A Vine Copula Panel Model for Day-Ahead Electricity Prices

Vienna Conference on Mathematical Finance 2019

September 9, 2019

Janus S. Valberg-Madsen, Esben Høg,
Troels S. Christensen, Anca Pircalabu

AALBORG UNIVERSITY
DENMARK

Agenda

Background
Modelling Day-Ahead Prices

Background

Modelling Day-Ahead Prices

- Every day, hourly prices for the following day are set based on trading activity

Background
Modelling Day-Ahead Prices

- Every day, hourly prices for the following day are set based on trading activity
- Prices are determined from the same information set $\implies$ panel data

Background
Modelling Day-Ahead Prices

- Every day, hourly prices for the following day are set based on trading activity
- Prices are determined from the same information set $\implies$ panel data
- Common to consider base prices, i.e. daily means

Our Goals

Our Goals

- Model the day-ahead prices of each hour individually, as daily time series

Our Goals

- Model the day-ahead prices of each hour individually, as daily time series
- Capture their joint behaviour with a copula-based model

Our Goals

- Model the day-ahead prices of each hour individually, as daily time series
- Capture their joint behaviour with a copula-based model
- Simulate the day-ahead price distribution from the full model

Recap on Copulas
Definition

Definition (Copula)

A *d-dimensional copula* is a function $C : [0,1]^d \to [0,1]$ such that

1. $C(u_1, \dots, u_d) = 0$ if $u_j = 0$ for at least one $j \in \{1, \dots, d\}$
2. $C\big(1, \dots, 1, u_j, 1, \dots, 1\big) = u_j$ for all $j \in \{1, \dots, d\}$
3. $V_C\big((\boldsymbol{a}, \boldsymbol{b}]\big) \geq 0$ for every d-box $(\boldsymbol{a}, \boldsymbol{b}] \subseteq [0,1]^d$

where $(\boldsymbol{a}, \boldsymbol{b}] \coloneqq (a_1, b_1] \times \cdots (a_d, b_d]$ and

$$V_C\big((\boldsymbol{a}, \boldsymbol{b}]\big) \coloneqq \sum_{\boldsymbol{v} \in \mathrm{ver}(\boldsymbol{a}, \boldsymbol{b}]} \mathrm{sign}(\boldsymbol{v})\, C(\boldsymbol{v})\,;$$

$$\mathrm{ver}\,(\boldsymbol{a}, \boldsymbol{b}] \coloneqq \{a_1, b_1\} \times \cdots \times \{a_d, b_d\}\,,$$

$$\mathrm{sign}(\boldsymbol{v}) \coloneqq \begin{cases} 1, & \text{if } v_j = a_j \text{ for an even number of indices,} \\ -1, & \text{otherwise.} \end{cases}$$

Recap on Copulas

Sklar's Theorem

Theorem (Sklar's Theorem)

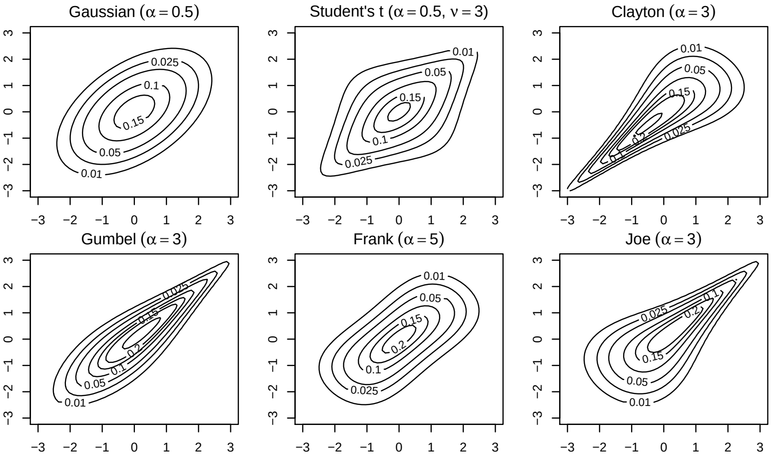
Let $\boldsymbol{X}$ be a d-dimensional random vector with joint distribution function F and marginals $F_1, \ldots, F_d$. Then there exists a d-copula C such that

$$F(\boldsymbol{x}) = C\big(F_1(x_1), \ldots, F_d(x_d)\big), \quad \forall \boldsymbol{x} \in \mathbb{R}^d$$

and if the marginals are all continous, then C is uniquely defined.

