

Causal Inference Course Final Project

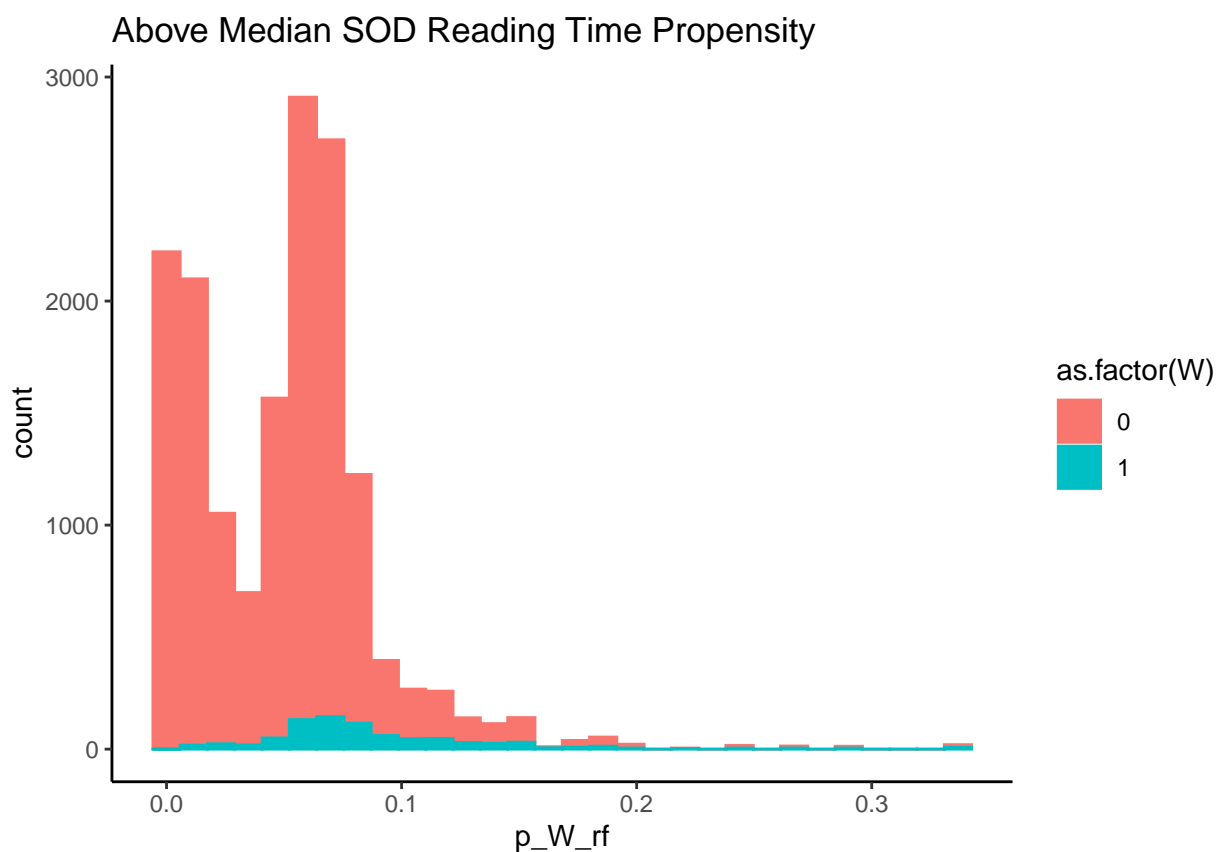
Ayush Kanodia and Mitchell Linegar

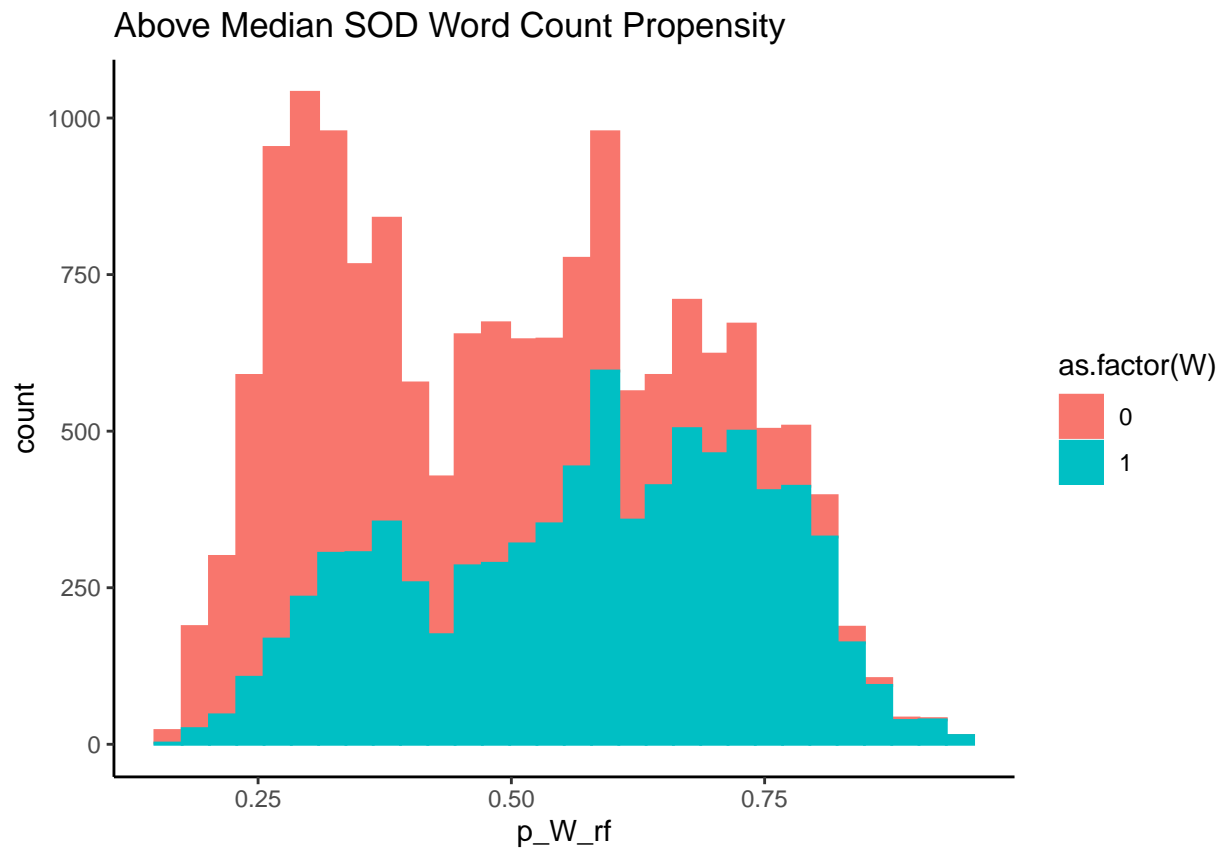
2020-06-05

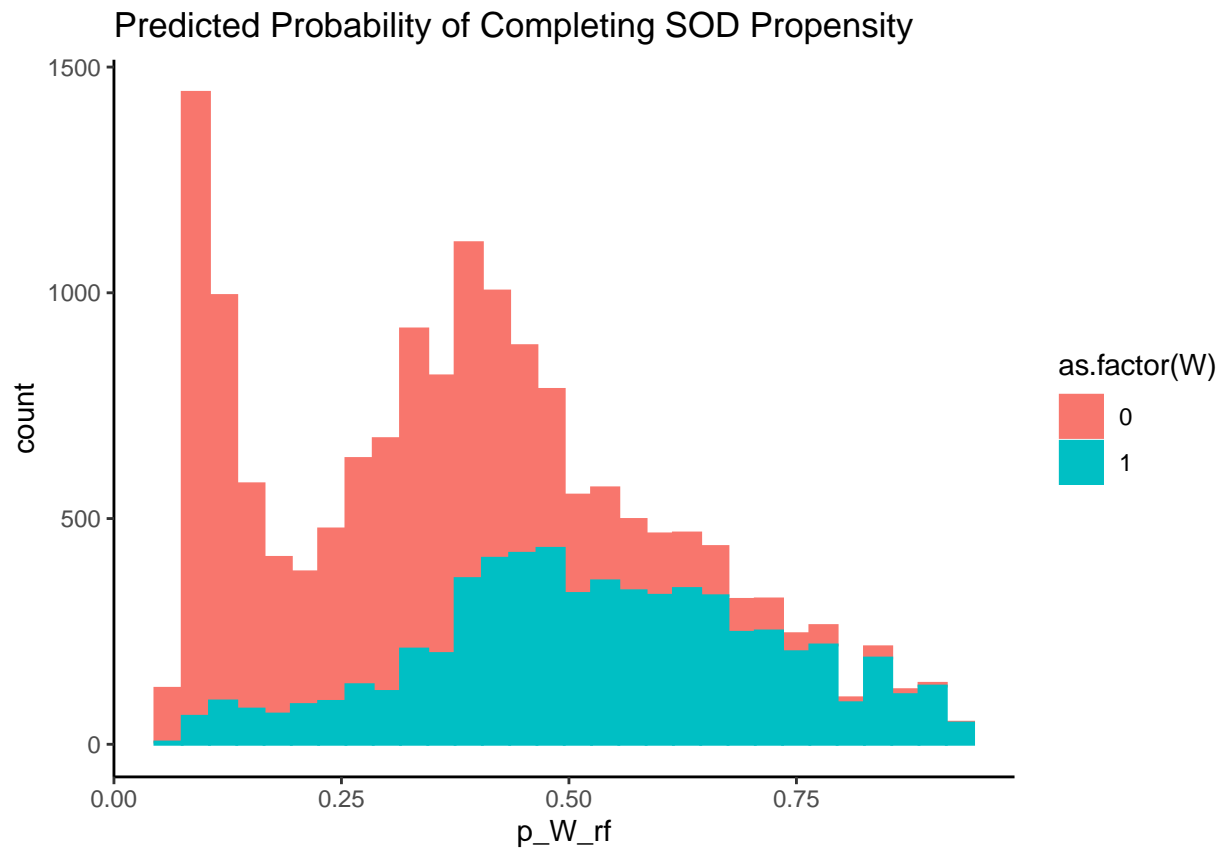
Contents

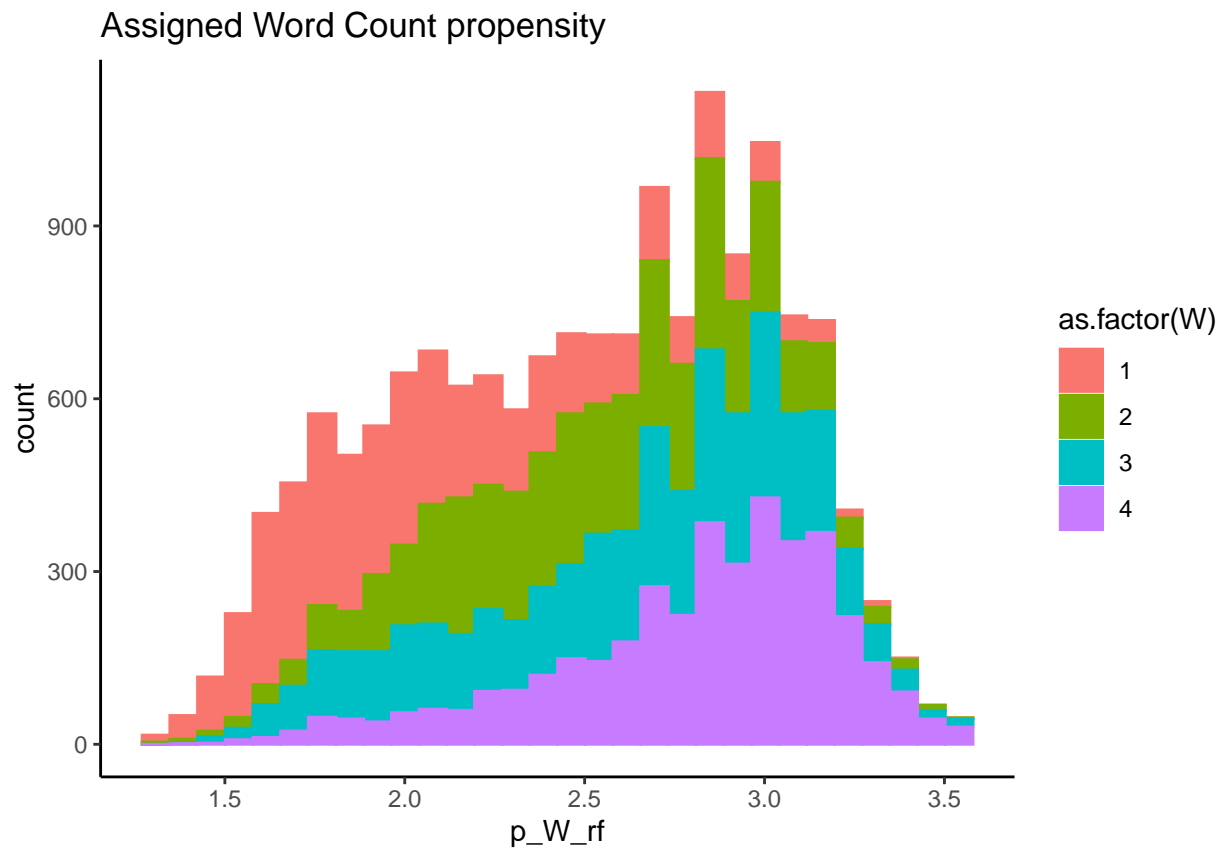
1	User-Session Level Analysis	9
1.1	Effect of Higher than Average Estimated Reading Time on SOD Completion	10
1.2	Effect of Higher than Average Estimated Word Count on SOD Completion .	11
1.3	Effect of Higher than Average Estimated Reading Time SOD on Time to Next Session	12
1.4	Effect of Higher than Average Word Count SOD on Time to Next Session . .	13
2	Number of Words Read	14
3	Word Count CATE on Time to Next Session	15
4	User-level Analysis (Averaged over User-Sessions)	15
4.1	User Average Effect of Higher than Average Estimated Reading Time on SOD Completion	16
4.2	User Average Effect of Higher than Average Estimated Word Count on SOD Completion	17

4.3	User Average Effect of Higher than Average Estimated Reading Time SOD on Time to Next Session	18
4.4	User Average Effect of Higher than Average Word Count SOD on Time to Next Session	19
5	Optimal Policy Trees	20
6	Introduction	21
7	Data Description	22
7.1	User Covariates:	22

