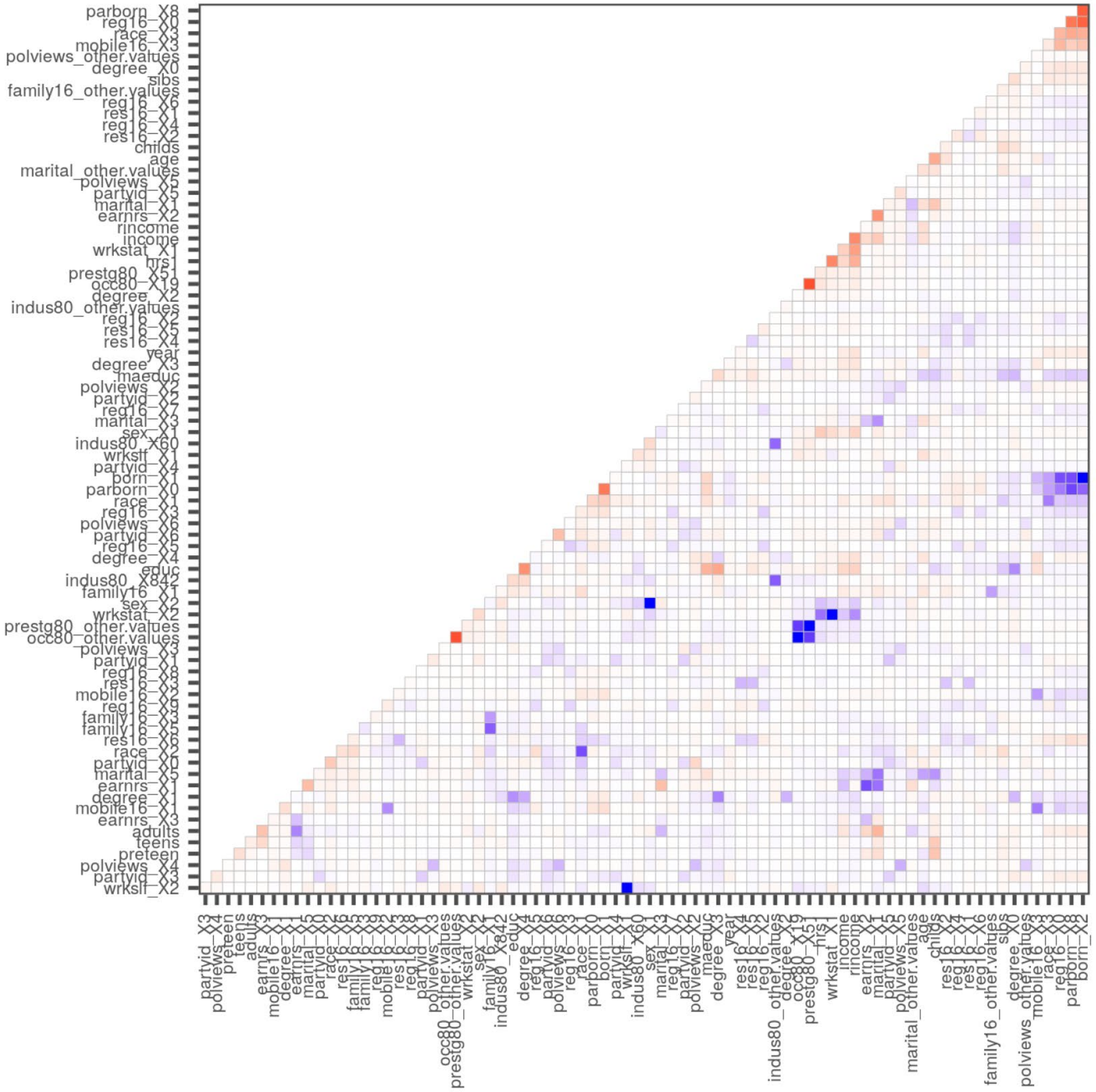
Causal Tree Algorithm

- Divide data into tree-building $\mathcal{S}^{tr}$ and estimation $\mathcal{S}^{est}$ samples
- Use a greedy algorithm to recursively partition covariate space $\mathcal{X}$ into a deep partition Π
  - At each node the split is selected as the one that minimizes our estimate of EMSE over all possible binary splits
  - Preserve minimum number of treated and control units in each child leaf
- Use cross-validation to select the depth d^* of the partition that minimizes an estimate of MSE of treatment effects, using left-out folds as proxies for the test set
- Select partition Π^* by pruning Π to depth d^*, pruning leaves that provide the smallest improvement in goodness of fit
- Estimate the treatment effects in each leaf of Π^* using the estimation sample $\mathcal{S}$

