SNSF Report

Chiara Gilardi, Leslie O'Bray, Carla Sch?rer and Tommaso Portaluri

13 June 2018

Introduction

The Swiss National Science Foundation (SNF) is a research funding agency which disseminates yearly, on behalf of the Swiss Government, billions of CHF to the best researchers in Switzerland. This report contains a statistical analysis performed on three data sets provided by SNF, containing information on the applications for funding received in 2016, the corresponding and the scores given by both internal and external evaluators.

The analysis performed for SNF had a three-fold aim, corresponding to the following three research questions: 1) Is gender bias occurring at any stage of the SNSF evaluation process? Is the gender of the main applicant influencing the rating of the application? 2) To what extent the different steps of the evaluation and the different criteria within each step determine the final funding decision? 3) When an application is approved, but the budget requested is cut, how can we explain this?

The SNSF evaluation procedure is a multi-step process (involving external reviewers, internal referees, and an internal board) which takes into consideration both the track record of the applicant and the quality of the project (see Appendix for a more detailed description of the evaluation procedure).

Several studies (Witteman et al., 2017; Solans-Domenech et al., 2017) have shown that female applicants' projects get higher score when the application is blinded. Moreover, female applicants receive usually higher grades for projects and lower grades for track record. Hence, after investigating the gender dimension to identify possible biases in the evaluation procedure, the focus of the analysis will be the relative importance of of the criteria for funding (applicant's track record vs. quality of the proposal) and, also, of each step of the evaluation procedure (which opinion is more likely to determine the final decision - the external referee's or the board's?). Possible interactions between the gender dimension and the second research question will also be investigated (for instance, by taking into account also the gender of evaluator or the percentage of female referees).

Data Description

We have three data sets: Applications, External Reviewers and Internal Referees. They contain respectively information about the SNSF project funding applications, the evaluation of the applications by external peer reviewers and the evaluation of the proposals by external the internal referee and co-referee (when available). For a full description of the data & variables, please see the Appendix.

Cleaning the Data

We decide to work with only complete applications, i.e. project for which we have information from all the three data sets.

To avoid a temporal trend, we are only considering application from 2016.

In both the external and internal step, we encountered applications which had several reviews per application. For the sake of our analysis, in these scenarios we computed the mean grade for each criteria, so that each application had a "single" score for each criteria assessed on. In doing so we also introduced a new variable, PercentFemale, which calculated the percent of female reviewers out of all reviewers of a single application (ranging from 0 to 1).

All applications with a grade were converted to an ordinal factor.

Specific to each data set, this are the detailed considerations:

Applications

We decide to consider only the MainDiscipline2 because for MainDiscipline we have 118 levels, while for the other only 21.

There is one application for which we do not know the gender of the applicant, and therefore we decided to omit that observation from the analysis.

We will also not consider the variables "CallTitle", "Professorship", "AcademicAge". The first one, because we consider it has nothing to add to the model. The two last, due to the fact that there are a considerable number of NA's on those variables (around 93% of the observations).

Type	Frequency
Assistant professor with tenure track	102
Assistant professor without tenure track	54
Associate professor	237
Full professor	512
Honorary professor or Titular professor	77
None	430
Professor at UAS / UTE	74
Visiting professor	4
NA	20037

External Reviewers

Reviewers always have the option to choose not to consider or to give the grade "0" when reviewing an application. Some might be mistakes, in others cases there might be a conflict of interest, or they might be very ambivalent about the project. Therefore, we did not considered observations with this grades.

One of the questions evaluated in the applications is "Broader impact (forms part of the assessment of scientific relevance, originality and topicality)". For the time frame we are considering, in all the applications this grade was NA. Hence, we omit this variable from our model.

- **ProposalCombined**: We created a new variable to summarize the assessment of the scientific proposal in the external review step. This is a simple mean of the grade given for Suitability and Scientific Relevance. This helped to isolate the effect of the grade given to the scientific proposal, versus the applicant track record, as well as to ensure easier comparison with the internal review step.
- **PercentFemale**: As previously mentioned, we introduced a new variable calculating the percent of reviewers of each application that is female.

Internal Referees

There were 22 observations (1 for the time frame we are dealing with) for which only demographic information was available, no grades were given. We decide to omit those observations.

Also we decide to not consider the Referee role as a variable in our model, as the majority of the evaluations has only one referee.

Type	Frequency
Applicant	2
Explicit inclusion	20
Recusal	8
Referee	15766
Second referee	870
NA	1276

- **PercentFemale**: As previously mentioned, we introduced a new variable calculating the percent of reviewers of each application that is female.

Exploratory Analysis

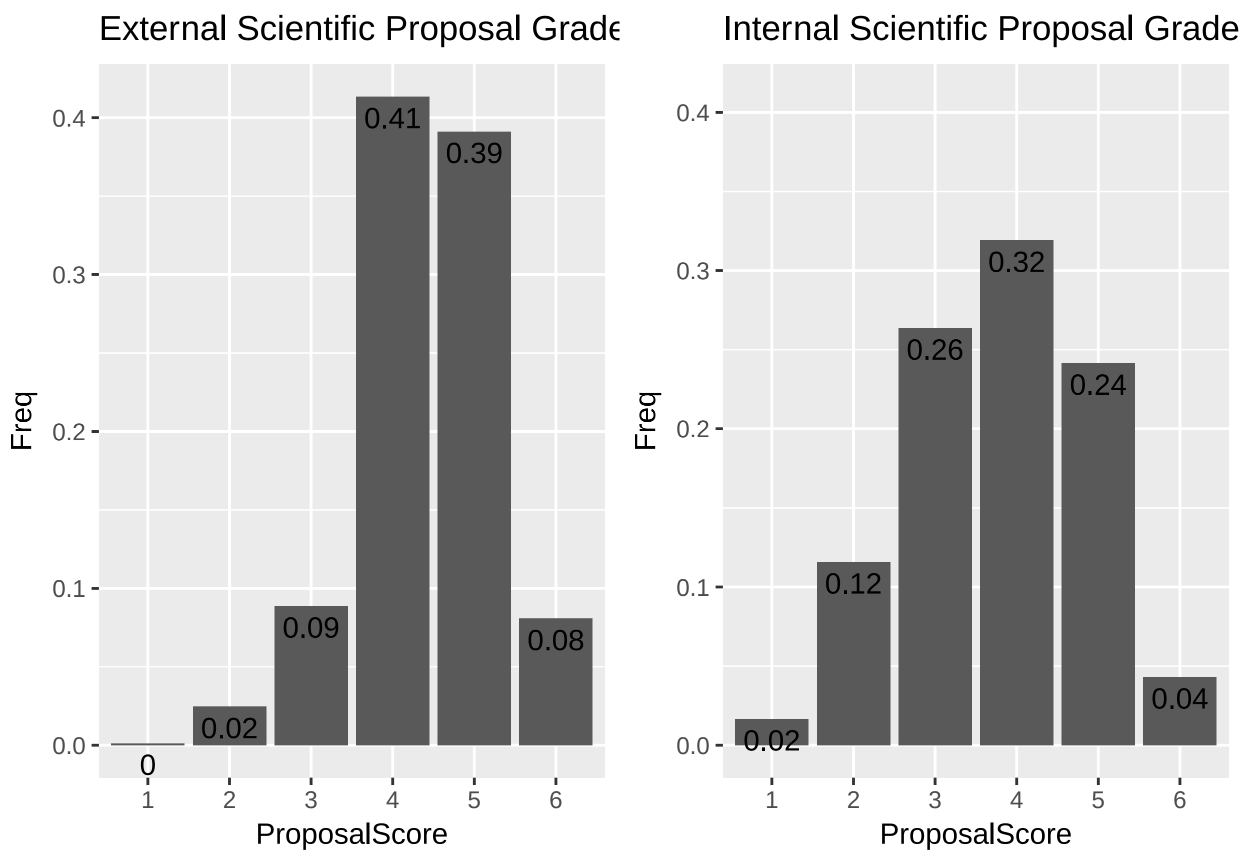
In our exploratory analysis, we discovered a few interesting insights, that relates to the findings we will discuss from our analysis.

