ADD PAGE NUMBERS!!!!!!

(Effectiveness analysis on tutors' interventions for an online math tutoring program)

OK as a working title. Final title should focus on the question or results, e.g. maybe "Jumps in learning due to instructor interventions" or something similar.

Yiwen Zhang, Naifei Pan, Jie Luo
04/21/2021

Abstract

In this study, we address the question of how effective tutors are on students' performances in an online intelligent tutoring system. We use the data from datashop which records the 195 students' learning progress in an online math tutor program. We utilize the AFM model - logistic regression models - in our analyses to examine the effects of tutors' help. Our results show that...

good start; obv more to come!

Is the help/intervention coming from the online tutor or from a human instructor, or ... ?

So far you have not cited the tech appx at all in the paper. Be sure to do so whereever the tech appedix will help clarify what you are doing in the main part of the paper.

1. introduction

In recent years, more and more educational institutions have incorporated new technologies with traditional education to improve the overall learning experiences. With the access to the internet and feasible devices, students can have a quality education wherever and whenever they want. On one hand, online education makes it easier to track students progress as it records their performances in each pre-designed problem with relative knowledge. On the other hand, teachers are able to monitor the class through the screen and decide if additional help is needed for specific students. Thus, the effectiveness of educators' help can also be reflected through the performance of students ~~student's performances~~ who receive those help. In this study, we seek to find out how educators' interventions affect students' learning progress on an online tutoring system. Specifically, we will address the following questions:

Can you add one or two citations from your reference list here?

1. Do these interventions put students on a different learning trajectory, with respect to the specific skills?
2. How can we measure the effect?
3. Do we see struggles before tutor interventions?

please add: for whom are you doing this research (who is the client and what is their organization) and what will they do with the results?

what is the name of the tutoring system? What grade level is targeted?

2. Data

1. what is data shop
2. how is it connected to your client
3. is there a web address or citation?

you have to define what this is

The Out-of-tutor event detection data we use is provided by Datashop. It consists of the 195 students' learning records on an online math tutoring system. There are 3 sub-datasets, organized by transaction, by student step, and by student-problem. We mainly focus on the Transaction dataset and Student Step dataset.

Spend some time here describing workspaces, problems, steps, opportunities, attempts, knowledge components, etc. so reader will have some idea how to read Table 1, Table 2, etc.

Rows in the Transaction dataset are ordered by student and the transaction time. It's worth noting that the tutor intervention event is only presented in this dataset. Detailed information for columns in Transaction dataset is as below.

What does this mean? Are you saying that the only thing in this data set is interventions, or that the only place to find interventions is in this data set, or something else?

The variables I've highlighted in yellow here require further explanation / description.

It would also help to say which variables are going to be important for your analyses (tell the reader)

I don't see anything in Table 1 that indicates interventions... please tell the reader which variable tells about the intervention events.

Table 1. Description of some variables in Transaction dataset

Column	Description
Row	A row counter
Sample Name	The sample that contains the transaction. If a transaction appears in multiple samples, the transaction will be repeated, but with a different sample name.
Transaction Id	A unique ID that identifies the transaction.
Anon Student Id	DataShop-generated anonymous student ID.
Session Id	A dataset-unique string that identifies the user's session with the tutor.
Time	Time the transaction occurred. The transaction time is at the point in which students press return.
Student Response Type	The type of attempt made by the student.
Student Response Subtype	A more detailed classification of the student attempt. For example, the CTAT software describes actions taken by the tutor on behalf of the student as having subtype "tutor-performed".
Tutor Response Type	The type of response made by the tutor.
Problem Name	The name of the problem.
KC	The knowledge component for this transaction.

The observations in Student Step dataset are ordered by student, time of the first correct attempt (encoded as "Correct Transaction Time") or, in the absence of a correct attempt, and the time of the final transaction on the step (encoded as "Step End Time"). However, the tutor intervention is missing in this dataset, which suggests that we need to combine the two datasets (Transaction dataset and Student Step dataset) for our analysis. Detailed information for columns in Student Step dataset is as below.

Same deal:

(1) variables highlighted in yellow require further explanation

(2) which variables are going to matter for your analyses (tell the reader)

Table 2. Description of some variables in Student Step dataset

Column	Description
Row	A row counter
Sample	The sample that includes this step. If you select more than one sample to export, steps that occur in more than one sample will be duplicated in the export.
Anon Student Id	The student that performed the step.
Problem Name	The name of the problem in which the step occurs.
First Attempt	The tutor's response to the student's first attempt on the step. Example values are "hint", "correct", and "incorrect".
Incorrects	Total number of incorrect attempts by the student on the step.
Hints	Total number of hints requested by the student for the step.
Corrects	Total correct attempts by the student for the step.
KC	Knowledge component(s) associated with the correct performance of this step. In the case of multiple KCs assigned to a single step, KC names are separated by two tildes ("~~").
Opportunity	An opportunity is the first chance on a step for a student to demonstrate whether he or she has learned the associated knowledge component. Opportunity number is therefore a count that increases by one each time the student encounters a step with the listed knowledge component. In the case of multiple KCs assigned to a single step, opportunity number values are separated by two tildes ("~~") and are given in the same order as the KC names.

be careful of language: do you mean the computer tutor or a human instructor?

Be careful of this through-out the document!

We bridged the tutor intervention as an indicator variable from the transaction dataset to Student Step dataset, which suggests whether a tutor has intervened in the learning process during this observation. Based on this indicator variable, we calculate the tutor intervention time for each student, which is under the first assumption (mentioned in Methods) and tutor intervention time by different KC for each student, which is under the second assumption (mentioned in Methods).

As indicated in Figure 2.1, the numbers of tutor intervention for each student on all KCs (Knowledge Component) vary from 0 to 37 and the distribution appears to be rightly skewed. There are 17 students who never get tutor intervention.

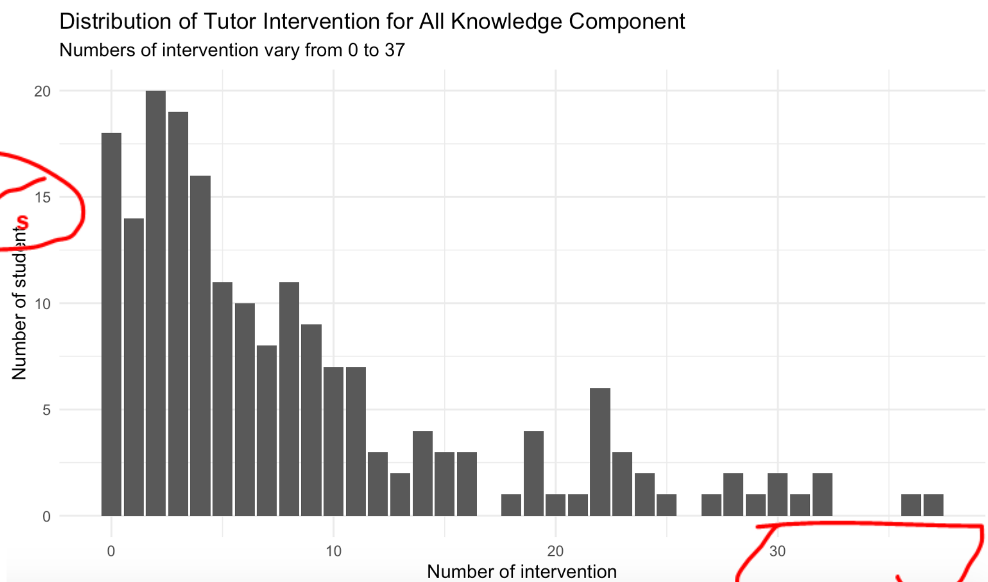


Figure2.1 Distribution of numbers of intervention across all KCs

Figure 2.2 shows the distribution of tutor intervention times for each KC. We see that "Combine_constant_terms'' and "Add/subtract_from_both_sides" are the two KC that tutors have helped with. This might relate to the difficulty of problems associated with these two KCs, or the amount of the problems associated with these KC's.

From fig 2.2, these are not the only two KC's that received help. What do you actually mean to say here?

How can there be both "only 37" interventions, and also "more than 600" interventions?

Please explain to the reader how both figures can be correct.

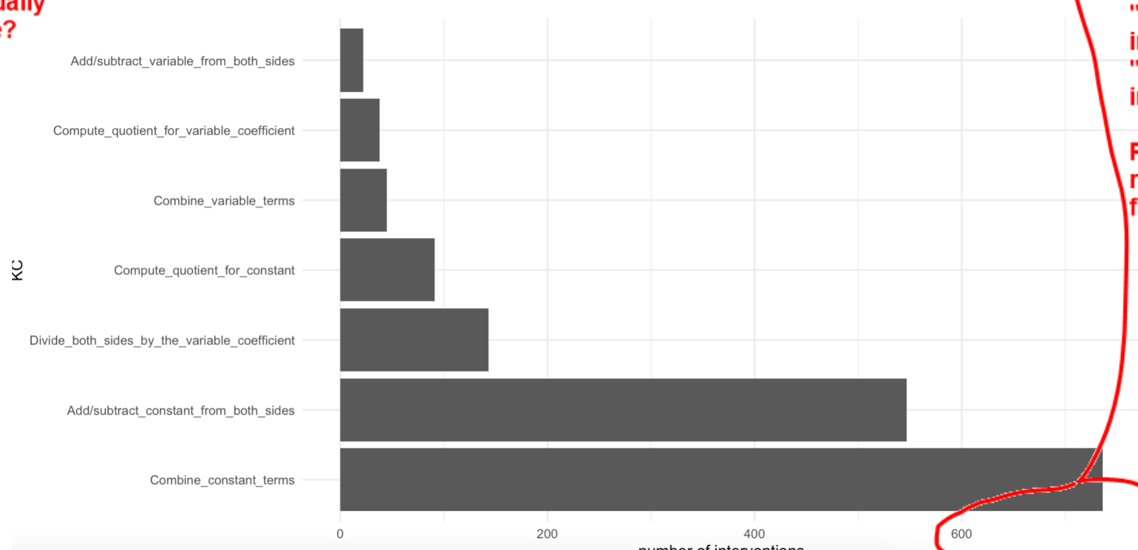


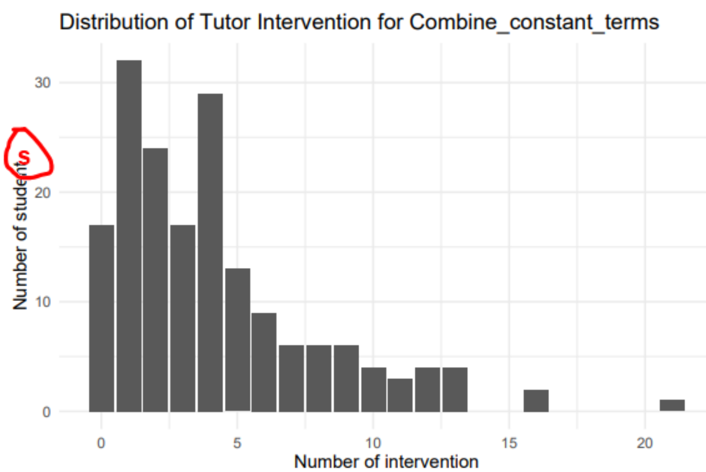
Figure2.2 Distribution of numbers of intervention for different KCs

(We then focus on a single KC, "Combine_constant_terms", and plot the distribution of tutor intervention numbers of this KC, which is shown in Figure 2.3. Apparently, the distribution is still rightly skewed, indicating that most students did not receive much tutor intervention on this KC.