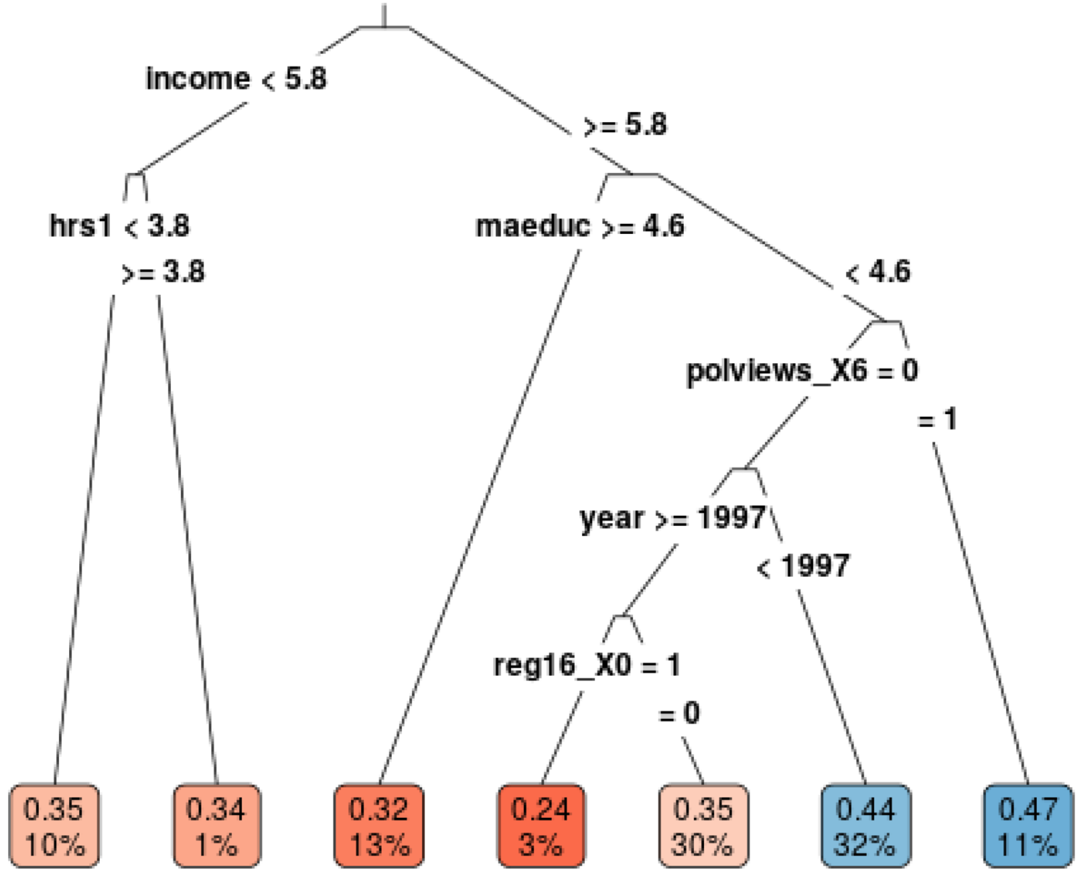
Causal Trees: GSS Survey Experiment

- Survey experiment in the General Social Survey conducted between 1986 and 2010 (Green and Kern, 2012)
- Respondents randomly assigned to one of two questions about public spending
- Experimental condition respondents were asked about "welfare" and in the other they were asked about "assistance to the poor"
- Outcome of interest: expressed support for government spending in this domain

