Fully flexible views: theory and practice

Attilio Meucci *proposes a unified methodology to input non-linear views from any number of users in fully general non-normal markets and perform, among others, stress testing, scenario analysis and ranking allocation. He walks the reader through the theory and details an efficient algorithm to easily implement this methodology under fully general assumptions*

Scenario analysis allows a practitioner to explore the implications for a given portfolio of a set of subjective views on possible market realisations (see, for example, Mina & Xiao, 2001). The ground-breaking approach pioneered by Black & Litterman (1990) generalises scenario analysis by adding uncertainty about the views and the reference risk model. Further generalisations have been proposed in recent years. Qian & Gorman (2001) provide a framework to stress test volatilities and correlations in addition to expectations. Pezier (2007) processes partial views on expectations and covariances based on least discrimination. Meucci (2008b) extends the above models to act on risk factors instead of returns, and thus covers highly non-linear derivatives markets and views about external factors that influence the profit and loss only statistically.

In the above techniques, the reference distribution of the risk factors is normal. The copula-opinion pooling (COP) in Meucci (2006) explores non-normal markets, but correlation stress testing and non-linear views are not allowed. Furthermore, the COP relies on *ad hoc* manipulations.

Here we present the entropy pooling (EP) approach, which fully generalises the above and related techniques. The inputs are an arbitrary market model, which we call 'prior', and fully general views or stress tests on that market. The output is a distribution, which we call 'posterior', that incorporates all the inputs and can be used for risk management and portfolio optimisation.

To obtain the posterior, we interpret the views as statements that distort the prior distribution in such a way that the least possible amount of spurious structure is imposed. The natural index for the structure of a distribution is its entropy. Therefore we define the posterior distribution as the one that minimises the entropy relative to the prior distribution. Then, by opinion pooling, we assign different confidence levels to different views and users.

Among others, the EP approach handles: non-normal markets; views on non-linear combinations of risk factors that affect the profit and loss directly or only statistically through correlations; views on expectations, but also medians, to handle fat tails; views on volatilities, correlations, tail behaviours, etc; lax views, such as ranking, on all of the above, thereby generalising Almgren & Chriss (2006); inputs from multiple users; and multiple confidence levels for different views.

Furthermore, in its most general implementation, the reference model is represented by Monte Carlo simulations, and the posterior that incorporates all the inputs is represented by the same simulations with new probabilities. Hence, the most complex securities can be handled without costly repricing.

The remainder of this article is organised as follows. We introduce the EP theoretical framework. Then we present an analytical formula, which generalises the previous results and provides a benchmark for the numerical implementation. Next, we discuss the numerical routine to implement the EP approach in full generality. Finally, we illustrate a case study: option trading in a non-normal environment with non-linear and ranking views on realised volatility, implied volatility and external macro factors. We then conclude, comparing the EP appoach with other related techniques.

Fully documented code for this and other case studies, such as portfolios from ranking, can be downloaded from www.symmys.com > Teaching > Matlab. The technical appendix referenced in this article is available at www.symmys.com > Research > Working Papers.

The entropy pooling approach

We consider a book driven by an N-dimensional vector of risk factors $\mathbf{X}$. In other words, denoting by t the current time, by $\mathcal{I}_t$ the information currently available and by τ the time to the investment horizon, there exists a deterministic function P that maps the realisations of $\mathbf{X}$ and the information $\mathcal{I}_t$ on to the price $P_{t+\tau}$ of each security in the book at the horizon:

$$P_{t+\tau} \equiv P(\mathbf{X}, \mathcal{I}_t) \quad (1)$$

This framework is completely general. For instance, in a book of options, $\mathbf{X}$ can represent the changes in all the underlyings and implied volatilities. In this case (1) is approximated by a second-order Taylor expansion whose coefficients are the deltas, vegas, gammas, vannas, volgas, etc. Also, $\mathbf{X}$ can represent a set of risk factors behind a computationally expensive full Monte Carlo pricing function, such as interest rate values at different monitoring times for mortgage derivatives. Furthermore, $\mathbf{X}$ can be augmented with a set of external risk factors that do not feed directly into the pricing function (1), but that still influence the profit and loss statistically through correlation. We explore a detailed example of these directions below. In any case, we emphasise that $\mathbf{X}$ can be, but is by no means restricted to, returns on a set of securities.

The reference model. We assume the existence of a risk model, that is, a model for the joint distribution of the risk factors, as represented by its probability density function:

$$\mathbf{X} \sim f_{\mathbf{X}} \tag{2}$$

In Black & Litterman (1990), this is the 'prior' factor distribution. More generally, this is a model that risk managers use to perform risk analyses, such as the computation of the volatility, tracking error, value-at-risk and expected shortfall of a portfolio, along with the contributions to such measures from the different sources of risk. Portfolio managers and traders, on the other hand, use this model to optimise their positions. They specify a subjective index of satisfaction $\mathcal{S}$, such as the mean-(conditional) VAR trade-off, or the certainty equivalent stemming from a utility function, or a spectral measure, etc (see examples in Meucci, 2005). Satisfaction depends both on the market distribution $f_{\mathbf{X}}$ through the prices (1) and on the positions in the book, represented by a vector $\mathbf{w}$. Then the optimal book $\mathbf{w}^*$ is defined as:

$$\mathbf{w}^* \equiv \underset{\mathbf{w}\in\mathcal{C}}{\arg\max}\left\{\mathcal{S}\left(\mathbf{w}; f_{\mathbf{X}}\right)\right\} \tag{3}$$

where $\mathcal{C}$ is a given set of investment constraints. The reference model (2) can be estimated from historical analysis, or calibrated to current market observables (see Meucci, 2008b).

