On the Bayesian Interpretation of Black-Litterman

Petter Kolm and Gordon Ritter
Courant Institute, NYU

petter.kolm@nyu.edu

This presentation is based on Kolm and Ritter (2017), "On the Bayesian Interpretation of Black-Litterman," *European Journal of Operational Research*, 258 (2), pp. 564-572.

ARPM Conference
August 12, 2018

Outline

Background & motivation

Modern portfolio theory (1/2)

Modern portfolio theory (Markowitz, 1952) provides a conceptual framework to

- Quantify risk and return
- Determine the trade-off between risk and return

Remarks

- For detailed treatment of portfolio allocation techniques, see Fabozzi et al. (2007) and Fabozzi, Focardi, and Kolm (2010)
- For recent trends and developments in portfolio construction, see Kolm, Tütüncü, and Fabozzi (2014)

Modern portfolio theory (2/2)

Investors allocate their wealth by solving the mean-variance optimization (MVO) problem

$$w^* = \underset{h \in C}{\operatorname{argmax}} \left\{ h' \mathbb{E}[R] - \frac{\delta}{2} h' \mathbb{V}[R] h \right\}$$

where

h	:	portfolio weights
$\mathbb{E}[R]$	:	expected returns
$\mathbb{V}[R]$	:	covariance matrix of returns
δ	:	risk aversion
C	:	set of constraints (linear equality/inequality constraints, etc.)

What makes a good portfolio construction model?

Outperforms simple benchmark portfolios (e.g. equal / equal vol-weighted)

- Compare portfolio weights, turnover and performance

Sensible, intuitive and robust allocations

- What is driving particular allocations?
  - Expected returns, correlations, volatilities
  - Portfolio constraints
- How do allocations change when input parameters change?
  - Sensitivity analysis of model parameters

Model transparency

- "The boardroom test"

Extendable

- Distributional assumptions
- Multi-period models: Alpha decay, impact costs and taxes

MVO is sensitive to its inputs

- Small changes in the expected returns and covariance matrix may result in large changes in the resulting portfolio weights
- Any estimate is subject to estimation error (sampling error). If the estimation error is large the output would be garbage at best

Consequences

- Simple portfolio rules (e.g. equal weighting) often outperforms (Jobson and Korkie (1981), Jorion (1985))
- Portfolio optimizers are often "error maximizers" (Michaud (1998))
- Optimal portfolios are not necessarily well diversified and result in "corner solutions" (Green and Hollifield (1992))

Portfolio optimization in practice (1/2)

Techniques aimed mitigating the sensitivity of MVO

- Constrain portfolio weights ("the practitioner's solution")
  - No short-selling constraints (see for example, Frost and Savarino (1988), Chopra (1991), Gupta and Eichhorn (1998), Grauer and Shen (2000), Jagannathan and Ma (2003))
  - "Diversification indicators" (Bouchard, Potters and Aguilar (1997))
- Improve estimation
  - Bayesian techniques:
    - James-Stein estimation (Jobson and Korkie (1981), Ledoit and Wolf (2003))
    - Black-Litterman model (Black and Litterman (1990))
  - Robust statistics (Trojani and Vanini (2002), DeMiguel and Nogales (2006))

Portfolio optimization in practice (2/2)

Techniques aimed mitigating the sensitivity of MVO:, cont'd

- Incorporate estimation error in portfolio allocation
  - Adjustment of risk aversion factor (Horst, Roon and Werker (2000))
  - Resampled efficiency (Michaud (1998), Jorion (1992), Scherer (2002), Markowitz and Usmen (2003))
  - Robust optimization (El Ghaoui and Lebret (1977), Ben-Tal and Nemirovski (1998, 1999))

The Black-Litterman model

The Black-Litterman model (BL):[1]

- "Market based" shrinkage estimator of the expected returns, referred to as Black-Litterman expected returns
- Resulting expected return estimates are "confidence weighted" averages of market equilibrium and investors' views
- Investors express "subjective" views on signals, strategies and/or portfolios
- No need to estimate expected returns or have views on all assets

BL is a Bayesian model, but the Bayesian nature of the model has been little explored

---

[1] See, Black and Litterman (1991).

