CLASS 2: ECONOMIC & FINANCIAL FORECASTING

Linear – Supervised - Unsupervised

Machine learning, data mining, predictive analytics, etc. all use data to predict some variable as a function of other variables.

- May or may not care about insight, importance, patterns
- May or may not care about inference---how y changes as some x changes

Econometrics & Finance: Use statistical methods for prediction, inference, causal, modeling of economic relationships.

- Hope for some sort of insight, inference is a goal
- In particular, causal inference is goal for decision making

THE PROBLEM:

85% of financial models fail to predict large changes in macro variables. While we've gotten ok with trends, we keep failing at inflection points. -*Austan Goldsbee, Chief Economist & Chairman, President's Council of Economic Advisors.*

75% of financial forecasters data users list "out of date," and 65% list "significant errors," as the top two problems they face when using financial data. - *World Economic Forum's World Risk Report 2016*.

65% of senior management at Fortune 500 reported being ill-equipped to meet growing compliance and reporting standards. - *World Economic Forum's World Risk Report 2016*.

Less than 15% of economists & financial analysts in corporate and financial America use predictive analytics, machine learning and real time sources of data.

What econometrics & financial forecasting can learn from machine learning

- "Big Data: New Tricks for Econometrics"
- train-test-validate to avoid overfitting
- cross validation
- nonlinear estimation (trees, forests, SVGs, neural nets, etc)
- bootstrap, bagging, boosting
- variable selection (lasso and friends)
- model averaging
- computational Bayesian methods (MCMC)
- tools for manipulating big data (SQL, NoSQL databases)
- textual analysis (not discussed)

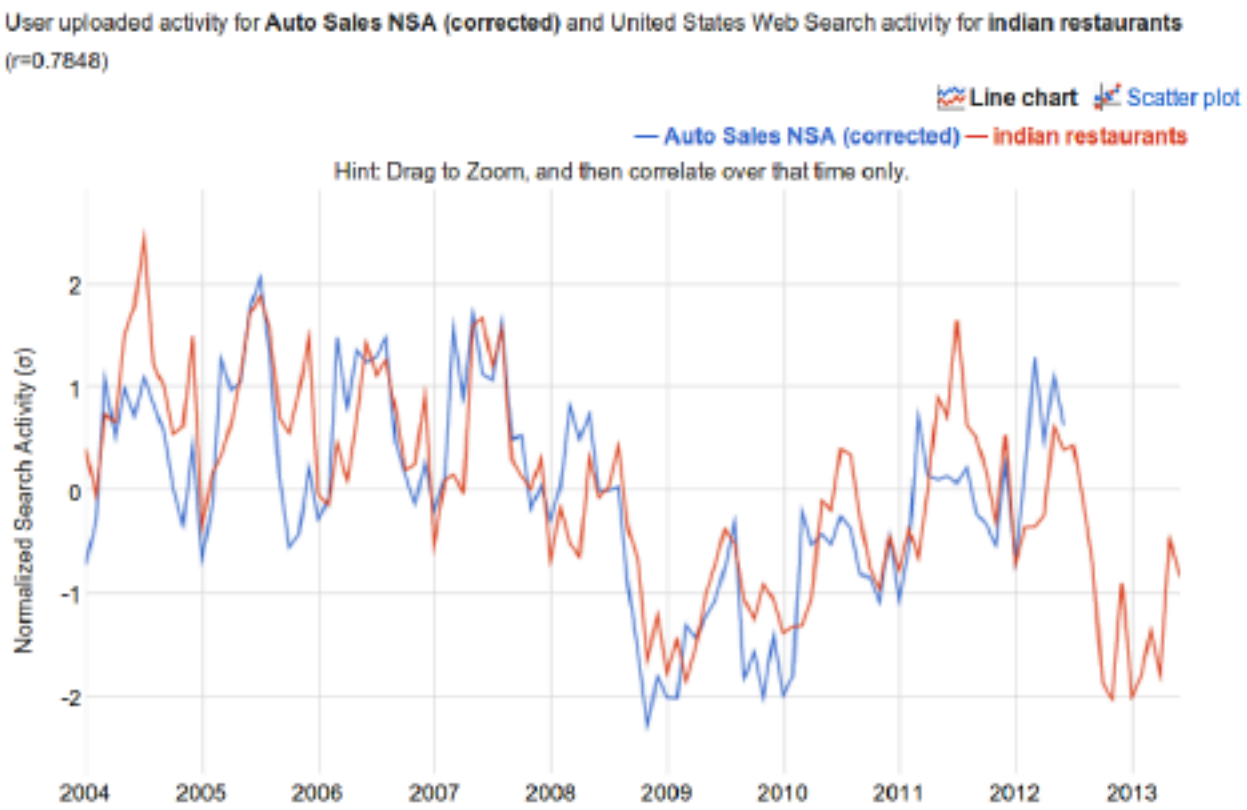
What machine learning can learn from econometrics & financial modeling

