Multifactors Risk Research of China Stock Listed Banks

Chuxuan Ma chm524@lehigh.edu
Lehigh University P.C. ROSSIN COLLEGE

Abstract

Research Goal
This research investigates the causal relationships between various financial factors and risk premiums in China's banking sector. Using the Weighted Average Cost of Capital (WACC) as a foundation, we decompose bank risk premiums into common (market-driven) and idiosyncratic (bank-specific) components and analyze how these are influenced by different risk types—default risk, liquidity risk, and market risk.

Methodology
Our methodology employs a Double Machine Learning framework to establish causal inference while controlling for confounding effects. We leverage a comprehensive dataset of 421 financial indicators across 42 Chinese listed banks from 2019 to 2022. The modeling approach combines advanced techniques including K-Shape clustering for risk classification, AdaBoost regression with recursive feature elimination, and Causal Forest models to capture treatment effect heterogeneity.

Objective
Through careful estimation of risk components and their interactions, we aim to provide insights into the fundamental risk structures of China's banking industry. This research contributes to the financial literature by offering a causal perspective on risk pricing mechanisms and providing practical implications for bank capital management, regulatory oversight, and investment strategies in Chinese financial markets.

Introduction

The Banking Sector in China
The banking sector forms a critical foundation of China's financial system. As China continues its market reforms and financial liberalization, understanding the risk factors affecting banking institutions has become increasingly important for regulators, investors, and financial stability.

Research Context
China's banking landscape presents a unique research context due to several distinctive characteristics:
1. Mixed ownership structures combining state and private capital
2. Ongoing financial reforms
3. Varying degrees of government influence
4. An evolving regulatory environment
These factors make Chinese banks fundamentally different from their Western counterparts and necessitate specialized research to understand their risk dynamics.

Research Focus
Our research focuses on the Weighted Average Cost of Capital (WACC) and particularly its risk premium component, which represents the additional return required by investors to compensate for uncertainty when investing in bank securities. By identifying the causal effects of various financial factors on this risk premium, we aim to provide insights into how different types of risks—default risk, liquidity risk, and market risk—impact the capital costs of Chinese banks. Risk premium is a common measure financial risk for a given interest rate, calculating how far the interest rate exceeds the risk-free interest rate and undergo the corresponding risk. We will refer the "risk" to "risk premium" in the following parts.

Materials

Our study utilizes a comprehensive dataset of China's banking sector, encompassing 42 listed banks tracked over 978 trading days from January 2019 to December 2022 using JoinQuant API. This period is particularly valuable as it captures both pre-pandemic market conditions and subsequent disruptions, providing a rich contextual environment for risk analysis.

Data Structure and Sources
The dataset is organized into several complementary components that together provide a holistic view of banking risk factors:
1. **Financial Indicators (421 factors) categorized into:**
   1. Technical market factors (price movements, volatility measures, trading patterns)
   2. Fundamental indicators (derived from bank financial statements)
   3. Macroeconomic environment indicators
   4. Securities margins trading data
   5. Money flow statistics
   6. Industry classification metrics
   7. Market index data
2. **Bank-specific Information:**
   1. Historical listing details (IPO dates and delisting dates where applicable)
   2. Enterprise value weights for sectoral aggregation
   3. Market capitalization data
   4. Return rates over various periods
3. **Financial Modeling Parameters:**
   1. Weighted Average Cost of Capital (WACC) components
   2. Risk-free interest rates (SHIBOR)
   3. Market returns
   4. Free Cash Flow (FCF) discounted model inputs

