Group 5

Rare Earth Elements Scarcity under Trade Restrictions: An Econometric and Machine Learning Approach

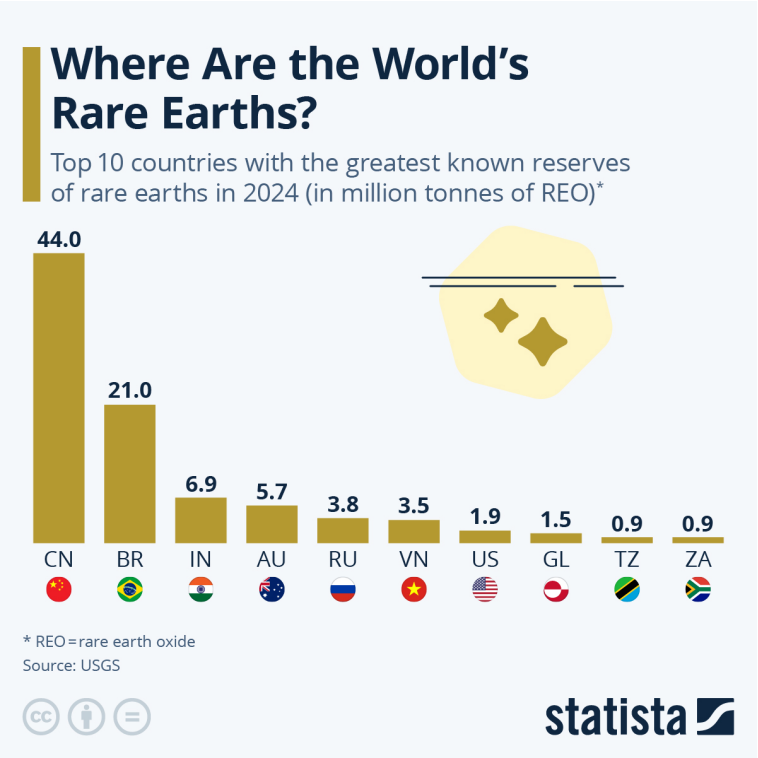
Presented by

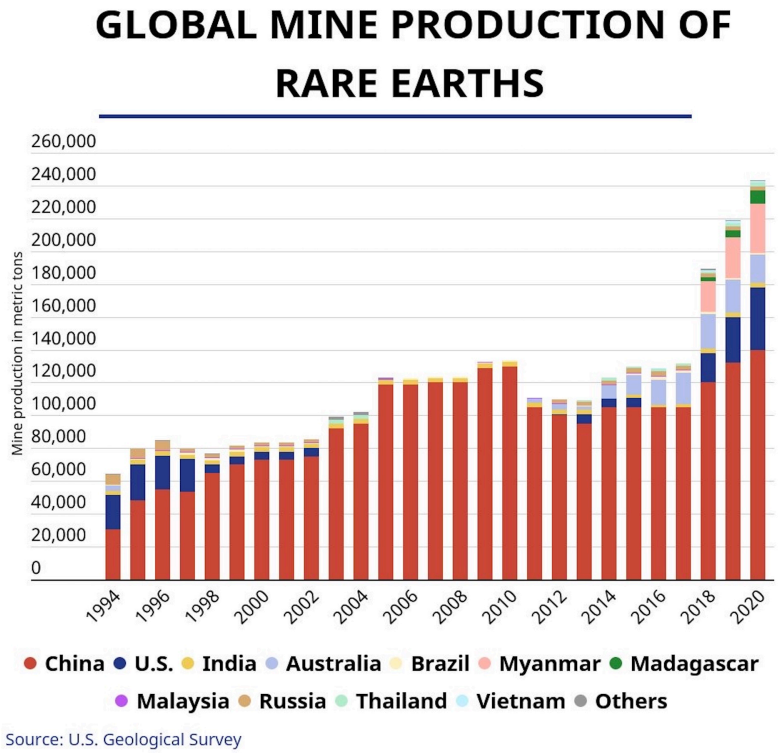
- Mr. Rahul Kumar, Research Scholar, BITS-Pilani, Hyderabad
- Dr. K. Suganya, Assistant Professor of Economics, Lady Doak College
- Dr. S. Senthamarai, Associate Professor of Economics, Holy Cross College
- Ms. Rinu Jayaprakash, Assistant Professor, TKM Institute of Management

