final_project_v2.R

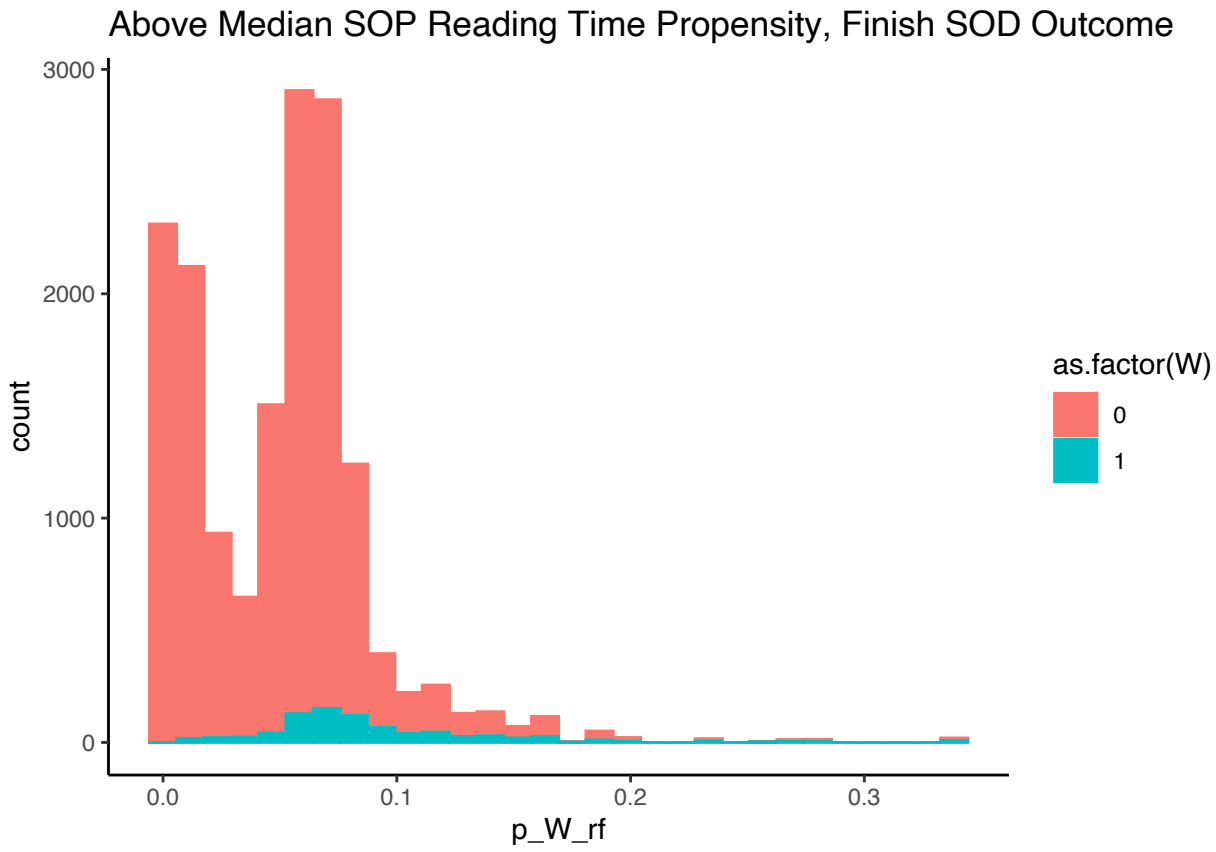
mlinegar

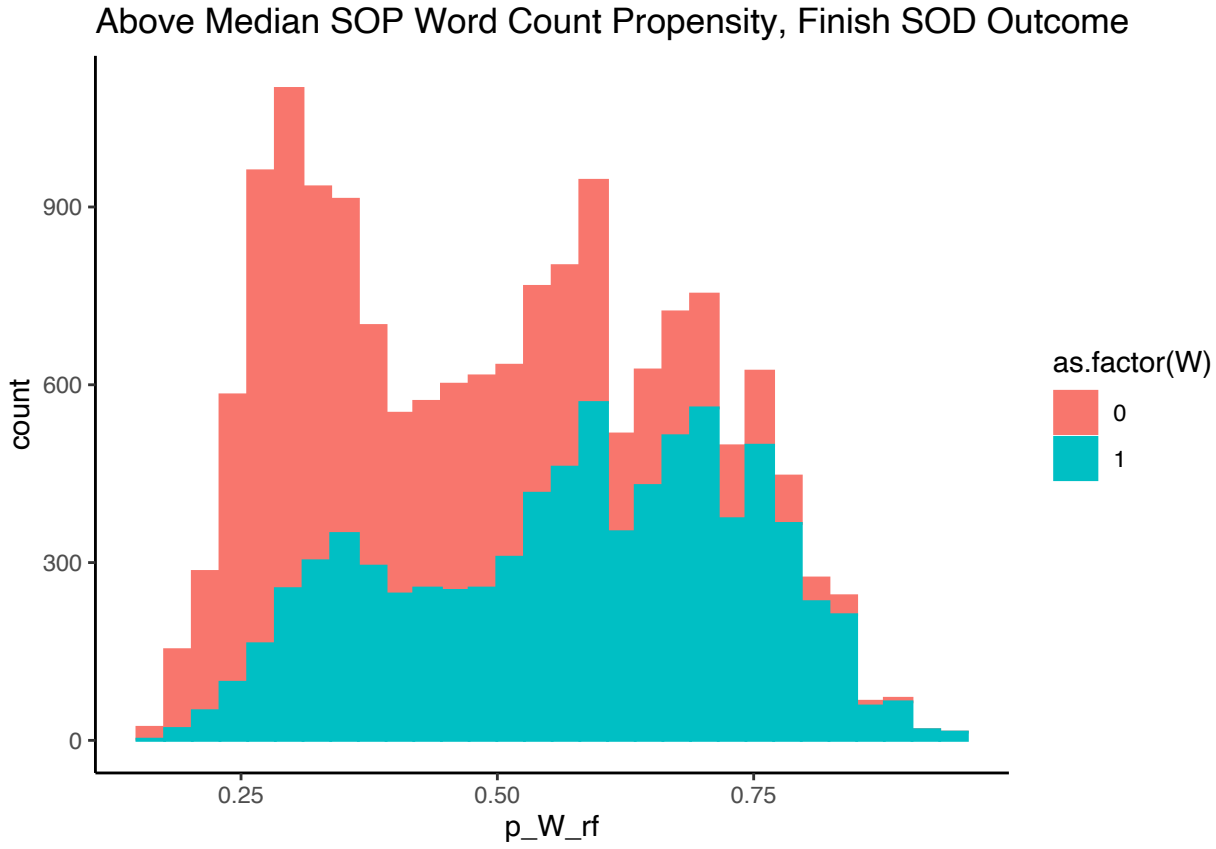
2020-06-05

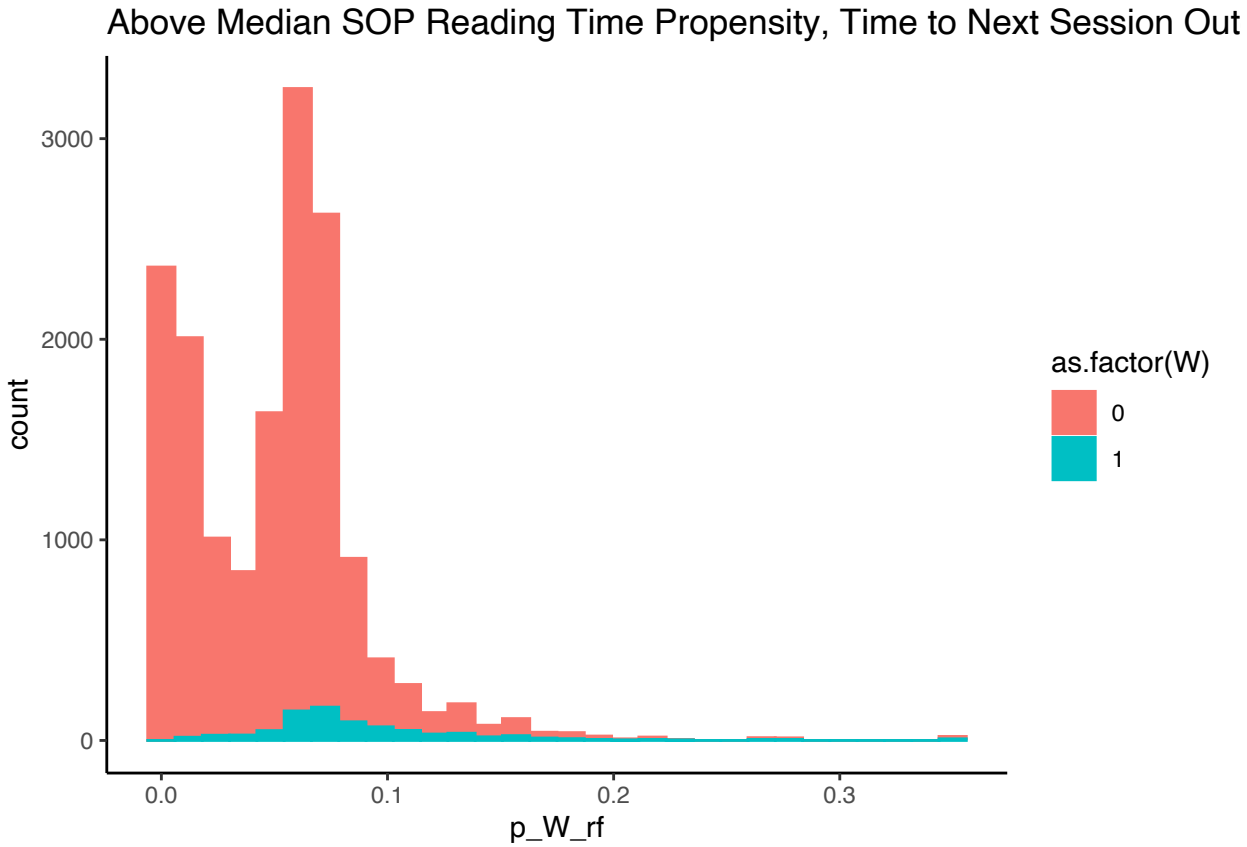
Contents

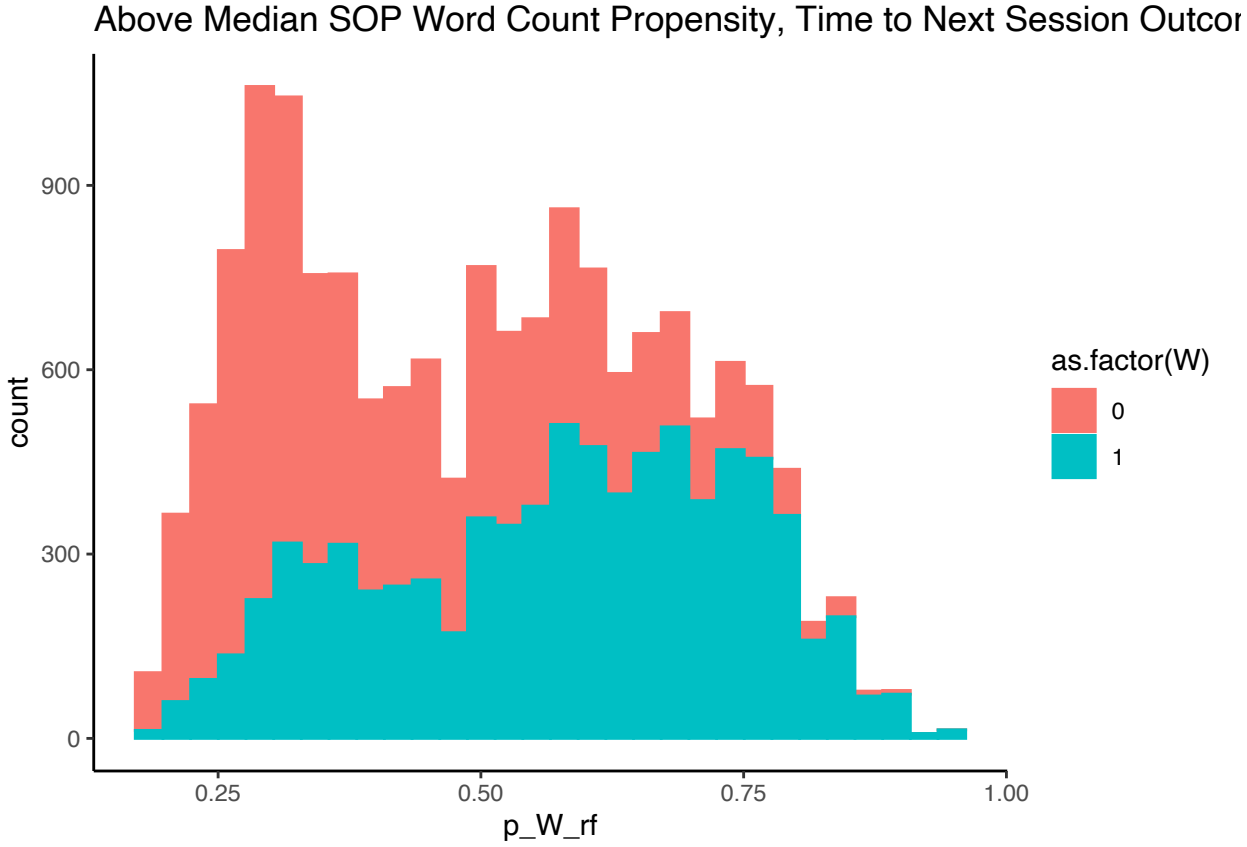

```
#### SETUP ####
```

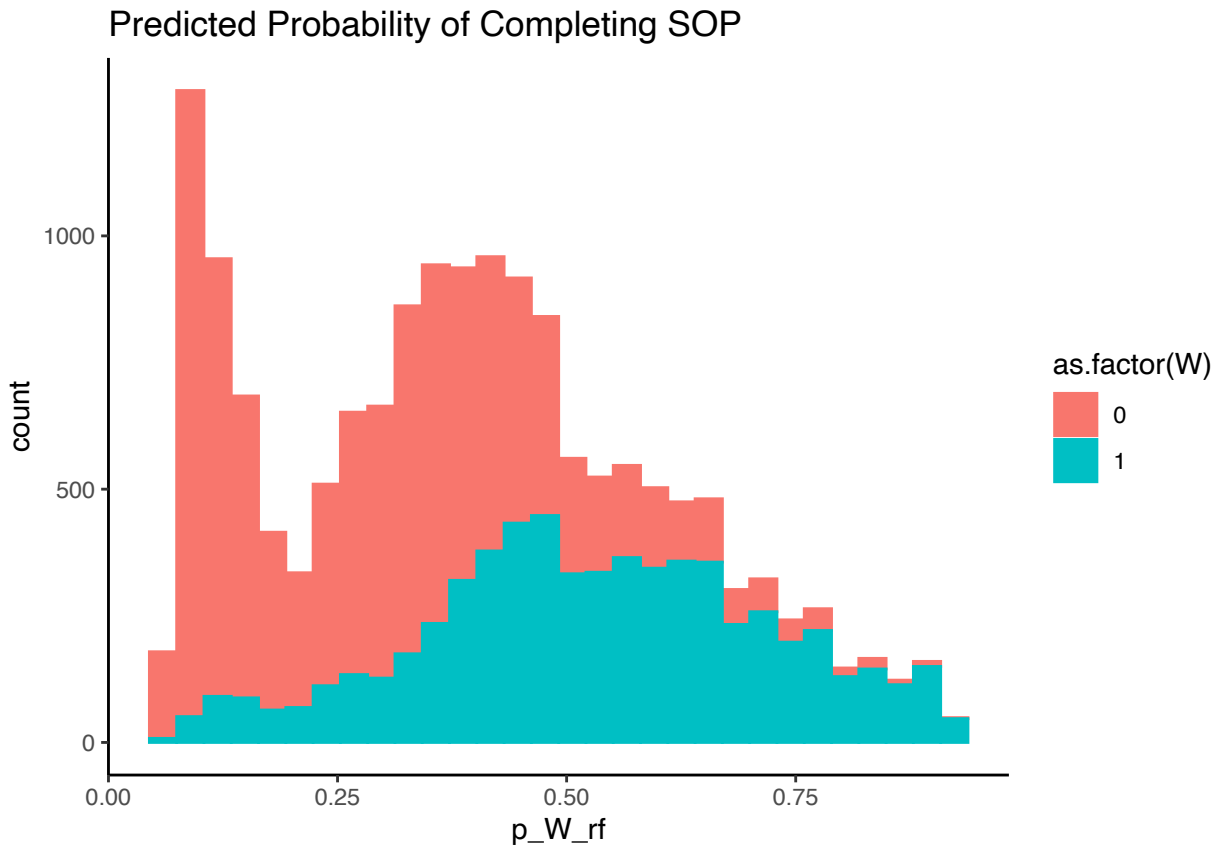
```
## Skipping install of 'amlinear' from a github remote, the SHA1 (83ee1d18) has not changed since last i
##   Use `force = TRUE` to force installation

## Skipping install of 'causalTree' from a github remote, the SHA1 (48604762) has not changed since last
##   Use `force = TRUE` to force installation

## Skipping install of 'sufrep' from a github remote, the SHA1 (317be9e7) has not changed since last ins
##   Use `force = TRUE` to force installation

## Skipping install of 'policytree' from a github remote, the SHA1 (3e42a771) has not changed since last
##   Use `force = TRUE` to force installation
```