Black-Litterman: A Bayesian(?) model

- Black and Litterman (1991) optimization seems to have generated more than its share of confusion over the years, as evidenced by articles with titles such as *“A demystification of the Black–Litterman model”* (Satchell and Scowcroft, 2000), etc.
- The method itself is often described as “Bayesian” but the original authors do not elaborate directly on connections with Bayesian statistics
- The paper by Litterman and He (1999) contains many references to a “prior” but only one mention of a “posterior” without details, and no mention of a “likelihood”

On the Bayesian interpretation of Black-Litterman

What we do

In this talk we:

- Introduce the Black-Litterman-Bayes (BLB) model, which is the most general model of the type considered by Black and Litterman (1991). In the process we also lay out the full set of assumptions made, some of which are often glossed over
- Fully clarify the exact nature of the Bayesian statistical model to which BL corresponds, in terms of the prior, likelihood, and posterior
- Work out the treatment of views on factor risk premia in the context of APT (Ross, 1976)

Suppose A and B are events such that $P(B) \neq 0$.

From the definition

$$P(A|B) \quad := \quad \frac{P(A \cap B)}{P(B)} \tag{1}$$

we immediately obtain *Bayes rule*

$$\begin{aligned} P(A|B) &= \frac{P(B|A)P(A)}{P(B)} \qquad (2) \\ &= \frac{P(B|A)P(A)}{P(B|A)P(A) + P(B|A^c)P(A^c)} \qquad (3) \end{aligned}$$

where $P(A^c)$ denotes the compliment of A

Bayes rule (continuous version)

In 1764 Thomas Bayes proved the continuous version of Bayes rule:

Suppose we have two random variables x and y, with conditional distribution $f(x|y)$ and marginal distribution $g(y)$.

Then the conditional distribution of y given x is

$$g(y|x) = \frac{f(x|y)g(y)}{\int f(x|y)g(y)dy}$$

This motivates the following definition:

Bayesian statistical model

Definition

A *Bayesian (parametric) statistical model* consists of:

1. A random variable x, distributed according to $f(x \mid \theta)$, where realizations of x have been observed and only the parameter θ (which belongs to a vector space Θ of finite dimension) is unknown, and
2. A prior density $\pi(\theta)$ on Θ

- $f(x \mid \theta)$ is called the *likelihood*
- The *posterior* is the density on Θ proportional to $f(x \mid \theta)\pi(\theta)$

The Bayesian approach to statistics is to base inference on the *posterior distribution*

$$\pi(\theta|x) = \frac{f(x \mid \theta)\pi(\theta)}{\int f(x \mid \theta)\pi(\theta)d\theta}$$

which follows from the continuous version of Bayes rule

Our plan: Bring MVO and the Bayesian perspective together

Investors allocate their wealth by solving the mean-variance optimization (MVO) problem

$$w^* = \underset{h \in C}{\operatorname{argmax}} \left\{ h' \mathbb{E}[R] - \frac{\delta}{2} h' \mathbb{V}[R] h \right\}$$

where

h	:	portfolio weights
$\mathbb{E}[R]$	:	expected returns under the posterior distribution
$\mathbb{V}[R]$	:	covariance matrix of returns under the posterior distribution
δ	:	risk aversion
C	:	set of constraints (linear equality/inequality constraints, etc.)

Black-Litterman Views

- Consider a portfolio manager's view: "The German equity market will outperform a capitalization-weighted basket of the rest of the European equity markets by 5%" (Litterman and He, 1999)
  - Equivalent to
  $$\mathbb{E}[p'R] = 5\% \tag{4}$$
  where $p \in \mathbb{R}^n$ is a portfolio, and R is the random vector of asset returns over some subsequent interval
- Given multiple views,
  $$\mathbb{E}[p_i'R] = q_i, \quad i = 1, \ldots, k$$
  then p_i are rows of a matrix P. Then the statement becomes
  $$\mathbb{E}[PR] = Q \text{ for } Q \in \mathbb{R}^k \tag{5}$$

Where is the likelihood?

