Context for writing sample: This task was to forecast Chile's inflation outlook (Core ex-volatiles) for the next 3 months and assess if the Banco Central De Chile were likely to turn more dovish or hawkish.

Abstract: VAR Model shows upward pressure on Chile's Core CPI in the next 3 months and if we couple that with a Central Bank that is on pause due to "a more dynamic than expected economy" and "observes higher than average wage growth" means we should see a less dovish CBC going forward.

Model – Approach and Results

Discussion of approach

- My forecasting approach focuses on Chile's Core CPI (excluding volatiles) as it represents the most relevant metric for monetary policy decisions. To enhance trend identification, I utilize seasonally adjusted time series data and apply a 3-month moving average to mitigate model confusion stemming from month-to-month volatility.
- Central banks typically employ three modern forecasting methodologies: Vector Auto Regression (VAR), Machine Learning approaches (including Gradient Boost and Random Forest algorithms), and Dynamic General Stochastic Equilibrium (DGSE) models. Given the constraints of publicly available data, I selected the VAR methodology as my primary approach.
- I acknowledge that forecast error in VAR models tends to amplify over longer timeframes and while my model projects forecasts for a six-month period to provide comprehensive insights, I place greater confidence in the initial three data points.
- For extended forecasting periods, a DGSE model would represent best practice as it incorporates economic theory; however, I currently lack the necessary parameters to implement this approach. Additionally, I have completed approximately 50% of the Machine Learning approach, designating the remainder as an extension project after reaching a stage requiring feeder models.

Variables/Features used

- From Chile CB website I retrieved the following monthly data points from Jan 2015-present. Besides Core-CPI, everything is a X variable.

Variable Name	Variable Name	Variable Name
Core CPI (ex Volatiles)	Overall PPI	Manufacturing PPI
Mining PPI	Utilities PPI	Agri PPI
IMACEC Total	IMACEC Production of Goods	IMACEC Commerce
IMACEC Services	IMACEC Factor Cost	USD/CLP
11 month expected MPR	Ex-Ante Real Rate	Real Wages
Unit Labor Costs	Nominal Wages	Unemployment Rate
M1	M2	REER
Headline CPI	IPE (External Price)	Oil Prices

- I differenced the data (month-over-month) to produce approximately stationary time series. When series remain non-stationary, I applied further differencing which is essential for VAR model accuracy.
- Due to Chile's economic release calendar, I receive non-CPI macroeconomic data with a one-month lag. For example, when forecasting February 2025 Core CPI, I only utilize data up to January 2025 for explanatory variables to ensure no look-ahead bias occurs.
- I expanded my original 24 variables by incorporating up to six lags for each variable, calculating moving averages (two and three months), and including lagged moving averages. This methodology yields a comprehensive dataset of 122×647 dimensions for forecasting Chilean inflation.
- Ideally, I would have included copper prices in my analysis; however, I was unable to locate a suitable publicly available data source with sufficient historical price information.