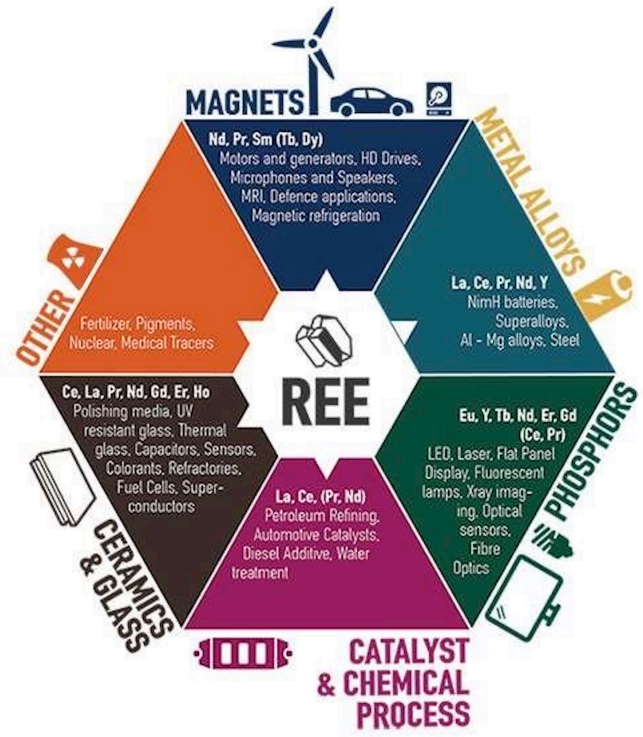
Introduction

Rare Earth Elements (REEs) are a group of 17 chemically similar metals essential for various advanced technologies, including electric vehicles, wind turbines, smartphones, and military equipment. Despite their name, these elements are relatively abundant in the Earth's crust but are rarely found in concentrated forms, making their extraction and processing economically challenging.

China maintains a dominant position in the global supply of several critical rare earth elements, particularly: **Dysprosium (Dy), Terbium (Tb), Neodymium (Nd) and Praseodymium (Pr), Samarium (Sm), Gadolinium (Gd), Lutetium (Lu), Scandium (Sc), and Yttrium (Y)**