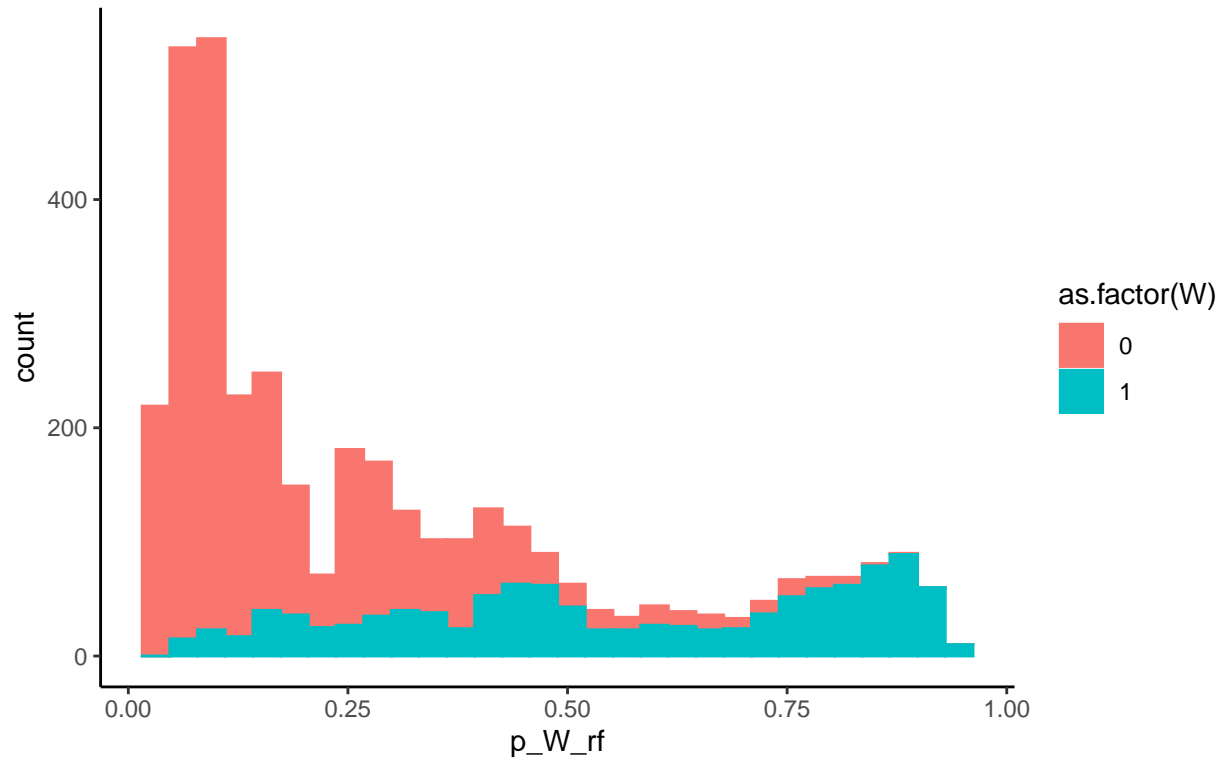




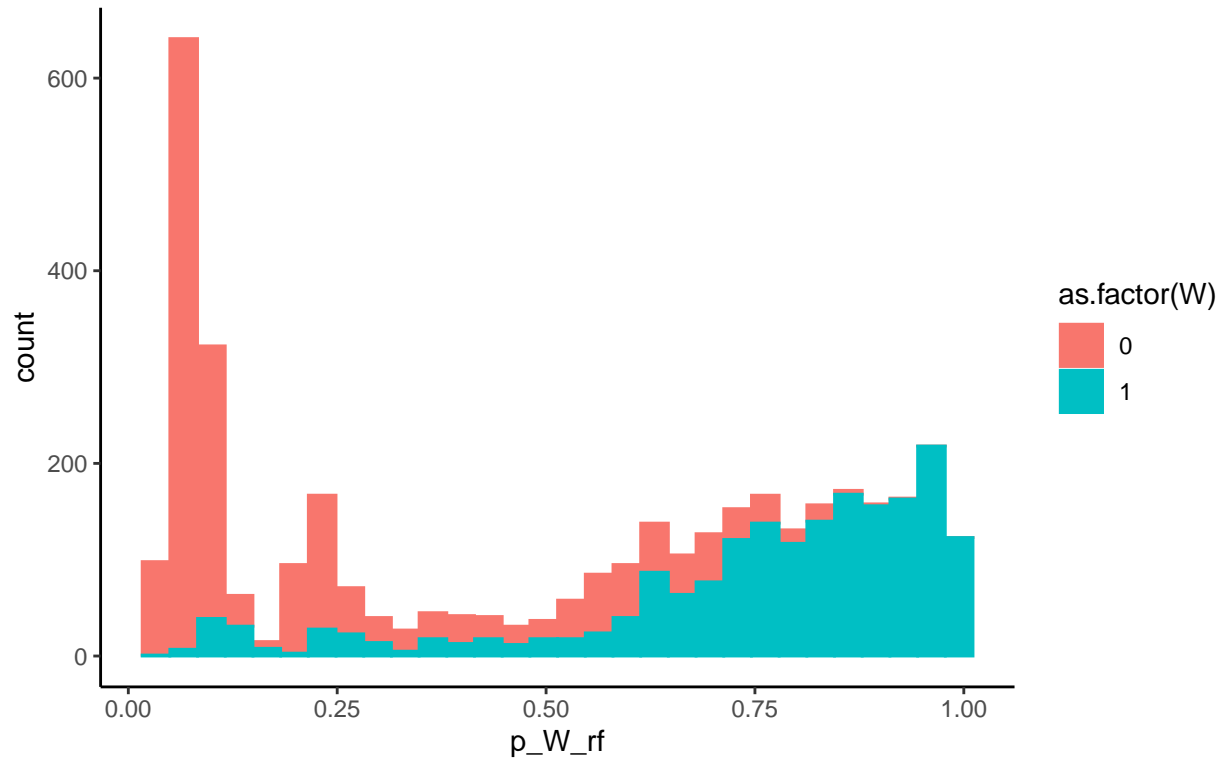


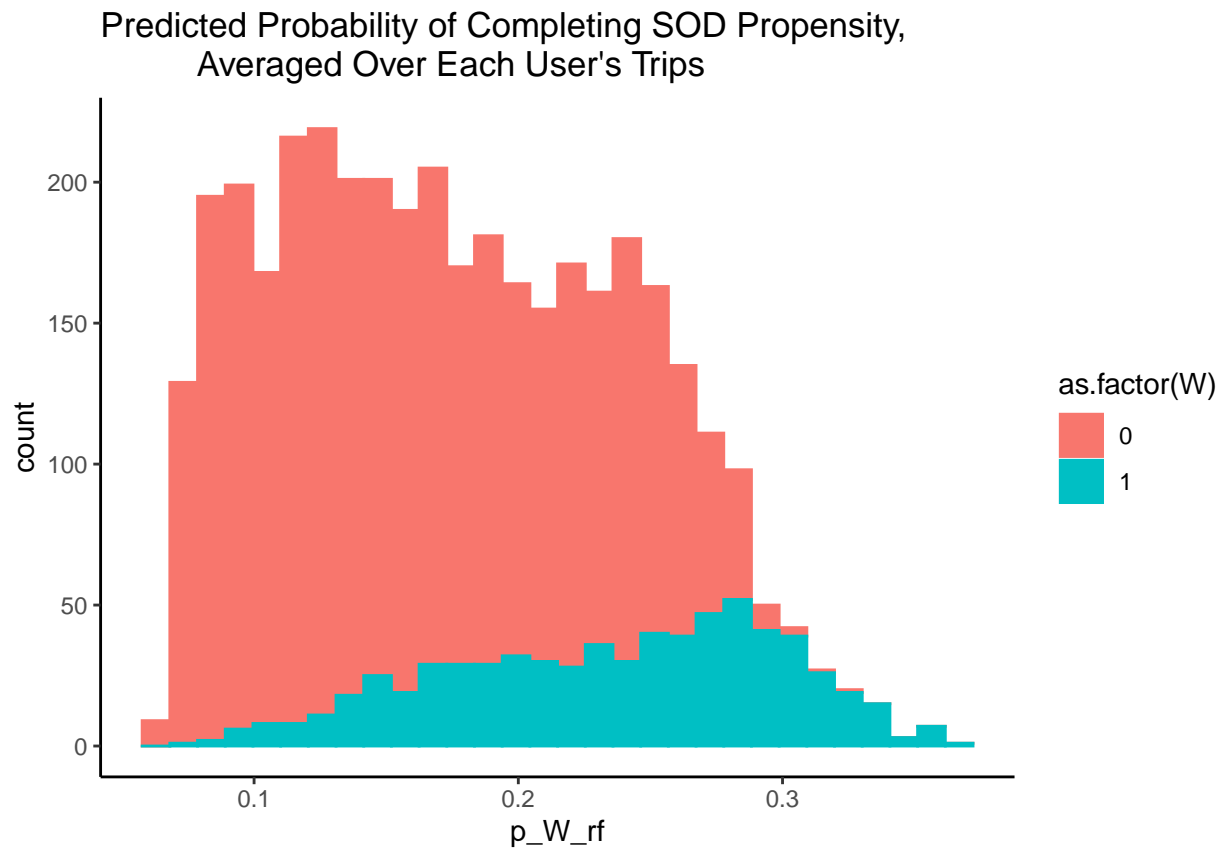


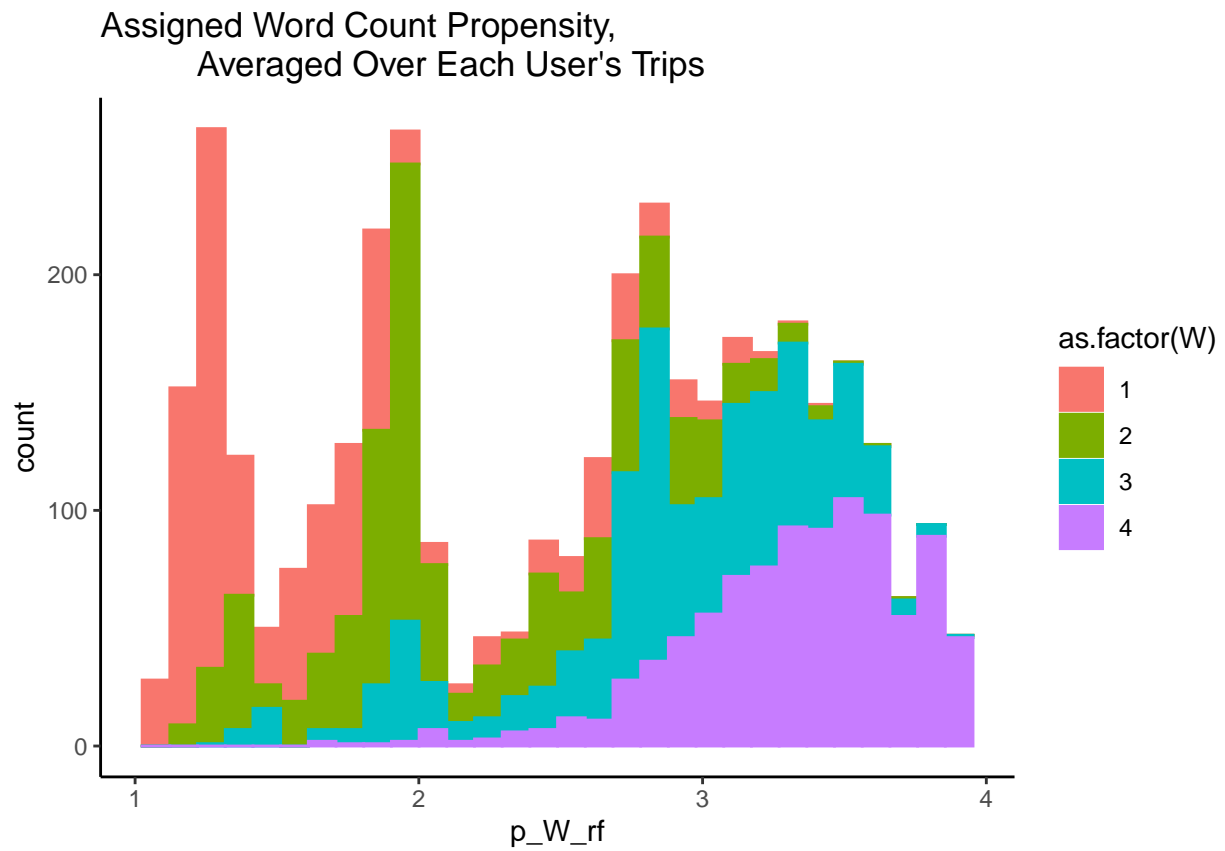
Above Median SOD Reading Time Propensity,
Averaged Over Each User's Trips



Above Median SOD Word Count Propensity,
Averaged Over Each User's Trips







1 User-Session Level Analysis

1.1 Effect of Higher than Average Estimated Reading Time on SOD Completion

[1] 801 [1] 12200

	ATE	lower_ci	upper_ci	ci_length
RCT_gold_standard	-0.103	-0.169	-0.038	0.131
logistic_propensity_weighted_regression	-0.034	-0.109	0.041	0.149
IPW_logistic	0.051	-0.132	0.234	0.366
AIPW_linear_plus_logistic	-0.050	-0.122	0.023	0.145
IPW_forest	-0.190	-0.370	-0.009	0.361
AIPW_ate_causal_forest	-0.050	-0.115	0.015	0.130
AIPW_linear_plus_forest	-0.050	-0.116	0.017	0.133

1.2 Effect of Higher than Average Estimated Word Count on SOD Completion

[1] 8002 [1] 8005

