Final Report: Swiss RE

Analytics Amigos: *Elisa Becker-Foss, Abhranil Chakrabarti, Bibhabasu Das, Arushi Jain, Jillian Puskas*

PROJECT OVERVIEW

Over the past two months, our team has worked with Swiss RE to help individuals hedge against commodity inflation. In our current inflationary environment, individual consumption is getting increasingly expensive. However, an individual's inflation exposure is determined by their unique commodity consumption patterns, and as such, no two individuals have the exact same inflation exposure. Thus, we envisioned that no two individuals should have the exact same inflation hedging strategy. ***By predicting future commodity prices with the publicly available market and macroeconomic data, our team was able to create an optimized portfolio of commodities that reduces an individual's inflation exposure by 82%.***

Summary of Accomplishments:

- ★ Conducted research to gain a deeper understanding of inflation and commodities trading
- ★ Identified, loaded, and tested 30+ macroeconomic variables in addition to various commodity spot price lags and futures to understand their ability to predict changes in commodity spot prices and analyzed the relationships among them
- ★ Iterated through several phases of predictive modeling
  - ○ Baseline: Used current commodity prices to predict commodity prices 3 months into the future
  - ○ Linear: Ran linear regression models to select the most effective features of our 30+ factors
  - ○ Ensemble: Ran ensemble models to make accurate commodity spot price forecasts
  - ○ Time Series: Ran several time series models including exponential smoothing, ARIMA, FB Prophet, and Auto TS to model spot price seasonality and trend more accurately, selected FB Prophet to predict the commodity price 3 months into the future within 10% MAE of the true dollar value
- ★ Built an optimization model that uses our predictions to create 'the perfect inflation hedge' commodity portfolios for individuals based on their unique consumption that reduces their inflation losses by 82%, an improvement of 30% over blind equal investment portfolio for a 3-month horizon

PURPOSE AND SCOPE

<u>The Problem</u>: Our current high inflationary environment is squeezing the disposable income of households. The effect of inflation on the wealth of an individual is twofold: First, the overall operational expenditure increases due to the rising costs for consumption. Second, overall return on investments decreases due to the erosion of real returns during inflation jumps. The net result is that inflation causes severe erosion of wealth for the common people. For Swiss RE, the current high-inflationary period in the market also has a severe downside impact on their existing portfolio. The purpose of this project is to help Swiss RE find a solution to protect customers' wealth against inflation costs.

<u>The Solution</u>: The proposed solution for Swiss RE involves:

1. <u>Analyzing relationships between commodity prices, inflation proxies, and macroeconomic factors</u>
   We analyzed the relationship between commodity prices, commodity futures, and several macroeconomic factors. This allowed us to test which proxies were most effective in modeling as well as to find unique data sources that help predict inflation outcomes.
2. <u>Predicting Commodity Inflation</u>
   With the understanding gained in Step 1, we then attempted to predict future inflation per commodity based on historic commodity prices, commodity futures, and macroeconomic factors through linear modeling, ensemble methods and most importantly, time series models.
3. <u>Recalibrating this index based on the individual consumption pattern for the optimal hedge</u>
   We eventually recommended an optimal trading strategy for a consumer personalized by their consumption pattern, budget constraints and risk appetite.

<u>Evaluation Metrics</u>: Our team created baselines and benchmarks for both our commodity price forecasts as well as the optimal portfolio construction. The baseline prediction model used 3 months old commodity prices as the current commodity price. The baseline investment model assumed that customers invest in all commodities equally irrespective of their consumption. The hypothetical benchmark for our model was a perfectly neutral net position regarding the change of commodity prices given a certain consumption pattern. Our model was backtested based on the historically available data.

<u>Scope Limitations</u>: Our team was limited by the predictive power of the open source data that we found, particularly, their frequency and "causal nature" in predicting price changes.

DATASET

Data Collection

Since most of the data related to commodity spot prices, futures, and various macroeconomic variables and inflation indicators is publicly available, our team's biggest challenge was identifying the relevant features from a seemingly infinite number of data sources and systematically extracting them to create a comprehensive dataset.

Commodity Spot Prices

There are over 30 commodities that are openly traded across several exchanges. However, it is infeasible to create a model for every single commodity. Hence, instead of working with individual commodities, we selected a handful of commodity subindices that directly impact the average consumer. For each commodity category, the respective subindices contain weighted combinations of individual commodities. We decided to take 5 commodity subindices (Agriculture, Livestock, Energy, Industrial Metals, and Precious Metals) over a period of 20 years from Bloomberg. Please see Table 1 for more detail.

Commodity Futures

Futures were a late, but critical component to our best-performing models, proving that they contain lots of additional information beyond the macroeconomic factors we were able to pull in. We pulled futures data from Bloomberg for the same 5 selected commodity subindices. Please see Table 1 for more detail.