- non IID data (time series, panel data) [research topic, not in textbooks]
- causal inference -- response to a treatment [manipulation, intervention]
- confounding variables
- natural experiments
- explicit experiments
- regression discontinuity
- instrumental variables

Credit: Google

Google

NSA auto sales and Google Correlate to 2012



Credit: Google

Core & Alternative Data

Illustrative Core Data: Bond Prices

Data Category	Examples
Economy / Tax Base	Tax Base Size Full Value Per Capita Wealth (Median Family) System Characteristics / Asset Conditions Economic Concentration Tax per capita greater than regional average Weather Volatility / Extreme Conditions Natural Endowments / Disasters FEMA / Resilience Programs
Financial Strength	Debt / Fund balance as % of Revenues Debt / Fund balance trend (5 year change) Cash balances as % of revenues Cash balances trends (5 year change) Debt to Operating Revenues Constrained Liquidity Oversized Transfers & Obligations Outsized Capital Needs
Government & Management	Institutional Framework Operating History Regulatory Compliance and Capital Planning
Debt & Legal Provisions	Rate Covenant Debt Service Reserve Requirements Unusual Debt Structure / Enhancements
Pensions	Pensions Liabilities to Full Value Pension Liabilities to Revenue Pension Liabilities (3 year average) to Full Value Pension Liabilities (3 year average) to Revenue

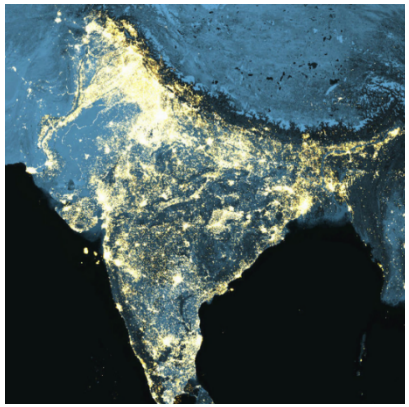
Illustrative Core Data Sources

Source	Example	Interface	Type	Update
Bloomberg	Pricing & Economic Data	API	Structured & Unstructured	Real Time
Reuters	News	API	Unstructured	Real Time
AB Systems (Data)	Portfolio Positions	API/Systems	Structured	Real Time
Research Group (Buy Side)	Research & Credit Opinions	Email, Manual	Structured & Unstructured	Weekly / Monthly / Quarterly
Research Group / Sell Side	Research & Credit Opinions	Email, Manual	Structured & Unstructured	Weekly / Monthly / Quarterly
News	Political & Economic Announcements	Web	Unstructured	Real Time
USG - Bureau of Labor Statistics	Economic Data, Unemployment, etc.	API, e-government portals	Structured	Weekly
Municipalities – Treasury & Finance Dept.	Financial Statements / Budgets	Email, Manual, Web, e-government portals	Structured & Unstructured	Quarterly
SEC	Enforcement Actions, Issuance Data	Email, Web, e-government portals	Structured & Unstructured	Daily
Accounting Firms	Audited Financial Statements	Email, Web, e-government portals	Structured & Unstructured	Quarterly
Consulting / Research Groups	Feasibility, Economic, Engineering Studies	Email, Manual	Structured & Unstructured	On Call
Rating Agencies	Research & Credit Opinions, Credit Ratings, Credit Actions	Email, Web.	Structured & Unstructured	Real Time

