Option pricing bottom up and top down

Liuren Wu

IAQF & Thalesians Seminar Series

Centralized *bottom-up* option pricing

- This classic approach is similar to that in classic physics
  - Identify the smallest *common denominator*
    - The risk-neutral dynamics
    - The pricing of Arrow-Debreu securities, characteristic function
  - *Centralize* the valuation of all contracts with the common denominator
    - Take expectation of future payoffs on the same dynamics
    - Price the payoffs as a basket of spanning instruments
- The approach offers *cross-sectional consistency*
  - The common denominator provides a *single yardstick* for valuing all derivatives built on it.
  - The valuations on different contracts are consistent with one another, in the sense that they are all derived from the same yardstick.
  - Even if the yardstick is wrong, the valuations remain" consistent" with one another —*They are just consistently wrong!*

Time-changed Lévy process as an assembly line

- It is an easy-to-use framework to assemble/integrate different parts together:
  - Model each economic shock or the innovation of each factor with a *Lévy process* (X_t^i)
  - Model the stochastic intensity or stochastic impact of each shock with *stochastic time change* ($X_{\mathcal{T}_t}^i$)
- Indulge in a researcher's ambition of building *a model of everything*: World $= \sum X_{\mathcal{T}_t}^i$
  - The key premise: *Everything is related*
  - The key objective: Identify the linkages, consolidate the information
- A good starting point for financing engineering of real applications
  - Build market making/valuation platforms
  - Extract key information from diverse sources
  - Understand information transmission mechanisms
  - Identify (cross-)market arbitrage opportunities

A foolish consistency is the hobgoblin of little minds ...

- It is difficult to get everything under one blanket
  - Pricing long-dated contracts requires unrealistically long projections.
  - Short-term variations of long-dated contracts often look incompatible with long-run (stationarity) assumptions.
- It is not always desirable to *chain* everything together
  - Error-contagion: Disturbance on one contract affects everything else.
  - Pooling is useful for information consolidation, but can be limiting for individual contract pricing/investment, which needs domain expertise.
    - *Jack of all trades, master of none.*
- Forcing "consistency" with a mis-specified model often leads to the *divergence and disengagement* between pricing and risk management:
  - The force-fitted "dynamics" look nothing like actual behaviors.
  - While pricing team works on a sophisticated model, risk management relies on simple but more robust BMS greeks.

Decentralized *top-down* option pricing

The approach starts top down with *a particular option contract* i,
not bottom up from some common denominator (dynamics):

1. *Represent* the value of the option contract via the Black-Merton-Scholes (BMS) pricing equation, $B^i(t, S_{t,i}, I_{t,i})$
   - BMS was also a bottom-up option pricing model, but we are not using it as a bottom-up model, but just as a *value representation* function.
   - The representation *does not decompose* the price into building blocks
   - but *summarizes/characterizes* its main sources of variation in *observable variables* $(S_{t,i}, I_{t,i})$
     — "symptoms" of a patient (medicine)
     — "features" of an object (machine learning)

The top-down P&L attribution

2. Perform *P&L attribution* against the observable variation sources (not against hidden factors/state variables):

$$
\begin{aligned}
dB^i\left(t, S_{t,i}, I_{t,i}\right) &= \left[B^i_t dt + B^i_S dS_{t,i} + B^i_I dI_{t,i}\right] \\
&+ \left[\frac{1}{2} B^i_{SS}\left(dS_{t,i}\right)^2 + \frac{1}{2} B^i_{II}\left(dI_{t,i}\right)^2 + B^i_{IS}\left(dS_{t,i} dI_{t,i}\right)\right] \cdots
\end{aligned}
$$

- Work in forward space to hide financing (rates)
- Normalize $S_{t,i} = 1$ to turn the excess P&L into *excess return*.
- Stop at 2nd order when the *next moves* on $S_{t,i}$ and $I_{t,i}$ are diffusive.
- Adding higher order price changes can potentially be useful to differentiate different types of skews/smiles, subject to identification...
- I use continuous time notation, but it can be just a Taylor expansion of discrete returns over some horizon...

