Semi-Static and Sparse Variance-Optimal Hedging

Martin Keller-Ressel

TU Dresden
Joint work with Paolo Di Tella and Martin Haubold

Advances in Stochastic Analysis for Risk Modeling, CIRM Marseille
November 17, 2017

Part I

Motivation

Semi-static replication of variance swaps

Variance-Swap replication (Neuberger 1994, Carr and Madan 2001)

Let S be a continuous local martingale. Then the variance swap $H_T^{\text{swap}} = \langle \log S, \log_S \rangle_T$ can be replicated by dynamic trading in S and static positions in European puts and calls, i.e.

$$H_T^{\text{swap}} = \underbrace{2\int_0^T \left(\frac{1}{S_t} - \frac{1}{S_0}\right) \mathrm{d}S_t}_{\text{dynamic part}} + \underbrace{2\int_0^{S_0} \frac{(K - S_T)^+}{K^2}\,\mathrm{d}K}_{\text{static part (puts)}} + \underbrace{2\int_{S_0}^{\infty} \frac{(S_T - K)^+}{K^2}\,\mathrm{d}K}_{\text{static part (calls)}}.$$

Note: *Infinitesimally small* positions in *infinitely many* options.

Semi-static replication of variance swaps (II)

Questions:

- What is the **optimal** semi-static hedge for finite number n (e.g. $n = 30$) of hedging assets?
- How many assets $d < n$ are enough for a 'reasonably small' hedging error?

Semi-static replication of variance swaps (II)

Questions:

- What is the **optimal** semi-static hedge for finite number n (e.g. $n = 30$) of hedging assets?
- How many assets $d < n$ are enough for a 'reasonably small' hedging error?

Remarks:

- Optimality criterion for first question: Variance-optimality.
- Second question: Related to variable selection in high-dimensional regression.

Part II

Variance-Optimal Semi-Static Hedging

(Classic) Variance-optimal hedging

Claim H^0 in $L^2(\mathbb{Q})$, to be hedged with underlying S.
Identify claim H^0 with martingale

$$H^0_t = \mathbb{E}\left[H^0 \middle| \mathcal{F}_t \right], \qquad t \in [0, T]$$

Set of all admissible dynamic strategies:

$$\mathrm{L}^2(S) := \left\{ \vartheta \text{ predictable and } \mathbb{R}\text{-valued: } \mathbb{E}\left[\int_0^T |\vartheta_t|^2 d\langle S, S\rangle_t \right] < +\infty \right\}.$$

Variance-optimal hedging (Föllmer and Schweizer 1986)

Variance-optimal hedge ϑ with initial capital c of claim H^0 is solution of the minimization problem

$$\epsilon^2 = \min_{\vartheta \in \mathrm{L}^2(S), c \in \mathbb{R}} \mathbb{E}\left[\left(c + \int_0^T \vartheta_t \mathrm{d}S_t - H^0_T \right)^2 \right].$$

Galtchouk-Kunita-Watanabe (GKW)-decomposition

The variance-optimal hedging problem is solved by the GKW-decomposition

$$H_t^0 = c + \int_0^t \vartheta_s^0 \mathrm{d}S_s + L_t^0 \,,$$

where L^0 is a local martingale orthogonal to S, i.e. $\langle S, L^0 \rangle = 0$.
Optimal strategy ϑ can be recovered as

$$\langle H^0, S \rangle_t = \int_0^t \vartheta_s^0 \, \mathrm{d}\langle S, S \rangle_s \,,$$

i.e. as the Radon-Nikodym derivative

$$\vartheta_t^0 = \frac{\mathrm{d}\langle H^0, S \rangle_t}{\mathrm{d}\langle S, S \rangle_t}.$$

Semi-static variance-optimal hedging

In addition to H^0: Hedging assets $H = (H^1, \ldots, H^n)$ with associated martingales

$$H^i_t = \mathbb{E}\left[H^i \middle| \mathcal{F}_t \right], \qquad t \in [0, T]$$

Use: dynamic position ϑ in S, static positions $v = (v_1, \ldots, v_n)$ in $H = (H^1, \ldots, H^n)$.

