Modeling Corporate Bond Returns Using IPCA

By: Marcus Walton and Neil Nieh

Table Of Contents

What Are Corporate Bonds?

Debt Instrument

Corporate bonds are debt securities companies use to raise capital without giving up ownership. They function like loans from investors to the company.

Investor Payments

- Bondholders receive fixed interest payments (called coupons) on a regular schedule, along with full repayment of the principal at maturity.

Risk & Return Tradeoff

- Corporate bonds usually offer higher yields than government bonds, but come with added credit and market risk depending on the issuer’s financial health.

How The Market Works

Primary vs. Secondary Market

- **Primary Market:** Companies issue new bonds to raise capital, typically through investment banks.
- **Secondary Market:** Investors trade existing bonds with one another. Prices fluctuate based on supply, demand, and market conditions.

Pricing Influences

Bond prices are shaped by interest rates, credit ratings, and macroeconomic sentiment. As interest rates rise, bond prices typically fall, and vice versa.

Key Participants

Institutional investors, market makers, dealers, and brokers all play roles in bond issuance and trading.

Microstructure of Bonds

OTC Trading (Over-the-Counter)

Most corporate bonds are traded OTC rather than on centralized exchanges. This means transactions are negotiated directly between parties, leading to limited price transparency.

Market Inefficiency

Corporate bond markets tend to be less liquid and more fragmented than equity markets. As a result, prices can deviate from fair value, especially for lower-volume bonds.

Role of Intermediaries

Dealers and brokers play a key role by facilitating trades and holding inventory. Their participation influences bid-ask spreads and execution speed.

Intro To IPCA

Instrumented Principal Component Analysis (IPCA) is a factor modeling technique designed to explain bond returns, also useful in financial markets where there are roch cross-sectional characteristics

What is IPCA?

- A factor model that uncovers hidden drivers of bond returns using both bond data and firm-level traits.

Why IPCA Over PCA?

- It captures time-varying relationships and uses observable features to improve accuracy—unlike PCA, which is static and purely statistical.

Application in Bonds

- Well-suited for modeling corporate bonds where risk exposure shifts over time across issuers.

Bryan Kelly's Paper

"Modeling Corporate Bond Returns" – Bryan Kelly, Diogo Palhares, Seth Pruitt

Purpose

Introduces the IPCA model to improve understanding of corporate bond return drivers through dynamic, factor-based modeling.

Key Findings

- Higher-than-expected credit risk premium
- Stronger bond-equity market integration
- IPCA outperforms traditional factor models in explaining returns

Application to Our Project

- Serves as the foundation for our modeling approach
- We benchmark our portfolio performance and Sharpe ratios against the paper's published results

Business Value

Buy-Side

Value

- Uncovers dynamic risk exposures, improving portfolio efficiency and precision in alpha generation.
- Enhances portfolio diversification by focusing on time-varying, firm-specific risk factors, reducing reliance on static models.
- Provides more transparent, interpretable factors, helping traditional PMs gain confidence in data-driven strategies.

Example

- A BlackRock portfolio manager could use IPCA to refine exposures in smart beta or actively managed equity funds, improving Sharpe ratios by focusing on impactful traits like quality, momentum, and size.

Business Value

Sell-Side

Value

- Enables more accurate pricing and reduces mispricing risks in structured products and bonds.
- Supports creation of innovative, factor-based ETFs, indexes, or notes using IPCA-identified signals.
- Strengthens research and client advisory with data-driven insights on true return drivers, enhancing client trust.

Example

- Goldman Sachs could use IPCA on their structured products desk to issue bond-linked notes that exploit IPCA-identified credit risk factors, offering clients differentiated exposure and fee revenue.

Preprocessing

Feature Engineering

- Created bond-based features (e.g., duration, spread, rating) and firm-level variables (e.g., market cap, profitability).
- Included volatility, skewness, and liquidity measures to capture risk factors relevant for IPCA.

Credit Excess Return Calculation

- Remove interest rate risk and isolate pure credit return.
- Calculated Credit Excess Return by subtracting the return of a duration-matched Treasury hedge portfolio from the bond's total return (coupon + price) – ICE's KRD approach
- This isolates credit risk exposure, providing clean inputs for the IPCA model.