The views. In the most general case, the user expresses views on generic functions of the market $g_1(\mathbf{X}), \dots, g_K(\mathbf{X})$. These functions constitute a K-dimensional random variable whose joint distribution is implied by the reference model (2):

$$\mathbf{V} \equiv \mathbf{g}(\mathbf{X}) \sim f_{\mathbf{V}} \tag{4}$$

We emphasise that, unlike in Black & Litterman, in EP we do not assume that the functions g_k are linear. Notice that, as a special case, one can also express views on the securities' values (1).

The views, or the stress tests, are statements on the variables (4) that can clash with the reference model. In a stochastic environment, this means statements on their distribution. Therefore, the most detailed possible view specification is a complete, subjective joint distribution for those variables:

$$\mathbf{V} \sim \tilde{f}_{\mathbf{V}} \neq f_{\mathbf{V}} \tag{5}$$

However, views in general are statements only on select features of the distribution of $\mathbf{V}$.

- The classical views in the manner of Black & Litterman are statements on $\tilde{\mathbb{E}}\{V_k\}$, the expectations of each of the V_ks according to the new distribution $\tilde{f}_{\mathbf{V}}$. Since for distributions such as stable distributions the expectation is not defined, in the EP approach we consider views about a more general location measure $\tilde{m}\{V_k\}$, which can be the expectation or the median. The views are then set as:

$$\tilde{m}\left\{V_k\right\} \gtreqless m_k, \qquad k = 1,\dots,K \tag{6}$$

The values m_k can be determined exogenously. If the user has only qualitative views, it is convenient to set as in Meucci (2008a):

$$m_k \equiv m\left\{V_k\right\} + \xi\sigma\left\{V_k\right\} \tag{7}$$

In this expression σ is a measure of volatility in the reference model, such as the standard deviation or, in fat-tailed markets with infinite variance, the interquartile range, and ξ is an *ad hoc* multiplier, such as –2, –1, 1 and 2 for 'very bearish', 'bearish', 'bullish' and 'very bullish', respectively.

- The generalised Black & Litterman views (6) are not necessarily expressed as equality constraints: EP can process views expressed as inequalities. In particular, EP can process ordering information, frequent in equity and in bond management:

$$\tilde{m}\left\{V_1\right\} \geq \tilde{m}\left\{V_2\right\} \geq \cdots \geq \tilde{m}\left\{V_K\right\} \tag{8}$$

- Views can be expressed on the volatilities. A convenient formulation reads:

$$\tilde{\sigma}\left\{V_k\right\} \gtreqless \xi\sigma\left\{V_k\right\}, \qquad k = 1,\dots,K \tag{9}$$

- Correlation stress tests are also views. Convenient specifications for the correlation matrix $\tilde{\mathbb{C}}\{\mathbf{V}\}$ are the homogeneous shrinkage:

$$\tilde{\mathbb{C}}\left\{\mathbf{V}\right\} \equiv \rho_1 \mathbf{I} + \rho_2 \mathbb{C}\left\{\mathbf{V}\right\} + \rho_3 \mathbf{1}\mathbf{1}' \tag{10}$$

where $0 \leq \rho_1, \rho_2, \rho_3 < 1$, $\rho_1 + \rho_2 + \rho_3 \equiv 1$, $\mathbf{I}$ is the identity matrix and $\mathbf{1}$ is a vector of ones. For different structures, see, for example, Brigo & Mercurio (2001).

- The user can input views on the lower (upper) tail behaviour, as represented, for example, by $\tilde{Q}_V(u)$, the quantile of V_k according to the new distribution $\tilde{f}_{\mathbf{V}}$, where the tail level u is close to zero (one). A convenient specification is:

$$\tilde{Q}_V(u) \gtreqless Q_V(u) \tag{11}$$

where Q_V is the reference quantile induced by $f_{\mathbf{V}}$, or alternatively benchmark quantiles such as the normal or the Student t.

- Lower (upper) tail co-dependence, as represented by $\tilde{C}_{\mathbf{V}}(\mathbf{u})$, the cumulative distribution function (CDF) of the copula of $\mathbf{V}$ at joint threshold levels $\mathbf{u}$ close to zero (one). A convenient specification reads:

$$\tilde{C}_{\mathbf{V}}(\mathbf{u}) \gtreqless \xi C_{\mathbf{V}}(\mathbf{u}) \tag{12}$$

where $C_{\mathbf{V}}$ is the reference copula CDF induced by $f_{\mathbf{V}}$, or alternatively benchmark copula CDFs such as the normal or the Student t.

The above is only a partial list of all the possible features on which the user may wish to express views, and which can be handled by the EP approach.