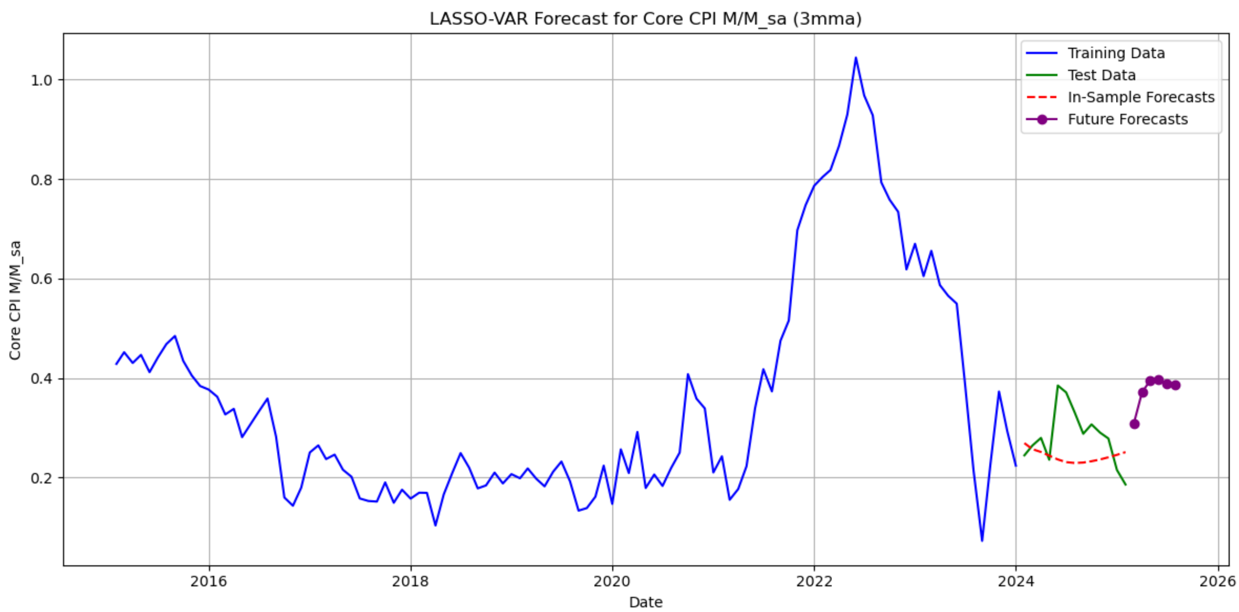
LASSO-VAR Approach

- A Vector Autoregression (VAR) model cannot effectively process 647 features due to its assumption of endogeneity across all variables.
- I implemented LASSO regularization as a feature selection technique to identify the most predictive 15 features from the original 647 variables.
- This approach yields a more parsimonious model that retains only the most statistically significant predictors.
- The VAR methodology effectively captures the complex feedback relationships between inflation and other macroeconomic indicators, including unemployment rates, ex-ante interest rates, and wage growth dynamics.
- By incorporating multiple interrelated variables with their appropriate lag structures, VAR models typically generate more accurate inflation forecasts compared to univariate modeling approaches.
- For rigorous validation, I restricted the model training to data through December 2023, reserving subsequent periods as an in-sample testing horizon to quantitatively assess predictive accuracy.

Discussion of Results

- The VAR model indicates increased inflationary pressures over the three-month forecast horizon (Figure 1).
- My assessment suggests that the optimal interpretation of these results is not to focus on the precise 31 basis point forecast for next month's Core CPI SA 3mma, but rather to recognize the high probability of an emerging upward trend from the current baseline.
- The model demonstrates reasonable performance in in-sample testing given data constraints. While month-to-month precision may show some variance, the model effectively captures the overall downward trend throughout the test period.
- A significant portion of forecast deviation in the test sample can be attributed to an anomalous higher core inflation print in mid-2024, which resulted in actual values following the same directionally downward trend but at consistently elevated levels compared to model projections.
- This single elevated reading subsequently influenced the trajectory of actual Core CPI readings relative to model forecasts due to the autoregressive component, whereby the three-month moving average of previous inflation prints serves as the strongest predictor of subsequent monthly Core CPI values.

Figure 1: Price pressures in Chile are rising. Note gaps in the lines are purely for visualization.



- A critical validation measure for model reliability is error decomposition analysis (Figure 2). The VAR model generates forecasts by simultaneously projecting both explanatory variables (Xs) and the target variable.
- By analyzing the test sample with actual explanatory variables, I can disaggregate forecast error into two components: (1) X error (inaccuracies in predicting future explanatory variables) and (2) inherent model error.
- This decomposition reveals that recent forecast deviations are predominantly attributable to X error rather than structural model deficiencies.

- This finding provides significant reassurance regarding the fundamental soundness of the model architecture, indicating that forecast accuracy would improve substantially with enhanced input data quality.
- I address specific model enhancement strategies in subsequent sections of this analysis.