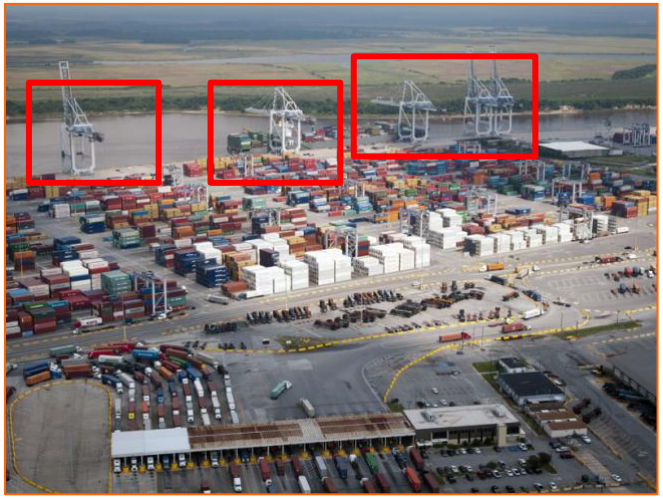
Illustrative Alternative Data

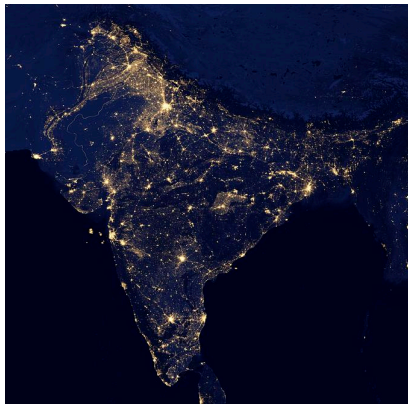
PALM OIL PRODUCTION, LAND USE

GDP OUTPUT PROXIES



SHIPS & ROUTES MONITORING





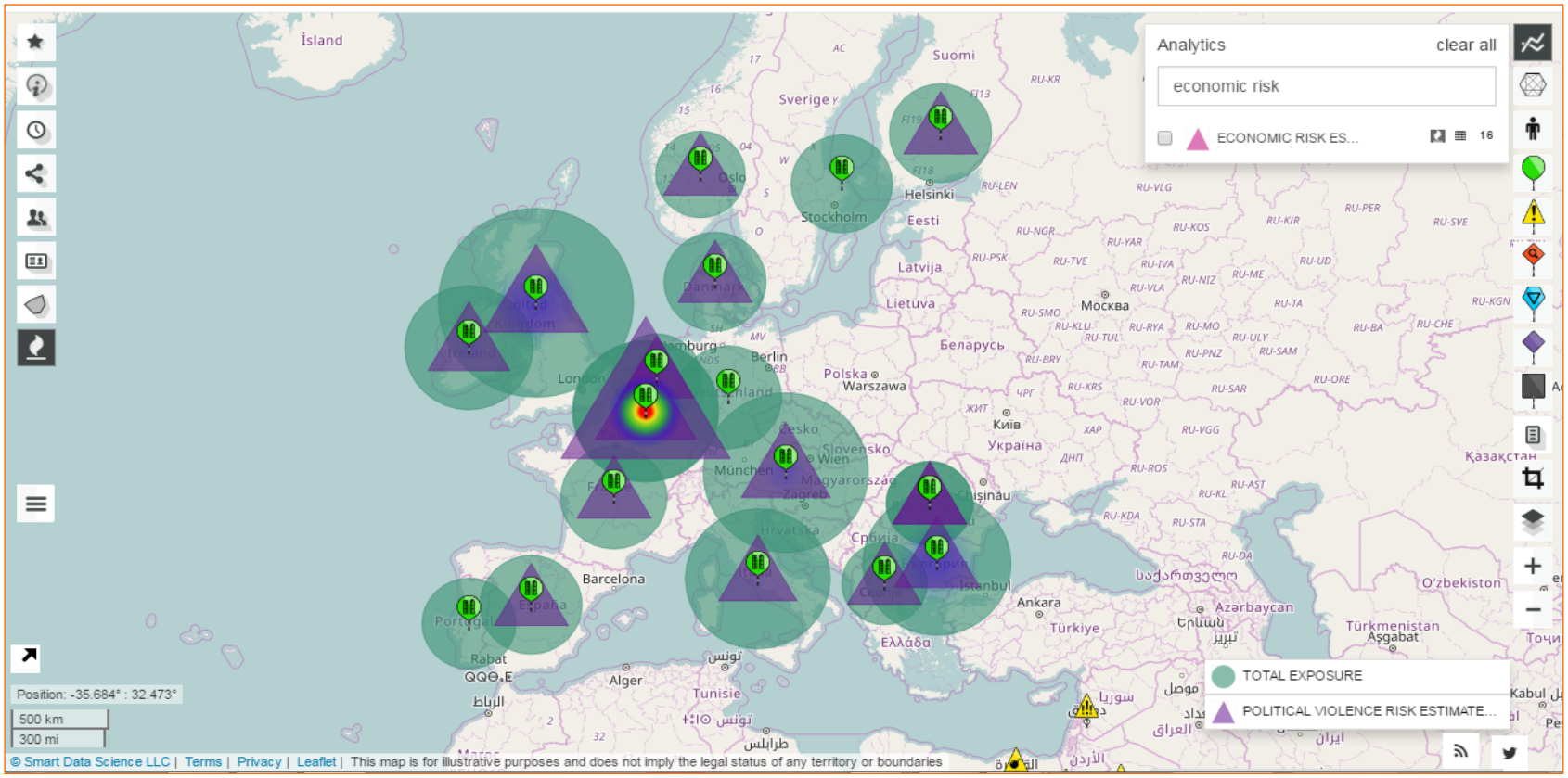
ALTERNATIVE DATA SAMPLE CORRECTION

Gen	Size	GDP/Inflation	Terrorism	P&Ls
1st	News & Web	X	X	
	Facebook		X	X
	Google	X	X	
	Twitter			
2nd	Satellites	X	X	X
	Drones		X	X
	Weather	X	X	X
	Terrain		X	
	Movements	X	X	X
3rd	Micro Satellites		X	tbc
	Micro Drones		X	tbc
	Email Inboxes	X		X
	Credit Cards	X		X
	App Data	X		X
	IoT	X		

Classification & Taxonomies

SIMILAR BUT FOR EUROPE

3) BREXIT SOCIAL UNREST FINANCIAL IMPACT



Credit: MaKro LLC / SDS LLC

Modeling

PROCESS

Data management and overall workflow
1. Divide the data into a training sample and a out of sample.

2. On the training sample, run each of the prediction algorithms
- No tuning / OLS
- Complexity / train empirically
- Ensembles weights, cross-validation, etc.
- Account for clustering or data behavior/patterns/anomalies
- Especially when merging alt+core datasets

Run and look for best R2 given set CI.

3. Tuning
- Impose / expand scarcity
- Complexity parameters - such as the depth of regression tree
- Separate into multiple sub samples

Run….

4. Train/select algorithm for preferred sub sample performance
- Preferred is not necessarily same as best performance metrics
- Fit the algorithm with desired tuning parameter on the full training sample.

Credit: Mullainathan / Spiess

Full pre adjustments: https://assets.aeaweb.org/assets/production/files/4212.pdf

Linear Regression (Skipping)

Dependent Variable

Current Values
- Rating / Default Probability
- Economic Data
- Financial Metrics
- Bankruptcy

Forecast Values
- Rating / Default Probability
- Economic Data
- Financial Metrics
- Bankruptcy