The posterior. The posterior distribution should satisfy the views without adding additional structure and should be as close as possible to the reference model (2).

The relative entropy between a generic distribution $\tilde{f}_{\mathbf{X}}$ and a reference distribution $f_{\mathbf{X}}$:

$$\mathcal{E}\left(\tilde{f}_{\mathbf{X}}, f_{\mathbf{X}}\right) \equiv \int \tilde{f}_{\mathbf{X}}(\mathbf{x})\left[\ln \tilde{f}_{\mathbf{X}}(\mathbf{x}) - \ln f_{\mathbf{X}}(\mathbf{x})\right] d\mathbf{x} \tag{13}$$

is a natural measure of the amount of structure in $\tilde{f}_{\mathbf{X}}$. Furthermore, it also measures how distorted $\tilde{f}_{\mathbf{X}}$ is with respect to $f_{\mathbf{X}}$. Indeed, if the two distributions coincide, relative entropy is zero; by imposing constraints on $\tilde{f}_{\mathbf{X}}$ this distribution departs from $f_{\mathbf{X}}$ and relative entropy increases.

Therefore, we define the posterior market distribution as:

$$\tilde{f}_{\mathbf{X}} \equiv \underset{f\in\mathbb{V}}{\arg\min}\left\{\mathcal{E}\left(f, f_{\mathbf{X}}\right)\right\} \tag{14}$$

where $f \in \mathbb{V}$ stands for all the distributions consistent with the view statements such as (6)–(12).

Entropy minimisation is widely applied in physics and statistics (see Cover & Thomas, 2006). For applications to finance, see, for example, Avellaneda (1999), D'Amico, Fusai & Tagliani (2003), Cont & Tankov (2007) and Pezier (2007). In our context, entropy minimisation is even more natural, as it generalises Bayesian updating (see Caticha & Giffin, 2006).

The confidence. One last step is required: the posterior $\tilde{f}_{\mathbf{X}}$ assumes that the practitioner has full confidence in his statements. If the confidence is less than full, the posterior distribution of the factors must shrink towards the reference factor distribution. This is easily achieved, as in Meucci (2006), by opinion-pooling the reference model and the full-confidence posterior:

$$\tilde{f}_{\mathbf{X}}^{c} \equiv (1-c) f_{\mathbf{X}} + c\tilde{f}_{\mathbf{X}} \tag{15}$$

The pooling parameter $c \in [0, 1]$ represents the confidence level in the views: in the extreme case when the confidence is total, the full-confidence posterior is recovered; on the other hand, in the absence of confidence, the reference risk model is recovered.

Opinion pooling becomes very useful in a multi-manager context. Indeed, consider S users that input their separate views on (possibly, but not necessarily) different functions of the market. As in (14), we obtain S full-confidence posterior distributions $\tilde{f}_{\mathbf{X}}^{(s)}$, $s = 1, \ldots, S$. Then the posterior distribution results naturally as the confidence-weighted average of the individual full-confidence posteriors:

$$\tilde{f}_{\mathbf{X}}^{c} \equiv \sum_{s=1}^{S} c_s \tilde{f}_{\mathbf{X}}^{(s)} \tag{16}$$

These confidence levels can be linked naturally to the track record of the respective manager, that is, the sth confidence c_s can be set as an increasing function of the number of past views (that is, seniority) and of the correlation of these views with the actual market realisation, in the same spirit as the 'skill' measure in Grinold & Kahn (1999).

The definitions (15)–(16) follow from a probabilistic interpretation of the confidence. One can easily specify different confidence levels for the different views of the same user and integrate these within a multi-user context. As it turns out, this amounts to specifying a probability measure on the power set of the views. We discuss these simple rules in Appendix A1.

We emphasise that, unlike in Black & Litterman, in EP the confidence in the views (15) and the views on volatility (9) are modelled separately. Indeed, being sure about future volatility and being uncertain about future market realisations are two very different issues.

Limit cases. If the practitioner has no views, that is, $\mathbb{V}$ is the empty set in (14), then the confidence-weighted posterior distribution equals the reference model $f_{\mathbf{X}}$.

At the other extreme, if the views fully specify a joint distribution (5), the minimisation (14) is not necessary. Indeed, consistent with the principle of minimum discrimination information, the full-confidence posterior follows from its conditional-marginal decomposition:

$$\tilde{f}_{\mathbf{X}}(\mathbf{x}) \equiv \int f_{\mathbf{X}|\mathbf{v}}(\mathbf{x}) \tilde{f}_{\mathbf{V}}(\mathbf{v}) d\mathbf{v} \tag{17}$$

In particular, this is the case in scenario analysis, where the user associates full probability with one single scenario $\mathbf{g}(\mathbf{X}) \equiv \tilde{\mathbf{v}}$. The views are represented with a Dirac delta centred on the scenario $\tilde{f}_{\mathbf{V}}(\mathbf{v}) \equiv \delta(\mathbf{v} - \tilde{\mathbf{v}})$, which, substituted in (17), yields $\tilde{f}_{\mathbf{X}} \equiv f_{\mathbf{X}|\tilde{\mathbf{v}}}$. In words, the full-confidence posterior distribution is simply the reference distribution, conditioned on $\mathbf{g}(\mathbf{X})$ assuming the scenario values $\tilde{\mathbf{v}}$. Therefore, EP includes full-distribution specification and standard scenario analysis as special cases.