The no dynamic arbitrage condition

3. Take risk-neutral expectation and apply the *no-dynamic-arbitrage condition* (NDA) that the expected risk-neutral excess return be 0:

$$-B_t^i = \frac{1}{2} B_{SS}^i S_t^2 v_{t,i} + B_I^i I_{t,i} \mu_{t,i} + \frac{1}{2} B_{II}^i I_{t,i}^2 \omega_{t,i} + B_{IS}^i I_{t,i} S_t \gamma_{t,i} \qquad (1)$$

- $(\mu_t^i, \sigma_t^2, \omega_t^i, \gamma_t^i)$ are the *time-t conditional risk-neutral moments*:

$$\mu_{t,i} \equiv \frac{1}{dt} \mathbb{E}_t^{\mathbb{Q}} \left[\frac{dI_{t,i}}{I_{t,i}} \right] /, \quad v_{t,i} \equiv \frac{1}{dt} \mathbb{E}_t^{\mathbb{Q}} \left[\left(\frac{dS_{t,i}}{S_{t,i}} \right)^2 \right],$$

$$\omega_{t,i} \equiv \frac{1}{dt} \mathbb{E}_t^{\mathbb{Q}} \left[\left(\frac{dI_{t,i}}{I_{t,i}} \right)^2 \right], \quad \gamma_{t,i} \equiv \frac{1}{dt} \mathbb{E}_t^{\mathbb{Q}} \left[\left(\frac{dS_{t,i}}{S_{t,i}}, \frac{dI_{t,i}}{I_{t,i}} \right) \right].$$

- (1) looks similar to a PDE from a bottom-up model, except
  - A bottom-up model solves the unknown value function from the PDE.
  - In (1), we start with a known BMS representation, with known sensitivities, for a given contract i.
- (1) is a *pricing equation* in terms of the *expected behaviors* $(\mu_{t,i}, v_{t,i}, \omega_{t,i}, \gamma_{t,i})$ of the observable risk sources $(S_{t,i}, I_{t,i})$
- NDA requires the contract be priced to balance its time decay with its risk-neutral expected excess returns from the 4 risk sources.

Breakeven contribution of each risk source

- Taking *statistical expectation*, we decompose the expected excess return (EER) into the summation of risk exposures × risk magnitudes

$$\text{EER} = B_t^i + \frac{1}{2}B_{SS}^i S_t^2 v_{t,i}^{\mathbb{P}} + B_I^i I_{t,i}\mu_{i,t}^{\mathbb{P}} + \frac{1}{2}B_{II}^i I_{t,i}^2 \omega_{t,i}^{\mathbb{P}} + B_{IS}^i I_{t,i} S_t \gamma_{t,i}^{\mathbb{P}}$$

  - If EER=0, we expect to *breakeven* on the contract.
    The *breakeven contribution* from each source:

$$\left(B_I^i I_{t,i}\mu_{i,t}^{\mathbb{P}}, \quad , \frac{1}{2}B_{SS}^i S_t^2 v_{t,i}^{\mathbb{P}}, \quad \frac{1}{2}B_{II}^i I_{t,i}^2 \omega_{t,i}^{\mathbb{P}}, \quad B_{IS}^i I_{t,i} S_t \gamma_{t,i}^{\mathbb{P}} \right)$$

  - When EER is non-zero, the risk premium can also be attributed to each risk source

$$\begin{aligned} \text{EER} &\equiv \tfrac{1}{dt}(\mathbb{E}^{\mathbb{P}} - \mathbb{E}^{\mathbb{P}}](dB^i) \quad \text{(risk premium)} \\ &= \tfrac{1}{2}B_{SS}^i S_t^2 \left(v_{t,i}^{\mathbb{P}} - v_{t,i}\right) + B_I^i I_{t,i}\left(\mu_{i,t}^{\mathbb{P}} - \mu_{t,i}\right) \\ &+ \tfrac{1}{2}B_{II}^i I_{t,i}^2 \left(\omega_{t,i}^{\mathbb{P}} - \omega_{t,i}\right) + B_{IS}^i I_{t,i} S_t \left(\gamma_{t,i}^{\mathbb{P}} - \gamma_{t,i}\right) \end{aligned}$$