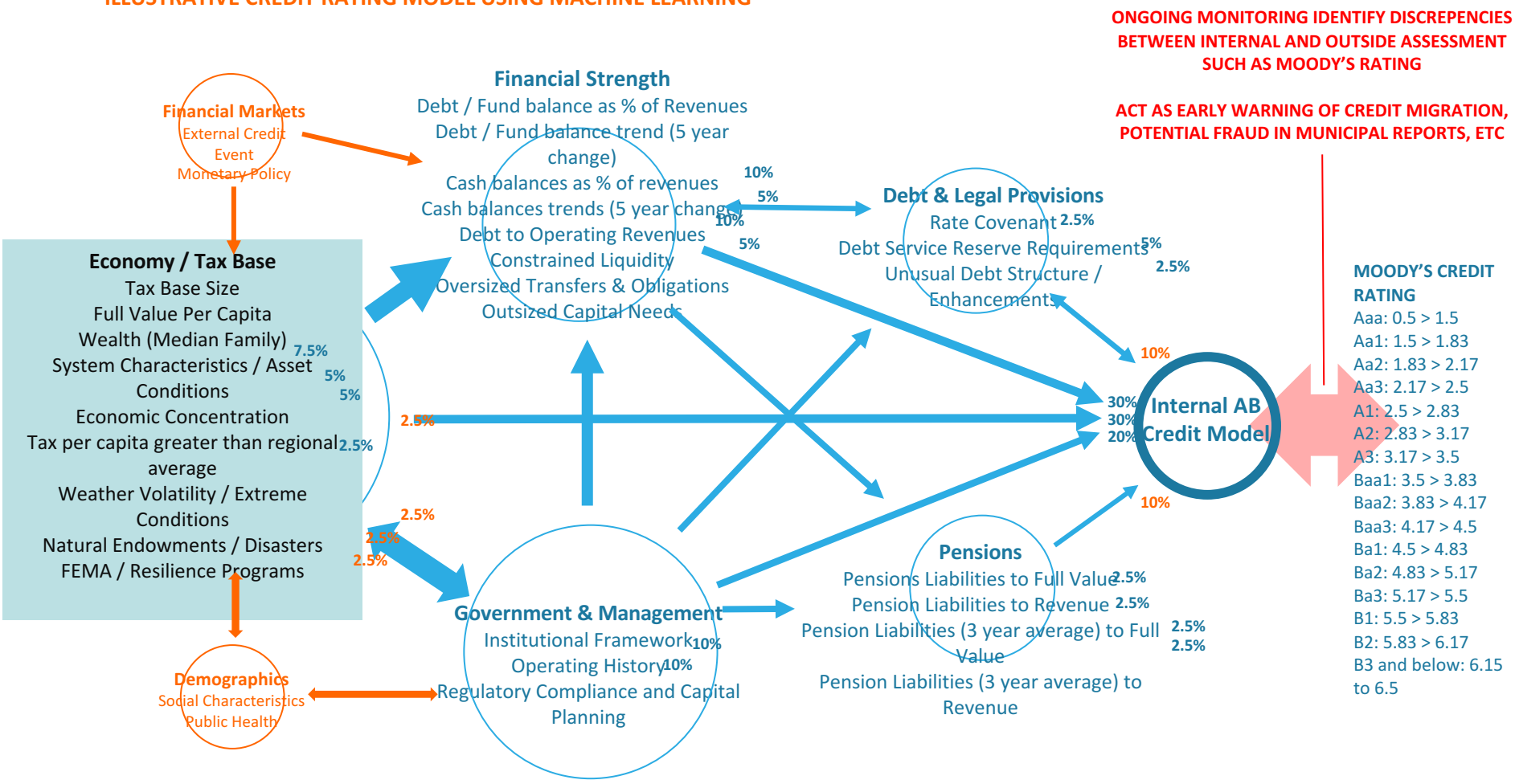
Logit		Firthlogit	Firthlogit
VARIABLES	No Fixed Effects	No Fixed Effects	With Fixed Effects
lnPOP	0.600***	0.583***	0.715***
(0.18)	(0.19)	(0.23)	
SEI	-0.112	-0.105*	-0.141**
(0.07)	(0.05)	(0.06)	
INT_BY_REV	16.48***	15.96***	22.34***
(2.96)	(3.20)	(6.00)	
REV_CHANGE	-4.386**	-4.067**	-3.444
(1.88)	(2.06)	(2.22)	
Constant	-7.213**	-7.118***	-9.056***
(3.01)	(2.54)	(3.47)	
Estimation strategy	logit	firthlogit	firthlogit
State fixed effects?	no	no	yes
Correctly classified	270	269	281
Observations	305	305	305
% correctly classified	88.5	88.2	92.1

Variable	Coefficient	Standard Error
Tax Levy per $1,000 Assessed Value	-0.00310	0.00247
Tax Levy per Capita	-0.00115	0.00108
Debt/Assessed Property Values	+0.3521	0.17000
Percentage of Current Taxes Delinquent	+0.07209	0.07277