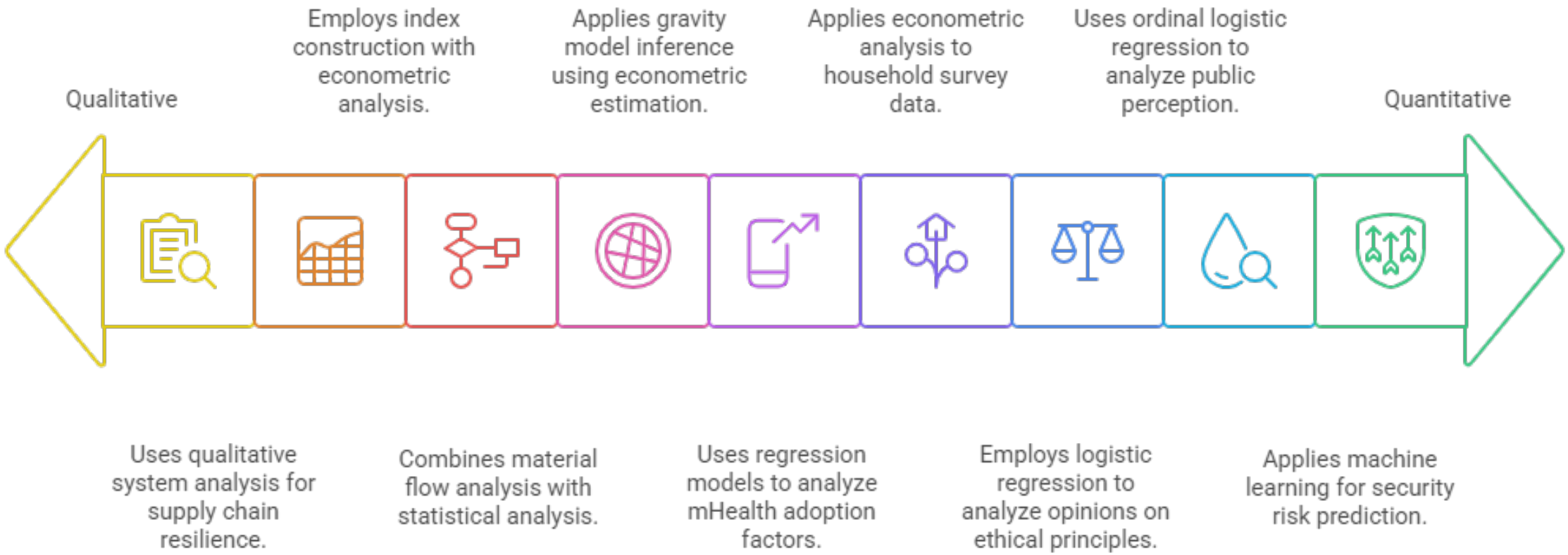
metallic state

oxidized state

Literature review

Paper (Citation)	Primary Focus	Scarcity/Criticality Concept Addressed	Core Methodology	Key Quantitative Techniques	Data Used
Liu, Y., Zhao, G., & Zhao, Y. (2022).	Public perceptions of water scarcity	Water scarcity perception	Survey / Analysis of public perception	Ordinal logistic regression model, Structural Equation Modeling (SEM)	Survey data on perceptions and individual characteristics (e.g., demographics, potentially tap water
Strech, D., Johann, S., & Marckmann, G. (2016).	Ethical frameworks for prioritizing patients	Scarcity of medical resources (e.g., donor organs, hospital beds during epidemic)	Survey / Analysis of opinions on ethical principles (fairness evaluations)	Logistic regression (to compare opinions between groups, yielding odds ratios)	Survey data on opinions regarding ethical principles for resource allocation (n=1,267).
Abate, G. T. (2016).	Impact of grazing/water scarcity on household welfare/food security	Grazing and water resource scarcity, impact on welfare	Econometric analysis of household survey data	Resource scarcity indicators (e.g., distance, potentially shadow prices); Regression analysis	Household survey data (similar studies use n≈500), food consumption expenditure (PCFE), resource collection time.
Alam, M. Z., Hu, W., & Barua, Z. (2020).	Adoption of mHealth services	Resource scarcity and accessibility (as external/ environmental factors); Medical resource scarcity	Analysis of factors influencing technology adoption	Regression models (e.g., SEM, multiple regression) testing direct & moderation effects	Data on mHealth adoption, user perceptions, user characteristics, external environment factors.
Bilen, A., & Özer, A. B. (2021).	Security risk / Crime prediction (focus on cyber-attacks in this example)	Risk assessment/prediction (motivated by efficient resource allocation)	Machine Learning application in risk context (Literature Review Context)	Logistic Regression, Support Vector Machines (SVM), Decision Trees (among others)	Varies; e.g., historical crime/cyber incident data, demographic data, social media data.
Anderson & Van Wincoop (2004)	Trade Costs	Economic Accessibility	Gravity Model Inference	Econometric Estimation (Theory-consistent, e.g., PPML implied), Inference of Latent Variables	Bilateral Trade Flows, GDP, Distance, Policy Proxies (e.g., tariffs, border effects)
Ciacci et al. (2022) & related MFA studies	Critical Metal Circularity	Physical Availability, Circularity Potential	Dynamic Material Flow Analysis (MFA) & CE Assessment	Simulation Modeling, Probabilistic Analysis (Uncertainty), Statistical Analysis of Flows/Stocks	Production/Consumption Data, Trade Data (e.g., BACI), Material Content Data, EoL Parameters (lifetimes, recycling rates)
Mancheri et al. (2019)	REE Supply Chain Resilience	Supply Chain Robustness, Policy Impact	Resilience Framework Analysis	Qualitative/Semi-Quantitative Systems Analysis, Conceptual Modeling, Scenario Analysis	Policy Documents, Trade Data, Prices, Expert Judgment, Case Study Comparison (Tantalum)
Batabyal & Nijkamp (2019) (Inferred)	REE Criticality in Asian Economies	Regional Vulnerability, Supply Risk, Econ. Imp.	Criticality Index Calculation & Regional Analysis	Index Construction, Econometric Analysis (Trade/Price Models?), Input-Output Modeling?	Trade Data, Production Data, Economic Indicators, Governance Indicators, Policy Information