It all goes from here

How to apply this decentralized top-down pricing relation in practice?

$$-B_t^i = \frac{1}{2}B_{SS}^i S_t^2 v_{t,i} + B_I^i I_{t,i}\mu_{t,i} + \frac{1}{2}B_{II}^i I_{t,i}^2 \omega_{t,i} + B_{IS}^i I_{t,i} S_t \gamma_{t,i}$$

- Like machine learning, the pricing relation uses a very simple "basis function" (the BMS function), but achieves great flexibility via localization.
- Too much flexibility: The time decay of a contract can be balanced by 4 sources of expected gains.
- Historical centralizing efforts fall on old habits:
  Impose *common risk-neutral implied vol dynamics* across contracts
  - Carr & Wu (2016): across the whole implied vol surface
  - Arsland et al (2009), Carr, Wu, Zhang (2022): per maturity
  - Carr & Wu (2020): nearby contracts

All men are not born equal

We unite the decentralized pricing not by assuming common dynamics, but by

1. *explicitly recognizing their behavior difference*
   - Use historical moment estimators as model inputs $(\widehat{\mu}_{t,i}, \widehat{v}_{t,i}, \widehat{\omega}_{t,i}, \widehat{\gamma}_{t,i})$
   - anchor the *breakeven contribution* of each risk source $(B_I^i I_{t,i}\widehat{\mu}_{t,i}, \quad \frac{1}{2}B_{SS}^i S_{t,i}^2 \widehat{v}_{t,i}, \quad \frac{1}{2}B_{II}^i I_{t,i}^2 \widehat{\omega}_{t,i}, \quad B_{IS}^i I_{t,i} S_{t,i} \widehat{\gamma}_{t,i})$
2. *pricing their differences fairly*
   - in accordance (in proportion) to the different risk levels
     - *Supply-demand* determines the pricing of a single product — Investors can add a *premium/discount* in the pricing on each risk source.
     - *Statistical arbitrage* trading across contracts can force the pricing on *each risk type/source* to converge in the limit (Ross, 76).
   - Devils are in the details — how to be fair...

Different but fair: A linear option pricing model

$$
\begin{aligned}
-B_t^i = \; & \left(\beta_{V,t}\widehat{\omega}_{t,i} + \beta_{\mu,t}\widehat{\mu}_{t,i}\right)\left[B_I^i I_{t,i}\right] \\
& + \beta_{G,t}\widehat{v}_{t,i}\left[\tfrac{1}{2}B_{SS}^i S_{t,i}^2\right] + \beta_{O,t}\widehat{\omega}_{t,i}\left[\tfrac{1}{2}B_{II}^i I_{t,i}^2\right] + \beta_{A,t}\sqrt{\widehat{v}_{t,i}\widehat{\omega}_{t,i}}\left[B_{IS}^i S_{t,i} I_{t,i}\right]
\end{aligned}
$$

- Sensitivities $\times$ historical moment estimators capture actual/historical *behavior differences* across contracts
- Common coefficients β impose *fair pricing* in proportion to risk levels:
  1. Volatility risk premium: $\mu_{t,i} = \mu_{t,i}^{\mathbb{P}} + \beta_{V,t}\widehat{\omega}_{t,i}$ for all i (proportional to risk)
  2. Statistical trend forecast: $\mu_{t,i}^{\mathbb{P}} = \beta_{\mu,t}\widehat{\mu}_{t,i}$ for all i (momentum/reversal)
  3. Risk-neutral return variance: $v_t = \beta_{G,t}\widehat{v}_{t,i}$ for all i (vary around historical estimators)
  4. Risk-neutral implied vol change variance: $\omega_{t,i} = \beta_{O,t}\widehat{\omega}_{t,i}$ for all i (vary around historical estimators)
  5. Risk-neutral covariance: $\gamma_{t,i} = \beta_{A,t}\sqrt{\widehat{v}_{t,i}\widehat{\omega}_{t,i}}$ for all i (proportional to historical variance)

Alternative representation in the implied volatility space

- The linear model is in the *expected excess return* space. People tend to be more familiar with the implied vol surface behavior.

