 and Figure 1.9, blue points are simulated data from FDG.frechet copula, t Copula, C-Vine and D-Vine copula respectively while red points are the original data. Red numbers represent the empirical Pearson correlation coefficients and blue numbers are estimated coefficients from copula models. From Figure 1.6, Figure 1.7, Figure 1.8, Figure 1.9, Table 1.10, Table 1.11, Table 1.13 and Table 1.14, FDG.frechet copula presents a stronger tail dependence. By comparison with the empirical tail dependence represented by Table 1.6 and Table 1.7, FDG.frechet copula is the only one that gives over-

Table 1.10: λ: Upper and Lower Tail Dependence Estimated from FDG.frechet Copula

	x1	x2	x3	x4	x5	x6
x1:North Central	1.00	0.86	0.87	0.84	0.87	0.80
x2:Southwest	0.86	1.00	0.88	0.85	0.88	0.81
x3:Southeast	0.87	0.88	1.00	0.86	0.90	0.82
x4:Eastern Canada	0.84	0.85	0.86	1.00	0.87	0.79
x5:Mid-Atlantic	0.87	0.88	0.90	0.87	1.00	0.82
x6:Western Canada	0.80	0.81	0.82	0.79	0.82	1.00

Table 1.11: λ: Upper and Lower Tail Dependence Estimated from t Copula

	x1	x2	x3	x4	x5	x6
x1:North Central	1.00	0.59	0.59	0.68	0.61	0.59
x2:Southwest	0.59	1.00	0.71	0.58	0.69	0.54
x3:Southeast	0.59	0.71	1.00	0.60	0.79	0.52
x4:Eastern Canada	0.68	0.58	0.60	1.00	0.61	0.58
x5:Mid-Atlantic	0.61	0.69	0.79	0.61	1.00	0.52
x6:Western Canada	0.59	0.54	0.52	0.58	0.52	1.00

estimated tail dependence while other copulas provide underestimated tail dependence. For example, the empirical upper and lower tail dependence of Southeast US market and Mid-Atlantic US market are 0.86 and 0.8, and estimated coefficients of tail dependence from FDG.frechet copula are 0.9 and 0.9, estimates from t copula are 0.78 and 0.78, estimates from C-Vine copula are 0.15 and 0.15, and estimates from D-Vine copula are 0.07 and 0.07. Therefore, we could conclude that FDG.frechet copula would be a better choice when the data set presents obviously high tail dependence.

Now we use the data simulated from Vine, FDG.frechet and t copulas to plot the graph investigating the relationship between $\beta = \frac{\partial x_i}{\partial x_j}$ and x_j, $i > j, i, j = 1, 2, \ldots, 6$. We use locally estimated scatterplot smoothing (LOESS) to plot the relationship between β and x_j's. From Figure 1.10, Figure 1.11, Figure 1.12 and Figure 1.13, we can see in most cases, when x_j falls

Table 1.12: Decomposition of Vine Copulas

D-Vine		C-Vine	
Tree	*Edge*	*Tree*	*Edge*
1	3,2	1	5,3
	1,6		5,4
	4,1		5,2
	5,3		5,1
	5,4		6,5
2	5,2;3	2	1,3;5
	4,6;1		1,4;5
	5,1;4		1,2;5
	4,3;5		6,1;5
3	4,2;5,3	3	2,3;1,5
	5,6;4,1		2,4;1,5
	3,1;5,4		6,2;1,5
4	1,2;4,5,3	4	4,3;2,1,5
	3,6;5,4,1		6,4;2,1,5
5	6,2;1,4,5,3	5	6,3;4,2,1,5

in the interval of [-0.2, 0.2], β is around 1, which implies the law of one price. It can also be concluded that β varies when there are different percentage changes in prices, that is, different percentage changes in prices lead to different degrees of market integration. As our simulated data from copula models has less points falling within [0.21, 0.3] and [-0.3, -0.21], the polynomials we obtained may not accurately represent the relationships when price change falls in those intervals. Therefore, we could use the tail dependence coefficients estimated by FDG.frechet to more precisely describe the relationship of markets when prices change a lot. From Table 1.10, we could conclude that Southeast US market and Mid-Atlantic US market has the largest degree of market integration when prices increase or decrease dramatically.

In summary, our results support the law of one price for six oriented strand board market in North America and most pairs of markets appear to be highly linked to each

Table 1.13: λ: Upper and Lower Tail Dependence Estimated from D-Vine Copula

	λ_U						λ_L					
	x1	x2	x3	x4	x5	x6	x1	x2	x3	x4	x5	x6
x1:North Central	1.00	0.00	0.01	0.03	0.00	0.75	1.00	0.07	0.01	0.03	0.19	0.75
x2:Southwest	0.00	1.00	0.00	0.01	0.00	0.62	0.07	1.00	0.06	0.01	0.20	0.62
x3:Southeast	0.01	0.00	1.00	0.02	0.07	0.72	0.01	0.06	1.00	0.02	0.07	0.72
x4:Eastern Canada	0.03	0.01	0.02	1.00	0.14	0.82	0.03	0.01	0.02	1.00	0.00	0.82
x5:Mid-Atlantic	0.00	0.00	0.07	0.14	1.00	0.57	0.19	0.20	0.07	0.00	1.00	0.57
x6:Western Canada	0.75	0.62	0.72	0.82	0.57	1.00	0.75	0.62	0.72	0.82	0.57	1.00

Table 1.14: λ: Upper and Lower Tail Dependence Estimated from C-Vine Copula

	λ_U						λ_L					
	x1	x2	x3	x4	x5	x6	x1	x2	x3	x4	x5	x6
x1:North Central	1.00	0.00	0.06	0.15	0.01	0.82	1.00	0.03	0.00	0.15	0.02	0.82
x2:Southwest	0.00	1.00	0.00	0.00	0.43	0.57	0.03	1.00	0.17	0.00	0.43	0.57
x3:Southeast	0.06	0.00	1.00	0.11	0.15	0.71	0.00	0.17	1.00	0.00	0.15	0.71
x4:Eastern Canada	0.15	0.00	0.11	1.00	0.34	0.55	0.15	0.00	0.00	1.00	0.34	0.55
x5:Mid-Atlantic	0.01	0.43	0.15	0.34	1.00	0.44	0.02	0.43	0.15	0.34	1.00	0.44
x6:Western Canada	0.82	0.57	0.71	0.55	0.44	1.00	0.82	0.57	0.71	0.55	0.44	1.00

Table 1.15: Statistics of Six Markets before COVID

	min	mean	max
$x1_{pc}$:North Central	-0.390	-0.00002	0.242
$x2_{pc}$:Southwest	-0.336	-0.00017	0.228
$x3_{pc}$:Southeast	-0.336	-0.0001	0.178
$x4_{pc}$:Eastern Canada	-0.381	-0.00006	0.255
$x5_{pc}$:Mid-Atlantic	-0.336	-0.0001	0.18
$x6_{pc}$:Western Canada	-0.372	-0.00019	0.303

Table 1.16: ADF test result before COVID

ADF test alternative hypothesis: stationary			
data	p value	data	p value
$PAAS_{pc}$:North Central	0.035	$x1_{pc}$	<0.01
$PACS_{pc}$:Southwest	0.024	$x2_{pc}$	<0.01
$PADS_{pc}$:Southeast	0.035	$x3_{pc}$	<0.01
$PAKU_{pc}$:Eastern Canada	0.048	$x4_{pc}$	<0.01
$PALB_{pc}$:Mid-Atlantic	0.029	$x5_{pc}$	<0.01
$PALI_{pc}$:Western Canada	0.037	$x6_{pc}$	<0.01

other. We use t Copula, Vine copulas and FDG.frechet copula to describe the dependence of markets, and FDG.frechet copula is better at modeling high tail dependence, which is a good feature for other markets to take precautionary actions in time if one abnormal increase or decrease in prices happens in one market, that is, when the price changes in an extreme case, FDG.frechet copula could capture the tail dependence more accurately.

1.3.4 Pre-COVID Results

To investigate the impact of COVID, we redo the model using the data excluding the prices after 2020. We use the footnote pc to represent the Pre-COVID variables.

Figure 1.14 illustrates the pre-COVID data, and Table 1.15 shows the statistics of $x1_{pc}$,

$x2_{pc}$, $x3_{pc}$, $x4_{pc}$, $x5_{pc}$, $x6_{pc}$. Comparing Table 1.15 with Table 1.1, the upper limit of price changes in Southeast US market and Mid-Atlantic US market are lower when we look at the pre-COVID data. Then the results of ADF test in Table 1.16 ensure that $x1_{pc}$, $x2_{pc}$, $x3_{pc}$, $x4_{pc}$, $x5_{pc}$, $x6_{pc}$ pass the stationary test.

Figure 1.16, Figure 1.17, Table 1.17 and Table 1.18 present the results that we use logistic distributions as marginals for the pre-covid data and the corresponding estimates of the parameters.

Table 1.19 and Table 1.20 present the empirical upper and lower tail dependence of $x1_{pc}$, $x2_{pc}$, $x3_{pc}$, $x4_{pc}$, $x5_{pc}$ and $x6_{pc}$. From the results, North Central US market and Eastern Canada market, and Southeast US market and Mid-Atlantic US market have higher upper and lower tail dependence, which means when the price increases a lot in one market, the price of other market has a high probability to increase as well; and when the price decreases a lot in one market, the price of the other market presumably would decrease correspondingly. This result is consistent with that obtained from the model using the full dataset.

Table 1.21 presents AIC and BIC criteria for the optimal bivariate copula specification and estimates results. All 15 pairs favor the t copula specification, and that is also consistent with the result we obtain from the model using the full dataset.

From Table 1.22, we can conclude that Vine copulas are better than t copula, which is different from what we have in previous sections. However, FDG.frechet copula is the one that could overestimate the tail dependence while others underestimate it. Therefore, we keep FDG.frechet and Vine copulas as our final model.

Then we use the data simulated from Vine and FDG.frechet copulas to plot the graph investigating the relationship between $\beta_{pc} = \frac{\partial x_{ipc}}{\partial x_{jpc}}$ and x_{jpc}, $i > j, i, j = 1, 2, \ldots, 6$. We again use LOESS to plot the relationship between β_{pc} and x_{jpc}'s.

Table 1.17: Goodness of Fit for Marginals before COVID

	Normal			Logistic			Optimal
	AIC	BIC	lglk	AIC	BIC	lglk	
$x1_{pc}$:North Central	-3708.786	-3698.631	1856.393	-3902.843	-3892.688	1953.422	Logistic
$x2_{pc}$:Southwest	-3669.083	-3658.928	1836.541	-3840.375	-3830.22	1922.188	Logistic
$x3_{pc}$:Southeast	-3598.109	-3587.954	1801.055	-3755.562	-3745.407	1879.781	Logistic
$x4_{pc}$:Eastern Canada	-3682.462	-3672.307	1843.231	-3890.13	-3879.975	1947.065	Logistic
$x5_{pc}$:Mid-Atlantic	-3616.381	-3606.226	1810.19	-3761.006	-3750.851	1882.503	Logistic
$x6_{pc}$:Western Canada	-3420.439	-3410.284	1712.22	-3581.36	-3571.205	1792.68	Logistic

Table 1.18: Parameter Estimates for Marginals before COVID

	location		scale	
	estimate	*std error*	*estimate*	*std error*
$x1_{pc}$:North Central	0.0011	0.0012	0.0252	0.0006
$x2_{pc}$:Southwest	0.0015	0.0013	0.0259	0.0006
$x3_{pc}$:Southeast	0.0021	0.0013	0.0269	0.0007
$x4_{pc}$:Eastern Canada	0.0006	0.0012	0.0252	0.0006
$x5_{pc}$:Mid-Atlantic	0.0018	0.0013	0.0270	0.0007
$x6_{pc}$:Western Canada	0.0003	0.0014	0.0291	0.0007

Table 1.19: λ^U: Upper Tail Dependence of Six Markets before COVID

	$x1_{pc}$	$x2_{pc}$	$x3_{pc}$	$x4_{pc}$	$x5_{pc}$	$x6_{pc}$
$x1_{pc}$:North Central	1.00	0.71	0.62	0.82	0.65	0.53
$x2_{pc}$:Southwest	0.71	1.00	0.79	0.65	0.76	0.53
$x3_{pc}$:Southeast	0.62	0.79	1.00	0.59	0.82	0.53
$x4_{pc}$:Eastern Canada	0.82	0.65	0.59	1.00	0.65	0.50
$x5_{pc}$:Mid-Atlantic	0.65	0.76	0.82	0.65	1.00	0.50
$x6_{pc}$:Western Canada	0.53	0.53	0.53	0.50	0.50	1.00

Table 1.20: λ^L: Lower Tail Dependence of Six Markets before COVID

	$x1_{pc}$	$x2_{pc}$	$x3_{pc}$	$x4_{pc}$	$x5_{pc}$	$x6_{pc}$
$x1_{pc}$:North Central	1.00	0.59	0.59	0.79	0.62	0.65
$x2_{pc}$:Southwest	0.59	1.00	0.76	0.53	0.74	0.53
$x3_{pc}$:Southeast	0.59	0.76	1.00	0.62	0.76	0.53
$x4_{pc}$:Eastern Canada	0.79	0.53	0.62	1.00	0.62	0.65
$x5_{pc}$:Mid-Atlantic	0.62	0.74	0.76	0.62	1.00	0.53
$x6_{pc}$:Western Canada	0.65	0.53	0.53	0.65	0.53	1.00

Table 1.21: Bivariate Copula Selection before COVID

markets	optimal copula	$\hat{\theta}_1$	$SE(\hat{\theta}_1)$	$\hat{\theta}_2$	$SE(\hat{\theta}_2)$	τ	λ_L	λ_U
($x1_{pc}$:North Central,$x2_{pc}$:Southwest)	t	0.85	0.01	3.39	0.45	0.64	0.58	0.58
($x1_{pc}$:North Central,$x3_{pc}$:Southeast)	t	0.84	0.01	3.94	0.59	0.64	0.55	0.55
($x1_{pc}$:North Central,$x4_{pc}$:Eastern Canada)	t	0.9	0.01	2.2	0.24	0.71	0.71	0.71
($x1_{pc}$:North Central,$x5_{pc}$:Mid-Atlantic)	t	0.86	0.01	4.83	0.79	0.65	0.53	0.53
($x1_{pc}$:North Central,$x6_{pc}$:Western Canada)	t	0.85	0.01	2.52	0.31	0.64	0.62	0.62
($x2_{pc}$:Southwest,$x3_{pc}$:Southeast)	t	0.92	0.01	2.17	0.25	0.74	0.73	0.73
($x2_{pc}$:Southwest,$x4_{pc}$:Eastern Canada)	t	0.83	0.01	3.29	0.44	0.62	0.56	0.56
($x2_{pc}$:Southwest,$x5_{pc}$:Mid-Atlantic)	t	0.91	0.01	2.76	0.35	0.72	0.69	0.69
($x2_{pc}$:Southwest,$x6_{pc}$:Western Canada)	t	0.8	0.01	5.64	1	0.59	0.43	0.43
($x3_{pc}$:Southeast,$x4_{pc}$:Eastern Canada)	t	0.84	0.01	3.9	0.58	0.63	0.54	0.54
($x3_{pc}$:Southeast,$x5_{pc}$:Mid-Atlantic)	t	0.95	0.00	2	0.31	0.81	0.81	0.81
($x3_{pc}$:Southeast,$x6_{pc}$:Western Canada)	t	0.78	0.01	5.94	1.11	0.57	0.39	0.39
($x4_{pc}$:Eastern Canada,$x5_{pc}$:Mid-Atlantic)	t	0.85	0.01	3.72	0.53	0.65	0.57	0.57
($x4_{pc}$:Eastern Canada,$x6_{pc}$:Western Canada)	t	0.83	0.01	4.02	0.62	0.62	0.52	0.52
($x5_{pc}$:Mid-Atlantic,$x6_{pc}$:Western Canada)	t	0.79	0.01	5.52	0.97	0.58	0.41	0.41

Table 1.22: Multivariate Copulas Comparison before COVID

Copulas	MAE_{ρ}	$MAE_{\lambda L}$	$ME_{\lambda L}$	$MAE_{\lambda U}$	$ME_{\lambda U}$
fdg.frechet	0.0302	0.1569	0.2081	0.1848	0.1983
fdg.cuadrasauge	0.0302	0.6333	-0.6333	0.2337	0.2337
fdg.sinus	0.4735	0.6333	-0.6333	0.3569	0.3569
Gaussian	0.0395	0.6333	-0.6333	0.6431	-0.6431
t	0.0299	0.0733	-0.0276	0.1069	-0.0374
CVine	0.017	0.0664	-0.0352	0.0781	-0.0064
RVine	0.018	0.0566	-0.0714	0.0861	-0.014

In summary, although we use Vine and FDG.frechet copulas to model the pre-COVID situation, which is different from the case using full dataset, our results from Figure 1.18 to Figure 1.20 still support the law of one price for six oriented strand board market in North America and most pairs of markets appear to be highly linked to each other.

1.4 Conclusion

To conclude, we apply FDG copulas to six markets for oriented strand board, which is a versatile structural wood panel widely traded in North America. We firstly use bivariate copulas to model the dependence of 15 pairs of markets, and t copulas are selected using Bayesian information criterion. Although bivariate t copula could provide good estimation for correlation coefficent ρ, it underestimates tail dependence in most cases. Therefore, for pairs of markets that have high tail dependence such as Southeast US market and Mid-Atlantic US market, a bivariate t copula is not a good choice. Then we use multivariate copulas to model the dependence of six markets more precisely. After comparing MAE_{ρ}, MAE_{λ} and ME_{λ}, we choose multivariate t copula, Vine copulas and FDG.frechet copula. While a t copula still underestimates the tail dependence in most cases, and Vine copulas

also provide underestimated tail dependence in some cases, FDG.frechet copula is better at describing the case when the tail dependence coefficient has a high value. Hence, we could use t copula and FDG.frechet copula together in practice to give us a range of the tail dependence.

Our results also reveal varying degrees of market integration when there are different percentage changes in prices. The results also provide support for law of one price in North American OSB markets. In most cases, the six markets appear to be strongly integrated, especially when the price changes are not large.

And we also do the estimation using the pre-COVID dataset. The copulas we choose for that data are Vine and FDG.frechet copulas. Although the final model is different, we still have the conclusion that our results support the law of one price in North American OSB markets.

We can learn from the results that the coefficients of tail dependence obtained from bivariate copula and multivariate copula are different, which is also the reason why we need a multivariate copula to investigate the dependence of markets. As for multivariate copula models, Vine copulas are flexible and can provide good estimates of association between markets while we need to specify the decomposition of density; FDG copula is convenient and flexible to use because we do not need to assume the dependence structure in advance and make a lot of assumptions. It works well when the coefficients of tail dependence are large. However, when we need to model a weak tail dependence, t copula could be a good choice.