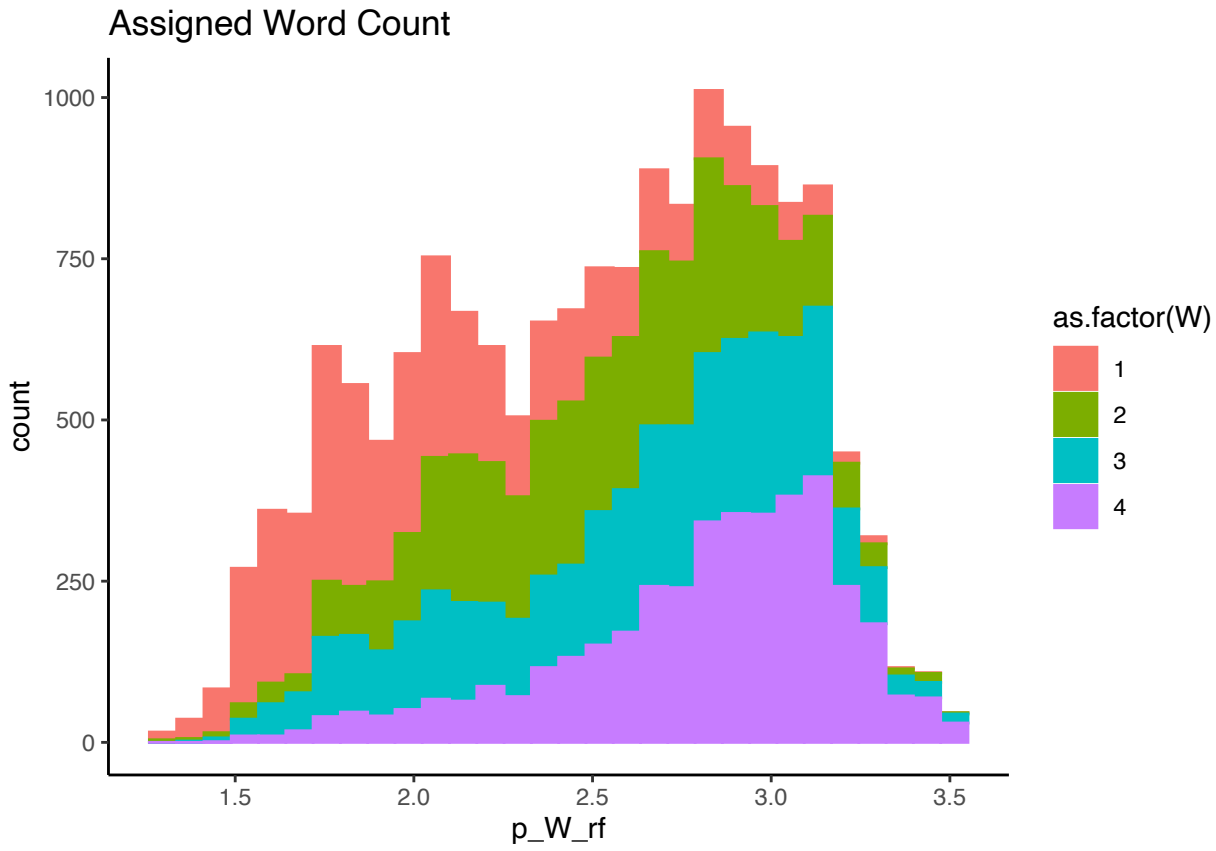

Above Median SOP Reading Time Propensity, Finish SOD Outcome
count
3000
2000
1000
0
0.0
0.1
0.2
0.3
p_W_rf
as.factor(W)
0
1

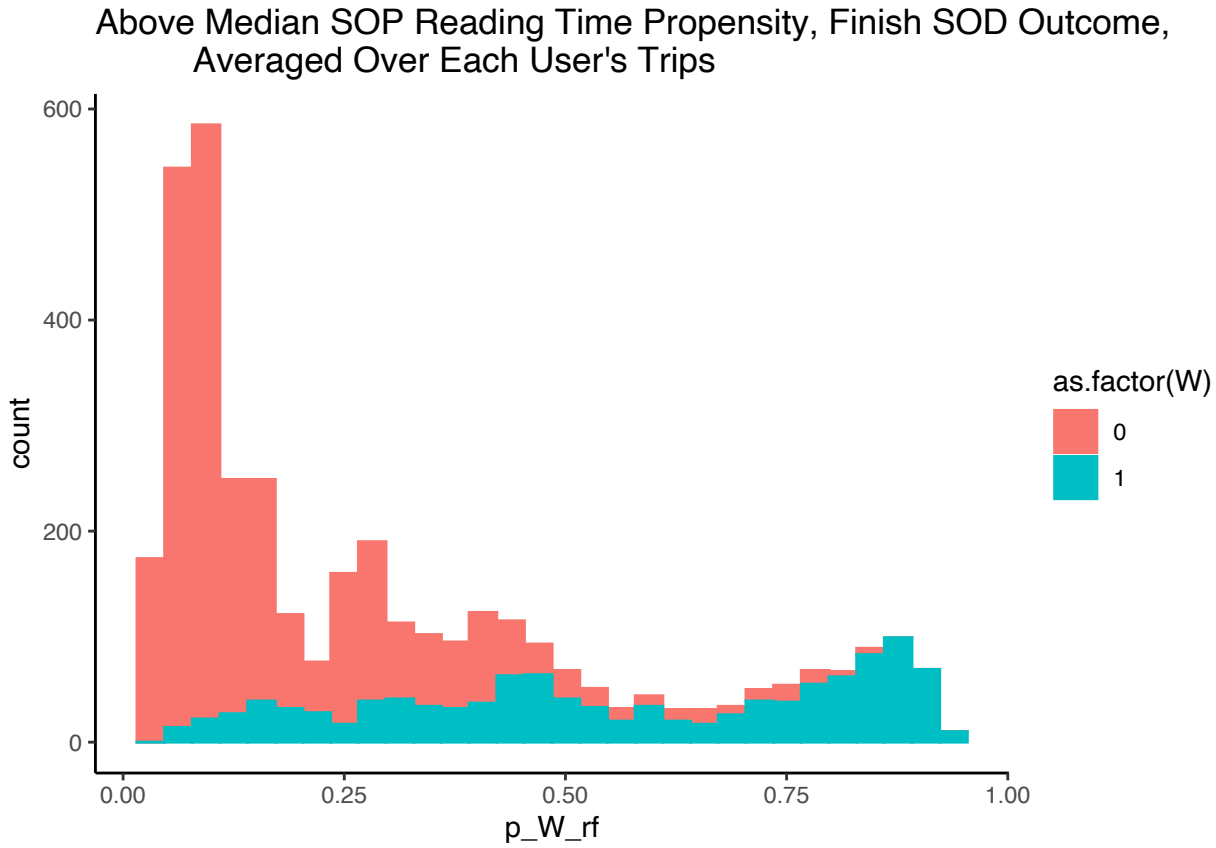

Above Median SOP Word Count Propensity, Finish SOD Outcome
count
900
600
300
0
0.25
0.50
0.75
p_W_rf
as.factor(W)
0
1

