

# ス 「 ス 人

Masters in Computation മ cienc D Univers T hicago



# INTRODUC TION N

## ARCH QUESTION:

ured pension funds ideology by fac

about influence liabiliti ugge

#### politic nvestment actor behavior

esearch, and Ing polarization to 9 state that to

## DATA

project uses sources: data from three primary

- Public annualized details on 170 large U.S. pension plans (95% of population) Plans Data: contains
- Years: 2001-2016.
- Variables: over 150
- **Topics** assumptions, alloca fund characteristics include: e: funding, accounting allocation, returns, an
- **N** 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
- policy, education, election information, government, public opinion, partisanship, and ideology Topics include: economic and fiscal
- Shor-McCarthy State Aggregate Ideology Data Legislative
- Years: 1993-2016
- Variables: 26
- Topics include: state political polarization ideology

RIBUT

SNOI

ĭ N

San

nples

Leaf:

0

Year

Return

True

False

urn <= 0.01

samples

mse = 0.0 samples = value = 0.

.049 = 10 0.35

value = 0.564

Dem

Heterogeneity

0.493

mse

0.01

samples value = (

s = 691 0.569

PA

RAMETERS:

Sen Rep Heterogeneity

0.551

mse

0.011

samples

739

value = 0.561

7

Туре	(one train/test split)	(Bootstrapped)	Mode	Network (2 folds)
MSE	0.0091	0.0073	0.0038	0.0074
Error	0.0956	0.0859	0.0621	0.0865

state

fund risk behavi

political ideology and pension risk behavior, albeit a small

or,

There

2

a relationship between

CONCLUSIONS

one.

Following

Following previous work on board ideology and wealth, the state political and economic environment

MODEL

RA

#### STEP \_\_ **VARIABLE** SELEC

Exploratory data analysis using Principal Component Analysis to variables (final independent var plots ar o select  $\supset$ 

# STEP 2: **PREPROCESSING**

Clean NAs and scale variables N(0,1)

STEP ω COMPETING MODELS

- -Deci ision Tree (one train-test plit)
- 2 Deci ision Tree (bootstrapped)
- 4 Random Forest Model
- 4 Neu Network with 4 fold cross

# RESULTS

- Model has of significant. 0.0038). the מ data formed owever, given the distribution , this rate is *only slightly* low error the best. This I model
- The decision below (MSE = three factor categories play partitioning the data. tree with five layers shows 0.0091) shows that all a role in
- other higher 5 associated trees not shown, i homogeneity in p with it appears politics is **=**: return.

Table the Random Forest

associations, but allocations, but in

associated with

Higher funded ratios

acios were generally lower equity

also relates to ri

sk-taking behavior.

with higher

equi

investment

t higher previous urns were associated uity allocations,

which is

an inco

nsistent finding.

**Limitations:** 

played a role in

the

models

Cost accounting

methods

also

Error rates, whil

e low, are not very

significant

Neural Net:

it is

difficult to

see

which variables

are

playing the

roles

Low interpretab

ility

of

RFM

and

general, from the higher below tree and

# ECISION

MODEL

#### 150 150 250 200 100 200 100 250 50 50 0 0.2 0.4 0.6 0.8 Senate Dem Heterogeneity 0.2 Prev Year Inv Return 0.0 0.2 mse = 0.001 samples = 11 value = 0.548 ax-depth: 4 mse = 0.006 samples = 12 value = 0.504 unded Ratio <= 0. mse = 0.009 samples = 38 value = 0.48 House member mse = 0.01 samples value mse = 0.009 samples = 15 value = 0.412

**Total Observations** 

Total Observations

s = 270.453

Taxes (%GSP) <= 4 mse = 0.008

.905

GASB Cost Method < mse = 0.021

samples

s = 89 0.521

value

samples = 602 value = 0.576

value

samples value = (

\$ = 202 0.553

mse = 0.006 samples = 400 value = 0.588

mse = 0.008 samples = 64 value = 0.565

mse = 0.036 samples = 25 value = 0.408

mse:

0.01

8

0

Funded

Ratio

0

200

300

Equities

0

0.6 Total

Total Observations

Total Observations

300

# Ideas largest

- dataset had 1,233 observations
- Use deep learning neural nets with

# for Future Work:

- Build with more data: the cleaned
- regularization

### NEURAL PAR AMET ERS

- Logistic Activation
- **LBFGS** Solver
- 4-Folds

#### CONI AC ME

PHONE: **GITHUB:** github.com/bethbailey EMAIL: baileyb@uchicago.edu (616) 402-7831



