Spread option pricing: implied volatility implied from implied correlation *(and its implications!)*

Michael Coulon

(m.coulon@sussex.ac.uk)

University of Sussex

(work with Elisa Alòs, Universitat Pompeu Fabra, Barcelona)

Dec 14th, 2017

Energy Finance Christmas Workshop 2017,
Krakow, Poland

Energy or Commodity Spread Options

A general spread option payoff (at time T) has the general form (but sometimes $a = 1, b = 1$ and/or $K = 0$):

$$(aX_T - bY_T - K)^+$$

where X_T and Y_T are different commodity prices (spot or forward): e.g.

Energy or Commodity Spread Options

A general spread option payoff (at time T) has the general form (but sometimes $a = 1, b = 1$ and/or $K = 0$):

$$(aX_T - bY_T - K)^+$$

where X_T and Y_T are different commodity prices (spot or forward): e.g.

- Input / Output (e.g., 'dark' if X_T is electricity, Y_T is coal)
- Input / Output (e.g., 'crack' if X_T is refined product, Y_T is crude)
- Calendar (e.g., X_T is Dec13 forward, Y_T is Jun13 forward)
- Locational (e.g., X_T is Henry Hub gas, Y_T is NorthEast gas)

Energy or Commodity Spread Options

A general spread option payoff (at time T) has the general form (but sometimes $a = 1, b = 1$ and/or $K = 0$):

$$(aX_T - bY_T - K)^+$$

where X_T and Y_T are different commodity prices (spot or forward): e.g.

- Input / Output (e.g., 'dark' if X_T is electricity, Y_T is coal)
- Input / Output (e.g., 'crack' if X_T is refined product, Y_T is crude)
- Calendar (e.g., X_T is Dec13 forward, Y_T is Jun13 forward)
- Locational (e.g., X_T is Henry Hub gas, Y_T is NorthEast gas)

Spread options are of utmost importance, due to their strong link with **physical assets** (hence hedging and valuation tools). Examples above:

- Coal power plant, Oil refinery, Gas storage facility, Pipeline, etc.

Optimal **unconstrained** operation mimics a LONG string of spread options.

Margrabe's Approach

Margrabe ('78) introduced a formula for exchange options ($K = 0$) assuming lognormal underlyings. e.g. for payoff $(F_T^{(1)} - F_T^{(2)})^+$:

$$V_t = \mathrm{e}^{-r(T-t)} \left[F_T^{(1)} \Phi(d_+) - F_T^{(2)} \Phi(d_-) \right]$$

where $\Phi(d_\pm) = \frac{\log\left(F_T^{(1)}/F_T^{(2)}\right) \pm \frac{1}{2}\sigma_{ratio}^2}{\sigma_{ratio}}$ and $\sigma_{ratio}^2 = \mathrm{Var}_t\{\log(F_T^{(1)}/F_T^{(2)})\}$ depends on our model. e.g. for GBMs $\sigma_{ratio}^2 = (\sigma_1^2 + \sigma_2^2 - 2\rho\sigma_1\sigma_2)(T - t)$.

Margrabe's Approach

Margrabe ('78) introduced a formula for exchange options ($K = 0$) assuming lognormal underlyings. e.g. for payoff $(F_T^{(1)} - F_T^{(2)})^+$:

$$V_t = \mathsf{e}^{-r(T-t)} \left[F_T^{(1)} \Phi(d_+) - F_T^{(2)} \Phi(d_-) \right]$$

where $\Phi(d_\pm) = \frac{\log\left(F_T^{(1)}/F_T^{(2)}\right) \pm \frac{1}{2}\sigma_{ratio}^2}{\sigma_{ratio}}$ and $\sigma_{ratio}^2 = \mathsf{Var}_t\{\log(F_T^{(1)}/F_T^{(2)})\}$ depends on our model. e.g. for GBMs $\sigma_{ratio}^2 = (\sigma_1^2 + \sigma_2^2 - 2\rho\sigma_1\sigma_2)(T-t)$.

What if $K \neq 0$?

Margrabe's Approach

Margrabe ('78) introduced a formula for exchange options ($K = 0$) assuming lognormal underlyings. e.g. for payoff $(F_T^{(1)} - F_T^{(2)})^+$:

$$V_t = \mathrm{e}^{-r(T-t)} \left[F_T^{(1)} \Phi(d_+) - F_T^{(2)} \Phi(d_-) \right]$$

where $\Phi(d_\pm) = \frac{\log\left(F_T^{(1)}/F_T^{(2)}\right) \pm \frac{1}{2}\sigma_{ratio}^2}{\sigma_{ratio}}$ and $\sigma_{ratio}^2 = \mathrm{Var}_t\{\log(F_T^{(1)}/F_T^{(2)})\}$ depends on our model. e.g. for GBMs $\sigma_{ratio}^2 = (\sigma_1^2 + \sigma_2^2 - 2\rho\sigma_1\sigma_2)(T-t)$.

What if $K \neq 0$? No explicit formulas but various approximations:

- Kirk's approximation - quite widely used, and can be understood as using Margrabe with σ_2 adjusted: $\tilde{\sigma}_2 = \sigma_2 \left(\frac{F_0^{(1)}}{F_0^{(2)} + K} \right)$.
- Levy approximation - similar, assumes $F_T^{(2)} + K$ lognormal

Margrabe's Approach

Margrabe ('78) introduced a formula for exchange options $(K = 0)$ assuming lognormal underlyings. e.g. for payoff $(F_T^{(1)} - F_T^{(2)})^+$:

$$V_t = \mathrm{e}^{-r(T-t)} \left[F_T^{(1)} \Phi(d_+) - F_T^{(2)} \Phi(d_-) \right]$$

where $\Phi(d_\pm) = \frac{\log\left(F_T^{(1)}/F_T^{(2)}\right) \pm \frac{1}{2}\sigma_{ratio}^2}{\sigma_{ratio}}$ and $\sigma_{ratio}^2 = \mathrm{Var}_t\{\log(F_T^{(1)}/F_T^{(2)})\}$ depends on our model. e.g. for GBMs $\sigma_{ratio}^2 = (\sigma_1^2 + \sigma_2^2 - 2\rho\sigma_1\sigma_2)(T-t)$.

What if $K \neq 0$? No explicit formulas but various approximations:

- Kirk's approximation - quite widely used, and can be understood as using Margrabe with σ_2 adjusted: $\tilde{\sigma}_2 = \sigma_2 \left(\frac{F_0^{(1)}}{F_0^{(2)} + K} \right)$.
- Levy approximation - similar, assumes $F_T^{(2)} + K$ lognormal

See Carmona, Durrlemann (2003), Swindle (2014) for more details/ideas

Implied Correlations

Implied correlation ρ^{imp} is the number which makes the Margrabe formula (or potentially Kirk, etc.) match the true (market) price! Any problems?

Implied Correlations

Implied correlation ρ^{imp} is the number which makes the Margrabe formula (or potentially Kirk, etc.) match the true (market) price! Any problems?

- As we know, the prices are not lognormal!

Implied Correlations

Implied correlation ρ^{imp} is the number which makes the Margrabe formula (or potentially Kirk, etc.) match the true (market) price! Any problems?