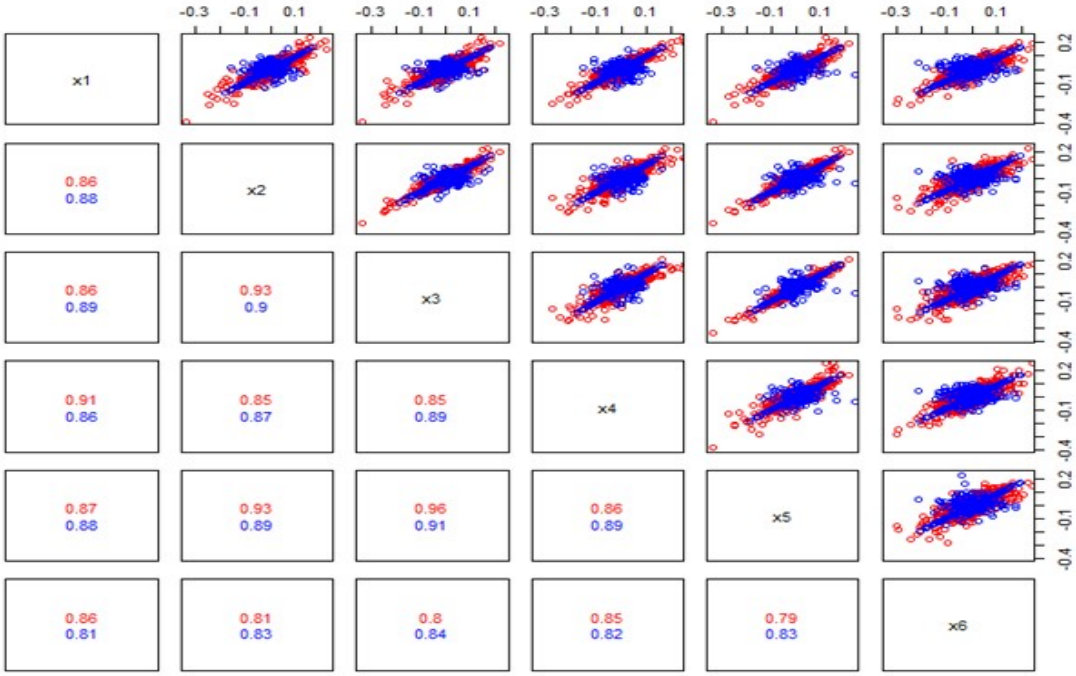


Figure 1.6: FDG.Frechet Copula

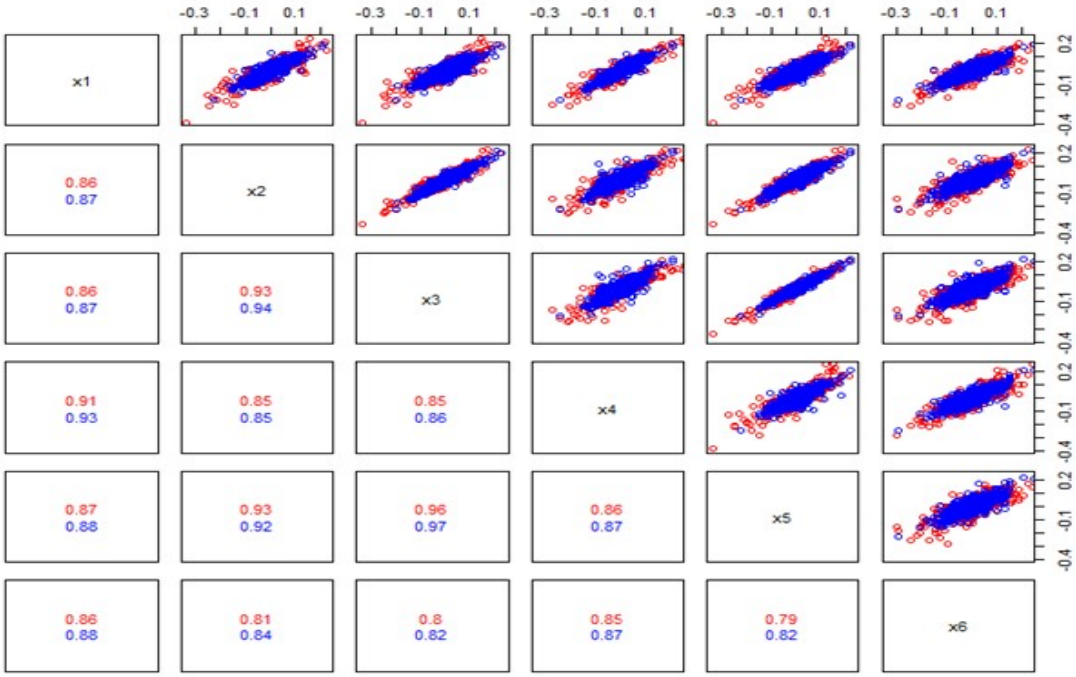


Figure 1.7: t Copula

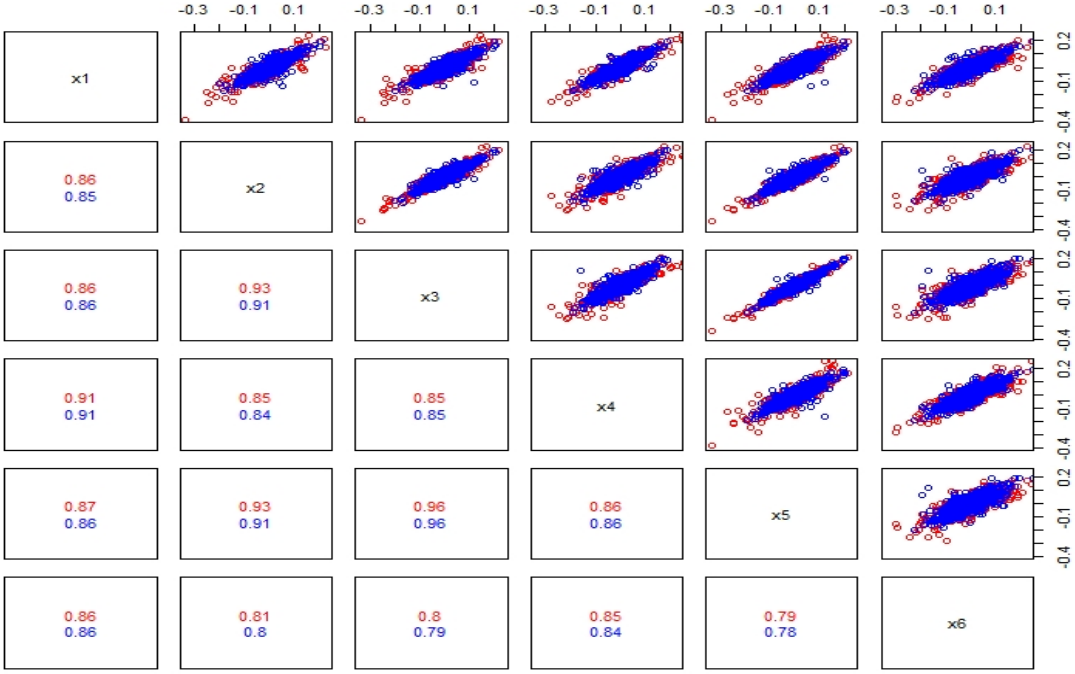


Figure 1.8: C-Vine Copula

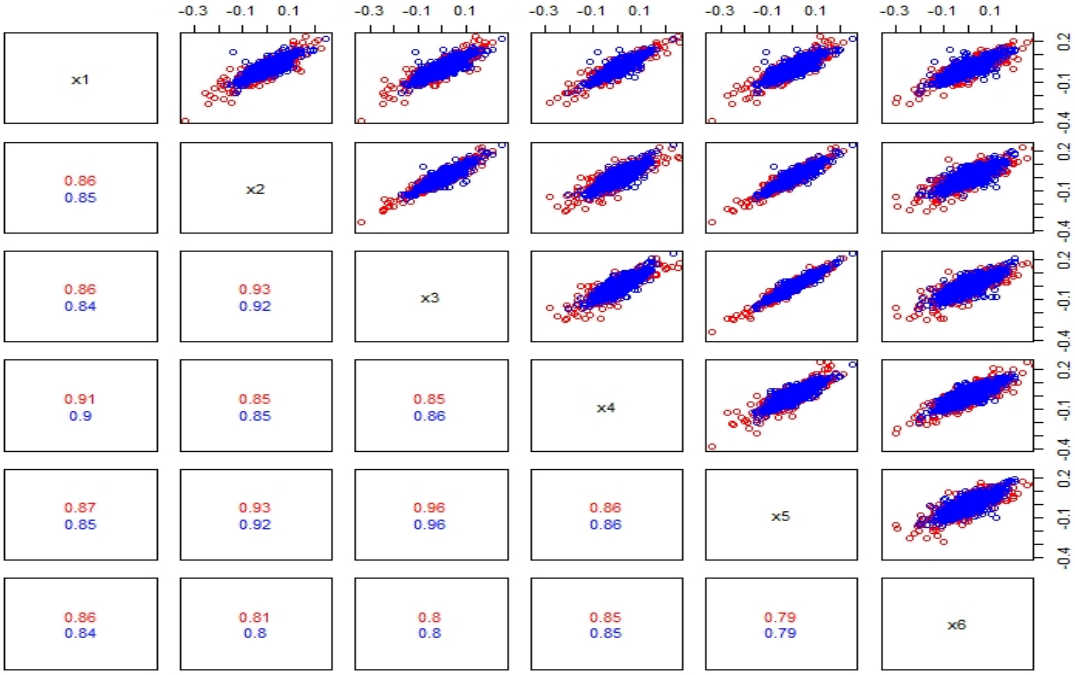


Figure 1.9: D-Vine Copula

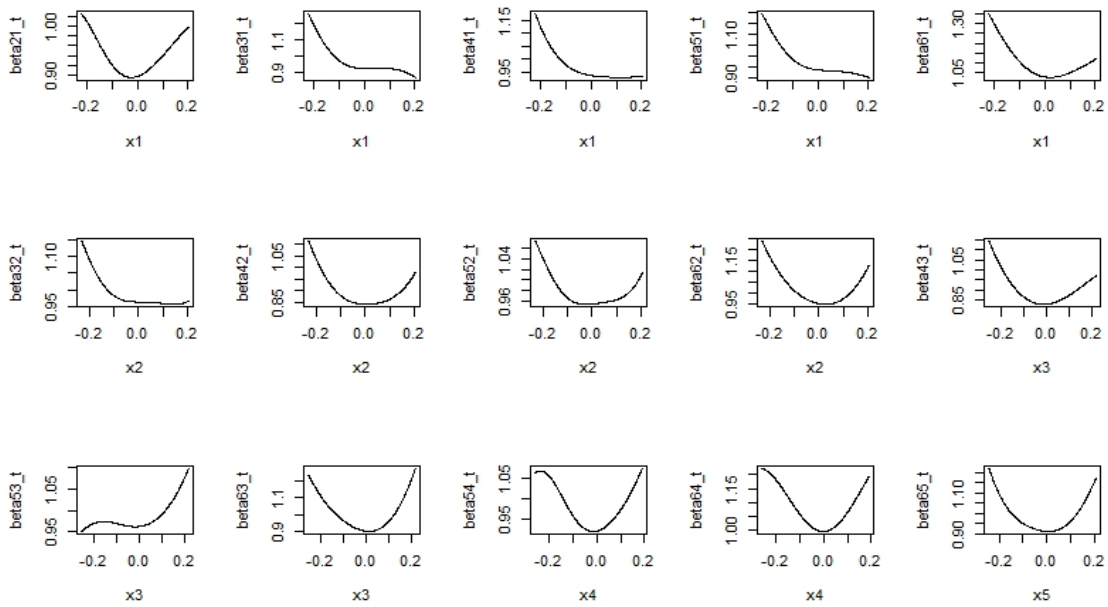


Figure 1.10: β Estimated from t Copula

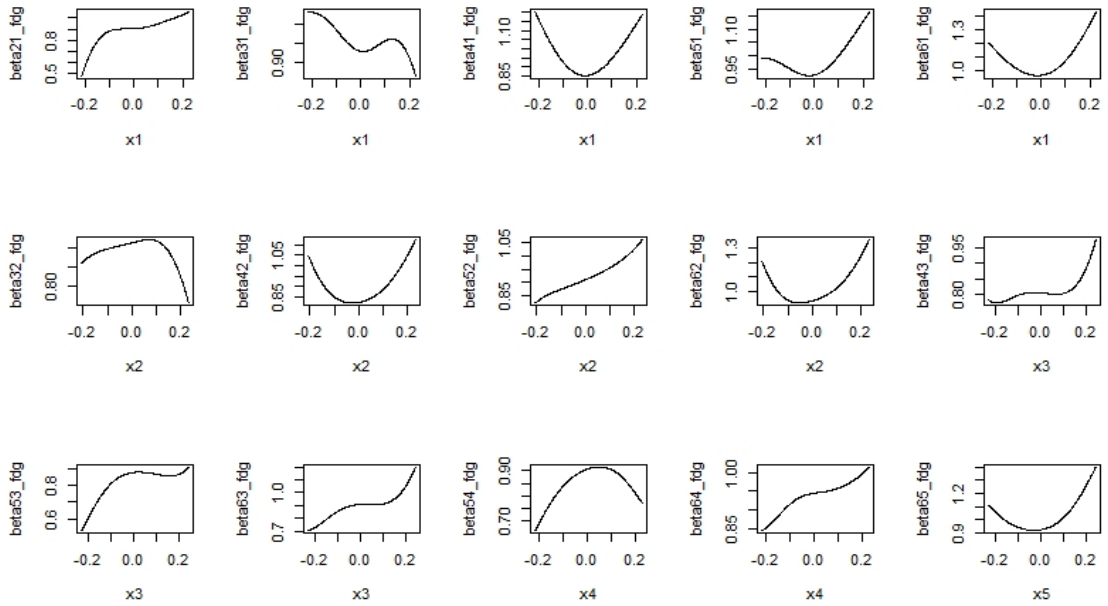


Figure 1.11: β Estimated from FDG.frechet Copula

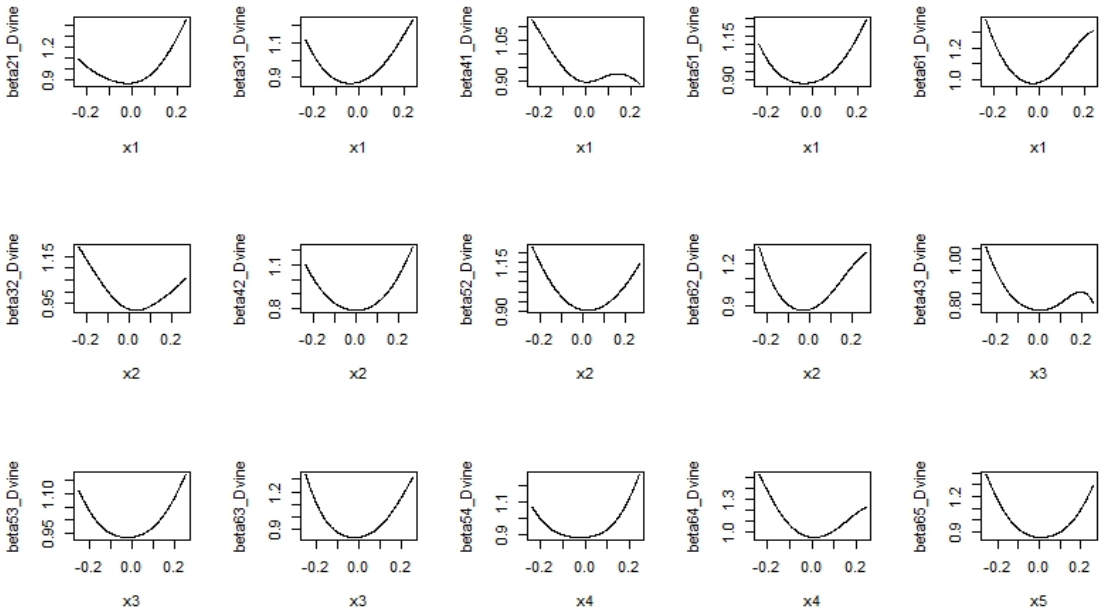


Figure 1.12: β Estimated from D-Vine Copula

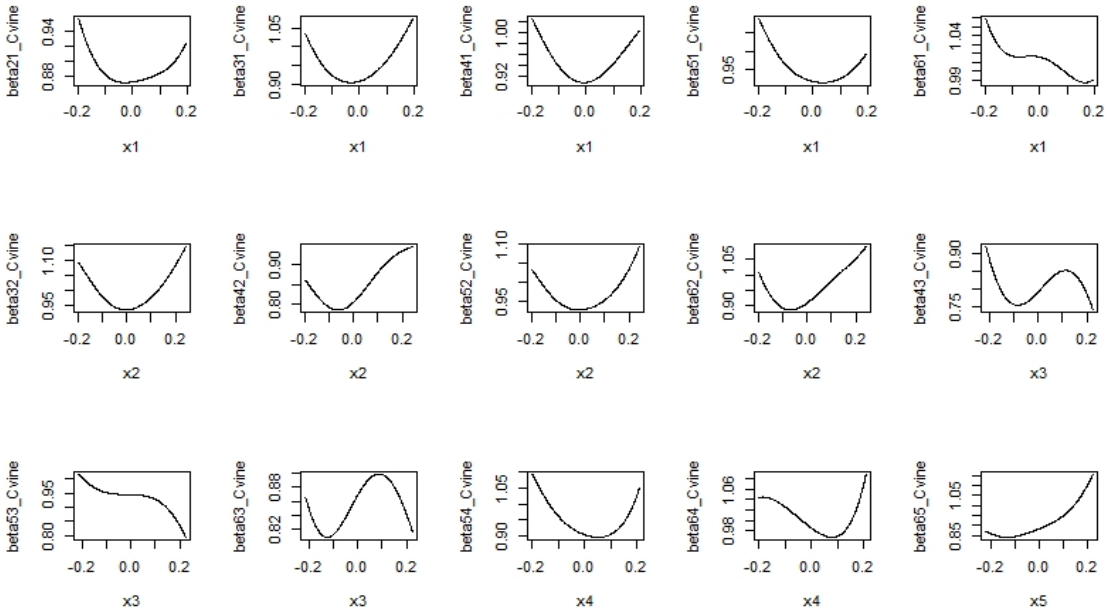


Figure 1.13: β Estimated from C-Vine Copula

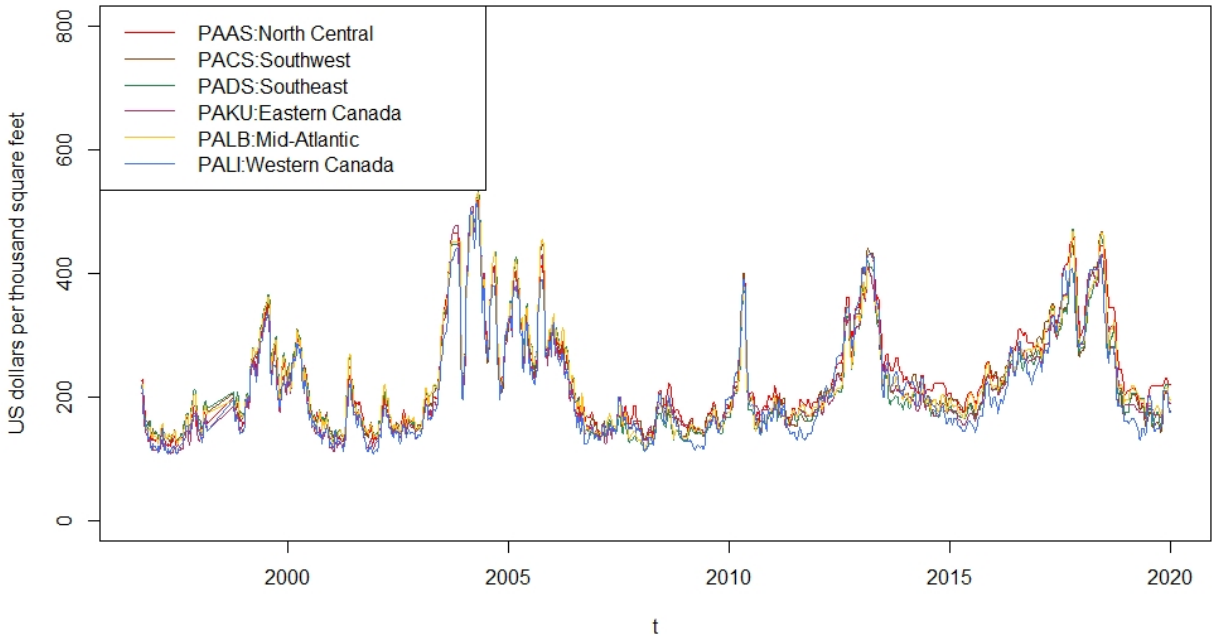


Figure 1.14: Prices before COVID

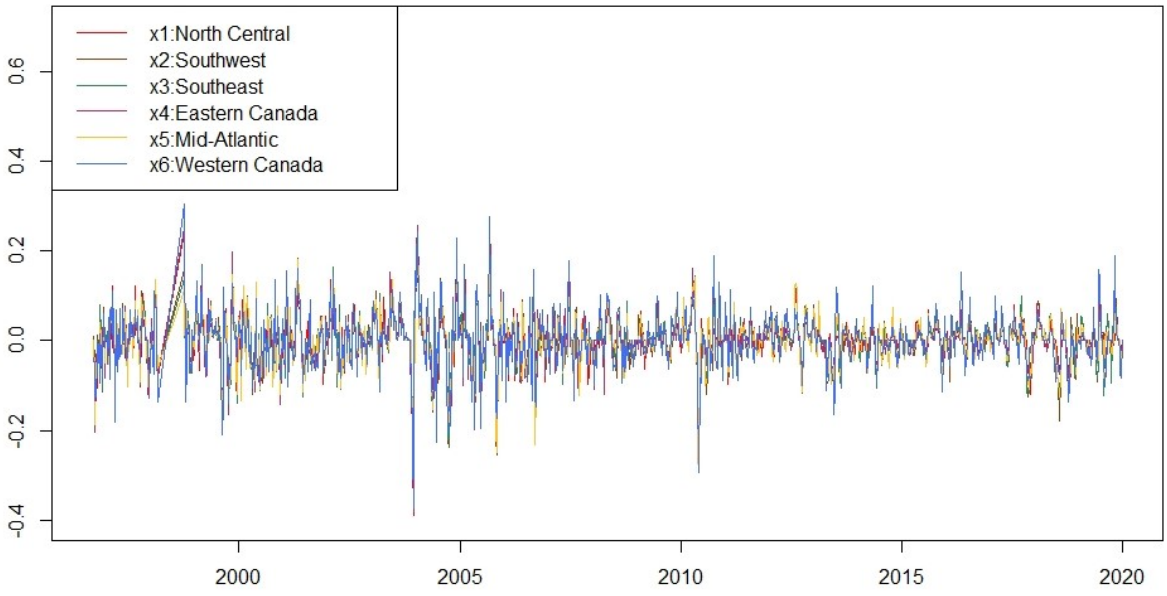


Figure 1.15: First Difference of Log Prices before COVID

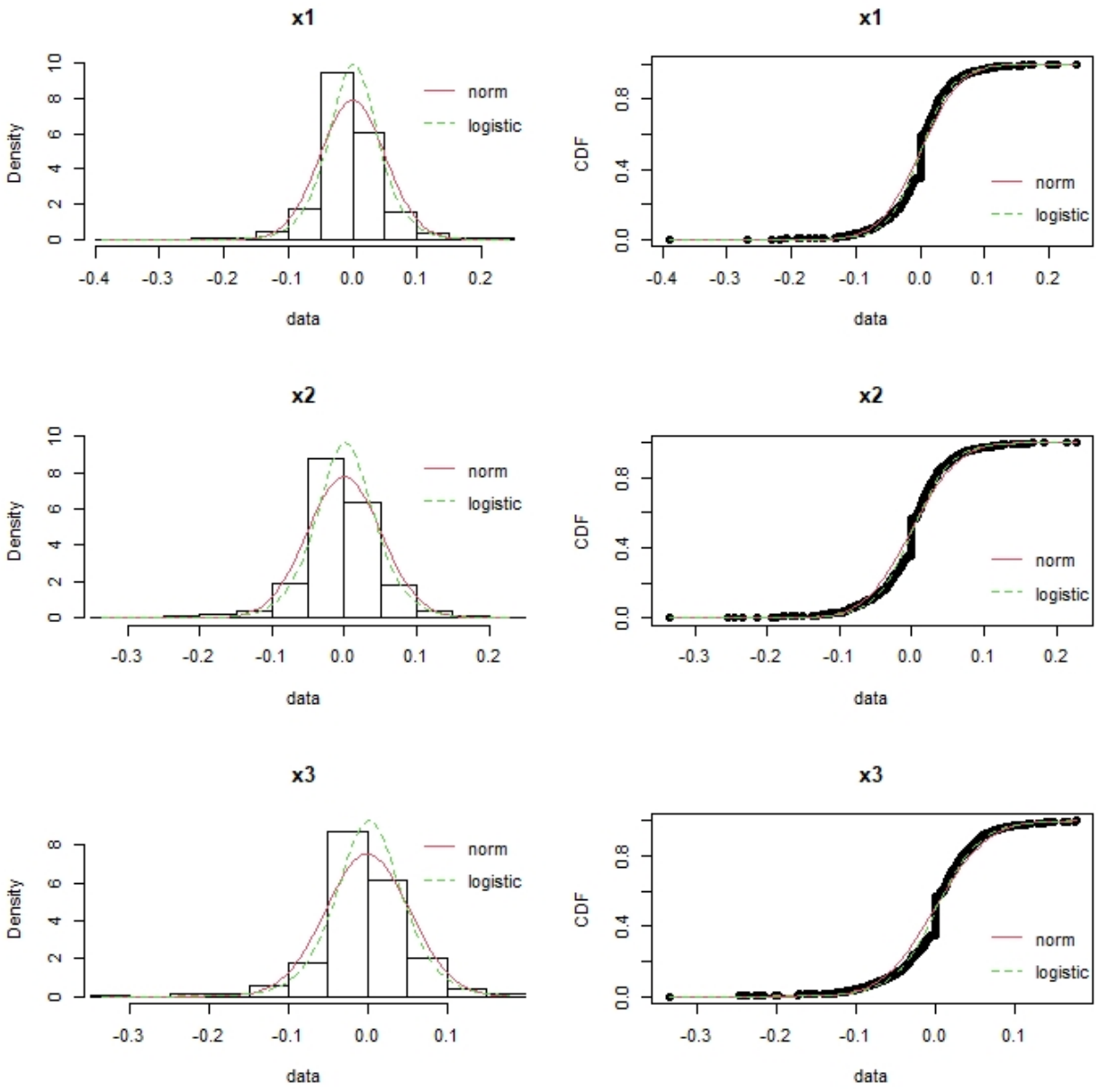


Figure 1.16: Marginals before COVID: $x1_{pc}$ $x2_{pc}$ $x3_{pc}$

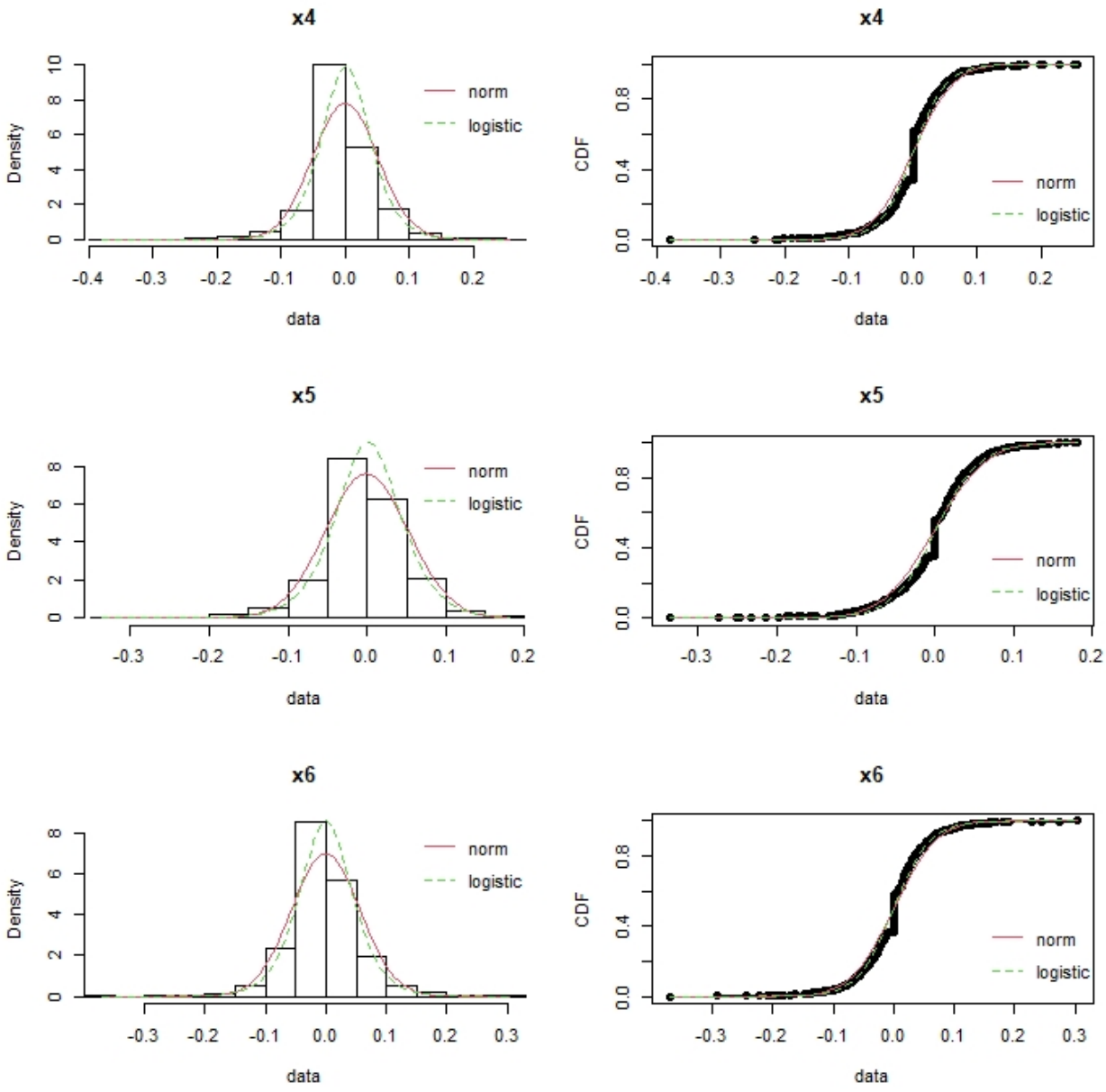


Figure 1.17: Marginals before COVID: $x4_{pc}$ $x5_{pc}$ $x6_{pc}$

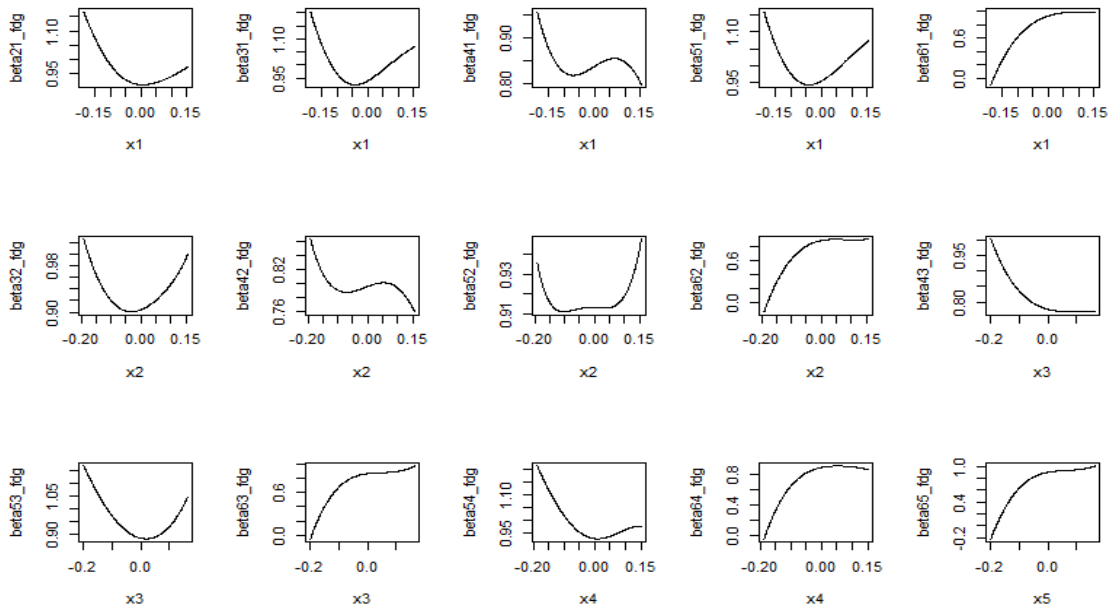


Figure 1.18: β Estimated from FDG.frechet Copula

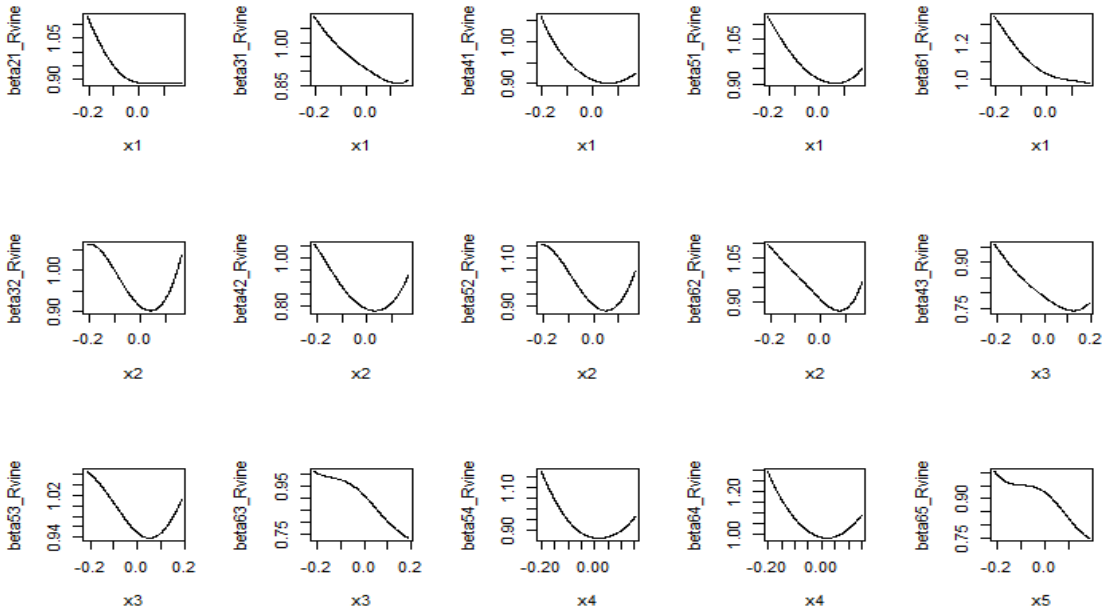


Figure 1.19: β Estimated from D-Vine Copula

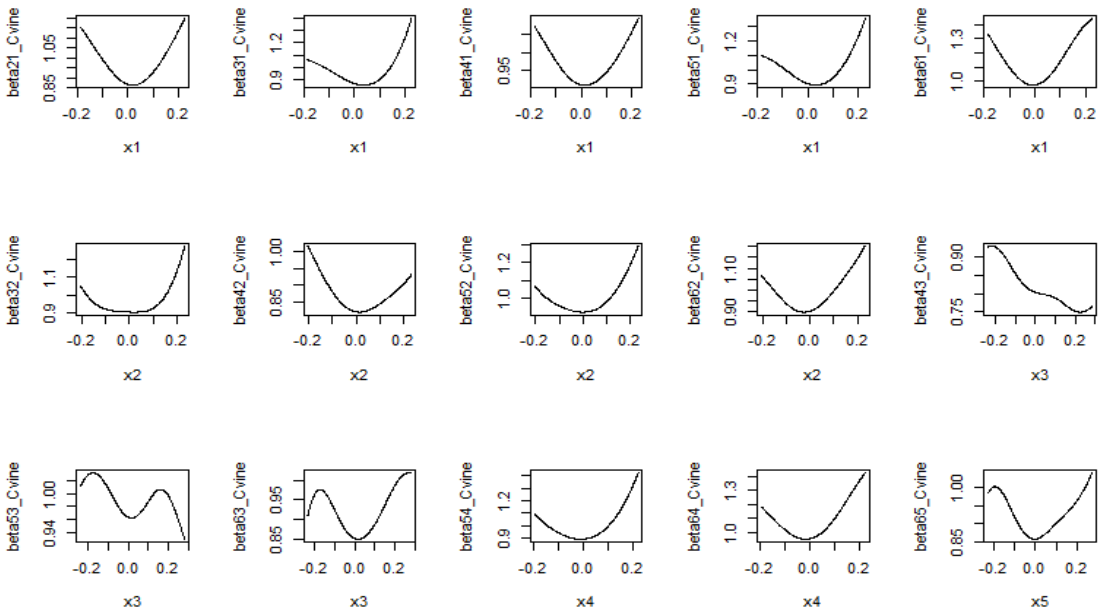


Figure 1.20: β Estimated from C-Vine Copula

CHAPTER

2

MODELING SPATIAL PRICE LINKAGES USING A THRESHOLD SDE MODEL

2.1 Introduction

Economists have devoted considerable attention to the notion of market integration and in the case of spatially separated markets, spatial arbitrage is the key mechanism that links markets. These linkages make price differentials be no more than the costs associated with arbitrage. Although arbitrage costs are important, it may be difficult to define and even harder to empirically measure them (Goodwin, Holt, and Prestemon (2021)). The

integration of spatial markets can have important implications for price discovery and the operation of the market since persistent deviations from integration may imply riskless profits through arbitrage and trade (Goodwin and Piggot (2001)). The fundamental condition of the integration is often called the "Law of One Price" (LOP), which is an essential ingredient in modern and classical models in economics. The LOP states that, abstracting from transactions costs, regional markets that are linked by trade and arbitrage will have a common, unique price (Goodwin and Fackler (2001)).

To test the LOP, early studies focused on correlations or linear time series analysis involving prices (Goodwin and Fackler (2001)). Recent research has turned to the use of models that can capture the regime-switching feature of price linkages. Those models have been used to represent such regime changes and the influences of unobserved transactions costs and other factors that may influence spatial price adjustments. The underlying motivation is that adjustments to equilibrium may not be linear, and this nonlinearity may be associated with the transactions costs that are associated with arbitrage (Goodwin, Holt, Prestemon, and Onel (2018)). Threshold models are one example of such regime-switching models, where thresholds that define alternative trading regimes are used to evaluate spatial price parity in a nonlinear fashion. To model the behavior of the price differences of two locations, Goodwin and Piggot (2001) used a threshold error correction model and firstly explicitly accounted for the "neutral band" resulting from transactions costs. Within the neutral band, prices are not linked to one another, while price equalizing arbitrage activities are triggered only when shocks result in price differences which exceed the neutral band. Price differences that fall outside of the limits of the band should be mean reverting, which is driven by arbitrage behavior. And price differences within the limits of the band may follow something very close to a unit root process without drift (Goodwin, Holt, Prestemon, and Onel (2018)).

Most of the models used in the literature are discrete time models. The most recent research in this area has applied smooth or discrete threshold time series models that typically consider refinements of autoregressive or vector error correction models in analyzing price relationships. However, as computational power increases, it becomes feasible to use more accurate differential equation models and to make more precise predictions. This paper considers a continuous time threshold model using stochastic differential equations (SDE). In particular, we use two different processes setting up the threshold model, Brownian motion and Ornstein-Uhlenbeck process. Brownian motion is the extension of a (discrete-time) random walk to a continuous-time process. The Ornstein–Uhlenbeck process can be considered as the continuous-time analogue of the discrete-time AR(1) process. This process tends to drift towards its mean function over time and has a tendency to move back towards a central location, with a greater attraction when the process is further away from the center. Thus this process is called a mean-reverting process. To be consistent with the description using the "neutral band", here we use the Ornstein-Uhlenbeck model (arithmetic Brownian motion) to characterise mean reversion when the price differences fall outside the limits of the band; and we use a Brownian motion model to characterise the behavior of price differences that fall within the limits of the band.

Specifically, we apply this model to examine price linkages among spatially distinct US markets for plywood. This product plays an essential role in the real property construction industries and is subject to significant influences from economic shocks. The recessions of 2007–2010 had notable impacts on the production of plywood. A prominent feature of such recessions was a significant decrease in the demand for construction materials. For example, housing starts dropped from 2.2 million in February of 2008 to 478 thousand in April of 2009 (Goodwin et al. (2021)). Plywood is produced widely and is typically traded across North American regions as well as in international markets. It has several properties: high

strength, flexibility, moisture resistance, impact resistance, fire resistance and insulation, which make it an important construction materials in the market. When sheathing and flooring panels are routinely covered with rain, snow, and ice during construction delays, plywood is preferred. Also, the National Tile Contractors Association and the Resilient Floor Covering Institute both recommend plywood for subflooring and underlayment, because it doesn't have the risk of swollen edges that other construction materials do. To sum up, plywood is a widely used, versatile sheet wood that can be used to construct fine furniture, cabinets, and boats and has many applications in the building and construction industry. Typically, users look for strength, rigidity, and dimensional stability with plywood products.

The data used here are plywood prices and they are observed in Western and three Southern regions. Western refers to the western producing region that includes Oregon, Washington, Idaho, Montana, and California. Southern-westside refers to a southern pine plywood producing zone that includes Texas, Louisiana, and Arkansas. Southern-central refers to a southern pine plywood producing zone that includes Mississippi and Alabama. Southern-eastside refers to a southern pine plywood producing zone that includes Georgia, Florida, and South Carolina, plus North Carolina and Virginia, where prices are usually higher.

The paper is organized as follows. In section 2.2, we present the framework of the paper. That is, we provide a brief introduction of Ornstein-Uhlenbeck model and Brownian Motion. In section 2.3, we apply the two types of stochastic differential equations to model the dependence of spatial price in a single West Coast market and three spatially-distinct Southern US markets using the data from 1995 to 2017, and we have the conclusion in section 2.4.

2.2 Empirical Framework

According to Goodwin et al. (2018), a homogeneous commodity traded in a common currency in two regional markets is represented, respectively, by location indices i and j. The individual market prices are denoted by P_i and P_j and the corresponding logarithms of prices are p_i and p_j. A simple model of the law of one price that incorporates the effects of transactions costs can be written as:

$$\tau_1 \leq p_i - p_j \leq \tau_2 \tag{2.1}$$

where τ_1 and τ_2 represent the transactions cost for trade between the markets. The implication from equation (2.1) is that within the band $[\tau_1, \tau_2]$, no profitable arbitrage activity will occur; and arbitrage is profitable when log price differentials, $p_i - p_j$, fall outside of the limits of the band.

The dynamic nature of trade and commodity arbitrage often allows for the price differential to exceed arbitrage costs in the shortrun, but such arbitrage opportunities should quickly dissipate as commodity exchange restores the equilibrium arbitrage condition of equation (2.1). The speed of adjustment to such deviations from arbitrage equilibrium is often taken as a measure of the degree of market integration (Goodwin et al. (2021)).

Such no-arbitrage conditions are typically represented using threshold models, where the threshold represents an empirical measure of the transactions cost, τ_1 and τ_2. Over time, we expect that log price differentials within the band may follow something very close to a unit root process and log price differences that fall outside of the limits of the band should be mean reverting. This transactions cost band, which has often been assumed in the literature, typically yields an empirical model consistent with the threshold models.

In this study, we use a combination of Brownian Motion and Ornstein-Uhlenbeck

Model to capture the movement of log price differentials within and outside of the band respectively.

2.2.1 Brownian Motion

The physical phenomenon of Brownian motion was discovered by Robert Brown, a 19th century scientist who observed through a microscope the random swarming motion of pollen grains in water, now understood to be due to molecular bombardment. So Brownian motion takes its name from this Scottish botanist.

The modern mathematical treatment of Brownian motion, also called the Wiener process is due to Wiener (1923). He proved that there exists a version of Brownian motion with continuous paths. Brownian motion is a Gaussian process, a Markov process, and a martingale. Hence it is important in the theory of stochastic process. It serves as a basic building block of stochastic calculus and therefore, the key to simulating stochastic processes. Although it is not easy to observe pure Brownian motions in real-world data, we can combine them and rescale them to build more complex processes that successfully approximate the data. In simple terms, Brownian motion is a continuous process such that its increments for any time scale are drawn from a normal distribution.

Brownian Motion $W = \{W(t), t \geq 0\}$ is a Gaussian process with continuous paths and with independent increments such that $W(0) = 0$ with probability 1, $EW(t) = 0$, and $Var(W(t) - W(s)) = t - s$ for all $0 \leq s \leq t$. In addition, $W(t) - W(s) \sim N(0, t-s)$, for $0 \leq s \leq t$ and on any two disjoint intervals, say $(t_1, t_2), (t_3, t_4)$ with $t_1 \leq t_2 \leq t_3 \leq t_4$, the increments $W(t_2) - W(t_1)$ and $W(t_4) - W(t_3)$ are independent.

2.2.2 Ornstein-Uhlenbeck Model

In mathematics, the Ornstein–Uhlenbeck process is a stochastic process with applications in financial mathematics and the physical sciences. It is named after Leonard Ornstein and George Eugene Uhlenbeck. In their famous paper, Ornstein and Uhlenbeck (1930) studied a free particle in Brownian motion, and it was a model for the velocity of a massive Brownian particle under the influence of friction. In order to understand the displacement process of the particle, they investigated the velocity process, which is now known as the (one-dimensional) Ornstein-Uhlenbeck process.

The Ornstein-Uhlenbeck process is the unique solution to the following stochastic differential equation

$$dX_t = (\theta_1 - \theta_2 X_t)dt + \theta_3 dW_t, \quad X_0 = x_0 \tag{2.2}$$

with $\theta_3 \in \mathbb{R}_+$ and $\theta_1, \theta_2 \in \mathbb{R}$. The model ($\theta_1 = 0$) was originally proposed by Ornstein and Uhlenbeck in the context of physics and then generalized by Vasicek to model interest rates (Iacus (2009)). $\frac{\theta_1}{\theta_2}$ is the mean-reversion level, θ_2 is the mean-reversion speed, θ_3 is the volatility and W_t is the Brownian motion. For $\theta_2 > 0$, this is a mean reverting process, which means that the process tends to oscillate around some equilibrium state. Another property of the Ornstein-Uhlenbeck process is that, contrary to the Brownian Motion, it is a process with finite variance for all $t \geq 0$.

2.2.3 The Threshold SDE Model

To test the validity of Law of One Price, we set up the model as follows:

$$d\,y_t = \begin{cases} (a_1 - a_2 y_t)d\,t + a_3 d\,W_t & y_t < c_1 \\ \theta_1 d\,W_t & c_1 \le y_t \le c_2 \\ (b_1 - b_2 y_t)d\,t + b_3 d\,W_t & y_t > c_2 \end{cases} \tag{2.3}$$

where y_t is the contemporaneous log price differentials of two regionally-distinct markets, that is, $y_t = p_t^1 - p_t^2$. W_t is a standard Brownian motion process, and c_1, c_2 are the lower bound and upper bound of the "neutral band", respectively. $\frac{a_1}{a_2}$ is the mean of the lower regime, and a_2 is the mean-reversion speed of the lower regime. Similarly, $\frac{b_1}{b_2}$ is the mean of the upper regime, and b_2 is the mean-reversion speed of the upper regime. If $a_2 > 0$ and $b_2 > 0$, then the lower and upper regimes are depicted by two mean reverting processes, and that means the LOP is valid.

To be more specific, when $y_t < c_1$, $d\,y_t = d\,y_t^{low}$. This y_t^{low} is described by a low-OU process. When $y_t > c_2$, $d\,y_t = d\,y_t^{up}$. This y_t^{up} is described by an up-OU process. When $c_1 \le y_t \le c_2$, $d\,y_t$ is decribed by a Brownian Motion process. The reason is that OU process is mean-reverting and we cannot guarantee that all possible realizations of the up-OU would be greater than c_2 or all possible realizations of the low-OU would fall below c_1 in our model. Therefore we can regard the middle part as the composition of low-OU and up-OU. Then y_t^{mid} turns out to be a Brownian Motion. That is, $d\,y_t = d\,y_t^{mid}$ when $c_1 \le y_t \le c_2$.

We expect that log price differentials within the limits of the band may follow something very close to a Brownian Motion. However, log price differences that fall outside of the limits of the band should be mean reverting, that is, should follow an Ornstein-Uhlenbeck process.

2.3 Empirical Application

2.3.1 Data

For the empirical analysis, we consider regional North American markets for a prominent traded commodity—plywood. Plywood as a building material is very widely used due to its many useful properties. It is an economical, factory-produced sheet of wood with precise dimensions that does not warp or crack with changes in atmospheric moisture. The price data were obtained from *Random Lengths*, an industry publication that reports weekly prices for most major softwood lumber and structural wood-based panel products used in the construction sector. The softwood plywood we consider in this paper was of a standard quality with 4-plies at $5/8^{ths}$ inches and was quoted on a fob basis at the mills in US dollars per thousand square feet.

The 1995-2017 data set, which contains 1179 weekly observations, consists of prices of plywood in a single West Coast market and three spatially-distinct Southern US markets. The four markets yield 6 pairs of bivariate price comparisons. X_t^1, X_t^2, X_t^3 and X_t^4 stand for the time series of prices in this four markets, where X_t^1 is the western price, X_t^2 is the southern-westside price, X_t^3 is the southern-central price and X_t^4 is the southern-eastside price. And we use the logarithmic price differentials in our model. That is, $p_t^i = log(X_t^i)$, then $y_t^s = p_t^i - p_t^j$, where $i, j = 1, \ldots, 4, i \neq j$ and $s = 1, \ldots, 6$.

Figure 2.1 shows the data of the original prices of four markets, where PAAE is the western market, PAGS is the southwestern market, PAHM is the southcentral market and PAIG is the southeastern market. Table 2.1 presents the statistics of four market and Table 2.2 presents the statistics of six pairs of logarithmic price differentials. From Table 2.1, we could see the western plywood price is higher than others and more volatile. And in Table 2.2, we could see the difference of logarithmic prices between western market and other

southern markets and the mean, and standard deviation of the differences.