This is a good idea but explain to the reader why you are doing it.

Is the right-skewing typical of all KCs? If so, say so.

If not, say what other distribution shapes you saw.

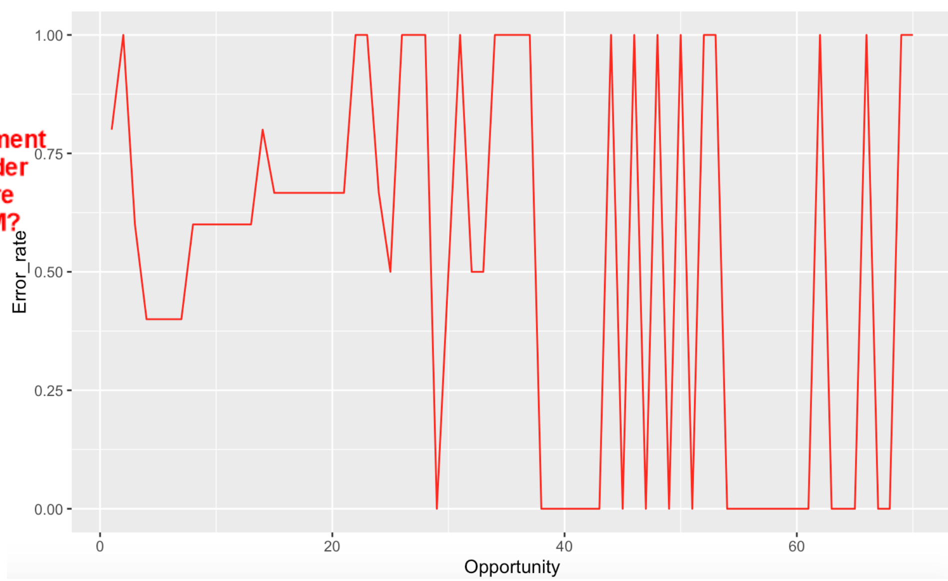


?

Figure2.3 Distribution of numbers of intervention for KC “Combine_constant_terms”

We also calculate the error rate for all KCs and a single student, and the line chart is shown in Figure 2.4. Based on ploy, we cannot observe any trend, as the error rate would sometimes switch between 0 and 1. Consequently, we would need a model to fit the raw error rate and that would be the AFM model (introduced in Methods).

I agree with this, but before this don't you need to make an argument that you should consider KC's individually before you talk about the AFM?



If this plot is for all KC's and one student, I don't really get it.

Each KC gets its own opportunity count. But what counts as an opportunity here?

What are you actually plotting?

Figure2.4 Error rate for one student)

3. Methods

3.1 Method 1

We have tried two methods so far to see if the tutor intervention accelerates the students' learning rate. The first method is fitting two AFM models for each knowledge component. The original AFM model is shown below.

what is m?

$$m(i, j \in KCs, n) = \alpha_i + \sum_{j \in KCs} (\beta_j + \gamma_j n_{i,j})$$

The AFM model is a logistic regression model that estimates the log ~~likelihood~~ odds that student gets step correct (Koedinger, 2010). The first variable is the student's initial proficiency, and the second variable is a combination of the ease of the KC and how much the student learned on prior opportunities for this KC. The first AFM model is fitted on the pre-tutor data. In this step, we want to know how the students learn before tutor intervention. And then we fit another AFM model on the post-tutor data. Since the intervention happens at different times for different students, we move all the data after intervention to the beginning. This means, if the intervention happens at opportunity M, then we treat opportunity M+1 as opportunity 1 and so on and so forth. This method gives us two AFM models. We plot them in one plot and see if there is a jump between these two cures. ? If the second AFM model gives us a lower error rate, that may imply that tutor intervention does accelerate the learning rate. However, in this method we assume that one intervention will influence all KCs. This assumption later got overturned since our client is more interested in having one model instead of two.

you haven't talked about pre-tutor data yet. This needs to be discussed in the "data" section above.

Also need to say what this is (in the data section).

inconsistent notation: you used n, not M, in the equation above for opportunity. Be consistent.

separate out these two ideas -- they are different ideas! -- and talk about them separately.

What is good or bad about each one?

3.2 Method 2

For method 2, we focused on improving the current AFM model. We used the idea of Performance Factors Analysis. The PFA model is shown below.

no \alpha_i ?

$$m(i, j \in KCs, s, f) = \sum_{j \in KCs} (\beta_j + \gamma_j s_{i,j} + \rho_j f_{i,j})$$

say what they are. (are there really 4? I am not getting that)

Compared to the AFM model, PFA has four new parameters which track the prior successes for the KC for the student, f tracks the prior failures for the KC for the student, and γ and ρ scale the effect of these observation counts. However, in our case, we say that s tracks whatever happens before the intervention and f tracks whatever happens after the intervention. This idea led us to our new model, a linear mixed-effects model.

nice idea

Indices not the same. Explain i, j, k, or use consistent notation.

$$AFM_k = \theta_i + \gamma_k N_{ik} + \phi_k N_{ik} I_{ik}\{Post\}$$

neat model. As I said in class, this allows for changes in slope on opportunity but not changes in overall difficulty. You could add another term, say \alpha_k I_{ik}{Post} to track change in difficulty.

Compared to the original AFM model, we still have a variable to represent the student's initial proficiency, a variable that represents the number of opportunities the student has tried, and a variable that indicates whether these opportunities are before tutor or after tutor. Also, in this method, we changed our assumption. Here, we assume one tutor intervention will only influence one KC. Since we change our assumption, we also have a different separation method. Below is a brief demonstration of how we define pre and post intervention opportunity.

This needs to come earlier in the paper, and with more explanation.

Student 1	**1**	**2**	**3**	**4**	**5**
Student 2	**1**	**2**	**3**	**4**	**5**
Student 3	**1**	**2**	**3**	**4**	**5**

We first subset by KC, and separate the data according to the first intervention for each student. For example, in the table above, for student 1, the first intervention happened after the first opportunity. Therefore, the first opportunity is pre tutor data and whatever after it are post tutor data. For student 2, the first intervention happened after the third opportunity, so the first to the third opportunities are pre tutor data and whatever after the third opportunity are post tutor data. For student 3, the first intervention happened after the second opportunity, so the first and the second opportunities are pre tutor data and whatever after the second opportunity are post tutor data.

3.3 Method 3

After presenting method 2 to our client, he suggests we try method 3 which is a combination of method 1 and 2. We first fit the original AFM model to all students. Then we fit the original AFM model to students who never get tutor intervention. Last, we fit the new AFM model we generated in method 2 to students who got tutor intervention. We compare the slopes and intercepts of these models and see if there are any useful insights we can derive from them.

you can do this in one model by adding a main effect for I{Post} as I suggested under method 2 above.

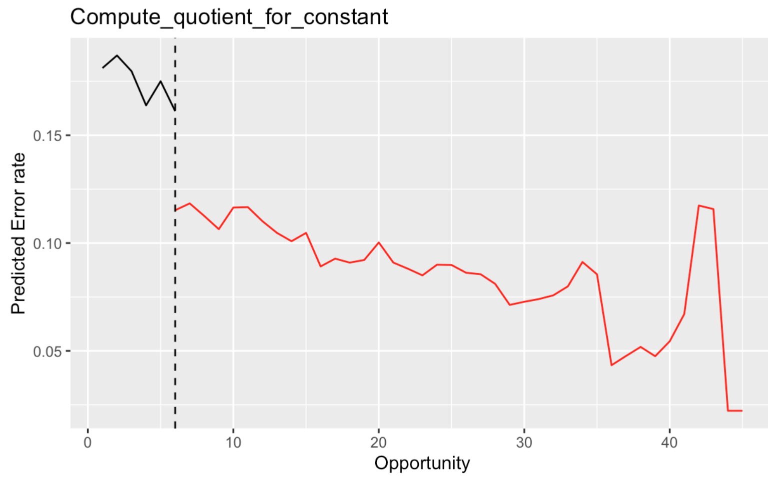
4. Results

4.1 Results from Method 1

From our initial data exploration, we found that for KC “Compute quotient for constant”, the most frequent first intervention time was around opportunity 7. Thus, we separated the data into two subsets according to this intervention point. Later, we fitted two AFM models and predicted the error rate. In figure 4.1, the black line indicated the pre-tutor error rate that is calculated using the pre-tutor subset. The red line represented the post-tutor error rate derived by the post-tutor subset. We observed a jump between the two lines, which implied that the first intervention at opportunity 6 improved students’ performance on Computing the quotient for constant.

sort of cool how well this worked!