	ATE	lower_ci	upper_ci	ci_length
RCT_gold_standard	-0.087	-0.115	-0.059	0.056
logistic_propensity_weighted_regression	-0.040	-0.070	-0.010	0.060
IPW_logistic	-0.010	-0.081	0.062	0.143
AIPW_linear_plus_logistic	-0.040	-0.069	-0.012	0.057
IPW_forest	-0.040	-0.110	0.029	0.139
AIPW_ate_causal_forest	-0.029	-0.058	0.000	0.058
AIPW_linear_plus_forest	-0.033	-0.060	-0.005	0.055

1.3 Effect of Higher than Average Estimated Reading Time SOD on Time to Next Session

[1] 801 [1] 12186

	ATE	lower_ci	upper_ci	ci_length
RCT_gold_standard	-0.022	-0.190	0.147	0.337
logistic_propensity_weighted_regression	0.137	-0.097	0.370	0.466
IPW_logistic	0.224	-0.083	0.531	0.614
AIPW_linear_plus_logistic	0.103	-0.119	0.325	0.444
IPW_forest	-0.065	-0.302	0.173	0.474
AIPW_ate_causal_forest	0.127	-0.045	0.298	0.343
AIPW_linear_plus_forest	0.077	-0.091	0.245	0.337

1.4 Effect of Higher than Average Word Count SOD on Time to Next Session

[1] 8002 [1] 8005