Distribution of Grades between the External & Internal Review Step

Since the external & internal step both assess candidates on the same criteria (the strength of the scientific proposal, and the strength of the applicant), on the same ordinal scale (from poor to outstanding), we were interested to see if the distribution of grades are the same. We would expect different distributions for the Overall Grade vs the Ranking, since those have two different measurements, however we were interested to see if for the same absolute ranking, the external and internal reviewers had different perspectives on the application. After combining the Suitability & Scientific Relevance grades given to a candidate in the external review step, we can compare the average grade given for the Scientific Proposal in the two steps, as well as the grade given for the Applicant Track Record in both steps.

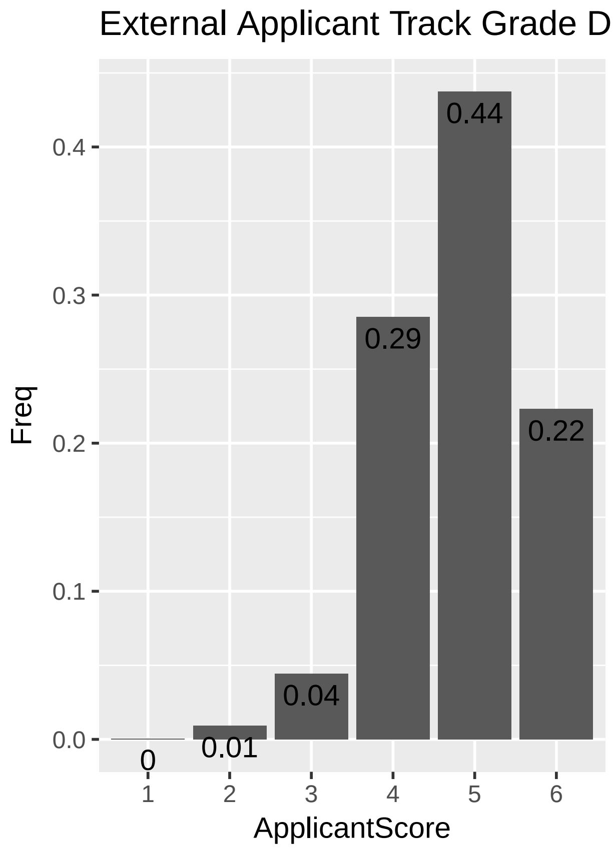
We see that the External Reviewers are more generous with their grades; for the strength of the Scientific Proposal, 48% of proposals are considered "excellent" or "outstanding", versus only 28% in the internal review step.

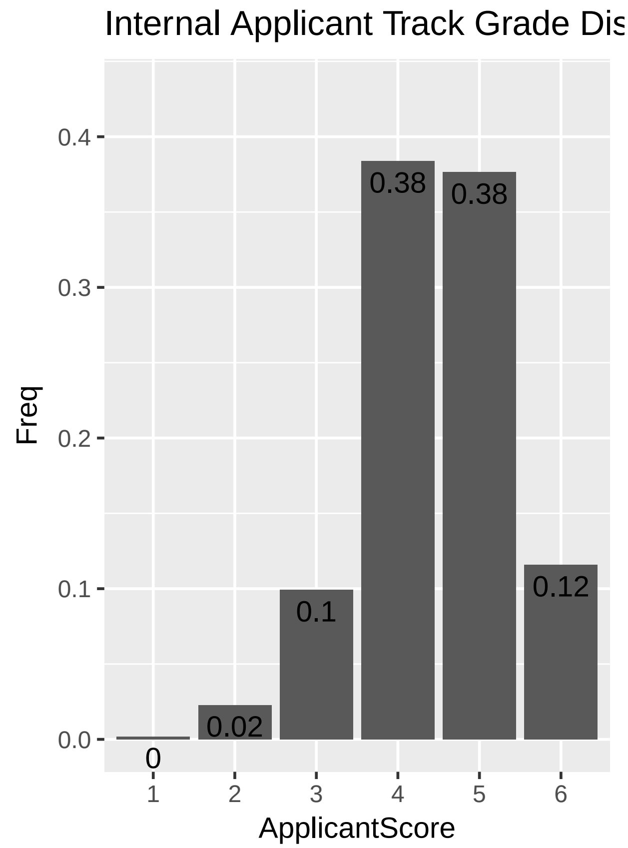
```
<ScaleContinuousPosition>
 Range:
 Limits:    0 -- 0.45
```



Similarly we see the same pattern with Applicant Track Record: 66% of Applicant Track records are considered “excellent” or “outstanding” by the External Reviewers, versus merely 50% by the Internal Reviewers.

```
<ScaleContinuousPosition>
 Range:
 Limits:    0 -- 0.45
```





Since we noticed this discrepancy, we wanted to quantify how differently the grades were to one another. To assess the agreement between the two steps, for the same criteria, we used Cohen's Kappa. Cohen's Kappa measures the proportion of agreement between two raters assessing something on an ordinal scale, accounting for the fact that there will always be some proportion by random chance. An important specification of Cohen's Kappa is the weight given to the measurements. If the external & internal reviewers both assessed the Applicant Track Record as "excellent", that would be considered full agreement. However, we want to allocate partial credit if the rating is a level close to it. We used a linear weight up to distance 2, and after that gave no credit. (In this example, if one rater gave an "outstanding" or "very good", that would be considered a distance of one and be weighted by 0.8. If the second rater assessed the Applicant Track to be "good", which is a distance of two away from excellent, that would be weighted as 0.6). Anything with a distance of 3 or more (in this example, if the second rater gave a rating of "average"), we allocated no weight, as the difference between average and excellent is quite large.

From this, we found that there was just moderate agreement between the two steps when using the weighted kappa, for both the grades given for the Scientific Proposal & the Applicant Track Record.