- In the language of statistics, the core idea of Black and Litterman (1991) is to treat the portfolio manager's views as noisy observations which are useful for performing statistical inference concerning the parameters in some underlying model for R
- If

$$R \sim N(\theta, \Sigma) \tag{6}$$

with Σ known, then the views (5) can be recast as "observations" relevant for inference on the parameter θ

Answer to "Where is the likelihood?"

- The portfolio manager must also specify a level of uncertainty for each view
- Error in a view is an independent source of noise from the volatility already accounted for in the return model $R \sim N(\theta, \Sigma)$
- With notation as in Litterman and He (1999):

$$P\theta = Q + \epsilon^{(v)}, \quad \epsilon^{(v)} \sim N(0, \Omega), \quad \Omega = \operatorname{diag}(\omega_1, \ldots, \omega_k) \tag{7}$$

- Statistical inference requires a likelihood and prior. Eq. (7) specifies the likelihood as

$$p(Q \mid \theta) \propto \exp\left[-\frac{1}{2}(P\theta - Q)'\Omega^{-1}(P\theta - Q) \right] \tag{8}$$

$\Rightarrow$ multiple linear regression problem with dependent variable Q and design matrix P.

$p > n$ regression

- Typical regression problem has p variables, n observations, and $p \ll n$. However, models with $p \gg n$ also arise in many applications (and can be handled by ridge and the lasso)
- BL with one single view represents a data-scarce situation: one observation and one parameter *per asset*: θ are the unobservable means of the asset returns
- More generally, we may be presented with no views, one, or very many. Views from diverse portfolio managers or economists may contain internal contradictions
- Bayesian regression is the ideal tool to deal with all such cases. Internal contradictions $\Rightarrow$ no exact (zero-residual) solution to the regression equations – the usual situation in linear regression when $p \ll n$

The prior (part 1)

- BL: In absence of any sort of information/views which have alpha over the benchmark, optimization should return global CAPM equilibrium portfolio, with weights denoted w_{eq}
- In absence of views, and with prior mean Π, the investor's model of the world is that

$$R \sim N(\theta, \Sigma), \quad \text{and} \quad \theta \sim N(\Pi, C) \tag{9}$$

  for some covariance C representing the amount of precision in the prior
- For any portfolio p, then, according to (9) we have

$$\mathbb{E}[p'R] = p'\Pi, \quad \text{and} \quad \mathbb{V}[p'R] = p'(\Sigma + C)p\,.$$

Conditional or unconditional variance?

- We must make a choice whether to use the conditional or unconditional variance in optimization: $\mathbb{V}(R \mid \theta) = \Sigma$ but $\mathbb{V}(R) = \Sigma + C$
- Investors are presumably concerned with unconditional variance of wealth, and therefore mean-variance optimization with risk-aversion parameter δ gives

$$w_{eq} = \delta^{-1}(\Sigma + C)^{-1}\Pi$$

  Any combination of Π, C satisfying this $\Rightarrow$ optimal portfolio with only the information given in the prior is the prescribed portfolio w_{eq}
- Taking $C = \tau\Sigma$ with $\tau > 0$, as did the original authors, leads to $\Pi = \delta(1 + \tau)\Sigma w_{eq}$

Likelihood and prior together

- We thus have the normal likelihood

$$p(Q \mid \theta) \propto \exp\left[-\frac{1}{2}(P\theta - Q)'\Omega^{-1}(P\theta - Q)\right]$$

and the normal prior

$$R \sim N(\theta, \Sigma), \qquad \text{and} \qquad \theta \sim N(\Pi, C)$$

which is a *conjugate prior* for that likelihood, meaning that the posterior is of the same family (i.e. also normal in this example)[2]

[2] A detailed discussion of conjugate priors is found in Robert (2007, Sec. 3.3).