Did it work on other KC's too?



need a legend indicating the meaning of red vs black lines

Figure 4.1. Pre vs post tutor error rate for KC “Compute quotient for constant”

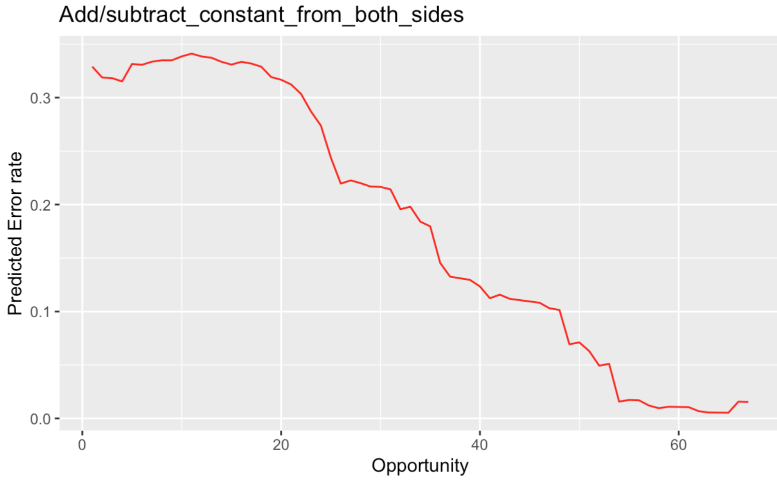
How many KC's did you apply this method to? How well did it work on each of them?

4.2 Results from Method 2

and this is different for every student, right? if so, please write that for the reader.

With the separation method mentioned in method 2, we managed to split the data into two groups with a variable indicating which observation is pre-tutor or post-tutor. As indicated in figure 4.3, we picked KC “Add/subtract constant from both sides” and predicted the error rate for all students who received tutors’ help at some point of their learning progress. Overall, we observed that with the opportunity goes up, students’ predicted error rate for this KC decreases.

Isn't the important thing here the value of \phi? What is that value and what is its standard error?



if this is aggregated across students and intervention points, I'm not sure it can tell us very much (since the interventions happen on different opportunities for different students)

If it is telling us something useful please say what that is.

Figure4.3. Predicted Error rate for all students in KC “Add/subtract constant from both side”

To get a clear idea of the tutor effect, we randomly picked a student whose first intervention in this KC happened at opportunity 15. As shown in figure 4.3, we noticed that before opportunity 5, there’s a dip in the predicted error rate. However, we did not observe a significant change on the slope before and after the intervention at opportunity 15.

Do you mean \hat\phi was not signif different from 0?

Show us.

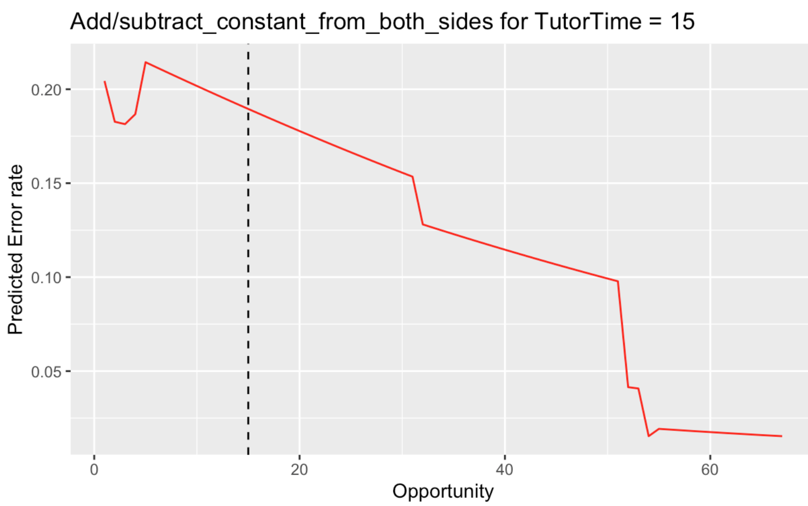


Figure4.3. Error rate for one student whose first intervention happens at opportunity 15

* At this point, we are still exploring our last method. After we decide the final method to use, we will present more statistical analysis(significance, coefficients,etc)

Great to see! I'm looking forward to your final paper (and also the single model that allows you to look at changes in slope and intercept in the same model)!

5. Discussion

For our results of Method 1, we observed a significant gap before and after tutor intervention in predicted error rate, which potentially suggests that the tutors' interventions are effective at improving students' performance. However, the assumption we held for this method has some flaws. That is, simply assuming one tutor intervention would influence all KCs would lead to imbalance sample size between pre and post tutor subset. For example, if a student encountered tutor intervention when they was taking KC1-related questions at Opportunity=2 for KC1, and then encountered tutor intervention when learning KC2-related questions at Opportunity=6 for KC2, our algorithm would then decide the tutor intervention time for this student as 2, instead of 6, allocating all observations after Opportunity=2 as post-tutor subset for this student. Besides, our client prefers the assumption in Method 2 than that in Method 1 as well. Our client also prefers an integrated model, rather than two separate models.

Method 2 is an improved version of method 1. The assumption is that one tutor intervention would only influence KC(s) related to the question that a student is taking within one observation(which would always be 1 or 2, no more than 2). Based on our results of Method 2, there's an increase in error rate followed by a sharp turn with decreasing error rate, which might suggest that tutors intervene after noticing the struggle. Nonetheless, we did not observe the expected changes in slopes before and after intervention for a single student, and the tutor intervention time does not match with the break point of the slopes as well. Consequently, we need another way to observe our predicted results.

This is a good start.

For the final paper, phrase things in terms of good and bad features of each of the methods, and results from each of the methods, rather than a conversation between you and the client.

These references are good.

1. make sure to cite each one at least once in the main body of the paper.

2. They are close to ASA style but not quite there yet. Please review ASA style and edit the references to conform more closely to that style.

References

Ido Roll, Ryan S. J. d. Baker, Vincent Aleven & Kenneth R. Koedinger (2014) On the Benefits of Seeking (and Avoiding) Help in Online Problem-Solving Environments, Journal of the Learning Sciences, 23:4, 537-560, DOI: 10.1080/10508406.2014.883977

Koedinger, K.R., Baker, R.S.J.d., Cunningham, K., Skogsholm, A., Leber, B., Stamper, J. (2010) A Data Repository for the EDM community: The PSLC DataShop. In Romero, C., Ventura, S., Pechenizkiy, M., Baker, R.S.J.d. (Eds.) Handbook of Educational Data Mining. Boca Raton, FL: CRC Press.

Aleven, V., & Koedinger, K. R. (2013). Knowledge component approaches to learner modeling. In R. Sottilare, A. Graesser, X. Hu, & H. Holden (Eds.), Design recommendations for adaptive intelligent tutoring systems (Vol. I, Learner Modeling, pp. 165-182). Orlando, FL: US Army Research Laboratory.

Pavlik Jr, Phil & Cen, Hao & Koedinger, Kenneth. (2009). Performance Factors Analysis - A New Alternative to Knowledge Tracing. Frontiers in Artificial Intelligence and Applications. 200. 531-538. 10.3233/978-1-60750-028-5-531.

Koedinger, K.R., Baker, R.S.J.d., Cunningham, K., Skogsholm, A., Leber, B., Stamper, J. (2010) A Data Repository for the EDM community: The PSLC DataShop. In Romero, C., Ventura, S., Pechenizkiy, M., Baker, R.S.J.d. (Eds.) Handbook of Educational Data Mining. Boca Raton, FL: CRC Press. For exploratory analysis, we used the PSLC DataShop, available at http://pslcdatashop.web.cmu.edu (Koedinger et al., 2010).

Cen H., Koedinger K., Junker B. (2008) Comparing Two IRT Models for Conjunctive Skills. In: Woolf B.P., Aïmeur E., Nkambou R., Lajoie S. (eds) Intelligent Tutoring Systems. ITS 2008. Lecture Notes in Computer Science, vol 5091. Springer, Berlin, Heidelberg. https://doi.org/10.1007/978-3-540-69132-7_111

Technical Appendix

HCI-Learning Discontinuity

2021/4/21

good start on tech appendix.

remember to put enough english before and after each analysis that reader can understand what and why you are doing things, and what the results are

```
library(tidyverse)
library(ggpubr)
```

Load data

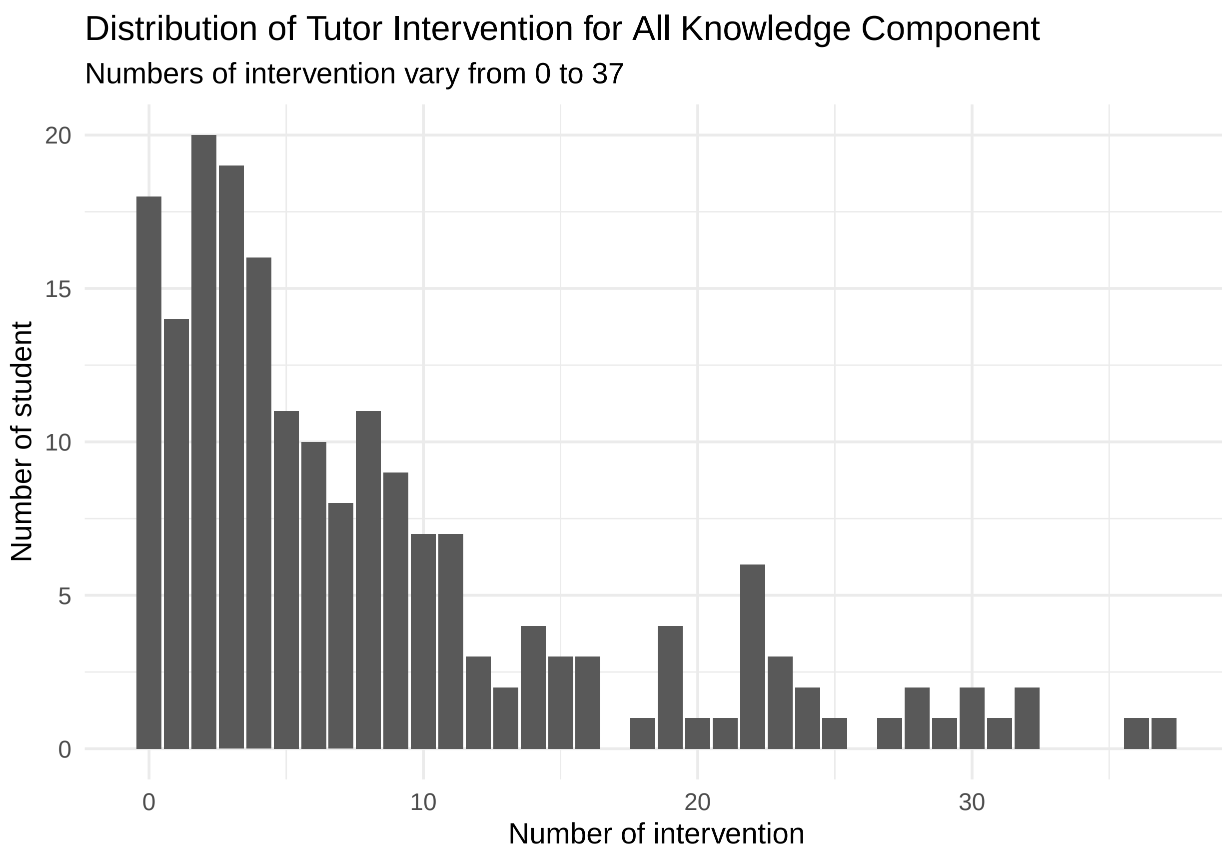
The loaded data is combined version of transaction dataset and studentstep data set. Due to the run-time limitation, we will show the R code of bridging the data along with our final products.

```
HCI_full <- read.csv("combined_195.csv")
HCI_full <- rename(HCI_full, c("IfTutor" ="V29"))
HCI_full$IfTutor <- ifelse(is.na(HCI_full$IfTutor),0, 1)
```