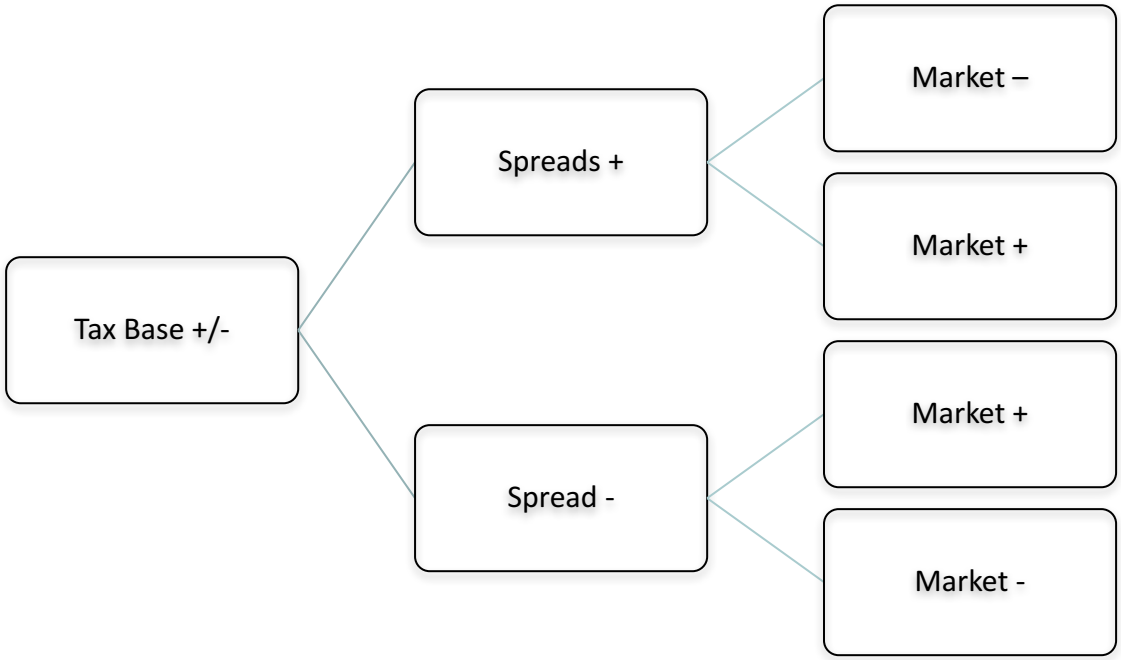
Variable	Coefficient	Standard Error
Log of 1932 Population	-0.07678	0.0321
Assessed Property Value per Capita in 1932	+0.0001585	0.0000523
Growth of Population from 1922 to 1932	-0.02146	0.0113
Growth of Debt Relative to Population Growth	-0.007912	0.00213
Debt/Assessed Property Values in 1932	+0.4885	0.258
Tax Levy per $1,000 Assessed Value in 1932	+0.00919	0.00242
Tax Levy per Capita in 1932	-0.007197	0.00322
Percentage of Current Taxes Delinquent in 1932	+0.2095	0.0962
Notes Outstanding per Capita in 1932	+0.009159	0.00246