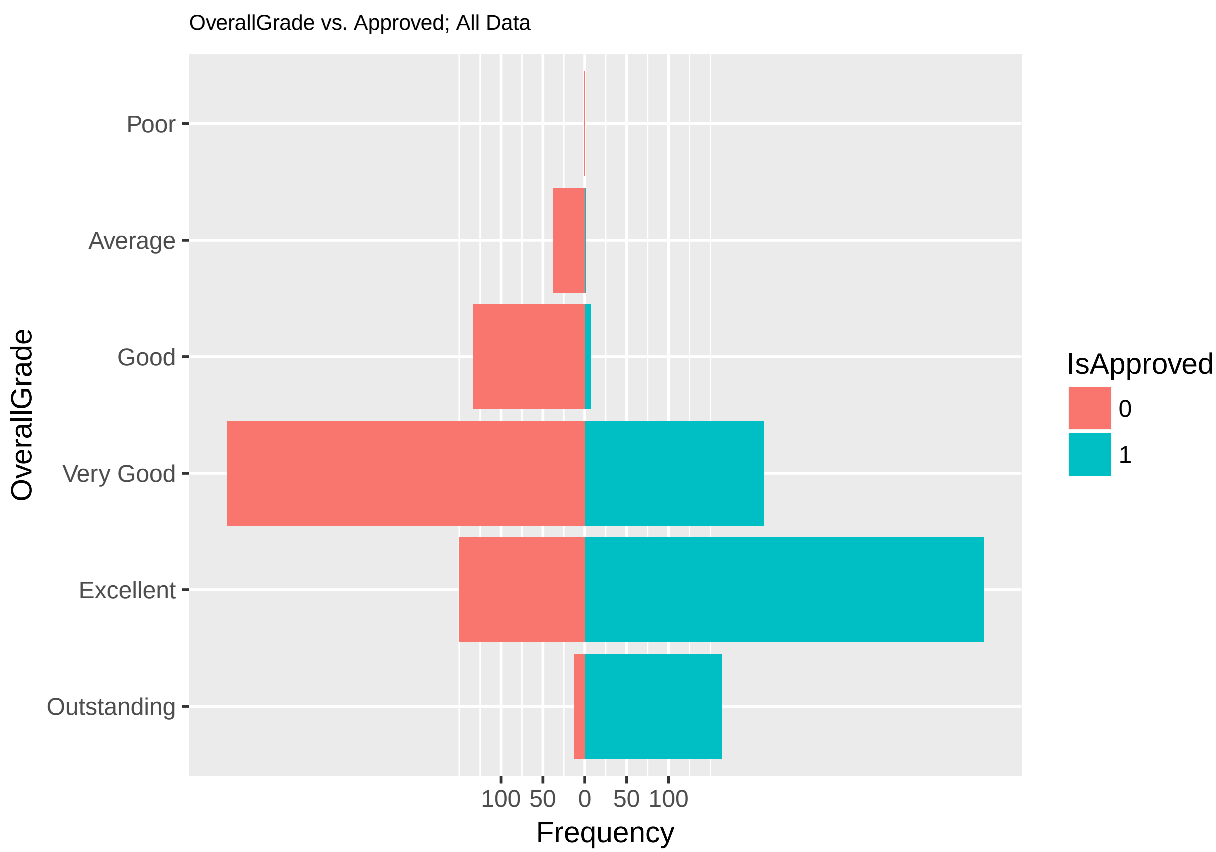
```
[1] "Cohen's Kappa for Applicant Track Record"

Call: cohen.kappa1(x = x, w = w, n.obs = n.obs, alpha = alpha, levels = levels)

Cohen Kappa and Weighted Kappa correlation coefficients and confidence boundaries
                 lower estimate upper
unweighted kappa  0.21     0.24  0.28
weighted kappa    0.35     0.42  0.49

 Number of subjects = 1623

[1] "Cohen's Kappa for Scientific Proposal"
```

```
Call: cohen.kappa1(x = x, w = w, n.obs = n.obs, alpha = alpha, levels = levels)

Cohen Kappa and Weighted Kappa correlation coefficients and confidence boundaries
                 lower estimate upper
unweighted kappa  0.21     0.24  0.28
weighted kappa    0.35     0.42  0.49

 Number of subjects = 1623
```

Impact of Internal Reviewers on Funding

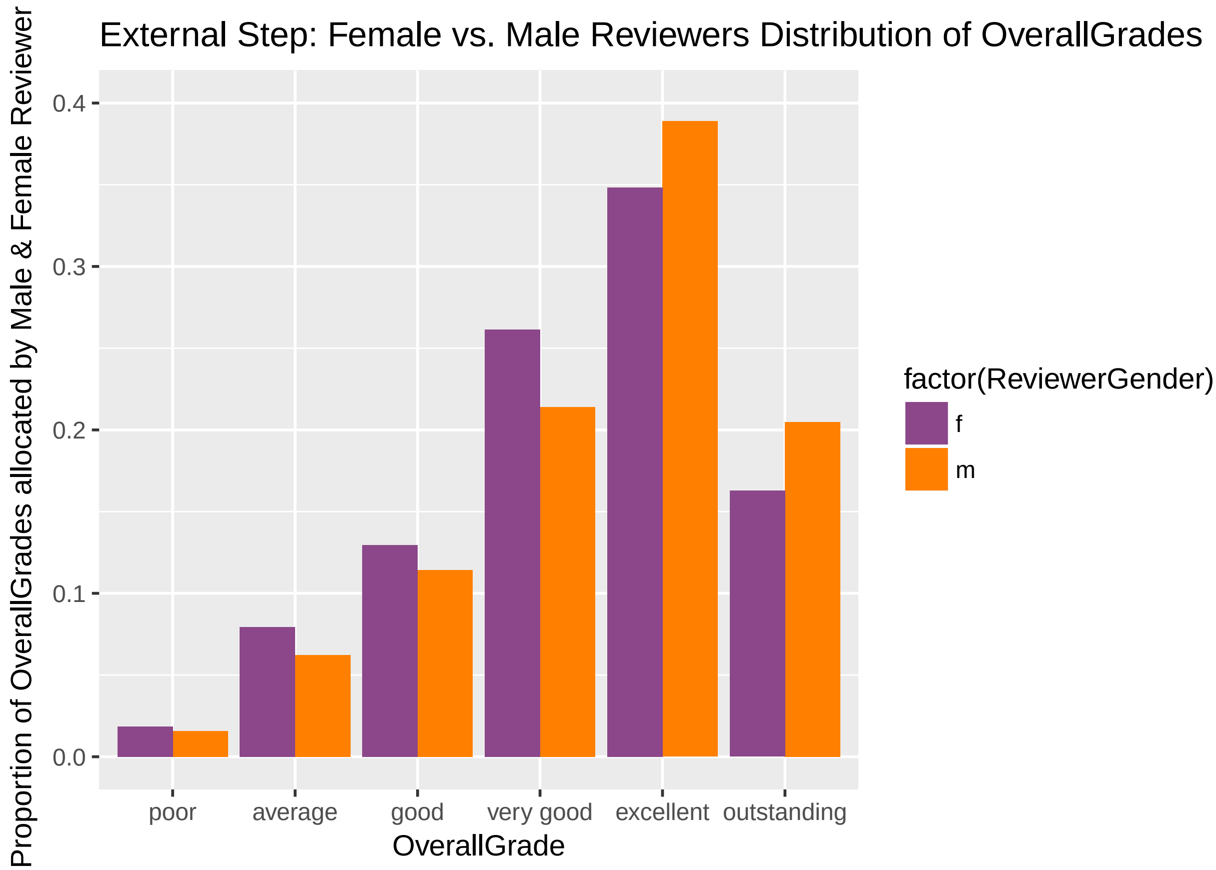
We wanted to understand if this discrepancy between grades had an impact on whether an application is funded. To do this, we visualized the summary grade given to an application, and whether that application is funded or not. As we can see here, there are several applications with an OverallGrade of "excellent" or "outstanding" that end up not approved. It highlights that not only do the Internal Reviewers give tougher grades in general than the external step, but they also consider some "excellent" and "outstanding" applications by the external reviewers to be not of the quality that deserves funding. This trend is true in all divisions and both genders, please refer to the appendix to see the specific graphic.