Recap on Copulas

Families of Copulas ($d = 2$)



Recap on Copulas
Basic Idea Behind Vine Copulas

Recap on Copulas

Basic Idea Behind Vine Copulas

- Factor full joint density into product of marginal densities and copula densities

Recap on Copulas

Basic Idea Behind Vine Copulas

- Factor full joint density into product of marginal densities and copula densities
- Represent with sequence of trees, such that
  - Marginals are nodes in the first tree
  - Pair-copulas between the marginals are edges in the first tree
  - Edges from the previous tree becomes nodes in the next
  - Edges in subsequent trees are conditional pair-copulas
  - Nodes in a tree can be joined by an edge, if the nodes share a node from the previous tree

Recap on Copulas

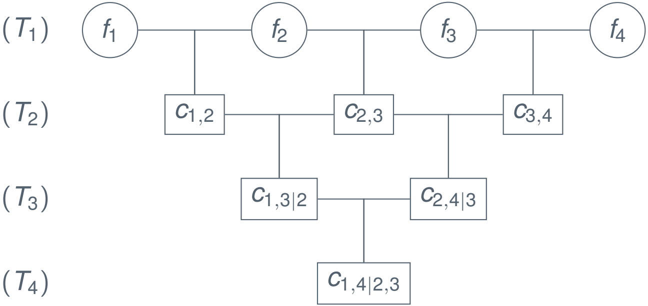
Basic Idea Behind Vine Copulas

- Factor full joint density into product of marginal densities and copula densities
- Represent with sequence of trees, such that
  - Marginals are nodes in the first tree
  - Pair-copulas between the marginals are edges in the first tree
  - Edges from the previous tree becomes nodes in the next
  - Edges in subsequent trees are conditional pair-copulas
  - Nodes in a tree can be joined by an edge, if the nodes share a node from the previous tree
- The chosen pair-copulas need not be the same

Recap on Copulas

Basic Idea Behind Vine Copulas: Example

Example in $d = 4$



Recap on Copulas

Estimating Vine Copulas

Sequential estimation

Maximum likelihood estimation

Recap on Copulas
Estimating Vine Copulas

Sequential estimation

- Estimate tree structure with maximum spanning tree algorithm, using e.g. Kendall's τ as edge weights

Maximum likelihood estimation

Recap on Copulas
Estimating Vine Copulas

Sequential estimation

- Estimate tree structure with maximum spanning tree algorithm, using e.g. Kendall's τ as edge weights
- For each edge, estimate a range of copulas with MLE, select one with e.g. BIC

Maximum likelihood estimation

Recap on Copulas

Estimating Vine Copulas

Sequential estimation

- Estimate tree structure with maximum spanning tree algorithm, using e.g. Kendall's τ as edge weights
- For each edge, estimate a range of copulas with MLE, select one with e.g. BIC
- Use estimated copulas to calculate pseudo-observations for the next tree

Maximum likelihood estimation

Recap on Copulas
Estimating Vine Copulas

Sequential estimation

- Estimate tree structure with maximum spanning tree algorithm, using e.g. Kendall's τ as edge weights
- For each edge, estimate a range of copulas with MLE, select one with e.g. BIC
- Use estimated copulas to calculate pseudo-observations for the next tree

Maximum likelihood estimation

- Full MLE using numerical optimisation methods

Recap on Copulas

Estimating Vine Copulas

Sequential estimation

- Estimate tree structure with maximum spanning tree algorithm, using e.g. Kendall's τ as edge weights
- For each edge, estimate a range of copulas with MLE, select one with e.g. BIC
- Use estimated copulas to calculate pseudo-observations for the next tree

Maximum likelihood estimation

- Full MLE using numerical optimisation methods
- Can use sequential estimation results as starting values

Data

German Day-Ahead Electricity Prices

Data
German Day-Ahead Electricity Prices

- Estimation data range: 2015-01-06 – 2018-12-31

Data
German Day-Ahead Electricity Prices

- Estimation data range: 2015-01-06 – 2018-12-31
- Downloaded from the ENTSO-E Transparency Platform
  - 2015-01-06 – 2018-09-30: DE-AT-LU bidding zone prices
  - 2018-10-01 – 2018-12-31: DE-LU bidding zone prices