Commodity Subindex	Description	Bloomberg Ticker (Spot)	Bloomberg Ticker (Futures)
Agriculture	Measure price movements of coffee, corn, cotton, soybeans, soybean oil, soybean meal, sugar, wheat	BCOMAGSP	BCOMAG
Livestock	Measure price movements of live cattle, lean hogs	BCOMLISP	BCOMLI
Energy	Measure price movements of crude oil, heating oil, unleaded gasoline, natural gas	BCOMENSP	BCOMEN
Industrial Metals	Measure price movements of aluminum, cobalt, palladium, copper, nickel and zinc	BCOMINSP	BCOMIN
Precious Metals	Measure price movements of gold, silver, platinum	BCOMPRSP	BCOMPR

Table 1: ***Commodity spot and futures pulled from Bloomberg***

Other Macro Economic Parameters

We also investigated a wide array of macroeconomic parameters to help our model predict commodity-based inflation. Macroeconomic data were selected using a combination of research and data availability. This paper by the Federal Reserve modeling commodity prices published in 2017 helped formulate many of our early ideas on what types of data could influence commodity inflation. Through our research, we hypothesized that the macroeconomic data given in Table 2 would help us predict commodity prices:

Variable Category	Macroeconomic Variables	Description	Bloomberg Ticker/ Source	Frequency
Inflation (Global inflation measures are leading indicators of commodity prices)	CPI (Consumer Price Index)	Consumer prices are a measure of prices paid by consumers for a market basket of consumer goods and services.	CPURNSA	Monthly
	Money Supply M1	M1 money supply includes those monies that are liquid such as cash, checkable (demand) deposits, and traveler's checks	FRED M1	Monthly
	Money Supply M2	M2 money supply is less liquid in nature and includes M1 plus savings and time deposits, and money market funds	FRED M2	Monthly
Interest Rates (High interest rates depress inflation and therefore commodity prices)	Real Yield 2Y	Real yield on 2-year US Treasury	RR2YCUS	Daily
	Real Yield 10Y	Real yield on 10-year US Treasury	RR10CUS	Daily

	US Treasury Bills	Nominal 13-week US Treasury Bill interest rate	[U.S. Department of Treasury]()	Daily
Demand and Growth (Increasing demand drives up inflation and commodity prices)	US GDP	US Real GDP (Annual YoY%)	EHGDUSY	Yearly
	GDP of importing countries	Canada, China, EU, Mexico, Japan, Germany, India, Colombia,, Saudi Arabia, Russia	[World Bank]()	Yearly
	Dollar Index	Broad Nominal Index of the dollar or dollar spot rate	BRDXY	Daily
	Unemployment	US Unemployment	[U.S. Bureau of Labor Statistics]()	Yearly
Commodities	Gold Spot Rate	Daily price of gold	XAUBGN	Daily
	Silver Spot Rate	Daily price of silver	XAGBGN	Daily
Health of Financial Markets (Market volatility and uncertainty impacts commodity prices)	MSCI Equity Index	Weighted market-cap index designed to accurately represent and measure global equity markets used to signal market health	MSCI	Daily
	VIX Index	The index signals the level of fear or stress in the stock market	VIX	Daily
	Equity Risk Premium (ERP)	The index measures the spread of returns of U.S. stocks over long term government bonds.	[Historical Equity Risk Premium]()	Yearly
Supply (Increasing supply decreases the commodity spot prices)	Industry Production	The industrial production (IP) index measures the real output of all relevant establishments located in the United States.	[FRED INDPRO]()	Monthly
	GDP of exporting countries	Canada, China, Mexico, Japan, Korea, Latin America	[World Bank]()	Yearly

Table 2: ***Macro economic data pulled from Bloomberg, World Bank and Federal Reserve (FRED)***

Data Cleaning

Considering that our dependent variable (commodity index spot) is high-frequency (daily) but our independent variables are mixed frequencies (daily, monthly, yearly), we needed to bring them to the same scale to be able to create a reasonable dataset. We believed that a daily frequency for modeling will be too volatile while a yearly frequency will be too low and hence, decided to bring all of our data to monthly level.

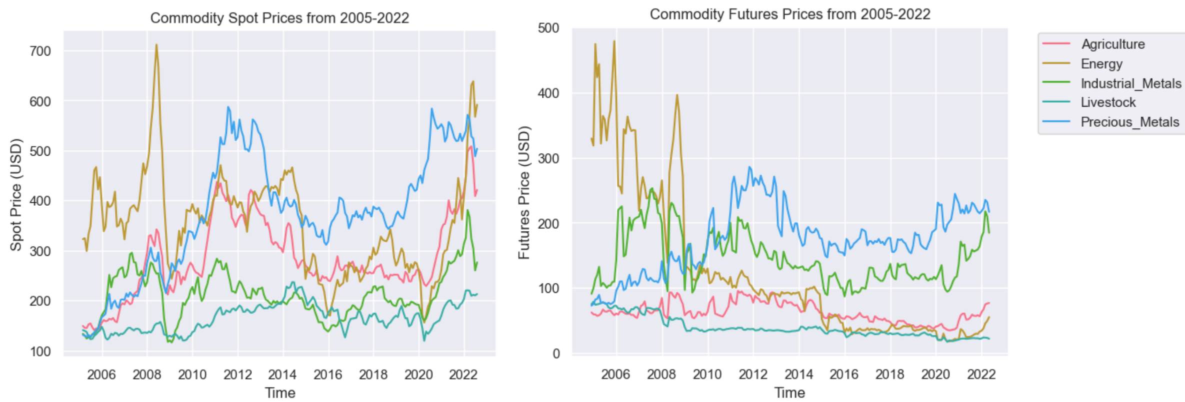
- Interpolating Yearly Data: The yearly data (such as GDP, Unemployment) was linearly interpolated to 12 months
- Smoothing Daily Data: The daily data (such as commodity prices, interest rates, financial indices) was smoothed across each month by taking a simple average
- Missing Data: Some of the independent variables were missing before 1st Jan 2005. Hence, to ensure uniformity, we removed all data prior to 1st Jan 2005
- Rates of Change: For some predictors such as CPI, we added their change over time as another dependent variable
- Normalization: We did a min-max normalization for all the predictors to bring them to the same scale