An analytical formula

Consider, as in Black & Litterman, a normal reference model:

$$\mathbf{X} \sim N(\boldsymbol{\mu}, \boldsymbol{\Sigma}) \tag{18}$$

Consider views on the expectations of arbitrary linear combinations $\mathbf{QX}$ and on the covariances of arbitrary, potentially different, linear combinations $\mathbf{GX}$:

$$V: \begin{cases} \tilde{\mathbb{E}}\{\mathbf{QX}\} \equiv \tilde{\boldsymbol{\mu}}_{\mathbf{Q}} \\ \tilde{\mathbb{C}}ov\{\mathbf{GX}\} \equiv \tilde{\boldsymbol{\Sigma}}_{\mathbf{G}} \end{cases} \tag{19}$$

where $\mathbf{Q}$, $\mathbf{G}$, $\tilde{\boldsymbol{\Sigma}}_{\mathbf{G}}$ and $\tilde{\boldsymbol{\mu}}_{\mathbf{Q}}$ are conformable matrices/vector.

As we show in Appendix A2, the full-confidence posterior distribution (14) is normal:

$$\mathbf{X} \sim N(\tilde{\boldsymbol{\mu}}, \tilde{\boldsymbol{\Sigma}}) \tag{20}$$

where:

$$\tilde{\boldsymbol{\mu}} \equiv \boldsymbol{\mu} + \boldsymbol{\Sigma}\mathbf{Q}'(\mathbf{Q}\boldsymbol{\Sigma}\mathbf{Q}')^{-1}(\tilde{\boldsymbol{\mu}}_{\mathbf{Q}} - \mathbf{Q}\boldsymbol{\mu}) \tag{21}$$

$$\tilde{\boldsymbol{\Sigma}} \equiv \boldsymbol{\Sigma} + \boldsymbol{\Sigma}\mathbf{G}'\left((\mathbf{G}\boldsymbol{\Sigma}\mathbf{G}')^{-1}\tilde{\boldsymbol{\Sigma}}_{\mathbf{G}}(\mathbf{G}\boldsymbol{\Sigma}\mathbf{G}')^{-1} - (\mathbf{G}\boldsymbol{\Sigma}\mathbf{G}')^{-1}\right)\mathbf{G}\boldsymbol{\Sigma} \tag{22}$$

Then the confidence-weighted posterior distribution (15) is a normal mixture:

$$\mathbf{X} \sim \begin{cases} N(\boldsymbol{\mu}, \boldsymbol{\Sigma}) & (probability: 1-c) \\ N(\tilde{\boldsymbol{\mu}}, \tilde{\boldsymbol{\Sigma}}) & (probability: c) \end{cases} \tag{23}$$

This distribution is suitable for instance to stress test market crashes, where high volatilities, high correlations and low expectations in $\tilde{\boldsymbol{\mu}}, \tilde{\boldsymbol{\Sigma}}$ are expected to occur with probability $c << 1$.

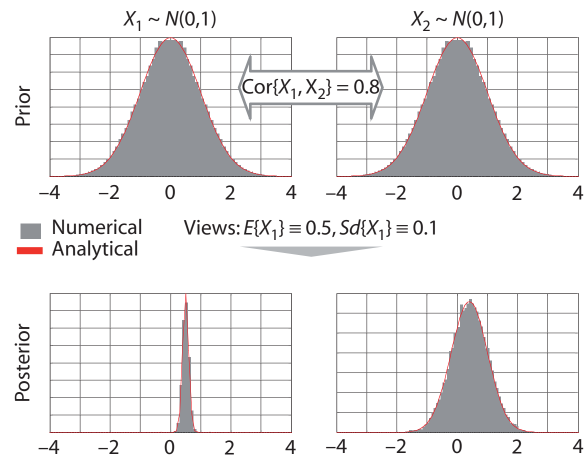
Formula (23) generalises results in Pezier (2007). Also, the special case of full-confidence $c \equiv 1$ on only one set of linear combinations $\mathbf{Q} \equiv \mathbf{G}$ yields the result in Qian & Gorman (2001). This is not surprising, as the authors' approach is equivalent to the decomposition (17). Finally, the further specialisation to null dispersion in the views $\tilde{\boldsymbol{\Sigma}}_{\mathbf{G}} \to \mathbf{0}$ yields scenario analysis as in Meucci (2005), which in turn generalises the standard regression-based approach that appears, for example, in Mina & Xiao (2001).

Numerical implementation

Except for the special case in the previous section, the EP approach cannot be implemented analytically. However, the numerical implementation of the EP approach in full generality is extremely simple and computationally efficient.