GSS Survey Experiment, 88 covariate correlation matrix

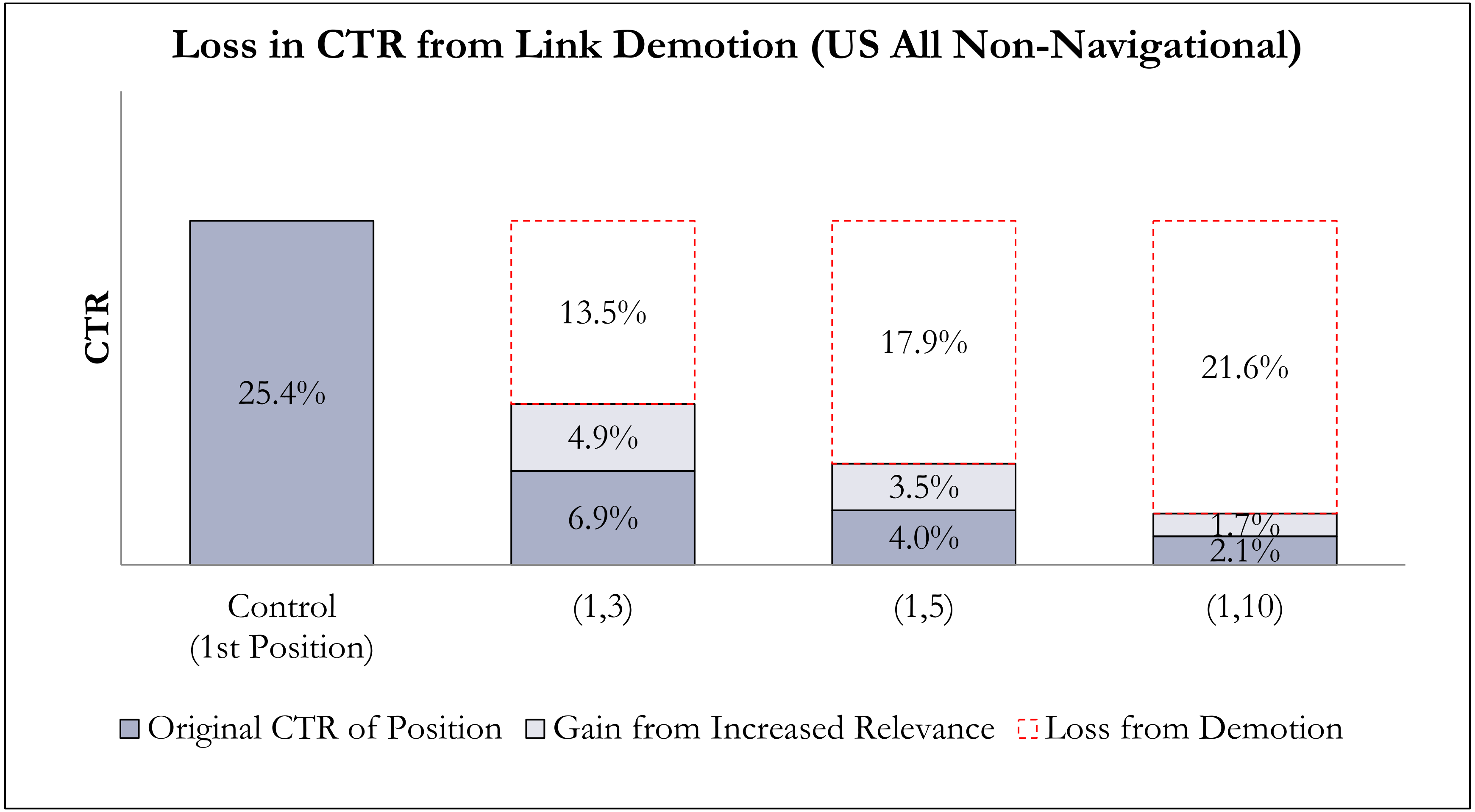


Causal Tree: Visualize the Tree

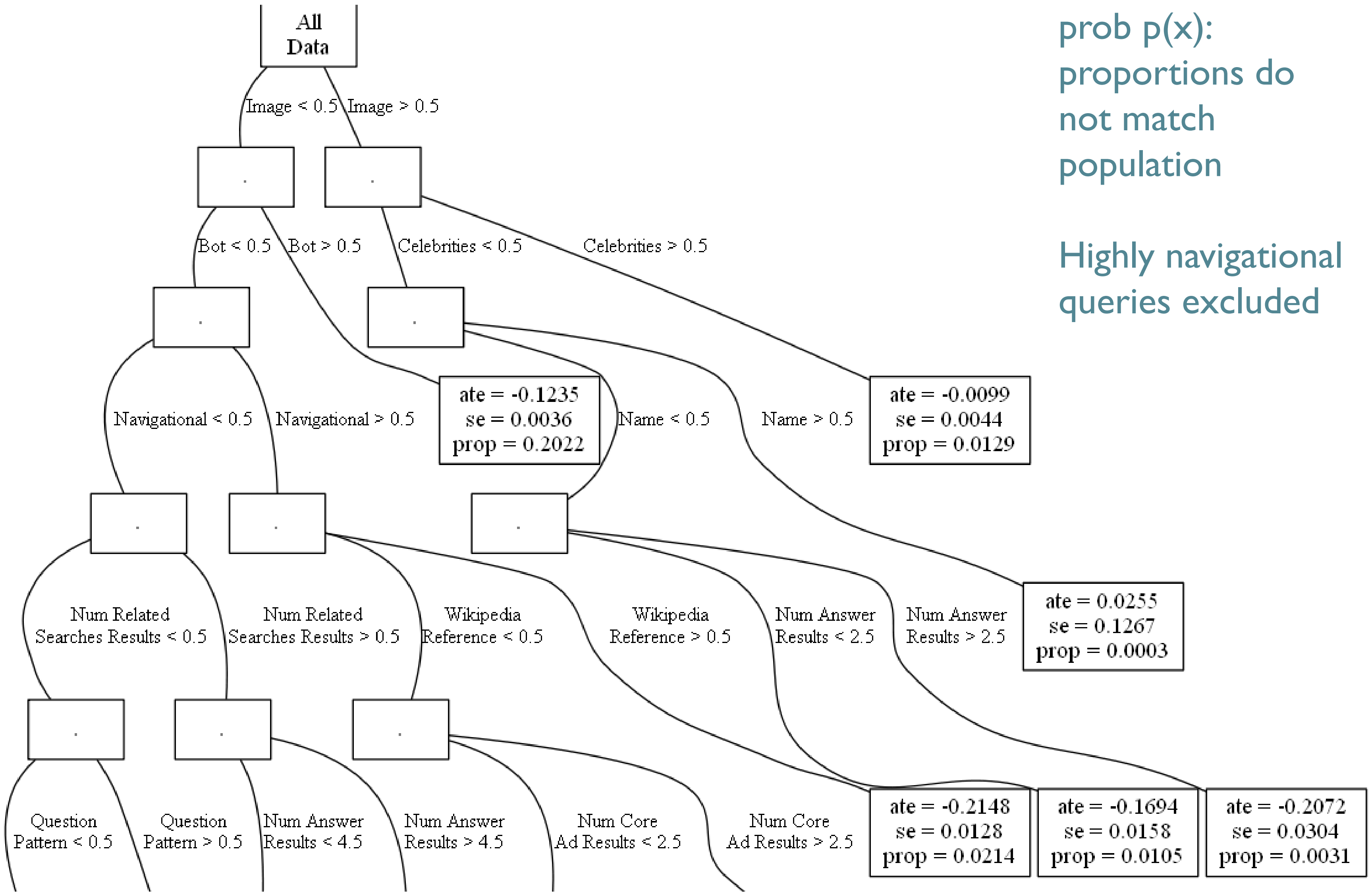



```
Linear hypothesis test

Hypothesis:
leaf1:W - leaf2:W = 0
leaf1:W - leaf3:W = 0
leaf1:W - leaf4:W = 0
leaf1:W - leaf5:W = 0
leaf1:W - leaf6:W = 0
leaf1:W - leaf7:W = 0

Model 1: restricted model
Model 2: Y ~ leaf + W:leaf

  Res.Df Df      F    Pr(>F)
1   5272
2   5266  6 4.4771 0.0001575 ***
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1
```