This rich dataset allows us to examine both system-wide patterns and bank-specific risk characteristics, providing a foundation for understanding how various financial factors contribute to the risk premiums in China's banking sector.
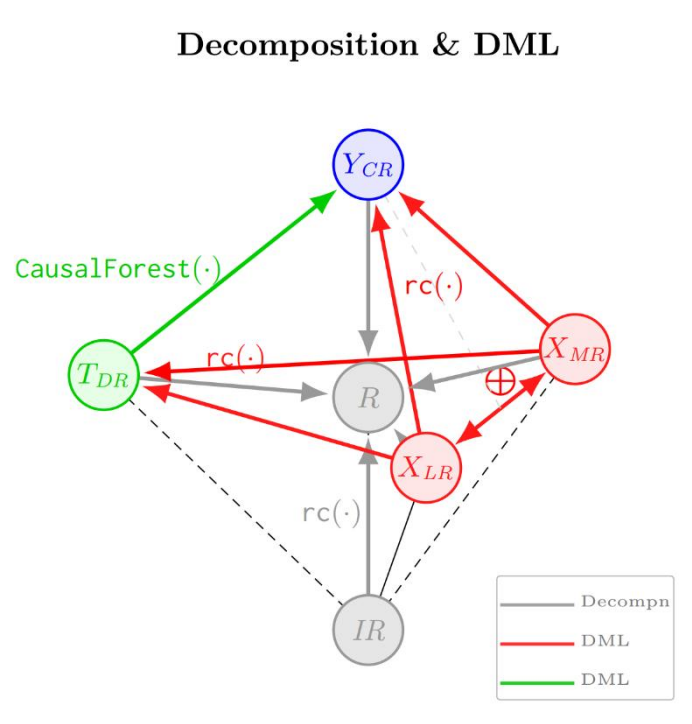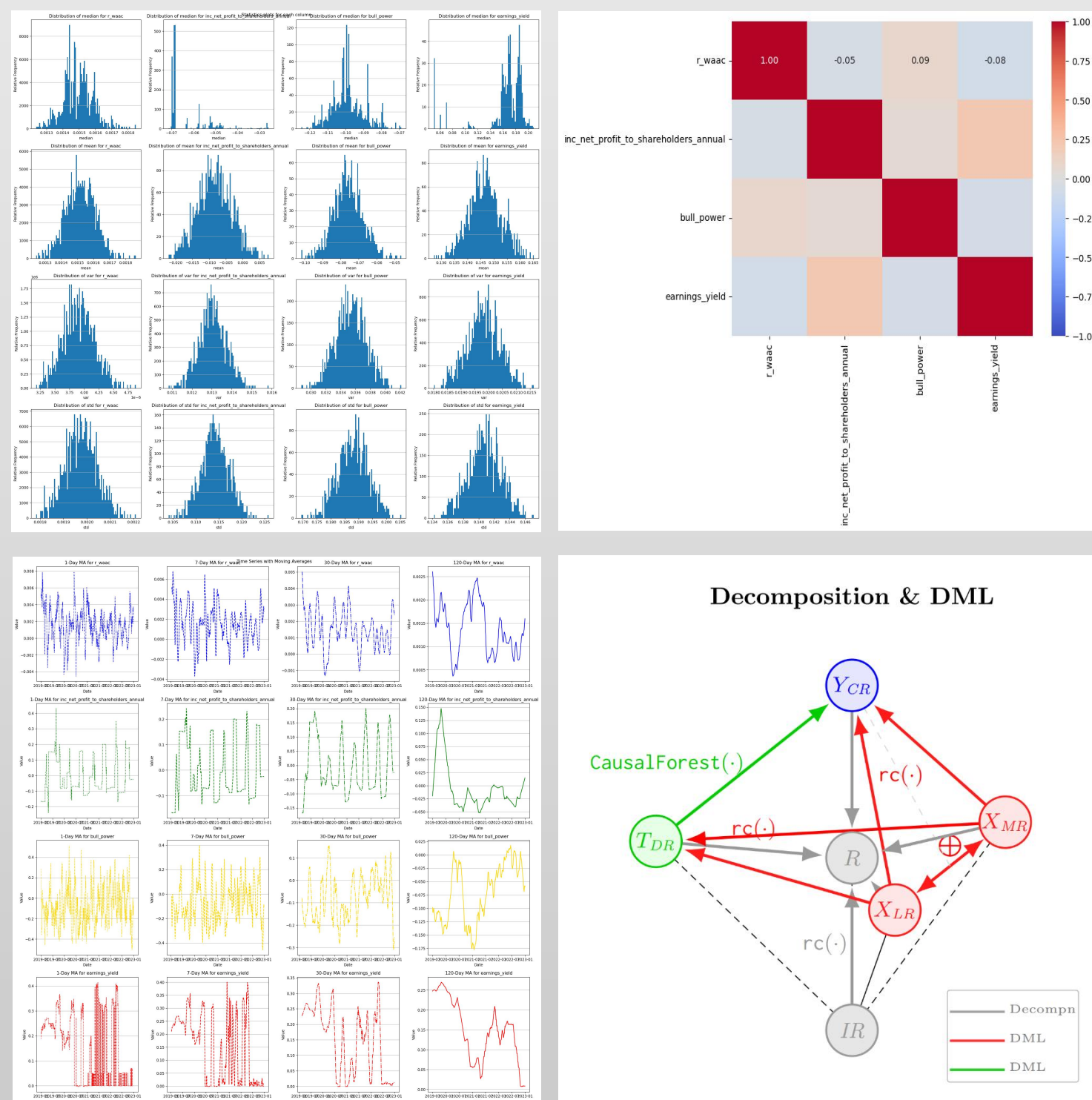
Methodology