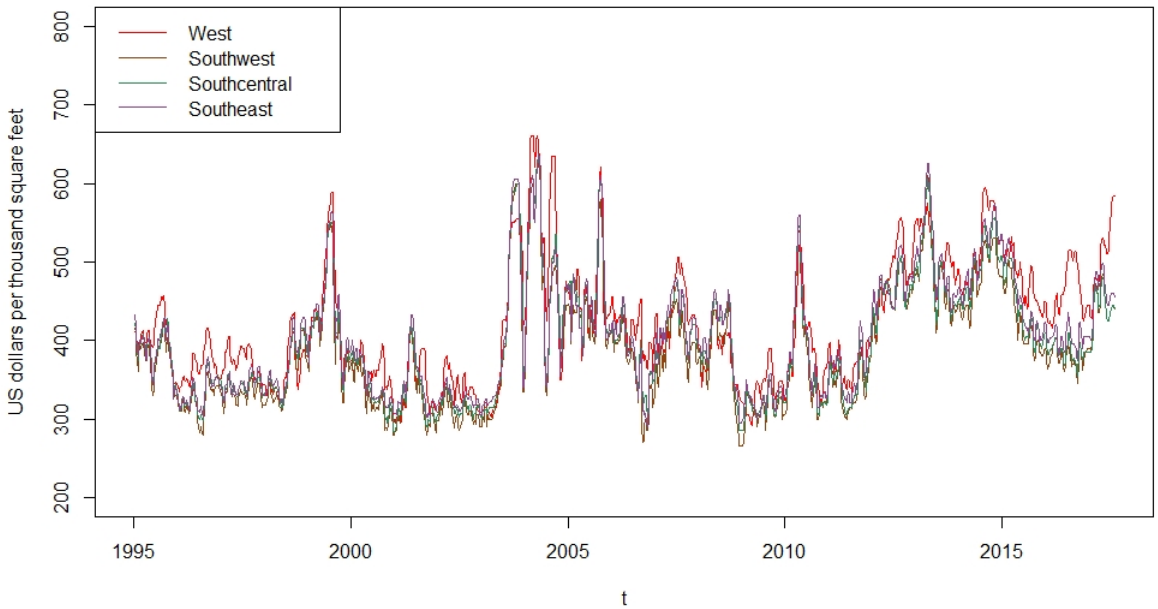


Figure 2.1: Plywood Prices

2.3.2 Thresholds Selection

We use the threshold model (2.3) to describe the log price differentials y_t^s and to see whether it can imply the empirical validity of the Law of One Price.

To estimate the two thresholds c_1 and c_2, we need to group the data. The data are categorized into three groups: the up-OU group ($y_t^s > c_2$), the mid-BM group ($c_1 < y_t^s < c_2$), and the low-OU group ($y_t^s < c_1$). Roughly speaking, we use the up-OU and mid-BM group to fit the up-OU model, use the mid-BM group to fit the mid-BM model and use the mid-BM and low-OU group to fit the low-OU model. Then we need to use the data in an appropriate

Table 2.1: Statistics of Four Markets: Prices

	min	mean	max	sd
West	290	420.3	660	76.57
Southwest	265	387.3	625	72.55
Southcentral	283	398	635	72.96
Southeast	290	405.2	637	73.84

Table 2.2: Statistics of Six Pairs of Log Price Differentials

	min	mean	max	sd
West-Southeast	-0.164	0.036	0.284	0.076
West-Southwest	-0.108	0.082	0.361	0.078
West-Southcentral	-0.170	0.055	0.333	0.076
Southeast-Southcentral	-0.118	0.018	0.084	0.026
Southwest-Southcentral	-0.188	-0.028	0.047	0.025
Southwest-Southeast	-0.159	-0.046	0.064	0.028

way to estimate our model. The idea is that we treat the mid-BM group as the result of the combination of up-OU and low-OU models. Then we need to reorganize the data for our study as follows:

$$y_t^{s_{up}} = \begin{cases} y_t^s & y_t^s > c_2 \\ \frac{1}{1+\alpha_{up}} y_t^s + \frac{\alpha_{up}}{1+\alpha_{up}} c_1' & c_1' \le y_t \le c_2 \end{cases} \tag{2.4}$$

where $c_1' = 2c_2 - max(y_t^s)$, $\alpha_{up} = k_{up} y_t^s + b_{up}$, $k_{up} = -\frac{1}{c_2 - c_1'}$, $b_{up} = \frac{c_2}{c_2 - c_1'}$.

$$y_t^{s_{low}} = \begin{cases} y_t^s & y_t^s < c_1 \\ \frac{1}{1+\alpha_{low}} y_t^s + \frac{\alpha_{low}}{1+\alpha_{low}} c_2' & c_1 \le y_t \le c_2' \end{cases} \tag{2.5}$$

where $c_2' = 2c_1 - min(y_t^s)$, $\alpha_{low} = k_{low} y_t^s + b_{low}$, $k_{low} = \frac{1}{c_2' - c_1}$, $b_{low} = -\frac{c_1}{c_2' - c_1}$.

$$y_t^{s_{mid}} = y_t^s, \quad c_1 \le y_t \le c_2 \tag{2.6}$$

Then we use the Pseudo-likelihood method, which is introduced in Guidoum and

Boukhetala's paper (2016), to estimate the parameters of the SDE models and obtain the AICs. After comparing the AICs, we choose the c_1 and c_2 that can minimize the summation of three AICs from the estimation of the three parts as the thresholds. In order to choose proper initial values for the parameters in estimation, we utilize the invariant law. For $\theta_2 > 0$ in (2.2), the Ornstein-Uhlenbeck process is also ergodic, and its invariant law is the Gaussian density with mean θ_1/θ_2 and variance $\theta_3^2/2\theta_2$. In our model, it means that $y_t^{s_{low}} \sim N(\frac{a_1}{a_2}, \frac{a_3^2}{2a_2})$ and $y_t^{s_{up}} \sim N(\frac{b_1}{b_2}, \frac{b_3^2}{2b_2})$.

To be more specific, in the estimation of up-OU model, the initial values of b_1, b_2 and b_3 are b_1^0, b_2^0 and b_3^0, where $b_1^0 = b_2^0 \cdot E\, y_t^{s_{up}}$ and $b_3^0 = \sqrt{2b_2^0 \cdot Var(y_t^{s_{up}})}$; in the estimation of low-OU model, the initial values of a_1, a_2 and a_3 are a_1^0, a_2^0 and a_3^0, where $a_1^0 = a_2^0 \cdot E\, y_t^{s_{low}}$ and $a_3^0 = \sqrt{2a_2^0 \cdot Var(y_t^{s_{low}})}$.

After we find the parameters and thresholds of our model, in order to see whether the data fits the model well, we compare the number of log price differentials that fall within the band $[c_1, c_2]$ of our data with the number of simulated log price differentials that fall within the band. We also compare the average *time period* that the log price differentials fall outside of the band of our data with the average *time period* that the simulated log price differentials fall outside of the band. In other words, this *time period* means that when the log price differential falls outside of the band, how long will it take to come back within the band.

2.3.3 Results

From Table 2.3 to Table 2.20, we can see the estimates of the six pairs of markets. In the West-Southeast markets, a_2 and b_2 are greater than 0 and $a_2 > b_2$ as $0.0923 > 0.0616$. And the "neutral band" $[c_1, c_2]$ is [-0.0212, 0.1104]. It means that when the log price differential is

in the lower regime, that is, $y_t < c_1$, the adjustment speed, which can be reflected by a_2, of pushing the log price differential back to the transaction costs band is faster than the speed, which can be reflected by b_2, when the log price differential is in the upper regime, that is, $y_t > c_2$. In other words, when the price in western market is around 11% higher than it in southeastern market, arbitrage activities are triggered and the adjustment speed is relatively lower; while the price in western market is around 2% lower than it in southeastern market, arbitrage activities are triggered and the adjustment speed is relatively higher. In addition, positive a_2 and b_2 implies LOP is valid between the western and southeastern markets. In the West-Southwest markets, the adjustment is relatively slower when the log price differential is in the upper regime as $a_2 = 0.0985$, $b_2 = 0.0652$. Both a_2 and b_2 are greater than 0, which supports the validity of LOP between the western and southwestern markets. And the similar situation is observed in the West-Southcentral markets. $a_2 = 0.0845$, $b_2 = 0.0598$, thus a_2 and b_2 are above 0 and $a_2 > b_2$. Also, the band is [-0.0114, 0.1413]. That is, when the price in the western market is about 14% higher than the price in the southcentral market or when the price in the western market is about 1% lower than it in the southcentral market, the arbitrage activities will be triggered and the adjustment speed is faster in the lower regime. Moreover, the positive a_2 and b_2 provides supportive evidence for the LOP between the western and southcentral markets.

In the Southeast-Southcentral markets, a_2 and b_2 are greater than 0, which implies the validity of LOP between the southeastern and the southcentral markets. $a_2 = 0.2062$, $b_2 = 0.2393$ and the "neutral band" is [-0.0317, 0.0432]. It means that when the price in the southeastern market is around 3% lower than it in the southcentral market or when the price in the southeastern market is around 4% higher than it in the southcentral market, arbitrage activities will be triggered and the adjustment rates are very close. In the Southwest-Southcentral markets, a_2 and b_2 are still positive, which supports the validity of

Table 2.3: Parameters of West-Southeast

	Estimate	**Std. Error**
a_1	0.0031	0.00094
a_2	0.0923	0.01299
a_3	0.0274	0.00060
θ_1	0.0217	0.00056
b_1	0.0022	0.00086
b_2	0.0616	0.01084
b_3	0.0259	0.00056

Table 2.4: Thresholds of West-Southeast

c_1	c_2
-0.0212	0.1104

LOP between the southwestern and the southcentral markets. And the adjustment speed of upper regime is significantly higher than it of lower regime, compared to the case in the Southeast-Southcentral markets. In the Southwest-Southeast markets, $a_2 = 0.2029$, $b_2 = 0.2174$. It shows that adjustment rates between two regimes are close while the log price differentials in the upper regime will trigger a little bit faster adjustment, and the result also supports LOP between the southwestern and the southeastern markets.

In general, when the log price differentials fall outside of the band triggering the spatial trades, the three pairs among three southern markets present faster adjustment. And from our results, we can conclude that the series of log price differentials are bounded by two mean-reverting Ornstein-Uhlenbeck processes, and it takes 3 or 4 weeks on average for the log price differentials to get back into the band.

Table 2.5: Comparison Between Real Data and Simulated Data of West-Southeast

	Number of Log Price Differentials within Band	**Time Period**
Real Data	740	4 (median)
Simulated Data	738 (mean)	3 (median)

Table 2.6: Parameters of West-Southwest

	Estimate	**Std. Error**
a_1	0.0070	0.00128
a_2	0.0985	0.01379
a_3	0.0264	0.00059
θ_1	0.0211	0.00053
b_1	0.0046	0.00112
b_2	0.0652	0.01081
b_3	0.0287	0.00060

Table 2.7: Thresholds of West-Southwest

c_1	c_2
0.0194	0.1564

Table 2.8: Comparison Between Real Data and Simulated Data of West-Southwest

	Number of Log Price Differentials within Band	**Time Period**
Real Data	791	3 (median)
Simulated Data	779 (mean)	3 (median)

Table 2.9: Parameters of West-Southcentral

	Estimate	**Std. Error**
a_1	0.0048	0.00107
a_2	0.0845	0.01226
a_3	0.0272	0.00058
θ_1	0.0224	0.00054
b_1	0.0029	0.00091
b_2	0.0598	0.01070
b_3	0.0256	0.00055

Table 2.10: Thresholds of West-Southcentral

c_1	c_2
-0.0114	0.1413

Table 2.11: Comparison Between Real Data and Simulated Data of West-Southcentral

	Number of Log Price Differentials within Band	**Time Period**
Real Data	832	4 (median)
Simulated Data	784 (mean)	3 (median)

Table 2.12: Parameters of Southeast-Southcentral

	Estimate	**Std. Error**
a_1	0.0054	0.00064
a_2	0.2062	0.01824
a_3	0.0145	0.00030
θ_1	0.0126	0.00028
b_1	0.0058	0.00064
b_2	0.2393	0.02156
b_3	0.0114	0.00026

Table 2.13: Thresholds of Southeast-Southcentral

c_1	c_2
-0.0317	0.0432

Table 2.14: Comparison Between Real Data and Simulated Data of Southeast-Southcentral

	Number of Log Price Differentials within Band	**Time Period**
Real Data	969	3 (median)
Simulated Data	912 (mean)	2 (median)

Table 2.15: Parameters of Southwest-Southcentral

	Estimate	**Std. Error**
a_1	-0.0030	0.00052
a_2	0.2147	0.01813
a_3	0.0156	0.00032
θ_1	0.0138	0.00031
b_1	-0.0064	0.00065
b_2	0.2619	0.02180
b_3	0.0116	0.00025

Table 2.16: Thresholds of Southwest-Southcentral

c_1	c_2
-0.0822	-0.0008

Table 2.17: Comparison Between Real Data and Simulated Data of Southwest-Southcentral

	Number of Log Price Differentials within Band	**Time Period**
Real Data	989	2 (median)
Simulated Data	938 (mean)	2 (median)

Table 2.18: Parameters of Southwest-Southeast

	Estimate	**Std. Error**
a_1	-0.0081	0.00087
a_2	0.2029	0.01808
a_3	0.0166	0.00035
θ_1	0.0136	0.00030
b_1	-0.0105	0.00103
b_2	0.2174	0.01930
b_3	0.0146	0.00031

Table 2.19: Thresholds of Southwest-Southeast

c_1	c_2
-0.0797	-0.0076

Table 2.20: Comparison Between Real Data and Simulated Data of Southwest-Southeast

	Number of Log Price Differentials within Band	**Time Period**
Real Data	981	3 (median)
Simulated Data	903 (mean)	2 (median)

2.3.4 Impulse Response Analysis

To interpret the dynamic interrelationships among prices at 6 pairs of markets, we utilize the generalized impulse response functions. Generalized impulse response function is defined as the difference between the a conditional and unconditional forecast, where the conditioning information set is the shock to the tth variable (Koop et al. (1996)). The shock originates from the stochastic error term in the VAR models and in the SDE, the shock is from the stochastic term. We adopt the nonlinear impulse response function approach of Koop, Pesaran, and Potter (1996), which defines responses, GI, on the basis of observed data, y_t, and a shock, v. We chose the last observation of our dataset to evaluate responses to four-standard deviation positive and negative shocks.

$$GI_{t+k}(v, y_t, y_{t-1}, \ldots) = E(y_{t+k}|y_t + v, y_{t-1}, \ldots) - E(y_{t+k}|y_t, y_{t-1}, \ldots) \tag{2.7}$$

We choose four-standard deviation as the shock because we would like to see when the size of the shock is large enough to push the price differential outside of the neutral band between the thresholds, what the adjustment path would be like. Figure 2.2 illustrates responses to four-standard deviation positive and negative shocks in each pair of the markets and the two grey lines provide a 90% confidence interval. Note that each plot illustrates responses to both positive and negative price shocks.

From Figure 2.2, we observe that west-southeast, west-southwest and west-southcentral

markets take longer to eliminate the shocks, and the southeast-southcentral, southwest-southcentral and southwest-southeast markets take significant less time. The positive shock or negative shock could be caused by an increase in prices in one market or a decrease in prices in the other, and we care about the dynamic adjustment path caused by the relative price increases or decreases in each pair of the markets.

The three pairs of markets in the southern region appear to fully adjust to four-standard deviation positive and negative shocks after about 25 weeks. However, it takes up to about 75 weeks to fully adjust to the shocks when western market is involved. The reason might be that the three southern regions are geographically closer to each other. It should be noted that four-standard deviation shocks are relatively large and are not likely to be observed in week to week price movements. When one considers the time required for long-distance arrangements to purchase and transport bulky commodities such as plywood, combining with the huge change in price differentials, adjustment period of at least 20 more weeks are not that surprising. The impulse responses also indicate significant asymmetries in the patterns of adjustment to positive and negative shocks among southern markets. Such asymmetries may reflect asymmetric (with respect to direction) transactions costs, that is, transactions costs are higher for trade in one direction than is the case for trade in the opposite direction (Goodwin et al. (2021)). In all, our results are consistent with efficient linkages among spatially distinct regional markets for plywood and exhibit support for law of one price.

2.4 Conclusion

We have evaluated spatial market linkages and weekly price dynamics among a set of 4 important US markets for plywoods utilizing threshold stochastic differential equation

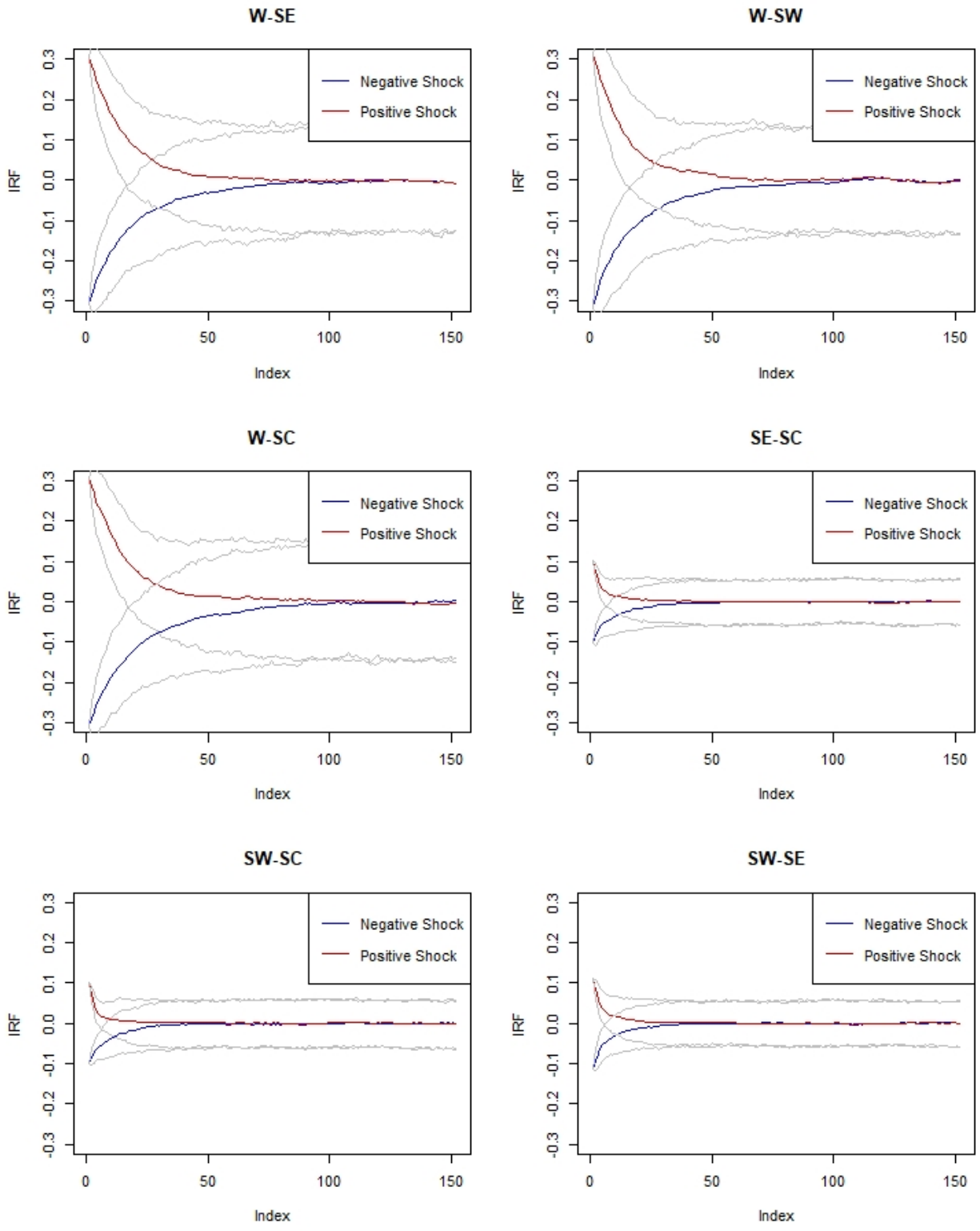


Figure 2.2: GIRFs of Six Pairs of Markets

models. And we find that the log price differentials are bounded by two mean reverting Ornstein-Uhlenbeck processes and the log price differentials falling within the neutral band can be described by a Brownian Motion. Our results indicate that southeast-southcentral, southwest-southcentral and southwest-southeast markets adjust faster to neutral band when arbitrage activities are triggered and the reason could be that the three markets are geographically closer to each other.

We then utilize generalized impulse response functions to evaluate the dynamic paths of adjustment to shocks. The adjustments confirm converging responses which is consistent with integration of markets. The three pairs of markets in the southern region appear to fully adjust to four-standard deviation shocks after about 25 weeks and the other three pairs of markets need about 75 weeks to get fully adjusted because of the geographical distance. Also, the impulse responses demonstrate siginificant asymmetries in price differentials responses to shocks in southeast-southcentral, southwest-southcentral and southwest-southeast markets.

The results are consistent with adherence to the Law of One Price and market integration in plywood markets in the United States. Also, the impulse reponses to shocks indicate that the responses may differ for positive and negative shocks, and the reason could be the asymmetric transactions costs.

CHAPTER

3

AN ANALYSIS OF THE RELATIONSHIP BETWEEN AGRICULTURAL COMMODITY MARKETS AND FINANCIAL MARKETS: A REGIME SWITCHING APPROACH

3.1 Introduction

A main function of futures markets is price discovery. Price discovery refers to the use of futures prices in determining cash prices. The significance of this contribution of futures markets depends on a close relationship between the futures and cash prices, which can reflect the impact of futures markets on cash market transactions (Garbade and Silber (1983); Schroeder and Goodwin (1991)). The futures price may be regarded as the market's expectation of the cash price in the future (Gardner (1976)) and the degree to which its changes is reflected in the cash market has important implications for hedgers and arbitragers who use these markets (Schroeder and Goodwin (1991)). This study investigates the short-run relationships between cash and futures prices and the long-run cointegration of the cash and futures prices for storable and nonstorable commodities.

There are many factors determining the long-run relationship between cash and futures prices. The futures market provides an important price discovery vehicle for the underlying assets and this is fulfilled by all market participants involved in futures trading. The actions of those participants guarantee the long-run relationship to some degree. For example, when speculators and hedgers trade futures contracts, they are acting on their own market analysis of supply and demand conditions in the underlying assets. Their trading, in turn, affects futures prices. The discovery of the underlying asset's price is completed as individuals' information and analysis become visible to the wider market through their trading in futures markets (Patel and Tkac (2007)). The introduction of futures contracts also provides a platform for arbitrageurs, who exploit pricing anomalies between spot and futures markets to produce a riskless profit. Arbitrage activities should therefore ensure that any information first involved in futures markets should immediately be reflected in the underlying spot markets, and vice versa. This has the effect of preventing prices in both

markets from diverging too far away from their equilibrium values.

The price discovery functions may be different for storable and nonstorable commodities. In general, commodity futures markets can perform both storage facilitation and forward pricing roles in their price discovery function. As storable commodities have been important in developing futures markets historically, both theoretical and empirical research have focused on the storage facilitation role (Yang, Bessler and Leatham (2001)). The storage facilitation role is important to price discovery because arbitrage may work through storage. As storage is the economic force that links cash and futures prices together, arbitrage may not work effectively without storage. And arbitrage is the underlying economic mechanism that maintains linkages and controls deviations from equilibrium (von Cramon-Taubadel and Goodwin (2021)).