Data

German Day-Ahead Electricity Prices

- Estimation data range: 2015-01-06 – 2018-12-31
- Downloaded from the ENTSO-E Transparency Platform
  - 2015-01-06 – 2018-09-30: DE-AT-LU bidding zone prices
  - 2018-10-01 – 2018-12-31: DE-LU bidding zone prices
- Missing Hour 3 observations (due to summer time) interpolated from hours 2 and 4

Data

German Day-Ahead Electricity Prices

- Estimation data range: 2015-01-06 – 2018-12-31
- Downloaded from the ENTSO-E Transparency Platform
  - 2015-01-06 – 2018-09-30: DE-AT-LU bidding zone prices
  - 2018-10-01 – 2018-12-31: DE-LU bidding zone prices
- Missing Hour 3 observations (due to summer time) interpolated from hours 2 and 4
- Duplicate Hour 3 observations (due to winter time) averaged

Data

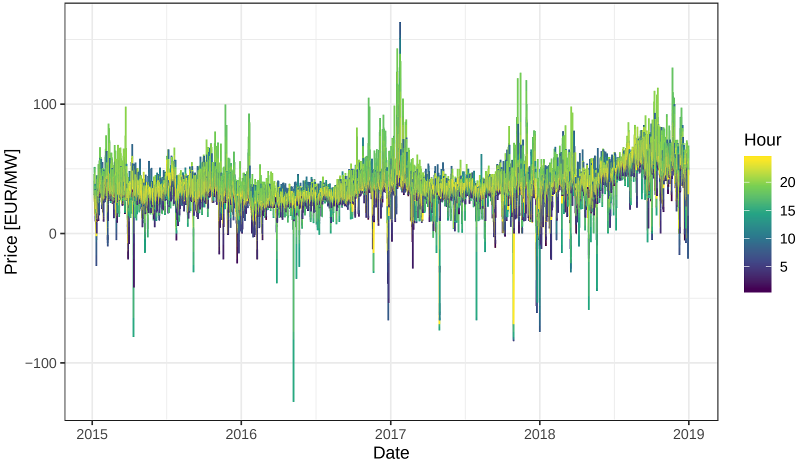
German Day-Ahead Electricity Prices

- Estimation data range: 2015-01-06 – 2018-12-31
- Downloaded from the ENTSO-E Transparency Platform
  - 2015-01-06 – 2018-09-30: DE-AT-LU bidding zone prices
  - 2018-10-01 – 2018-12-31: DE-LU bidding zone prices
- Missing Hour 3 observations (due to summer time) interpolated from hours 2 and 4
- Duplicate Hour 3 observations (due to winter time) averaged

$\Longrightarrow$ 24 daily time series of 1668 observations

Data

German Day-Ahead Electricity Prices



Marginal Models
Overview

With $t \in \{1, \ldots, 1668\}$ and $h \in \{1, \ldots, 24\}$,

$$p_{t,h} := \log\big(P_{t,h} + K\big) = s_{t,h} + X_{t,h},$$

where

- $K = 500$ (chosen as such from bid restrictions)
- $s_{t,h}$ is a deterministic component for capturing season
- $X_{t,h}$ is a stochastic component for capturing serial behaviour

Marginal Models
Seasonal Component

Seasonal model specification

$$s_{t,h} = \alpha_{0,h} + \alpha_{1,h} \cdot t + \sum_{\phi \in \Phi} \left(\beta_{1,\phi,h} \sin(2\pi t \phi) + \beta_{2,\phi,h} \cos(2\pi t \phi) \right) + \gamma_{w,h} w_t,$$

where

- $\Phi = \{1/365, 2/365\}$ is a set of frequencies
- w_t is a factor variable for weekdays

Marginal Models
Seasonal Component

Seasonal model specification

$$s_{t,h} = \alpha_{0,h} + \alpha_{1,h} \cdot t + \sum_{\phi \in \Phi} \left(\beta_{1,\phi,h} \sin(2\pi t \phi) + \beta_{2,\phi,h} \cos(2\pi t \phi) \right) + \gamma_{w,h} w_t,$$

where

- $\Phi = \{1/365, 2/365\}$ is a set of frequencies
- w_t is a factor variable for weekdays

Estimated with a linear model, residuals used for serial model

Marginal Models
Serial Component

Serial model specification

Each model is chosen among a panel of ARMA-GARCH models:

- ARMA orders: $p, q \in \{0, 1\}$
- GARCH orders: $p, q \in \{0, 1\}$ such that at least one is nonzero
- GARCH types: GJR-GARCH and EGARCH
- Conditional distributions: Student's t-distribution and its skewed variant

Marginal Models
Serial Component

Serial model specification

Each model is chosen among a panel of ARMA-GARCH models:

- ARMA orders: $p, q \in \{0, 1\}$
- GARCH orders: $p, q \in \{0, 1\}$ such that at least one is nonzero
- GARCH types: GJR-GARCH and EGARCH
- Conditional distributions: Student's t-distribution and its skewed variant

$\implies$ 48 different models for each series;
estimated with MLE, model chosen with BIC

Marginal Models
Serial Component

Hour	Model
1	ARMA(1,1)-GJR-GARCH(1,1)-st
2	ARMA(1,1)-GJR-GARCH(1,1)-st
3	ARMA(1,1)-E-GARCH(1,1)-st
4	ARMA(1,1)-GJR-GARCH(1,1)-st
5	ARMA(1,1)-E-GARCH(1,1)-st
6	ARMA(1,1)-GJR-GARCH(1,1)-st
7	ARMA(1,1)-E-GARCH(1,1)-st
8	ARMA(1,1)-E-GARCH(1,1)-st
9	ARMA(1,1)-E-GARCH(1,1)-st
10	ARMA(1,1)-E-GARCH(1,1)-st
11	ARMA(1,1)-E-GARCH(1,1)-st
12	ARMA(1,1)-E-GARCH(1,1)-st

Hour	Model
13	ARMA(1,1)-GJR-GARCH(1,0)-st
14	ARMA(1,1)-GJR-GARCH(1,0)-st
15	AR(1)-GJR-GARCH(1,0)-st
16	ARMA(1,1)-GJR-GARCH(1,0)-st
17	ARMA(1,1)-GJR-GARCH(1,1)-st
18	ARMA(1,1)-E-GARCH(1,1)-st
19	ARMA(1,1)-GJR-GARCH(1,1)-t
20	ARMA(1,1)-GJR-GARCH(1,1)-st
21	ARMA(1,1)-GJR-GARCH(1,1)-st
22	ARMA(1,1)-GJR-GARCH(1,1)-st
23	ARMA(1,1)-E-GARCH(1,1)-st
24	ARMA(1,1)-E-GARCH(1,1)-st

Joint Model
Estimation of Vine Copula

Joint Model

Estimation of Vine Copula

1. Transform the standardised residuals from each marginal model with its distribution function, i.e. *t*- and skew-*t*-distributions with estimated shape and skew parameters

Joint Model
Estimation of Vine Copula

1. Transform the standardised residuals from each marginal model with its distribution function, i.e. *t*- and skew-*t*-distributions with estimated shape and skew parameters
2. Sequentially fit vine copula, where

Joint Model

Estimation of Vine Copula

1. Transform the standardised residuals from each marginal model with its distribution function, i.e. *t*- and skew-*t*-distributions with estimated shape and skew parameters
2. Sequentially fit vine copula, where
   - Edge weights used for MST algorithm are empirical Kendall's τ's

Joint Model
Estimation of Vine Copula

1. Transform the standardised residuals from each marginal model with its distribution function, i.e. *t*- and skew-*t*-distributions with estimated shape and skew parameters
2. Sequentially fit vine copula, where
   - Edge weights used for MST algorithm are empirical Kendall's τ's
   - Each edge is tested for independence

Joint Model
Estimation of Vine Copula

1. Transform the standardised residuals from each marginal model with its distribution function, i.e. *t*- and skew-*t*-distributions with estimated shape and skew parameters
2. Sequentially fit vine copula, where
   - Edge weights used for MST algorithm are empirical Kendall's τ's
   - Each edge is tested for independence
   - Pair-copulas considered are: Gaussian, Student's *t*, Clayton, Gumbel, Frank, and Joe, and their rotations

