Interim Results Report

Group 7: Victor Xiao, Zhaochen Jiang, Changle Li

Apr 13th 2025

Introduction

This report presents our ongoing effort to develop a multi-asset allocation framework that integrates heterogeneous forecasting techniques through an *Agent-Based Black-Litterman* (ABBL) model. Our key objective is to extend the classic Black-Litterman approach (Black and Litterman, 1992; He and Litterman, 1999)—which fuses a market equilibrium prior with subjective views—into a multi-agent environment. Each specialized agent captures distinct return-generating mechanisms in equities, fixed income, foreign exchange (FX), and commodities, while a Bayesian aggregator reconciles their forecasts in a unified portfolio optimization scheme.

Conventional portfolio approaches often rely on a single model or homogeneous market view, potentially overlooking the complex, regime-dependent drivers that differ across asset classes. In contrast, the ABBL framework integrates specialized agents that harness unique signals (e.g., yield-curve factors for bonds, momentum and fundamentals for equities, uncovered interest parity and macro surprises for FX, and basis plus roll yields for commodities). By assigning confidence levels to each agent's forecast and dynamically updating those weights, the model aims to produce robust and adaptive multi-asset allocations.

Data Collection and Cleaning

The multi-asset nature of our framework requires a comprehensive dataset that spans equities, bonds, commodities, and FX, along with relevant macroeconomic indicators. To this end, we draw on three primary sources:

- **WRDS (CRSP and Compustat):** Historical equity returns, corporate balance sheet and income statement fundamentals, and certain fixed-income data.
- **Bloomberg:** Coverage of government bond yields, foreign exchange rates, futures contracts, implied volatility indices, and real-time news or economic releases.
- **FRED:** Supplementary macroeconomic data (e.g., GDP, inflation, employment) with monthly or quarterly frequencies.

Because these sources differ in granularity (daily, monthly, or quarterly) and time span, we have implemented a rolling-window structure that accommodates frequent updates in market-facing agents (e.g., daily or weekly for FX and equities) and less frequent updates for macro-based signals (e.g., monthly or quarterly). The following steps summarize our data handling:

- **Frequency Alignment:** We standardize most asset prices and returns to a daily frequency, carrying forward the latest value of lower-frequency series (macroeconomic indicators) until their next update. This ensures consistent timestamps across equities, bonds, commodities, and FX.
- **Outlier Treatment:** Sudden, large price jumps or erroneous ticks are addressed through a combination of statistical outlier detection and winsorization. We aim to preserve genuine tail events (e.g., a market crash) without allowing spurious data points to skew model estimations.
- **Corporate Action Adjustments:** Equity data is adjusted for dividends, splits, and similar corporate actions. Using CRSP ensures that total returns properly reflect investors' real gains or losses, rather than raw price changes.
- **Data Merging and Splicing:** Overlapping time intervals are carefully unified. If one dataset starts later or ends earlier than another, we splice the overlapping date ranges to form a common analysis window. We also impose strict *no-look-ahead* constraints, making sure that each data point is only matched with information available at that historical moment.

This integrated data environment serves as the foundation for calibrating each specialized agent. When an agent estimates its forecast (e.g., an FX agent blending interest rate differentials with momentum factors), it pulls from a time-synchronized dataset to avoid misalignment or bias.

Exploratory Data Analysis

In this section, we summarize the main findings from our exploratory data analysis, which helped shape our multi-agent design. We focus on the dynamic correlations among asset classes, momentum patterns, and macro-sensitivity, illustrating how a single homogenous forecasting model might miss crucial nuances.

Empirical Techniques

Equities:

We model excess equity returns using linear predictors:

$$R_{t+1}^{\text{excess}} = R_{t+1}^{\text{equity}} - R_{t+1}^{\text{rf}},$$

$$R_{t+1}^{\text{excess}} = \alpha + \beta_1 \left(\text{Valuation}_t\right) + \beta_2 \left(\text{Momentum}_t\right) + \epsilon_{t+1},$$

where Valuation_t can be metrics like dividend yield or earnings–price ratio, and Momentum_t is a short-term return over a recent look-back period.