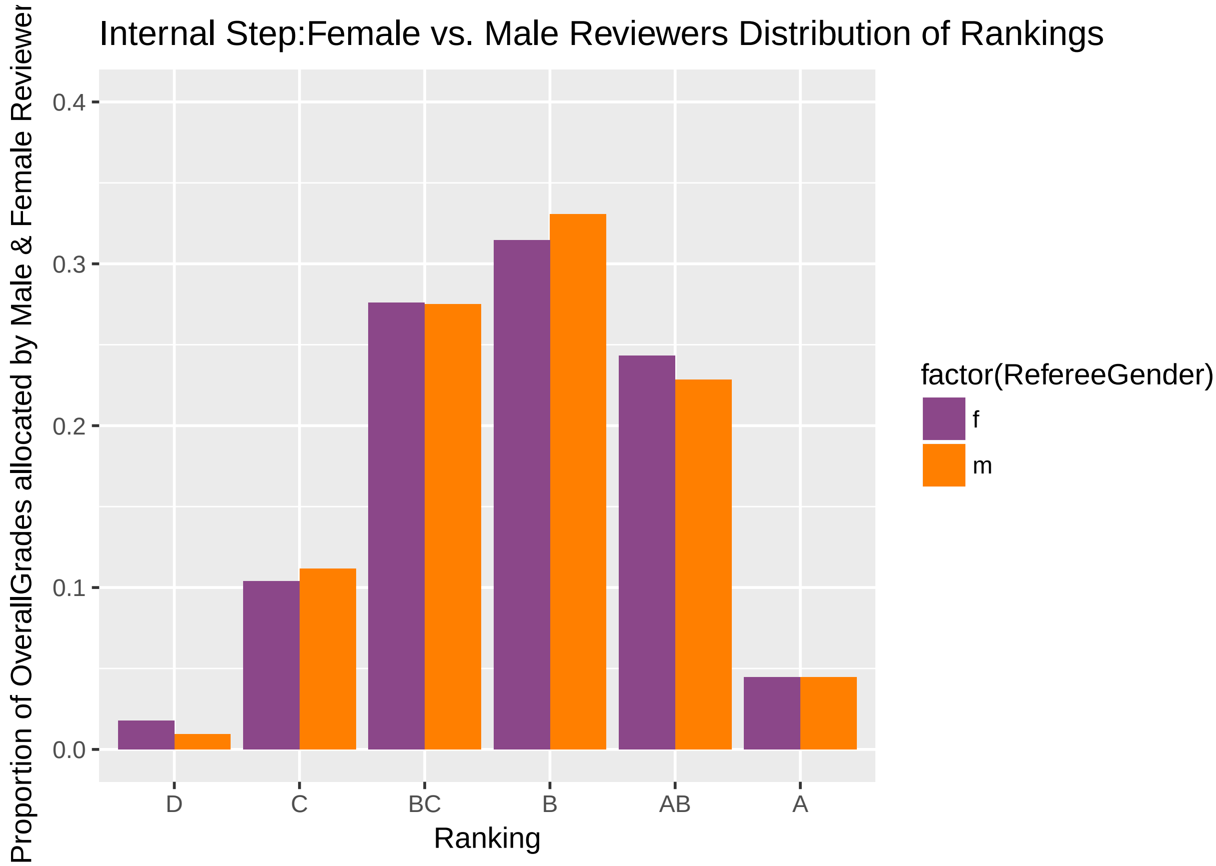
Our conclusion for this is that the internal step is very consequential, and the difference in the rating they give translates into differences in whether an application gets funded or not.



Distribution of Grades by the Gender of the Reviewer

The third interesting insight we found was when we investigated the impact of the gender of the person reviewing the data. We look at the relative frequencies of grades given by male and female reviewers, to applicants, regardless of gender. Within the external step in particular, we found that female reviewers give proportionally fewer "excellent" and "outstanding" grades, compared to their male counterparts. Within the internal step, we did not notice a particular difference, though we will consider the impact of the gender of the reviewer more rigorously in our analysis.





Gender Bias

To see if gender has an influence in any of the steps of the evaluation process, we did several things. For the external and internal steps, we first fit a logistic regression with the function glm in R, where we used IsApproved (a binary variable) as a response and demographic information of the applicant, project information and the given grades as predictors. The aim of this regression is to see if gender has an influence on the final decision from the perspective of each step.

As the final decision is determined by the different grades in the process, in order to see if gender has an influence on any of them, we fitted an Ordinal regression with the function clm of the package "ordinal" on each grade with demographic data and project information as predictors.

Finally, we estimated the relative importance of each of the predictors in the model as a last check, in order to see the importance of gender in all the models.

Analysis

External Step

Logistic Regression

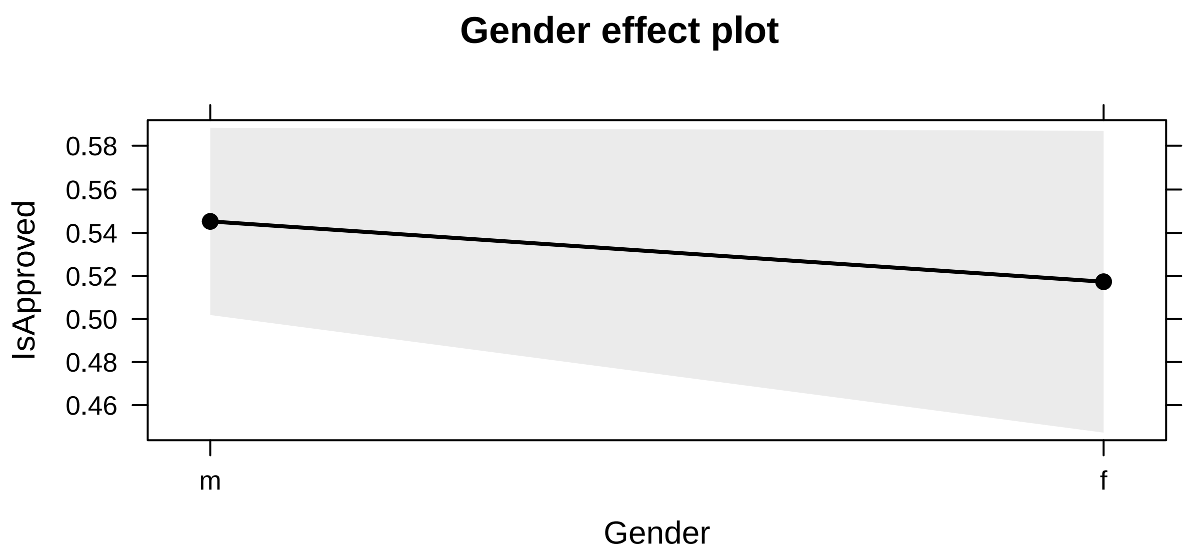
Regression data: To perform the analysis, we combined in one data frame information about the applications (IsApproved, Age, Gender, Division, IsContinuation, PreviousRequest, InstType, log(AmountRequested), Semester) and about the grades given by the external reviewers (ApplicantTrack, ScientificRelevance, Suitability, OverallGrade, ProposalCombined, PercentFemale).

As there are almost no application approved that have grades smaller than “good”, we decide to aggregate grades “poor”, “average” and “good” to avoid perfect separation problems. All grades are considered as ordered factors.

We first fitted a full model with all the variables and the interactions between Gender and Division, PercentFemale and ApplicantTrack. Also we considered the interaction between InstType and Division. we didn’t considered OverallGrade as it is highly correlated with the grades of the applicant and the project. we achieved a pseudo-R^2 value of 0.4225, indicating that percent of the variation in Y can not be explained very well by the model.