Completing the squares

- The negative log posterior is thus proportional to (neglecting terms that don't contain θ):

$$
\begin{aligned}
&(P\theta - Q)'\Omega^{-1}(P\theta - Q) + (\theta - \Pi)'C^{-1}(\theta - \Pi) \\
&= \theta'P'\Omega^{-1}P\theta - \theta'P'\Omega^{-1}Q - Q'\Omega^{-1}P\theta \\
&\qquad\qquad + \theta'C^{-1}\theta - \theta'C^{-1}\Pi - \Pi'C^{-1}\theta \\
&= \theta'[P'\Omega^{-1}P + C^{-1}]\theta - 2(Q'\Omega^{-1}P + \Pi'C^{-1})\theta
\end{aligned}
$$

Lemma (Completing the squares)

If a multivariate normal random variable x has density $p(x)$ and if

$$-2\log p(x) = x'Hx - 2\eta'x + (\textit{terms without } x)$$

then $\mathbb{V}[x] = H^{-1}$ and $\mathbb{E}x = H^{-1}\eta$.

The prior (part 2)

- Hence the posterior has mean

$$v = [P'\Omega^{-1}P + C^{-1}]^{-1}[P'\Omega^{-1}Q + C^{-1}\Pi] \tag{10}$$

and covariance

$$H^{-1} = [P'\Omega^{-1}P + C^{-1}]^{-1} \tag{11}$$

- Part of the beauty is simplicity: Really just a few lines of algebra

The prior (part 3)

- Investors with CARA utility of final wealth will want to solve

$$w^* = \underset{h}{\text{argmax}} \left\{ \mathbb{E}[h'R] - \frac{\delta}{2}\mathbb{V}[h'R] \right\}$$

  where $\mathbb{E}[R]$ and $\mathbb{V}[R]$ denote, respectively, the unconditional mean and covariance of R under the posterior
- Unconditional covariance is a sum of variance due to parameter uncertainty, and variance due to randomness in R:

$$\mathbb{V}[h'R] = h'[P'\Omega^{-1}P + C^{-1}]^{-1}h + h'\Sigma h$$

- The optimal portfolio with both types of variance is therefore

$$w^* = \delta^{-1}[H^{-1} + \Sigma]^{-1}H^{-1}[P'\Omega^{-1}Q + C^{-1}\Pi]$$

Black-Litterman-Bayes model

Definition

A *Black-Litterman-Bayes* (BLB) model consists of:

(a) A parametric statistical model for asset returns $p(R \mid \theta)$ with finite-dimensional parameter vector θ

(b) A prior $\pi(\theta)$ on the parameter space

(c) A likelihood function $f(Q \mid \theta)$ where θ is any parameter vector appearing in a parametric statistical model for asset returns, and Q is a vector of views supplied by the portfolio manager

(d) A utility function $u(w)$ of final wealth in the sense of Arrow (1971) and Pratt (1964)

The optimal decision

- Note that the first two items simply state that one has a Bayesian statistical model in the sense we discussed above
- Under such a model, Decision Theory teaches us that the optimal decision is the one maximizing posterior expected utility[3]
- This leads us to...

[3] See, Robert (2007, Ch. 2) and references therein.

The Black-Litterman-Bayes optimal portfolio

Definition

Given a BLB model as per the definition above, the associated BLB optimal portfolio is defined to be

$$h^* \in \operatorname*{argmax}_h \mathbb{E}[u(h'R) \mid Q]$$

where $\mathbb{E}[\cdot \mid Q]$ denotes the expectation with respect to the posterior predictive density for the random variable R.

In other words, h^* maximizes posterior expected utility. Explicitly, the posterior predictive density of R is given by

$$\begin{aligned} p(R \mid Q) &= \int p(R \mid \theta) p(\theta \mid Q) d\theta \quad \text{where} \\ p(\theta \mid Q) &= \frac{f(Q \mid \theta)\pi(\theta)}{\int f(Q \mid \theta)\pi(\theta) d\theta} \end{aligned}$$