Researchers have had difficulty visualizing futures trading for nonstorable commodities because of the lack of a concept similar to a carrying charge for storable commodities (Purcell and Hudson (1985)). As summarized in previous studies (Kamara (1982); Purcell and Hudson (1985)), relatively little evidence has been found in favor of price discovery functioning for nonstorable commodities, but there is more evidence for price discovery functioning for storable commodities. In fact, because cash and futures prices of commodities are usually nonstationary, many researchers have used notions of cointegration (Engle and Granger (1987)) to investigate price discovery in futures markets, and some studies (Covey and Bessler, 1995; Randall and Zapata, 1993) have argued that cointegration may depend on asset storability and they have argued that researchers should not expect cointegration between cash and futures prices for nonstorable commodities but should expect it for storable commodities.

The dependences of price discovery on the types of commodities and the nonstationarities of commodity prices imply that the cash prices are related to futures prices in complex

patterns. Some empirical results suggest cash and nearby futures prices in livestock markets are strongly related (Leuthold (1979); Oellermann and Farris (1985); Oellermann et al. (1989)). However, research also shows there is a lack of conintegration in the live hogs market (Schroeder and Goodwin (1991)). The relationship between cash and futures prices for storable commodities has been tested in past studies and the empirical results confirm the presence of a long-run cointegrating relationship (Kuiper et al. (2002); Adämmer and Bohl (2018); Wu et al. (2018)). Besides the long-run relationship, existing research provides the evidence of asymmetric price transmission (APT) in the futures and cash prices (Chang et al. (2012); Wu et al. (2018)).

There is considerable research about vertical and spatial APT, which describing the asymmetric adjustment toward equilibrium (Meyer and von Cramon-Taubadel (2004); von Cramon-Taubadel and Goodwin (2021)). A positive price asymmetry occurs when a decrease in prices is not fully or immediately transmitted, but an increase passes on more quickly or thoroughly, while a negative price asymmetry occurs when an increase in prices is not fully or immediately transmitted, but a decrease passes on more quickly or thoroughly.

Inventory management is one possible reason for APT. Production lags and finite inventories imply that negative shocks to the future optimal consumption path can be accommodated more quickly than positive shocks (Borenstein, Cameron, and Gilbert (1997)). The APT may also come from leverage effects, that is, the effect of news is greater if the previous shock is negative. Market participants tend to react more to negative (bad) than positive (good) news of equal magnitude. Moreover, the adjustment process for bad news can take more time as people tend to overreact to them. Another reason of APT could be the transaction costs. Rising transaction costs during times of high volatility can reduce contract liquidity and may hinder the adjustment process. More liquid contracts induce

greater symmetry in the adjustment process, while less liquid contracts infuse asymmetry, leading to slower adjustments during lower transactions periods. Also, government regulations could lead to APT (Chang et al. (2012)). For commodities, some producer groups have claimed that grain merchants are quick to adjust local spot prices when the futures price drops but slow to transmit increases in the futures price. Grain farmers may not be as price sensitive during periods of increasing commodity prices, allowing merchants to delay strengthening the local basis. On the other hand, a rapid increase in the futures price may have prompted grain merchants to accelerate the local price as a means to obtain product for sale. The direction of the asymmetry may depend on whether the crop is primarily exported or used within the local area (Wu et al. (2018)). To test for the presence of APT, many forms and variants of the cointegration and error correction models have been applied (Chang et al. (2012); Wu et al. (2018)).

The World Health Organization has declared the COVID-19 disease a global pandemic and the pandemic may change the patterns in which the cash prices are related to futures prices. Recent studies find that sudden panic demand has dramatically increased short-term price volatility in agricultural commodities, and COVID-19 has had a disruptive impact on prices and supply chains in agricultural markets (Adewopo et al. (2021); Zeng and Lu (2022)).

From January to April, 2020, the May futures price for corn has fallen by 15%. Ethanol has kept corn in high demand, and made it the most-planted U.S. crop. Today's corn crop is mainly used for biofuels (roughly 40% of U.S. corn is used for ethanol) and as animal feed (roughly 36% of U.S. corn, plus distiller's grains left over from ethanol production, is fed to cattle, pigs and chickens). Much of the rest is exported. Only a tiny fraction of the national corn crop is directly used for food for Americans, much of that for high-fructose corn syrup. When consumption of gasoline had plummeted due to COVID-19-related travel

restrictions, it was leading to significant declines in both ethanol demand and prices. Thus ethanol and corn prices have fallen concurrently.

The May futures price for soybeans has fallen by 10%. The shutdown of travel had resulted in a substantial fall in demand for fuel for transport, which accounted for two-thirds of global oil demand. Many countries had implemented wide-ranging travel bans and stay-at-home orders, sharply reducing travel. Lower demand and prices for oil also reduced the price of crops used for biofuels, such as corn and soybean oil.

The higher demand for wheat-based products in U.S. groceries and the fact that winter wheat has already been planted help to support wheat prices. The May wheat futures price was down only 3%. And a spike in domestic retail flour, bread, and wheat-based product sales, which are related to greatly increased expenditures on food eaten at home, contributed to the observed wheat price increase.

The May futures price for cotton, a product heavily dependent on manufacturing capacity abroad, declined by nearly 30%. Global cotton mill use sustained unprecedented decline because of the pandemic. During the COVID-19 pandemic, consumers deferred purchases of items such as clothing, and the associated industries adjusted operations accordingly. For the textile and apparel industry, these adjustments had included temporary closures or substantial reductions in manufacturing operations, as postponement or cancellation of orders had ripple effects throughout the entire supply chain, from raw fiber procurement to retail sales. The developing global economic slowdown had significantly reduced world cotton demand.

As for livestock futures, June lean hog futures price has declined by nearly 40%. Closures and production slowdowns of several hog-processing facilities beginning in early April due to workforce absences caused by the COVID-19 virus decreased processor demand for hogs, reducing weekly hog slaughter numbers and driving April hog prices (prices paid to hog

producers by processors) more than 35% below prices in April 2019. As of late April, the closures comprised more than a third of the industry's processing capacity. At the time of the first plant closing, pork demand was already low, as much of the hotel, restaurant, and institutional sector had been shut down to block transmission of the virus.

The economic uncertainty related to the impact of COVID-19 on the global economy and the demand destruction for many agricultural products contributed to significant price declines for crops and animal proteins. The decline in futures prices likely coincided with declines in cash market prices as well. Therefore, it is worth comparing the relationships between futures and cash prices before and after COVID.

In this paper, we apply a cointegration model to examine the long-run relationship, and a TVECM (Threshold Vector Error Correction Model) to examine the short-run relationship between spot and futures prices, using the basis, expressed as the error correction term, as forcing variable (Chang et al. (2012); Wu et al. (2018)), to decribe the asymmetric price transmission for corn, soybean, wheat, cotton and lean hogs. Then we use the impulse response to analyze the impacts of positive and negatives shocks from cash and futures prices on the adjustment pattern of the basis in five markets. The results reaffirm the cointegration relationship and the nonlinear dynamic relationship between spot and futures prices in agricultral commodity markets using threshold models. We then compare the results from the data before and after COVID to present the impacts of the pandemic on the cointegration and price transmission.

3.2 Empirical Framework

3.2.1 Engle–Granger two-step cointegration model

Before performing a cointegration analysis, augmented Dickey–Fuller (ADF) test is used to check for stationarity in the spot and futures price series. Then the Engle and Granger (1987) two-step procedure is used to assess the linear and symmetric adjustment toward long-run equilibrium in the spot and futures prices. The first step is to use ordinary least squares to estimate the long-run equilibrium relationship between the logarithms of spot and futures prices for corn, soybean, wheat, cotton and lean hogs. The equation is given by:

$$s_t = \beta_0 + \beta_1 f_t + b_t \tag{3.1}$$

where s_t and f_t are integrated of order 1; b_t is the disturbance term, that is, the basis; β_0 and β_1 are parameters. According to Engle and Granger (1987), the basis b_t should be stationary if a stable long-run relationship between the spot and futures prices exists. The second step is to apply the ADF test to verify the stationarity of residuals $\hat{b}_t$ by estimating the following equation:

$$\Delta b_t = \rho\, b_{t-1} + \sum_{p=1}^{k} \phi_p \Delta b_{t-p} + \upsilon_t \tag{3.2}$$

where υ_t is a white noise error term, $\sum_{p=1}^{k} \phi_p \Delta b_{t-p}$ is the sum of lagged differenced terms to correct for autocorrelation.

A test of cointegration is a test of whether $\hat{b}_t$ is stationary. Thus the spot price and futures price form a long-run cointegrated relationship if the null hypothesis $\rho = 0$ is rejected in favor of $\rho < 0$.

3.2.2 Threshold Vector Error Correction Model

The Granger representation theorem states that an error correction representation exists for any cointegrated series (Engle and Granger (1987)). The threshold cointegration model, which was introduced by Balke and Fomby (1997), allows for non-linear adjustment to long-run equilibrium. In this paper, we use a two-regime vector error correction model with a single cointegrating vector and a threshold effect in the error-correction term (ECT). The TVECM is used to determine the dynamics by which the spot and futures price revert to the long-run equilibrium level. A TVECM that reveals differential adjustments to positive and negative short-term deviations from the equilibrium can be represented as:

$$\Delta x_t = \begin{cases} A_1' X_{t-1}(\beta) + u_t & \text{if } w_{t-1}(\beta) \leq \gamma \\ A_2' X_{t-1}(\beta) + u_t & \text{if } w_{t-1}(\beta) > \gamma \end{cases} \tag{3.3}$$

where

$$X_{t-1}(\beta) = \begin{pmatrix} 1 \\ w_{t-1}(\beta) \\ \Delta x_{t-1} \\ \Delta x_{t-2} \\ \vdots \\ \Delta x_{t-l} \end{pmatrix}$$

x_t is a 2-dimensional time series, which is cointegrated with one 2×1 cointegrating vector β. In this study, x_t contains the cash price and futures price of an agricultural commodity. $w_{t-1}(\beta) = \alpha + \beta' x_{t-1}$ denote the error-correction term (ECT) (i.e., b_t), and γ is the threshold parameter. The coefficient matrices A_1 and A_2 govern the dynamics in these two regimes, and u_t is the error term with finite covariance matrix $\Sigma = E(u_t u_t')$. The coefficients on

the ECT are the speed of adjustment parameters, representing the speed at which the cointegrated series move back toward their long-run equilibrium relationship. Under the assumption that the errors u_t are iid Gaussian, the likelihood is

$$L_n(A_1, A_2, \Sigma, \beta, \gamma) = -\frac{n}{2} log|\Sigma| - \frac{1}{2}\sum_{t=1}^{n} u_t(A_1, A_2, \beta, \gamma)'\Sigma^{-1} u_t(A_1, A_2, \beta, \gamma) \tag{3.4}$$

The MLE $(\hat{A}_1, \hat{A}_2, \hat{\Sigma}, \hat{\beta}, \hat{\gamma})$ are the values which maximize $L_n(A_1, A_2, \Sigma, \beta, \gamma)$.

In the low regime, that is, the cash price is relatively lower than the futures price, arbitrageurs will buy the spot commodity and sell futures contracts to profit from the difference. On the contrary, the high regime, that is, the cash price is higher than the futures price, arbitrageurs will profit from selling the spot commodity and buying the futures contracts. As we mentioned before, the transaction cost could lead to APT in the commodity market and the theory of storage (Fama and French (1987)) could be used to explain the asymmetric adjustment pattern to some degree. The theory claims that the return from purchasing the commodity today and selling it for delivery later (the so-called basis) equals the interest forgone by storing the commodity plus the marginal storage cost less the marginal convenience yield from an additional unit of inventory. The relationship can be represented as:

$$s_t - f_t = -(r_t + u_t - c\,y_t)(T - t) \tag{3.5}$$

where s_t and f_t are the logarithm of the spot and futures prices, r_t is the interest rate, u_t is the marginal storage cost and $c\,y_t$ is the marginal convenience yield. T is the delivery date. Let's define the right side of equation (3.5), $(r_t + u_t - c\,y_t)(T - t)$, as the "net cost". Combining equation (3.5) with equation (3.1), we can regard β_0 as the negative "average net cost" when β_1 is close to 1 in equilibrium. Therefore, when the threshold value γ is close to 0, the low regime describes the situation when the difference between the futures price and

the cash price is greater than the “average net cost”, and the high regime depicts the case when the difference between the futures price and the cash price is less than the “average net cost”.

3.3 Empirical Application

3.3.1 Data

For the empirical analysis, we consider corn, soybean, wheat and cotton as our storable commodities; lean hogs as our nonstorable commodity. These five markets are selected because of their important roles in US agricultural commodity markets.

Corn is one of world’s most popular crops. Today the US is considered a world leader in corn exports (responsible for 20 percent of the world’s corn), as well as the leader in corn production (accounting for 39 percent of the world’s cultivation). One major part of corn production goes for livestock feed, and corn is also widely-applied for ethanol and food production. Corn futures are traded on the Chicago Mercantile Exchange (CME) and the contract size comprises 5000 bushels.

The US is the world’s leading soybean producer and the second-leading exporter and soybeans comprise about 90 percent of US oilseed production. They serve as a central ingredient in many food products, and are increasingly being seen as a renewable resource with many industrial applications. Soybeans are highly appreciated by farmers because they naturally improve the soil’s nitrogen level and clear the fields for other cultivated crops. Soybean futures are traded on the Chicago Mercantile Exchange (CME) and the contract size comprises 5000 bushels.

Wheat is one of the most widely planted crops in the US. The primary use for wheat is

flour, and wheat byproducts are also used in livestock feeds. Wheat has significant industrial applications, too, as an ingredient in starches, adhesives, and coatings. While there are several wheat futures classifications, Chicago Soft Red Winter (SRW) wheat sets the global industry standards for wheat price risk management, with it being the most liquid wheat futures contract in the world. Wheat futures are traded on the Chicago Mercantile Exchange (CME) and the contract size comprises 5000 bushels.

Cotton is naturally grown in different regions with the US being positioned among the top producing countries in the world. It is heavily used in the textile industry for clothing, and it is also used in the production of coffee filters, tents, etc. Cotton No. 2 is the futures contract where cotton is the underlying commodity, traded at the Intercontinental Exchange (ICE). The reference asset is cotton that fulfils the minimum defined standards in terms of basis grade and sample length. The Cotton No.2 futures settlement is executed with physical delivery and the futures have a contract size of 50000 pounds.

The lean hog futures contract, which is a contract for the hog's carcass, trades on the Chicago Mercantile Exchange (CME) and is used primarily by producers of lean hogs, both domestic and international, and pork importers and exporters. The lean hog contract is a fairly new addition to the CME, which was launched in 1997 as a replacement after the live hog futures contract was retired. The lean hog futures is financially settled and the futures have a contract size of 40000 pounds.

The daily cash and futures prices data are obtained from the Commodity Research Bureau (CRB Infotech CD). The data we use in this study consists of daily cash price and futures price from 1/4/2016 to 10/30/2020 (1145 observations). As with most existing studies, to avoid expiration effects (i.e., liquidation bias), we roll over to the next nearest contract one month prior to expiration of the contract. We use the 2020 data as the prices during COVID and the remainder as the before-COVID prices. The Chow test results in

Table 3.1: Chow Test Results

null hypothesis: 1/2/2020 is not a break date				
corn	soybean	wheat	cotton	lean hogs
<0.01	<0.01	<0.01	<0.01	<0.01

Table 3.2: Statistics of Five Markets: Prices

	Before COVID				COVID			
	min	mean	max	sd	min	mean	max	sd
corn cash	311.0	368.0	482.8	28.83	303.8	359.2	432.8	33.39
corn futures	301.5	369.4	454.8	22.32	302.8	351.1	419.2	29.98
soybean cash	745.0	908.2	1150.5	74.41	809.0	893.0	1070.5	64.80
soybean futures	802.5	943.8	1177.8	69.26	821.8	908.8	1087.8	67.57
wheat cash	348.0	473.3	623.0	52.26	484.0	565.5	652.0	42.55
wheat futures	361.0	469.8	574.5	41.72	474.0	539.5	632.8	33.71
cotton cash	54.41	71.10	94.21	8.11	44.66	59.27	69.03	5.60
cotton futures	55.93	72.57	95.21	8.16	48.41	62.58	72.11	5.50
lean hogs cash	45.44	67.28	92.84	10.59	44.55	58.99	78.69	9.46
lean hogs futures	41.10	68.10	92.75	10.47	37.33	58.41	72.80	8.18

Table 3.1 support the split of data .

We define cc, sc, wc, ctc, lhc as corn, soybean, wheat, cotton and lean hogs cash (spot) prices respectively, and cf , sf, wf, ctf, lhf as corresponding futures prices; lcc, lsc, lwc, $lctc$, $llhc$ as logarithms of the cash prices of corn, soybean, wheat, cotton and lean hogs respectively, and lcf,lsf,lwf,$lctf$,$llhf$ as corresponding logarithms of futures prices. Figure 3.1 to Figure 3.5 present the futures and cash prices of corn, soybean, wheat, cotton and lean hogs markets. Table 3.2 and Table 3.3 list the statistics of the five markets before and during COVID.

From the statistics of the five markets, we can see that corn, soybean, cotton and lean hogs markets have higher cash and futures prices before COVID while wheat market exhibits higher cash and futures prices during COVID. Before COVID, corn, soybean, cotton and

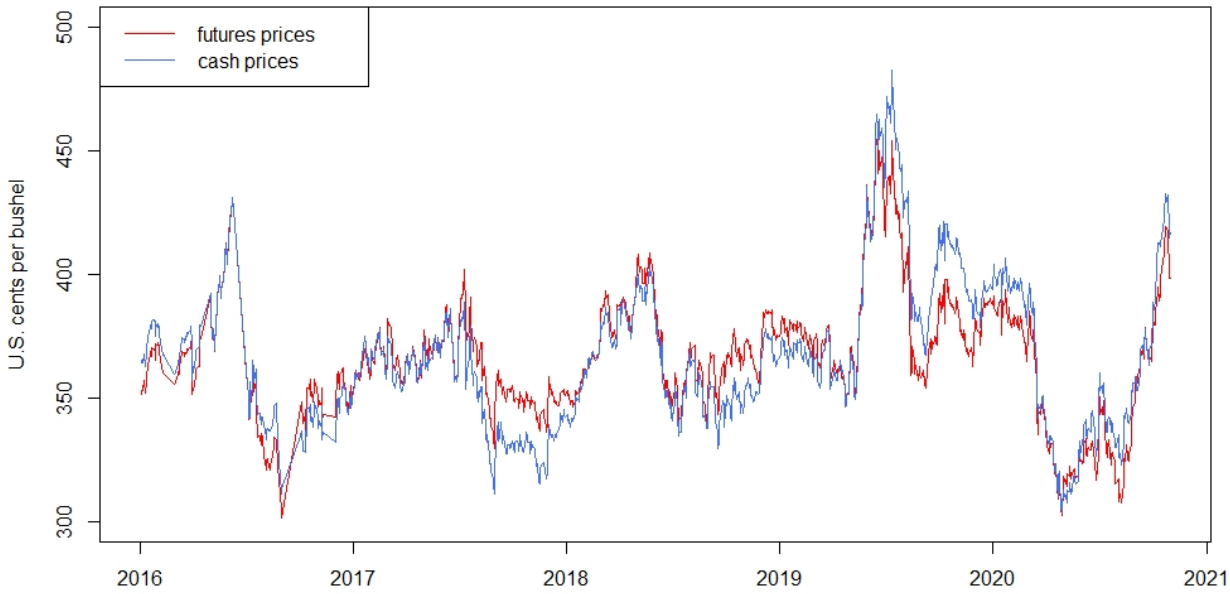


Figure 3.1: Corn Prices

Table 3.3: Statistics of Five Markets: Log Prices

	Before COVID				COVID			
	min	mean	max	sd	min	mean	max	sd
corn cash	5.740	5.905	6.179	0.076	5.716	5.880	6.070	0.092
corn futures	5.709	5.910	6.120	0.059	5.713	5.857	6.038	0.085
soybean cash	6.613	6.808	7.048	0.082	6.696	6.792	6.976	0.070
soybean futures	6.688	6.847	7.071	0.073	6.711	6.810	6.992	0.072
wheat cash	5.852	6.154	6.435	0.110	6.182	6.335	6.480	0.075
wheat futures	5.889	6.148	6.354	0.089	6.161	6.289	6.450	0.062
cotton cash	3.997	4.258	4.546	0.114	3.799	4.077	4.235	0.097
cotton futures	4.024	4.278	4.556	0.113	3.880	4.132	4.278	0.090
lean hogs cash	3.997	4.258	4.546	0.114	3.799	4.077	4.235	0.097
lean hogs futures	3.816	4.197	4.531	0.157	3.797	4.065	4.366	0.158