This spectrum categorizes research papers based on their methodological approach, ranging from qualitative, conceptual analyses to quantitative, data-driven modeling.



Objectives of the study

- Estimate the probability of REE scarcity resulting from high tariffs and related trade restrictions.

Data Type	Source
Stock Market Performance	Compustat, Yahoo Finance, EDGAR
Trade Data (Imports/ Exports)	World Bank WDI:
Macroeconomic Indicators	FRED (GDP series by country)
REE Price and Scarcity Proxies	World Bank Pink Sheet, USGS
Policy Simulation Variables	Custom simulated panel data on trade restrictions and responses

1. Dependent Variables

These represent the outcomes the model seeks to explain or predict.

- **Scarcity Binary:** A simplified dichotomous indicator where 1 = High scarcity, 0 = otherwise.
- **Scarcity Level (Ordered):** Categorized into *Low*, *Moderate*, and *High* levels of REE scarcity.
- **Policy Response (Unordered):** Categorical variable representing government response types: *No Action*, *Stockpiling*, *Trade Agreements*, *Domestic Investment*.

2. Explanatory Variables

A. Trade Restriction Indicators (Main Independent Variables)

- **Trade Restriction Index (0–10):** Composite measure capturing the intensity of trade barriers.
- **Tariff Rate (%):** Import/export tariff rate on REEs.
- **Export Quota (Binary):** 1 = Quotas imposed, 0 = No quotas.
- **Import License (Binary):** 1 = Licensing required, 0 = No licensing restrictions.

B. Control Variables

- **Domestic Production (log):** Log of national REE production volume.
- **Consumption-to-Production Ratio:** Indicates dependency on external supply.
- **Stockpile Months:** Number of months REEs can be supplied via reserves.
- **Price Volatility:** Standard deviation or % change in REE price index over time.

✅ 1. Binary Probit Model

Dependent Variable:

•$Scarcity\_Binary_{it}$ = 1 if high REE scarcity, 0 otherwise.

Equation:

$$P(Scarcity\_Binary_{it} = 1) = \Phi(\beta_0 + \beta_1 \cdot TRI_{it} + \beta_2 \cdot Tariff_{it} + \beta_3 \cdot Quota_{it} + \beta_4 \cdot License_{it} + \gamma \cdot Z_{it})$$

Where:

- ❑ Φ is the cumulative distribution function (CDF) of the standard normal distribution
- ❑ Z_{it} is the vector of control variables:
  - Domestic Production (log)
  - Consumption-to-Production Ratio
  - Stockpile Months
  - Price Volatility

✅ 2. Binary Logit Model

Equation:

$$P(Scarcity\_Binary_{it} = 1) = 1 / (1 + e^{(-(\beta_0 + \beta_1 \cdot TRI_{it} + \beta_2 \cdot Tariff_{it} + \beta_3 \cdot Quota_{it} + \beta_4 \cdot License_{it} + \gamma \cdot Z_{it}))})$$

Where:

- ❑ e is the base of the natural logarithm
- ❑ Z_{it} is the vector of control variables:
  - Domestic Production (log)
  - Consumption-to-Production Ratio
  - Stockpile Months
  - Price Volatility

✅ 3. Ordered Probit Model

Dependent Variable: $\text{Scarcity\_Level}_{it} \in \{\text{Low, Moderate, High}\}$ (Ordinal outcome)

We define a continuous variable S_{it}^* such that:

$$S_{it}^* = \beta_0 + \beta_1 \cdot \text{TRI}_{it} + \beta_2 \cdot \text{Tariff}_{it} + \beta_3 \cdot \text{Quota}_{it} + \beta_4 \cdot \text{License}_{it} + \gamma \cdot Z_{it} + \varepsilon_{it}$$