- For vanilla options, risk exposures are all proportional to cash gamma.Divide the relation by cash gamma leads to an implied volatility representation:

$$\begin{aligned} I_{t,i}^2 &= \beta_{G,t}\widehat{v}_{t,i} \\ &+ 2\left(\beta_{V,t}\widehat{\omega}_{t,i} + \beta_{\mu,t}\widehat{\mu}_{t,i}\right) I_{t,i}^2\,\tau \\ &+ \beta_{O,t}\widehat{\omega}_{t,i}(z_+ z_-) + 2\beta_{A,t}\sqrt{\widehat{v}_{t,i}\widehat{\omega}_{t,i}}\,(z_+) \end{aligned}$$

  - Implied variance = risk-neutral return variance at short-term $\tau \downarrow 0$ and at-the-money $z_+ = 0$, ... plus adjustments
  - along maturity for volatility risk premium and trend,
  - along moneyness ($z_\pm = \ln K/S_t \pm \frac{1}{2} I_t^2 \tau$) for implied vol variation (quadratic in $z_+ z_-$) and covariation (linear in z_+)

- The common pricing coefficients β_t generate *broad, overall* adjustments on the implied volatility *level, slopes, and curvature*.

- The moment estimators $(\widehat{\omega}_{t,i}, \widehat{\mu}_{t,i})$ are the key drivers on the fine details of the implied volatility (term structure, skew/smile) shape.

Linear regression pricing of options

$$y_{t,i} = X_{t,i}\mathbf{b}_t + e_{t,i}$$

- The dependent variable is the time decay, $y_{t,i} = -B_t^i$.
- The regressors are historical risk estimator-adjusted BMS sensitivities: $X_{t,i} =$
$\left[\begin{array}{ccccc} B_I^i I_{t,i}\widehat{\omega}_{t,i}, & B_I^i I_{t,i}\widehat{\mu}_{t,i}, & \frac{1}{2}B_{SS}^i S_{t,i}^2 \widehat{v}_{t,i}, & \frac{1}{2}B_{II}^i I_{t,i}^2 \widehat{\omega}_{t,i}, & B_{IS}^i S_{t,i} I_{t,i}\sqrt{\widehat{v}_{t,i}\widehat{\omega}_{t,i}} \end{array}\right]$.
- The slope estimates at each date $\mathbf{b}_t = [\beta_{V,t}, \beta_{\mu,t}, \beta_{G,t}, \beta_{O,t}, \beta_{A,t}]$ — the *common* coefficients across all contracts on that date.
- Under the null of no premium, no prediction, and no forecasting bias, the *null breakeven values* are $\mathbf{b}_t^0 = \left[\begin{array}{ccccc} 0 & 0 & 1 & 1 & \widehat{\rho}_t \end{array}\right]$.
- The time-t *common market pricing* of risk:

$$\eta_t = \mathbf{b}_t - \mathbf{b}_t^0 = [\beta_{V,t}, \quad \beta_{\mu,t}, \quad \beta_{G,t} - 1, \quad \beta_{O,t} - 1, \quad \beta_{A,t} - \widehat{\rho}_t]$$

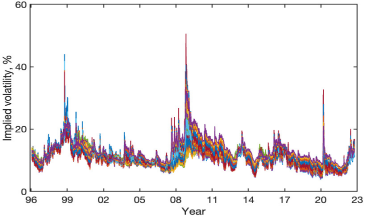
- The regression residual $e_{t,i}$ captures the relative mispricing on each contract in the expected excess return space.