This meant that our 20 years of data quickly became just 210 data points, which made our modelling more challenging from the perspective of overfitting and complexity.

Basic EDA

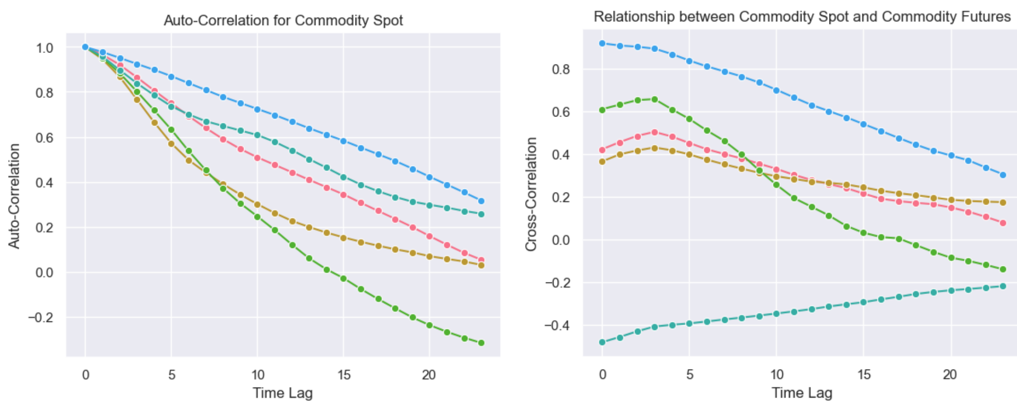
We analyzed the time-series data and the relationship among various time-series across different time lags to understand our data better. We tried looking for seasonality, general trend and strong relationships/ correlations among commodities spot, commodities futures and certain predictors.

Visualizing Commodity Spots and Futures across Time



Inference: We don't observe any strong seasonal trend across commodity spots and futures

Visualizing Relationship among Commodity Spots and Commodity Futures



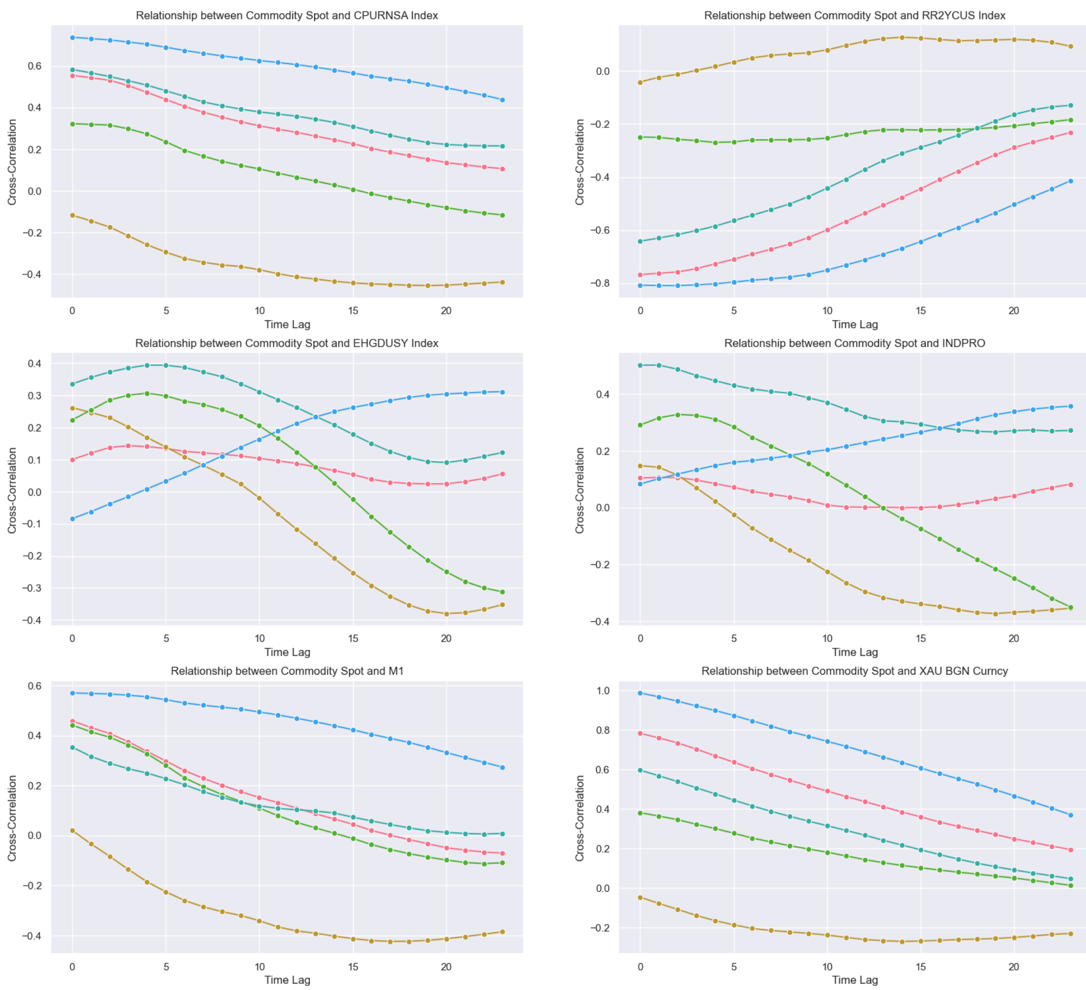
Inference: The auto-correlation is strong for the first 6 months for all the 5 commodities. However, we don't see any peaks after certain time periods which confirms the lack of proper seasonality. The cross-correlations among the commodity spot prices and the respective futures prices are strong for the first 3 months.