Model Performance

In-Sample Results

Num of Factors	Intercept	Total R^2	Time Series R^2	Cross Sectional R^2
1	True	0.443	0.042	0.329
	False	0.434	0.013	0.305
2	True	0.488	0.306	0.368
	False	0.482	0.290	0.355
3	True	0.505	0.314	0.392
	False	0.502	0.292	0.393
4	True	0.575	0.343	0.435
	False	0.568	0.308	0.420
5	True	0.582	0.337	0.442
	False	0.579	0.319	0.437

Factor-Count Effects:

- Rising Total R^2: As K goes from 1 to 5 factors,Total R^2 climbs from ~44% to ~58%.
- Sharply Up at K=4

Diminishing Returns:

- Gains taper off beyond 4 factors, suggesting that the marginal benefit of each extra factor is smaller.
- Kelly's report has similar pattern in-sample: Total R^2 goes from ~42% at K=1 to ~50% atK=5 .

Model Performance

Out-of-Sample Results

	Training Window	Lag	Factors	Intercept	Total R^2
Kelly	Expanding	Unspecified	5	No	50.7
Our Model	120	25	5	Yes	36.5
				No	38.7
			4	Yes	39.2
			3	Yes	39.3
			1	Yes	39.1

Improvements of IPCA

Can we improve the model by adding new features?

1 Delta Distance-to-Default

Kelly shows that D2D is one of the most important features

2 Spread per Duration

Pure “value” signal: bonds with high spreads relative to their duration are cheap on a risk-adjusted basis, while low values flag potentially overvalued bonds.

Model	Number Factors	Total R^2 (In-Sample)	Time Series R^2	Cross Sectional R^2
Original	5	0.578	0.316	0.422
New Features		0.579	0.319	0.424

Machine Learning Implementation

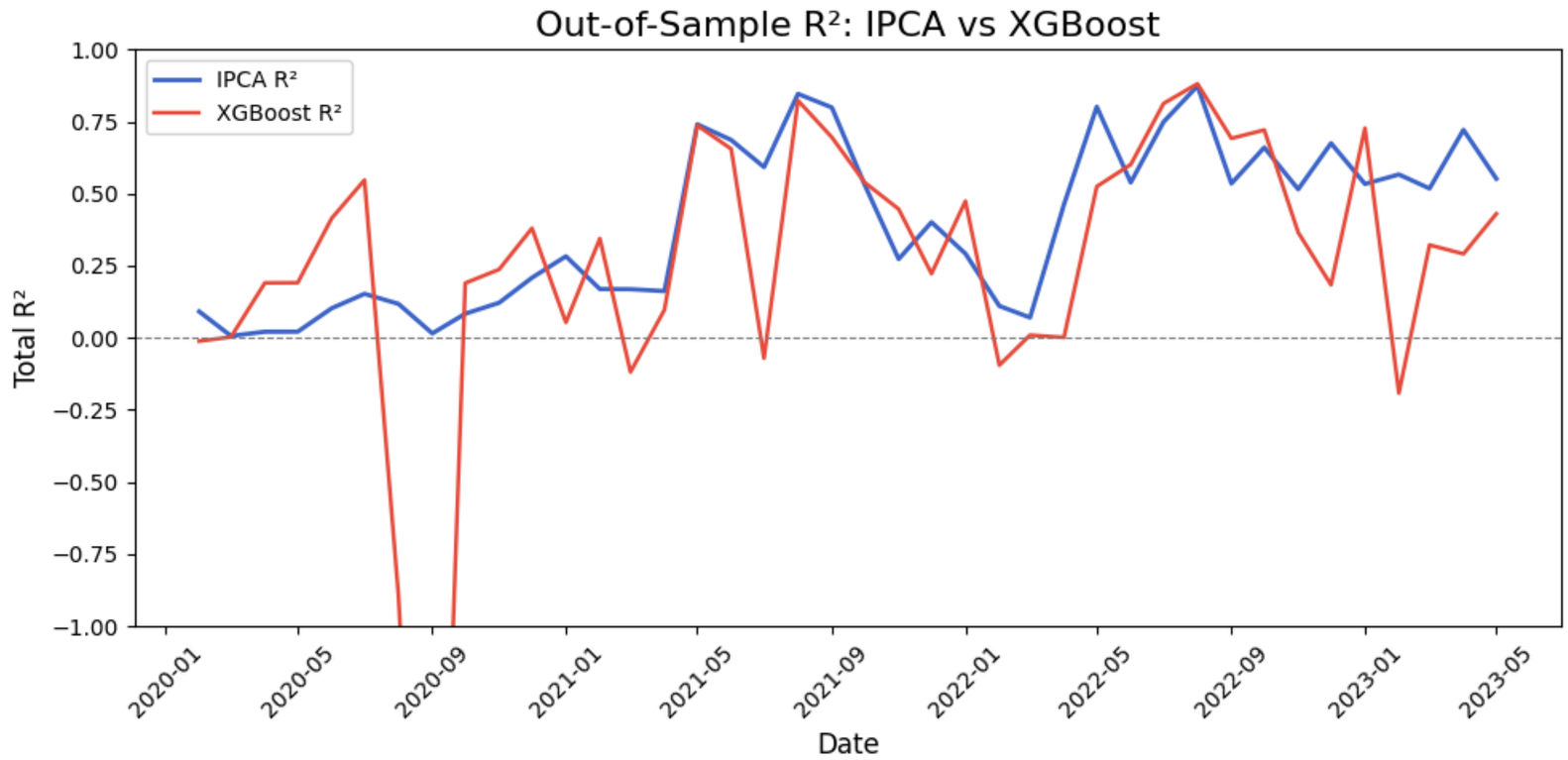
We'll be trying XG Boost and LSTM model and compare the performance of IPCA