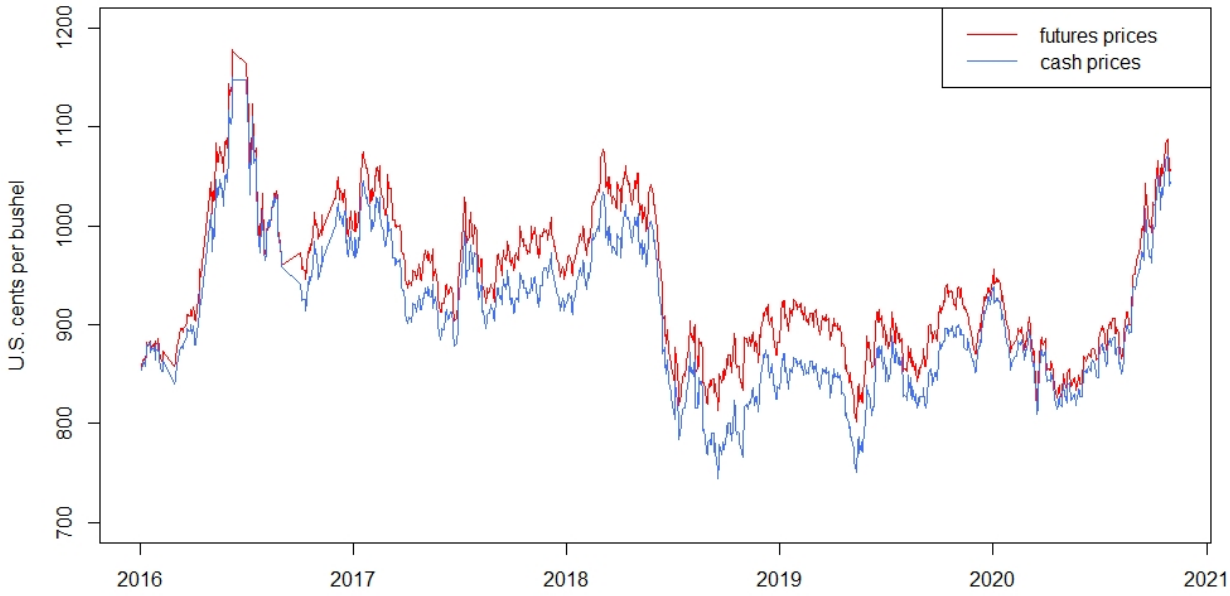


Figure 3.2: Soybean Prices

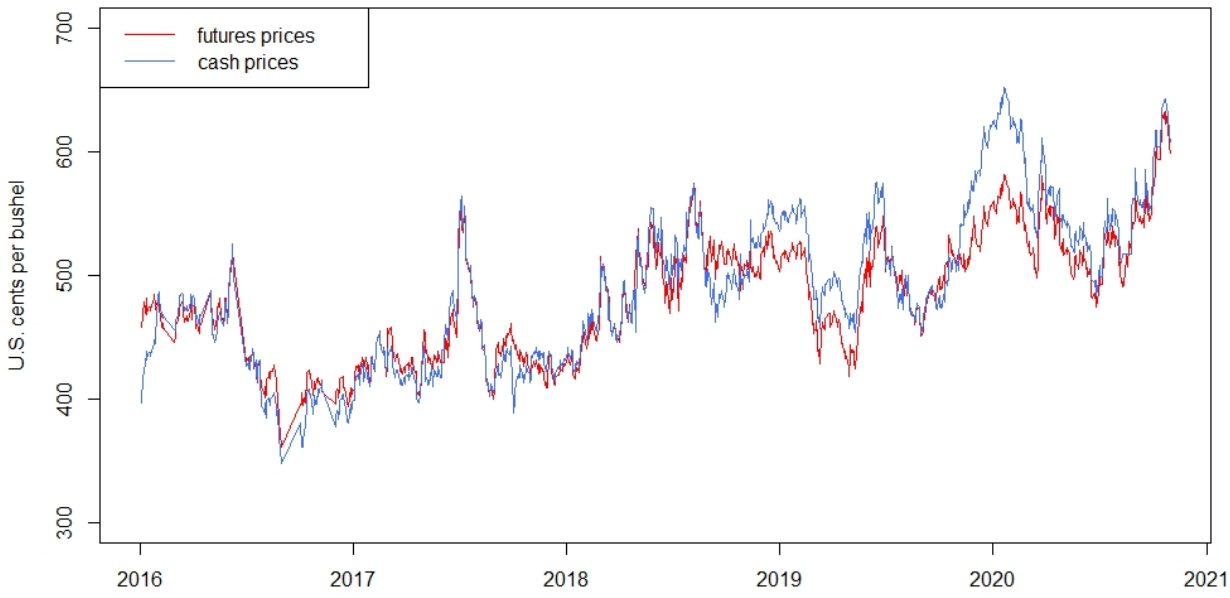


Figure 3.3: Wheat Prices

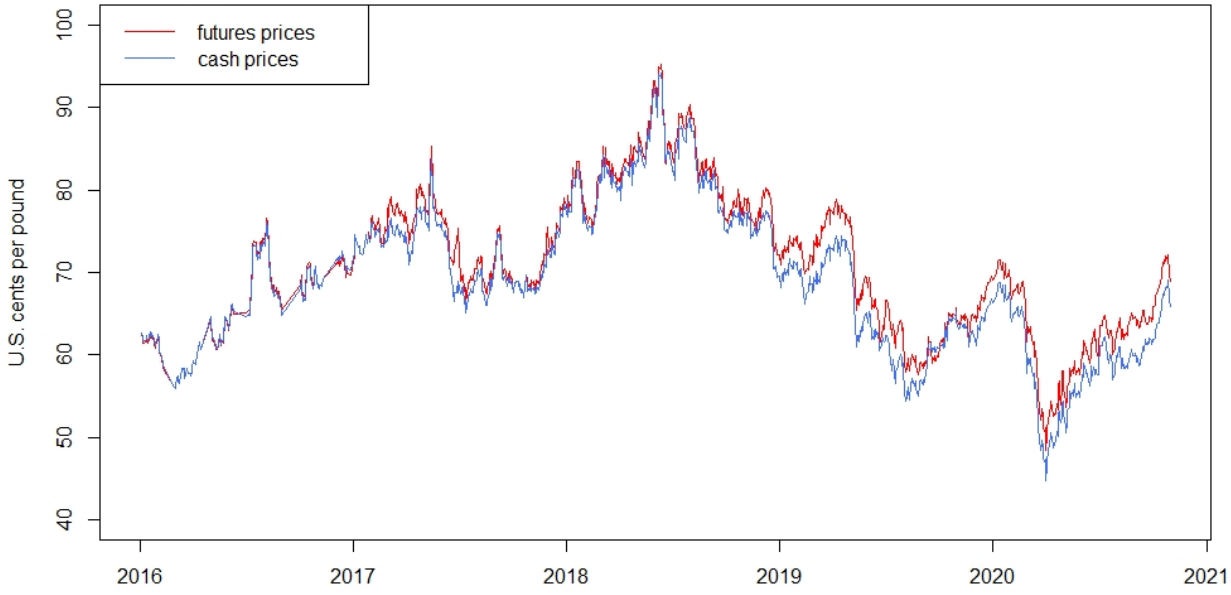


Figure 3.4: Cotton Prices

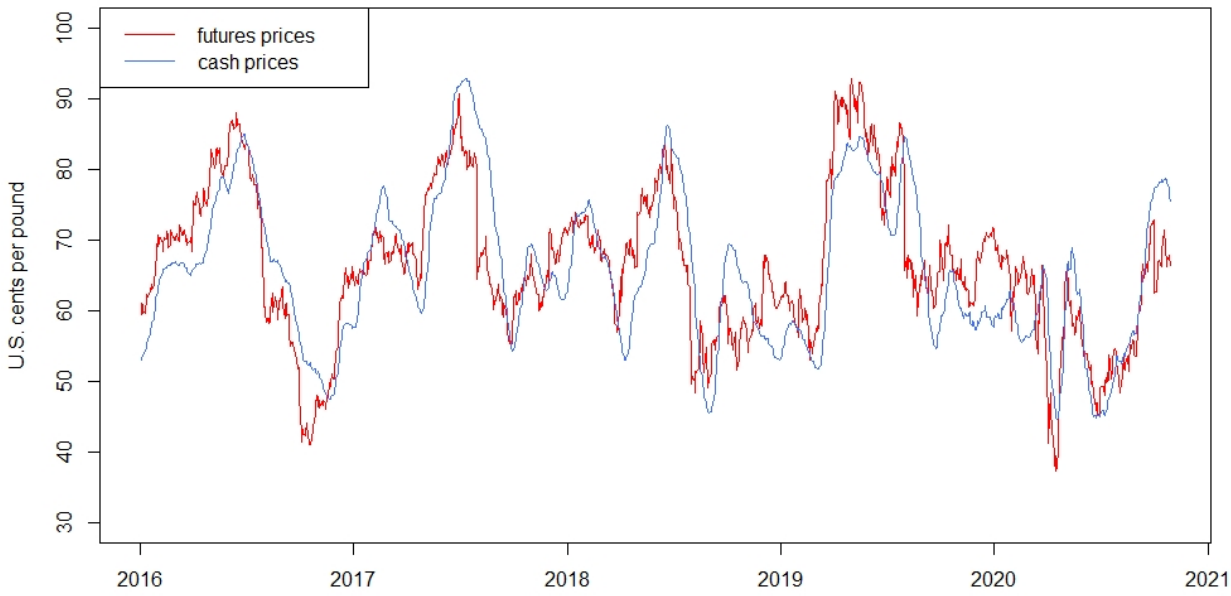


Figure 3.5: Lean Hogs Prices

lean hogs markets have higher futures prices relative to cash prices while wheat market has higher cash price relative to futures price. However, corn, wheat and lean hogs markets have higher cash prices relative to futures prices during COVID, while soybean and cotton markets still have higher futures prices relative to cash prices. We also find out that the corn market experiences a higher price volatility during COVID, while soybean, wheat, cotton and lean hogs markets present a lower price volatility during COVID.

3.3.2 Cointegration Test

The results of the ADF test used to check stationarity of logarithms of cash and futures prices are shown in Table 3.4. The null hypothesis cannot be rejected at any standard level of significance for almost all price level series. However, the null hypothesis is rejected at the 1% level of significance for the first-differenced data. Therefore, the ADF test indicates that the cash and the futures prices for corn, soybean, wheat, cotton and lean hogs are integrated of order 1.

Since the data series are integrated of order 1, it is possible that a cointegration relationship exists between the pair of cash price and futures price series. From Figure 3.1 to Figure 3.5, corn, soybean, wheat, cotton and lean hogs markets present the overall long-term co-movements among the futures and cash prices. So we implement a cointegration test to see if the two prices are cointegrated in each market and the results are presented in Table 3.5. And the estimates of the cointegration models of five markets are listed in Table 3.6.

3.3.3 Threshold VECM

Then we apply the threshold vector error correction model to corn, soybean, wheat, cotton and lean hogs markets.

Table 3.4: ADF Test for Prices Series: p value

alternative hypothesis: stationary				
	Before COVID		COVID	
	Level	First Difference	Level	First Difference
corn cash	0.2057	≤0.01	0.960	≤0.01
corn futures	0.0101	≤0.01	0.933	≤0.01
soybean cash	0.102	≤0.01	0.697	≤0.01
soybean futures	0.0385	≤0.01	0.672	≤0.01
wheat cash	0.0428	≤0.01	0.413	≤0.01
wheat futures	0.0100	≤0.01	0.755	≤0.01
cotton cash	0.346	≤0.01	0.795	≤0.01
cotton futures	0.288	≤0.01	0.759	≤0.01
lean hogs cash	0.010	≤0.01	0.174	≤0.01
lean hogs futures	0.046	≤0.01	0.247	≤0.01

Table 3.5: Cointegration Test Results

Engle-Granger Cointegration Test alternative: cointegrated		
data	Before COVID p value	COVID p value
corn	≤0.01	0.0209
soybean	≤0.01	≤0.01
wheat	≤0.01	0.0821
cotton	≤0.01	<0.01
lean hogs	≤0.01	0.0183

Table 3.6: Cointegration Estimates

	Before COVID		COVID	
	β_0	β_1	β_0	β_1
corn	-1.002	1.169	-0.3752	1.0678
soybean	-0.7211	1.0996	0.1702	0.9724
wheat	-0.9296	1.1520	-0.4701	1.0821
cotton	0.01299	0.99214	-0.3437	1.0699
lean hogs	0.9876	0.7624	0.7194	0.8246

Let $w_t^i(\beta) = \alpha + \beta' x_t^i$ (i.e., b_t^i) denote the error-correction term (ECT), then the linear cointegration model is given as follows:

$$\Delta x_t^i = A^{i'} X_{t-1}^i(\beta) + u_t^i \tag{3.6}$$

where

$$X_{t-1}^i(\beta) = \begin{pmatrix} 1 \\ w_t^i(\beta) \\ \Delta x_{t-1}^i \\ \Delta x_{t-2}^i \\ \vdots \\ \Delta x_{t-l}^i \end{pmatrix}$$

x_t^i is a 2-dimensional time series, which is cointegrated with one 2×1 cointegrating vector β, containing the cash price and futures price of market i, where $i =$ corn, soybean, wheat, cotton and lean hogs. A^i is a matrix of parameters and u_t^i is the error term.

The threshold cointegration model is given as follows:

$$\Delta x_t^i = \begin{cases} A_1^{i'} X_{t-1}^i(\beta) + u_t^i & \text{if } w_{t-1}^i(\beta) \leq \gamma \\ A_2^{i'} X_{t-1}^i(\beta) + u_t^i & \text{if } w_{t-1}^i(\beta) > \gamma \end{cases} \tag{3.7}$$

where γ is the threshold parameter and the error correction term $w_{t-1}^i(\beta)$ is the forcing variable. A_1^i and A_2^i are matrices of parameters depending on different states, and u_t^i is the error term. In equation (3.6), $i =$ corn, soybean, wheat, cotton and lean hogs so that when we analyze the situation before COVID, $x_t^{corn} = \begin{pmatrix} lccb \\ lcfb \end{pmatrix}$, $x_t^{soybean} = \begin{pmatrix} lscb \\ lsfb \end{pmatrix}$, $x_t^{wheat} =$

$\begin{pmatrix} lwcb \\ lwfb \end{pmatrix}$, $x_t^{cotton} = \begin{pmatrix} lctcb \\ lctfb \end{pmatrix}$, $x_t^{leanhogs} = \begin{pmatrix} llhcb \\ llhfb \end{pmatrix}$; when we analyze the situation during COVID, $x_t^{corn} = \begin{pmatrix} lcc \\ lcf \end{pmatrix}$, $x_t^{soybean} = \begin{pmatrix} lsc \\ lsf \end{pmatrix}$, $x_t^{wheat} = \begin{pmatrix} lwc \\ lwf \end{pmatrix}$, $x_t^{cotton} = \begin{pmatrix} lctc \\ lctf \end{pmatrix}$, $x_t^{leanhogs} = \begin{pmatrix} llhc \\ llhf \end{pmatrix}$.

The estimation results are listed Table 3.8-3.17.

3.3.4 Impulse Response Analysis

We utilize the generalized impulse response functions to interpret the dynamic adjustments of cash and futures prices among corn, soybean, wheat, cotton and lean hogs markets. Generalized impulse response function is defined as the difference between the a conditional and unconditional forecast, where the conditioning information set is the shock to the tth variable. In this study, we adopt the nonlinear impulse response function approach of Koop, Pesaran, and Potter (1996), which defines responses, GI, on the basis of observed data, y_t, and a shock, v. We choose the last observation to evaluate responses to two-standard deviation positive and negative shocks.

$$GI_{t+k}(v, y_t, y_{t-1}, \ldots) = E(y_{t+k}|y_t + v, y_{t-1}, \ldots) - E(y_{t+k}|y_t, y_{t-1}, \ldots) \tag{3.8}$$

We choose two-standard deviation as the shock because we would like to see when the size of the shock is large enough to evoke the dynamic adjustments below or above the thresholds, what the path would be like. Figures 3.6 to Figure 3.15 illustrate responses to two-standard deviation positive and negative shocks in each market before and during

Table 3.7: F Test for Threshold Effect

null hypothesis: VECM					
	corn	soybean	wheat	cotton	lean hogs
Before COVID					
p value	0.092	0.089	0.006	0.090	0.002
COVID					
p value	0.036	0.086	0.035	0.001	0.080

COVID.

Table 3.8: Estimates of Corn Market before COVID

	$w^i_{t-1}(\beta) \leq -0.045$				$w^i_{t-1}(\beta) > -0.045$			
	cash		futures		cash		futures	
w^i_{t-1}	0.040	(0.020)**	0.057	(0.021)***	0.178	(0.077)**	0.203	(0.081)**
Const	0.004	(0.002)**	0.006	(0.002)***	0.004	(0.002)**	0.006	(0.002)***
c_{t-1}	0.065	(0.100)	0.194	(0.104)*	0.065	(0.100)	0.194	(0.104)*
f_{t-1}	-0.037	(0.095)	-0.135	(0.099)	-0.037	(0.095)	-0.135	(0.099)

Standard errors are in parentheses.
***, **, and * denote significance level of 1%, 5% and 10%, respectively.

3.4 Results

3.4.1 Cointegration Test

Tables 3.5 presents the Engle and Granger (1987) cointegration tests for corn, soybean, wheat, cotton and lean hogs markets respectively. Based on the Engle and Granger (1987) cointegration test, the results of the five markets reject the null hypothesis of no cointe-