Figure 2: Error mostly comes from X variables rather than model

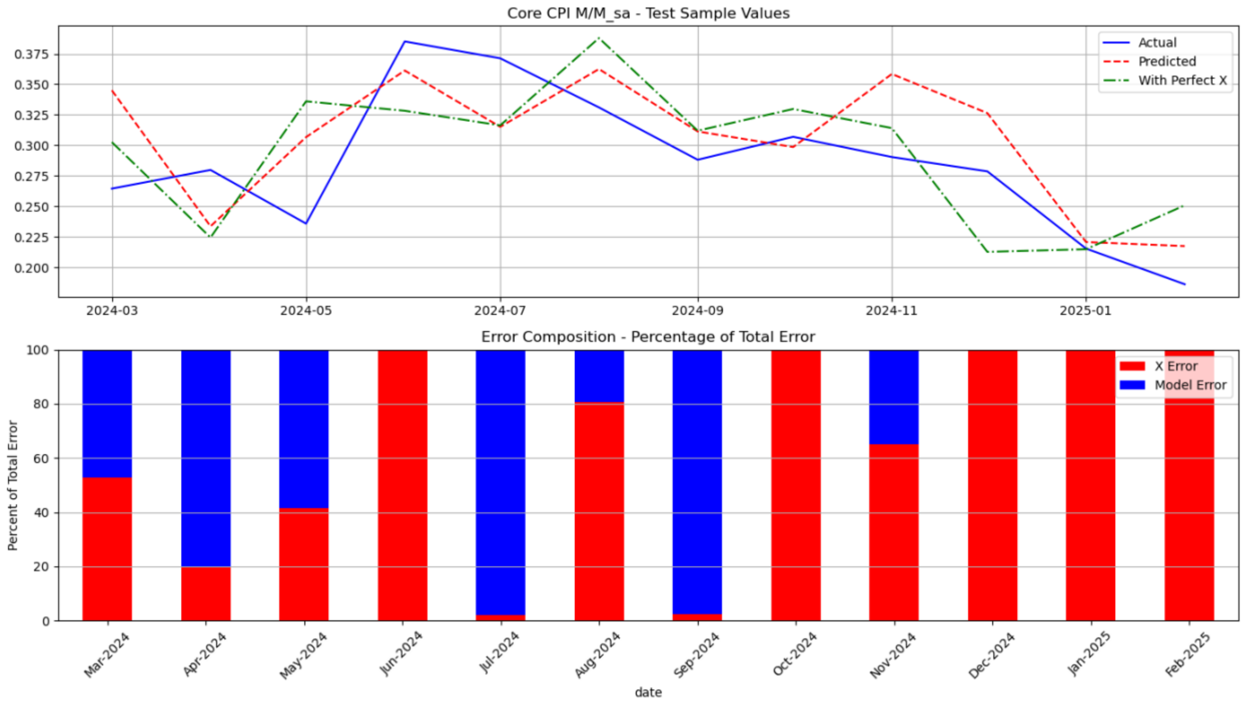
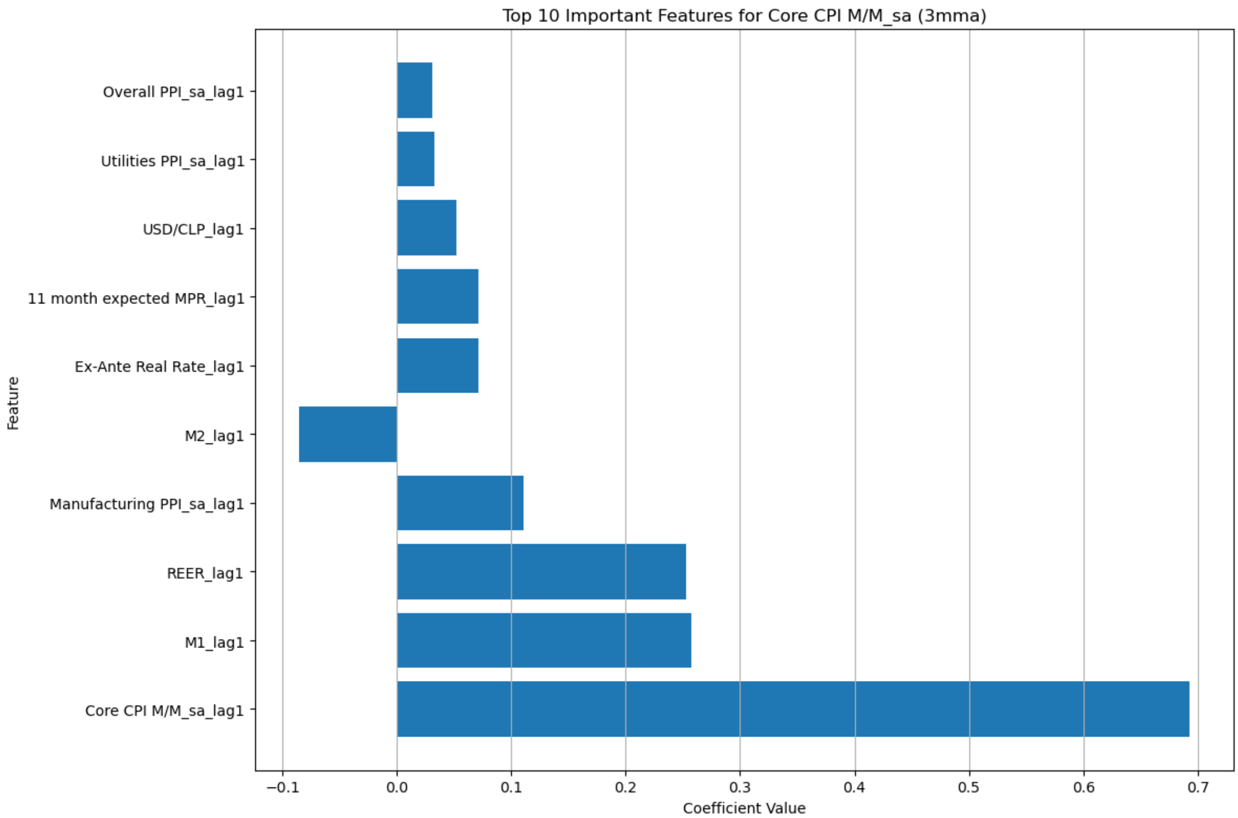