EDA

Distribution of the number of intervention

```
tutor_num <- HCI_full %>%
  select(Anon.Student.Id, IfTutor) %>%
  group_by(Anon.Student.Id) %>%
  summarise(num_intervention = sum(IfTutor)) %>%
  arrange(num_intervention)

num_tutor <- tutor_num %>%
  select(num_intervention, Anon.Student.Id) %>%
  group_by(num_intervention) %>%
  summarise(number_student = n()) %>%
  add_row(num_intervention=0, number_student=18, .before = 1)

ggplot(data=num_tutor, aes(x=num_intervention, y= number_student))+
  geom_bar(stat="identity") +
  labs(title="Distribution of Tutor Intervention for All Knowledge Component" , y="Number of student",
       x= "Number of intervention", subtitle = "Numbers of intervention vary from 0 to 37")+
  theme_minimal()
```

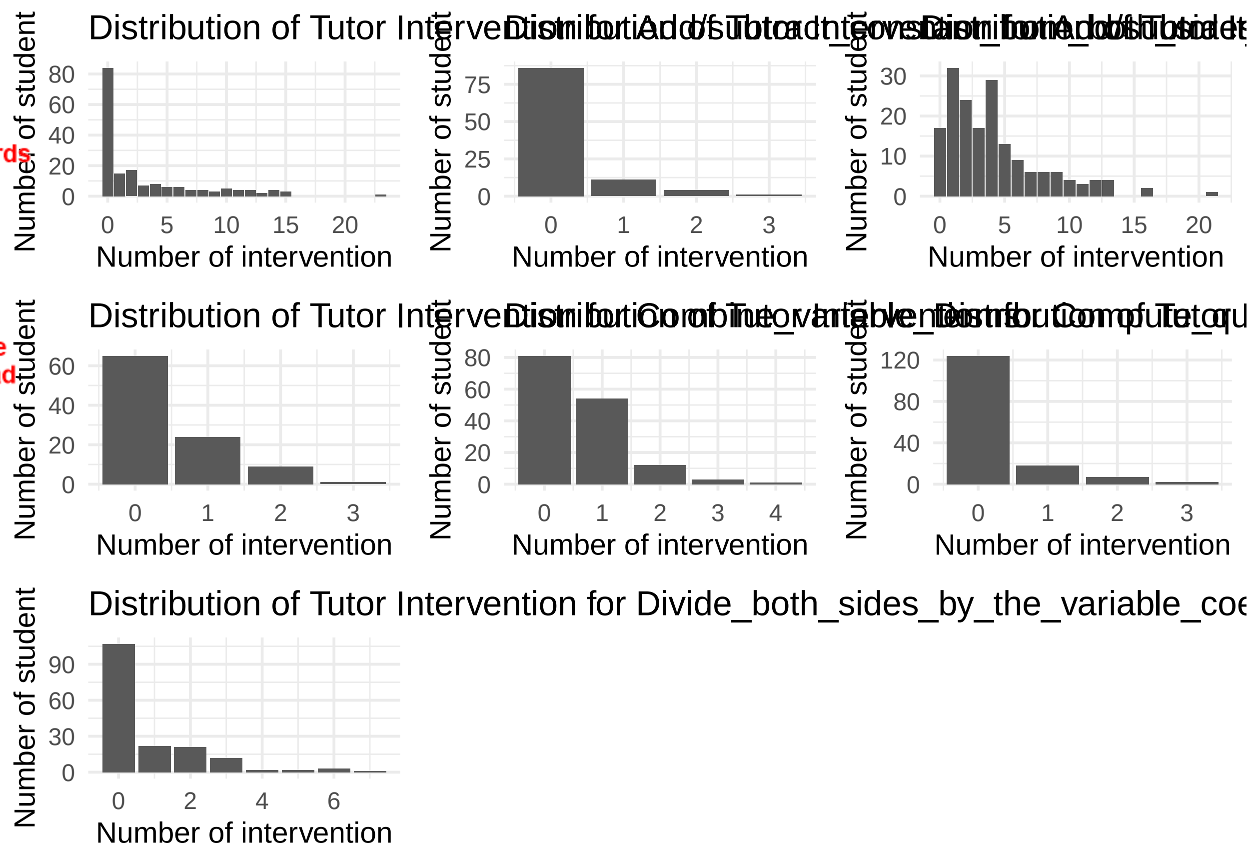



```
kc <- c("Add/subtract_constant_from_both_sides",
        "Add/subtract_variable_from_both_sides",
        "Combine_constant_terms",
        "Combine_variable_terms",
        "Compute_quotient_for_constant",
        "Compute_quotient_for_variable_coefficient",
        "Divide_both_sides_by_the_variable_coefficient")

for(k in 1:7){
  df <- HCI_full %>%
    filter(grepl(kc[k], KC..Default.)) %>%
    select(Anon.Student.Id, IfTutor) %>%
    group_by(Anon.Student.Id) %>%
    summarise(num_intervention = sum(IfTutor)) %>%
    arrange(num_intervention)   %>%
    select(num_intervention, Anon.Student.Id) %>%
    group_by(num_intervention) %>%
    summarise(number_student = n())

    assign(paste0("kc_",k), ggplot(data=df, aes(x=num_intervention, y= number_student))+
                      geom_bar(stat="identity") +
                      labs(title=paste("Distribution of Tutor Intervention for", kc[k]),
                           y="Number of student",
                           x= "Number of intervention")+
                      theme_minimal())
}
```

```
ggarrange(kc_1,kc_2,kc_3,kc_4,kc_5,kc_6,kc_7)
```

Use fewer words and smaller type so the titles don't crash into one another.

An unreadable graph is as bad as no graph at all.



Pick a student and look at the true error rate with respected first tutor intervention time.

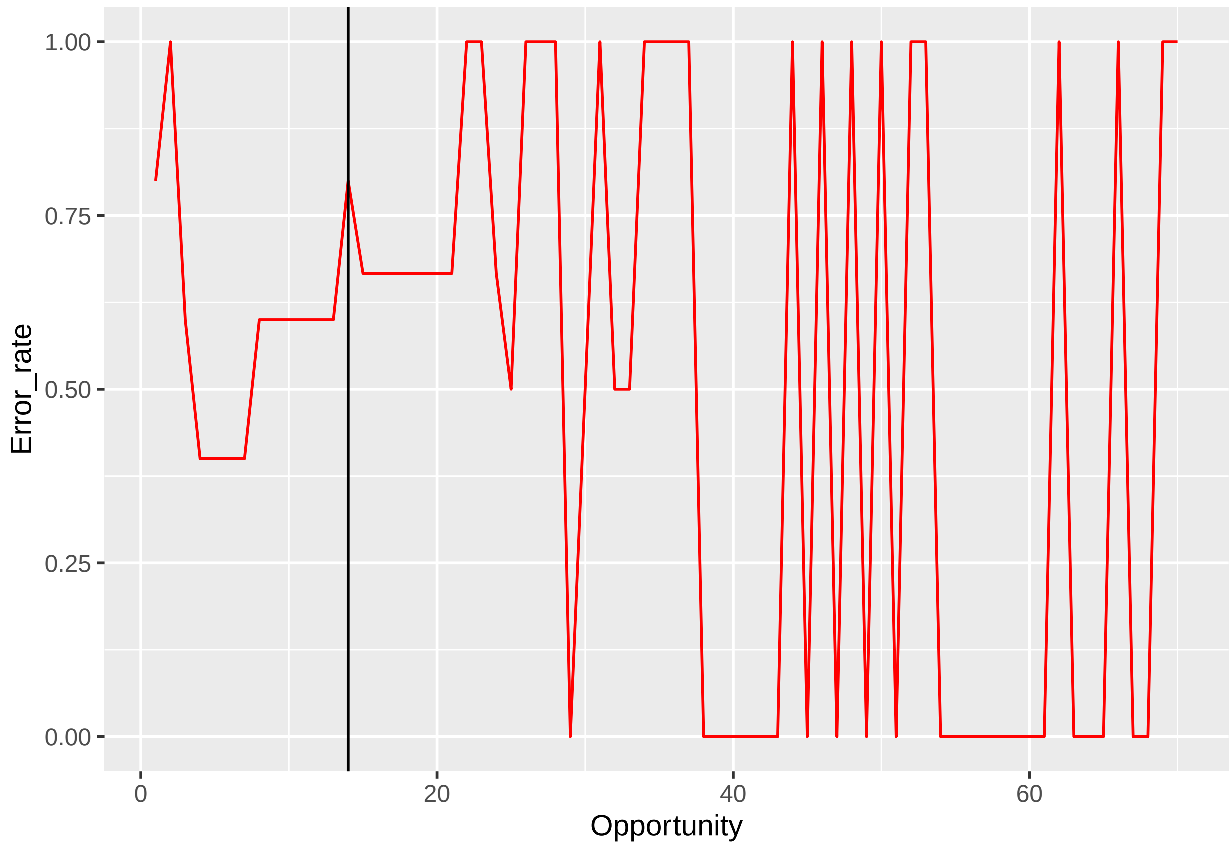
```
stu1 <- HCI_full %>%
  filter(Anon.Student.Id == 'Stu_2095f540e7a586c41339d8a7be3ea8e2')
```