- As we know, the prices are not lognormal!
- Furthermore, prices may be outside Margrabe's $\rho \in [-1, 1]$ range

Implied Correlations

Implied correlation ρ^{imp} is the number which makes the Margrabe formula (or potentially Kirk, etc.) match the true (market) price! Any problems?

- As we know, the prices are not lognormal!
- Furthermore, prices may be outside Margrabe's $\rho \in [-1, 1]$ range

$$\implies \quad \rho^{imp} \text{ may be undefined!}$$

Implied Correlations

Implied correlation ρ^{imp} is the number which makes the Margrabe formula (or potentially Kirk, etc.) match the true (market) price! Any problems?

- As we know, the prices are not lognormal!
- Furthermore, prices may be outside Margrabe's $\rho \in [-1, 1]$ range

$$\Longrightarrow \quad \rho^{imp} \text{ may be undefined!}$$

- Finally, ρ^{imp} depends on choice of (implied) volatilities σ_1, σ_2

Implied Correlations

Implied correlation ρ^{imp} is the number which makes the Margrabe formula (or potentially Kirk, etc.) match the true (market) price! Any problems?

- As we know, the prices are not lognormal!
- Furthermore, prices may be outside Margrabe's $\rho \in [-1, 1]$ range

$$\Longrightarrow \quad \rho^{imp} \text{ may be undefined!}$$

- Finally, ρ^{imp} depends on choice of (implied) volatilities σ_1, σ_2

$$\Longrightarrow \quad \text{a 'strike convention' } k_1, k_2 \text{ is needed}$$

Implied Correlations

Implied correlation ρ^{imp} is the number which makes the Margrabe formula (or potentially Kirk, etc.) match the true (market) price! Any problems?

- As we know, the prices are not lognormal!
- Furthermore, prices may be outside Margrabe's $\rho \in [-1, 1]$ range

$$\Longrightarrow \quad \rho^{imp} \text{ may be undefined!}$$

- Finally, ρ^{imp} depends on choice of (implied) volatilities σ_1, σ_2

$$\Longrightarrow \quad \text{a `strike convention' } k_1, k_2 \text{ is needed}$$

Building on Rebonato's well known quote, here we have something like:

Implied Correlations

Implied correlation ρ^{imp} is the number which makes the Margrabe formula (or potentially Kirk, etc.) match the true (market) price! Any problems?

- As we know, the prices are not lognormal!
- Furthermore, prices may be outside Margrabe's $\rho \in [-1, 1]$ range

$$\Longrightarrow \qquad \rho^{imp} \text{ may be undefined!}$$

- Finally, ρ^{imp} depends on choice of (implied) volatilities σ_1, σ_2

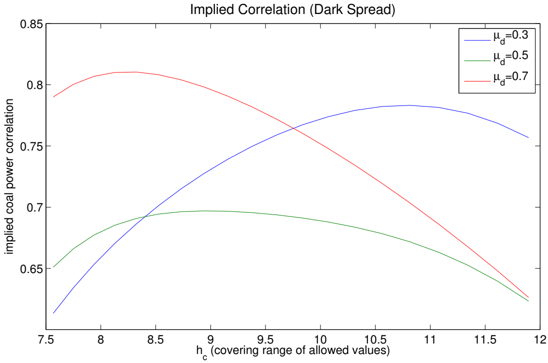
$$\Longrightarrow \qquad \text{a 'strike convention' } k_1, k_2 \text{ is needed}$$

Building on Rebonato's well known quote, here we have something like:

Implied correlation is the ***wrong*** *number put in the* ***wrong*** *formula along with two other* ***wrong*** *numbers from* ***wrong*** *formulas to get the* ***right*** *price.*

Implied Correlation for Structural Power Models

Example of results from Carmona, Coulon & Schwarz (2013) for spark spread options using a complex structural power price model:



Implied correlation 'frowns' are observed in many cases (in simpler models, and in the market), BUT does this correlation structure make any sense?

The Strike Convention

- **Goal:** to price an exchange option via Magrabe, using the implied volatilities I_1, I_2 of vanilla options:

$$BS\,(T, x, y, \gamma)\,, \quad \text{where } \gamma = \sqrt{I_1^2 + I_2^2 - 2\rho I_1 I_2}$$

  where BS denotes the classical Black-Scholes function in terms of the log-prices $x := \log S_0^1$, $y := \log S_0^2$.

The Strike Convention

- **Goal:** to price an exchange option via Magrabe, using the implied volatilities I_1, I_2 of vanilla options:

$$BS\left(T, x, y, \gamma\right), \quad \text{where } \gamma = \sqrt{I_1^2 + I_2^2 - 2\rho I_1 I_2}$$

  where BS denotes the classical Black-Scholes function in terms of the log-prices $x := \log S_0^1$, $y := \log S_0^2$.
- **Problem:** $I_1 = I(x, k_1)$ and $I_2 = I_2(y, k_2)$ depend on the strike!!!

Therefore, what is the optimal strike choice?

The Strike Convention

- **Goal:** to price an exchange option via Magrabe, using the implied volatilities I_1, I_2 of vanilla options:

$$BS\,(T, x, y, \gamma)\,, \quad \text{where } \gamma = \sqrt{I_1^2 + I_2^2 - 2\rho I_1 I_2}$$

where BS denotes the classical Black-Scholes function in terms of the log-prices $x := \log S_0^1$, $y := \log S_0^2$.

- **Problem:** $I_1 = I(x, k_1)$ and $I_2 = I_2(y, k_2)$ depend on the strike!!!

Therefore, what is the optimal strike choice?

- **Solution:** Defining $\hat{\gamma}$ via the 'true' (model) price V_0:

$$V_0 = \mathrm{e}^{-rT} E(S_T^1 - S_T^2)^+ = BS\,(T, x, y, \hat{\gamma}(x, y))\,,$$

our problem reduces to find k_1, k_2 such that

$$\gamma(x, y) = \hat{\gamma}(x, y).$$

Implied Correlation and the Strike Convention

Note: As implied correlation $\hat{\rho}$ is defined directly from 'implied gamma':

$$\hat{\gamma}(x, y) = \sqrt{I_1^2 + I_2^2 - 2\hat{\rho} I_1 I_2}$$

our problem is equivalent to finding k_1, k_2 such that

$$\rho(x, y) = \hat{\rho}(x, y).$$

Implied Correlation and the Strike Convention

Note: As implied correlation $\hat{\rho}$ is defined directly from 'implied gamma':

$$\hat{\gamma}(x,y) = \sqrt{I_1^2 + I_2^2 - 2\hat{\rho} I_1 I_2}$$

our problem is equivalent to finding k_1, k_2 such that

$$\rho(x,y) = \hat{\rho}(x,y).$$

Key idea: Choose the strike convention such that spread price is consistent with known (historically estimated) correlation, if possible also:

Implied Correlation and the Strike Convention

Note: As implied correlation $\hat{\rho}$ is defined directly from 'implied gamma':

$$\hat{\gamma}(x, y) = \sqrt{I_1^2 + I_2^2 - 2\hat{\rho} I_1 I_2}$$

our problem is equivalent to finding k_1, k_2 such that

$$\rho(x, y) = \hat{\rho}(x, y).$$

Key idea: Choose the strike convention such that spread price is consistent with known (historically estimated) correlation, if possible also:

- keeping methodology as model-independent as possible

Implied Correlation and the Strike Convention

Note: As implied correlation $\hat{\rho}$ is defined directly from 'implied gamma':

$$\hat{\gamma}(x, y) = \sqrt{I_1^2 + I_2^2 - 2\hat{\rho} I_1 I_2}$$

our problem is equivalent to finding k_1, k_2 such that

$$\rho(x, y) = \hat{\rho}(x, y).$$

Key idea: Choose the strike convention such that spread price is consistent with known (historically estimated) correlation, if possible also:

- keeping methodology as model-independent as possible
- pricing consistently across all moneyness... no more FROWNS! :)

Implied Correlation and the Strike Convention

Note: As implied correlation $\hat{\rho}$ is defined directly from 'implied gamma':

$$\hat{\gamma}(x,y) = \sqrt{I_1^2 + I_2^2 - 2\hat{\rho} I_1 I_2}$$

our problem is equivalent to finding k_1, k_2 such that

$$\rho(x,y) = \hat{\rho}(x,y).$$

Key idea: Choose the strike convention such that spread price is consistent with known (historically estimated) correlation, if possible also:

- keeping methodology as model-independent as possible
- pricing consistently across all moneyness... no more FROWNS! :)

Methodology: Expand γ and $\hat{\gamma}$ (or ρ and $\hat{\rho}$) to first order as functions of y, in the short-time limit. Then match terms to solve for k_1, k_2.

The Underlying Model

Stochastic volatility model for the two assets (with $r = 0$ for simplicity):

$$\frac{dS_t^1}{S_t^1} = \lambda_1 \sigma_t dW_t^{(1)}, \quad x = \log(S_0^1)$$
$$\frac{dS_t^2}{S_t^2} = \lambda_2 \sigma_t dW_t^{(2)}, \quad y = \log(S_0^2)$$

with $\lambda_1, \lambda_2 > 0$ and σ_t is a non-negative square integrable process driven by another Brownian motion Z_t, with correlations

$$\langle W_t^1, Z \rangle = \rho_1, \langle W_t^2, Z \rangle = \rho_2, \langle W_t^1, W_t^2 \rangle = \rho.$$

(note: can use fractional volatility models as in Alòs, León, Vives (2007))

The Underlying Model

Stochastic volatility model for the two assets (with $r = 0$ for simplicity):

$$\frac{dS_t^1}{S_t^1} = \lambda_1 \sigma_t dW_t^{(1)}, \quad x = \log(S_0^1)$$
$$\frac{dS_t^2}{S_t^2} = \lambda_2 \sigma_t dW_t^{(2)}, \quad y = \log(S_0^2)$$

with $\lambda_1, \lambda_2 > 0$ and σ_t is a non-negative square integrable process driven by another Brownian motion Z_t, with correlations

$$\left\langle W_t^1, Z \right\rangle = \rho_1, \left\langle W_t^2, Z \right\rangle = \rho_2, \left\langle W_t^1, W_t^2 \right\rangle = \rho.$$

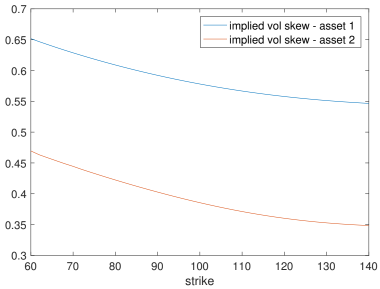
(note: can use fractional volatility models as in Alòs, León, Vives (2007))

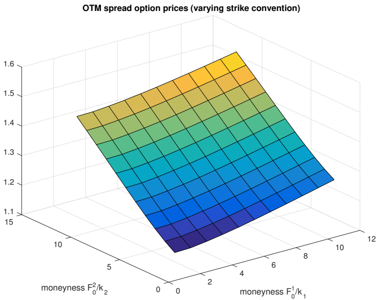
But for now a simple numerical example using the Heston Model:

$$d\sigma_t^2 = \kappa \left(\theta - \sigma_t^2\right) dt + \nu \sqrt{\sigma_t^2} dZ_t$$

Implied Correlation and the Strike Convention

Example: $(F_T^{(1)} - F_T^{(2)})^+$, with $T = 0.05$,
Heston params $\kappa = 1.5, \theta = 0.15, \nu = 0.5, \sigma_0 = 0.15, \lambda_1 = 1.5, \lambda_2 = 1$,
and correlations $\rho = 0.5, \rho_1 = -0.4, \rho_2 = -0.5$:

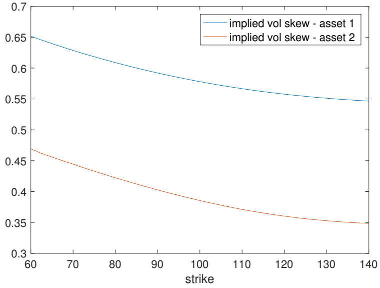


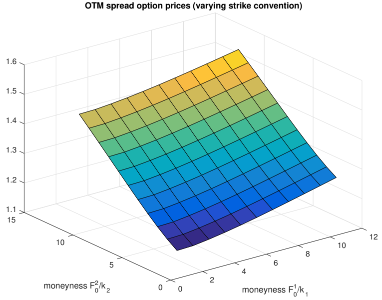


Implied vol skews (left) and spread prices (right) for OTM: $F_0^{(1)} = 90, F_0^{(2)} = 100$

Implied Correlation and the Strike Convention

Example: $(F_T^{(1)} - F_T^{(2)})^+$, with $T = 0.05$,
Heston params $\kappa = 1.5, \theta = 0.15, \nu = 0.5, \sigma_0 = 0.15, \lambda_1 = 1.5, \lambda_2 = 1$,
and correlations $\rho = 0.5, \rho_1 = -0.4, \rho_2 = -0.5$:





Implied vol skews (left) and spread prices (right) for OTM: $F_0^{(1)} = 90, F_0^{(2)} = 100$

Key challenge: A big range of spread option prices possible... how to pick?

A Brief Intermission (sorry, no coffee... but lunch is soon!)

At this point, you may be wondering about one (or all) of the following:

A Brief Intermission (sorry, no coffee... but lunch is soon!)

At this point, you may be wondering about one (or all) of the following:

1. What are we doing again?

A Brief Intermission (sorry, no coffee... but lunch is soon!)

At this point, you may be wondering about one (or all) of the following:

1. What are we doing again?
2. What is the point of it anyway?

A Brief Intermission (sorry, no coffee... but lunch is soon!)

At this point, you may be wondering about one (or all) of the following:

1. What are we doing again?
2. What is the point of it anyway?
3. What do people ‘normally’ do?

1. What are we doing again?

An intuitive way of thinking about the aims / benefits:

- ‘Improving’ Margrabe to make it more consistent, both:
    - between different options (across strikes)
    - with historical correlation estimates

1. What are we doing again?

An intuitive way of thinking about the aims / benefits:

- 'Improving' Margrabe to make it more consistent, both:
  - between different options (across strikes)
  - with historical correlation estimates
- Without fitting a full model, we can translate information contained in the two observed volatility smiles / skews into spread option pricing.