Semi-static variance-optimal hedging

Variance-optimal semi-static hedge $(\vartheta, v) \in \mathrm{L}^2(S) \times \mathbb{R}^n$ with initial capital c of claim H^0 is solution of the minimization problem

$$\epsilon^2 = \min_{(\vartheta,v)\in \mathrm{L}^2(S)\times\mathbb{R}^n, c\in\mathbb{R}} \mathbb{E}\left[\left(c - v^\top \mathbb{E}\left[H_T\right] + \int_0^T \vartheta_t \mathrm{d}S_t - (H^0_T - v^\top H_T) \right)^2 \right].$$

Semi-static variance-optimal hedging (II)

The semi-static problem can be decomposed into inner and outer problem:

$$(\star)\begin{cases}\epsilon^2(v) = \min_{\vartheta\in \mathrm{L}^2(S), c\in\mathbb{R}} \mathbb{E}\left[\left(c - v^\top \mathbb{E}\left[H_T\right] + \int_0^T \vartheta_t \mathrm{d}S_t - (H_T^0 - v^\top H_T)\right)^2\right], \\ \epsilon^2 = \min_{v\in\mathbb{R}^n} \epsilon(v)^2. \end{cases}$$

- Inner problem is a classic variance-optimal hedging problem
- Outer problem is a finite-dimensional quadratic optimization problem

Semi-static variance-optimal hedging (II)

The semi-static problem can be decomposed into inner and outer problem:

$$(\star)\begin{cases}\epsilon^2(v) = \min_{\vartheta \in \mathrm{L}^2(S), c \in \mathbb{R}} \mathbb{E}\left[\left(c - v^\top \mathbb{E}\left[H_T\right] + \int_0^T \vartheta_t \mathrm{d}S_t - (H_T^0 - v^\top H_T)\right)^2\right], \\ \epsilon^2 = \min_{v \in \mathbb{R}^n} \epsilon(v)^2. \end{cases}$$

- Inner problem is a classic variance-optimal hedging problem
- Outer problem is a finite-dimensional quadratic optimization problem

Denote the GKW-decompositions of $(H^1, \ldots, H^n)$ by

$$H_t^i = c + \int_0^t \vartheta_s^i \mathrm{d}S_s + L_t^i$$

and introduce vector notation

$$L_t = (L_t^1, \ldots L_t^n), \quad \vartheta_t = (\vartheta_t^1, \ldots, \vartheta_t^n).$$

Semi-static variance-optimal hedging (III)

Theorem (Di Tella, Haubold and K.-R. (2017))

Consider the variance-optimal semi-static hedging problem and set

$$A := \mathrm{Var}[L^0_T], \qquad B := \mathrm{Cov}[L_T, L^0_T], \qquad C := \mathrm{Cov}[L_T, L_T] \,.$$

Under a non-redundancy condition, C is invertible and the unique solution of the semi-static hedging problem is given by

$$c = \mathbb{E}\left[H^0_T\right], \qquad v = C^{-1}B, \qquad \vartheta^v = \vartheta^0 - v^\top \vartheta.$$

The minimal squared hedging error is given by

$$\varepsilon^2 = A - B^\top C^{-1} B.$$

...

Semi-static variance-optimal hedging (IV)

Theorem (continued.)

Moreover, the elements of A, B and C can be expressed as

$$\mathbb{E}\left[{L^i}_T {L^j}_T\right] = \mathbb{E}\left[\langle H^i, H^j \rangle_T - \int_0^T \vartheta^i_t \vartheta^j_t \,\mathrm{d}\langle S, S\rangle_t\right], \quad i,j = 0,\dots,n\,.$$

Still: Challenging to compute $\mathbb{E}\left[{L^i}_T {L^j}_T\right]$ numerically.