Visualizing Relationship among Commodity Spots and Predictors

Inference: We compute the cross-correlations of all commodity spot prices with specific predictors:

- CPURNSA Index (CPI): Strong positive correlations in the first 6 months with Agriculture, Livestock and Precious Metals
- RR2YCUS Index (Real Yield 2 Year): Strong negative correlations in the first 12 months with Agriculture, Livestock and Precious Metals
- EHGDUSY Index (US GDP): Moderate positive correlation in the first 12 months with Livestock
- INDPRO (Industrial Production): Moderate positive correlation with Industry Metals in the first 3 months and strong positive correlation with Livestock in the first 6 months
- M1 (Money Supply): Strong positive correlations in the first 3 months with Agriculture, Livestock, Industry Metals and Precious Metals
- XAU BGN (Gold Price): Strong positive correlations in the first 12 months with Agriculture, Livestock, Industry Metals and Precious Metals

Note: All cross-correlations have commodity prices as the lagging variable



Based on the EDA, we found that the dependent variable (spot prices) should be predicted at a gap of atmost 3 months to maintain the predictive power of our independent variables.

METHODOLOGY

In order to recommend a perfect hedge based on individual consumption, we first need to predict the future prices of the commodities. Hence, our modeling consists of 2 major steps:

1. Predicting future commodity indices spots based on past prices, futures, and macroeconomic indicators
2. Solving a robust optimization problem that balances individual consumption with the recommended investment using the predicted commodity index prices

Predicting the Future Commodity Spot Indices

For **each of the 5 commodity indices,** we train various models to predict their spot prices 3 months into the future as accurately as possible. We will first give an overview of all the techniques that we tried and then summarize the results and findings towards the end.

Baseline

The baseline model simply uses past commodity prices (t-3) as the prediction for the current commodity price (t). It doesn't use any macroeconomic information or commodity futures information.

To improve over the baseline, we will train regression models that use all of our features as well as futures data to predict the commodity spot indices 3 months into the future. To achieve this, we will lag all of our independent variables by 3 months so that our dependent variable is predicted using 3-month-old data. For each model, we will follow the below sets to make the comparison fair:

- *Training*: 1st Jan 2005 to 1st Aug 2020

- *Testing*: 1st Sep 2020 to 1st Aug 2022
- *Number of Models*: 5 (Agriculture, Energy, Industry metals, Precious metals, Livestock)

Linear Modeling

We first tried a linear model to understand variable importance among the 30+ macroeconomic factors and remove any correlated or insignificant features (p-value < 0.1).

Initial Features**:** RR2YCUS (Real Yield), CPURNSA (CPI), Delta CPURNA (Change in CPI), EHGDUSY (US GDP), M1 Money Supply, M2 Money Supply, INDPRO (Industrial Production), 13-week Bill Rate, VIX Index (Volatility), MSCI US Equity, BBDXY (Broad Dollar), XAU BGN Curncy (Gold Spot), XAG BGN Curncy (Silver Spot), Unemployment Rate, ERP (Equity Risk Premium), GDPs for Russia, Canada, China, India, Mexico, Japan, Saudi Arabia, and European Union + Respective futures (Agriculture, Energy, Industry metals, Precious metals, Livestock)

Feature Selection**:** Across all 5 models, we noticed that no GDP data beyond the US GDP had any predictive power over commodity prices. Our research suggested that the economic growth of countries heavily involved in either producing or consuming the given commodity could be highly predictive, but we did not find that to be the case. It was probably because the GDP data of other countries is more of a noise rather than a signal in driving the commodity prices in the future since it encapsulates much more than mere import or export data. A more accurate metric would have been US exports to these respective countries and US imports from these respective countries in order to understand the demand-supply curve within the US that actually determined the market prices of the commodities in the US.
We also found that CPURNSA Index, EHGDUSY Index, and RR2YCUS Index were important in predicting almost all of the future commodity indices.