Empirical analysis on currency options

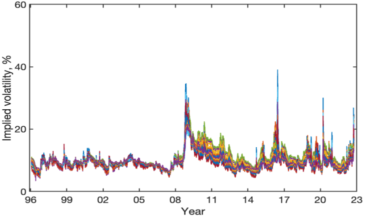
OTC option implied volatility quotes on dollar prices of *yen and pound*

- 1996/1/24 — 2022/10/13, 6,702 business days
- 5 delta (10, 25, 50, 75, 90) $\times$ 5 maturities (1, 3, 6, 9, 12 months)

yen



pound

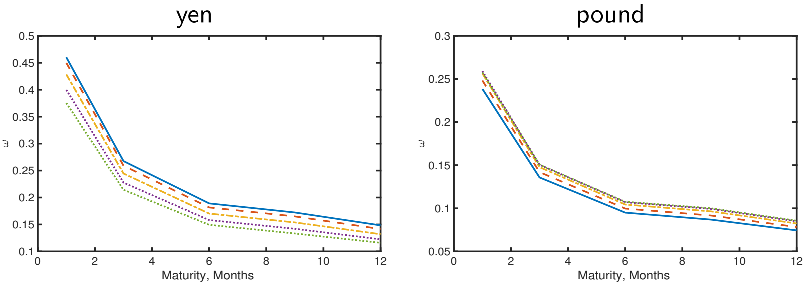


- Volatilities tend to spike up, and then calm down slowly.
- Yen option implied volatilities are on average higher, and show larger cross-contract dispersion.
- The level and dispersion of pound option implied volatilities picked up in the second half.

Construct historical moment estimators

- Create *daily* percentage changes on *fixed-contract* forward price and implied volatility $(dS_{t,i}, dI_{t,i})$ from *floating* series
  - Need interpolation ... more accurate over short-term change
- Construct historical moment estimators with horizons of 1,3,12 months.
- Cross-contract smoothing/averaging: Implied vol surface should be smooth, so should their moment estimators.
  - Moment estimators are embedded in the regressors. Large estimation errors can cause identification issues.
- Under different market conditions, future moments are predicted the best with historical estimators of different horizons.
  - *Market knows the best*: At each date, pick the horizon with the best pricing performance — mainly to better accommodate market expectation of the *term structure* variation.
  - If you have better forecasts (better methods), we can use them.

The average implied volatility surface variation ($\widehat{\omega}$)



- The different lines in each panel are for different delta (moneyness).
- At the same maturity, the variance estimates are similar across delta.
- The average variance of implied vol change $\widehat{\omega}$ declines strongly with maturity, with a strong curvature.
  - This $\widehat{\omega}$ term structure pattern determines the shape of the implied volatility term structure via $\left(\beta_{V,t}\widehat{\omega}_{t,i} + \beta_{\mu,t}\widehat{\mu}_{t,i}\right) I_{t,i}^2\,\tau$
  - It also contributes to the shape of the smile/skew term structure shape via $\beta_{O,t}\widehat{\omega}_t^i(z_+z_-)$ and $\beta_{A,t}\sqrt{\widehat{\omega}_t^i\widehat{v}_{t,i}}(z_+)$

Pricing performance

- Daily cross-sectional OLS regressions on the pricing model, from 1997 to 2022, for 6,450 business days
- Median percentage variance contribution from each risk source: $\text{VC}_{k,t} = [\mathbf{b}_t]_k \, [(X_t^\top X_t)\mathbf{b}_t]_k / (\mathbf{y}_t^\top \mathbf{y}_t)$, sum to R^2

	yen	pound
Gamma	81.95	79.40
Vega	21.58	24.79
Trend	-7.71	-6.18
Volga	4.28	3.43
Vanna	0.58	-0.15
R^2	**99.91**	**99.94**

- Gamma has the largest variance contribution, followed by vega
- The model explains the cross-sectional variation of expected option excess return extremely well, with median $R^2 > 99.9\%$.
- Pricing performance is better than most jump-diffusion stochastic vol models...