Relevance v. Position

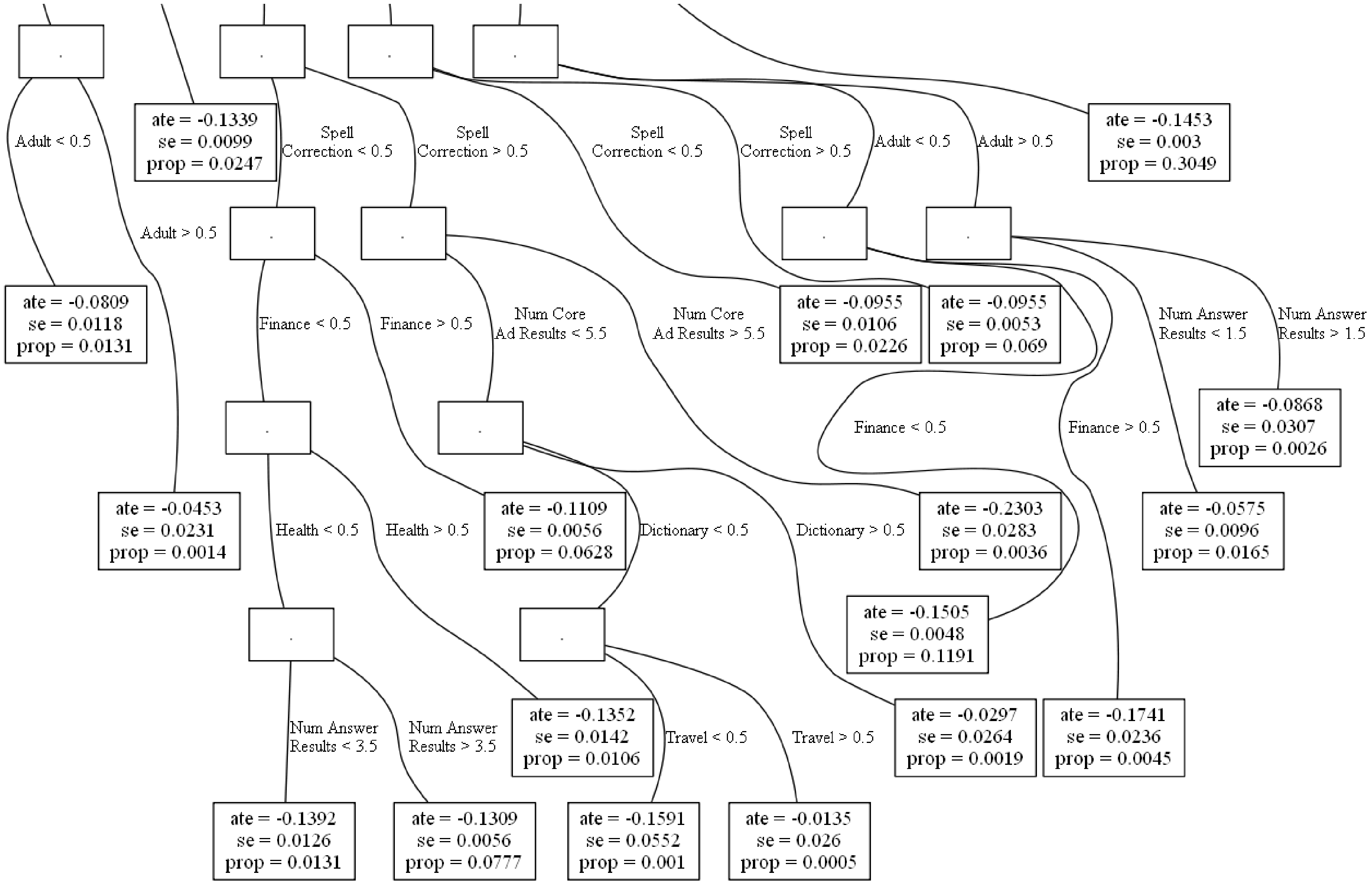


Search Experiment Tree: Effect of Demoting Top Link (Test Sample Effects)