1. What are we doing again?

An intuitive way of thinking about the aims / benefits:

- 'Improving' Margrabe to make it more consistent, both:
  - between different options (across strikes)
  - with historical correlation estimates
- Without fitting a full model, we can translate information contained in the two observed volatility smiles / skews into spread option pricing.
- Aside: Is an implied correlation frown logical? Swindle (2014) argues...

1. What are we doing again?

An intuitive way of thinking about the aims / benefits:

- ‘Improving’ Margrabe to make it more consistent, both:
  - between different options (across strikes)
  - with historical correlation estimates
- Without fitting a full model, we can translate information contained in the two observed volatility smiles / skews into spread option pricing.
- Aside: Is an implied correlation frown logical? Swindle (2014) argues...

variations in implied correlation are *“purely an artifact of the interaction of skew with the Margrabe formulation."*

2. What is the point of it anyway?

These are European spread options... why try to use Margrabe anyway?

2. What is the point of it anyway?

These are European spread options... why try to use Margrabe anyway?

- Simpler approaches are still common in industry, and convenient

2. What is the point of it anyway?

These are European spread options... why try to use Margrabe anyway?

- Simpler approaches are still common in industry, and convenient
- Computation times can still be a big problem:
  - if maturity frequency is high (e.g. hourly)
  - or if a single valuation is only the first step (e.g., risk management)

2. What is the point of it anyway?

These are European spread options... why try to use Margrabe anyway?

- Simpler approaches are still common in industry, and convenient
- Computation times can still be a big problem:
    - if maturity frequency is high (e.g. hourly)
    - or if a single valuation is only the first step (e.g., risk management)

Thompson (2016) uses a PDE approach to study valuation and associated credit risk (PFE and CVA) of natural gas storage facilities. He notes that

2. What is the point of it anyway?

These are European spread options... why try to use Margrabe anyway?

- Simpler approaches are still common in industry, and convenient
- Computation times can still be a big problem:
  - if maturity frequency is high (e.g. hourly)
  - or if a single valuation is only the first step (e.g., risk management)

Thompson (2016) uses a PDE approach to study valuation and associated credit risk (PFE and CVA) of natural gas storage facilities. He notes that

for a moderately sized firm, *"the number of individual deal level valuations that need to be done **each day** to price and manage counter-party credit risk can easily number into the many **trillions**."*

3. What do people 'normally' do?

Recalling that $x = \log(S_0^1), y = \log(S_0^1)$, what 'strike conventions' might we try? (for strikes e^{k_1}, e^{k_2})

3. What do people 'normally' do?

Recalling that $x = \log(S_0^1), y = \log(S_0^1)$, what 'strike conventions' might we try? (for strikes e^{k_1}, e^{k_2})

1. $k_1 = x$, $k_2 = y$ (use ATM vol for both assets)

3. What do people ‘normally’ do?

Recalling that $x = \log(S_0^1), y = \log(S_0^1)$, what ‘strike conventions’ might we try? (for strikes e^{k_1}, e^{k_2})

1. $k_1 = x$, $k_2 = y$ (use ATM vol for both assets)
2. $k_1 = y$, $k_2 = x$ (fix opposite leg as strike to look up the imp vol)

3. What do people 'normally' do?

Recalling that $x = \log(S_0^1), y = \log(S_0^1)$, what 'strike conventions' might we try? (for strikes e^{k_1}, e^{k_2})

1. $k_1 = x$, $k_2 = y$ (use ATM vol for both assets)
2. $k_1 = y$, $k_2 = x$ (fix opposite leg as strike to look up the imp vol)
3. $k_1 = k_2 = \frac{x+y}{2}$ (a compromise?)

Which seems more logical?

3. What do people 'normally' do?

Recalling that $x = \log(S_0^1), y = \log(S_0^1)$, what 'strike conventions' might we try? (for strikes e^{k_1}, e^{k_2})

1. $k_1 = x$, $k_2 = y$ (use ATM vol for both assets)
2. $k_1 = y$, $k_2 = x$ (fix opposite leg as strike to look up the imp vol)
3. $k_1 = k_2 = \frac{x+y}{2}$ (a compromise?)

Which seems more logical? Hard to say! What do people do?

3. What do people 'normally' do?

Recalling that $x = \log(S_0^1), y = \log(S_0^1)$, what 'strike conventions' might we try? (for strikes e^{k_1}, e^{k_2})

1. $k_1 = x$, $k_2 = y$ (use ATM vol for both assets)
2. $k_1 = y$, $k_2 = x$ (fix opposite leg as strike to look up the imp vol)
3. $k_1 = k_2 = \frac{x+y}{2}$ (a compromise?)

Which seems more logical? Hard to say! What do people do?
Very little academic literature on this issue:

- Alexander and Venkatramanan (2011): mention using first approach above, and testing a number of alternatives, but in a different context.

3. What do people 'normally' do?

Recalling that $x = \log(S_0^1), y = \log(S_0^1)$, what 'strike conventions' might we try? (for strikes e^{k_1}, e^{k_2})

1. $k_1 = x$, $k_2 = y$ (use ATM vol for both assets)
2. $k_1 = y$, $k_2 = x$ (fix opposite leg as strike to look up the imp vol)
3. $k_1 = k_2 = \frac{x+y}{2}$ (a compromise?)

Which seems more logical? Hard to say! What do people do?
Very little academic literature on this issue:

- Alexander and Venkatramanan (2011): mention using first approach above, and testing a number of alternatives, but in a different context.
- Swindle (2014): discusses issue at length, suggests 'vol look up heutristic' (2nd above) and investigates size of impact. He notes:

"skew risk can manifest itself as spurious correlation risk simply due to the look-up heuristic."

3. What do people 'normally' do?

Recalling that $x = \log(S_0^1), y = \log(S_0^1)$, what 'strike conventions' might we try? (for strikes e^{k_1}, e^{k_2})

1. $k_1 = x$, $k_2 = y$ (use ATM vol for both assets)
2. $k_1 = y$, $k_2 = x$ (fix opposite leg as strike to look up the imp vol)
3. $k_1 = k_2 = \frac{x+y}{2}$ (a compromise?)

Which seems more logical? Hard to say! What do people do?
Very little academic literature on this issue:

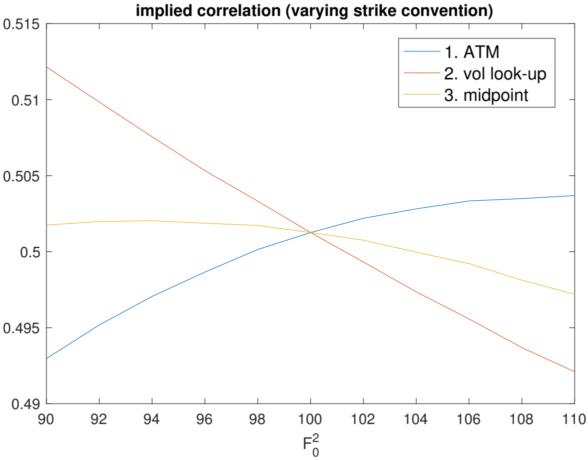
- Alexander and Venkatramanan (2011): mention using first approach above, and testing a number of alternatives, but in a different context.
- Swindle (2014): discusses issue at length, suggests 'vol look up heutristic' (2nd above) and investigates size of impact. He notes:

"skew risk can manifest itself as spurious correlation risk simply due to the look-up heuristic."