Part IV

Sparse Semi-static hedging

Sparse Semi-static Hedging problem

On $\mathbb{R}^n$ define

- Usual ℓ_1-norm: $\|v\|_1 = \sum_{i=1}^n |v_i|$
- (Non-convex) ℓ_0-quasinorm $\|v\|_0$: counts the number of non-zero elements of v.

Sparse Variance-Optimal Semi-Static Hedging

The *sparse variance-optimal semi-static hedge* $(\vartheta, v) \in \mathrm{L}^2(S) \times \mathbb{R}^n$ with effective portfolio size $d < n$ and its optimal initial capital $c \in \mathbb{R}$ are the solution of the minimization problem $(\star)$, with the outer problem replaced by

$$\epsilon^2 = \min_{v \in \mathbb{R}^n, v \geq 0} (v^\top C v - 2 v^\top B + A), \quad \text{subj. to} \quad \|v\|_0 \leq d.$$

ℓ_1-relaxed problem

ℓ_1-relaxed Sparse Semi-static Hedging Problem

For the ℓ_1-relaxation of the sparse hedging problem the outer problem in $(\star)$ is replaced by

$$\epsilon^2 = \min_{v \in \mathbb{R}^n, v \geq 0} (v^\top C v - 2 v^\top B + A) + \lambda \|v\|_1 ,$$

where $\lambda > 0$ is a tuning parameter that replaces d.

In both problems, we allow for long/short contains of the form

$$p^\top v \geq 0,$$

for some $p \in \mathbb{R}^n$.

Sparse Semi-static Hedging problems (II)

Remarks:

- ℓ_0-problem:
  - Non-convex; hard to solve exactly for large n
  - Equivalent to *variable selection problem* in high-dimensional linear regression
- ℓ_1-problem:
  - Convex, efficient solvers; often a good approximation to ℓ_0-problem
  - Equivalent to *LASSO regression* (Tibshirani (1996))

Sparse Semi-static Hedging problems (III)

Solution/Approximation methods for the ℓ_0-problem:

- **Brute-Force:** Iterate over all $\binom{d}{n}$ subsets of size d; not efficient and completely infeasible for large n.
- **Leaps-and-Bounds:** Branch-and-bound algorithm introduced by Furnival and Wilson (1974). Gives exact solution to ℓ_0-problem without testing all possible subsets.
- **Greedy Forward Selection:** Assume that the optimal subsets of different cardinality are nested. Solve first for $d = 1$ then for $d = 2$, etc. Fast, but no guarantee of being close the exact solution can be given.

Part V

Numerical computation of the GKW-decomposition

Numerical computation of GKW-decomposition

Problem: Given European payoffs $H^i = f^i(S_T)$ with associated martingales H^i_t, how can we numerically compute

$$\vartheta^i, \qquad \text{and} \qquad \mathbb{E}\left[L^i_T L^j_T\right] \quad ?$$

Idea: Combine Fourier representation

$$f^i(u) = \frac{1}{2\pi\,\mathrm{i}} \int_{-\infty}^{+\infty} \exp(ux)\tilde{f}^i(x)\mathrm{d}x$$

with a model S^X where the characteristic function is explicitly known:

- Lévy processes (Hubalek, Kallsen and Krawczyk (2006))
- Affine processes (Kallsen and Pauwels (2006, 2010))
- Semimartingale and Markovian Semimartingale models (Di Tella, Haubold, K.-R. (2017))

Technically challenging!