Some data excluded with prob p(x): proportions do not match population

Highly navigational queries excluded



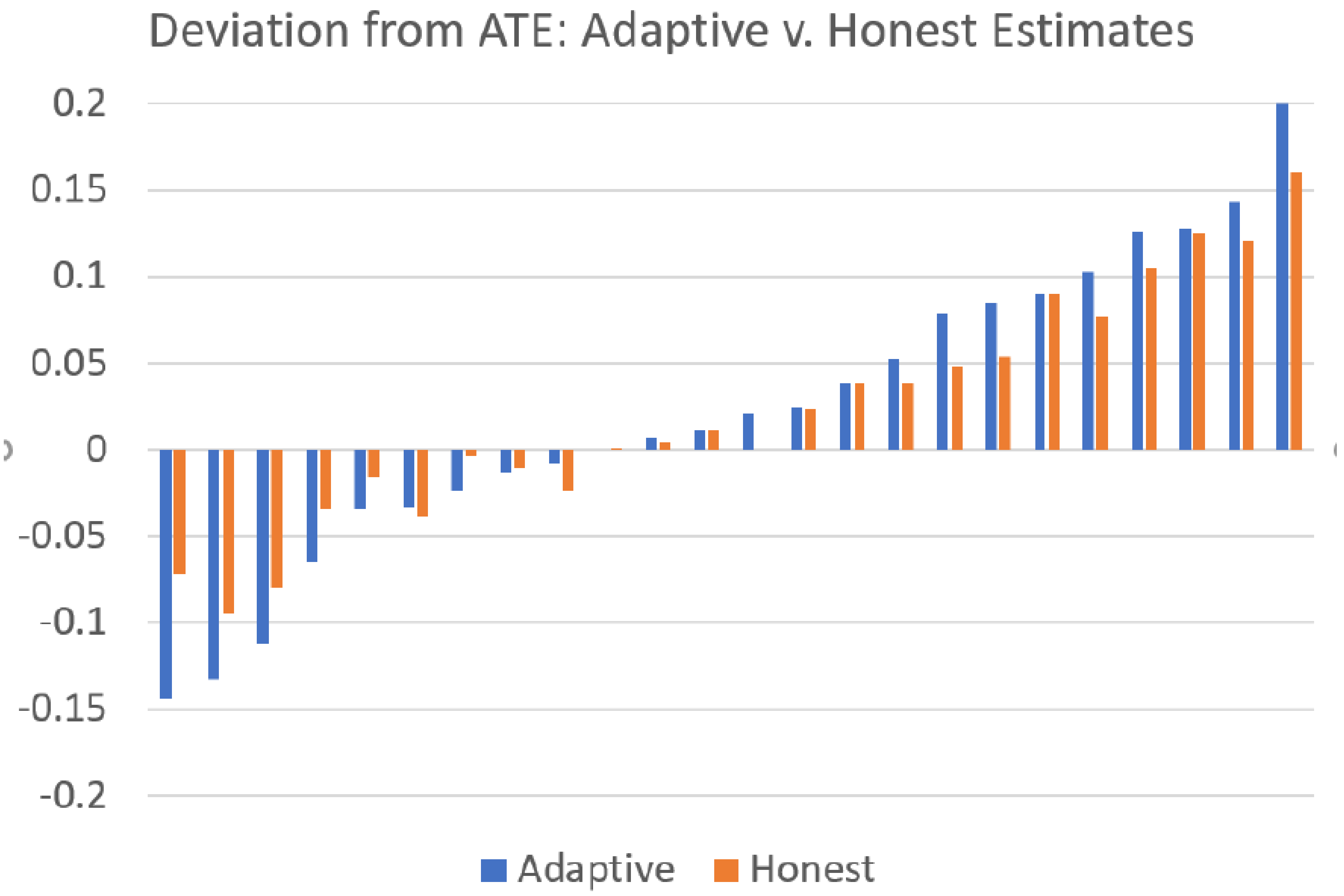
Adult < 0.5
Adult > 0.5
ate = -0.1339
se = 0.0099
prop = 0.0247
Spell Correction < 0.5
Spell Correction > 0.5
Spell Correction < 0.5
Spell Correction > 0.5
Adult < 0.5
Adult > 0.5
ate = -0.1453
se = 0.003
prop = 0.3049
ate = -0.0809
se = 0.0118
prop = 0.0131
Finance < 0.5
Finance > 0.5
Num Core Ad Results < 5.5
Num Core Ad Results > 5.5
ate = -0.0955
se = 0.0106
prop = 0.0226
ate = -0.0955
se = 0.0053
prop = 0.069
Num Answer Results < 1.5
Num Answer Results > 1.5
Finance < 0.5
Finance > 0.5
ate = -0.0868
se = 0.0307
prop = 0.0026
ate = -0.0453
se = 0.0231
prop = 0.0014
Health < 0.5
Health > 0.5
ate = -0.1109
se = 0.0056
prop = 0.0628
Dictionary < 0.5
Dictionary > 0.5
ate = -0.2303
se = 0.0283
prop = 0.0036
ate = -0.0575
se = 0.0096
prop = 0.0165
ate = -0.1505
se = 0.0048
prop = 0.1191
Num Answer Results < 3.5
Num Answer Results > 3.5
ate = -0.1352
se = 0.0142
prop = 0.0106
Travel < 0.5
Travel > 0.5
ate = -0.0297
se = 0.0264
prop = 0.0019
ate = -0.1741
se = 0.0236
prop = 0.0045
ate = -0.1392
se = 0.0126
prop = 0.0131
ate = -0.1309
se = 0.0056
prop = 0.0777
ate = -0.1591
se = 0.0552
prop = 0.001
ate = -0.0135
se = 0.026
prop = 0.0005