Can we investigate this in the Heston model example?

Implied Correlation and the Strike Convention

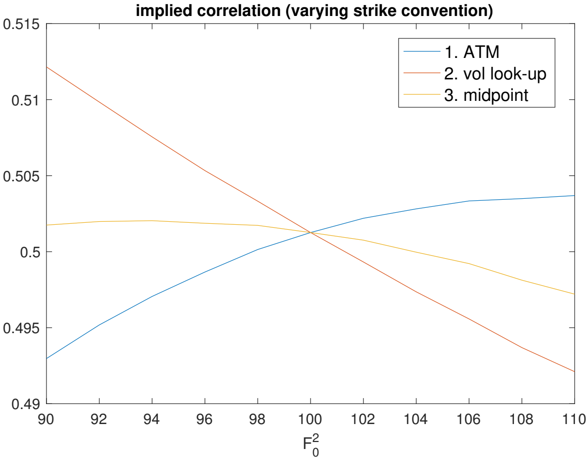
Comparing the three strike conventions for different F_0^2 (moneyness):



So perhaps the midpoint idea is the best?

Implied Correlation and the Strike Convention

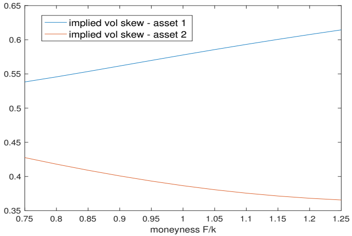
Comparing the three strike conventions for different F_0^2 (moneyness):

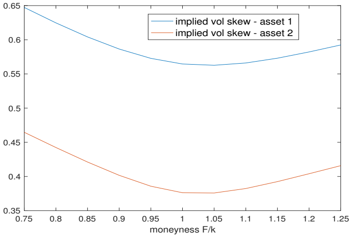


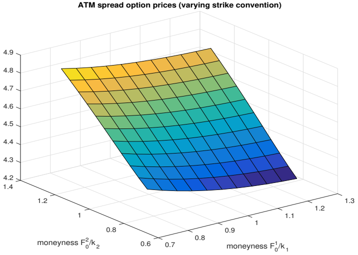
So perhaps the midpoint idea is the best? NOT ALWAYS!

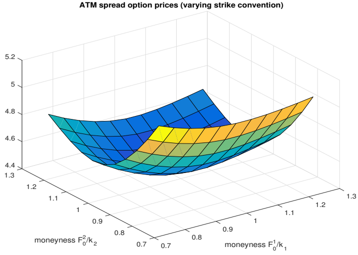
What about other parameter sets?

(i) Left: $\rho_1 = -0.5, \rho_2 = 0.4$; (ii) Right: $\rho_1 = \rho_2 = 0.1, \nu = 1.5$





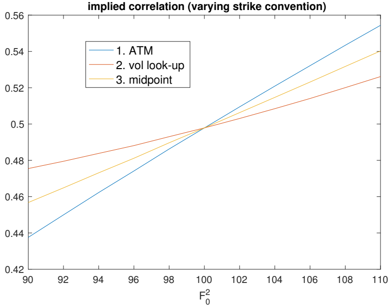


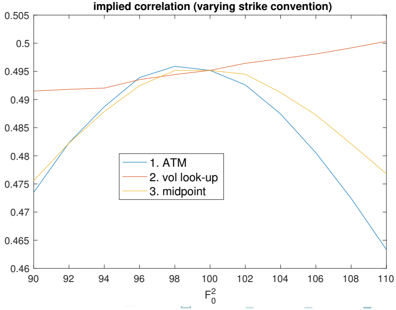


Implied Correlation and the Strike Convention

Key Observations:

- The best choice of convention can vary significantly across cases
- In some cases, best choice is NOT between the two simplest choices (ATM vs vol look up)
- High vol of vol case (steep smiles) leads to implied correlation frowns.
- Pricing differences can end up large, especially for OTM case.





Towards an Optimal Strike Convention

- Recall idea: seek $k_1(x,y)$, $k_2(x,y)$ such that $\hat{\rho}(k_1,k_2)=\rho$.
- A first order Taylor expansion (around the ATM point $x=y$) gives us:

$$\hat{\rho}\left(T,x,y,k_1,k_2\right) \approx \hat{\rho}\left(T,x,x,k_1,k_2\right) + \frac{\partial \hat{\rho}}{\partial y}\left(T,x,y,k_1,k_2\right)\Big|_{x=y}(y-x)$$

Towards an Optimal Strike Convention

- Recall idea: seek $k_1(x,y)$, $k_2(x,y)$ such that $\hat{\rho}(k_1,k_2)=\rho$.
- A first order Taylor expansion (around the ATM point $x=y$) gives us:

$$\hat{\rho}\left(T,x,y,k_1,k_2\right)\approx\hat{\rho}\left(T,x,x,k_1,k_2\right)+\frac{\partial\hat{\rho}}{\partial y}\left(T,x,y,k_1,k_2\right)\big|_{x=y}\left(y-x\right)$$

- Results for short T ATM vol imply $\lim_{T\to 0}\hat{\rho}\left(T,x,x,k_1,k_2\right)=\rho$.
- Hence seek to minimize the quantity $\lim_{T\to 0}\frac{\partial\hat{\rho}}{\partial y}\left(T,x,y,k_1,k_2\right)\big|_{x=y}$.

Or equivalently rewrite the problem in terms of γ instead of ρ...

Short-time expansion for γ (and $\hat{\gamma}$)

Using the fact that

$$\lim_{T\to 0} I_1(x,x) = \lambda_1\sigma_0, \quad \lim_{T\to 0} I_2(x,x) = \lambda_2\sigma_0$$

and a first-order Taylor expansion we get (with $C = \sqrt{\lambda_1^2 + \lambda_2^2 - 2\rho\lambda_1\lambda_2}$):

$$\gamma(x,y) \approx C\sigma_0 + \frac{1}{C}\left[(\lambda_1 - \rho\lambda_2)\frac{\partial I_1}{\partial z}\frac{\partial k_1}{\partial y} + (\lambda_2 - \rho\lambda_1)\left(\frac{\partial I_2}{\partial z}\frac{\partial k_2}{\partial y} + \frac{\partial I_2}{\partial y}\right)\right](x-y),$$

Short-time expansion for γ (and $\hat{\gamma}$)