Benchmark-optimal priors

Definition

Given a benchmark portfolio with holdings h_B (eg. the market portfolio), and given a BLB model, the prior $\pi(\theta)$ is said to be *benchmark-optimal* if h_B maximizes expected utility of wealth, where the expectation is taken with respect to the a priori distribution on asset returns $p(R) = \int p(R \,|\, \theta)\pi(\theta)d\theta$, so

$$h_B \in \operatorname*{argmax}_h \int u(h'R)p(R \,|\, \theta)\pi(\theta)d\theta \qquad (12)$$

The classic Black-Litterman model is a special case

The model of Black and Litterman (1991) is the special case where:

- $f(\cdot \mid \cdot)$ is the normal likelihood for a regression of the portfolio manager's views
- The utility of final wealth is the CARA function

$$u(w) = -e^{-\delta w}$$

- The prior is the unique normal distribution which is benchmark-optimal with respect to the market portfolio

BLB for factor models

Generalizing further, θ could represent means (and covariances) of unobservable latent factors in an APT model (eg. Ross (1976) and Roll and Ross (1980))

$$R = Xf + \epsilon, \quad \mathbb{E}[\epsilon] = 0, \quad \mathbb{V}[\epsilon] = D \tag{13}$$

where

- R is the cross-section of asset returns in excess of the risk-free rate
- X is a (non-random) $n \times k$ matrix that is known before R
- Also, ϵ follows a mean-zero distribution with diagonal variance-covariance

$$D = \text{diag}(\sigma_1^2, \dots, \sigma_n^2) \;\; \text{with all} \;\; \sigma_i^2 > 0 \tag{14}$$

APT & factor models (2/2)

The variable f in (13) denotes a k-dimensional random vector process which cannot be observed directly. Therefore,

- Information about the f-process must be obtained via statistical inference
- Specifically, we assume that the f-process has finite first and second moments given by

$$\mathbb{E}[f] = \mu_f, \quad \text{and} \quad \mathbb{V}[f] = F \tag{15}$$

APT and dimension reduction

- The model given by (13),(14) and (15) entails

$$\mathbb{E}[R] = X\mu_f, \quad \text{and} \quad \Sigma := \mathbb{V}[R] = D + XFX' \qquad (16)$$

  where X' denotes the transpose of X. The elements of μ_f are called *factor risk premia*
- Assume $k \ll n$. Then (16) reduces the number of parameters necessary to describe the density $p(R)$ from $O(n^2)$ down to the k parameters in μ_f, the $k(k+1)/2$ parameters in F, and n parameters in D, for a total of $n + k(k+3)/2$ parameters
- Models of the form (13) are ubiquitous in practice, and for good reason: In equity markets n is too large to allow direct estimation of Σ

We are free to choose θ as any vector of parameters appearing in a parametric statistical model for asset returns

- (13)-(15) is such a model, so as a starting point, choose $\theta = \mu_f$, the k parameters describing the factor risk premia
- For simplicity we treat F as a constant matrix, just as the original Black-Litterman model treats Σ as a constant matrix

What kinds of views on factor risk premia would arise in practice?

- Most parsimonious case is a view on each factor risk premium that is independent of our views on other factors
- A quantitative portfolio manager might have
  - A view on the value premium, and, separately from that
  - A view on the momentum premium
- It is not typical for portfolio managers to have views on, say, the sum or difference of the value and momentum premia
- Take the likelihood function to be

$$f(Q \mid \theta) = \prod_{i=1}^{k} \exp[-\frac{1}{2\omega_i^2}(\theta_i - Q_i)^2] \tag{17}$$

Types of prior

The choice of prior $\pi(\theta)$ is interesting. We discuss two types:

- Data-driven priors
- The benchmark-optimal prior: Driven by the desire to have some specific benchmark be optimal under the model of the prior as in the definition above

- If the random process driving the unobservable factor returns f_t is stationary (i.e. μ_f, F are approximately constant over time) then we obtain a prior for $\theta = \mu_f$ by taking the posterior from a Bayesian time-series model for the factor returns f_t
- In particular, the historical mean of the OLS estimates $\hat{f}_t = (X_t' X_t)^{-1} X_t' R_{t+1}$ could be taken as the prior mean
- This prior does not require a benchmark $\Rightarrow$ Nice for absolute return strategies where the effective benchmark is cash
- It is common in Bayesian statistics for the posterior from one analysis to become the prior for subsequent analysis