	ATE	lower_ci	upper_ci	ci_length
RCT_gold_standard	-0.012	-0.085	0.061	0.146
logistic_propensity_weighted_regression	0.100	0.019	0.182	0.162
IPW_logistic	0.129	0.027	0.230	0.204
AIPW_linear_plus_logistic	0.094	0.015	0.172	0.157
IPW_forest	0.082	-0.009	0.174	0.183
AIPW_ate_causal_forest	0.066	-0.014	0.146	0.160
AIPW_linear_plus_forest	0.091	0.021	0.161	0.140

2 Number of Words Read

[1] 8002 [1] 8005

	ATE	lower_ci	upper_ci	ci_length
RCT_gold_standard	-0.012	-0.085	0.061	0.146
logistic_propensity_weighted_regression	0.100	0.019	0.182	0.162
IPW_logistic	0.129	0.027	0.230	0.204
AIPW_linear_plus_logistic	0.094	0.015	0.172	0.157
IPW_forest	0.082	-0.009	0.174	0.183
AIPW_ate_causal_forest	0.057	-0.021	0.135	0.155
AIPW_linear_plus_forest	0.091	0.021	0.161	0.140

We now estimate the CATE, and use it to construct quartiles. We then report the ATE as estimated with AIPW from our causal forest estimate across quartiles.

