

RISK BEHAVIOR IN PUBLIC PENSION FUNDS

Bethany Bailey

Masters in Computational Social Science, University of Chicago



INTRODUCTION

RESEARCH QUESTION:

How do fund characteristics and state economic and political ideology factors influence risk (measured by equity allocation) in public pension funds in the United States?

Fund characteristics: accounting assumptions and performance

Previous literature suggests that fund performance, board and member composition, and accounting assumptions about liabilities and returns might influence risk in investment behavior. Thus, we will use measures of these in our model.

Why would state economic and political factors effect risky investment behavior in funds?

Literature suggests that investment behavior on pension boards and by individual investors is influenced by political motivation and economic stability. This study draws on that research, using measures of state ideology and polarization to try to predict fund risk-taking behavior.

DATA

This project uses data from three primary data sources:

- 1. **Public Plans Data:** contains annualized details on 170 large U.S. pension plans (95% of population)
 - Years: 2001-2016.
 - Variables: over 150
 - Topics include: funding, accounting assumptions, allocation, returns, and fund characteristics
- 2. **Correlates of State Policy Project (MSU):** contains annualized state-level data on all 50 U.S. states
 - Years: 1900-2016 (with missingness)
 - Variables: over 900
 - Topics include: economic and fiscal policy, education, election information, government, public opinion, partisanship, and ideology
- 3. **Shor-McCarthy State Legislative Aggregate Ideology Data**
 - Years: 1993-2016
 - Variables: 26
 - Topics include: state ideology and political polarization

KEY VARIABLE DISTRIBUTIONS

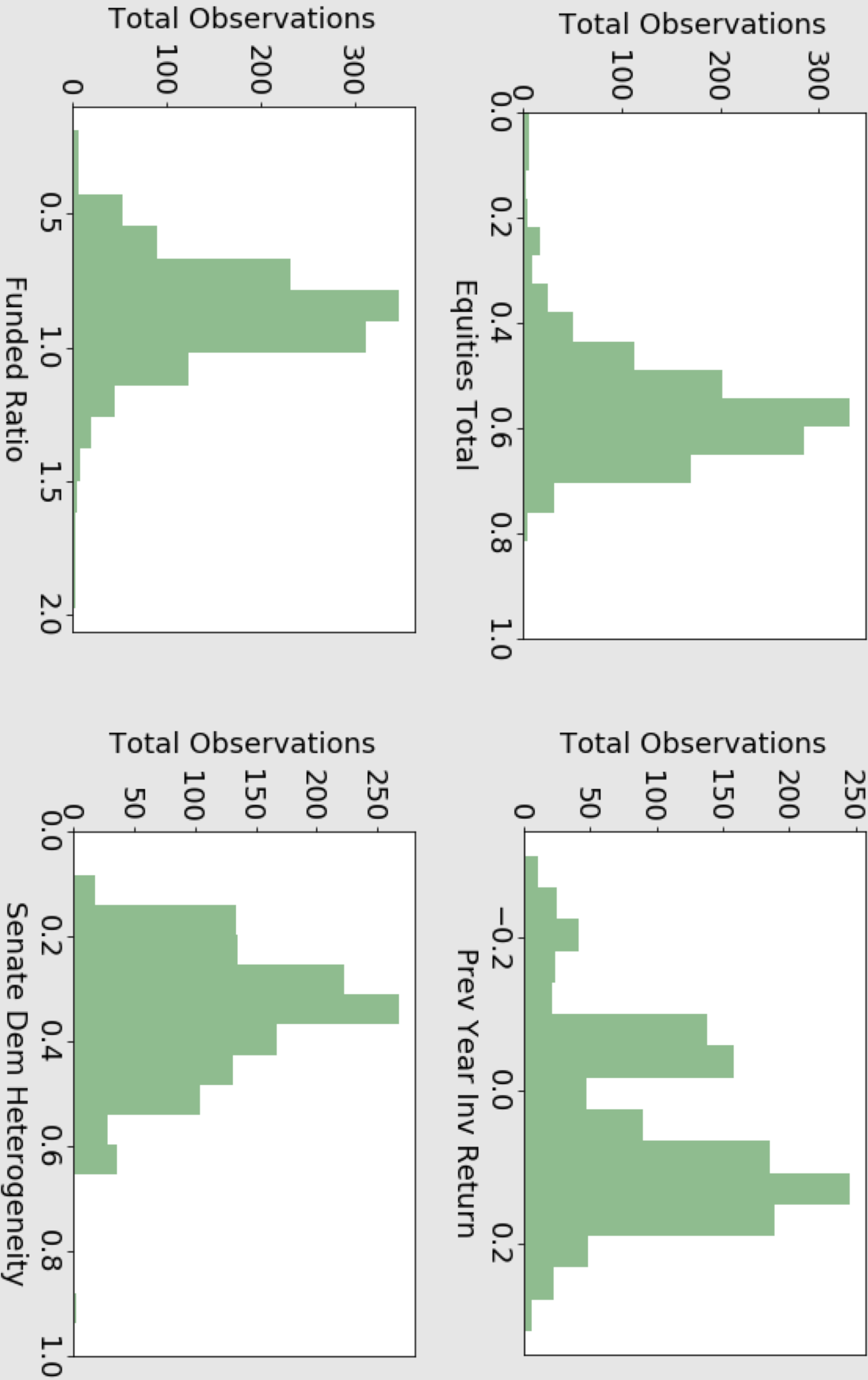


TABLE 1: MODEL ERROR RATES

Model Type	Decision Tree (one train/test split)	Decision Tree (Bootstrapped)	Random Forest Model	Neural Network (4-folds)
MSE	0.0091	0.0073	0.0038	0.0074
Error	0.0956	0.0859	0.0621	0.0865

METHODS

RESULTS

STEP 1: VARIABLE SELECTION

Exploratory data analysis using plots and Principal Component Analysis to select variables (final independent var n =19)

STEP 2: PREPROCESSING

Clean NAs and scale variables ~ $N(0,1)$

STEP 3: COMPETING MODELS

1. Decision Tree (one train-test split)
2. Decision Tree (bootstrapped)
3. Random Forest Model
4. Neural Network with 4-fold cross-validation

- As seen in Table 1, the *Random Forest Model performed the best*. This model has a fairly low error rates ($MSE = 0.0038$). However, given the distribution of the data, this rate is *only slightly significant*.
- The decision tree with five layers shown below ($MSE = 0.0091$) shows that all *three factor categories play a role in partitioning the data*.
- In general, from the below tree and other trees not shown, it appears that *higher homogeneity in politics is associated with a higher return*.

FIVE-LAYER DECISION TREE MODEL