Numerical computation of GKW-decomposition (II)

Theorem (Di Tella, Haubold and K.-R. (2017))

Consider stoch. vol model $(S = e^X, V)$ *and let* H^0 *be a variance swap and let* $(H^1, \ldots, H^n)$ *be European options with Fourier representations. Assume that* (S, V) *are continuous square-integrable semi-martingales and that there exist functions* $h(u, t, V_t)$, $\gamma(t, V_t)$, *such that*

$$H_t(u) := \mathbb{E}\left[e^{uX_T} \middle| \mathcal{F}_t \right] = e^{uX_t} h(u, T-t, V_t),$$
$$F_t := \mathbb{E}\left[[X,X]_T - [X,X]_t \middle| \mathcal{F}_t \right] = \gamma(T-t, V_t).$$

Then the following holds true:

$$A = \mathbb{E}\left[\mathcal{A}_T\right]$$
$$B_i = \int_{\mathcal{S}(R_i)} \mathbb{E}\left[\mathcal{B}_T(u)\right] \tilde{f}_i(u) \mathrm{d}u,$$
$$C_{ij} = \int_{\mathcal{S}(R_i)} \int_{\mathcal{S}(R_j)} \mathbb{E}\left[\mathcal{C}_T(u_1, u_2)\right] \tilde{f}_i(u_1) \tilde{f}_j(u_2) \mathrm{d}u_1 \mathrm{d}u_2,$$

Numerical computation of GKW-decomposition (III)

Theorem (continued)

The processes $\mathcal{A}, \mathcal{B}, \mathcal{C}$ can be written as

$$\mathrm{d}\mathcal{A}_t = (\partial_v \gamma(T-t, V_t))^2 \, dQ_t,$$
$$\mathrm{d}\mathcal{B}_t(u) = e^{uX_t} \, \partial_v h(u, T-t, V_t) \, \partial_v \gamma(T-t, V_t) \, dQ_t$$
$$\mathrm{d}\mathcal{C}_t(u_1, u_2) = e^{(u_1+u_2)X_t} \, \partial_v h(u_1, T-t, V_t) \, \partial_v h(u_2, T-t, V_t) \, dQ_t$$

where

$$dQ = \mathrm{d}[V, V] - \frac{\mathrm{d}[X, V]}{\mathrm{d}[X, X]} \mathrm{d}[X, V].$$

Part VI

Numerical Results

Model Setup (I)

We consider the Heston model $S = e^X$, where

$$\mathrm{d}X_t = -\frac{1}{2}V_t\mathrm{d}t + \sqrt{V_t}\mathrm{d}W_t^1,$$
$$\mathrm{d}V_t = -\lambda(V_t - \kappa)\mathrm{d}t + \sigma\sqrt{V_t}\mathrm{d}W_t^2,$$

with parameters calibrated in Gatheral (2006):

$$\kappa = 0.0354 \qquad \lambda = 1.3253 \qquad \rho = -0.7165$$
$$\sigma = 0.3877 \qquad V_0 = 0.0174$$

and $S_0 = 100$.

Model Setup (II)

The supplementary assets are 21 OTM-puts and OTM-calls with strikes ranging from

$$K_{\min} = 50 \text{ to } K_{\max} = 150 \text{ in steps of } \Delta K = 5.$$

To be hedged: Variance swap with price given by

$$\mathbb{E}\left[\langle X, X\rangle_T\right] = \int_0^T \mathbb{E}\left[V_t\right] dt = \kappa T + (V_0 - \kappa)\frac{1 - e^{-\lambda T}}{\lambda} = 0.025427.$$

Hedging quality assessed by relative hedging error ϵ / k_* in percentage points.

Goals of the numerical example

Goals:

- Comparing the different methods for sparse semi-static hedging problem:
  - Greedy forward selection (with and without short-sale constraints)
  - Leaps-and-Bounds (with and without short-sale constraints)
  - LASSO
- Analyze dependency of hedging error and portfolio composition on effective portfolio size d;
- Analyze dependency of hedging error and optimal portfolio composition on the leverage parameter ρ.