Results**:** This linear model had too many features and performed only modestly better than our baseline, performing best on Precious Metals and worst on Industrial Metals. This is because we only had 187 rows of training data (monthly frequency) and 10+ features which results in heavy overfitting for the linear model.
The findings are summarized in Table 3. Overall, the linear model only served to bring down the number of features and pave way for more advanced models with customized features for each commodity.

Ensemble Modeling

Having used the linear model to better understand which macroeconomic factors are most effective at predicting all the commodity indices, we moved on to more complex modeling. Our linear model didn't utilize the time-series information of current spot prices in making future predictions. This was partly intentional because we wanted to understand the predictive power of our other features instead of focusing on accurate predictions. However, now we want to customize the features for each of the 5 sub-indices to predict their spot prices in the future as accurately as possible.

We created the time-series information of our current spot prices lagged by 3 months (t-3) as an additional feature for each of the 5 respective models. We use the same training and testing data as the linear model for comparison. However, the testing is done in a rolling fashion here where at each test data point at time t, it is re-trained till time t-3 to make predictions for t-2, t-1, and t. Hence, using this model will require retraining at least every 3 months for the best results.

Feature Selection**:** For each model, we use a combination of forward and backward selection technique to identify the best-performing features. We started with the initial features of lagged spot prices, futures prices, CPURNSA (CPI), EHGDUSY (US GDP), and RR2YCUS (Real Yields) and then added and removed further features based on the model performance on the training set.
The findings and selected features are summarized in Table 4.

Results: XGBoost did outperform our baseline and linear model but there still exists scope for improvement.

Time Series Modeling

Since we are dealing with time-series data in all of our features as well as the dependent variable, leveraging time-series modeling is the most obvious final step. The linear model and the ensemble methods cannot model trends and seasonality across time as well as a time-series model can. We tried the following time-series models:

- ARMA - Autoregressive models assume that there is a linear relationship between current returns and their own history. Moving average models capture the fact that returns depend not only on current information, but also on signals that have arrived over a previous stretch of time. ARMA combines the Auto-Regressive and Moving Average component but is limited in its application as it assumes a weakly stationary process which is seldom satisfied in practical applications.
- Exponential Smoothing - Exponential smoothing is a heuristic approach for smoothing time series data using the exponential window function. In this technique, exponential functions are used to assign exponentially decreasing weights over time, giving greater importance to more recent observations.
- FbProphet: Prophet is an open-source time series forecasting algorithm designed by Facebook for ease of use without any expert knowledge in statistics or time series forecasting. Prophet builds a model by finding the best smooth line which can be represented as a sum of the overall growth trend and seasonality components using the given time series as well as any additional regressors (futures and macroeconomic variables in our case).

- AutoTS: Automatic Time Series is an advanced model that uses genetic programming to train multiple time series models to find the most optimal forecasts. It uses an ensemble of diverse time-series models.

Each of these models was tested by re-training till t-3 periods in a rolling fashion similar to the Ensemble Model testing to ensure that the models are not predicting too much into the future and have recent information available. The feature selection technique was also similar to the Ensemble Modelling.

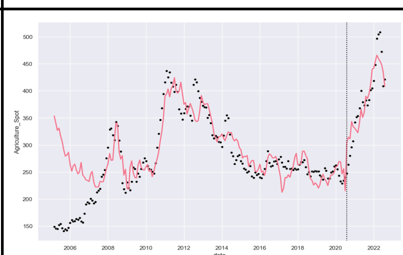
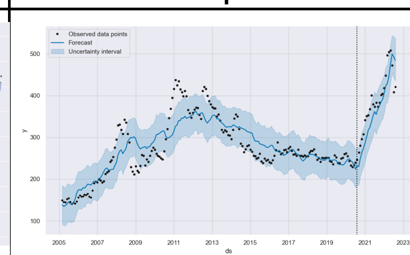
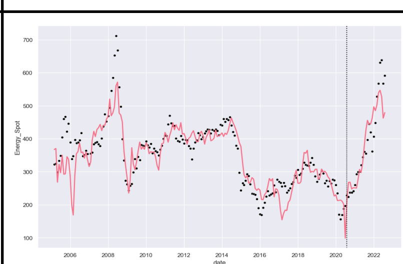
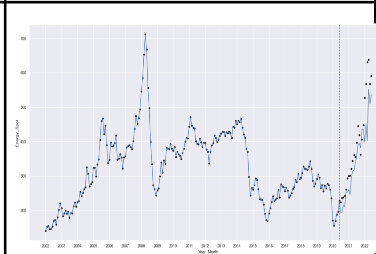
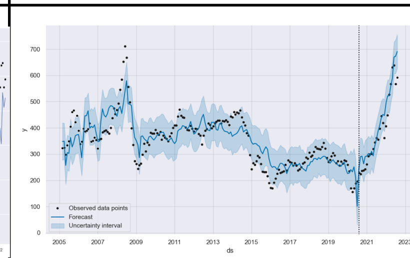
Results: ARMA and Exponential Smoothing were performing worse than our Linear Model, likely because we could not model additional regressors such as futures and macroeconomic variables. However, FbProphet was very accurate in making future forecasts, beating the baseline, linear, and ensemble methods by a good margin. AutoTS had similar results as FbProphet but it was a more complex model which took much longer to train and hence, we selected FbProphet as our final model. The findings and selected features for all our models are summarized in Table 3, 4, 5.