Using the fact that

$$\lim_{T\to 0} I_1(x,x) = \lambda_1\sigma_0, \quad \lim_{T\to 0} I_2(x,x) = \lambda_2\sigma_0$$

and a first-order Taylor expansion we get (with $C = \sqrt{\lambda_1^2 + \lambda_2^2 - 2\rho\lambda_1\lambda_2}$):

$$\gamma(x,y) \approx C\sigma_0 + \frac{1}{C}\left[(\lambda_1 - \rho\lambda_2)\frac{\partial I_1}{\partial z}\frac{\partial k_1}{\partial y} + (\lambda_2 - \rho\lambda_1)\left(\frac{\partial I_2}{\partial z}\frac{\partial k_2}{\partial y} + \frac{\partial I_2}{\partial y}\right)\right](x-y),$$

Malliavin Calculus techniques (see Alòs, León and Vives (2007)) give

$$\lim_{T\to 0} T^{\alpha}\frac{\partial I_1}{\partial y}(x,x) = -\frac{\rho_1}{2\lambda_1^2\sigma_0^2}\lim_{T\to 0}\frac{1}{T^{2-\alpha}}E\left(\int_0^T\int_s^T D_s^Z\sigma_u^2 du ds\right),$$

where $\alpha := H - 1/2$ (for diffusion models $H = 1/2$ and $\alpha = 0$) and D denotes the Malliavin derivative operator (e.g. $D = \frac{\nu}{2}$ for Heston).
Then via similar expressions for $\lim_{T\to 0} T^{\alpha}\frac{\partial I_2}{\partial y}(x,x), \lim_{T\to 0} T^{\alpha}\frac{\partial \hat{\gamma}}{\partial y}(x,x)$...

Optimal Strike Convention - Key Result

Main theoretical result: By matching terms in the expansions for γ and $\hat{\gamma}$, we have (under suitable integrability conditions, and for $\lambda_1 \neq 0$)

$$\frac{\partial k_1}{\partial y}\frac{\partial I_1}{\partial z}\left(1 - \frac{\rho\lambda_2}{\lambda_1}\right) + \frac{\partial k_2}{\partial y}\frac{\partial I_2}{\partial z}\left(\frac{\lambda_2}{\lambda_1} - \rho\right) = \frac{\partial I_1}{\partial z} - \rho\frac{\partial I_2}{\partial z}$$

Optimal Strike Convention - Key Result

Main theoretical result: By matching terms in the expansions for γ and $\hat{\gamma}$, we have (under suitable integrability conditions, and for $\lambda_1 \neq 0$)

$$\frac{\partial k_1}{\partial y}\frac{\partial I_1}{\partial z}\left(1 - \frac{\rho\lambda_2}{\lambda_1}\right) + \frac{\partial k_2}{\partial y}\frac{\partial I_2}{\partial z}\left(\frac{\lambda_2}{\lambda_1} - \rho\right) = \frac{\partial I_1}{\partial z} - \rho\frac{\partial I_2}{\partial z}$$

Notice that:

- This equation relating $\frac{\partial k_1}{\partial y}$ and $\frac{\partial k_2}{\partial y}$ does not depend on the specific model for σ_t (can even include fractional volatility).

Optimal Strike Convention - Key Result

Main theoretical result: By matching terms in the expansions for γ and $\hat{\gamma}$, we have (under suitable integrability conditions, and for $\lambda_1 \neq 0$)

$$\frac{\partial k_1}{\partial y}\frac{\partial I_1}{\partial z}\left(1 - \frac{\rho\lambda_2}{\lambda_1}\right) + \frac{\partial k_2}{\partial y}\frac{\partial I_2}{\partial z}\left(\frac{\lambda_2}{\lambda_1} - \rho\right) = \frac{\partial I_1}{\partial z} - \rho\frac{\partial I_2}{\partial z}$$

Notice that:

- This equation relating $\frac{\partial k_1}{\partial y}$ and $\frac{\partial k_2}{\partial y}$ does not depend on the specific model for σ_t (can even include fractional volatility).
- Can substitute for key market observables from vanilla options, $\frac{\partial I_1}{\partial z}$, $\frac{\partial I_2}{\partial z}$ from ATM skews, and $\frac{\lambda_2}{\lambda_1}$ from ratio of ATM vol levels.

Optimal Strike Convention - Key Result

Main theoretical result: By matching terms in the expansions for γ and $\hat{\gamma}$, we have (under suitable integrability conditions, and for $\lambda_1 \neq 0$)

$$\frac{\partial k_1}{\partial y}\frac{\partial I_1}{\partial z}\left(1 - \frac{\rho\lambda_2}{\lambda_1}\right) + \frac{\partial k_2}{\partial y}\frac{\partial I_2}{\partial z}\left(\frac{\lambda_2}{\lambda_1} - \rho\right) = \frac{\partial I_1}{\partial z} - \rho\frac{\partial I_2}{\partial z}$$

Notice that:

- This equation relating $\frac{\partial k_1}{\partial y}$ and $\frac{\partial k_2}{\partial y}$ does not depend on the specific model for σ_t (can even include fractional volatility).
- Can substitute for key market observables from vanilla options, $\frac{\partial I_1}{\partial z}, \frac{\partial I_2}{\partial z}$ from ATM skews, and $\frac{\lambda_2}{\lambda_1}$ from ratio of ATM vol levels.

Next step: Choose an appropriate form for $k_1(x, y), k_2(x, y)$ and simplify the expression above.

Optimal Log-Linear Strike Conventions

We assume a symmetric log-linear strike convention of the form:

$$k_1(x, y) = (1 - a)\, x + ay$$
$$k_2(x, y) = ax + (1 - a)\, y,$$

typically for $a \in [0, 1]$, (but can also let $a \in \mathbb{R}$).

Optimal Log-Linear Strike Conventions

We assume a symmetric log-linear strike convention of the form:

$$k_1(x,y) = (1-a)\,x + ay$$
$$k_2(x,y) = ax + (1-a)\,y,$$

typically for $a \in [0,1]$, (but can also let $a \in \mathbb{R}$).

Earlier result simplifies to give an optimal choice of a (if finite) given by:

$$a = \frac{(\lambda_1\rho_1 - \lambda_2\rho_2)}{\rho_1\,(\lambda_1 - \rho\lambda_2) - \rho_2\,(\lambda_2 - \rho\lambda_1)}$$

(recall that this choice ensures $\lim_{T\to 0} \frac{\partial\hat{\rho}}{\partial x}\,(T,x,y,k_1,k_2)|_{x=y} = 0$).

Optimal Linear Strike Conventions

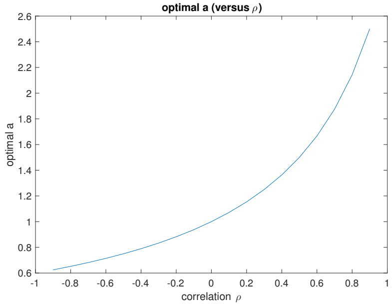
What does a look like in a simple case?