First, we represent the reference distribution (2) of the market $\mathbf{X}$ in terms of a $J \times N$ panel $\mathcal{X}$ of simulations. The generic jth row of $\mathcal{X}$ represents one in a very large number of joint scenarios for the N variables $\mathbf{X}$, whereas the generic nth column of $\mathcal{X}$ represents the marginal distribution of the nth factor X_n. With the scenarios we associate the $J \times 1$ vector of the respective probabilities $\mathbf{p}$, whose each entry typically, but not necessarily, equals $1/J$. See Glasserman & Yu (2005) for a variety of methods to determine $\mathbf{p}$.

We assume that each of the joint scenarios in $\mathcal{X}$ has been mapped into the respective joint price scenarios for the I securities in the market considered by the user, by means of the potentially costly function (1), thereby generating a $J \times I$ panel of prices $\mathcal{P}$. The panel of the security prices $\mathcal{P}$, along with the respective probabilities $\mathbf{p}$, is



1 Entropy pooling: numerical approach matches analytical solution

then analysed for risk management purposes, or it is fed into an optimisation algorithm to perform the asset allocation step (3).

The user expresses views on generic non-linear functions of the market (4). Their distribution as implied by the reference model is readily represented by the $J \times K$ panel $\mathcal{V}$ defined entry-wise as follows:

$$\mathcal{V}_{j,k} \equiv g_k\left(\mathcal{X}_{j,1},\dots,\mathcal{X}_{j,N}\right) \tag{24}$$

To represent the posterior distribution of the market that includes the views, instead of generating new simulations, we use the same scenarios with different probabilities $\tilde{\mathbf{p}}$. Then, as we show in Appendix A3, general views such as (6)–(12) can be written as a set of linear constraints on the new, yet to be determined, probabilities:

$$\underline{\mathbf{a}} \leq \mathbf{A}\tilde{\mathbf{p}} \leq \overline{\mathbf{a}} \tag{25}$$

where $\mathbf{A}$, $\underline{\mathbf{a}}$ and $\overline{\mathbf{a}}$ are simple expressions of the panel (24). For instance, for standard views on expectations $\mathbf{A} \equiv \mathcal{V}'$ and $\underline{\mathbf{a}} \equiv \overline{\mathbf{a}}$ quantify the views.

Furthermore, the relative entropy (13) becomes its discrete counterpart:

$$\mathcal{E}(\tilde{\mathbf{p}},\mathbf{p}) \equiv \sum_{j=1}^{J} \tilde{p}_j\left[\ln\left(\tilde{p}_j\right) - \ln\left(p_j\right)\right] \tag{26}$$

Therefore, the full-confidence posterior distribution (14) is defined as:

$$\tilde{\mathbf{p}} \equiv \underset{\underline{\mathbf{a}} \leq \mathbf{A}\mathbf{f} \leq \overline{\mathbf{a}}}{\arg\min}\left\{\mathcal{E}(\mathbf{f},\mathbf{p})\right\} \tag{27}$$

This optimisation can be solved very efficiently: as we show in Appendix A4, the dual formulation is a simple linearly constrained convex program in a number of variables equal to the number of views, not the number of Monte Carlo simulations, which can be kept large. Therefore, we can achieve excellent accuracy even under extreme views (see figure 1).

Now we can compute the opinion-pooling, confidence-weighted posterior (15). This is represented by $(\mathcal{X}, \mathbf{p}_c)$, the same simulations as for the reference model, but with new probabilities:

$$\mathbf{p}_c \equiv (1-c)\mathbf{p} + c\tilde{\mathbf{p}} \tag{28}$$

A similar expression holds for the more general multi-user, multi-confidence posterior discussed in Appendix A1.

Since the posterior factor distribution is obtained by tweaking the relative probabilities of the scenarios $\mathcal{X}$ without affecting the scenarios themselves, the posterior distribution of the market prices is represented by $(\mathcal{P}, \mathbf{p}_c)$, the original panel of joint prices and the new probabilities. Hence, no repricing is necessary to process views and stress tests.

Case study: option trading

As in Meucci (2008b), we consider a trader of butterflies, defined as long positions in one call and one put with the same strike, underlying and time to maturity. The price $P_{t+\tau}$ of the butterfly at the investment horizon can be written in the format (1) as a deterministic non-linear function of a set of risk factors and current information. Indeed:

$$P_{t+\tau} = BS\left(y_t e^{X_y}, h\left(y_t e^{X_y}, \sigma_t + X_\sigma, K, T-\tau\right); K, T-\tau, r\right) \tag{29}$$

where τ is the investment horizon, y_t is the current value and $X_y \equiv \ln(y_{t+\tau}/y_t)$ is the log-change of the underlying, σ_t is the current value and $X_\sigma \equiv \sigma_{t+\tau} - \sigma_t$ is the change in at-the-money implied volatility. BS is the Black-Scholes formula:

$$\begin{aligned} &BS(y,\sigma;K,T,r) \\ &\equiv y\left[\Phi(d_1) - \Phi(-d_1)\right] - Ke^{-rT}\left[\Phi(d_2) - \Phi(-d_2)\right] \end{aligned} \tag{30}$$

where Φ is the standard normal CDF, K is the strike, T is the time to expiry, r is the risk-free rate, $d_1 \equiv (\ln(y/K) + (r + \sigma^2/2)T)/\sigma\sqrt{T}$, $d_2 \equiv d_1 - \sigma\sqrt{T}$ and h is a skew/smile map:

$$h(y,\sigma;K,T) \equiv \sigma + \alpha\frac{\ln(y/K)}{\sqrt{T}} + \beta\left(\frac{\ln(y/K)}{\sqrt{T}}\right)^2 \tag{31}$$

for coefficients α and β, which depend on the underlying and are fitted empirically, similarly to Malz (1997). If the investment horizon τ is short, a delta-gamma-vega approximation of (29) would suffice. However, we leave the exact formulation to demonstrate how the present approach does not require costly repricing.