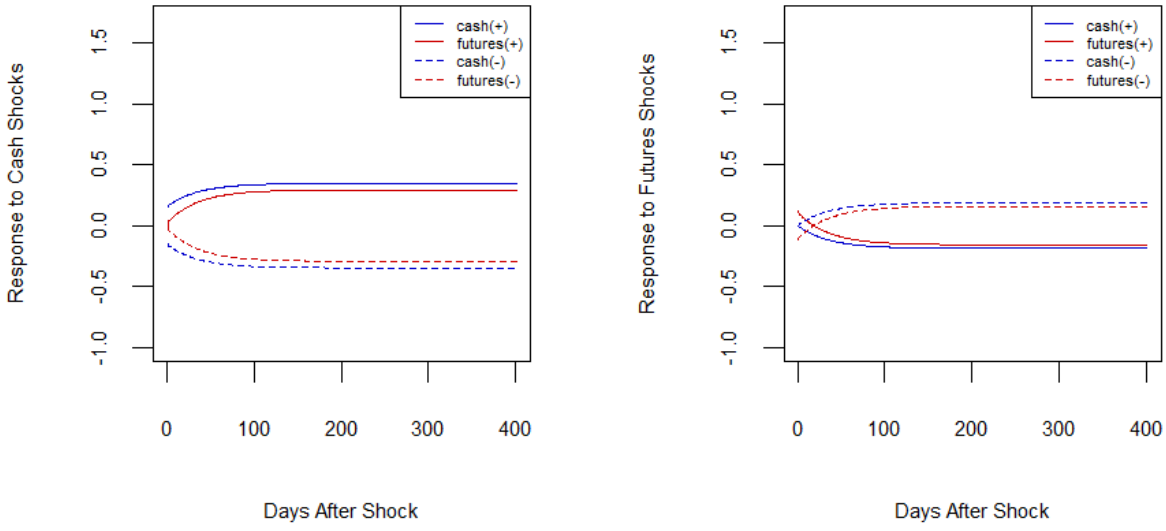


Figure 3.6: Responses in Corn Market: Before COVID

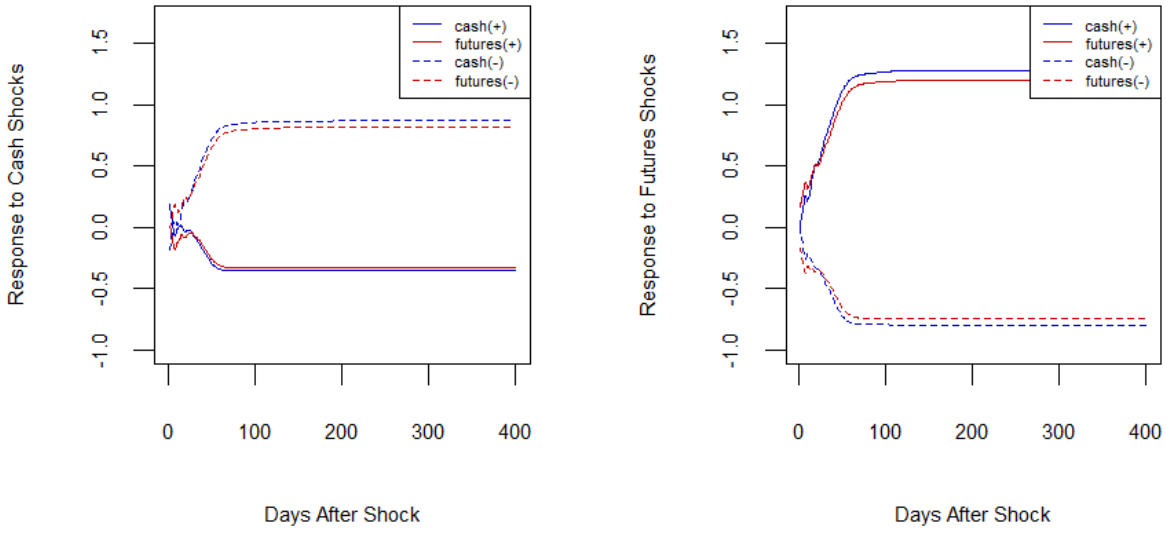


Figure 3.7: Responses in Corn Market: During COVID

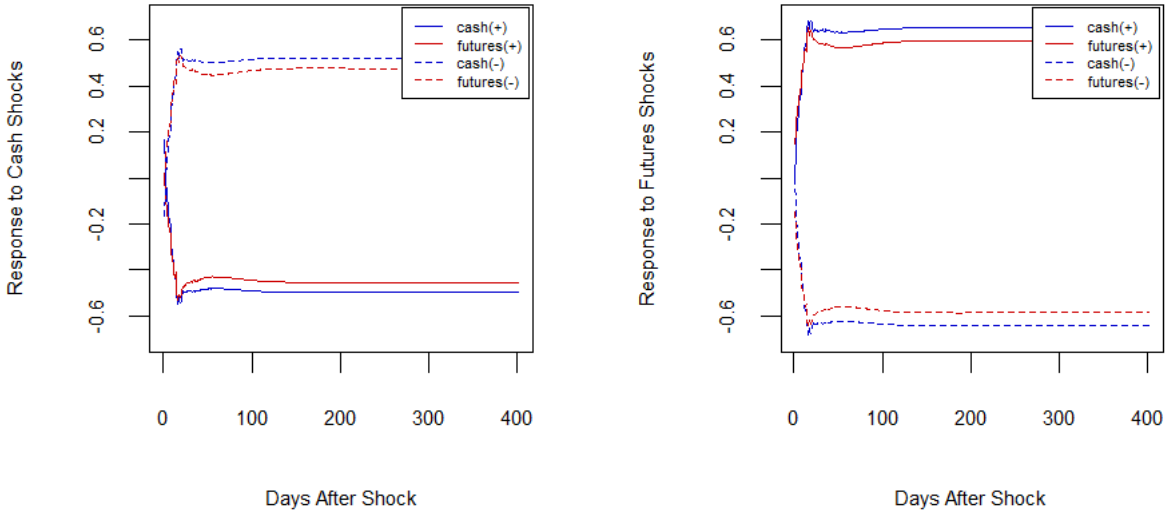


Figure 3.8: Responses in Soybean Market: Before COVID

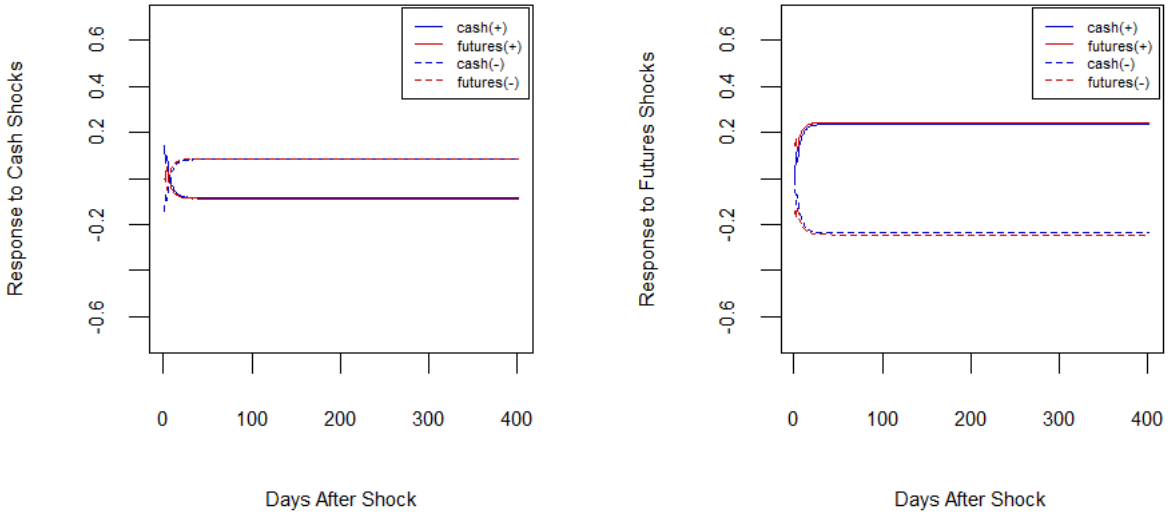


Figure 3.9: Responses in Soybean Market: During COVID

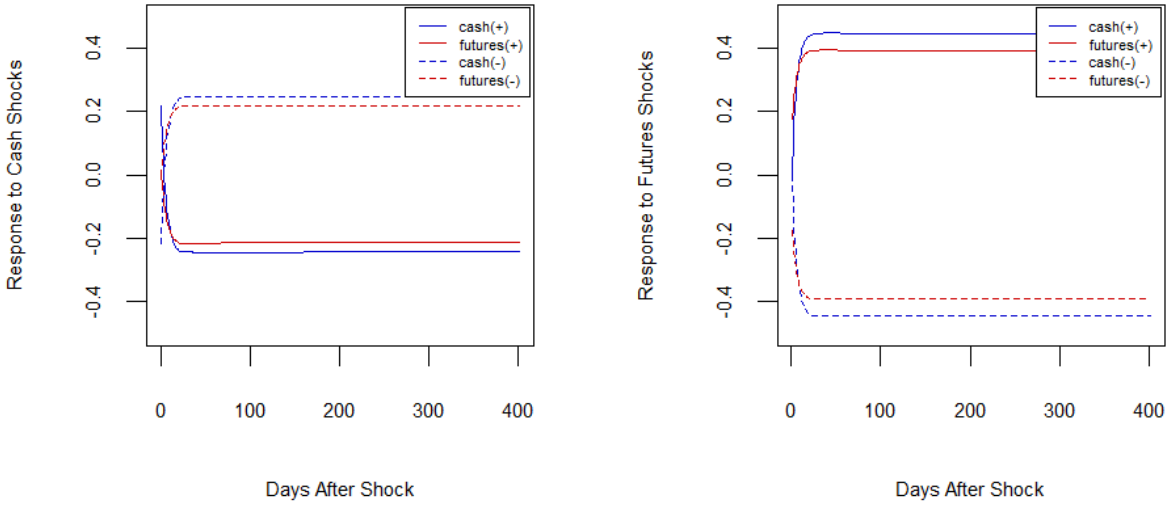


Figure 3.10: Responses in Wheat Market: Before COVID

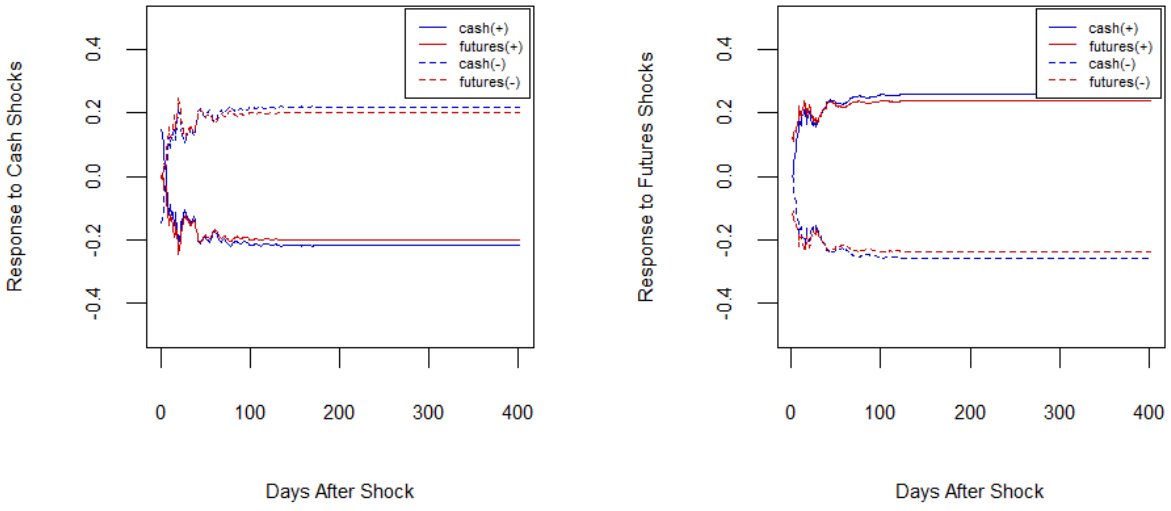


Figure 3.11: Responses in Wheat Market: During COVID

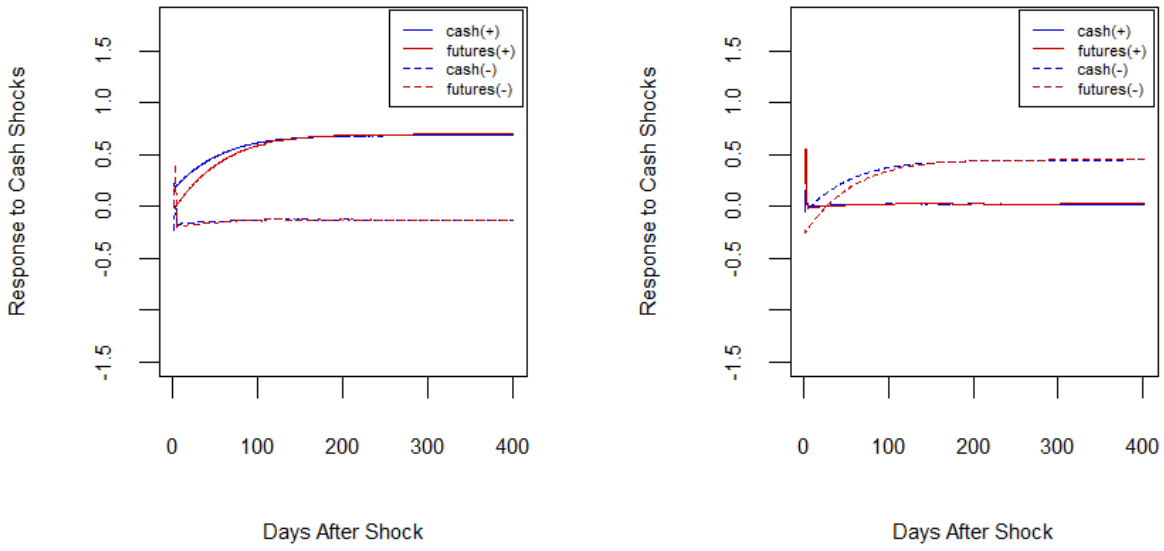


Figure 3.12: Responses in Cotton Market: Before COVID

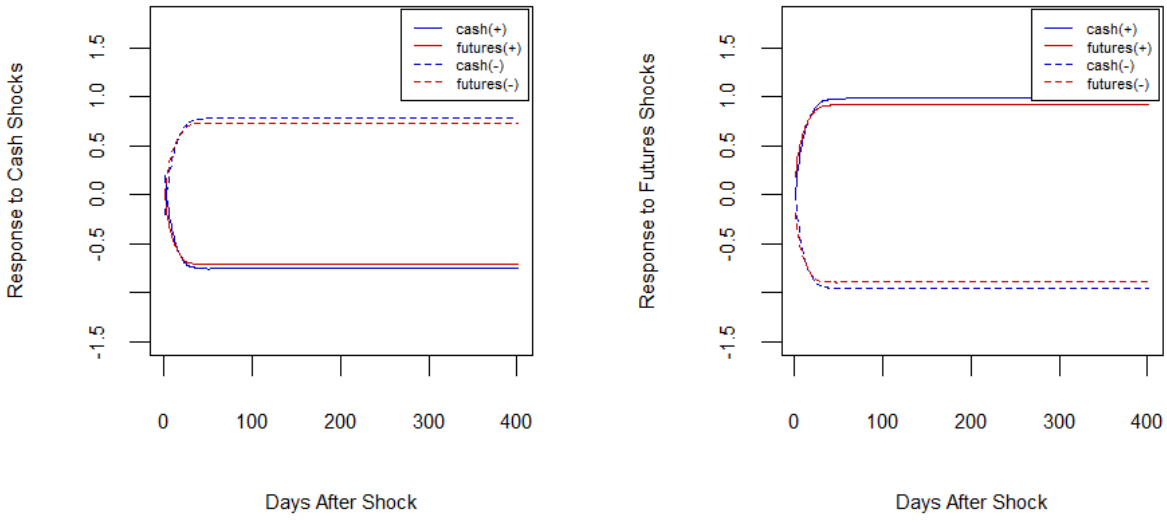


Figure 3.13: Responses in Cotton Market: During COVID

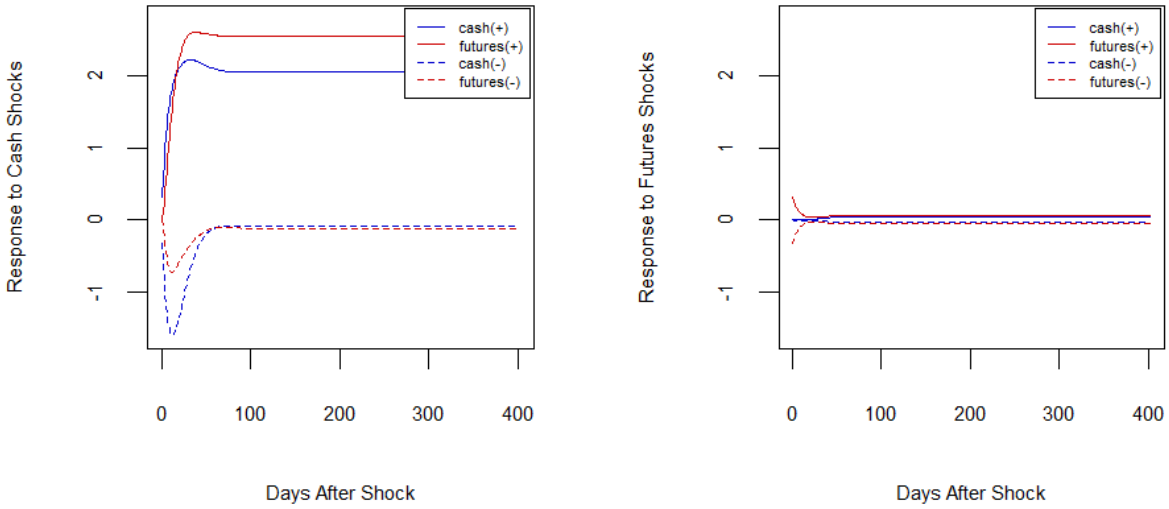


Figure 3.14: Responses in Lean hogs Market: Before COVID

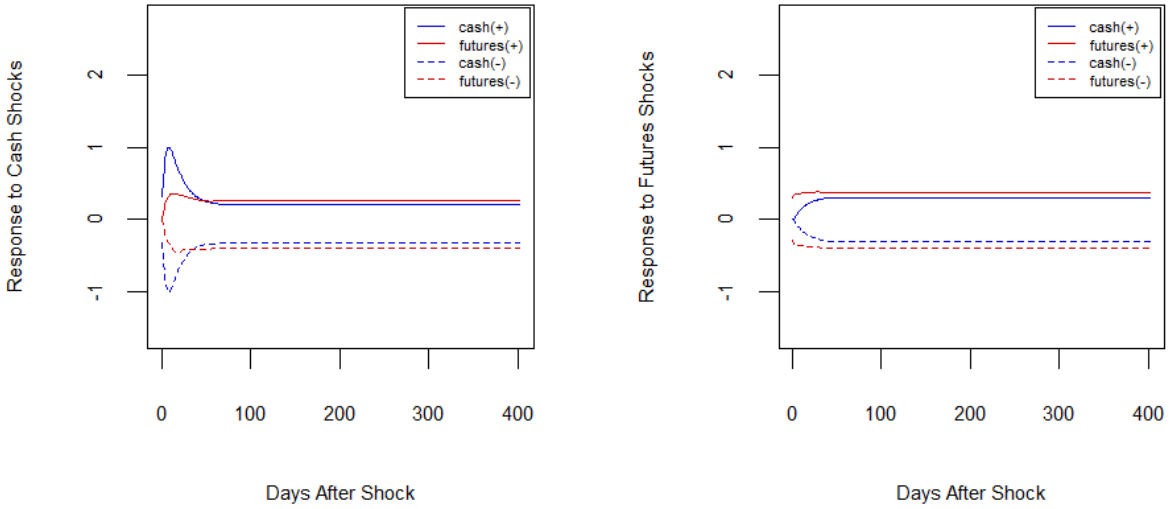


Figure 3.15: Responses in Lean hogs Market: During COVID

Table 3.9: Estimates of Corn Market during COVID

	$w^i_{t-1}(\beta) \leq 0.006$				$w^i_{t-1}(\beta) > 0.006$			
	cash		futures		cash		futures	
w^i_{t-1}	-0.271	(0.112)**	-0.157	(0.113)	0.153	(0.105)	0.195	(0.107)*
Const	-0.003	(0.002)*	-0.002	(0.002)	-0.003	(0.002)*	-0.002	(0.002)
c_{t-8}	0.403	(0.185)**	0.336	(0.187)*	0.403	(0.185)**	0.336	(0.187)*
f_{t-8}	-0.381	(0.183)**	-0.349	(0.185)*	-0.381	(0.183)**	-0.349	(0.185)*

Standard errors are in parentheses.
***, **, and * denote significance level of 1%, 5% and 10%, respectively.

gration at the 1% level of significance using the before-COVID data, which suggests that cash and futures prices are cointegrated significantly before COVID for both storable and nonstorable commodities. When we look at the COVID data, the cash and futures prices of soybean and cotton are cointegrated significantly during COVID, while corn and lean hogs markets reject the null hypothesis of no cointegration at the 5% level of significance and wheat market rejects the null hypothesis at the 10% level. This result shows that COVID has larger impact on the conintegration relationships in corn, lean hogs and wheat markets.

Table 3.6 shows the estimates of the cointegration models of five markets. The cointegrating parameter β_1 is close to 1 in all four storable commodities markets so that β_0 can be regarded as the "net cost" as we discussed before. In the corn market, the "net cost" before COVID is 1.002 and it is 0.3752 during COVID. The "net cost" before COVID is 0.7211 and it is -0.1702 during COVID in the soybean market. For wheat, the "net cost" is 0.9296 before COVID and it is 0.4701 during COVID. In the cotton market, the "net cost" is -0.01299 before COVID and it is 0.3437 during COVID. For storable commodities, when the "net cost" is greater than 0, it means the interest forgone by storing the commodity plus the marginal storage cost is greater than the marginal convenience yield from an additional unit of inventory, and the convenience yield is inversely related to inventories. In other words,

Table 3.10: Estimates of Soybean Market before COVID

	$w^i_{t-1}(\beta) \leq -0.011$				$w^i_{t-1}(\beta) > -0.011$			
	cash		futures		cash		futures	
w^i_{t-1}	-0.577	(0.232)**	-0.400	(0.225)*	0.026	(0.030)	0.049	(0.029)*
c_{t-8}	-0.186	(0.122)	-0.227	(0.119)*	-0.186	(0.122)	-0.227	(0.119)*
f_{t-8}	0.143	(0.126)	0.177	(0.123)	0.143	(0.126)	0.177	(0.123)
c_{t-9}	-0.090	(0.123)	-0.136	(0.119)	-0.090	(0.123)	-0.136	(0.119)
f_{t-9}	0.156	(0.127)	0.202	(0.123)*	0.156	(0.127)	0.202	(0.123)*
c_{t-11}	-0.155	(0.123)	-0.193	(0.119)	-0.155	(0.123)	-0.193	(0.119)
f_{t-11}	0.188	(0.127)	0.218	(0.123)*	0.188	(0.127)	0.218	(0.123)*
c_{t-14}	-0.314	(0.123)**	-0.335	(0.119)***	-0.314	(0.123)**	-0.335	(0.119)***
f_{t-14}	0.307	(0.127)**	0.331	(0.123)***	0.307	(0.127)**	0.331	(0.123)***
c_{t-15}	-0.285	(0.123)**	-0.323	(0.120)***	-0.285	(0.123)**	-0.323	(0.120)***
f_{t-15}	0.254	(0.127)**	0.296	(0.123)**	0.254	(0.127)**	0.296	(0.123)**
c_{t-16}	0.194	(0.124)	0.116	(0.120)	0.194	(0.124)	0.116	(0.120)
f_{t-16}	-0.212	(0.128)*	-0.144	(0.124)	-0.212	(0.128)*	-0.144	(0.124)

Standard errors are in parentheses.
***, **, and * denote significance level of 1%, 5% and 10%, respectively.

the higher the level of stored commodities the lower the value from storing an additional unit (Ziegelbäck and Kastner (2013)).

For corn, soybean and wheat markets, inventories were decreasing in COVID so that the "net cost" is lower during COVID. For the cotton market, the demand for cotton and cotton products, like apparel, decreased due to COVID-19 stay-at-home orders and economic recession, so that the inventories were increasing which lead to higher "net cost". For the nonstorable commodity, lean hogs, the "net cost" is -0.9876 before COVID and it is -0.7194 during COVID. As the demand for hogs decreased in COVID, the "net cost" is higher during pendamic period because of increasing inventories. One thing needs to be noticed is the unit of trading is 40000 pounds of lean value hog carcasses in the lean hog futures, so it could be stored to some degree. Existing literature argues that the storage activity links futures and cash prices through the cost of carry, but if inventories do not play a role, as

Table 3.11: Estimates of Soybean Market during COVID

	$w^i_{t-1}(\beta) \leq 0.006$				$w^i_{t-1}(\beta) > 0.006$			
	cash		futures		cash		futures	
w^i_{t-1}	-1.278	(0.464)***	-0.874	(0.478)*	-0.360	(0.175)**	-0.129	(0.180)
Const	0.004	(0.002)**	0.002	(0.002)	0.004	(0.002)**	0.002	(0.002)
c_{t-1}	-0.163	(0.215)	0.126	(0.222)	-0.163	(0.215)	0.126	(0.222)
f_{t-1}	0.425	(0.210)**	0.057	(0.217)	0.425	(0.210)**	0.057	(0.217)
c_{t-3}	0.334	(0.180)*	0.402	(0.185)**	0.334	(0.180)*	0.402	(0.185)**
f_{t-3}	-0.396	(0.187)**	-0.472	(0.192)**	-0.396	(0.187)**	-0.472	(0.192)**

Standard errors are in parentheses.
***, **, and * denote significance level of 1%, 5% and 10%, respectively.

Table 3.12: Estimates of Wheat Market before COVID

	$w^i_{t-1}(\beta) \leq 0.016$				$w^i_{t-1}(\beta) > 0.016$			
	cash		futures		cash		futures	
w^i_{t-1}	-0.444	(0.115)***	-0.201	(0.108)*	0.026	(0.018)	0.053	(0.017)***
Const	-0.002	(0.002)	-0.004	(0.001)***	-0.002	(0.002)	-0.004	(0.001)***
c_{t-1}	-0.123	(0.056)**	0.066	(0.053)	-0.123	(0.056)**	0.066	(0.053)
f_{t-1}	0.129	(0.060)**	-0.042	(0.056)	0.129	(0.060)**	-0.042	(0.056)

Standard errors are in parentheses.
***, **, and * denote significance level of 1%, 5% and 10%, respectively.

in the case of a nonstorable commodity, then the link between futures and cash prices is broken. But it is really an empirical question (Carter and Mohapatra (2008)). The result we have here is consistent with the work of Yang et al. (2001), who also find cointegration for nonstorable commodities. They claim that this suggests that both cash and futures prices for nonstorable commodities share certain long-run information and asset storability may not affect the existence of cointegration.

Table 3.13: Estimates of Wheat Market during COVID

	$w^i_{t-1}(\beta) \leq 0.003$				$w^i_{t-1}(\beta) > 0.003$			
	cash		futures		cash		futures	
w^i_{t-1}	-0.293	(0.069)***	-0.204	(0.070)***	2.055	(0.622)***	1.551	(0.631)**
Const	-0.019	(0.004)***	-0.013	(0.005)***	-0.019	(0.004)***	-0.013	(0.005)***
c_{t-1}	0.218	(0.176)	0.263	(0.179)	0.218	(0.176)	0.263	(0.179)
f_{t-1}	-0.255	(0.180)	-0.346	(0.183)*	-0.255	(0.180)	-0.346	(0.183)*
c_{t-16}	0.224	(0.175)	0.249	(0.178)	0.224	(0.175)	0.249	(0.178)
f_{t-16}	-0.342	(0.178)*	-0.335	(0.181)*	-0.342	(0.178)*	-0.335	(0.181)*
c_{t-17}	-0.364	(0.175)**	-0.417	(0.178)**	-0.364	(0.175)**	-0.417	(0.178)**
f_{t-17}	0.135	(0.178)	0.253	(0.181)	0.135	(0.178)	0.253	(0.181)
c_{t-19}	-0.282	(0.181)	-0.326	(0.184)*	-0.282	(0.181)	-0.326	(0.184)*
f_{t-19}	0.151	(0.183)	0.201	(0.186)	0.151	(0.183)	0.201	(0.186)

Standard errors are in parentheses.
***, **, and * denote significance level of 1%, 5% and 10%, respectively.

3.4.2 Threshold VECM

Table 3.7 presents the p value of F test for the threshold effect and the null hypothesis is that the model is a symmetric vector error correction model (VECM) and we use AIC, BIC and SSR to determine the optimal lags. From Table 3.7, we could see there exist significant threshold effect in all five markets. Table 3.8 to Table 3.17 present the results we obtain from the TVECM, which allows us to examine the cash and futures price responses to corn, soybean, wheat, cotton and lean hogs basis. In particular, it allows us to determine which price, the cash or futures price, moves to restore the long-run cointegrating equilibrium. For example, suppose that the coefficient on the ECT in the equation for the cash price is negative and statistically significant, while the coefficient on the ECT in the equation for the futures price is zero or statistically insignificant. This may indicate that the cash price reacts to the deviation to restore the long-run equilibrium in the agricultral commodity markets.

Table 3.14: Estimates of Cotton Market before COVID

	$w^i_{t-1}(\beta) \leq -0.014$				$w^i_{t-1}(\beta) > -0.014$			
	cash		futures		cash		futures	
w^i_{t-1}	0.095	(0.454)	1.127	(0.464)**	0.040	(0.024)*	0.059	(0.024)**
Const	0.003	(0.009)	0.017	(0.009)*	-0.002	(0.001)	-0.002	(0.001)**
c_{t-1}	-1.021	(0.909)	-2.712	(0.929)***	-0.264	(0.091)***	-0.151	(0.093)
f_{t-1}	0.669	(0.949)	2.346	(0.970)**	0.278	(0.089)***	0.169	(0.090)*

Standard errors are in parentheses.
***, **, and * denote significance level of 1%, 5% and 10%, respectively.

Table 3.15: Estimates of Cotton Market during COVID

	$w^i_{t-1}(\beta) \leq -0.007$				$w^i_{t-1}(\beta) > -0.007$			
	cash		futures		cash		futures	
w^i_{t-1}	-0.433	(0.175)**	-0.323	(0.170)*	0.631	(0.171)***	0.742	(0.166)***
Const	-0.004	(0.001)**	-0.004	(0.001)***	-0.004	(0.001)**	-0.004	(0.001)***
c_{t-1}	-0.289	(0.228)	-0.311	(0.222)	-0.289	(0.228)	-0.311	(0.222)
f_{t-1}	0.303	(0.231)	0.340	(0.225)	0.303	(0.231)	0.340	(0.225)

Standard errors are in parentheses.
***, **, and * denote significance level of 1%, 5% and 10%, respectively.