Bonds:

To forecast bond *returns* rather than just yields, we employ a factor approach (e.g., Cochrane–Piazzesi), letting

$$R_{t+1}^{\text{excess}} = \alpha + \beta\, F_t + \epsilon_{t+1},$$

where F_t can be a forward-rate-based factor derived from yield-curve data. Alternatively, the yield curve itself may be parameterized by a three-factor model (level, slope, curvature), then mapped to bond returns in a rolling regression framework.

FX:

FX forecasting incorporates interest-rate differentials and short-term momentum:

$$R_{t+1}^{\text{FX}} = \alpha + \beta_1\, \Delta i_t + \beta_2\, m_t + \varepsilon_{t+1},$$

where R_{t+1}^{FX} denotes currency (or currency ETF) returns, Δi_t is the interest-rate differential (foreign minus domestic), and m_t captures recent momentum. Some specifications also incorporate event surprises from macro announcements.

Commodities:

Commodity price forecasts frequently use the basis (futures minus spot), plus momentum or seasonality:

$$r_{t+1} = \alpha + \beta\, \big(f_t - s_t\big) + \gamma\, (\text{Momentum}_t) + \epsilon_{t+1},$$

where f_t and s_t are the futures and spot prices, respectively.

Rolling Correlations Across Asset Classes

A key observation in multi-asset contexts is that equity–bond correlation may change over time, alternating between negative (traditional "flight to quality") and positive (periods of synchronized price movements). We computed rolling-window correlations using:

$$\rho_{i,j}(t) \;=\; \frac{\sum_{\tau=t-L+1}^{t} \big(R_{i,\tau} - \overline{R}_{i,t}\big)\,\big(R_{j,\tau} - \overline{R}_{j,t}\big)}{\sqrt{\sum_{\tau=t-L+1}^{t} \big(R_{i,\tau} - \overline{R}_{i,t}\big)^2}\;\sqrt{\sum_{\tau=t-L+1}^{t} \big(R_{j,\tau} - \overline{R}_{j,t}\big)^2}}. \tag{1}$$

Here, $R_{i,t}$ and $R_{j,t}$ denote daily returns on assets i and j, respectively, and $\overline{R}_{i,t}, \overline{R}_{j,t}$ are their mean returns over the preceding L days. Visual inspection of $\rho_{i,j}(t)$ revealed regime-dependent fluctuations, supporting an agent-based paradigm where a specialized bond agent might respond differently to macro signals than an equity agent responding to earnings or sector-based indicators.

Momentum and Seasonality Effects

For commodities and, to a lesser degree, FX, short-term momentum and seasonal cycles emerged as strong patterns. We captured momentum through a k-day look-back:

$$\text{Momentum}_t^{(k)} = \frac{P_t - P_{t-k}}{P_{t-k}}, \tag{2}$$

where P_t is the asset's (or currency's) price at time t. Commodities such as crude oil and agricultural products often displayed seasonal supply–demand cycles (e.g., harvest schedules or weather shifts), which amplify momentum at particular times of the year. Consequently, our *Commodity Agent* prioritizes a combination of momentum and basis signals (futures vs. spot) to better capture these recurring features.

Implications for Agent-Based Forecasting

Taken together, these EDA results confirm that each market segment exhibits unique dynamics and correlation patterns. A single all-encompassing model would likely fail to capture the nuances of the yield curve for bonds, the cyclicalities and basis structures in commodities, and the event-driven nature of FX. By building specialized agents that internalize these features and feeding their forecasts into a Bayesian aggregator, we aim to construct a more robust, flexible, and interpretable multi-asset portfolio allocation scheme.

Interim Empirical Analysis

We conducted exploratory analysis across fixed income, commodities, and FX, uncovering distinct predictive patterns that justify our multi-agent design.

- **Fixed Income (U.S. Treasuries):** Yield curve shifts—level, slope, and curvature—respond strongly to macroeconomic announcements. Cochrane–Piazzesi factor shows moderate, time-varying predictive power in rolling-window regressions. Different tenors display heterogeneous sensitivities to policy expectations and inflation dynamics.
- **Commodities (Agriculture & Energy):** Seasonal price behavior is tied to predictable supply-demand imbalances and geopolitical events. Basis and roll yields correlate with inventory levels and forward demand expectations. Commodities with low inventories tend to experience stronger forward returns.
- **Foreign Exchange (FX):** Interest rate differentials and macro surprises drive short-term return swings. Momentum effects persist across major currency pairs, particularly with wide rate spreads. Cross-currency basis movements reflect policy divergence and market stress.
- **Cross-Asset Patterns:** Equity–bond correlations flip with risk-on/risk-off regimes. FX returns partially co-move with equity risk sentiment. Volatility spikes occur in commodities (supply shocks), FX (monetary divergence), and bonds (rate repricing).
- **Feature Construction:** Weekly-level data includes ETF returns, sovereign yields, macro rates, momentum (1m/3m/12m), and volatility (VIX, MOVE). Asset-specific features are fed into modular LLM agents for tailored forecasting.