- Given a benchmark portfolio h_B, we can search for a benchmark-optimal prior
- Say $\pi(\theta) \sim N(\xi, V)$ with $\xi \in \mathbb{R}^k$ and $V \in S^k_{++}$, the set of symmetric positive definite $k \times k$ matrices
- Choosing a prior amounts to choosing ξ and V, which are constrained by

$$h_B \in \operatorname*{argmax}_h \int u(h'R)p(R \,|\, \theta)\pi(\theta)d\theta \tag{18}$$

- The first step in evaluating (18) is to compute the a priori density on returns, $\int p(R \,|\, \theta)\pi(\theta)d\theta$. Since $\pi(\theta)$ and $p(R \,|\, \theta)$ are both Gaussian, this is another completion of squares

- Skipping a straightforward calculation, the *a priori* second moments of the returns are

$$\mathbb{V}_\pi[R] = (\Sigma^{-1} + \Sigma^{-1} X H^{-1} X' \Sigma^{-1})^{-1}$$

Similarly, (referring to the Lemma above) the mean is

$$\begin{aligned} \mathbb{E}_\pi[R] &= \mathbb{V}_\pi[R]\Sigma^{-1} X H^{-1} V^{-1} \xi \\ &= (\Sigma^{-1} + \Sigma^{-1} X H^{-1} X' \Sigma^{-1})^{-1} \Sigma^{-1} X H^{-1} V^{-1} \xi \end{aligned}$$

where $H = V^{-1} + X'\Sigma^{-1} X$

The *a priori* optimal portfolio

- The *a priori* optimal portfolio under CARA utility is then

$$(\delta \mathbb{V}_\pi[R])^{-1}\mathbb{E}_\pi[R] = \delta^{-1}\Sigma^{-1}XH^{-1}V^{-1}\xi \tag{19}$$

- Unlike the classical case, it is not true that any arbitrary benchmark portfolio can be realized as an *a priori* optimal portfolio
- Eq. (19) provides a simple characterization of those that can:
  - They are precisely of the form $\delta^{-1}\Sigma^{-1}\Pi$ where Π is some linear combination of the columns of X
  - They are portfolios which are optimal with respect to a set of individual asset risk premia that come from the factor model (not a real restriction; residuals to X should be i.i.d. noise after all)

Can the market portfolio be the benchmark?

- Not every possible portfolio is realizable as *a priori* optimal, but what about the market portfolio, which is the benchmark for many strategies?
- If the market is in a CAPM equilibrium and if one of the columns of X contains the CAPM betas, then the individual asset risk premia will be proportional to that column of X, and then the market portfolio *will* be realizable as *a priori* optimal, as per (19)

Calculating the *a posteriori* optimal portfolio

Now continue with calculating the *a posteriori* optimal portfolio, i.e. the portfolio which takes into account the views on the factor risk premia. This calculation proceeds in three steps:

1. Calculate the posterior distribution of θ, after the views are taken into account
2. Calculate the *a posteriori* distribution of asset returns (also called the posterior predictive density), given by

$$p(R \mid Q) = \int p(R \mid \theta) p(\theta \mid Q) d\theta \tag{20}$$

3. Calculate the mean-variance optimal portfolio under $p(R \mid Q)$

We refer to our paper for the details of these calculations

Explicit formula for *a posteriori* optimal portfolio

The a posteriori optimal portfolio is

$$h^* \;=\; \delta^{-1}\Sigma^{-1}\Pi$$

with

$$\begin{aligned} \Pi &:= X\tilde{\mu}_f \\ \tilde{\mu}_f &:= (V^{-1}+\Omega^{-1}+X'\Sigma^{-1}X)^{-1}(V^{-1}\xi+\Omega^{-1}Q) \end{aligned}$$