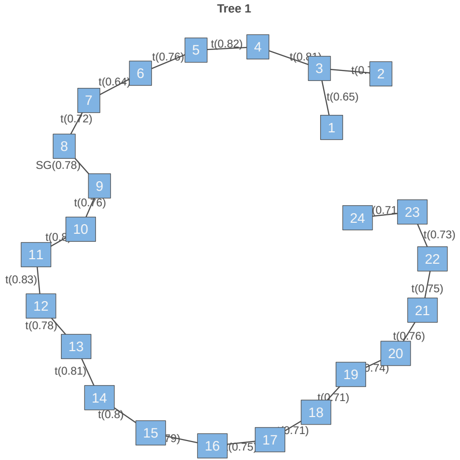
Joint Model
Estimation of Vine Copula

1. Transform the standardised residuals from each marginal model with its distribution function, i.e. *t*- and skew-*t*-distributions with estimated shape and skew parameters
2. Sequentially fit vine copula, where
   - Edge weights used for MST algorithm are empirical Kendall's τ's
   - Each edge is tested for independence
   - Pair-copulas considered are: Gaussian, Student's *t*, Clayton, Gumbel, Frank, and Joe, and their rotations
   - Pair-copulas are chosen with BIC

Joint Model
Estimation of Vine Copula

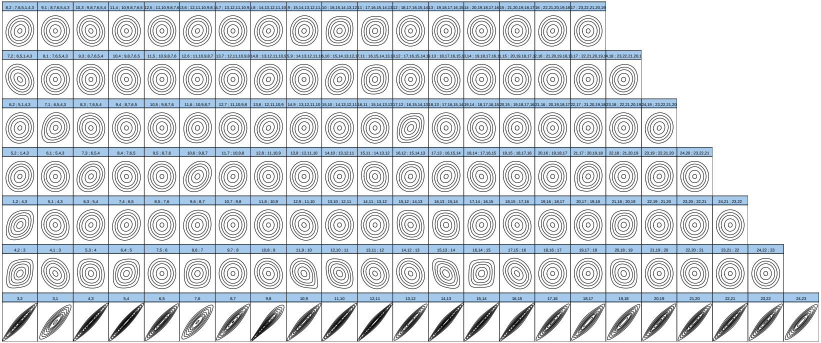
1. Transform the standardised residuals from each marginal model with its distribution function, i.e. *t*- and skew-*t*-distributions with estimated shape and skew parameters
2. Sequentially fit vine copula, where
   - Edge weights used for MST algorithm are empirical Kendall's τ's
   - Each edge is tested for independence
   - Pair-copulas considered are: Gaussian, Student's *t*, Clayton, Gumbel, Frank, and Joe, and their rotations
   - Pair-copulas are chosen with BIC
3. Reestimate with MLE

Joint Model
Summary of Results



Joint Model

Summary of Results



Joint Model
Summary of Results

Joint Model
Summary of Results

- Very close to D-vine structure

Joint Model
Summary of Results

- Very close to D-vine structure
- Many (but but not all) *t*-copulas in first tree

Joint Model
Summary of Results

- Very close to D-vine structure
- Many (but but not all) t-copulas in first tree
- Strong, positive dependence in first tree, including tail dependence

Joint Model
Summary of Results

- Very close to D-vine structure
- Many (but but not all) *t*-copulas in first tree
- Strong, positive dependence in first tree, including tail dependence
- A lot of independence or "near-independence" in later trees, in particular tree 19, 21, 22, 23, and 24
  - For tree 5 and later, over half of the edges are independence copulas
  - In total, about 53% of the copulas are independence copulas
  - On the other hand, dependence persists until tree 20

Simulation of Price Distribution

Prediction Framework

Simulation of Price Distribution

Prediction Framework

1. Draw N samples, $u_{n,h}$ from the vine copula model, which will be an $N \times 24$ matrix, uniformly distributed (column-wise)

Simulation of Price Distribution

Prediction Framework

1. Draw N samples, $u_{n,h}$ from the vine copula model, which will be an $N \times 24$ matrix, uniformly distributed (column-wise)
2. Transform each column with the estimated quantile function for the residuals of the ARMA-GARCH models, producing t- or skew-t-distributed draws, $z_{n,h}$

Simulation of Price Distribution

Prediction Framework

1. Draw N samples, $u_{n,h}$ from the vine copula model, which will be an $N \times 24$ matrix, uniformly distributed (column-wise)
2. Transform each column with the estimated quantile function for the residuals of the ARMA-GARCH models, producing t- or skew-t-distributed draws, $z_{n,h}$
3. Use tail of input values for the ARMA-GARCH models together with $z_{n,h}$ to calculate $X_{n,h}$