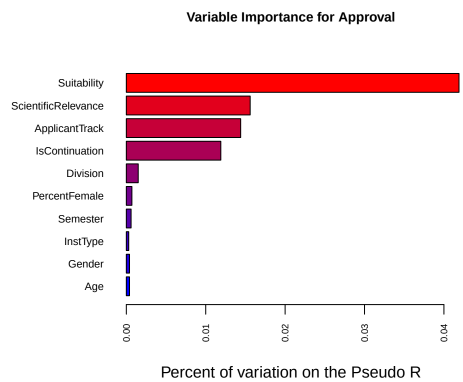
When selecting the variables with the AIC criteria in order to work with a small and effective model, we end up with the following predictors: ApplicantTrack, ScientificRelevance, Suitability, PercentFemale, Age, Gender, Division, IsContinuation, InstType, Semester, Division:InstType. No interaction with Gender where significant. The pseudo R^2 for this model is 0.4203 , i.e. this smaller model explains basically the same variance of the data than the former one. Non of them reveal that the information from the external reviewers explain the final decision correctly. This fact will be explored in more detail later on.

As we are interested in the effect of gender in each step of the evaluation process, we looked at the difference of the probability of being accepted between male and female given the final model. Female have a probability 0.517, while male have 0.545. Although there is a small difference, our analysis suggest that this difference is not significant.



In order to understand which of the variables in this model are the more influential, we ran a permutation test. We randomly permuted the values of all predictors (one at the time) and refit the the model. We did this 1000 times for each predictor, and calculated the average change in pseudo R^2 when we permuted that particular variable. If permuting a variable changes the pseudo r^2 a lot, this means that that variable was an important predictor in our regression. As shown in the next table, Gender is the second least important variable in this model.

	Feature	Importance
5	Age	0.0004
6	Gender	0.0004
8	InstType	0.0003
10	Semester	-0.0006
4	PercentFemale	-0.0007
7	Division	-0.0015
9	IsContinuation	-0.0119
1	ApplicantTrack	-0.0144
2	ScientificRelevance	-0.0156
3	Suitability	-0.0419



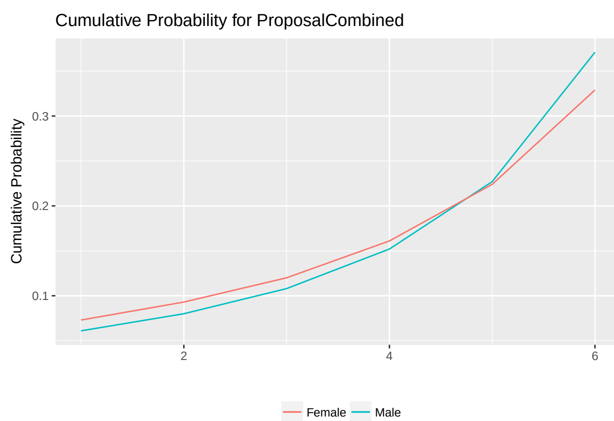
Ordinal Regression

To look into the variables that influence the different grades in this part of the evaluation process, and see if the gender of the main applicant has an influence on it, we ran an ordinal regression with the function clm of the package ordinal in R. We did this for both, the applicant grade (ApplicantTrack), and the average of the grades given to the project (ProposalCombined). Also, we want to investigate which grades are more influential in the Overall Grade, which is the summary grade given by the external reviewers. In the next few paragraphs we will look into the analysis of this two regressions.

- **Project Assessment**: After fitting a full model with ProposalCombined as a response variable and different interactions, and then selecting from this model the significant variables with the AIC criteria and the help of the drop1() function in R, we end up with a model with the following predictors: Gender, Division, PercentFemale, IsContinuation, InstType and log(AmountRequested). If we fit the same model without Gender and compare it with the anova() function to the one with gender, we get a p.value of 0.1064133, meaning that for the grades given to the project, gender is not important. This is to be expected, as the project is being evaluated and not the applicant.

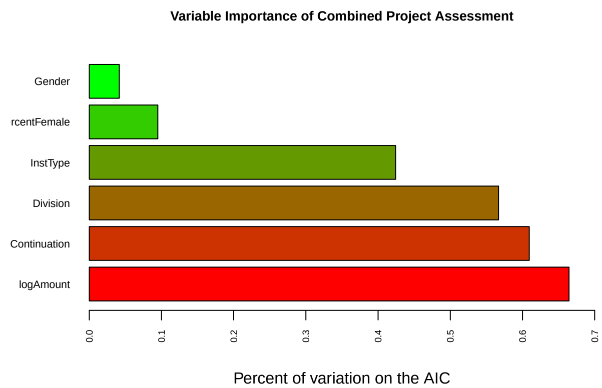
  Nevertheless we kept gender as a predictor to be able to show the effect it has in the grades. For this purpose we estimated the cumulative probability of falling in the different grades for each gende. The result is presented in the next table. Overall the average difference is 0.015. We see here as well, that there is no evidence of gender influencing the probability of achiving a certain grade.

	Male	Female	Difference
poor	0.061	0.073	-0.012
average	0.080	0.093	-0.012
good	0.108	0.120	-0.012
very good	0.152	0.161	-0.009
excellent	0.227	0.224	0.003
outstanding	0.371	0.329	0.042



We did also here our permutation test where we permuted each variable 1000 times, and used the percentage of variation of the AIC as a measure of good fit. Gender is the least important variable, in average it increments the AIC 0.04%.

	feature	Diff.AIC	Percent.AIC
6	logAmount	26.095959	0.6644556
4	IsContinuation	23.933805	0.6094028
2	Division	22.260672	0.5668015
5	InstType	16.666777	0.4243697
3	PercentFemale	3.727439	0.0949081
1	Gender	1.629380	0.0414873

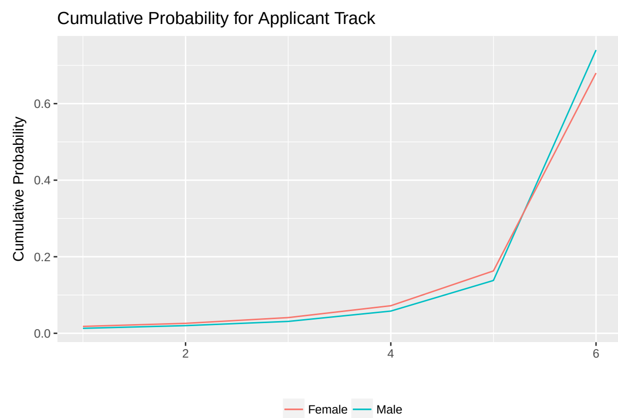


- **Applicant Track assessment**: The final model we used has ApplicantTrack as a response variable, and the following predictors: Gender, Division, PercentFemale, IsContinuation, InstType, log(AmountRequested) and the interaction between Gender and PercentFemale. Again we fitted the same model without Gender and compare it with the anova() function to the one with gender, we get a p.value of 0.0014313, meaning that for the grades given to the main applicant, gender needs to be considered in the model. In the next table we present part of the summary for this model, to see the full summary refer to the Appendix.