Use Test Sample for Segment Means & Std Errors to Avoid Bias

Variance of estimated treatment effects in training sample 2.5 times that in test sample (adaptive estimates biased)

Honest Estimates			Adaptive Estimates		
Treatment Effect	Standard Error	Proportion	Treatment Effect	Standard Error	Proportion
-0.124	0.004	0.202	-0.124	0.004	0.202
-0.134	0.010	0.025	-0.135	0.010	0.024
-0.010	0.004	0.013	-0.007	0.004	0.013
-0.215	0.013	0.021	-0.247	0.013	0.022
-0.145	0.003	0.305	-0.148	0.003	0.304
-0.111	0.006	0.063	-0.110	0.006	0.064
-0.230	0.028	0.004	-0.268	0.028	0.004
-0.058	0.010	0.017	-0.032	0.010	0.017
-0.087	0.031	0.003	-0.056	0.029	0.003
-0.151	0.005	0.119	-0.169	0.005	0.119
-0.174	0.024	0.005	-0.168	0.024	0.005
0.026	0.127	0.000	0.286	0.124	0.000
-0.030	0.026	0.002	-0.009	0.025	0.002
-0.135	0.014	0.011	-0.114	0.015	0.010
-0.159	0.055	0.001	-0.143	0.053	0.001
-0.014	0.026	0.001	0.008	0.050	0.000
-0.081	0.012	0.013	-0.050	0.012	0.013
-0.045	0.023	0.001	-0.045	0.021	0.001
-0.169	0.016	0.011	-0.200	0.016	0.011
-0.207	0.030	0.003	-0.279	0.031	0.003
-0.096	0.011	0.023	-0.083	0.011	0.022
-0.096	0.005	0.069	-0.096	0.005	0.070
-0.139	0.013	0.013	-0.159	0.013	0.013
-0.131	0.006	0.078	-0.128	0.006	0.078

Devia


Deviation from ATE: Adaptive v. Honest Estimates
0.2
0.15
0.1
0.05
0
-0.05
-0.1
-0.15
-0.2
Adaptive
Honest

Low-Dimensional Representations v. Fully Nonparametric Estimation

Causal Trees

- Move the goalpost, but get guaranteed coverage
- Easy to interpret, easy to mis-interpret
- Can be many trees
- Leaves differ in many ways if covariates correlated; describe leaves by means in all covariates

Causal Forests

- Attempt to estimate $\tau(x)$
- Can estimate partial effects
- In high dimensions, still can have omitted variable issues
- Confidence intervals lose coverage in high dimensions (bias)