Where:

S_{it}^* is unobserved

$\varepsilon_{it} \sim N(0,1)$ (standard normal distribution)

Z_{it} = control variables as before

Observed categories:

$$\text{Scarcity\_Level}_{it} = \begin{cases} \text{Low} & \text{if } S_{it}^* \leq \tau_1 \\ \text{Moderate} & \text{if } \tau_1 < S_{it}^* \leq \tau_2 \\ \text{High} & \text{if } S_{it}^* > \tau_2 \end{cases}$$

Where τ_1 and τ_2 are cutoff thresholds estimated from the data

✅ 4. Ordered Logit Model

The structure is similar, but now ε_{it} follows a logistic distribution:

$$\mathbf{S_{it}^* = \beta_0 + \beta_1 \cdot TRI_{it} + \beta_2 \cdot Tariff_{it} + \beta_3 \cdot Quota_{it} + \beta_4 \cdot License_{it} + \gamma \cdot Z_{it} + \varepsilon_{it}}$$

Category probabilities:

$$P(Scarcity\_Level_{it} = Low) = \Lambda(\tau_1 - X_{it} \cdot \beta)$$
$$P(Scarcity\_Level_{it} = Moderate) = \Lambda(\tau_2 - X_{it} \cdot \beta) - \Lambda(\tau_1 - X_{it} \cdot \beta)$$
$$P(Scarcity\_Level_{it} = High) = 1 - \Lambda(\tau_2 - X_{it} \cdot \beta)$$

Where:

$\Lambda(z) = 1 / (1 + e^{(-z)})$ is the logistic CDF

$X_{it} \cdot \beta$ is the linear predictor

✅ Multinomial Logit Model

Let $\text{Policy\_Response}_{it} \in \{0, 1, 2, 3\}$ (unordered categories)

The probability that government i at time t chooses policy response j is:

$$P(\text{Policy\_Response}_{it} = j) = \exp(X_{it} \cdot \beta_j) / \Sigma_k \exp(X_{it} \cdot \beta_k), \quad \text{for } j \in \{1, 2, 3\}$$

For the baseline category (e.g., j = 0 = No Action):

$$P(\text{Policy\_Response}_{it} = 0) = 1 / \Sigma_k \exp(X_{it} \cdot \beta_k), \quad \text{where } k \in \{0, 1, 2, 3\} \text{ and } \beta_0 = 0$$

Variables:

- X_{it}: Vector of independent variables (e.g., Trade Restriction Index, Tariff Rate, Quota, License, controls)
- β_j: Coefficient vector for category j
- $\exp(\cdot)$: Exponential function
- Σ_k: Sum across all categories

Table 1: Characteristics of the Sample

Characteristic	Value	Notes
Sample Size	256	Cross-sectional observations
Time Period	2015-2024	Quarterly data across 10 years
Countries/Regions	16	Major REE producers and consumers
Variables	12	8 predictors, 4 control variables
Missing Values	2.3%	Primarily in early periods
Outliers Removed	5	Extreme tariff/quota values

- **Sample Size**: 256 cross-sectional observations from 16 countries/regions, covering data from 2015 to 2024
- **Variables**: 12 variables, including 8 predictors and 4 control variables
- **Missing Values**: 2.3%, mainly in earlier periods
- **Outliers**: 5 extreme tariff/quota values removed