Results: Relative Hedging Error

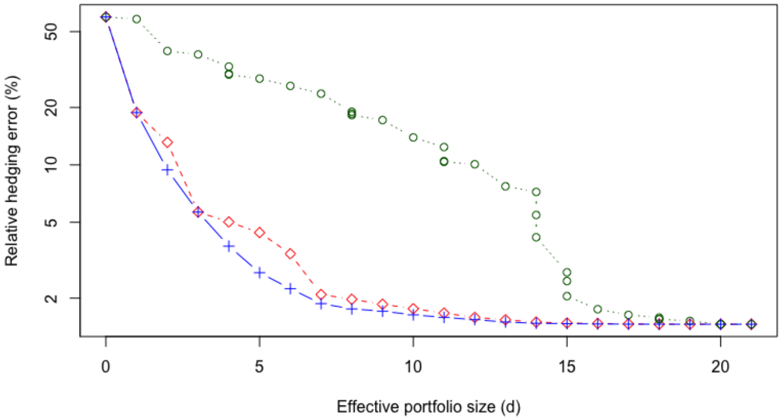

Figure: Shown is relative hedging error (on log-scale) of LASSO (green circles), Greedy forward selection (red diamonds) and Leaps-and-Bounds (blue x's).

Comparison of Methods

Observations:

- LASSO surprisingly bad
- Greedy forward selection very fast and surprisingly good
- Leaps-and-Bounds gives exact solutions; runtime is tolerable but model-sensitive

Reason for LASSO's bad performance: Bad condition of matrix C (reciprocal condition number: 1.11×10^{-6}).
Options with neighboring strikes are highly correlated!

Results: Relative Hedging Error (II)

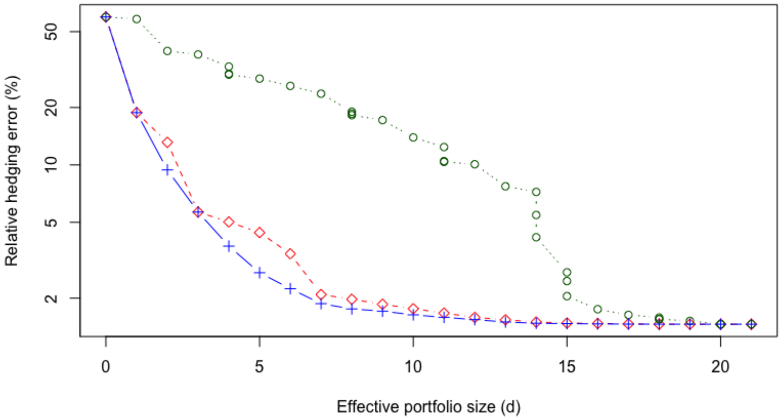

Figure: Shown is relative hedging error (on log-scale) of LASSO (green circles), Greedy forward selection (red diamonds) and Leaps-and-Bounds (blue x's).

Relative Hedging Error vs. Effective Portfolio Size

Reduction of hedging error:

Effective portfolio size	Relative hedging error
$d = 0$ (no static pos.)	59.7 %
$d = 3$	5.7 %
$d = 6$	3.4 %
$d = 21$ (all avail. assets)	1.6 %
$d \to \infty$, $K_{\text{min}} = 0$, $K_{\text{max}} = \infty$	0 %

Portfolio Composition

Focus on optimal portfolio composition and compare:

- Leaps-and-Bounds without short-sale constraints
- Leaps-and-Bounds with short-sale constraints
- Greedy-forward selection with short-sale constraints
- LASSO (without constraints)