Commodity	Last Spot Price	Baseline	Linear	Ensemble (XGBoost)	Timeseries (FbProphet)		
		MAE($)	MAE($)	MAE($)	MAE($)	%MAE	vs. Baseline
Agriculture	$420.94	43	40	40	36	9%	17%
Energy	$591.26	59	55	53	38	6%	35%
Industrial Metals	$276.04	31	38	34	21	8%	32%
Livestock	$212.78	15	14	12	8	4%	45%
Precious Metals	$503.32	26	20	19	21	4%	20%

Table 3: Results on the Test Set: 1st Sep 2020 - 1st Aug 2022 including % improvement over the baseline

Commodity	Linear	XGBoost	FbProphet
Agriculture	All except Spot	*Spot, Futures*	Spot, Futures, CPI, M1, Real Yields 2Y
Energy	All except Spot	*Spot, Futures, VIX, Industry Production, T-Bill Rate*	Spot, Futures, CPI, Real Yields 2Y, Industry Production, M1
Industrial Metals	All except Spot	*Spot, Futures*	Spot, Futures, US GDP, CPI, Industry Production, M1, Gold spot price
Livestock	All except Spot	*Spot, Futures, VIX, Industry Production, T-Bill Rate*	Spot, Futures, Real Yields 2Y, CPI Change, Industry Production, M1
Precious Metals	All except Spot	*Spot, VIX*	Spot, Futures, US GDP, CPI, Gold spot price, Unemployment Rate

Table 4: Final Features Used for Each Model

Commodity	Linear	XGBoost	FbProphet
Agriculture			
Energy			

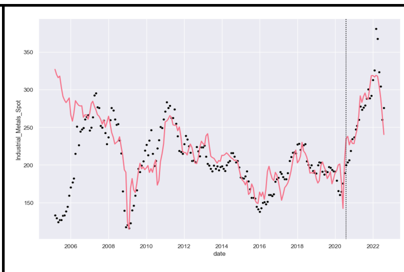
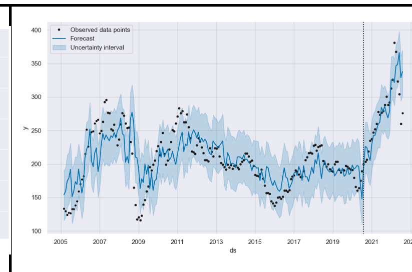
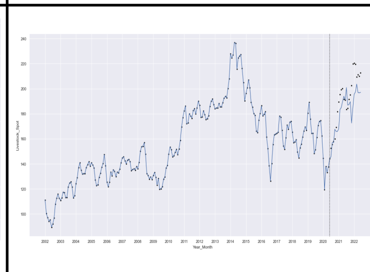
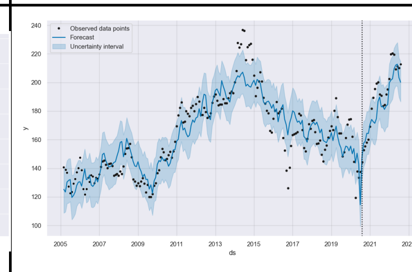
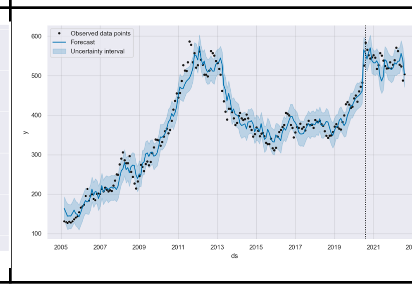
Industrial Metals			
Livestock			
Precious Metals			

*Table 5: **Commodity-wise Predictions for various models**: The dots represent the actual observed values. The lines represent the predicted values. The dotted vertical line separates the training and testing period.*

Saving Dollars: Recommending The Perfect Hedge!

Once we had the commodity spot index predictions for 3 months into the future, we used these predictions to construct personalized portfolios based on different consumptions for each commodity index with the aim to hedge against the predicted rise in prices in the future. To do so, we decided to solve a robust linear optimization problem.