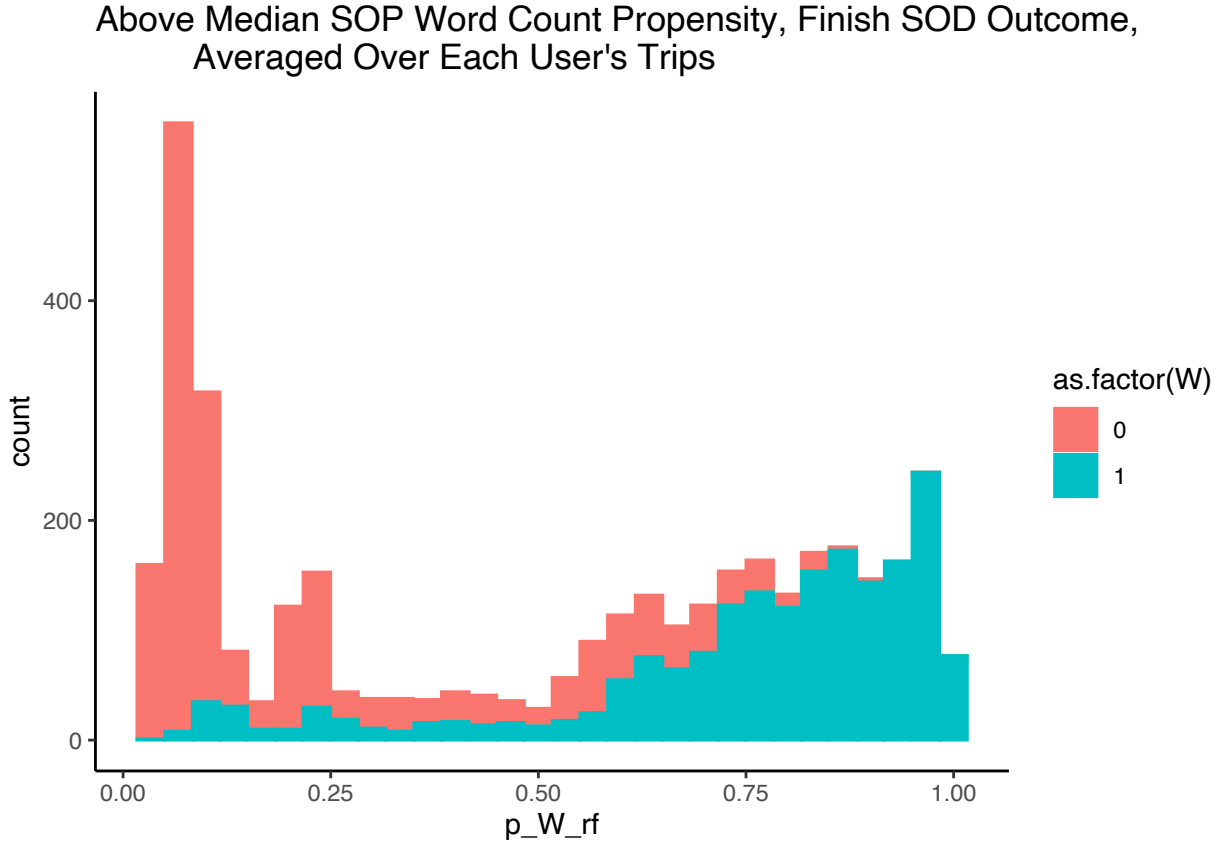

Above Median SOP Reading Time Propensity, Time to Next Session Out
count
3000
2000
1000
0
0.0
0.1
0.2
0.3
p_W_rf
as.factor(W)
0
1

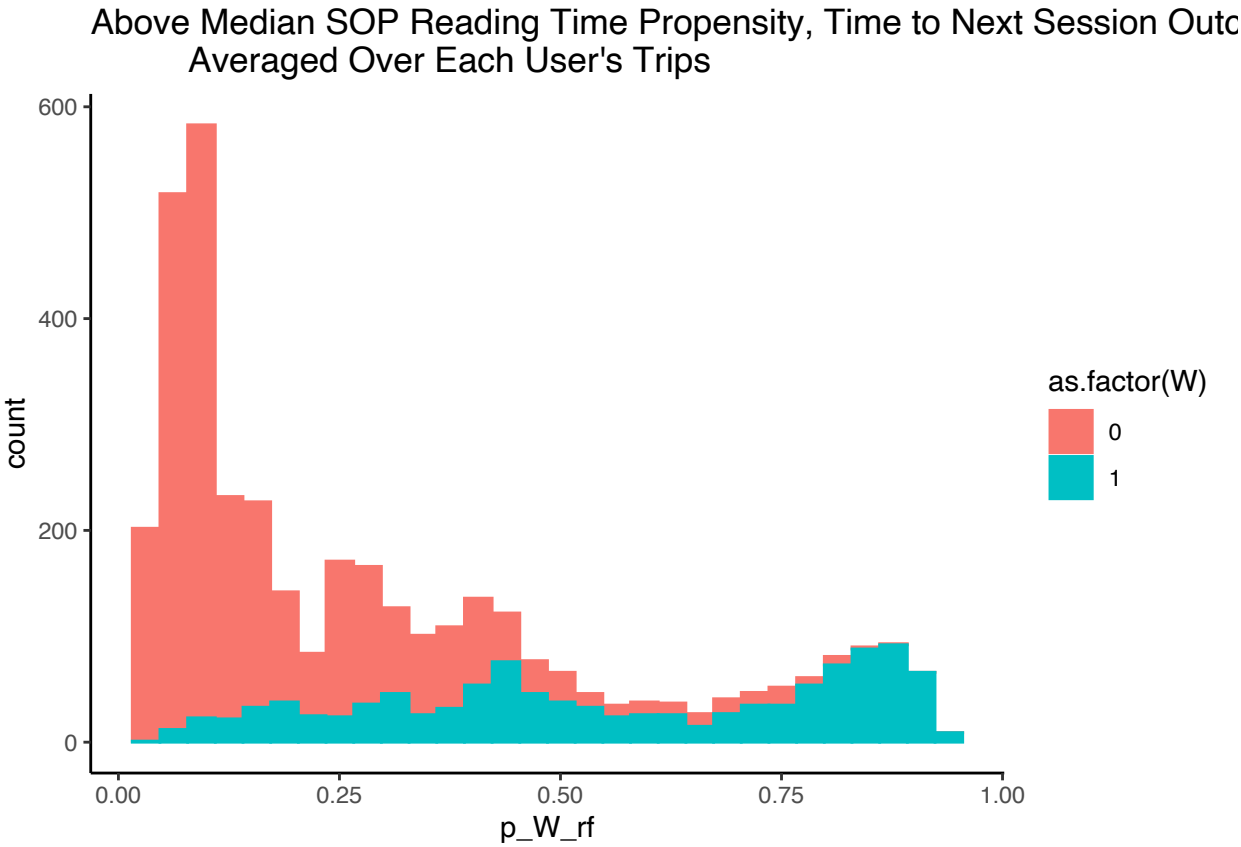

Above Median SOP Word Count Propensity, Time to Next Session Outcor
count
900
600
300
0
0.25
0.50
0.75
1.00
p_W_rf
as.factor(W)
0
1