Optimal Linear Strike Conventions

What does a look like in a simple case? Notice that $\rho = 0$ gives (why?):

$$a = \frac{(\lambda_1 \rho_1 - \lambda_2 \rho_2)}{\rho_1 (\lambda_1 - \rho \lambda_2) - \rho_2 (\lambda_2 - \rho \lambda_1)} = \frac{(\lambda_1 \rho_1 - \lambda_2 \rho_2)}{(\lambda_1 \rho_1 - \lambda_2 \rho_2)} = 1$$

If instead $\rho_1 = 0$, then $a = \frac{\lambda_1}{\lambda_1 - \rho \lambda_2}$. Intuition when $\rho > 0$? Or $\rho < 0$?

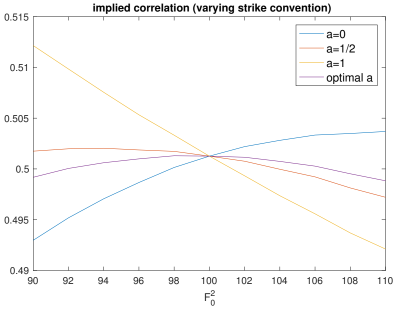


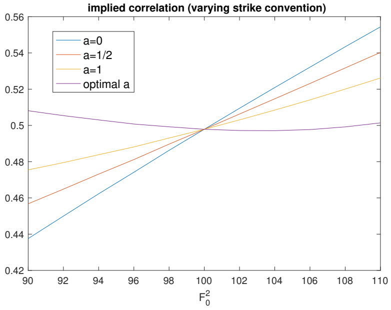
Numerical Tests of Optimal Strike Convention

Adding the optimal a line (in purple) to our plots from earlier:

- Example 1 (left): two downward skews $\Longrightarrow$ $a = 0.364$
- Example 2 (right): one downward, one upward $\Longrightarrow$ $a = 1.917$

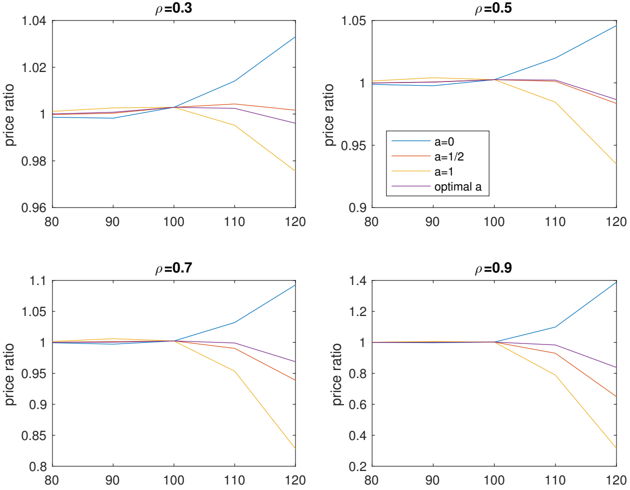
As hoped, ρ^{imp} is close to flat across moneyness and matching 'true' ρ.





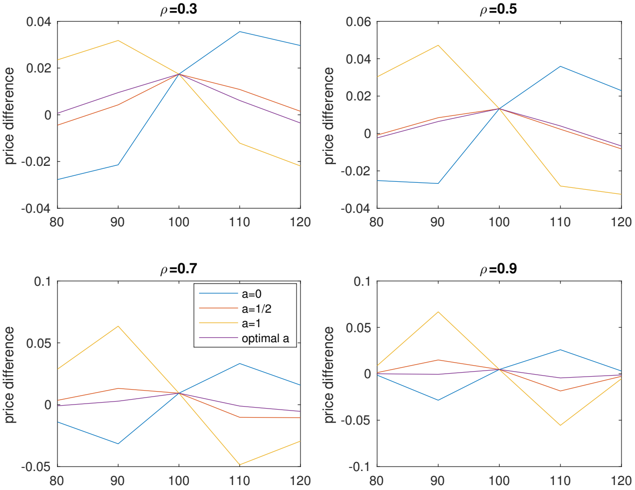
More Numerical Tests - Price Ratios (% pricing errors)

OTM % errors dominate (here $\rho_1 = -0.7$, $\rho_2 = -0.8$, varying ρ):



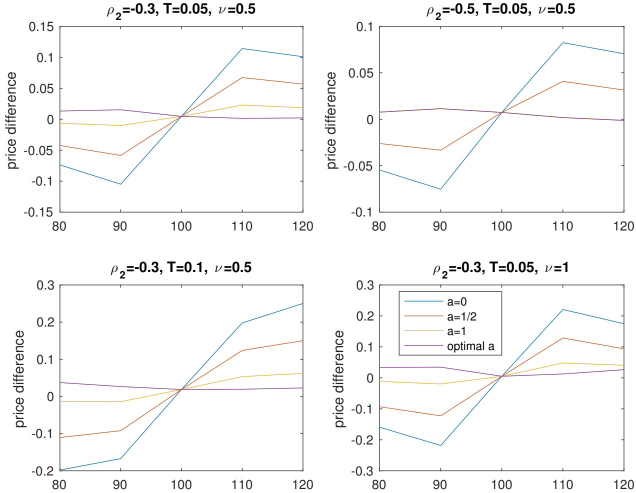
More Numerical Tests - Price Differences

ITM and OTM absolute errors similar (same parameters as above):



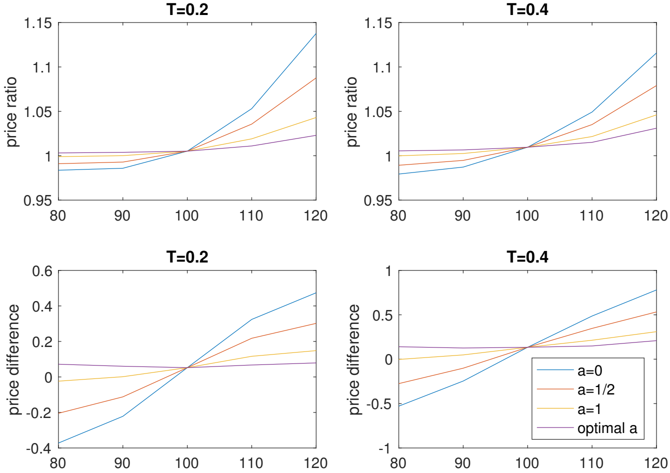
More Numerical Tests - Varying ρ_2, T, and ν

Higher ν (vol of vol) doesn't change the optimal a, but does increase error:



More Numerical Tests - Varying T

Extending to longer maturities also produces encouraging results so far...