Formulation

Parameters:

- ΔC_i - Forecasted change in i^{th} Commodity
- B – Investment budget
- μ – Risk Tolerance
- σ_i – Price Fluctuations for i^{th} Commodity
- l_i, u_i – Consumption Ranges for i^{th} Commodity

Decision Variable:

- X_i - Investment in the i^{th} Commodity (in units)
- λ_i - Assumed Consumption (in units)

Objective:

$$\min_{X} \max_{\lambda} \sum_{i} \mu(\lambda_i - X_i)^2 * \Delta C_i + (1-\mu) * X_i * \sigma_i$$

Constraints:

- Budget Constraint:

$$\sum_{i} X_i * C_i <= B$$

- Consumption Uncertainty (Consumption varies within a range):

$$l_i \leq \lambda_i \leq u_i, \quad \forall i \subset [n]$$

- Shorting is not permitted:

$$X_i >= 0, \quad \forall i \subset [n]$$

Decision: We want to calculate the amount of units for each commodity we should buy

Objective: Primarily, we want to minimize the exposure to inflation 3 months down the line. However, different investors may have different risk appetite and hence, we may not want to recommend high-risk investments to investors who want low-risk and vice versa. Here we define risk as the amount of fluctuation in the real price of a given commodity over a period of 12 months. Thus, we formulate the problem as a multi-objective optimization problem where we want to minimize the exposure to inflation while also placing importance on having a low risk. This risk tolerance adjusts the two competing objectives.

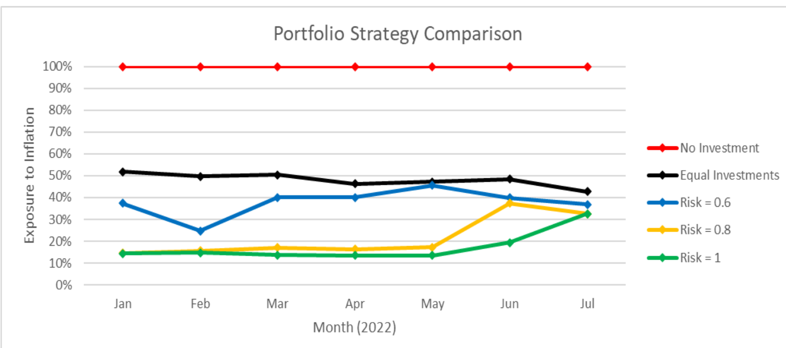
Parameters: We are given the investment budget, the commodity price change forecasts, consumption range for each commodity (in units) and risk tolerance of a given investor

Consumption Constraint: We account for the uncertainty in the consumptions by allowing the consumption to lie in a range and hedge for the worst-case consumption scenario making the solution robust (Note that this leds to a min-max objective which is not straightforward to solve and needs to be solved over a range of consumption values - refer Appendix)
Budget Constraint: We ensure that the investment does not exceed the budget of the individual investor
No Shorting: Finally, we don't permit negative investment (selling) if we predict the commodity prices to go down

Results

Investment Budget = $500					
	Baseline		**Robust Optimization**		
Month	**No Investment**	**Equal Investments**	**Risk = 0.6**	**Risk = 0.8**	**Risk = 1**
Jan	21.75	11.29	8.14	3.15	3.11
Feb	26.8	13.31	6.64	4.2	3.97
Mar	26.06	13.16	10.46	4.43	3.55
Apr	22.49	10.41	9.03	3.66	3.02
May	24.26	11.44	11.08	4.19	3.26
Jun	47.69	23.12	19.05	17.84	9.26
Jul	30.74	13.12	11.33	10.02	10.02

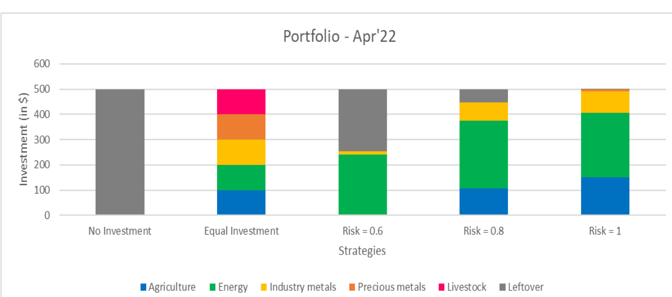
*Table 6: **Dollar Amount Lost to Inflation for Various Investment Strategies***



While testing on the data from Jan'22 to Jul'22, we observe that for our consumption range, one would have lost close to $28 on an average per month as a result of inflation had they been holding cash. With an equal investment of $100 in each commodity (baseline), this exposure reduces to $13.7. Using our approach and a budget of $500, the exposure is down to $5.2 and most of the inflation is hedged against. This is also explained by the graph.

Comments and Inferences:

- No investment baseline comes from our baseline price predictions where our 3-month future prediction is same as current prediction. According to this baseline, the prices remain the same and hence, no investment is the reasonable solution.
- The weight of the inflation in the formulation is termed as risk and a risk factor of 1 would completely ignore the volatility in the commodity prices while a lower risk factor would increasingly weigh the variance.
- Our investment strategy consistently outperforms the baseline of equal investment by a huge margin till June and a risk factor of 1 performs the best, reaching very close (18% exposure) to the benchmark of 0% exposure (inflation neutral).
- The inflation exposure slowly catches up with the baseline however, for July (due to high spikes in real prices). This can be countered if we increase our investment budget.
- A risk factor of 1 will perform the best as long as the predictions are not significantly off from the actual.