Model	Reasons	Total R^2	Time Series R^2	Cross Sectional R^2
XG Boost	Highly accurate, great with structured data, fine-grained control	0.264	0.254	0.207
LSTM	Excellent for time series forecasting, captures long-term patterns	0.021	0.014	0.077
XG Boost	After hypertuning	0.372	0.355	0.269

IPCA vs. XG Boost

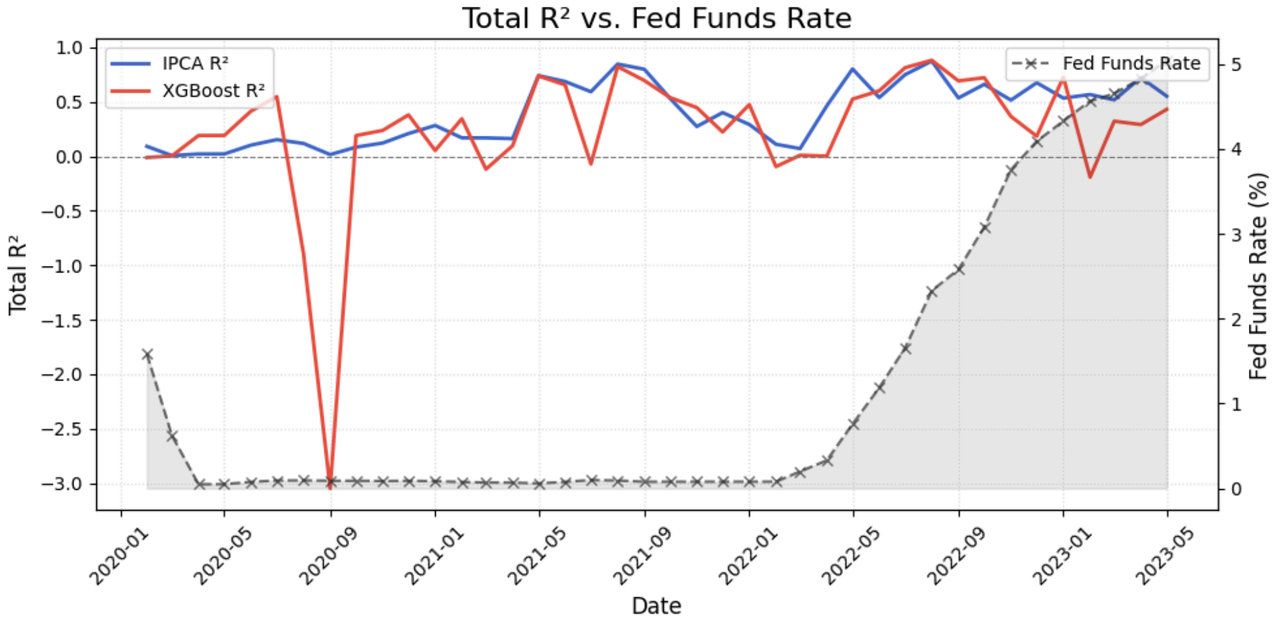
Comparison

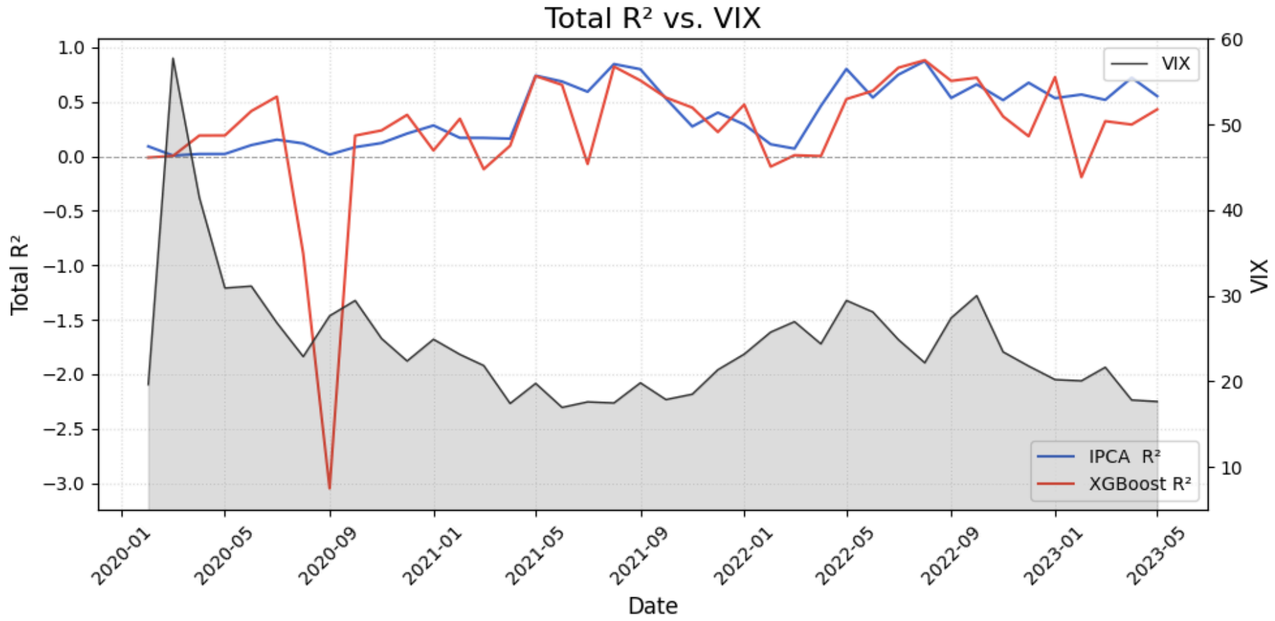
We then perform OOS with same traing window(= 120) and lag (= 25)