Preliminary Results

We implemented an LLM-based forecasting pipeline with structured prompts and JSON outputs, producing return, confidence, and rationale for each asset and time horizon.

- **Forecasting Framework:** GPT-4o models receive weekly macro and market features, and output deterministic (temperature = 0) predictions. Outputs follow a strict schema for interpretability and downstream portfolio integration.
- **Forecast Behavior:** LLMs tend to avoid extreme views. Forecasts are conservative and centered around moderate expectations. This behavior ensures stability but may underreact to sharp macro shifts.
- **Portfolio Integration (Black-Litterman):** Forecasts feed into Black-Litterman models, combining LLM views with historical priors under quadratic regularization. This reduces overfitting and volatility, yielding stable portfolio weights.

- **Performance Highlights:** Fixed income BL portfolio shows a Sharpe ratio of 0.76 with low drawdown and moderate turnover. FX BL portfolio delivers modest returns with strong risk control, outperforming equal-weighted benchmarks. FX directional accuracy is around 59%, with USD/CAD and USD/CHF leading.
- **Insights on Model Confidence:** Confidence does not always align with accuracy. Mid-confidence (0.50–0.60) bins outperform high-confidence predictions. This suggests a need for improved calibration or the use of ensemble models.

Robustness and Causality

Although a detailed robustness analysis is still under development, we plan to use the following approaches to ensure reliability and assess potential causal links:

- **Regime-Switching Checks:** We will examine the performance of the ABBL framework under different volatility regimes (e.g., *low-vol* vs. *crisis* periods) to confirm that specialized agents can adapt to structural breaks or macroeconomic shocks.
- **Out-of-Sample Tests:** We will conduct rolling out-of-sample forecasts and trading simulations to validate whether the aggregated portfolio indeed generates risk-adjusted returns superior to standard benchmarks (e.g., equal-weighted or naive momentum).
- **VAR or Lagged Regression:** For certain assets (particularly bonds or FX), we will investigate lagged relationships between model predictors (e.g., yield-curve slope, macro announcements) and realized returns, aiming to distinguish causal factors from correlations.
- **Sensitivity Analysis:** By systematically varying each agent's confidence levels or excluding certain signals (like momentum in commodities), we can assess which components materially influence the final portfolio, thereby testing the stability of our architecture.

Plans for the Next Step

Moving forward, our focus lies on refining and expanding the ABBL model:

1. **Enhancing Agent Forecasting:** Incorporate additional features such as news sentiment or central bank communications for the FX agent, yield-curve decomposition for the bond agent (e.g., Nelson–Siegel), and advanced factor models for the equity agent.
2. **Dynamic Confidence Scoring:** Develop a scheme where each agent's confidence weight adapts based on recent predictive accuracy or root mean squared error (RMSE), allowing high-performing agents to exert greater influence on the posterior returns.
3. **Tool-Calling and Reinforcement Learning (Optional):** Explore the use of large language models (LLMs) with external data queries (e.g., real-time data or news APIs) and reinforcement learning to further customize or automate the agent-based framework.
4. **Extended Backtesting and Stress Testing:** Conduct deeper analyses over multiple sample periods, including stress scenarios such as financial crises or major geopolitical events, to confirm the resilience and adaptiveness of the ABBL allocations.

Overall, these steps are designed to evolve the ABBL concept into a fully functional, data-rich system capable of delivering robust multi-asset portfolios. By systematically combining agent-specific expertise and Bayesian aggregation, we aim to exhibit improved risk-adjusted performance, minimal drawdowns, and responsiveness to rapidly changing market conditions.

References

Black, F. and Litterman, R. (1992). Global portfolio optimization. *Financial Analysts Journal*, 48(5):28–43.

He, G. and Litterman, R. (1999). *The intuition behind the Black-Litterman model portfolios.* Goldman Sachs Asset Management.