Consider a portfolio represented by the vector $\mathbf{w}$, whose generic ith entry is the number of contracts in the respective butterfly. The profit and loss then reads:

$$\Pi_{\mathbf{w}} \equiv \sum_{i=1}^{I} w_i\left(P_i(\mathbf{X},\mathcal{I}_t) - P_{i,t}\right) \tag{32}$$

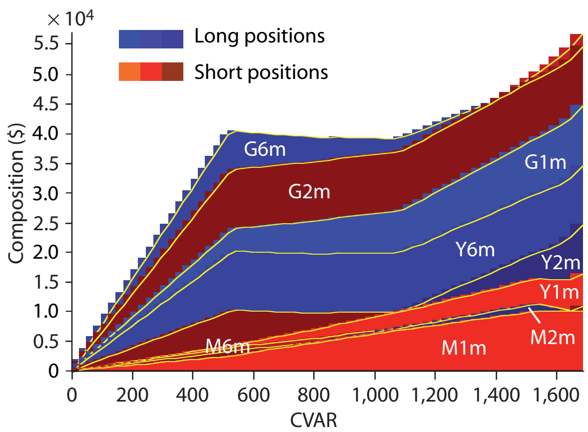
where $P_i(\mathbf{X}, \mathcal{I}_t)$ is the price at the horizon (29) and $P_{i,t}$ is the currently traded price of the ith butterfly. We assume that, to account for market asymmetries and downside risk, the trader optimises the mean-CVAR trade-off. Therefore, (3) becomes:

$$\mathbf{w}_\lambda \equiv \underset{\underline{\mathbf{b}} \leq \mathbf{B}\mathbf{w} \leq \overline{\mathbf{b}}}{\arg\max}\left\{\mathbb{E}\{\Pi_{\mathbf{w}}\} - \lambda CVAR_\gamma\{\Pi_{\mathbf{w}}\}\right\} \tag{33}$$

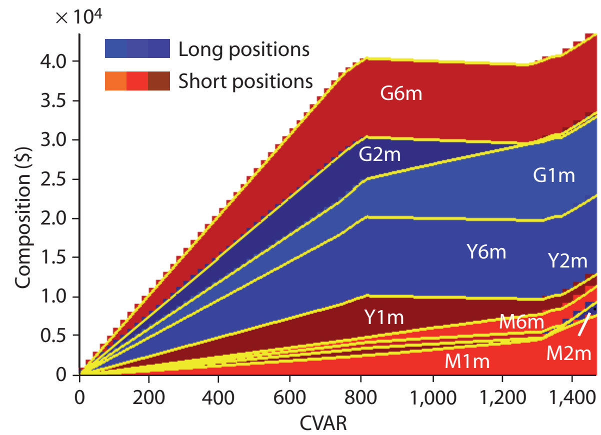
where γ is the CVAR tail level, and $\mathbf{B}$, $\underline{\mathbf{b}}$ and $\overline{\mathbf{b}}$ are a matrix and vectors that represent investment constraints.

To illustrate, we set $\gamma \equiv 95\%$, and we impose that the long-short positions offset to a zero delta and a zero initial budget, and that the absolute investment in each option does not exceed a fixed threshold. We set the investment horizon as $\tau \equiv 1$ day. We consider a limited market of $I \equiv 9$ securities: one-month, two-month and six-month butterflies on the three technology stocks Microsoft (M), Yahoo (Y) and Google (G).

In addition to the respective underlyings and implied volatili-



2 Mean-CVAR long-short efficient frontier: reference risk model



3 Mean-CVAR long-short efficient frontier: view on Google six-month/Google two-month spread

ties, we include the possibility of views on growth or inflation, as represented by the slope of the interest rate curve. Therefore we add the changes in the two- and 10-year points of the curve, for a total of $N \equiv 14$ factors:

$$\mathbf{X} \equiv \left(X^M, X^M_{1m}, X^M_{2m}, X^M_{6m}, \ldots, X^G_{6m}, X_{2y}, X_{10y} \right)' \qquad (34)$$

To determine the reference distribution (2) of these factors, we consider the panel of joint observations of the factors over a three-year horizon. This amounts to 700 observations. To achieve $J \equiv 10^5$ joint simulations, we kernel-bootstrap the historical scenarios. For each historical observation $\mathbf{x}_t$, we draw $10^5/700$ observations from the multivariate normal distribution $N(\mathbf{x}_t, \varepsilon\hat{\mathbf{\Sigma}})$, where $\hat{\mathbf{\Sigma}}$ is the sample covariance, and we set $\varepsilon \equiv 0.15$. The juxtaposition of the above simulations yields the desired $J \times N$ panel $\mathcal{X}$, where each scenario has equal probability $p_j \equiv 1/J$.