	Estimate	Std. Error	z value	Pr(>|z|)
1|2	-0.7523	1.4897	-0.5050	0.6136
2|3	2.0350	1.1325	1.7969	0.0724
3|4	3.8149	1.1106	3.4349	0.0006
4|5	6.1707	1.1115	5.5515	0.0000
5|6	8.2987	1.1208	7.4041	0.0000
Genderf	-0.5133	0.1610	-3.1880	0.0014
DivisionDiv 2	0.5132	0.1407	3.6482	0.0003
DivisionDiv 3	-0.1914	0.1297	-1.4754	0.1401
PercentFemale	-0.7975	0.2354	-3.3886	0.0007
IsContinuation1	0.4731	0.1173	4.0330	0.0001
InstTypeOther	-0.5593	0.2364	-2.3657	0.0180
InstTypeUAS/UTE	-1.2653	0.2029	-6.2369	0.0000
InstTypeUni	-0.2382	0.1307	-1.8227	0.0683
logAmount	0.5520	0.0864	6.3861	0.0000
Genderf:PercentFemale	1.0440	0.4142	2.5207	0.0117

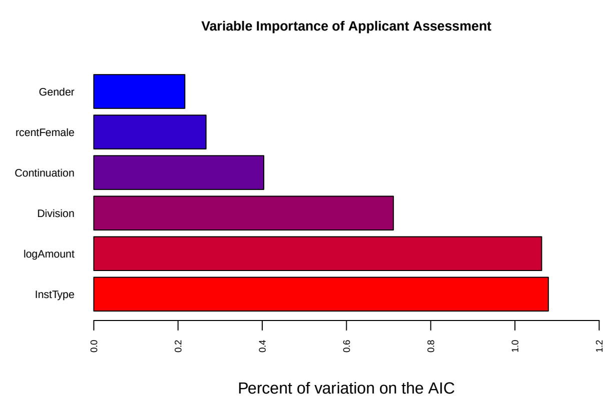
The predicted probabilities of achieving certain grade for male and female is shown in the next table. The average difference of the cumulative probability is here as well close to zero,0.01983, but the greater difference are in grades “outstanding”, which are the ones that are more likely to be approved.

	Male	Female	Difference
poor	0.013	0.018	-0.004
average	0.020	0.026	-0.006
good	0.031	0.041	-0.009
very good	0.058	0.072	-0.015
excellent	0.138	0.163	-0.025
outstanding	0.740	0.680	0.060



Though, gender is significant to the model, when estimating its relative importance in comparison with the other variables, by permuting each variable 1000 times, and using the average percent variation on the AIC as a measure of goodness of fit, turns out that gender is the least important variable in the model, in average it increments the AIC 0.22%.

	Feature	Diff.AIC	Percent.AIC
5	InstType	41.25026	1.0793737
6	logAmount	40.64019	1.0634102
2	Division	27.18788	0.7114108
4	IsContinuation	15.42453	0.4036055
3	PercentFemale	10.18358	0.2664683
1	Gender	8.26095	0.2161599

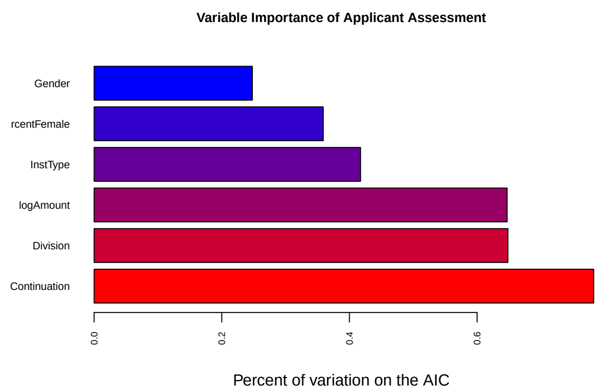


- **Overall Grade**: To see which variable is the most important in determining the OverallGrade, we fitted a model, after variable selection, with OverallGrade as a Response, and predictors: ApplicantTrack, ProposalCombined, PercentFemale, Gender, IsContinuation and interaction between Gender and

PercentFemale. Comparing this model with the one without gender suggest that gender is not significant to the model, p.value of 0.0182634.

Again, we estimate the relative importance of the variables base on the percentage change of the AIC when permuting the variable. We did this permutation 1000 times for each variable. Gender is the second least important variable in the model, in average it increments the AIC 0.42%.

	Feature	Diff.AIC	Percent.AIC
4	IsContinuation	31.304189	0.7827131
2	Division	25.928311	0.6482976
6	logAmount	25.879597	0.6470795
5	InstType	16.694106	0.4174105
3	PercentFemale	14.355237	0.3589306
1	Gender	9.913537	0.2478728



Internal Step

Logistic Regression

Regression data: as explained before, in order to perform the analysis, we combined in a single data frame all the information about applications (IsApproved, Age, Gender, Division, IsContinuation, PreviousRequest, InstType, log(AmountRequested), Semester) and about grades given by the internal referees (ApplicantTrack, ProjectAssessment, Ranking, PercentFemale).

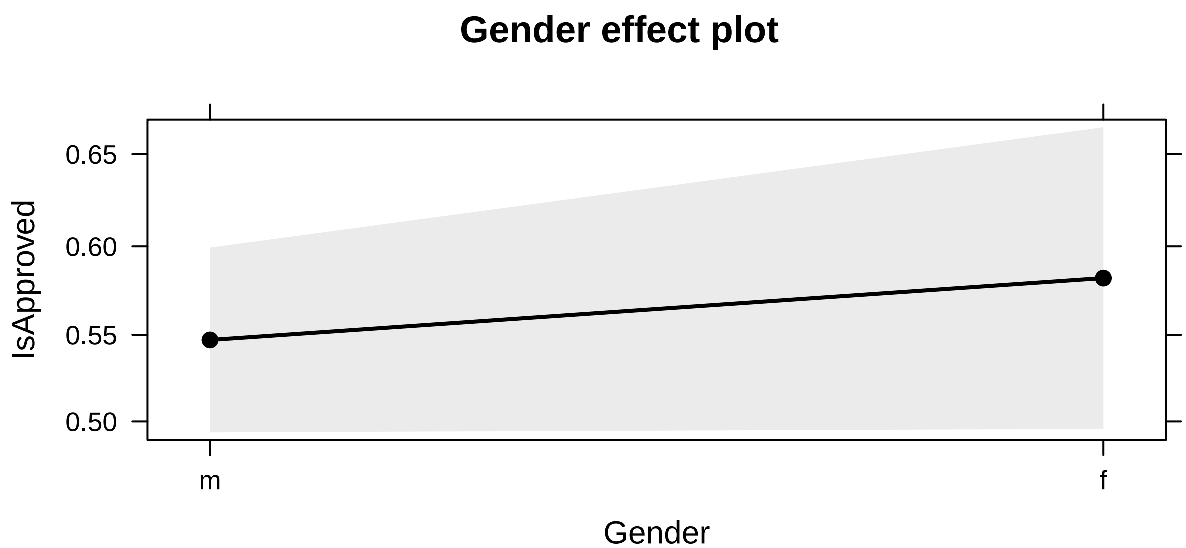
We had again a perfect separation problem, due to the fact that there are very few approved applications with grades worse than "good". We aggregated the grades which were in category "poor", "average" and "good" in a unique category for both the applicant track record and the project assessment. All grades are considered as ordered factors.

We first fitted a full model with all the available variables and the interactions between Gender and Division, Gender and PercentFemale, Gender and ApplicantTrack. Moreover, we considered the interaction between InstType and Division. We didn't include the Ranking grades into the model, since they are highly correlated with the single grades for the applicant and the project. We achieved a pseudo-R^2 value of 0.6923, meaning that the variation in the binary variable Y (approved or not) can be explained for more than half by this model including only the internal step information.