Sample portfolio of the dollar amount invested in Apr'22

A lower risk factor however, helps during times of uncertainty and we can see in the graph below that the dollar amount invested reduces as risk tolerance reduces and only a part of the budget gets invested in commodities which have lower volatility. For risk factors below 0.6, we observed that we perform worse than the baseline for this reason.

CONCLUSION AND NEXT STEPS

In this project we have worked out a hedging strategy against commodity inflation that takes into account an individual consumption pattern. After conducting initial research on inflation patterns and the drivers of commodity prices we were able to find leading indicators that predict changes of commodity prices. Among these variables were current spot prices as well as futures prices and other macroeconomic factors, e.g. real yields, industry production and money supply. We then built a time-series model to predict commodity prices about 17-45% better than the baseline model that uses the current commodity price as a predictor for the commodity price 3 months ahead. Based on our best performing model, we constructed a portfolio that is optimized given the consumption of an individual, their budget constraints and their risk appetite.

In a hypothetical world, an investor choosing our solution in the period between January and July 2022 would have saved 82% of inflation losses that would have occurred without investing in our solution. An investor in the baseline model with equal investments would have only saved 52% in the same time period. We conclude that our solution provides a reasonable motivation for Swiss RE to further develop an inflation hedging solution.

There are several assumptions we made in the course of this project that could be subject to further investigation:

The choice of variables: As commodity prices can be perceived as a leading indicator for future overall inflation, one of our challenges was to find leading indicators for commodity prices themselves. We looked at various macroeconomic factors that indicate a commodity price shock before an inflation shock. Tracing back the impact of demand and supply on commodity prices we therefore assessed explicitly the money supply and industry production. We believe that these kinds of variables have a high information gain on future commodity spot prices. However, we suggest to further collect data on supply chains as well as import and export of commodities.

The choice of commodities: We currently used commodity indices instead of actual commodities to build a portfolio. However, it is difficult to model consumption in terms of an index rather than individual commodity. Doing further analysis by breaking down the indices into individual weighted commodities and then using true commodity-wise consumption can be an important next direction.

Baseline model: We defined a naive baseline model using the current commodity price as a predictor for the commodity price 3 months ahead. The choice of the baseline model has a severe impact on our results. The more sophisticated we pick our baseline model to be, the harder it will get to beat this baseline model. Slight variations in the baseline model towards more complexity would be an interesting next step to further push up expectations on our model.

Trading costs: As of now our project does not address trading costs which needs to be taken into account in order to deliver a feasible trading strategy. Trading costs will potentially cause less rebalancing in our portfolio construction as the cost of trading outreaches the reduction of inflation loss. Also, we highlight the question through which instruments our trading strategy is investible which goes hand in hand with the question of how liquid our strategy would be. An investor can get exposure to the commodity market through, e.g., Futures, Forwards as well as more exotic products. One of the next steps would be to replicate our proposed trading strategy with real-life trading instruments and closely monitor the impact of trading costs on the overall portfolio performance.

Data on consumption:

One of the most interesting next steps would be to increase the granularity in terms of consumption data to further individualize our hedging solution. As of now, we optimized a portfolio given generic consumption that is bounded from above and below. We assume that the target group of our solution are potentially made of retail investors that differentiate themselves based on the nature of their work, life habits and overall wealth. It would be preferable to aggregate more data in the form of personas that represent a certain target group of Swiss RE.

Overall this project can be understood as a viability analysis of constructing a hedging portfolio that takes consumption into account. We see that if an investor is willing to invest a certain amount of money into our hedging solution, he can be protected against losses that occur not only through erosion of investment return due to inflation, but also against increasing prices of goods he is consuming.

APPENDIX

The original min-max problem is reformulated as: (using Wald's maximin model)

Parameters:

- ΔC_i - Forecasted change in i^{th} Commodity
- B – Investment budget
- μ – Risk Tolerance
- σ_i – Price Fluctuations for i^{th} Commodity
- l_i, u_i – Consumption Ranges for i^{th} Commodity

Decision Variable:

- X_i - Investment in the i^{th} Commodity (in units)
- λ_i - Assumed Consumption (in units)

Objective:

$$\max_{V,X} \mathrm{V}$$

Constraints:

- Budget Constraint:

$$\sum_i X_i * C_i <= B$$

- Consumption Uncertainty (Consumption varies within a range):

$$V \leq -\sum_i \mu(\lambda_i - X_i)^2 * \Delta C_i - (1-\mu) * X_i * \sigma_i, \qquad \forall \lambda \in \mathrm{U}$$

  Where U denotes the set of possible values of λ

- Shorting is not permitted:

$$X_\mathrm{i} >= 0, \qquad \forall i \subset [n]$$