1. **Risk Classification System**
   We employ a dual-approach to risk classification:
   1. Initial Domain Knowledge Classification: Financial indicators are initially categorized into three fundamental risk types based on banking finance theory: Default Risk, Liquidity Risk, and Market Risk.
   2. Re-clustering by K-Shape: After the cluster or classification initialized, we use a time-series clustering technique: K-Shape, to classify the risk based on real industry data, to update our knowledge and belief.

2. **Risk Premium Decomposition**
   We decompose the **outcome** risk premium (WACC minus risk-free rate) into distinct components, which is the prediction of a specific risk type of factors to market returns risk premium:
   - **Common Risk**: Prediction based on WAAC
   - **Idiosyncratic Risk:** Residuals of the above prediction
   We also decompose different premium of **treatment** risks (**Default, Liquidity, Market**) based on their risk factors, and we estimate their contribution to both the common and idiosyncratic risk components.

3. **Machine Learning Architecture**
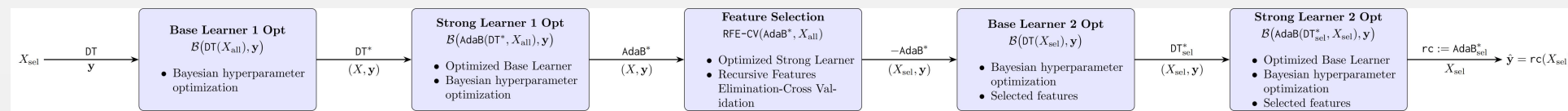
4. **Causal Inference Framework**
   We implement Double Machine Learning to address endogeneity concerns when estimating the causal effect of treatment risk on outcome risk:
   1. First estimate the effect of confounding variables (other risk types)
   2. Residualize both treatment and outcome risk
   3. Cross-estimate the conditional average treatment effect (CATE) using causal forest models (CausalForestDML) across folds
   This allows us to measure heterogeneous treatment effects across time periods, revealing how the impact of different risk types varies throughout our sampling period.