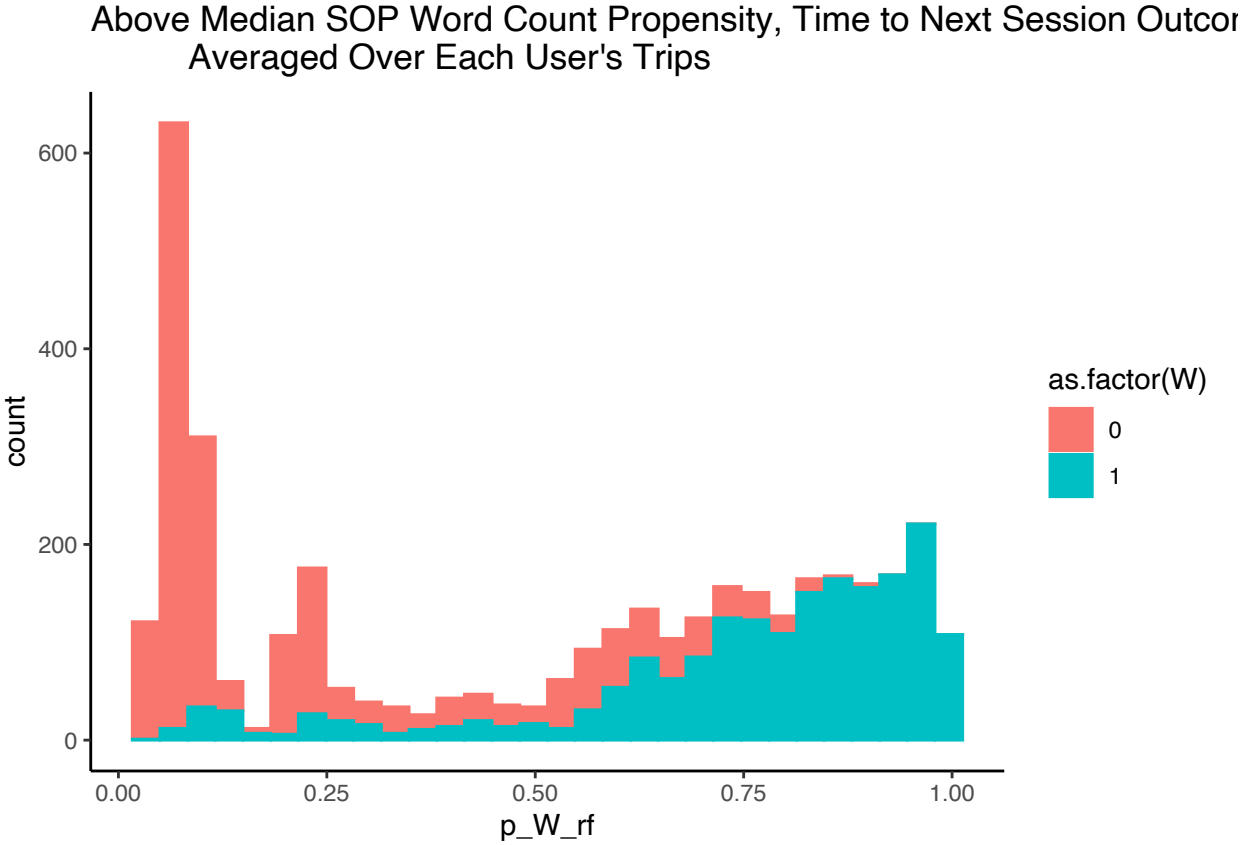

Predicted Probability of Completing SOP
count
1000
500
0
0.00
0.25
0.50
0.75
p_W_rf
as.factor(W)
0
1