MODEL: Graphical – Business Rules

ILLUSTRATIVE CREDIT RATING MODEL USING MACHINE LEARNING



…..Into decision tree



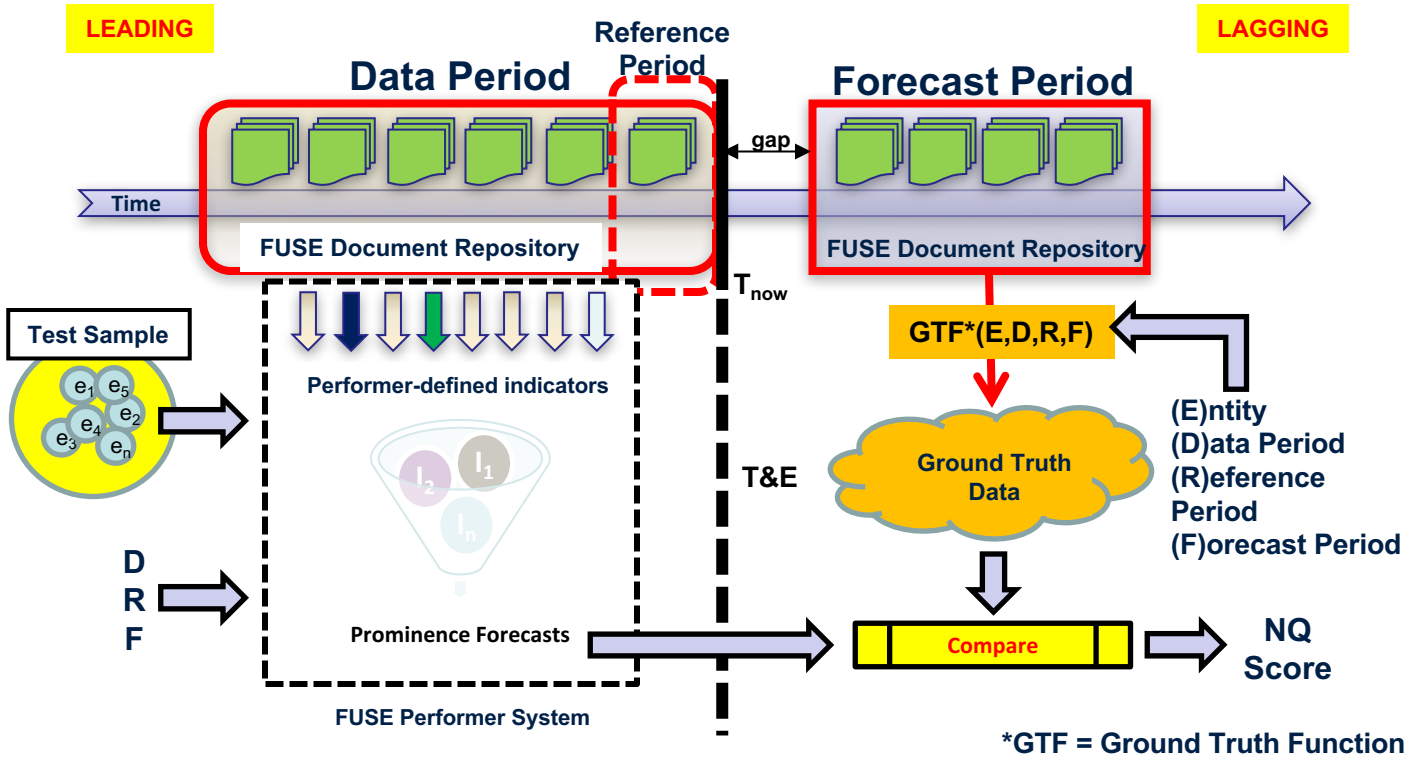
Economy / Tax Base
Tax Base Size
Full Value Per Capita
Wealth (Median Family)
System Characteristics / Asset Conditions
Economic Concentration
Tax per capita greater than regional average
Weather Volatility / Extreme Conditions
Natural Endowments / Disasters
FEMA / Resilience Programs

Financial Strength
Debt / Fund balance as % of Revenues
Debt / Fund balance trend (5 year change)
Cash balances as % of revenues
Cash balances trends (5 year change)
Debt to Operating Revenues
Constrained Liquidity
Oversized Transfers & Obligations
Outsized Capital Needs

Debt & Legal Provisions
Rate Covenant
Debt Service Reserve Requirements
Unusual Debt Structure / Enhancements

MOODY'S CREDIT RATING
Aaa: 0.5 > 1.5
Aa1: 1.5 > 1.83
Aa2: 1.83 > 2.17
Aa3: 2.17 > 2.5
A1: 2.5 > 2.83
A2: 2.83 > 3.17
A3: 3.17 > 3.5
Baa1: 3.5 > 3.83
Baa2: 3.83 > 4.17
Baa3: 4.17 > 4.5
Ba1: 4.5 > 4.83
Ba2: 4.83 > 5.17
Ba3: 5.17 > 5.5
B1: 5.5 > 5.83
B2: 5.83 > 6.17
B3 and below: 6.15 to 6.5

Phase 2 Evaluation: Nomination Test



Credit: DARPA/IARPA Unclassified

NEXT:

1- GROUPS

2- ASSIGNMENT DISCUSSION