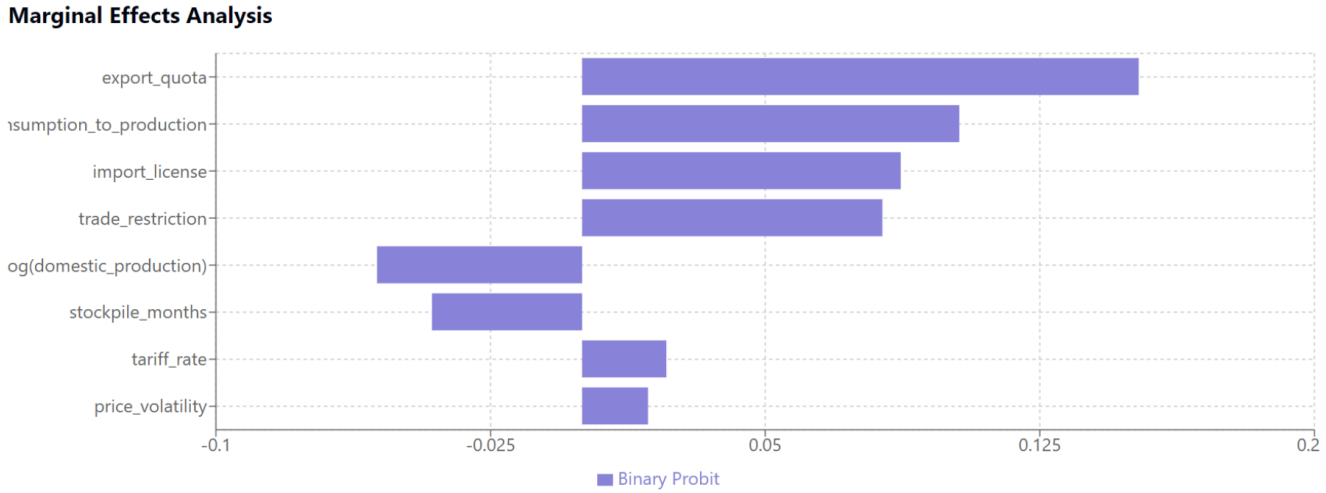
Table 2: Model Fit Statistics

Measure	Binary Probit	Binary Logit	Ordered Probit	Ordered Logit	Multinomial Logit
Log-likelihood	-351.73	-352.66	-618.36	-620.97	-927.62
AIC	723.45	725.32	1246.71	1251.94	1875.23
BIC	762.18	764.05	1295.38	1300.61	1972.56
McFadden's R^2	0.311	0.308	0.283	0.279	0.246
Count R^2	0.765	0.758	0.723	0.712	0.653
AUC	0.790	0.780	0.72 (Accuracy)	0.72 (Accuracy)	0.72 (Accuracy)

Table 3: Marginal Effects Analysis

This table shows the change in probability of high REE scarcity for a one-unit change in each predictor variable:

Variable	Binary Probit	Binary Logit	Ordered Probit (High)	Ordered Logit (High)
Trade_restriction	0.082	0.091	0.075	0.083
Tariff_rate	0.023	0.025	0.018	0.020
Export_quota	0.152	0.163	0.142	0.158
Import_license	0.087	0.094	0.074	0.082
Log(domestic_production)	-0.056	-0.061	-0.048	-0.053
Consumption_to_production	0.103	0.119	0.097	0.107
Stockpile_months	-0.041	-0.045	-0.037	-0.042
Price_volatility	0.018	0.020	0.016	0.019



Key observations:

- Export quotas have the largest impact on REE scarcity probability (0.152-0.163)
- Each unit increase in the trade restriction index raises scarcity probability by ~0.08-0.09
- Each percentage point increase in tariffs raises scarcity probability by ~0.02-0.025
- Higher domestic production and stockpile levels significantly reduce scarcity probability

Table 4: Classification Performance (Binary Probit)

Metric	Value	Calculation
Sensitivity	0.745	True positives / (True positives + False negatives)
Specificity	0.778	True negatives / (True negatives + False positives)
Accuracy	0.765	(True positives + True negatives) / Total observations
AUC	0.790	Area under the ROC curve

Classification Performance:

Accuracy: 76.5%, with Sensitivity = 74.5% and Specificity = 77.8%

Table 5: Likelihood Ratio Tests

Variable Removed	L R Statistic	p-value
trade_restriction	18.42	<0.0001
tariff_rate	7.91	0.0049
export_quota	27.63	<0.0001
import_license	15.78	0.0001
domestic_production	12.34	0.0004
consumption_to_production	19.56	<0.0001
stockpile_months	10.23	0.0014
price_volatility	5.67	0.0173