Define the error rate function

```
Error_calculate <- function(ds){
  oppo_vec <- as.numeric(as.character(unique(ds$Opportunity..Default.)))
  oppo_vec <- oppo_vec[!is.na(oppo_vec)]
  error_rate <- rep(0,length(oppo_vec))
  for (i in oppo_vec) {
    stu_oppo <- ds %>%
      filter(Opportunity..Default. == i)
    errors <- rep(0, nrow(stu_oppo))
    for (j in 1:nrow(stu_oppo)) {
      if(stu_oppo$Incorrects[j] == 0 || stu_oppo$Hints[j] == 0)
        errors[j] <- 0
      if(stu_oppo$Incorrects[j] > 0 || stu_oppo$Hints[j] > 0)
        errors[j] <- 1
    }
    error_rate[i] <- sum(errors)/nrow(stu_oppo)
  }
```

```r
  rate <- data.frame(Opportunity = oppo_vec, Error_rate = error_rate)
  return(rate)
}
```

```r
rate <- Error_calculate(stu1)
tutor_oppo <- min(as.numeric(as.character(stu1[stu1$IfTutor == 1,]$Opportunity..Default.)))
ggplot(aes(y = Error_rate, x = Opportunity), data = rate) + geom_line(col = 'red') + geom_vline(xinterce
```



Method 1

```r
AnonId <- as.character(unique(HCI_full$Anon.Student.Id))
m <- length(AnonId)

# Replace NA with 0
HCI_full$IfTutor[is.na(HCI_full$IfTutor)] <- 0 #IfTutor = V29 here

tutor_time <- rep(0,m)
for (i in 1:m) {
  # Find the tutor intervention time (opportunity)
  rows <- which(HCI_full$Anon.Student.Id == AnonId[i])
  stu <- HCI_full[rows,]
  tutor_oppo <- min(as.numeric(as.character(stu[stu$IfTutor == 1,]$Opportunity..Default.)))
```

```r
  tutor_time[i] <- tutor_oppo
}
tutor_time[is.na(tutor_time)] <- 0

# Add a column showing the intervention time (opportunity) for this student
TutorTime <- rep(0,nrow(HCI_full))
for (i in 1:m) {
  rows <- which(HCI_full$Anon.Student.Id == AnonId[i])
  TutorTime[rows] <- tutor_time[i]
}
HCI_full$TutorTime <- TutorTime
```

```r
ds5 <- HCI_full %>%
  filter(grepl('Compute_quotient_for_constant', KC..Default.))
# Pre-tutor: Intervention time after 6
# Post-tutor: Intervention time before 6 (including 6)
ds5_pre <- ds5 %>%
  filter(TutorTime > 6)
ds5_post <- ds5 %>%
  filter(TutorTime <= 6)

L1 <- length(ds5_pre$Anon.Student.Id)
Success1 <- vector(mode="numeric", length=L1)
Success1[ds5_pre$First.Attempt=="correct"] <- 1
model1.glm <- glm(Success1 ~ Anon.Student.Id + 1:Opportunity..Default., data=ds5_pre, family=binomial())

summary(model1.glm)
```

```
##
## Call:
## glm(formula = Success1 ~ Anon.Student.Id + 1:Opportunity..Default.,
##     family = binomial(), data = ds5_pre)
##
## Deviance Residuals:
##      Min        1Q    Median        3Q       Max
## -2.66659   0.00005   0.00005   0.32029   1.93484
##
## Coefficients:
##                                                      Estimate Std. Error
## (Intercept)                                         3.526e+00  1.015e+00
## Anon.Student.IdStu_0c62473254fc9dea3b71514a2cb75664  1.704e+01  1.024e+04
## Anon.Student.IdStu_13a914262c53c434b0f00957a85afb98 -2.833e+00  1.334e+00
## Anon.Student.IdStu_1669ccda209abd6a8d16dab09902dc35 -1.655e+00  1.267e+00
## Anon.Student.IdStu_185ebbc0b8afcbecd83e6328ca07afeb -1.447e+00  1.468e+00
## Anon.Student.IdStu_1bad963b134c86be6400445393bbd086 -2.833e+00  1.103e+00
## Anon.Student.IdStu_1fae116ff2f7973fc9232ca0b8d6f9c3 -1.329e+00  1.259e+00
## Anon.Student.IdStu_2095f540e7a586c41339d8a7be3ea8e2 -1.735e+00  1.270e+00
## Anon.Student.IdStu_26a567c6ce161f3a5b356a793f261522 -3.526e+00  1.741e+00
## Anon.Student.IdStu_28eeaa94e65be324be699009dda3939c -4.220e+00  1.590e+00
## Anon.Student.IdStu_309cb9fb08b6b2bc1d31fc8934af7a10  1.704e+01  3.477e+03
## Anon.Student.IdStu_3297ad412d5314032ac1009bdc74e4c0  1.704e+01  3.184e+03
## Anon.Student.IdStu_456727c07d0cd25f85e8cee94706b64b -2.833e+00  1.590e+00
## Anon.Student.IdStu_4b11834453ca35d57884f03f858e28b9  1.704e+01  3.041e+03
```