Then we input each scenario of $\mathcal{X}$ into the pricing function (30), obtaining the joint profit and loss scenarios $\mathcal{P}$ with equal probabilities $\mathbf{p}$. The sample counterpart of the mean-CVAR efficient frontier (33) reads:

$$\mathbf{w}_\lambda \equiv \underset{\underline{\mathbf{b}} \le \mathbf{B}\mathbf{w} \le \bar{\mathbf{b}}}{\arg\max} \left\{ (\mathbf{w}'\mathcal{P}'\mathbf{p}) + \lambda \frac{[\mathbf{p}]'[\mathcal{P}\mathbf{w}]}{[\mathbf{p}]'[\mathbf{1}]} \right\} \qquad (35)$$

where the operator $[\mathbf{x}]$ selects in the generic vector $\mathbf{x}$ only the entries that correspond to the $(1-\gamma)J$ smallest entries of $\mathcal{P}\mathbf{w}$. If J is not too large, this can be solved by linear programming, as in Rockafellar & Uryasev (2000). For very large J, we solve this heuristically, as in Meucci (2005), by a two-step approach: first determine the mean-variance efficient frontier, then perform a univariate grid search for the optimal trade-off (35).

In figure 2, we display the frontier ensuing from the reference market model in our example. For the extreme case of zero risk appetite, not investing at all is optimal. As the risk appetite increases, leverage increases, always respecting the constraint of a zero net initial investment, as well as delta-neutrality. When the risk appetite increases further, the remaining constraints enter the picture.

Now we consider the views of three distinct analysts. The first one is bearish about the two-month/six-month implied volatility spread for Google. From (6)–(7), this means:

$$\tilde{\mathbb{E}}\left\{ X^G_{6m} - X^G_{2m} \right\} \le \mathbb{E}\left\{ X^G_{6m} - X^G_{2m} \right\} - \sigma\left\{ X^G_{6m} - X^G_{2m} \right\} \qquad (36)$$

This view is represented in the form (25) as:

$$\sum_{j=1}^{J} \tilde{p}^{(1)}_j \left(X^G_{j,6m} - X^G_{j,2m} \right) \le \hat{m}_{6|2} - \hat{\sigma}_{6|2} \qquad (37)$$

where $\hat{m}_{6|2}$ and $\hat{\sigma}_{6|2}$ are the sample counterparts of the respective terms in (36). We can calculate $\tilde{\mathbf{p}}^{(1)}$ as in (27), under the constraint (37). To illustrate, we show in figure 3 the mean-CVAR efficient frontier (35) when this view is processed. As expected, the Google six-month/Google two-month spread, previously long, is now short.

The second analyst is bullish on the realised volatility of Microsoft, defined as $|X^M|$, the absolute log-change in the underlying. This is the variable such that, if larger than a threshold, a long position in the butterfly turns into a profit. Since this variable displays thick tails and the expectation might not be defined (see, for example, Rachev, 2003), we issue a relative statement on the median, comparing it with the third quintile implied by the reference market model:

$$\tilde{\mathbb{M}}\left\{ \left| X^M \right| \right\} \ge Q_{|X^M|}\left(\frac{3}{5} \right) \qquad (38)$$

This view is represented in the form (25) as:

$$\sum_{j \in \tilde{J}} \tilde{p}^{(2)}_j \le \frac{1}{2} \qquad (39)$$

where $\tilde{J}$ is the set of indexes j such that $|X^M_j|$ is smaller than the sample third quintile of $|X^M|$. Now we can calculate $\tilde{\mathbf{p}}^{(2)}$ as in (27) under the constraint (39).

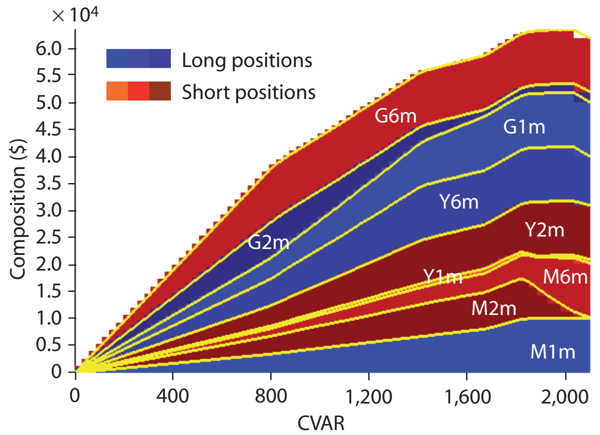
The third analyst believes that the slope of the curve will increase by 5 basis points. Therefore, he formulates the view in the manner of Black & Litterman, using in (6) expectations and binding constraints:

$$\sum_{j=1}^{J} \tilde{p}^{(3)}_j \left(X_{j,10y} - X_{j,2y} \right) \equiv 0.0005 \qquad (40)$$

and $\tilde{\mathbf{p}}^{(3)}$ can be calculated as in (27).