Risk-targeting portfolios

- The linear model structure makes it extremely simple to create option portfolios to either *hedge or target* a specific risk exposure.

- Let H_t denote is the $(N \times 6)$ matrix of *risk-adjusted* exposures and errors,

$$H_t \equiv \left[\begin{array}{cc} X_t & \mathbf{e}_t \end{array}\right]$$

  - BMS greeks tend to be more stable than exposure estimates from a complicated jump-diffusion stochastic vol model...
  - X_t adjusts BMS greeks with observed risk magnitudes for each contract — works well for actual risk management

- Minimize portfolio residual return variance, subject to the targeted exposure

$$\min_{\mathbf{w}_t} \quad \mathbf{w}_t^\top \Sigma_t \mathbf{w}_t \quad \text{s.t.} \quad H_t^\top \mathbf{w}_t = \mathbf{d},$$

$$\Rightarrow \mathbf{w}_t = \Sigma_t^{-1} H_t \left(H_t^\top \Sigma_t^{-1} H_t\right)^{-1} \mathbf{d}.$$

  - Assume $\Sigma_t = \sigma_t^2 I$ on the errors, we have $\mathbf{w}_t = H_t \left(H_t^\top H_t\right)^{-1} \mathbf{d}$.

Statistical arbitrage trading against regression residuals

- Setting $\mathbf{d} = \mathbf{1}_6$ constructs a portfolio that is neutral to all risk exposures, but with a unit short exposure to the pricing error.
- The portfolio amounts to sell(buy) over(under)-priced options while maintaining neutral to all risk exposures.
- At each date, we apply the portfolio weight to the delta-hedged excess returns *over the next business day* from shorting each contract, $r_{t+1,6} = \mathbf{w}_{t,6}^\top \mathbf{r}_{t+1}$.

Statistics	A. Fixed error		B. Fixed notional	
	yen	pound	yen	pound
Mean	4.48	5.65	0.56	0.44
Std	0.81	0.88	0.11	0.08
IR	5.56	6.43	5.00	5.81

- The high IR suggests that the residuals are highly reverting, validating the statistical arbitrage theory.

Predict risk portfolio returns with market pricing estimates

- Setting $\mathbf{d} = \mathbf{I}_k$, for $k = 1, \cdots, 5$ generates risk-targeting portfolios for each of the risk exposures.
- Market pricing estimates η_t reflect whether the market is charging a premium or discount on the risk dimension relative to the breakeven contribution.
- We expect the market pricing estimates to predict the ex post realized excess returns on the corresponding risk portfolio:

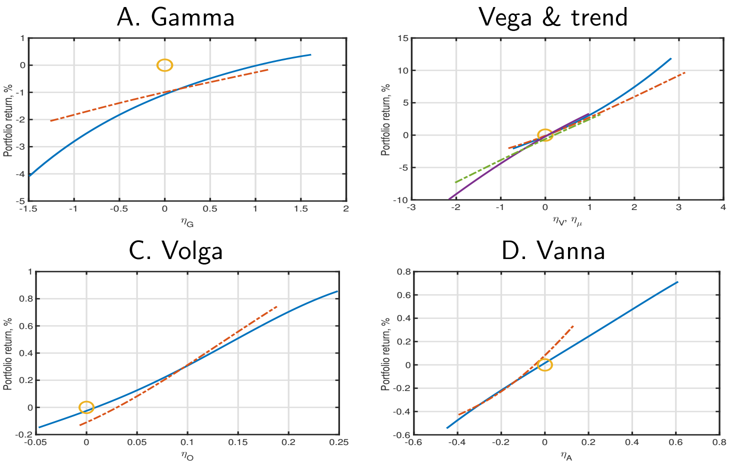
$$r_{t+1,k} = a_k + b_k \eta_{t,k} + e_{t+1,k}.$$

with the null: $a_k = 0$, $b_k > 0$.