Results

- By K-Shape, we update the prior risk cluster, as the first figure shown in several risk example.
- Using the CausalForestDML, we are able to draw several significant figures for each treatment risk to outcome risk (we choose market risk to idiosyncrasy risk here for illustration), displaying significant facts about the causal effect of the former risk to the latter, in different perspectives.
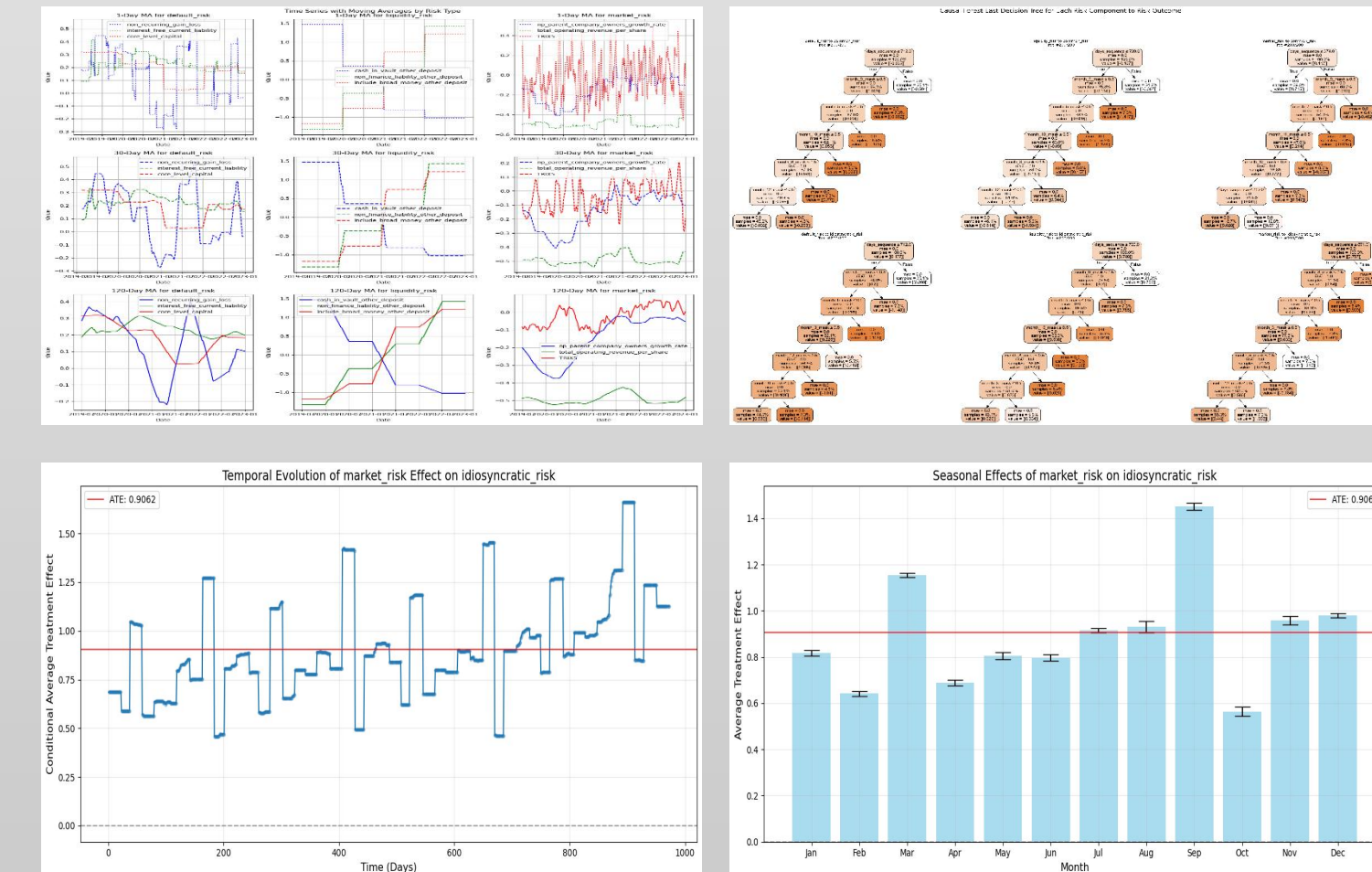
Conclusion

CATE Patterns
Directional Effects:
Most risk relationships show positive CATEs, with market risk → Common risk exhibiting the strongest and most consistent positive effects, followed by default risk, while liquidity risk shows more variable effects.
Magnitude Distribution: Effects typically range from -0.1 to +0.3, clustered mostly between 0.05-0.15, indicating modest but economically significant causal relationships.
Stability Differences: Market risk → Common risk shows remarkable stability across conditions, while Liquidity risk → Idiosyncratic risk displays high variability with frequent sign changes.
Temporal Heterogeneity
Effect Evolution:
Most relationships show stronger effects in earlier periods with gradual decay over time, suggesting evolving risk transmission mechanisms in the banking sector.
Structural Breaks: Clear temporal breakpoints exist where causal mechanisms change substantially, particularly visible in the liquidity risk relationships.
Key Anomalies:
Market risk uniquely maintains consistent effect strength throughout the sample period
Liquidity risk → idiosyncratic risk shows dramatic temporal sign inversions
Default risk → idiosyncratic risk displays complex multi-factor temporal dependencies
Seasonal Heterogeneity
Reporting Cycles: Quarter-end months (March, June, September, December) frequently appear as important decision nodes, suggesting financial reporting period effects.
Annual Patterns: January-February and July-August often serve as distinctive seasonal breakpoints for different risk types.
Notable Anomalies:
Market risk shows unexpectedly strong effects during typically low-volatility summer months
Default risk demonstrates amplified December-January effects, suggesting sensitivity to year-end assessments
Liquidity risk exhibits a unique pattern of negative mid-year effects but positive year-end effects
Integrated Temporal-Seasonal Effects
Increasing Seasonality: Risk relationships show stronger seasonal dependency in later time periods, particularly for default and liquidity risks.
Evolving Synchronization: Banking sector risk transmission has become increasingly synchronized with regulatory and reporting cycles over time.
These patterns point to a complex, time-dependent risk structure in China's banking sector where risk management must account for not just risk types but also their temporal and seasonal context.

Acknowledgements