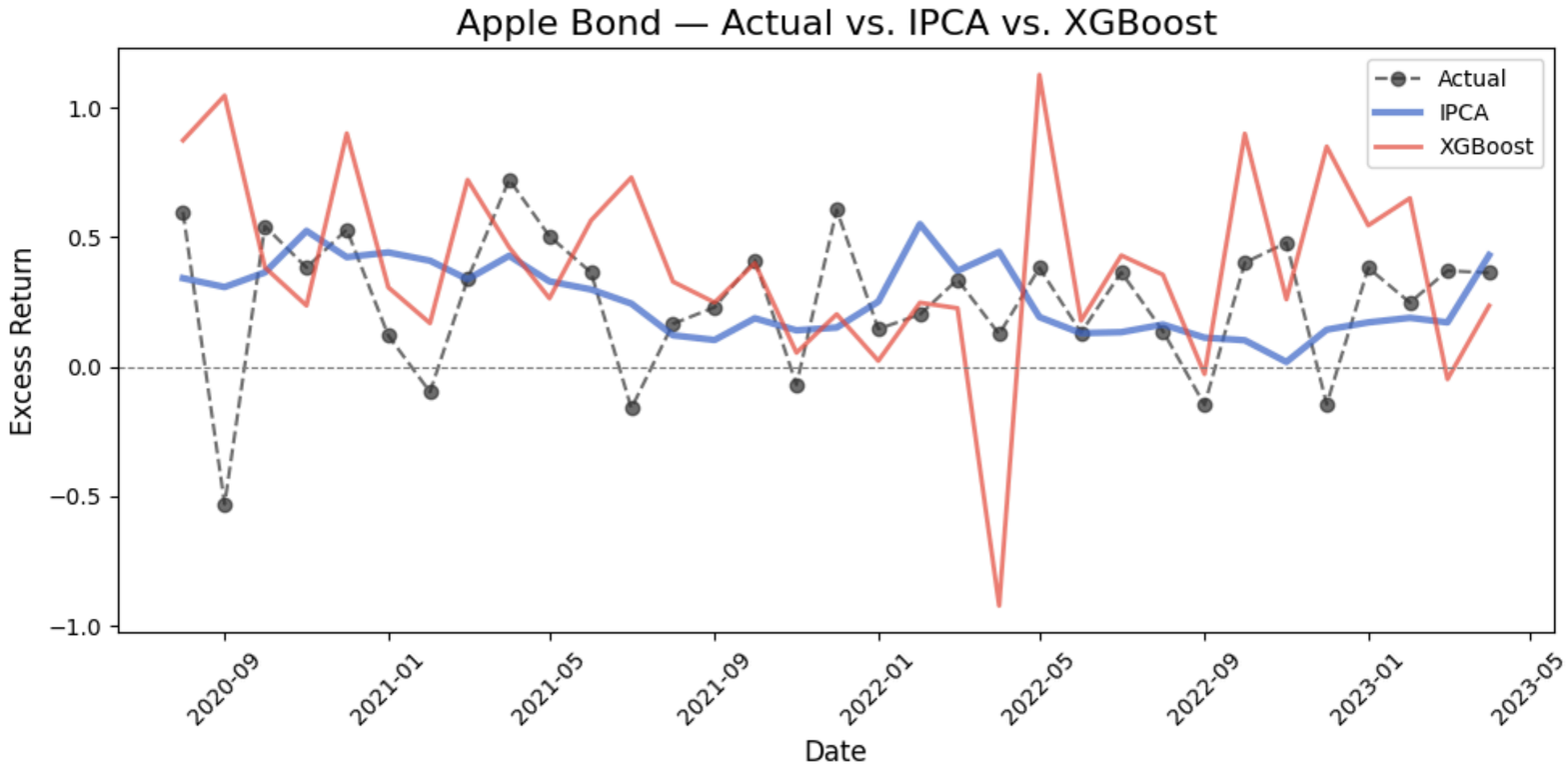
IPCA vs. XG Boost Comparison

Do Fed Fund Rates/VIX impact on prediction power?