- Risk premia $\Pi = X\tilde{\mu}_f$ for each asset are expressed in terms of the (predictable part of the) factor model $R = Xf + \epsilon$, as one would expect
- Think of $\tilde{\mu}_f$ as a set of factor risk premia adjusted to take account of the views
- Adjustments (unsurprisingly) tend to give more weight to factors which have high prior mean-variance ratios ξ_i/V_{ii} and/or high expected return-uncertainty ratios Q_i/ω_i^2

Example: BLB for risk premia portfolios

Data sources & risk premia

Data sources:

- Merged CRSP / IBES, top 2000 common US stocks by market cap, 1992-2015, daily
- "Major group" of the Standard Industrial Classification (SIC) system, about 70 industries

Risk premia we construct:[4]

- Market
- Size
- Value
- Momentum
- Volatility

[4] For details about data processing and risk premia construction, please refer to our paper. For general discussions of these risk premia, see Fama and French (1992), Fama and French (1993), Ang et al. (2006), Connor, Goldberg, and Korajczyk (2010), Menchero, Morozov, and Shepard (2008), and Asness, Moskowitz, and Pedersen (2013).

Factor returns

Figure 1: Cumulative factor returns to risk premia $\hat{f}_t = (X_t' X_t)^{-1} X_t' r_{t+1}$. The two positive-drift risk premia are momentum and value; the two negative-drift risk premia are size and volatility. The exposures to each premia are "gaussianized."

Factor priors: μ_f and F

We assume the random process driving the unobservable factor returns f_t is stationary (i.e. μ_f and F are approximately constant over time)[5] and set our priors as follows:

- μ_f: Components set to zero except for the following four risk premia
  - Momentum = 1.0 bps/day
  - Size = -0.5 bps/day
  - Volatility = -1.5 bps/day
  - Value = 0.5 bps/day
- F: Diagonal matrix with
  - Industry variances set to (10 bps/day)^2
  - Risk premia variances estimated using long-range moving windows
- D: Diagonal matrix estimated as the residual variance from time-series regressions of each stock's return against the S&P 500 return over a rolling two-year window[6]

[5] Note: (1) None of what follows depends on having very precise forecasts of μ_f, F, and (2) the resulting portfolios don't really depend on the industry block of F.

[6] We winsorize low values of residual variance to assure D is well-conditioned.

Factor views: q and Ω

- $q_{i,t}$: One-period-ahead forecast of the i-th risk premium's factor return from a univariate ARIMA model
- ARIMA model selected and fit on an expanding window using AIC$_c$ on a rolling basis (using only data known at time t)[7]
- Ω: For simplicity we set $\Omega = F$

[7] AIC$_c$ is AIC with a correction for small sample sizes.

Markowitz and BLB portfolios

For each day t, the Markowitz and BLB portfolios are computed via

$$\begin{aligned} h_{\text{mar}} &= (\delta\Sigma_t)^{-1} X_t \mu_f \\ h_{\text{blb}} &= (\delta\Sigma_t)^{-1} X_t (V_t^{-1} + \Omega_t^{-1} + X_t' \Sigma_t^{-1} X_t)^{-1} (V_t^{-1}\xi + \Omega_t^{-1} q_t) \end{aligned}$$

where

$$\Sigma_t = X_t F_t X_t' + D_t$$

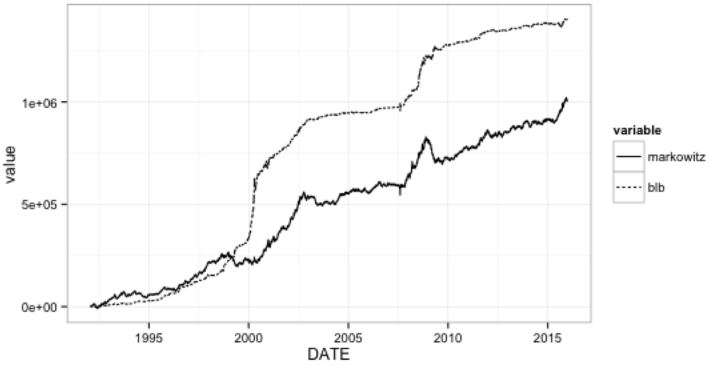

Figure 2: Cumulative returns of the Markowitz portfolio (solid line, IR = 1.3) and the BLB portfolio (by (21)) (dotted line, IR = 1.8).