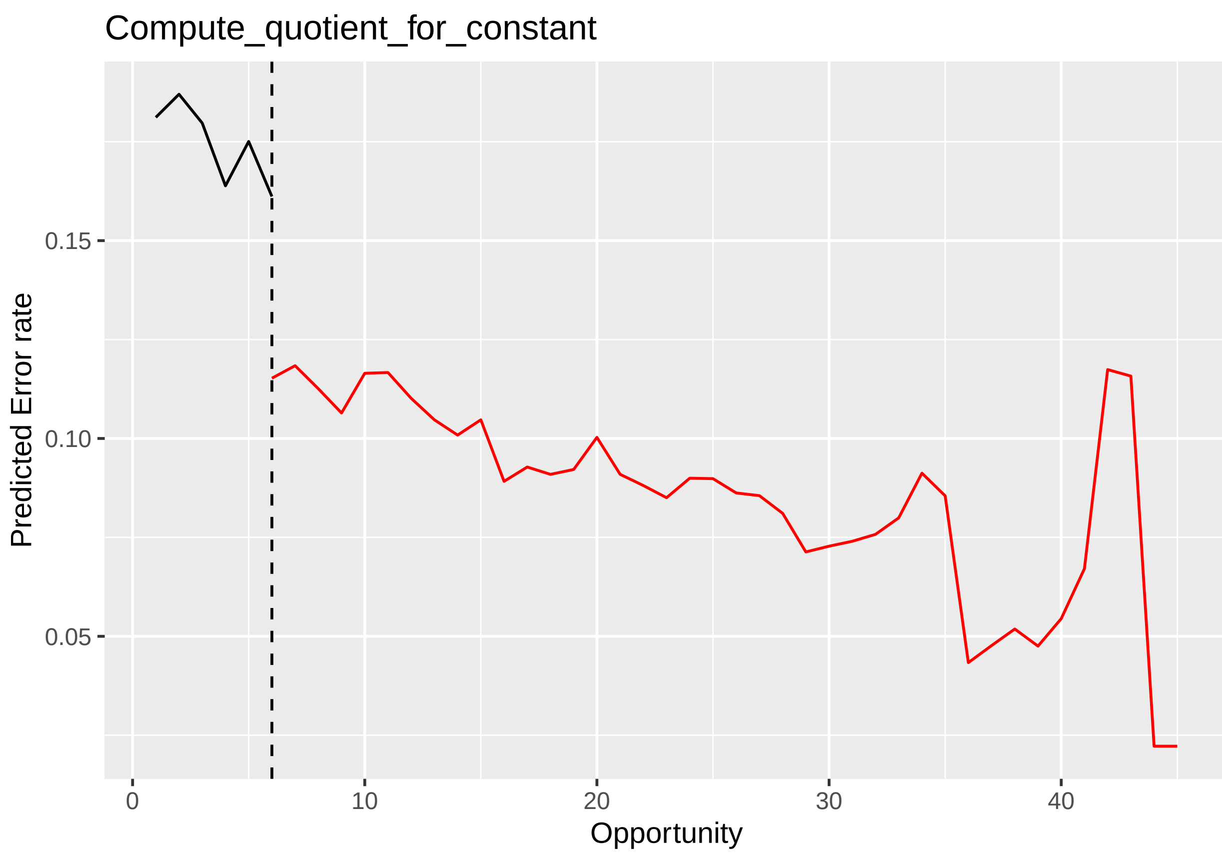
```
## Anon.Student.IdStu_518cb92a0ae4cc6a4c63fac064080aa0 -2.833e+00  1.334e+00
## Anon.Student.IdStu_52ec3719d321213f2617c4344b7f4bd5  1.704e+01  3.412e+03
## Anon.Student.IdStu_5648ddfacac53f5d057053ce66761573  1.704e+01  3.619e+03
## Anon.Student.IdStu_58fa0fa9dce9c2569b616c48173af0c3  1.704e+01  3.134e+03
## Anon.Student.IdStu_5a38fe2305dc34585d95471c53ae3266  1.704e+01  2.614e+03
## Anon.Student.IdStu_5d2be16526925841f53bf1d33d974d54 -1.252e-01  1.436e+00
## Anon.Student.IdStu_68d3d97a5265d1cd4ad5b8d7f06434db  1.704e+01  5.346e+03
## Anon.Student.IdStu_6c1cf94bb0925a34fda6b982dea87663 -2.610e+00  1.315e+00
## Anon.Student.IdStu_6f2533c91c4dd77bb7c3b004fc8b64df -2.274e+00  1.293e+00
## Anon.Student.IdStu_6f51d792b9f63033ee2a6643006c11c3 -9.237e-02  1.436e+00
## Anon.Student.IdStu_7093bb8cb97aabc2f3e4fd07d14c6d31 -1.191e+00  1.181e+00
## Anon.Student.IdStu_7173892326345cedac5bdd38804c959e -5.819e-01  1.247e+00
## Anon.Student.IdStu_751ad052d8a1405eb244ee839cce443e -3.526e+00  1.196e+00
## Anon.Student.IdStu_898f700e93d0075a9b1d7c4d3272c339 -2.140e+00  1.510e+00
## Anon.Student.IdStu_8f655a22c047ac0188c23033bf768244 -2.409e+01  6.701e+03
## Anon.Student.IdStu_95fdb9fc04eeb1d28137f232a2ae9512  1.704e+01  4.068e+03
## Anon.Student.IdStu_9b5de282662e21d7e8dab1dab44df6dc -1.252e-01  1.436e+00
## Anon.Student.IdStu_b3fa0b5ca7805966d1c1a0e8cfe47d8d -5.231e+00  1.273e+00
## Anon.Student.IdStu_b56bfb0a88008b152d7ca31925035585 -3.056e+00  1.164e+00
## Anon.Student.IdStu_c502bf88de8226f7f2929c31263a8c86 -3.201e+00  1.078e+00
## Anon.Student.IdStu_c9d92b3901959bf36c2f285ef6f6cf2e  1.704e+01  8.865e+03
## Anon.Student.IdStu_d1500ee7ebf4823f7b1a7614157a9916  1.704e+01  3.134e+03
## Anon.Student.IdStu_dfcad7876ca509486227c83277674f33 -1.041e+00  1.454e+00
## Anon.Student.IdStu_e3f2abc4baa6bb57d741853466310c1b  1.704e+01  2.559e+03
## Anon.Student.IdStu_e82fd4792d5f687fa1eb763956b41302 -1.917e+00  1.493e+00
## Anon.Student.IdStu_f294f51ad79e0bb784cbc591eabc09ea  1.704e+01  2.769e+03
## Anon.Student.IdStu_f66b03e7d166a2ce61ef0c6f035260c2  1.704e+01  7.238e+03
## Anon.Student.IdStu_f7805b8c0ecc73d172236b53c0f6fdf0  1.704e+01  2.876e+03
## Anon.Student.IdStu_fdab289cba1ce5e9e900da87b10a4708  1.704e+01  8.865e+03
##                                                     z value Pr(>|z|)
## (Intercept)                                           3.476  0.00051 ***
## Anon.Student.IdStu_0c62473254fc9dea3b71514a2cb75664   0.002  0.99867
## Anon.Student.IdStu_13a914262c53c434b0f00957a85afb98  -2.124  0.03368 *
## Anon.Student.IdStu_1669ccda209abd6a8d16dab09902dc35  -1.305  0.19174
## Anon.Student.IdStu_185ebbc0b8afcbecd83e6328ca07afeb  -0.986  0.32424
## Anon.Student.IdStu_1bad963b134c86be6400445393bbd086  -2.568  0.01022 *
## Anon.Student.IdStu_1fae116ff2f7973fc9232ca0b8d6f9c3  -1.056  0.29108
## Anon.Student.IdStu_2095f540e7a586c41339d8a7be3ea8e2  -1.366  0.17197
## Anon.Student.IdStu_26a567c6ce161f3a5b356a793f261522  -2.026  0.04276 *
## Anon.Student.IdStu_28eeaa94e65be324be699009dda3939c  -2.653  0.00798 **
## Anon.Student.IdStu_309cb9fb08b6b2bc1d31fc8934af7a10   0.005  0.99609
## Anon.Student.IdStu_3297ad412d5314032ac1009bdc74e4c0   0.005  0.99573
## Anon.Student.IdStu_456727c07d0cd25f85e8cee94706b64b  -1.781  0.07484 .
## Anon.Student.IdStu_4b11834453ca35d57884f03f858e28b9   0.006  0.99553
## Anon.Student.IdStu_518cb92a0ae4cc6a4c63fac064080aa0  -2.124  0.03368 *
## Anon.Student.IdStu_52ec3719d321213f2617c4344b7f4bd5   0.005  0.99602
## Anon.Student.IdStu_5648ddfacac53f5d057053ce66761573   0.005  0.99624
## Anon.Student.IdStu_58fa0fa9dce9c2569b616c48173af0c3   0.005  0.99566
## Anon.Student.IdStu_5a38fe2305dc34585d95471c53ae3266   0.007  0.99480
## Anon.Student.IdStu_5d2be16526925841f53bf1d33d974d54  -0.087  0.93055
## Anon.Student.IdStu_68d3d97a5265d1cd4ad5b8d7f06434db   0.003  0.99746
## Anon.Student.IdStu_6c1cf94bb0925a34fda6b982dea87663  -1.985  0.04717 *
## Anon.Student.IdStu_6f2533c91c4dd77bb7c3b004fc8b64df  -1.758  0.07872 .
## Anon.Student.IdStu_6f51d792b9f63033ee2a6643006c11c3  -0.064  0.94870
```

```
## Anon.Student.IdStu_7093bb8cb97aabc2f3e4fd07d14c6d31  -1.008  0.31328
## Anon.Student.IdStu_7173892326345cedac5bdd38804c959e  -0.467  0.64082
## Anon.Student.IdStu_751ad052d8a1405eb244ee839cce443e  -2.949  0.00318 **
## Anon.Student.IdStu_898f700e93d0075a9b1d7c4d3272c339  -1.417  0.15634
## Anon.Student.IdStu_8f655a22c047ac0188c23033bf768244  -0.004  0.99713
## Anon.Student.IdStu_95fdb9fc04eeb1d28137f232a2ae9512   0.004  0.99666
## Anon.Student.IdStu_9b5de282662e21d7e8dab1dab44df6dc  -0.087  0.93055
## Anon.Student.IdStu_b3fa0b5ca7805966d1c1a0e8cfe47d8d  -4.110 3.96e-05 ***
## Anon.Student.IdStu_b56bfb0a88008b152d7ca31925035585  -2.626  0.00863 **
## Anon.Student.IdStu_c502bf88de8226f7f2929c31263a8c86  -2.970  0.00298 **
## Anon.Student.IdStu_c9d92b3901959bf36c2f285ef6f6cf2e   0.002  0.99847
## Anon.Student.IdStu_d1500ee7ebf4823f7b1a7614157a9916   0.005  0.99566
## Anon.Student.IdStu_dfcad7876ca509486227c83277674f33  -0.717  0.47368
## Anon.Student.IdStu_e3f2abc4baa6bb57d741853466310c1b   0.007  0.99469
## Anon.Student.IdStu_e82fd4792d5f687fa1eb763956b41302  -1.284  0.19920
## Anon.Student.IdStu_f294f51ad79e0bb784cbc591eabc09ea   0.006  0.99509
## Anon.Student.IdStu_f66b03e7d166a2ce61ef0c6f035260c2   0.002  0.99812
## Anon.Student.IdStu_f7805b8c0ecc73d172236b53c0f6fdf0   0.006  0.99527
## Anon.Student.IdStu_fdab289cba1ce5e9e900da87b10a4708   0.002  0.99847
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## (Dispersion parameter for binomial family taken to be 1)
##
##     Null deviance: 529.34  on 844  degrees of freedom
## Residual deviance: 290.50  on 802  degrees of freedom
## AIC: 376.5
##
## Number of Fisher Scoring iterations: 19
```

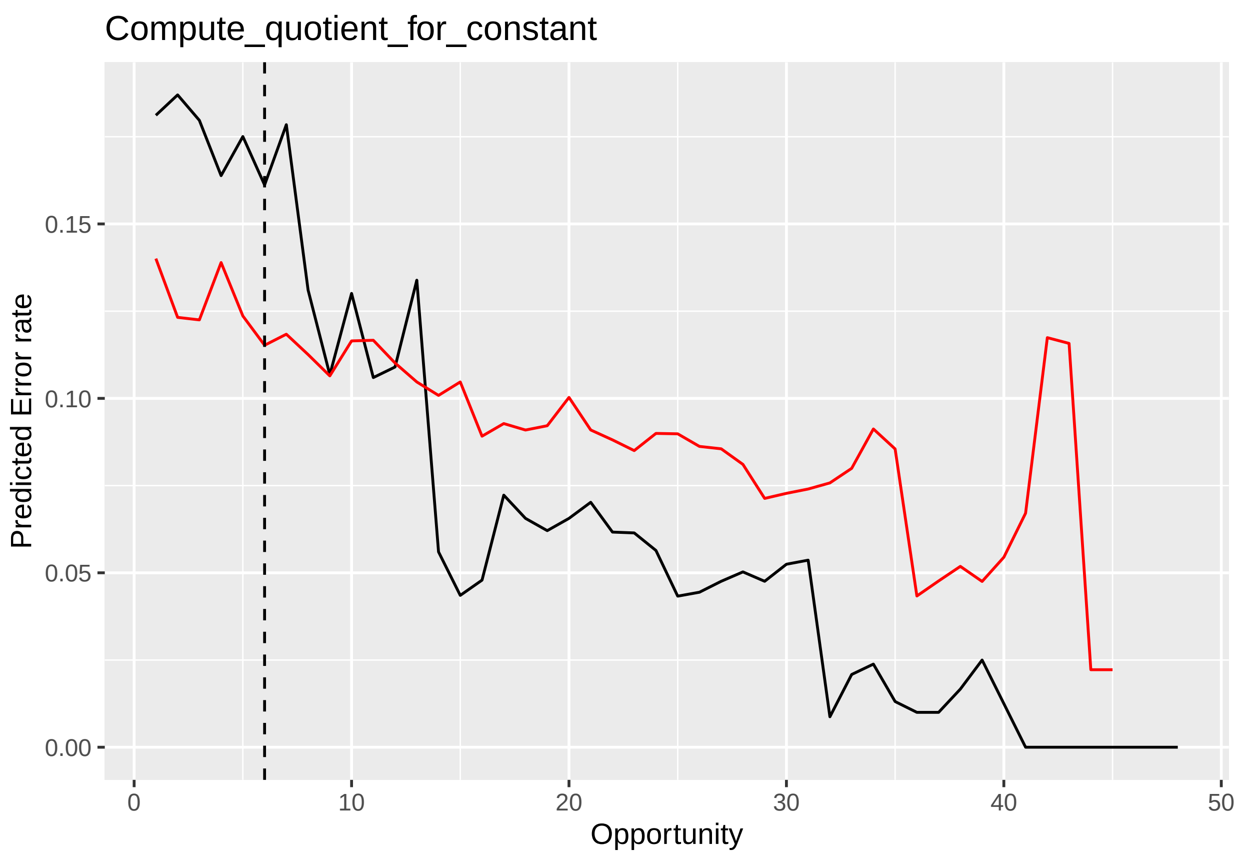
```r
pre.pred <- predict(model1.glm, ds5_pre, type="response")
df <- as.data.frame(cbind(as.numeric(as.character(ds5_pre$Opportunity..Default.)), pre.pred)) %>%
  group_by(V1) %>%
  summarise(error= 1 - mean(pre.pred))
L2 <- length(ds5_post$Anon.Student.Id)
Success2 <- vector(mode="numeric", length=L2)
Success2[ds5_post$First.Attempt=="correct"] <- 1
model2.glm <- glm(Success2 ~ Anon.Student.Id + 1:Opportunity..Default., data=ds5_post, family=binomial()
```

```r
post.pred <- predict(model2.glm, ds5_post, type="response")
df2 <- as.data.frame(cbind(as.numeric(as.character(ds5_post$Opportunity..Default.)), post.pred)) %>%
  group_by(V1) %>%
  summarise(error= 1 - mean(post.pred))
```

```r
#Plot results
df_before6 <- df %>%
  filter(V1 <= 6)
df_after6 <- df2 %>%
  filter(V1 >= 6)
ggplot(df_before6, aes(x=V1, y=error))+
  geom_line() +
  geom_line(data=df_after6, aes(x=V1, y=error), col="Red")+
  geom_vline(xintercept = 6, linetype="dashed") +
  labs(title="Compute_quotient_for_constant", x="Opportunity", y="Predicted Error rate")
```




```
ggplot(df, aes(x=V1, y=error))+
  geom_line() +
  geom_line(data=df2, aes(x=V1, y=error), col="Red")+
  geom_vline(xintercept = 6, linetype="dashed") +
  labs(title="Compute_quotient_for_constant", x="Opportunity", y="Predicted Error rate")
```