3 Word Count CATE on Time to Next Session

We now estimate the CATE, and use it to construct quartiles of user-sessions. We then report the ATE as estimated with AIPW from our causal forest estimate across quartiles.

ntile	avg_cf_cate	aipw_estimate	aipw_std.err
1	-0.139	-0.123	0.076
2	-0.006	0.021	0.032
3	0.054	0.011	0.051
4	0.366	0.382	0.111

4 User-level Analysis (Averaged over User-Sessions)

4.1 User Average Effect of Higher than Average Estimated Reading Time on SOD Completion

[1] 1114 [1] 2444

	ATE	lower_ci	upper_ci	ci_length
RCT_gold_standard	-0.129	-0.176	-0.082	0.095
logistic_propensity_weighted_regression	-0.073	-0.139	-0.007	0.131
IPW_logistic	-0.097	-0.194	0.000	0.193
AIPW_linear_plus_logistic	-0.080	-0.139	-0.022	0.116
IPW_forest	-0.200	-0.278	-0.121	0.157
AIPW_ate_causal_forest	-0.101	-0.180	-0.022	0.158
AIPW_linear_plus_forest	-0.078	-0.127	-0.029	0.099

4.2 User Average Effect of Higher than Average Estimated Word Count on SOD Completion

[1] 1716 [1] 1575

	ATE	lower_ci	upper_ci	ci_length
RCT_gold_standard	-0.124	-0.175	-0.073	0.102
logistic_propensity_weighted_regression	-0.112	-0.216	-0.008	0.207
IPW_logistic	-0.216	-0.366	-0.065	0.301
AIPW_linear_plus_logistic	-0.088	-0.182	0.007	0.189
IPW_forest	0.009	-0.077	0.095	0.172
AIPW_ate_causal_forest	-0.115	-0.206	-0.023	0.183
AIPW_linear_plus_forest	-0.076	-0.132	-0.020	0.112

4.3 User Average Effect of Higher than Average Estimated Reading Time SOD on Time to Next Session

[1] 1114 [1] 2444

	ATE	lower_ci	upper_ci	ci_length
RCT_gold_standard	-1.107	-1.309	-0.905	0.404
logistic_propensity_weighted_regression	-0.552	-0.807	-0.296	0.511
IPW_logistic	-0.639	-1.022	-0.256	0.767
AIPW_linear_plus_logistic	-0.558	-0.787	-0.328	0.458
IPW_forest	-0.772	-1.126	-0.417	0.709
AIPW_ate_causal_forest	-0.030	-0.382	0.322	0.704
AIPW_linear_plus_forest	-0.376	-0.612	-0.141	0.471

4.4 User Average Effect of Higher than Average Word Count SOD on Time to Next Session

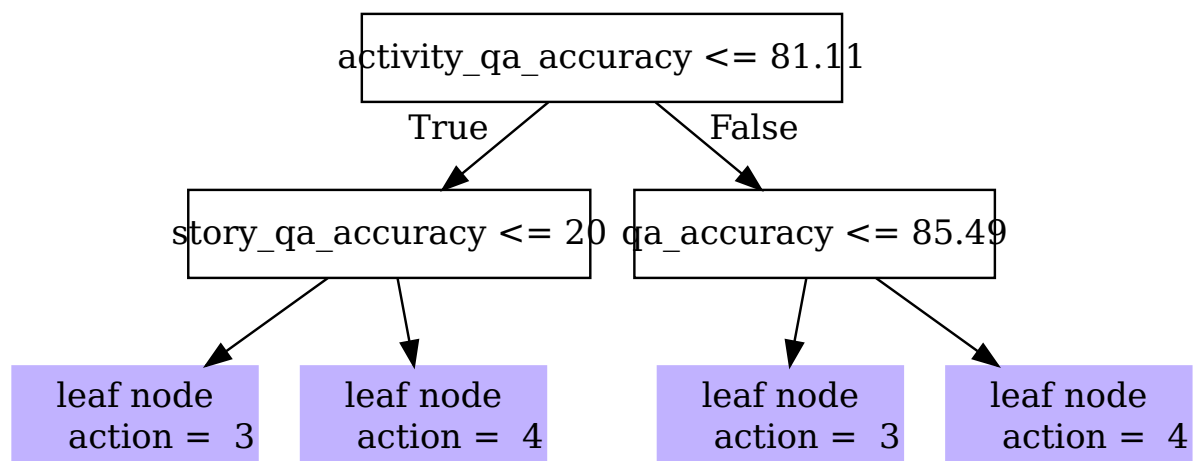
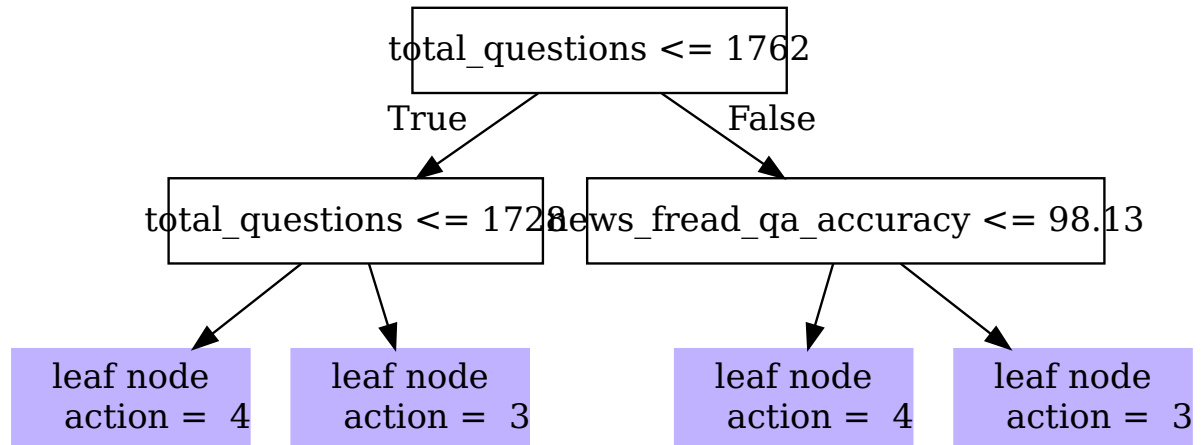
[1] 1710 [1] 1575