	yen					pound				
	a_k		b_k		R^2	a_k		b_k		R^2
Gamma	-1.76	(-5.67)	1.31	(6.60)	1.00	-1.00	(-5.97)	1.50	(13.45)	2.47
Vega	-1.08	(-1.63)	5.24	(25.97)	9.32	0.03	(0.06)	3.37	(19.97)	4.61
Trend	0.34	(0.65)	5.39	(28.60)	11.27	-0.68	(-1.46)	3.20	(18.04)	4.15
Vanna	-0.03	(-0.59)	1.02	(10.80)	1.78	0.13	(3.30)	1.45	(11.35)	1.98
Volga	0.02	(0.38)	2.56	(10.47)	1.55	-0.00	(-0.09)	3.49	(18.68)	3.30

Slopes are strongly positive for all risk portfolios (t-values in parentheses). Intercepts are mostly insignificant, with a few exceptions.

Visualize the predictive relations



- Market pricing estimates on all risk exposures *positively* predict future excess returns of the corresponding risk-targeting portfolio, with little bias.
  - The historical moment estimators create useful *breakeven anchors* for identifying the market pricing on each risk source.

Harvesting the time-varying risk premium

based on the market pricing estimates

- The market pricing estimate on each risk source represents the excess return investors *expect* to make on the risk-targeting portfolio.
  - We can decide which direction to take on each portfolio at each date based on the sign/size of the market pricing estimate.
- We can harvest the time-varying risk premium on each risk source by timing the risk portfolio investment based on the market pricing estimates.
  - Normalize the risk-target portfolio and the market pricing estimates to fixed notional amount ($100).
  - Set the median notional (absolute) weight to $100 corresponding to the median value of the absolute market pricing estimates.
  - Set the notional weight at each date proportional to the normalized market pricing estimate: $n_{k,t} = \eta_{k,t}/\overline{|\eta|}_{k,t}$.
  - Limit the investment to be within ± 2 times the median notional amount.
  - No forecasting regression involved in this exercise.

Harvesting risk premiums from risk-targeting portfolios

Risk Target	A. Fixed notional short			B. Market timing		
	Mean	Std	IR	Mean	Std	IR
			yen			
Vega	0.29	0.24	1.20	0.80	0.39	*2.08*
Trend	-0.14	0.19	-0.74	0.77	0.34	*2.26*
Gamma	-0.34	0.24	-1.42	0.34	0.38	*0.91*
Volga	0.20	0.20	1.00	0.51	0.31	*1.65*
Vanna	0.07	0.20	0.35	0.49	0.33	*1.48*
			pound			
Vega	0.17	0.16	1.10	0.47	0.24	*1.97*
Trend	0.01	0.13	0.05	0.44	0.21	*2.05*
Gamma	-0.24	0.16	-1.52	0.28	0.26	*1.05*
Volga	0.24	0.13	1.91	0.43	0.18	*2.34*
Vanna	-0.05	0.13	-0.39	0.32	0.20	*1.62*

- Market timing turns all risk portfolios to profitability, with good IR.
- Stat arb trading on pricing errors and risk-premium harvesting on market pricing estimates can both be highly profitable.

Concluding remarks

- History repeats itself in cycles of centralization for consolidation and decentralization for flexibility.
- We build a bridge between the two and centralize the decentralized pricing of option contracts via sequential application of two arbitrage arguments:
  - No dynamic arbitrage between an option and its underlying $\rightarrow$ a decentralized pricing relation on a single option contract
  - No statistical arbitrage across option contracts on a given risk type $\rightarrow$ unite the decentralized pricing of all contracts with common market pricing for each decentralized risk source
- The framework captures the strength of both:
  - *Decentralization*: Unique features of different contracts, domain expertise of different investors
  - *Centralization*: Not by disregarding dynamics differences, but by imposing pricing consistency
- These efforts lead to a simple linear option pricing relation that can both *price* and *risk manage* the variation of vanilla/exotic option contracts.