In general, the threshold values of all five markets are close to 0, thus we could regard the low regime as the case when the difference between cash and futures prices is below the negative "net cost", β_0, and the high regime as the case when the difference between cash and futures prices is above the negative "net cost". In other words, the low regime refers to the situation that the difference between futures and cash prices is higher than the "net cost" so that people would buy spot commodities and sell futures. This leads to the increase in cash prices and decrease in futures prices, so the difference between futures and cash prices becomes smaller. On the other hand, the high regime refers to the situation that the difference between futures and cash prices is lower than the "net cost" so that people would buy futures and sell spot commodities as it is expensive to store the commodities. This leads to the decrease in cash prices and increase in futures prices, so the difference

Table 3.16: Estimates of Lean Hogs Market before COVID

	$w^i_{t-1}(\beta) \leq -0.060$				$w^i_{t-1}(\beta) > -0.060$			
	cash		futures		cash		futures	
w^i_{t-1}	-0.012	(0.002)***	-0.023	(0.009)**	0.0003	(0.010)	0.179	(0.056)***
Const	-0.002	(0.000)***	-0.004	(0.002)**	-0.002	(0.000)***	-0.004	(0.002)**
c_{t-1}	0.840	(0.015)***	0.170	(0.083)**	0.840	(0.015)***	0.170	(0.083)**
f_{t-1}	0.018	(0.006)***	-0.0002	(0.033)	0.018	(0.006)***	-0.0002	(0.033)

Standard errors are in parentheses.
***, **, and * denote significance level of 1%, 5% and 10%, respectively.

Table 3.17: Estimates of Lean Hogs Market during COVID

	$w^i_{t-1}(\beta) \leq 0.067$				$w^i_{t-1}(\beta) > 0.067$			
	cash		futures		cash		futures	
w^i_{t-1}	-0.040	(0.037)	0.315	(0.128)**	-0.017	(0.007)**	0.011	(0.025)
Const	0.002	(0.001)*	-0.002	(0.004)	0.002	(0.001)*	-0.002	(0.004)
c_{t-1}	0.750	(0.046)***	0.228	(0.160)	0.750	(0.046)***	0.228	(0.160)
f_{t-1}	0.037	(0.021)*	0.167	(0.072)**	0.037	(0.021)*	0.167	(0.072)**

Standard errors are in parentheses.
***, **, and * denote significance level of 1%, 5% and 10%, respectively.

between cash and futures prices becomes smaller. And this is the arbitrage activity that keeps the cash and futures prices in line.

In the corn market, we find that before COVID, both cash price and futures price react to restore the long-run equilibrium. In the low regime, where the cash price is lower than the futures price, the futures price changes faster as $0.057 > 0.040$. The same case happens when it is in the high regime as $0.203 > 0.178$. Both cases will restore the equilibrium and we do observe the asymmetry of adjustment as $0.040 < 0.178$ and $0.057 < 0.203$. While during COVID, cash price reacts to restore the equilibrium when the basis is small, that is, when the cash price is relatively lower than the futures price. When the basis is large, the futures price moves to restore the equilibrium.

We observe the cash price has a feedback effect on the futures price before COVID and

bidirectional information flow between cash and futures prices during COVID. And the result also shows the asymmetric adjustment in low and high regimes. As we discussed in section 3.4.1, the "net cost" is higher before COVID because the inventory level is higher according to USDA data. So the "net cost", represented by β_0, used to define the low and high regimes, is lower during COVID.

Before COVID, when the basis is small in the soybean market, that is, the cash price is relatively less than the futures price, both cash price and futures price react to the disequilibrium with different speed of the adjustment. Based on the coefficients of ECT, $|-0.577| > |-0.400|$, the cash price changes more than the futures price when the basis is small. When the basis is large, the futures price behaves significantly in bringing back the equilibrium. During COVID, both cash price and futures price move to restore the equilibrium at higher speed compared to it before COVID. Again, the cash price changes more, $|-1.278| > |-0.874|$, when the soybean market is in contango, that is, the cash price is relatively lower than the futures price. When the soybean market is in backwardation, that is, the cash price is relatively higher than the futures price, cash price reacts to the deviation more significantly. We confirm the presence of a significant effect from futures to futures price as well as from cash to futures price before COVID. And we observe a short-run effect from futures to cash price and bidirectional relationships between cash the futures prices both before and during COVID. Again, the result shows the asymmetric adjustment in low and high regimes. Similar to the corn market, the "net cost" we used to define the low and high regimes is lower during COVID.

In the wheat market, both cash and futures prices move to re-establish the equilibrium and cash price adjusts faster based on the coefficients $|-0.444| > |-0.201|$ when the wheat market is in contango before COVID and only the futures price moves to restore the equilibrium when the market is in backwardation. The cash price of wheat is affected by its

own lagged changes and there is a short-run effect from futures to cash price before COVID. During COVID, both cash and futures prices react to the disequilibirum when the market is in contango and the cash price changes more as $|-0.293| > |-0.204|$. When the market is in backwardation, we need to look at the coefficients of ECT combining with the constants as these two work together to keep the cash and futures prices staying in line. The futures price is affected by its own lagged changes and by the cash price during COVID. And we also observe the bidirectional information flow between cash and futures prices during COVID as well as the asymmetry in price adjustment in low and high regimes. Just like what we have found for corn and soybean market, the "net cost" we used to specify the low and high regimes for wheat market is lower during COVID.

In the cotton market, both cash and futures prices react to the deviation when the market is in backwardation and the futures price moves more significantly to restore the equilibrium when the market is in contango before COVID. And the existence of short-run effect from futures to cash price and the feedback efffect of the cash price on the futures price are confirmed before COVID. During COVID, both cash price and futures price react to the deviation of equilibrium and the cash price adjusts more when the cotton market is in contango based on the coefficients $|-0.433| > |-0.323|$. The result is in line with the existence of asymmetric adjustment in low and high regimes. For the cotton market, as the demand for cotton decreased during COVID and the inventories were increasing, we have a higher "net cost" to separate the low and high regimes.

And in the lean hogs market, we need to take both the coefficients of ECT and the constants into account when we investigate the adjustment path back to equilibrium before COVID. While during COVID, the cash price reacts to restore the equilibrium when the basis is large, and the futures price reacts when the basis is small. We also find the cash price of lean hogs is affected by its own lagged changes both before and during COVID. As

the demand for hogs decreased in COVID, the “net cost” we use to describe low and high regimes is higher during COVID.

3.4.3 Impulse Response Analysis

Figure 3.6 to Figure 3.15 illustrate the adjustment patterns of futures and cash prices responding to the positive and negative shocks from cash and futures prices in corn, soybean, wheat, cotton and lean hogs markets. We choose two-standard deviation as the shock and add it to the last observation. In general, the responses are consistent with long-run cointegration in all five markets because the permanent shocks on cash and futures prices present comovements.

In the corn market, although the adjustment takes longer to settle down before COVID, the size of permanent shock is larger during COVID. And we could notice the positive shock from the cash price during COVID needs even less adjustment time which is consistent with the estimates we obtain for the corn market. The reason for faster adjustment during COVID could be that arbitrage happens when abnormalities occur in the price pattern. As people notice the different price pattern during COVID compared to the period before COVID, more people would like to take advantage of this abnormalities, thus the adjustment speed is higher. For the larger permanent shock during COVID, especially in the futures market, the reason may be that futures market could overreact to new information because of the lower transaction costs.

In the soybean market, the adjustment time is longer before COVID and the larger permanent shock happens in the futures market both before and after COVID. As corn and soybean futures are both high liquid commodity futures, the reason for the less time needed for adjustment in soybean market during COVID could also be that more arbitrage activities

happen when abnormalities in the price pattern occur, which leads to a faster adjustment. And larger size of permanent shock also results from the overreaction to information in the futures market.

In the wheat market, the COVID significantly increases the adjustment time for both positive and negative shocks. The reason could be that the COVID-19 pandemic lockdown policy and travel restriction to mitigate the spread of the virus disrupted the supply chain. Similiar to the soybean market, the shocks from futures prices have larger impacts on the cash and futures prices both before and during COVID and the reason could be the overreaction to information induced by the lower transaction cost in futures markets.

In the cotton market, the shocks from futures prices have a larger impact on the cash and futures prices compared to the cases before COVID. And we also notice the shorter adjustment time to shocks during COVID, which could be caused by increasing number of arbitragers who are attracted by the abnormal price pattern in COVID. Another thing is that the positive shock from the futures price has no permanent impact on both cash and futures prices before COVID, but it does have permanent influence on the prices during COVID.

In the lean hogs market, the shocks from cash prices have larger impacts on futures and cash prices both before and during COVID compared to the storable commodities. The positive shocks from cash prices have smaller impacts on futures and cash prices, while the negative shocks from cash prices have larger impacts on the futures and cash prices, and it could be caused by the decrease in demand. During the pre-COVID periods, the U.S. consumer spent $137.4 billion per month on food relative to $105 billion food spending in the first quarter of 2020. Total food spending in the U.S. rose in the second and third quarter of 2020 but was relatively lower compared to the same period in 2019 (Ahmed and Sarkodie (2021)). Also, the shocks from futures prices have larger impacts on cash and futures prices

during COVID.

3.5 Conclusion

To conclude, we apply a cointegration model and a Threshold Vector Error Correction Model (TVECM) using the basis (i.e., the the error correction term) as the forcing variable to examine the relationship between cash and futures prices. This study explores the long run equilibrium and the price transmission dynamics between the futures and cash prices for both storable and nonstorable commodities, i.e., corn, soybean, wheat, cotton and lean hogs markets. This paper confirms a long-run relationship and a short-run effect of the futures prices on the spot prices and provides the evidence supporting the asymmetric price adjustment behaviour for the five commodity markets. We also compare the results of pre-COVID and the COVID periods in order to see the impacts of COVID on the relationships of the five markets.

From the results, we find that when the deviation from the long run equilibrium of the financial-commodity market lies in different states, which is described by the threshold effect, it shows different adjustment patterns. And the two regimes can be described as the case when the difference between futures and cash prices is higher or lower than the "net cost". The asymmetry of the adjustment process is estimated with the threshold vector error correction model, which assesses the dynamic response in the cash and futures prices to the basis. According to the cointegration test and the estimates of the TVECM, we can conclude that the conintegration relationships of cash and futures prices in wheat and corn markets are more vulnerable to the impact of the COVID. And the asymmetry of adjustment back to equilibrium captured by the coefficients of ECT. In particular, the cash price and the futures price of soybeans adjust faster to restore the equilibrium during COVID and that

could be the reason why the price volatility in the soybean market are even lower compared to it before the COVID.

The impulse response results also support the cointegration relationships both in pre-COVID and COVID periods, and we observe some changes of adjustment patterns such as longer adjustment time and different size of impacts caused by the same size of shocks. The reasons could be that all countries around the world took strict isolation measures in order to control the spread of COVID-19, which affected the global grain supply system; more arbitrage activities occurs as people regard the price pattern during COVID as an abnormal pattern so that they would like to profit from it; and the changes of adjustment patterns could also be caused by the decrease in inventories and the decrease in the demand for foods during 2020. According to USDA, the consumer's expenditure on food was decreasing in 2020 because of COVID.

REFERENCES

Aas, K., Czado, C., Frigessi, A., and Bakken, H. (2009). Pair-copula constructions of multiple dependence. *Insurance: Mathematics and economics*, 44(2):182–198.

Adämmer, P. and Bohl, M. T. (2018). Price discovery dynamics in european agricultural markets. *Journal of Futures Markets*, 38(5):549–562.

Adewopo, J. B., Solano-Hermosilla, G., Colen, L., and Micale, F. (2021). Using crowd-sourced data for real-time monitoring of food prices during the covid-19 pandemic: insights from a pilot project in northern nigeria. *Global Food Security*, 29:100523.

Ahmed, M. Y. and Sarkodie, S. A. (2021). Covid-19 pandemic and economic policy uncertainty regimes affect commodity market volatility. *Resources policy*, 74:102303.

Balke, N. S., Fomby, T. B., et al. (1997). Threshold cointegration. *International Economic Review*, 38(3):627–645.

Bekiros, S. D. and Diks, C. G. (2008). The relationship between crude oil spot and futures prices: Cointegration, linear and nonlinear causality. *Energy Economics*, 30(5):2673–2685.

Borenstein, S., Cameron, A. C., and Gilbert, R. (1997). Do gasoline prices respond asymmetrically to crude oil price changes? *The Quarterly journal of economics*, 112(1):305–339.

Carter, C. A. and Mohapatra, S. (2008). How reliable are hog futures as forecasts? *American Journal of Agricultural Economics*, 90(2):367–378.

Chang, C.-L., Chen, L.-H., Hammoudeh, S., and McAleer, M. (2012). Asymmetric adjustments in the ethanol and grains markets. *Energy economics*, 34(6):1990–2002.

Covey, T. and Bessler, D. A. (1995). Asset storability and the informational content of intertemporal prices. *Journal of Empirical Finance*, 2(2):103–115.

Cramon-Taubadel, S. V. (1998). Estimating asymmetric price transmission with the error correction representation: An application to the german pork market. *European review of agricultural economics*, 25(1):1–18.

Dissmann, J., Brechmann, E. C., Czado, C., and Kurowicka, D. (2013). Selecting and estimating regular vine copulae and application to financial returns. *Computational Statistics & Data Analysis*, 59:52–69.

Engle, R. F. and Granger, C. W. (1987). Co-integration and error correction: Representation, estimation, and testing. *Econometrica*, 55(2):251–76.

Engle, R. F. and Kroner, K. F. (1995). Multivariate simultaneous generalized arch. *Econometric theory*, 11(1):122–150.

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

Fama, E. F. and French, K. R. (1987). Commodity futures prices: Some evidence on forecast power, premiums, and the theory of storage. *Journal of Business*, 60:55–73.

Garbade, K. and Silber, W. L. (1983). Price movements and price discovery in futures and cash markets. *The Review of Economics and Statistics*, 65(2):289–97.

Gardner, B. L. (1976). Futures prices in supply analysis. *American Journal of Agricultural Economics*, 58(1):81–84.

Goodwin, B. K., Holt, M. T., Önel, G., and Prestemon, J. P. (2018). Copula-based nonlinear modeling of the law of one price for lumber products. *Empirical Economics*, 54(3):1237–1265.

Goodwin, B. K., Holt, M. T., and Prestemon, J. P. (2011). North american oriented strand board markets, arbitrage activity, and market price dynamics: A smooth transition approach. *American Journal of Agricultural Economics*, 93(4):993–1014.

Goodwin, B. K., Holt, M. T., and Prestemon, J. P. (2021). Semi-parametric models of spatial market integration. *Empirical Economics*, 61(5):2335–2361.

Goodwin, B. K. and Piggott, N. E. (2001). Spatial market integration in the presence of threshold effects. *American Journal of Agricultural Economics*, 83(2):302–317.

Guidoum, A. C. and Boukhetala, K. (2016). Estimation of stochastic differential equations with sim.diffproc package version 3.3.

Hansen, B. E. and Seo, B. (2002). Testing for two-regime threshold cointegration in vector error-correction models. *Journal of econometrics*, 110(2):293–318.

Iacus, S. M. (2008). *Simulation and inference for stochastic differential equations: with R examples*, volume 486. Springer.

Ihle, R. and von Cramon-Taubadel, S. (2008). A comparison of threshold cointegration and markov-switching vector error correction models in price transmission analysis.

Jacobsen, M. (1996). Laplace and the origin of the ornstein-uhlenbeck process. *Bernoulli*, 2(3):271–286.

Joe, H. (1997). *Multivariate models and multivariate dependence concepts*. CRC press.

Kamara, A. (1982). Issues in futures markets: a survey. *The Journal of Futures Markets (pre-1986)*, 2(3):261.

Koop, G., Pesaran, M. H., and Potter, S. M. (1996). Impulse response analysis in nonlinear multivariate models. *Journal of econometrics*, 74(1):119–147.

Krupskii, P. and Joe, H. (2013). Factor copula models for multivariate data. *Journal of Multivariate Analysis*, 120:85–101.

Kuiper, W. E., Pennings, J. M., and Meulenberg, M. T. (2002). Identification by full adjustment: evidence from the relationship between futures and spot prices. *European Review of Agricultural Economics*, 29(1):67–84.

Leuthold, R. M. (1979). An analysis of the futures-cash price basis for live beef cattle. *North Central Journal of Agricultural Economics*, Jan 1:47–52.

Liebscher, E. (2008). Construction of asymmetric multivariate copulas. *Journal of Multivariate analysis*, 99(10):2234–2250.

Mazo, G., Girard, S., and Forbes, F. (2015). A class of multivariate copulas based on products of bivariate copulas. *Journal of Multivariate Analysis*, 140:363–376.

Mazo, G., Girard, S., and Forbes, F. (2016). A flexible and tractable class of one-factor copulas. *Statistics and Computing*, 26(5):965–979.

Meyer, J. and von Cramon-Taubadel, S. (2004). Asymmetric price transmission: a survey. *Journal of agricultural economics*, 55(3):581–611.

Nelsen, R. B. (2006). *An Introduction to Copulas*. Springer, New York.

Oellermann, C. M. and Farris, P. L. (1985). Futures or cash: which market leads live beef cattle prices? *The Journal of Futures Markets (pre-1986)*, 5(4):529.

Oellermann, C. M., Wade, B. B., and Farris, P. L. (1989). Price discovery for feeder cattle. *The Journal of Futures Markets (1986-1998)*, 9(2):113.

Patel, S. and Tkac, P. (2007). The past, present, and future of futures. *EconSouth*, 9(3).

Pindyck, R. S. (2004). Volatility and commodity price dynamics. *Journal of Futures Markets: Futures, Options, and Other Derivative Products*, 24(11):1029–1047.

Purcell, W. D. and Hudson, M. A. (1985). The economic roles and implications of trade in livestock futures. *Futures markets: Regulatory issues*, Mar:329–376.

Randall, F. T. and Zapata, H. O. (1993). An examination of cointegration relations between futures and local grain markets. *The Journal of Futures Markets (1986-1998)*, 13(8):921.

Sawyer, K. (1993). 27 continuous time financial models: Statistical applications of stochastic processes. In *Econometrics*, volume 11 of *Handbook of Statistics*, pages 739–764. Elsevier.

Schmidt, R. and Stadtmüller, U. (2006). Non-parametric estimation of tail dependence. *Scandinavian journal of statistics*, 33(2):307–335.

Silvapulle, P. and Moosa, I. A. (1999). The relationship between spot and futures prices: evidence from the crude oil market. *Journal of Futures Markets: Futures, Options, and Other Derivative Products*, 19(2):175–193.

Simioni, M., Gonzales, F., Guillotreau, P., and Le Grel, L. (2013). Detecting asymmetric price transmission with consistent threshold along the fish supply chain. *Canadian Journal of Agricultural Economics/Revue canadienne d'agroeconomie*, 61(1):37–60.

Sklar, A. (1959). Distribution functions in n dimensions and their margins. *Statistics Publications, University of Paris*, 8:229–231.

Tsay, R. S. (2013). *Multivariate time series analysis: with R and financial applications*. John Wiley & Sons.

Uhlenbeck, G. E. and Ornstein, L. S. (1930). On the theory of the brownian motion. *Physical review*, 36(5):823.

von Cramon-Taubadel, S. and Goodwin, B. K. (2021). Price transmission in agricultural markets. *Annual Review of Resource Economics*, 13:65–84.

Wiener, N. (1923). Differential-space. *Journal of Mathematics and Physics*, 2(1-4):131–174.

Wu, Z., Maynard, A., Weersink, A., and Hailu, G. (2018). Asymmetric spot-futures price adjustments in grain markets. *Journal of Futures Markets*, 38(12):1549–1564.

Yang, J., Bessler, D. A., and Leatham, D. J. (2001). Asset storability and price discovery in commodity futures markets: a new look. *Journal of Futures Markets: Futures, Options, and Other Derivative Products*, 21(3):279–300.

Zeng, H. Y. and Lu, R. (2022). Vix and major agricultural future markets: Dynamic linkage and time-frequency relations around the covid-19 outbreak. *SSRN Electronic Journal*.

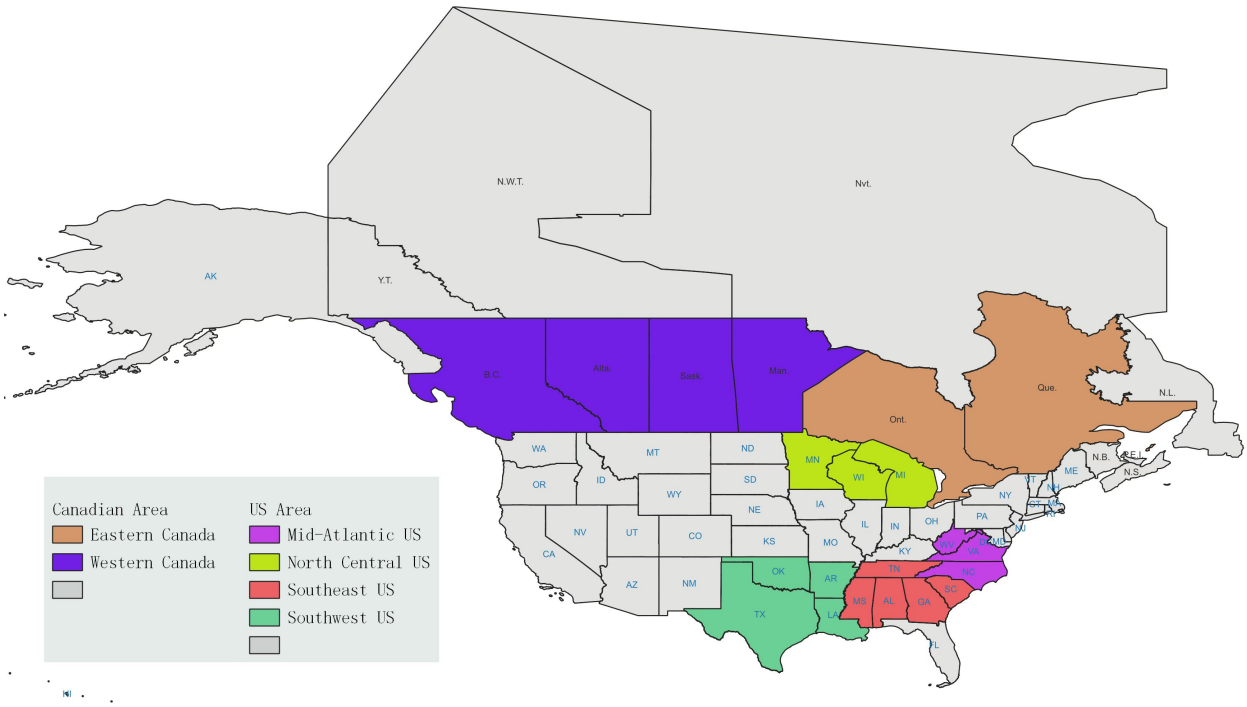
Ziegelbäck, M. and Kastner, G. (2013). Arbitrage hedging in markets for the us lean hogs and the eu live pigs. *Agricultural Economics (AGRICECON)*, 59(11):505–511.

APPENDIX

APPENDIX

A

MAP OF OSB MARKET ZONES



Canadian Area
- Eastern Canada
- Western Canada

US Area
- Mid-Atlantic US
- North Central US
- Southeast US
- Southwest US

Figure A.1: Map of OSB Markets in US and Canada