Method 2

```
HCI_full$TutorTime <- 0
# Add a column showing the intervention time (opportunity) for this student and this KC
for (i in 1:m) {
  rows <- which(HCI_full$Anon.Student.Id == AnonId[i])
  HCI_subset <- HCI_full[rows,]
  # The last KC: 'Divide_both_sides_by_the_variable_coefficient'
  rows7 <- grepl('Divide_both_sides_by_the_variable_coefficient', HCI_subset$KC..Default.)
  HCI7 <- HCI_subset[rows7,]
  if(nrow(HCI7) > 0){
    tutor_oppo7 <- min(as.numeric(as.character(HCI7[HCI7$IfTutor == 1,]$Opportunity..Default.)))
    HCI_subset[rows7, ]$TutorTime <- tutor_oppo7
  }
  # 6th KC: 'Compute_quotient_for_variable_coefficient'
  rows6 <- grepl('Compute_quotient_for_variable_coefficient', HCI_subset$KC..Default.)
  HCI6 <- HCI_subset[rows6,]
  if(nrow(HCI6) > 0){
    tutor_oppo6 <- min(as.numeric(as.character(HCI6[HCI6$IfTutor == 1,]$Opportunity..Default.)))
    HCI_subset[rows6, ]$TutorTime <- tutor_oppo6
  }
  # 5th KC: 'Compute_quotient_for_constant'
  rows5 <- grepl('Compute_quotient_for_constant', HCI_subset$KC..Default.)
```

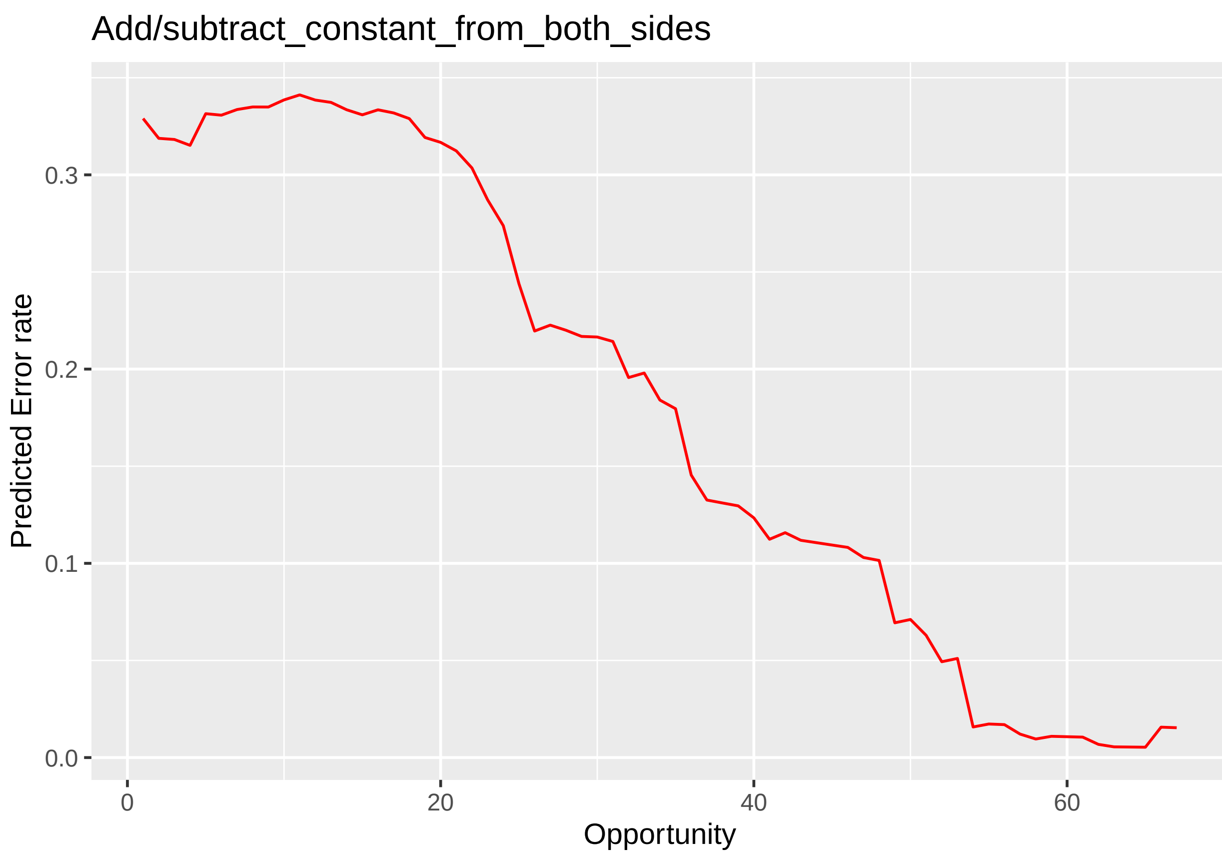
```r
  HCI5 <- HCI_subset[rows5,]
  if(nrow(HCI5) > 0){
    tutor_oppo5 <- min(as.numeric(as.character(HCI5[HCI5$IfTutor == 1,]$Opportunity..Default.)))
    HCI_subset[rows5, ]$TutorTime <- tutor_oppo5
  }
  # 4th KC: 'Combine_variable_terms'
  rows4 <- grepl('Combine_variable_terms', HCI_subset$KC..Default.)
  HCI4 <- HCI_subset[rows4,]
  if(nrow(HCI4) > 0){
    tutor_oppo4 <- min(as.numeric(as.character(HCI4[HCI4$IfTutor == 1,]$Opportunity..Default.)))
    HCI_subset[rows4, ]$TutorTime <- tutor_oppo4
  }
  # 3rd KC: 'Combine_constant_terms'
  rows3 <- grepl('Combine_constant_terms', HCI_subset$KC..Default.)
  HCI3 <- HCI_subset[rows3,]
  if(nrow(HCI3) > 0){
    tutor_oppo3 <- min(as.numeric(as.character(HCI3[HCI3$IfTutor == 1,]$Opportunity..Default.)))
    HCI_subset[rows3, ]$TutorTime <- tutor_oppo3
  }
  # 2nd KC: 'Add/subtract_variable_from_both_sides'
  rows2 <- grepl('Add/subtract_variable_from_both_sides', HCI_subset$KC..Default.)
  HCI2 <- HCI_subset[rows2,]
  if(nrow(HCI2) > 0){
    tutor_oppo2 <- min(as.numeric(as.character(HCI2[HCI2$IfTutor == 1,]$Opportunity..Default.)))
    HCI_subset[rows2, ]$TutorTime <- tutor_oppo2
  }
  # First KC: 'Add/subtract_constant_from_both_sides'
  rows1 <- grepl('Add/subtract_constant_from_both_sides', HCI_subset$KC..Default.)
  HCI1 <- HCI_subset[rows1,]
  if(nrow(HCI1) > 0){
    tutor_oppo1 <- min(as.numeric(as.character(HCI1[HCI1$IfTutor == 1,]$Opportunity..Default.)))
    HCI_subset[rows1, ]$TutorTime <- tutor_oppo1
  }
  # Insert Tutortime to HCI
  HCI_full[rows,]$TutorTime <- HCI_subset$TutorTime
}
```