Portfolio Composition: Leaps-and-Bounds I

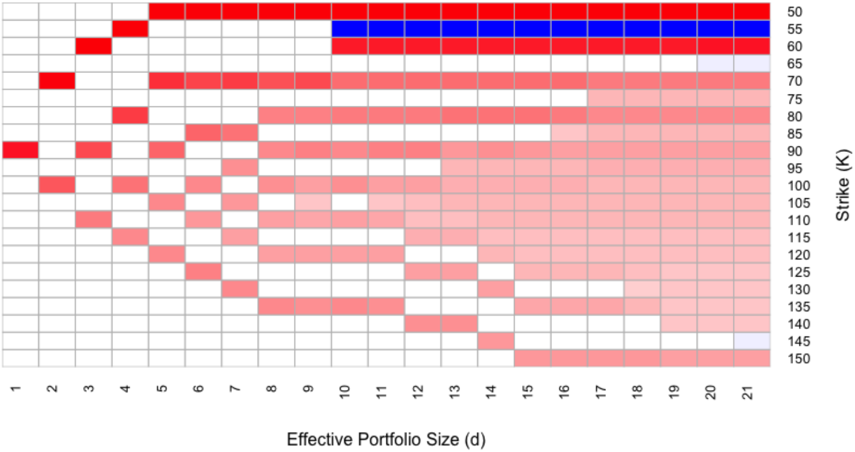

Figure: Leaps-and-Bounds without short-sale constraints

Portfolio Composition: Leaps-and-Bounds II

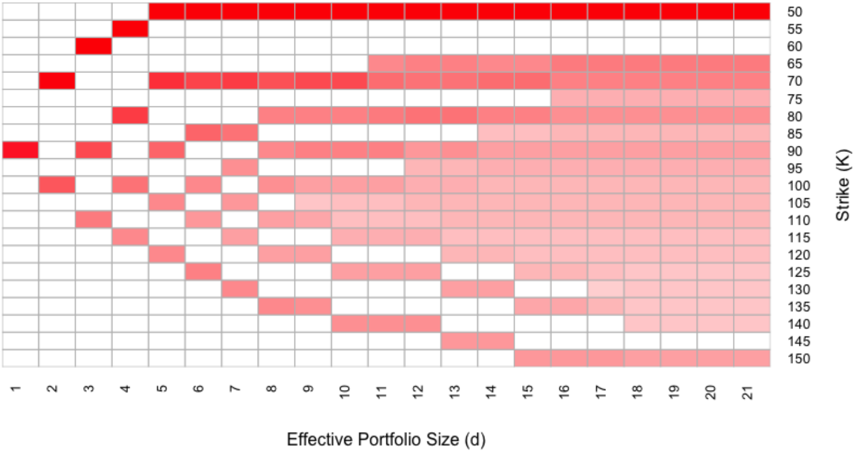

Figure: Leaps-and-Bounds with short-sale constraints

Portfolio Composition: Greedy forward selection

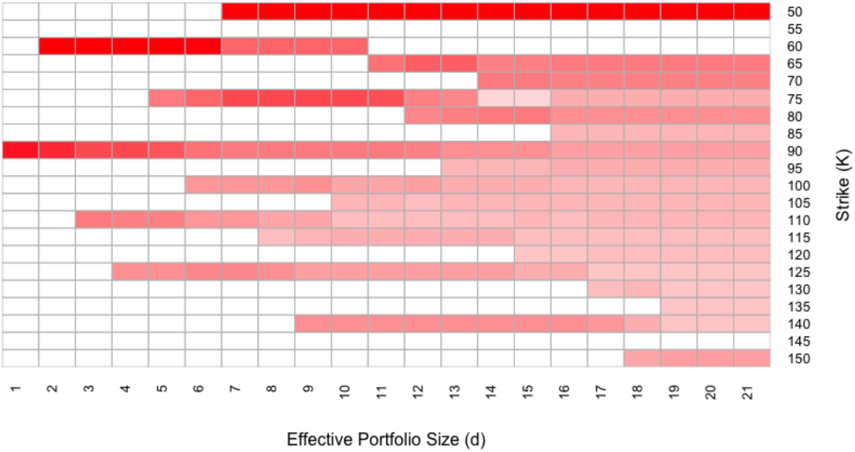

Figure: Greedy forward selection with short-sale constraints

Portfolio Composition: LASSO

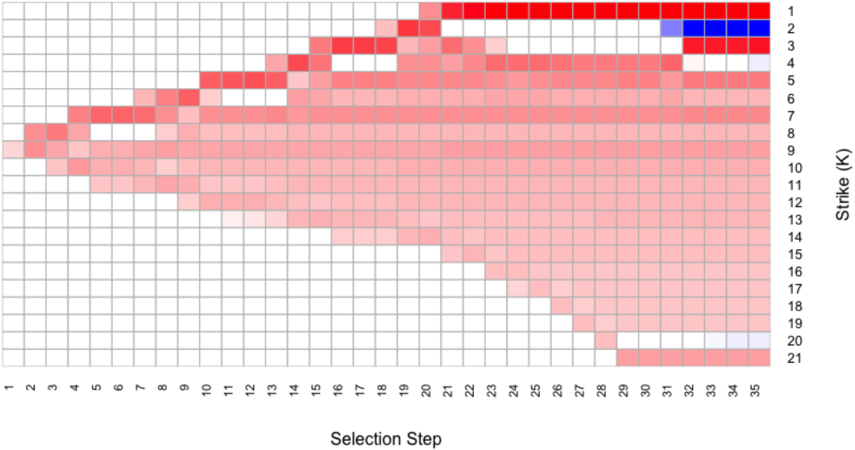

Figure: LASSO (no constraints)

Portfolio Composition (summary)

Observations:

- With the exception of put $K = 55$ only long positions are observed;
- Positions in OTM puts ($K < 100$) are larger than in OTM calls ($K > 100$), in line with Neuberger's replicating portfolio;
- General pattern (going from $d = 1$ to 21):
  - Start with (approximately) ATM option
  - Proceed by selecting both OTM puts and calls, going outwards as d increases until the limit $K_{\text{min}} = 50$ is reached
  - Put more weight on OTM puts ($K < 100$)
  - Continue by adding OTM calls and by filling up the gaps from earlier stages.

Comparison with infinitesimal portfolio (I)