Figure 3: Top features are mostly economically intuitive



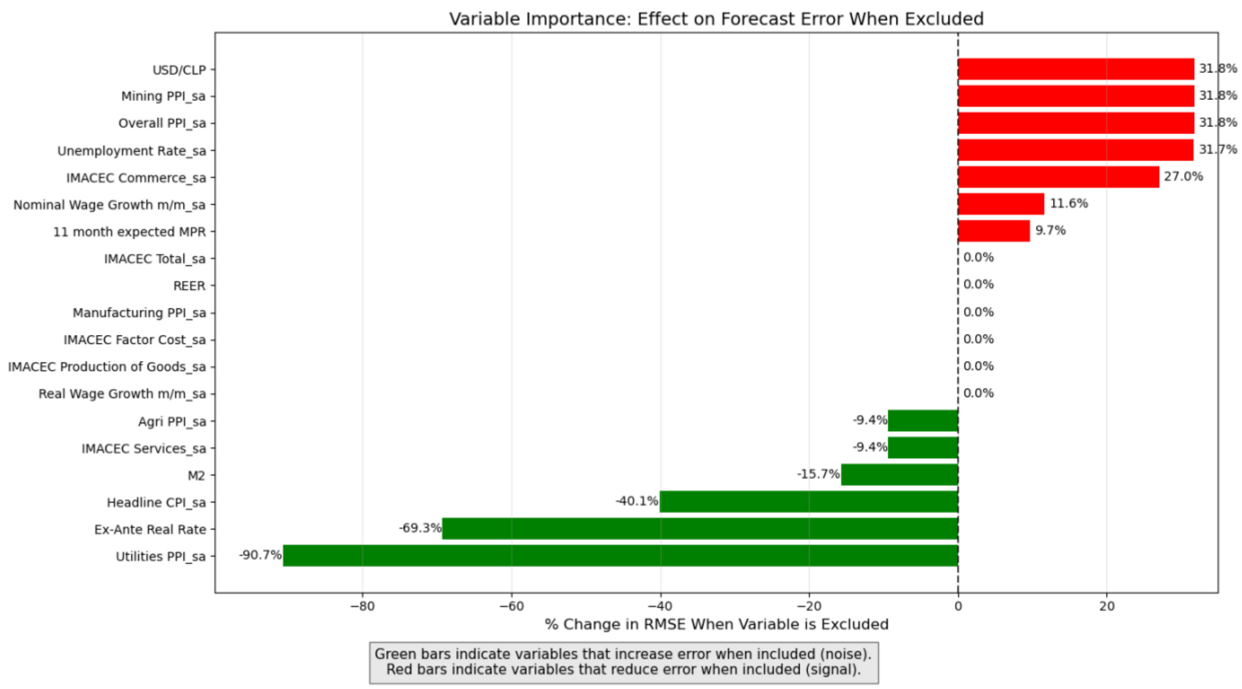
Room for improvement in modelling

- Several predictor variables exhibit economically counterintuitive coefficient signs, which raises methodological concerns.
    - For eg., the Real Effective Exchange Rate (REER) theoretically should display a negative coefficient as currency strengthening typically reduces imported inflation.
    - Further investigation into the Central Bank of Chile's REER methodology is warranted, with consideration given to potentially excluding this variable if it generates spurious relationships.
- Implementation of high-frequency data would significantly enhance month-to-month forecast precision.
    - I have identified several robust databases that could provide this granularity, though they remain behind paywalls (accessible through platforms such as Macrobond or CEIC).
- Development of an ensemble forecasting methodology would be preferable to exclusive reliance on VAR modeling, as multiple methodological approaches would provide confirmatory or contradictory signals that strengthen overall forecast reliability.
- Future improvements should include construction of dedicated models for more accurately predicting explanatory variables, or alternatively, incorporation of Central Bank forecasts or consensus economist projections to minimize X error in the inflation model.
- I also evaluated a Random Forest approach, which yielded promising test results (provided as proof of concept later on). However, unlike VAR modeling, machine learning approaches require explicitly forecasted explanatory variables to generate predictions for March 2025 onward, which I have designated as a future extension of this research.

Sensitivity Analysis

The 2 main ways to think about sensitivity in a VAR model are (1) Variable Inclusion/Exclusion tests and (2) Hyperparameter tuning.

Figure 4: Promising results in variable importance tests



- The model validation procedure involves systematically excluding individual variables from the VAR specification, re-estimating the complete model, and quantitatively comparing forecast accuracy against the original specification.
- Utilities PPI and Headline CPI, while appearing among the more problematic variables in statistical terms, do not present significant concerns as their coefficient magnitudes are negligible within the model structure.
- Regarding the Ex-Ante real interest rate, despite statistical evidence suggesting its inclusion worsens model performance, further empirical investigation led me to conclude that its retention is warranted based on economic theory. The forecast error improvements resulting from its exclusion appear to be economically counterintuitive, prompting me to exercise some judgment in maintaining this variable in the specification.

- I have implemented comprehensive hyperparameter optimization procedures (systematically evaluating training sample size, lag structure configurations, and LASSO regularization intensity) with the results presented throughout this document reflecting only the optimal model configuration from this tuning process.

Chile Monetary Policy Trajectory

- The current monetary policy stance exhibits only mild restrictiveness, with the real ex-ante policy rate at 1.2% (calculated from the 5.0% nominal rate and 3.8% one-year ahead inflation expectations per the latest economist survey) marginally exceeding the Central Bank's 1.1% neutral rate estimate.
- The absence of excessive monetary restraint suggests no immediate imperative for policy rate reduction, particularly given the favourable economic growth environment.
- The recently published Monetary Policy Report contains significant upward revisions to key macroeconomic indicators:
    - Annual GDP growth forecast elevated to 1.75%-2.75%, surpassing the previous projection of 1.5%-2.5%
    - Upward adjustments to domestic demand, private consumption, and export forecasts
    - End-2025 annual inflation expectations increased from 3.6% to 3.8%
- The Economic Expectations Survey (EES) conducted by the Central Bank indicates market consensus anticipates policy rate stability over the next five-month horizon, with expectations of modest monetary accommodation totalling 50 basis points within the coming year.

Machine Learning Extension

Figure 5: ML Approach performs well in the test set.