Cumulative returns to Markowitz and BLB portfolios (2/2)

Remark:

- Sometimes the BLB portfolio actually benefits when the Markowitz portfolio has a drawdown
- This is because the negative trend is picked up quickly enough by the ARIMA model generating the views, and the factor having the drawdown is given a negative view

Conclusions

Conclusions

In this talk we:

- Presented the Black-Litterman-Bayes model, which is the most general model of the type considered by Black and Litterman (1991). In the process we also laid out the full set of assumptions made, some of which are often glossed over
- Fully clarified exact nature of the Bayesian statistical model to which BL corresponds, in terms of the prior, likelihood, and posterior
- Worked out the treatment of views on factor risk premia in the context of APT (Ross, 1976)
- This framework is very general; it allows for tremendous flexibility in expressing priors and views

References I

Ang, Andrew et al. (2006). "The cross-section of volatility and expected returns". In: *The Journal of Finance* 61.1, pp. 259–299.

Arrow, Kenneth J (1971). "Essays in the theory of risk-bearing". In:

Asness, Clifford S, Tobias J Moskowitz, and Lasse Heje Pedersen (2013). "Value and momentum everywhere". In: *The Journal of Finance* 68.3, pp. 929–985.

Black, Fischer and Robert B Litterman (1991). "Asset allocation: Combining investor views with market equilibrium". In: *The Journal of Fixed Income* 1.2, pp. 7–18.

Connor, Gregory, Lisa R Goldberg, and Robert A Korajczyk (2010). *Portfolio risk analysis*. Princeton University Press.

Fabozzi, Frank J, Sergio M Focardi, and Petter N Kolm (2010). *Quantitative equity investing: Techniques and Strategies*. John Wiley & Sons.

Fabozzi, Frank J et al. (2007). *Robust portfolio optimization and management*. John Wiley & Sons.

Fama, Eugene F and Kenneth R French (1992). "The cross-section of expected stock returns". In: *the Journal of Finance* 47.2, pp. 427–465.

— (1993). "Common risk factors in the returns on stocks and bonds". In: *Journal of financial economics* 33.1, pp. 3–56.

Kolm, Petter and Gordon Ritter (2017). "On the Bayesian interpretation of Black–Litterman". In: *European Journal of Operational Research* 258.2, pp. 564–572.

References II

Kolm, Petter N, Reha Tütüncü, and Frank J Fabozzi (2014). "60 Years of portfolio optimization: Practical challenges and current trends". In: *European Journal of Operational Research* 234.2, pp. 356–371.

Litterman, Robert and Guangliang He (1999). "The intuition behind Black-Litterman model portfolios". In: *Goldman Sachs Investment Management Series*.

Markowitz, Harry (1952). "Portfolio selection". In: *The Journal of Finance* 7.1, pp. 77–91.

Menchero, Jose, Andrei Morozov, and Peter Shepard (2008). "The Barra Global Equity Model (GEM2)". In: *MSCI Barra Research Notes*, p. 53.

Pratt, John W (1964). "Risk aversion in the small and in the large". In: *Econometrica: Journal of the Econometric Society*, pp. 122–136.

Robert, Christian (2007). *The Bayesian choice: from decision-theoretic foundations to computational implementation*. Springer Science & Business Media.

Roll, Richard and Stephen A Ross (1980). "An empirical investigation of the arbitrage pricing theory". In: *The Journal of Finance* 35.5, pp. 1073–1103.

Ross, Stephen A (1976). "The arbitrage theory of capital asset pricing". In: *Journal of economic theory* 13.3, pp. 341–360.

Satchell, Stephen and Alan Scowcroft (2000). "A demystification of the Black–Litterman model: Managing quantitative and traditional portfolio construction". In: *Journal of Asset Management* 1.2, pp. 138–150.