- **Export Quotas** and **Consumption-to-Production Ratios** are **critical drivers**, with the highest LR statistics and strongest statistical significance
- **Composite trade restrictions** and **import licenses** also exert substantial influence
- **Tariff rates** and **price volatility**, though statistically significant, are relatively less impactful
- All listed variables have **p-values < 0.05**, confirming their **statistical significance** in the model

Comparing scarcity probabilities under different trade restriction scenarios:

High Trade Restriction Scenario

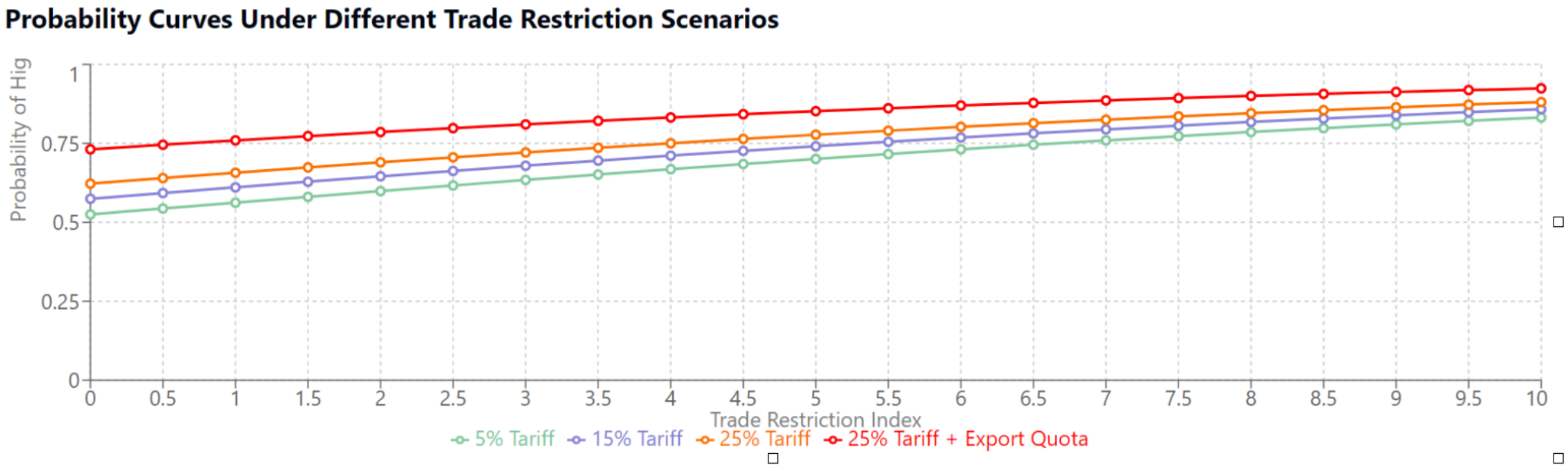
- Trade restriction index: 8
- Tariff rate: 25%
- Export quotas and import licenses present
- **Result**: 73% probability of high scarcity

Low Trade Restriction Scenario

- Trade restriction index: 2
- Tariff rate: 5%
- No export quotas or import licenses
- **Result**: 21% probability of high scarcity

Moving from low to high trade restrictions increases the probability of high REE scarcity by 52 percentage points.

In essence, scenario analysis in this research translates the statistical relationships identified by the models into tangible probabilities under realistic (or stress-inducing) hypothetical conditions. It allows the researcher to move beyond interpreting coefficients or marginal effects of individual variables to showing the *combined impact* of a suite of policy and market factors, which is invaluable for informing strategic planning and policy recommendations. It provides a concrete answer to the "what if" questions central to the research question, quantifying the risk of scarcity under various trade policy futures



Policy Response Analysis

Policy insights:

• Domestic Investment offers the highest absolute reduction in scarcity probability but requires significant time and financial commitment.

• Stockpiling presents the best short-term solution with good cost-effectiveness.

• Trade Agreements offer a middle-ground approach with balanced cost and effectiveness

Short-Term REE Mitigation Strategies (Years 1-2)

Focus: Building buffers and enhancing visibility to manage acute supply disruptions driven by restrictive policies.