More Extensive Numerical Tests - Parameters

Instead of picking sample cases, we now attempt to test a wide range of parameters:

- $S_0^1 = 100,\ S_0^2 \in [80, 84, \ldots, 100, \ldots, 116, 120]$
- $\lambda_1 = 1,\ \lambda_2 = [0.72, 1.02, 1.32]$
- Heston parameters (as before): $\kappa = 1.5, \theta = 0.15, \nu = 0.5, \sigma_0 = 0.15$
- $\rho \in [-0.5, -0.3, -0.1, 0.1, 0.3, 0.5, 0.7]$
- $\rho_1 \in [-0.72, -0.42, -0.12, 0.18, 0.48]$
- $\rho_2 \in [-0.61, -0.31, -0.01, 0.29, 0.59]$
- $T = 0.1$

More Extensive Numerical Tests - Parameters

Instead of picking sample cases, we now attempt to test a wide range of parameters:

- $S_0^1 = 100$, $S_0^2 \in [80, 84, \ldots, 100, \ldots, 116, 120]$
- $\lambda_1 = 1$, $\lambda_2 = [0.72, 1.02, 1.32]$
- Heston parameters (as before): $\kappa = 1.5, \theta = 0.15, \nu = 0.5, \sigma_0 = 0.15$
- $\rho \in [-0.5, -0.3, -0.1, 0.1, 0.3, 0.5, 0.7]$
- $\rho_1 \in [-0.72, -0.42, -0.12, 0.18, 0.48]$
- $\rho_2 \in [-0.61, -0.31, -0.01, 0.29, 0.59]$
- $T = 0.1$

Note: challenging to pick a 'fair' set, because of how our optimal a can change quickly, especially if its denominator gets close to zero. Also need positive definite correlation matrices (some cases thus omitted).

More Extensive Numerical Tests - Average Errors

Mean Absolute Errors (MAE) for different cases (best convention in red):

	ρ	$a = 0$	$a = 1/2$	$a = 1$	Optimal a	ATM error
$\lambda_2 = 0.72$	-0.5	0.2279	0.0738	0.1275	**0.048**	0.0449
	-0.3	0.2255	0.0967	0.0907	**0.0558**	0.0542
	-0.1	0.2333	0.1177	0.0644	**0.0552**	0.0541
	0.1	0.2496	0.1437	**0.0584**	0.0709	0.0443
	0.3	0.2499	0.1642	0.088	**0.0868**	0.0412
	0.5	0.2224	0.1662	**0.1195**	0.1268	0.0441
	0.7	0.2136	0.1758	0.1538	**0.0977**	0.0383
$\lambda_2 = 1.32$	-0.5	0.2625	0.117	0.158	**0.0949**	0.0956
	-0.3	0.262	0.14	0.1326	**0.1076**	0.1061
	-0.1	0.2677	0.1579	**0.111**	0.1176	0.1085
	0.1	0.2806	0.1778	0.1042	**0.1028**	0.0993
	0.3	0.2886	0.1986	**0.1227**	0.1303	0.0936
	0.5	0.2764	0.2053	0.1504	**0.1411**	0.0841
	0.7	0.2809	0.2218	0.194	**0.1175**	0.0707

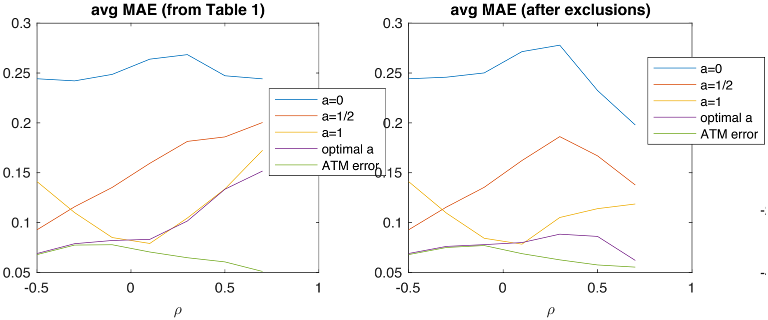
Note: right most column (ATM error) is the 'best case' we can hope for.

More Extensive Numerical Tests - Impact of Exteme a

Recall that

$$a = \frac{(\lambda_1\rho_1 - \lambda_2\rho_2)}{\rho_1(\lambda_1 - \rho\lambda_2) - \rho_2(\lambda_2 - \rho\lambda_1)}$$

How much are we hurt by cases which produce high $|a|$? Left plot averages all results while right plot excludes cases with $a < -1$ or $a > 2$:



Note: more could also be done to improve extrapolation of OTM implied vols

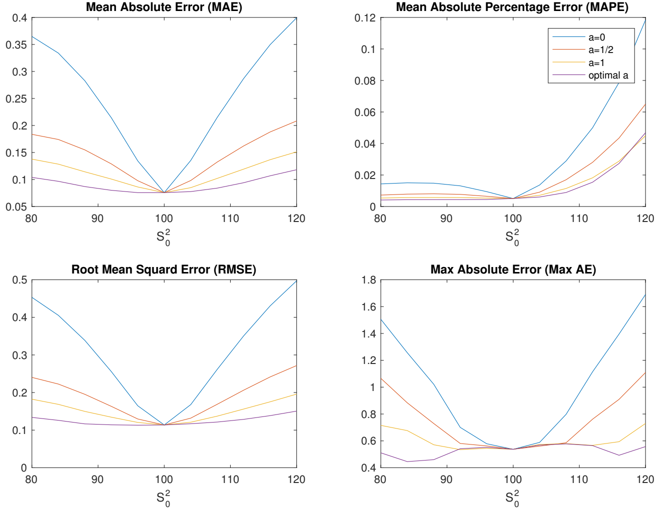
Average Errors for Different Error Measures

Results for Mean Absolute Errors (MAE), Mean Absolute Percentage Error (MAPE), Root Mean Squared Error (RMSE), Maximum Absolute Error (Max AE), and mean standard deviation of errors across moneyness grid (MStd):

	Measure	$a = 0$	$a = 1/2$	$a = 1$	Optimal a	ATM error
no exclusion	MAE	0.2512	0.153	0.118	**0.1**	0.0672
	MAPE	0.0275	0.0183	0.0141	**0.013**	0.0051
	RMSE	0.3264	0.1927	0.1476	**0.1352**	0.081
	Max AE	1.0097	0.5714	**0.4221**	0.5319	0.1574
	MStd	0.2727	0.1408	0.0912	**0.0316**	n/a
$-1 \leq a \leq 2$	MAE	0.2457	0.1424	0.1073	**0.0771**	0.0663
	MAPE	0.0258	0.0156	0.0114	**0.008**	0.0051
	RMSE	0.3194	0.1798	0.1342	**0.093**	0.0806
	Max AE	0.9897	0.5421	0.3864	**0.2464**	0.1571
	MStd	0.2727	0.1408	0.0912	**0.0316**	n/a

Average Errors versus Moneyness

Here a clear consistency benefit when using the optimal a. (purple lines)



Conclusions and Further Work

Key contributions:

- A rigorously justified optimal choice of strike convention, which adapts to the covariance structure of the two assets and volatility process.
- Model-independent inputs via market observables (and historical ρ)
- A tool to correct for the mispecification caused by Margrabe and skew.
- Numerical experiments to test and confirm results.
- Investigations into when the issue matters most.

Conclusions and Further Work

Key contributions:

- A rigorously justified optimal choice of strike convention, which adapts to the covariance structure of the two assets and volatility process.
- Model-independent inputs via market observables (and historical ρ)
- A tool to correct for the mispecification caused by Margrabe and skew.
- Numerical experiments to test and confirm results.
- Investigations into when the issue matters most.

Many further questions to explore and more work possible:

- Error bounds on the price; relevance of second order terms
- Investigation of possible extensions of results to other cases:
  - Kirk instead of Margrabe
  - behaviour for larger T
- Empirical analysis, if reliable spread data is available!