The management committee attributes $c_1 \equiv 0.20$, $c_2 \equiv 0.25$ and $c_3 \equiv 0.20$ confidence to the analysts' views, the remaining portion being attributed to the reference model. Then the uncertainty-weighted posterior probabilities read:

4 Mean-CVAR long-short efficient frontier: all views



$$\tilde{\mathbf{p}}_c \equiv \sum_{s=0}^{3} c_s \tilde{\mathbf{p}}^{(s)} \quad (41)$$

where $c_0 \equiv 1 - c_1 - c_2 - c_3$ and $\tilde{\mathbf{p}}^{(0)} \equiv \mathbf{p}$. We show in figure 4 the combined effects of all the views on the frontier (35).

We emphasise that in this case study the market has a non-parametric, thick-tailed, non-normal distribution; two views are expressed as inequalities; one view acts on a non-linear function, the absolute value, of a factor; the slope of the curve in one view is an external factor that appears nowhere in the pricing function of the securities; features different from expectations are being assessed, namely the median; and no repricing was ever necessary.

Conclusions

We have presented the EP approach, a unified framework to perform trading, portfolio management and generalised stress testing in markets with complex derivatives driven by non-normal factors. The inputs are a possibly non-normal reference market model and a set of very general equality or inequality views on a variety of features of the market. The output is a posterior distribution that incorporates all the inputs. As it turns out, the EP approach avoids costly repricing by representing the posterior distribution in terms of the same scenarios as the reference model, but with different probabilities whose computation is extremely efficient.

We summarise in table A the capabilities of the EP approach as compared with Black & Litterman (1990), Almgren & Chriss (2006), Qian & Gorman (2001), Pezier (2007), Meucci (2008b) and the COP in Meucci (2006). ■

A. Capabilities of entropy pooling compared with other approaches

	BL	AC	QG	P	M	COP	EP
Normal market and linear views	✓		✓	✓	✓	✓	✓
Scenario analysis			✓	✓	✓	✓	✓
Correlation stress test			✓	✓	✓		✓
Trading desk: non-linear pricing					✓	✓	✓
External factors: macro, etc					✓	✓	✓
Partial specifications				✓			✓
Non-normal market						✓	✓
Multiple users						✓	✓
Non-linear views							✓
Trading desk: costly pricing							✓
Lax constraints: ranking		✓					✓

Attilio Meucci is head of portfolio research at Bloomberg. He would like to thank Paul Glasserman, Sridhar Gollamudi, Ninghui Liu and an anonymous referee for helpful feedback, and aorda.com for providing Portfolio Safeguard to benchmark numerical computations. Email: a.meucci@bloomberg.net

References

Almgren R and N Chriss, 2006
Optimal portfolios from ordering information
The Journal of Risk 9, pages 1–47

Avellaneda M, 1999
Minimum-entropy calibration of asset-pricing models
International Journal of Theoretical and Applied Finance 1, pages 447–472

Black F and R Litterman, 1990
Asset allocation: combining investor views with market equilibrium
Goldman Sachs Fixed Income Research

Brigo D and F Mercurio, 2001
Interest rate models
Springer

Caticha A and A Giffin, 2006
Updating probabilities
Bayesian Inference and Maximum Entropy Methods in Science and Engineering, AIP Conference Proceedings

Cont R and P Tankov, 2007
Recovering exponential Lévy models from option prices: regularization of an ill-posed inverse problem
SIAM Journal on Control and Optimization 45, pages 1–25

Cover T and J Thomas, 2006
Elements of information theory
Wiley, second edition

D'Amico M, G Fusai and A Tagliani, 2003
Valuation of exotic options using moments
Operational Research 2, pages 157–186

Glasserman P and B Yu, 2005
Large sample properties of weighted Monte Carlo estimators
Operations Research 53, pages 298–312

Grinold R and R Kahn, 1999
Active portfolio management. A quantitative approach for producing superior returns and controlling risk
McGraw-Hill, second edition

Malz A, 1997
Option-implied probability distributions and currency excess returns
Federal Reserve Bank of New York

Meucci A, 2005
Risk and asset allocation
Springer

Meucci A, 2006
Beyond Black-Litterman in practice: a five-step recipe to input views on non-normal markets
Risk September, pages 114–119

Meucci A, 2008a
The Black-Litterman approach: original model and extensions
Encyclopedia of Quantitative Finance, extended version available at *www.symmys.com > Research > Working Papers*

Meucci A, 2008b
Enhancing the Black-Litterman and related approaches: views and stress-test on risk factors
Forthcoming in Journal of Asset Management, available at above address

Mina J and J Xiao, 2001
Return to RiskMetrics: the evolution of a standard
RiskMetrics

Pezier J, 2007
Global portfolio optimization revisited: a least discrimination alternative to Black-Litterman
ICMA Centre Discussion Papers in Finance

Qian E and S Gorman, 2001
Conditional distribution in portfolio theory
Financial Analyst Journal 57, pages 44–51

Rachev S, 2003
Handbook of heavy tailed distributions in finance
Elsevier/North-Holland

Rockafellar R and S Uryasev, 2000
Optimization of conditional value-at-risk
The Journal of Risk 2, pages 21–41