How does the variance-optimal portfolio compare with Neuberger's infinitesimal portfolio?
Neuberger's replicating portfolio places infinitesimal weight of

$$v(K)dK = \frac{1}{K^2}dK$$

on each option with strike K. In doubly logarithmic coordinates, this becomes

$$\log v(K) = -2 \log K,$$

i.e. the portfolio weights should form a line of downward slope -2.
Compare with sparse variance-optimal portfolio of size $d = 3, 6, 12$.

Comparison with infinitesimal portfolio (II)

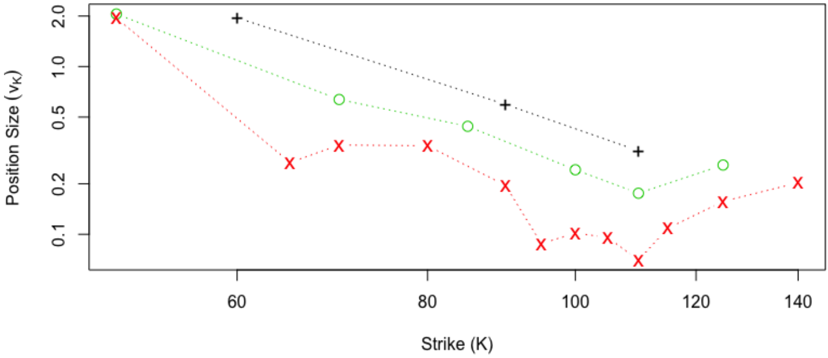

Figure: Portfolio weights v_K in the optimal hedging portfolios of effective size $d = 3$ (black crosses), $d = 6$ (green circles) and $d = 12$ (red x's) in doubly logarithmic coordinates.

Summary

- Framework to solve semi-static hedging problem under variance-optimality criterion
- Sparsity constraints allow to select small subset of suitable hedging instruments
- Tractable implementation in the Heston model
- Under realistic assumptions a variance swap can reasonably hedged with 3 puts/calls (+ underlying)
- Extensions: Other models (non-affine, rough), other payoffs, minimize $\mathbb{P}$-error instead of $\mathbb{Q}$-error, include transaction costs in dynamic strategy?

Thank you for your attention!