Adding a column showing pre and post tutor observations

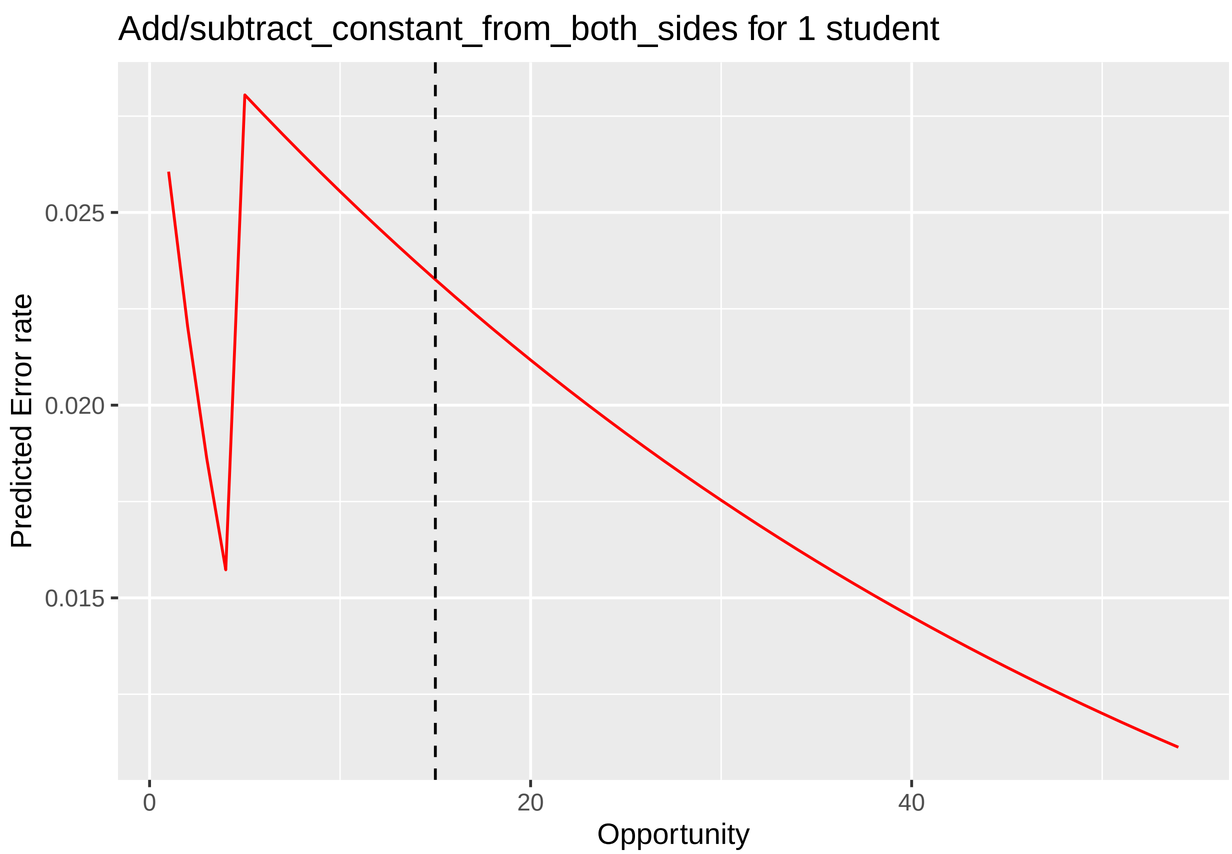
```r
HCI <- read.csv('HCI_final.csv')
tutor.indicator <- rep(0, nrow(HCI))
for (i in 1:nrow(HCI)){
  if(as.numeric(as.character(HCI$Opportunity..Single.KC.[i])) < HCI$TutorTime[i])
    tutor.indicator[i] <- 0
  if(as.numeric(as.character(HCI$Opportunity..Single.KC.[i])) >= HCI$TutorTime[i])
    tutor.indicator[i] <- 1
}
HCI$Post <- tutor.indicator
```

AFM for Add/subtract__constant__from__both__sides

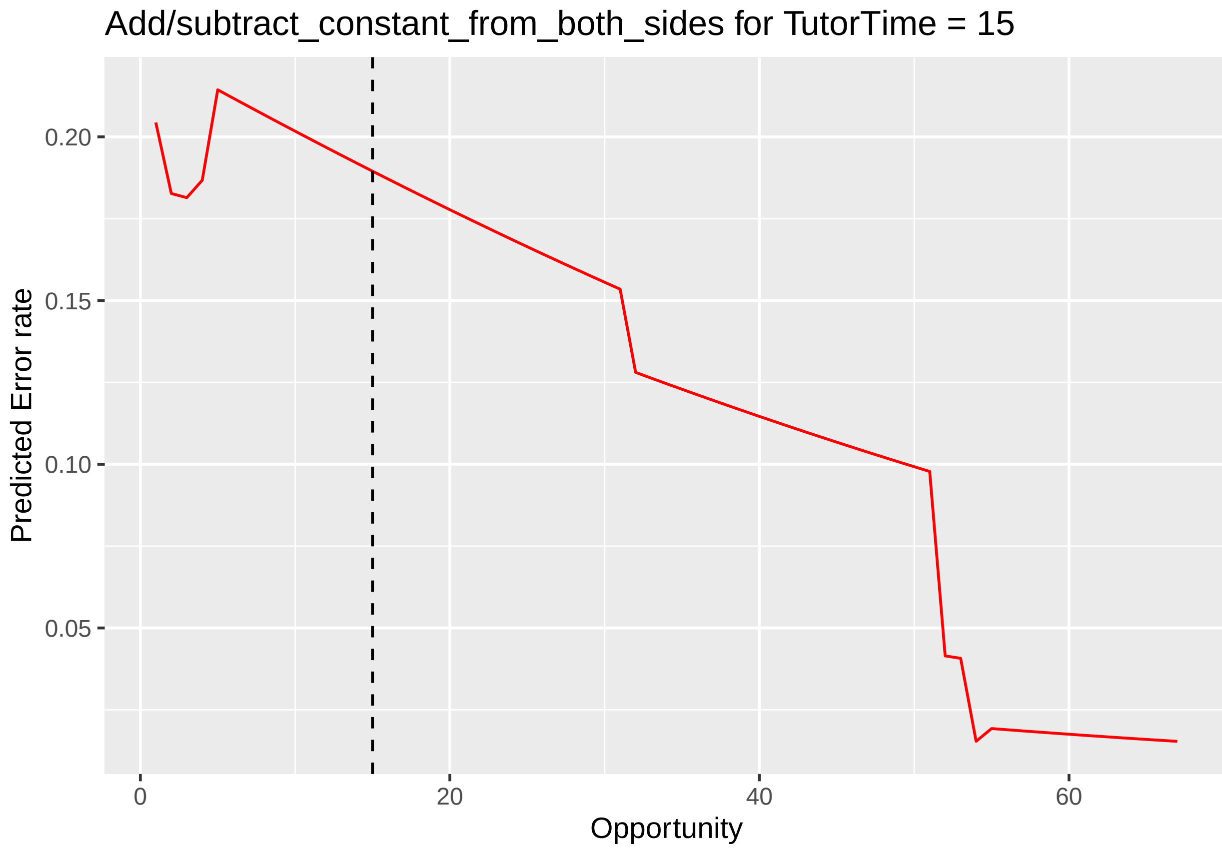
```
HCI1 <- HCI %>%
  filter(grepl('Add/subtract_constant_from_both_sides', KC..Default.))
L1 = length(HCI1$Anon.Student.Id)
Success1 = vector(mode="numeric", length=L1)
Success1[HCI1$First.Attempt=="correct"]=1
Oppo_num <- as.numeric(as.character(HCI1$Opportunity..Default.))
Oppo_character <- as.character(HCI1$Opportunity..Default.)
rows_multi <- which(grepl('~~', Oppo_character))
oppo_first <- rep(0, length(rows_multi))
for (i in 1:length(rows_multi)) {
  string <- strsplit(Oppo_character[rows_multi[i]], split = '~~')
  oppo_first[i] <- as.numeric(string[[1]][1])
}
Oppo_character[rows_multi] <- oppo_first
Oppo_num <- as.numeric(Oppo_character)
HCI1$Opportunity_Numeric <- Oppo_num

AFM1 <- glm(Success1 ~ Anon.Student.Id + Opportunity_Numeric + Opportunity_Numeric:Post, family=binomial
pred1 <- predict(AFM1, HCI1, type="response")
HCI1$Pred <- pred1

AFM1_new <- glm(Success1 ~ Anon.Student.Id + Opportunity_Numeric*Post, family=binomial(), data= HCI1)
pred1_new <- predict(AFM1, HCI1, type="response")
HCI1$Pred <- pred1_new

#df1 <- as.data.frame(cbind(as.numeric(as.character(HCI1$Opportunity..Default.)), pred1)) %>%
#  group_by(V1) %>%
#  summarise(error= 1 - mean(pred1))

df1 <- data.frame(Opportunity = HCI1$Opportunity_Numeric, Pred = HCI1$Pred) %>%
  group_by(Opportunity) %>%
  summarise(error= 1 - mean(Pred))


# Pick one student
HCI1_stu1 <- HCI1 %>%
  filter(Anon.Student.Id == 'Stu_e76a9a1b8a3445439f1eb05d2a79adb6')
#df1_stu1 <- as.data.frame(cbind(as.numeric(as.character(HCI1_stu1$Opportunity..Default.)), HCI1_stu1$P

# Pick students with Tutortime = 15
HCI1_stu15 <- HCI1 %>%
  filter(TutorTime == 15)
df15 <- data.frame(Opportunity = HCI1_stu15$Opportunity_Numeric, Pred = HCI1_stu15$Pred) %>%
  group_by(Opportunity) %>%
  summarise(error= 1 - mean(Pred))


ggplot(df1, aes(x= Opportunity, y = error))+
  geom_line(col = 'red')  +
  labs(title = 'Add/subtract_constant_from_both_sides',x="Opportunity", y="Predicted Error rate")
```




```
ggplot(HCI1_stu1, aes(x= Opportunity_Numeric, y= 1 - Pred))+
  geom_line(col = 'red')  +
  geom_vline(xintercept = HCI1_stu1$TutorTime[1], linetype="dashed") +
  labs(title = 'Add/subtract_constant_from_both_sides for 1 student ', x="Opportunity", y="Predicted Er
```




```
ggplot(df15, aes(x= Opportunity, y = error))+
  geom_line(col = 'red')  +
  geom_vline(xintercept = 15, linetype="dashed") +
  labs(title = 'Add/subtract_constant_from_both_sides for TutorTime = 15', x="Opportunity", y="Predicte
```


Add/subtract_constant_from_both_sides for TutorTime = 15
Predicted Error rate
0.20
0.15
0.10
0.05
0
20
40
60
Opportunity