Then, we did variable selection using the AIC criteria in order to obtain a small and effective model as we did for the external step. The remaining predictors are Gender, Age, Semester, IsContinuation, PercentFemale, ApplicantTrack, ProjectAssessment and log(AmountRequested). Interactions were removed from the model because not significant.

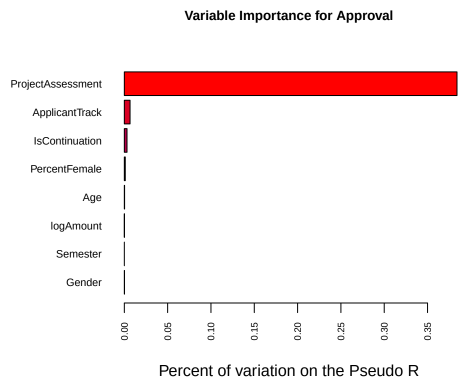
The pseudo R^2 for this model is 0.686 , i.e. this reduced model explains basically the same proportion of variance of the data as the previous model. Even if we removed some variables, the model still explains almost 70% of the variation of the variable IsApproved. It therefore seems that the internal grades are significant predictors for the final funding decision.

As said before, our focus is the effect of gender in each step of the evaluation process. We checked the difference in the probability of being accepted between male and female: women have probability 0.582 of the project being approved, while men have 0.547. Surprisingly, the probability of being funded seems to be higher for female applicants, even if the difference is too small to be significant. We can clearly see that from the Gender effect plot, where the confidence intervals are overlapping and the line is almost horizontal. From this initial analysis we can say that it doesn't seem that referees are biased against women.



We did also a further check through a permutation test, in order to better understand which of the variables are the most influencial in the model. We randomly permuted the values of all predictors (one by one) and refitted the model. We did this 1000 times for each predictor, and calculated the average change in pseudo R^2 when we permuted that particular variable. If permuting a variable reduces the pseudo R^2 a lot, this means that that variable was an important predictor in our regression. As shown in the next table, gender is the least important variable in the model and when we permute its values the pseudo R^2 increases slightly, which means that the model becomes a bit better than before.

	feature	importance
6	Gender	0.0002
8	Semester	0.0000
3	logAmount	-0.0001
5	Age	-0.0003
4	PercentFemale	-0.0011
7	IsContinuation	-0.0031
1	ApplicantTrack	-0.0066
2	ProjectAssessment	-0.3840


Variable Importance for Approval
ProjectAssessment
ApplicantTrack
IsContinuation
PercentFemale
Age
logAmount
Semester
Gender
0.00
0.05
0.10
0.15
0.20
0.25
0.30
0.35
Percent of variation on the Pseudo R

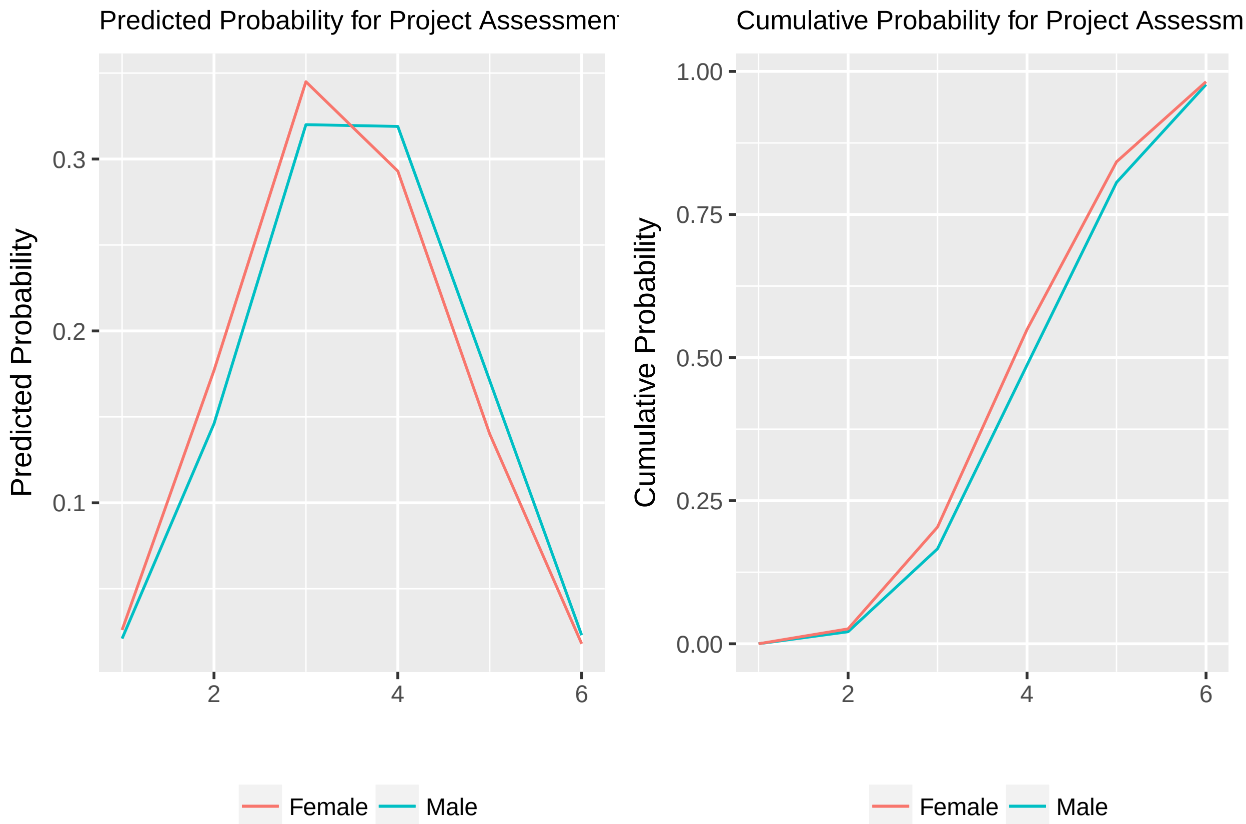
Ordinal Regression

To see which variables influence the different grades in the second step of the evaluation process and check if the gender of the main applicant has an influence on it, we ran an ordinal regression with the function clm of the package ordinal in R. We did this for both, the applicant grade (ApplicantTrack) and grade given to the project (ProjectAssessment). In order to avoid a multiple testing problem, we applied Bonferroni correction: since we performed 3 tests on the same dataset, we multiply the p-value obtained from the Likelihood Ratio Test to the number of tests performed.

- **Project Assessment**: we fitted the full model with ProjectAssessment as response variable and then selected from the significant variables with the AIC criteria and the help of the drop1() function in R. We end up with a model with the following predictors: Gender, Division, PercentFemale, Age, IsContinuation, InstType and log(AmountRequested). If we fit the same model without Gender and compare it to the one with gender with the anova() function, we get a p.value of 0.0756534, meaning that Gender doesn't seem to be a significant predictor for the grades given to the project by the internal referees. We kept gender as a predictor in our model to be able to show the effect it has on the grades. For this purpose, we estimated the probability of falling in the different grades for each gender and the difference between men and women. The result is presented in the following table.