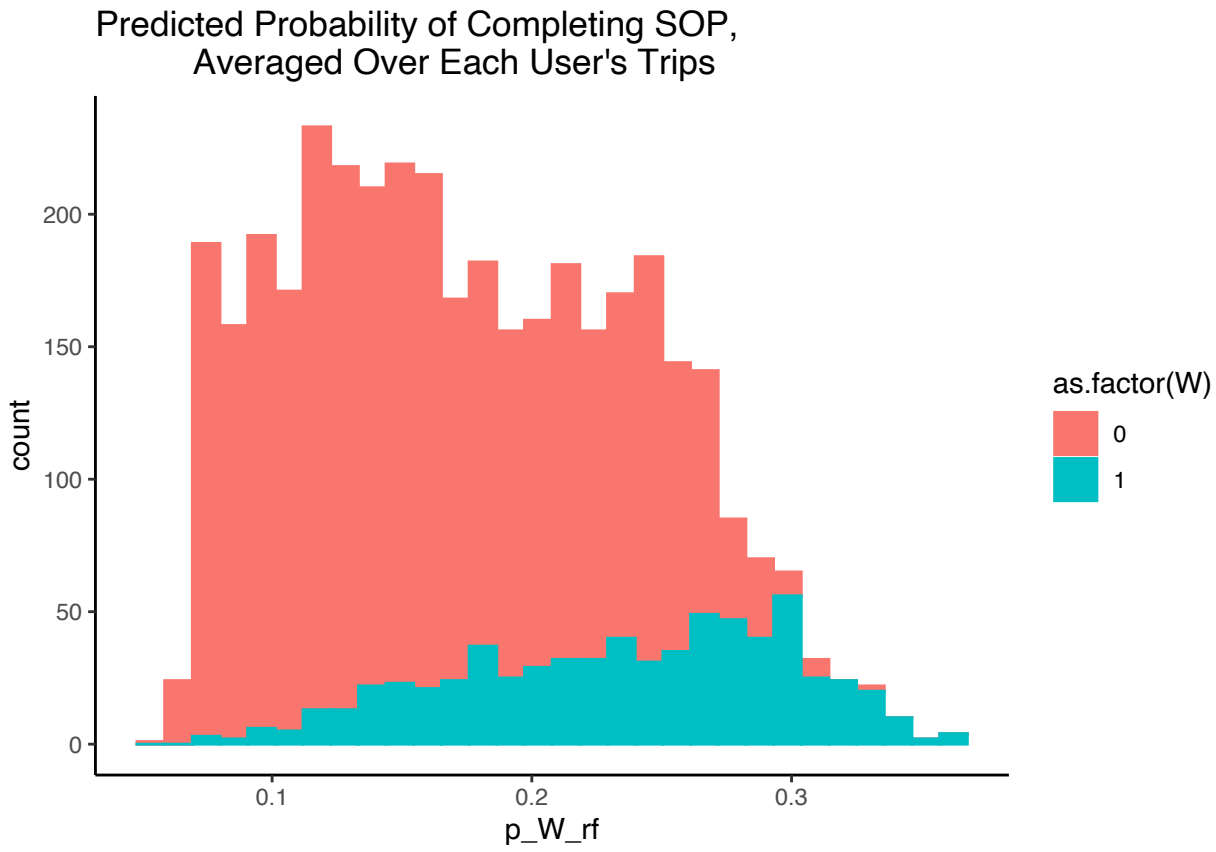

Assigned Word Count
count
1000
750
500
250
0
1.5
2.0
2.5
3.0
3.5
p_W_rf
as.factor(W)
1
2
3
4

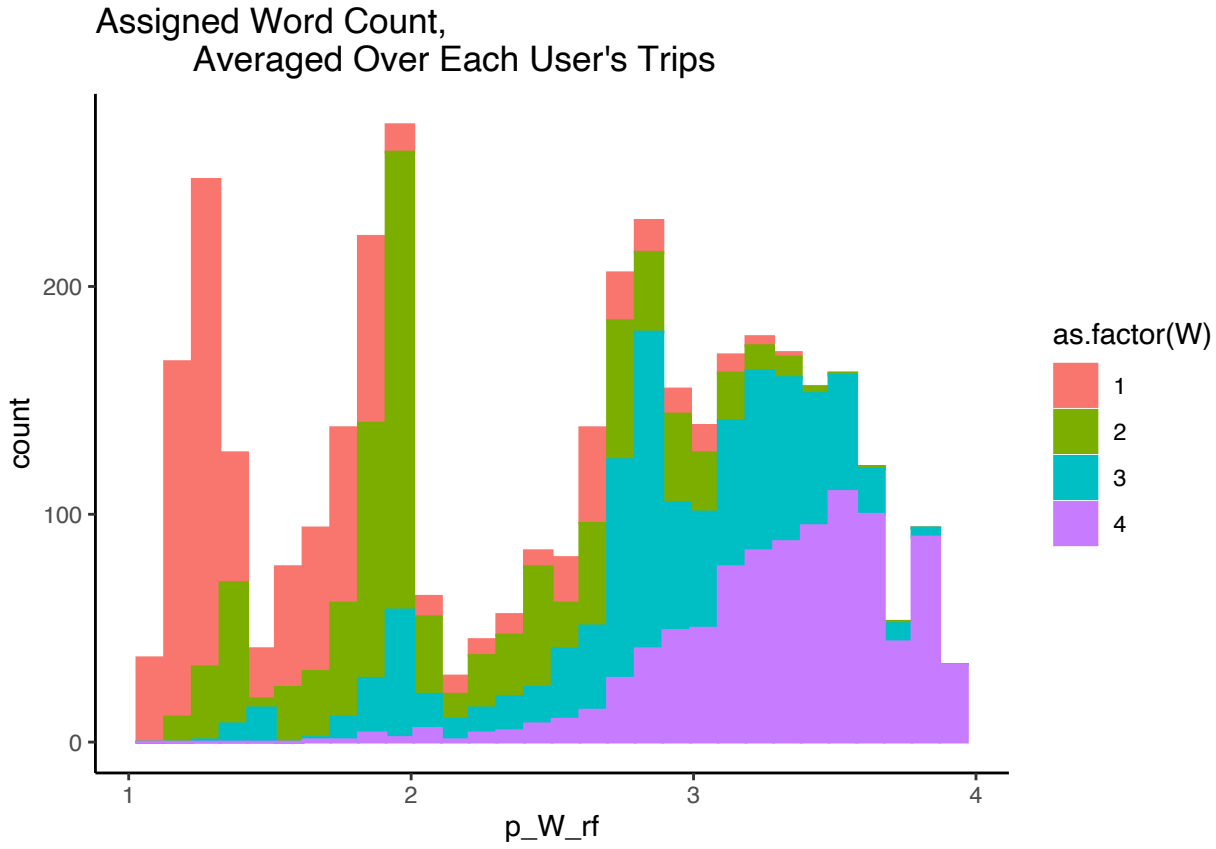

Above Median SOP Reading Time Propensity, Finish SOD Outcome,
Averaged Over Each User's Trips
count
600
400
200
0
0.00
0.25
0.50
0.75
1.00
p_W_rf
as.factor(W)
0
1


Above Median SOP Word Count Propensity, Finish SOD Outcome,
Averaged Over Each User's Trips
count
400
200
0
0.00
0.25
0.50
0.75
1.00
p_W_rf
as.factor(W)
0
1


Above Median SOP Reading Time Propensity, Time to Next Session Outc
Averaged Over Each User's Trips
600
400
200
0
count
0.00
0.25
0.50
0.75
1.00
p_W_rf
as.factor(W)
0
1


Above Median SOP Word Count Propensity, Time to Next Session Outcor
Averaged Over Each User's Trips
600
400
200
0
count
0.00
0.25
0.50
0.75
1.00
p_W_rf
as.factor(W)
0
1


Predicted Probability of Completing SOP,
Averaged Over Each User's Trips
count
200
150
100
50
0
0.1
0.2
0.3
p_W_rf
as.factor(W)
0
1


Assigned Word Count,
Averaged Over Each User's Trips
count
200
100
0
1
2
3
4
p_W_rf
as.factor(W)
1
2
3
4

1 Rate at Which Users Finish Stories of the Day by Length

Table 1:

	ATE	lower_ci	upper_ci	ci_length
RCT_gold_standard	-0.103	-0.134	-0.072	0.062
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	-0.196	-0.377	-0.016	0.362
AIPW_ate_causal_forest	-0.056	-0.120	0.008	0.127
AIPW_linear_plus_forest	-0.047	-0.113	0.020	0.133

2 Rate at Which Users Finish Stories of the Day by Word Count

Table 2:

	ATE	lower_ci	upper_ci	ci_length
RCT_gold_standard	-0.087	-0.118	-0.056	0.062
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	-0.038	-0.107	0.032	0.139
AIPW_ate_causal_forest	-0.030	-0.059	-0.001	0.058
AIPW_linear_plus_forest	-0.032	-0.060	-0.005	0.055

3 Rate at Which Users Finish Stories of the Day by Length

Table 3:

	ATE	lower_ci	upper_ci	ci_length
RCT_gold_standard	-0.022	-0.104	0.059	0.163
logistic_propensity_weighted_regression	0.137	-0.096	0.371	0.467
IPW_logistic	0.224	-0.083	0.532	0.615
AIPW_linear_plus_logistic	0.103	-0.119	0.326	0.445
IPW_forest	-0.097	-0.328	0.134	0.462
AIPW_ate_causal_forest	0.125	-0.055	0.305	0.360
AIPW_linear_plus_forest	0.063	-0.102	0.228	0.330

4 Rate at Which Users Finish Stories of the Day by Word Count

[1] 8002 [1] 8005

Table 4:

	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	-0.010	0.173	0.183
AIPW_ate_causal_forest	0.068	-0.010	0.146	0.157
AIPW_linear_plus_forest	0.091	0.021	0.161	0.140

5 Number of Words Read

[1] 8002 [1] 8005

Table 5:

	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	-0.010	0.173	0.183
AIPW_ate_causal_forest	0.068	-0.011	0.146	0.158
AIPW_linear_plus_forest	0.091	0.021	0.161	0.140

6 Effect of Finishing SOD on Time to Next Session

Table 6:

	ATE	lower_ci	upper_ci	ci_length
RCT_gold_standard	0.043	-0.037	0.123	0.160
logistic_propensity_weighted_regression	-0.097	-0.176	-0.018	0.157
IPW_logistic	-0.129	-0.227	-0.031	0.197
AIPW_linear_plus_logistic	-0.116	-0.192	-0.039	0.153
IPW_forest	-0.091	-0.182	0	0.182
AIPW_ate_causal_forest	-0.020	-0.109	0.069	0.178
AIPW_linear_plus_forest	-0.037	-0.107	0.032	0.139

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.

Table 7:

	ntile	avg_cf_cate	aipw_estimate	aipw_std.err
1	1	-0.276	0.060	0.127
2	2	-0.039	-0.090	0.056
3	3	0.018	0.067	0.055
4	4	0.157	-0.088	0.065

7 Rate at Which Users on Their Average Trip Finish Stories of the Day by Length

[1] 1114 [1] 2444

Table 8:

	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	0.193
AIPW_linear_plus_logistic	-0.080	-0.139	-0.022	0.116
IPW_forest	-0.203	-0.280	-0.125	0.155
AIPW_ate_causal_forest	-0.111	-0.188	-0.034	0.154
AIPW_linear_plus_forest	-0.080	-0.129	-0.031	0.097

8 Rate at Which Users on Their Average Trip Finish Stories of the Day by Word Count

[1] 1712 [1] 1575

Table 9:

	ATE	lower_ci	upper_ci	ci_length
RCT_gold_standard	-0.124	-0.175	-0.074	0.102
logistic_propensity_weighted_regression	-0.112	-0.216	-0.008	0.208
IPW_logistic	-0.217	-0.368	-0.067	0.302
AIPW_linear_plus_logistic	-0.088	-0.183	0.007	0.190
IPW_forest	0.013	-0.072	0.099	0.171
AIPW_ate_causal_forest	-0.104	-0.192	-0.016	0.175
AIPW_linear_plus_forest	-0.074	-0.130	-0.019	0.111

9 Rate at Which Users on Their Average Trip Finish Stories of the Day by Length

[1] 1114 [1] 2444

Table 10:

	ATE	lower_ci	upper_ci	ci_length
RCT_gold_standard	-1.110	-1.310	-0.905	0.404
logistic_propensity_weighted_regression	-0.552	-0.807	-0.296	0.511
IPW_logistic	-0.639	-1.020	-0.256	0.767
AIPW_linear_plus_logistic	-0.558	-0.787	-0.328	0.458
IPW_forest	-0.804	-1.160	-0.451	0.706
AIPW_ate_causal_forest	-0.011	-0.339	0.316	0.656
AIPW_linear_plus_forest	-0.380	-0.614	-0.146	0.468

10 Rate at Which Users on Their Average Trip Finish Stories of the Day by Word Count

[1] 1712 [1] 1575

Table 11:

	ATE	lower_ci	upper_ci	ci_length
RCT_gold_standard	0.077	-0.165	0.318	0.484
logistic_propensity_weighted_regression	-0.040	-0.575	0.495	1.070
IPW_logistic	-0.477	-1.210	0.255	1.470
AIPW_linear_plus_logistic	0.055	-0.477	0.588	1.060
IPW_forest	0.582	0.059	1.100	1.040
AIPW_ate_causal_forest	0.854	0.225	1.480	1.260
AIPW_linear_plus_forest	0.192	-0.189	0.572	0.761

11 Effect of Finishing SOD on Time to Next Session

[1] 669 [1] 3010

Table 12:

	ATE	lower_ci	upper_ci	ci_length
RCT_gold_standard	-0.081	-0.391	0.228	0.619
logistic_propensity_weighted_regression	-0.104	-0.457	0.248	0.706
IPW_logistic	-0.187	-0.640	0.267	0.907
AIPW_linear_plus_logistic	-0.223	-0.552	0.107	0.659
IPW_forest	-0.592	-0.962	-0.222	0.740
AIPW_ate_causal_forest	-0.170	-0.475	0.135	0.609
AIPW_linear_plus_forest	-0.217	-0.491	0.056	0.547

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.

Table 13:

	ntile	avg_cf_cate	aipw_estimate	aipw_std.err
1	1	-0.440	0.398	0.379
2	2	-0.267	0.241	0.340
3	3	-0.154	-0.209	0.284
4	4	0.026	-1.120	0.239

12 Optimal Policy

```
## policy_tree object
## Tree depth:  2
## Actions:  1: 0 2: 1
## Variable splits:
```

```
## (1) split_variable: reading_qa_accuracy  split_value: 72.88
##   (2) split_variable: reading_qa_accuracy  split_value: 51.43
##     (4) * action: 1
##     (5) * action: 2
##   (3) split_variable: grade_level  split_value: 3
##     (6) * action: 2
##     (7) * action: 1
```

13 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 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.

14 Data Description

14.1 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.