IPCA vs. XG Boost Comparison



	MAE
IPCA	0.225
XG Boost	0.334

CUSIP: 037833DF

Long-Short Portfolio Explained

Explanation

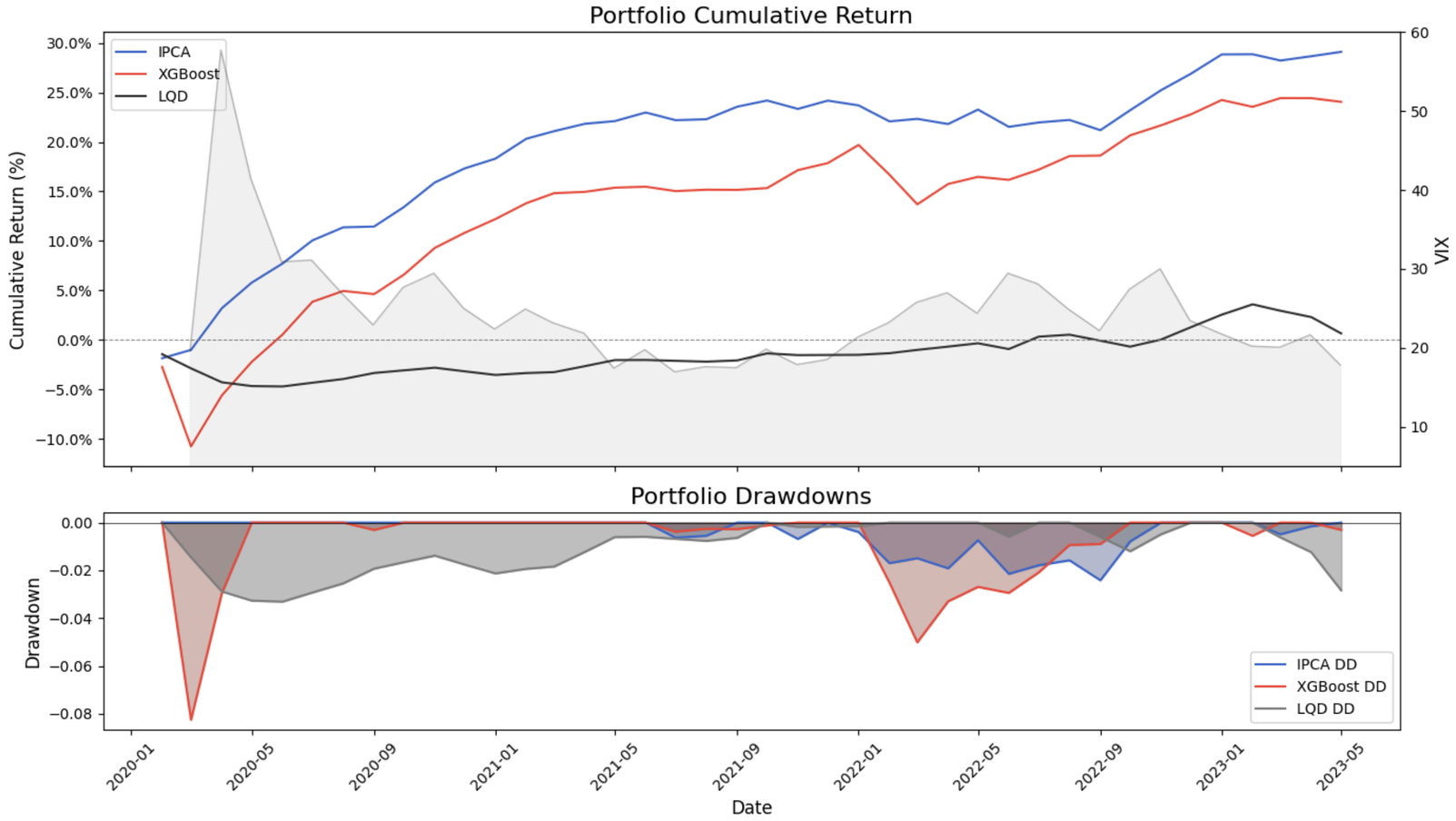
- Rank Bonds by Predicted Excess Return
- Form a Equal-Weight Long-Short Portfolio
  - Long top 10%, short the bottom 10%
- Incorporate Trading Cost
  - 17/19 bps one-way cost

Why a Long-Short Portfolio?

- **Isolates Alpha:** By offsetting longs and shorts, we strip out common drivers (e.g., shifts in overall yield levels or duration)
- P&L reflects purely the cross-sectional “spread” between under- and over-priced bonds.

	Annualized Return	Sharpe Ratio (Net Cost)	t-stat
IPCA	8.05%	1.62	12.4
XG Boost	6.97%	0.94	5.9

Long-Short Portfolio Examination



Conclusions

Findings & Possible Improvements

- **Recap of Objectives:**
  - Define Business Value
  - Validate Kelly's Results
  - Benchmark Against ML Models
- **Future Improvements**
  - Find Reasons Causing OOS R^2 Difference
  - Hybrid ML Approach
  - Augmented Feature Discovery

Model	Total R^2 (In-Sample)	OOS R^2	Portfolio Sharpe
Kelly's	0.49	0.51	1.94
Our Model	0.57	0.38	1.62

Thank You