Simulation of Price Distribution

Prediction Framework

1. Draw N samples, $u_{n,h}$ from the vine copula model, which will be an $N \times 24$ matrix, uniformly distributed (column-wise)
2. Transform each column with the estimated quantile function for the residuals of the ARMA-GARCH models, producing t- or skew-t-distributed draws, $z_{n,h}$
3. Use tail of input values for the ARMA-GARCH models together with $z_{n,h}$ to calculate $X_{n,h}$
4. Calculate $s_{n,h}$ and add to $X_{n,h}$ to obtain $p_{n,h}$

Simulation of Price Distribution

Prediction Framework

1. Draw N samples, $u_{n,h}$ from the vine copula model, which will be an $N \times 24$ matrix, uniformly distributed (column-wise)
2. Transform each column with the estimated quantile function for the residuals of the ARMA-GARCH models, producing t- or skew-t-distributed draws, $z_{n,h}$
3. Use tail of input values for the ARMA-GARCH models together with $z_{n,h}$ to calculate $X_{n,h}$
4. Calculate $s_{n,h}$ and add to $X_{n,h}$ to obtain $p_{n,h}$
5. Prices in EUR/MW are then $P_{n,h} = \exp p_{n,h} - K$

Simulation of Price Distribution

Example

Predicting the payoff distribution of a forward contract

Let $F(t, t_1, t_2)$ be a price in EUR/MW determined today (at time t) for delivery of power in the period $[t_1, t_2]$. Payoff on a long position is then

$$\sum_{s=t_1}^{t_2} \sum_{h \in H(s)} \left(P_{s,h} - F(t, t_1, t_2) \right),$$

where $H(s) \subseteq \{1, \ldots, 24\}$ are the hours on day s for which the forward is in effect.

Simulation of Price Distribution

Example

Example simulation

1. Prices simulated for the period 2019-01-01 – 2019-07-31
   $\Longrightarrow$ 212 observations $\times$ 24 hours.
2. Payoff calculated for some contract $F(t, t_1, t_2)$
3. Steps 1–2 repeated 500 times
   $\Longrightarrow$ simulated payoff distribution

Simulation of Price Distribution
Example

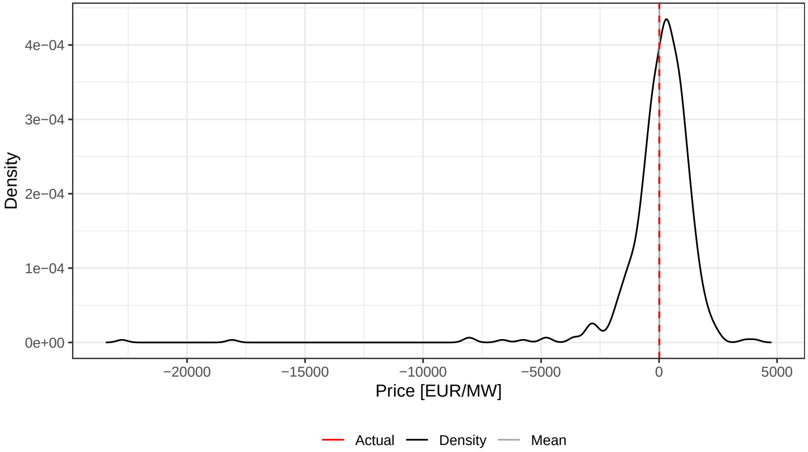
Example simulation

1. Prices simulated for the period 2019-01-01 – 2019-07-31
   $\Longrightarrow$ 212 observations $\times$ 24 hours.
2. Payoff calculated for some contract $F(t, t_1, t_2)$
3. Steps 1–2 repeated 500 times
   $\Longrightarrow$ simulated payoff distribution

For example, if the delivery interval is February 2019, on hours 2–4 and 16–18, for 40 EUR/MW, the simulated mean payoff is **14.10**, while the payoff using actual prices is **5.05**.

Simulation of Price Distribution

Example



Simulation of Price Distribution

Example of Single Draw

Future Considerations

Future Considerations

- Compare results with univariate model for average prices

Future Considerations

- Compare results with univariate model for average prices
- Majority of pair-copulas in first tree are t-copulas—model could be compared to a 24-dimensional t-copula

Future Considerations

- Compare results with univariate model for average prices
- Majority of pair-copulas in first tree are *t*-copulas—model could be compared to a 24-dimensional *t*-copula
- Simulation occasionally yields infinite values; about 2.5% of samples contained infinite values, and out of those, these values comprised about 0.5% of the sample—could be kinks in the code or parts of the model itself

Thanks for listening!