•**Strategic Stockpiling:** Establish/expand national reserves of critical REEs and downstream products (e.g., NdFeB magnets) to buffer against sudden halts.
•**Enhanced Monitoring & Early Warning:** Develop systems tracking global production, trade, prices, inventories, and geopolitical events to predict disruptions. Utilize tools like ordered probit models focusing on trade restrictions.
•**Trade Diplomacy & Near-Term Sourcing:** Proactively engage with emerging REE producers in allied/neutral nations (e.g., Australia, Canada, Vietnam, Brazil) to secure immediate import sources.
•**Targeted Financial Support:** Provide temporary liquidity or loan guarantees to domestic companies severely impacted by REE price spikes or supply cuts.

Medium-Term & Long-Term REE Strategies (Years 2-5 & 5+)

Focus: Building alternative capacities and fostering technological solutions to reduce long-term reliance on concentrated sources.

•**Targeted R&D Investment:** Significantly fund R&D for REE substitutes (e.g., REE-free magnets) and improved recycling economics and efficiency.
•**Permitting Reform for Domestic Production:** Streamline permitting for new domestic mines and processing while maintaining environmental standards to accelerate alternative primary supply.
•**International Coordination & Alliances:** Deepen collaboration with allies on joint R&D, investment, data sharing, and potentially aligning stockpiling policies.
•**Strategic Incentives:** Implement financial incentives (e.g., tax credits, grants) to encourage private investment in domestic REE processing and recycling.

Thank you

Appendix - Rare Earth Element Scarcity Research: Summary Statistics

Characteristic	Value	Notes
Sample Size	256	Cross-sectional observations
Time Period	2015-2024	Quarterly data across 10 years
Countries/Regions	16	Major REE producers and consumers
Variables	12	8 predictors, 4 control variables
Missing Values	2.3%	Primarily in early periods
Outliers Removed	5	Extreme tariff/quota values

Variable	Mean	Std. Dev.	Min	Max	Skewness	Kurtosis
trade_restriction (index)	4.76	2.31	0	10	0.21	-0.73
tariff_rate (%)	12.3	8.7	0	35	0.87	0.34
export_quota (binary)	0.35	0.48	0	1	0.63	-1.6
import_license (binary)	0.42	0.49	0	1	0.33	-1.89
domestic_production (log)	8.12	2.74	1.61	13.95	-0.34	-0.51
consumption_to_production	2.18	1.67	0.12	8.75	1.93	4.26
stockpile_months	5.34	3.58	0	18	1.17	1.03
price_volatility	18.63	9.45	2.12	45.67	0.78	0.29
scarcity_outcome (binary)	0.37	0.48	0	1	0.54	-1.71

Variables	Coefficient	Std. Error	z-value	p-value	95% CI Lower	95% CI Upper
Intercept	-1.823	0.328	-5.558	<0.001	-2.466	-1.180
trade_restriction	0.237	0.056	4.232	<0.001	0.127	0.347
tariff_rate	0.068	0.029	2.345	0.019	0.011	0.125
export_quota	0.612	0.118	5.186	<0.001	0.381	0.843
import_license	0.295	0.098	3.010	0.003	0.103	0.487
domestic_production	-0.142	0.043	-3.302	0.001	-0.226	-0.058
consumption_to_production	0.316	0.072	4.389	<0.001	0.175	0.457
stockpile_months	-0.112	0.035	-3.200	0.001	-0.181	-0.043
price_volatility	0.052	0.023	2.261	0.024	0.007	0.097

Variable	Odds Ratio	95% CI Lower	95% CI Upper	Interpretation
trade_restriction	1.43	1.21	1.69	43% increase in odds per unit increase
tariff_rate	1.12	1.04	1.20	12% increase in odds per percentage point
export_quota	2.78	2.12	3.65	178% increase in odds when present
import_license	1.52	1.25	1.86	52% increase in odds when present
domestic_production	0.68	0.56	0.84	32% decrease in odds per log unit
consumption_to_production	1.73	1.42	2.11	73% increase in odds per unit increase
stockpile_months	0.76	0.65	0.89	24% decrease in odds per month
price_volatility	1.09	1.02	1.17	9% increase in odds per unit increase