	ATE	lower_ci	upper_ci	ci_length
RCT_gold_standard	0.079	-0.163	0.321	0.484
logistic_propensity_weighted_regression	-0.038	-0.574	0.498	1.071
IPW_logistic	-0.478	-1.211	0.255	1.466
AIPW_linear_plus_logistic	0.058	-0.475	0.590	1.065
IPW_forest	0.522	0.024	1.020	0.996
AIPW_ate_causal_forest	0.837	0.222	1.451	1.228
AIPW_linear_plus_forest	0.173	-0.187	0.534	0.721

We now estimate the CATE, and use it to construct quartiles. We then report the ATE as estimated with AIPW from our causal forest estimate across quartiles.

ntile	avg_cf_cate	aipw_estimate	aipw_std.err
1	-0.820	-1.571	0.508
2	-0.181	-0.160	0.458
3	0.232	-0.202	0.394
4	0.983	1.886	0.509

5 Optimal Policy Trees



6 Introduction

Improvements to childhood literacy have been linked to numerous positive outcomes, including economic and social benefits (cite). In this paper, we use data from a mobile application, aimed at improving childhood reading outcomes. School going children from junior kindergarten until grade 3 use the app to read stories, among other things.

This work takes advantage of the application’s “Story of the Day” (hereafter SOD) feature. Stories of the Day are featured prominently on the app, and users read the Story of the Day on approximately 33% of days they use the app. Our analysis focuses entirely on these stories. Several Stories of the Day are available to be assigned to users on each day, and vary primarily by estimated reading time and word count. The assignment of story of the day is generic and not personalised, and different stories are shown everyday. As a result, if we consider a user opening the app as an exogenous random decision on a given day, since the stories shown to students are different each day, this gives us exogenous treatments for length of stories shown to students in terms of reading time and number of words in story. We stress that this is a reasonable assumption since students cannot observe the length of a story before logging in. We measure the effect of this treatment on reading outcomes. Our identifying motivation is that longer stories reduce the probability of a child reading a story. Even if this effect is true on average, this does not mean that longer stories have negative effects on all users. As such, we examine CATEs across a variety of groups in Section @ref(sod_length_cate).

Finally, in Section @ref(optimal_policy) we attempt to maximize aggregate reading time by identifying the optimal policy, and summarize possible gains from targeted assignment of Stories of the Day by length.

7 Data Description

7.1 User Covariates:

In our dataset, we have a bunch of covariates describing users. These include age, grade, statistics about usage such as books read, experience points gained on the app while using it, total time spent on the app, covariates about how well a child answered questions related to their readings, and their reading interests ## Treatment Definitions: We use the following treatments for analysis: Suggested Reading Time for a given story: The app includes suggested reading times for a story in one of five choices. We examine only stories where the estimated reading time was either 7.5 or 12.5 minutes, as estimated reading time is not continuous. These two values of estimated reading time account for 95% of all user-trips. Number of words for a given story: We parsed the stories shown on the app to get the number of words in each story, and we use this as a treatment variable. We divide observations into those where the number of words is below and above the median, giving us a treatment and a control. We also do another analysis with multiple treatments where the treatment is characterised by the quantile in which its number of words falls; giving us 4 treatment conditions We recall that as the assignment of the Story of the Day is at-random, so is the assignment of word count and estimated reading time. Note: We examine only users-trips where the user started reading the Story of the Day, as otherwise the user would have no estimate of the length or time required to read the story. # Outcomes We examine three outcomes in our analysis: whether the user finished reading their assigned Story of the Day, the length of time until their next session, and the estimated number of words they read.