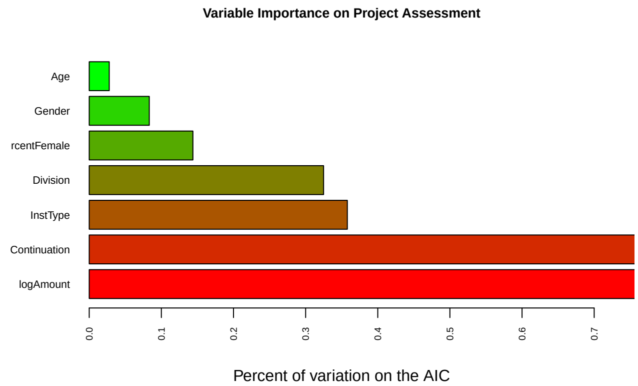
	Male	Female	Difference
poor	0.021	0.026	-0.006
average	0.146	0.177	-0.031
good	0.320	0.345	-0.024
very good	0.319	0.293	0.026
excellent	0.171	0.140	0.031
outstanding	0.023	0.018	0.005

Overall the average difference is really small: 0.0205. This seems to suggest that there is no evidence of gender influencing the probability of achiving a certain grade. We also represented in the plots below the probabilty and cumulative probability curves of getting each grade for male and female: they follow more or less the same trend and the only difference, as we've seen from the table above, is that women are slightly more likely to get a "very good" rather than an "excellent".



We did again a permutation test: we shuffled each variable 1000 times and we used the percentage of variation of the AIC as a measure of goodness fit. Gender is the second least important variable in the model and permuting its values, on average, it makes the AIC increase by 0.08%.

	feature	Diff.AIC	Percent.AIC
7	logAmount	40.525106	0.8558318
4	IsContinuation	40.057715	0.8459611
6	InstType	16.943383	0.3578198
2	Division	15.383656	0.3248806
3	PercentFemale	6.806232	0.1437378
1	Gender	3.943082	0.0832722
5	Age	1.314962	0.0277701

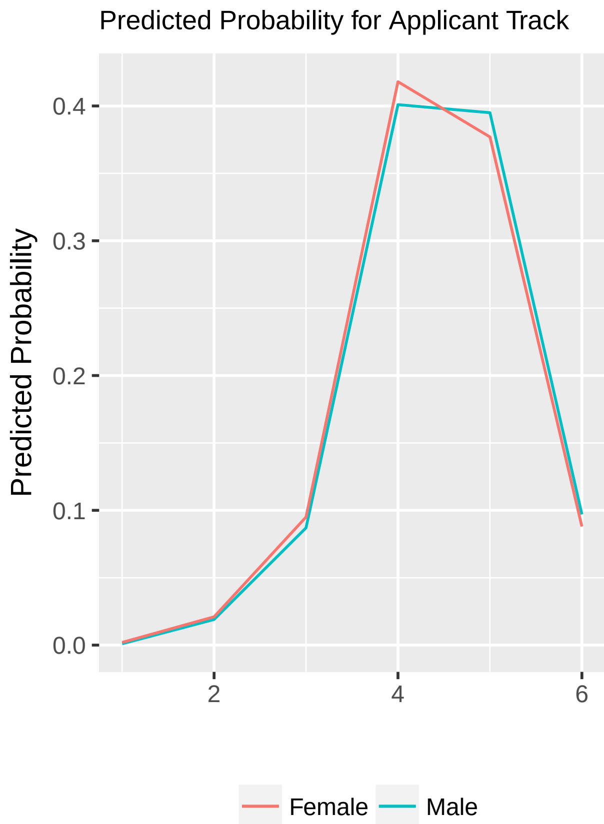


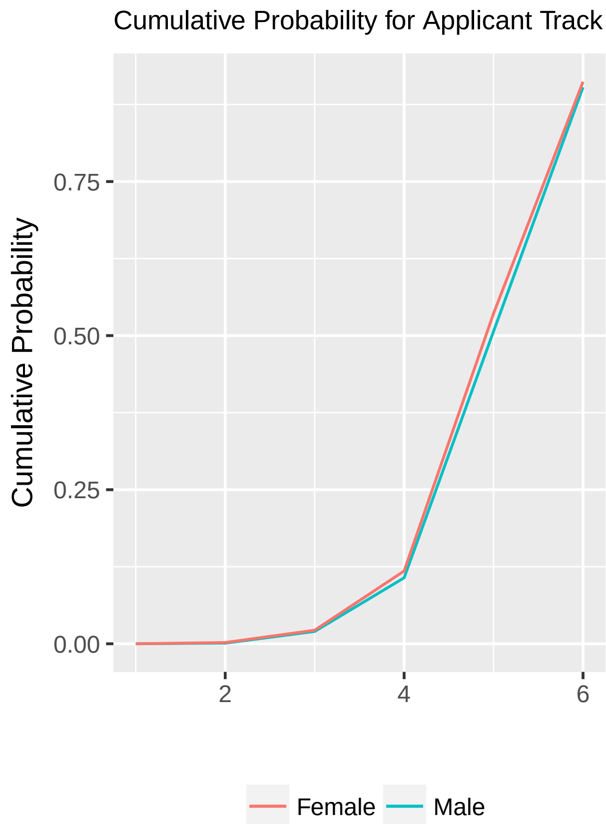
- **Applicant Track assessment**: the model we used has ApplicantTrack as a response variable and the following predictors: Gender, Division, PercentFemale, IsContinuation, InstType, log(AmountRequested), Semester, the interaction between Gender and Divion and the interaction between Division and PercentFemale. Again we fitted the same model without Gender and compare it with the anova() function to the one with gender, we get a p.value of 0.001, meaning that for the grades given to the main applicant track record, gender needs to be considered in the model. Before deciding whether there is enough evidence of gender bias, we did some more analysis.

  We computed the difference in probability of getting a specific grade for both male and female.

	Male	Female	Difference
poor	0.001	0.002	0.000
average	0.019	0.021	-0.002
good	0.087	0.095	-0.009
very good	0.401	0.418	-0.017
excellent	0.395	0.377	0.019
outstanding	0.097	0.088	0.010

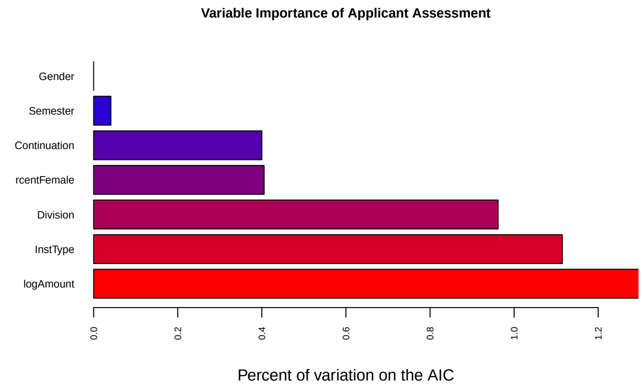
In the table above, we see the probabilty of getting each grade for both male and female and the difference between the two. The average difference of the cumulative probability is here 0.0095, very close to zero. From the plot below we see that there is almost no difference between women and men probabilities. This seems to suggest that the small p-value from the likelihood ratio test is not really reliable to establish whether there is gender bias.





So we proceeded with our analysis estimating the relative importance of the variables included in the model. We used a permutation test, as before, and it turned out that gender is not at all an important variable in the model, in fact when we shuffle the values referring to the variable Gender the AIC doesn't increase.

	feature	Diff.AIC	Percent.AIC
7	logAmount	71.913023	1.7602963
6	InstType	45.535382	1.1146210
2	Division	39.294631	0.9618591
3	PercentFemale	16.564343	0.4054641
4	IsContinuation	16.344027	0.4000712
5	Semester	1.676251	0.0410315
1	Gender	0.000000	0.0000000

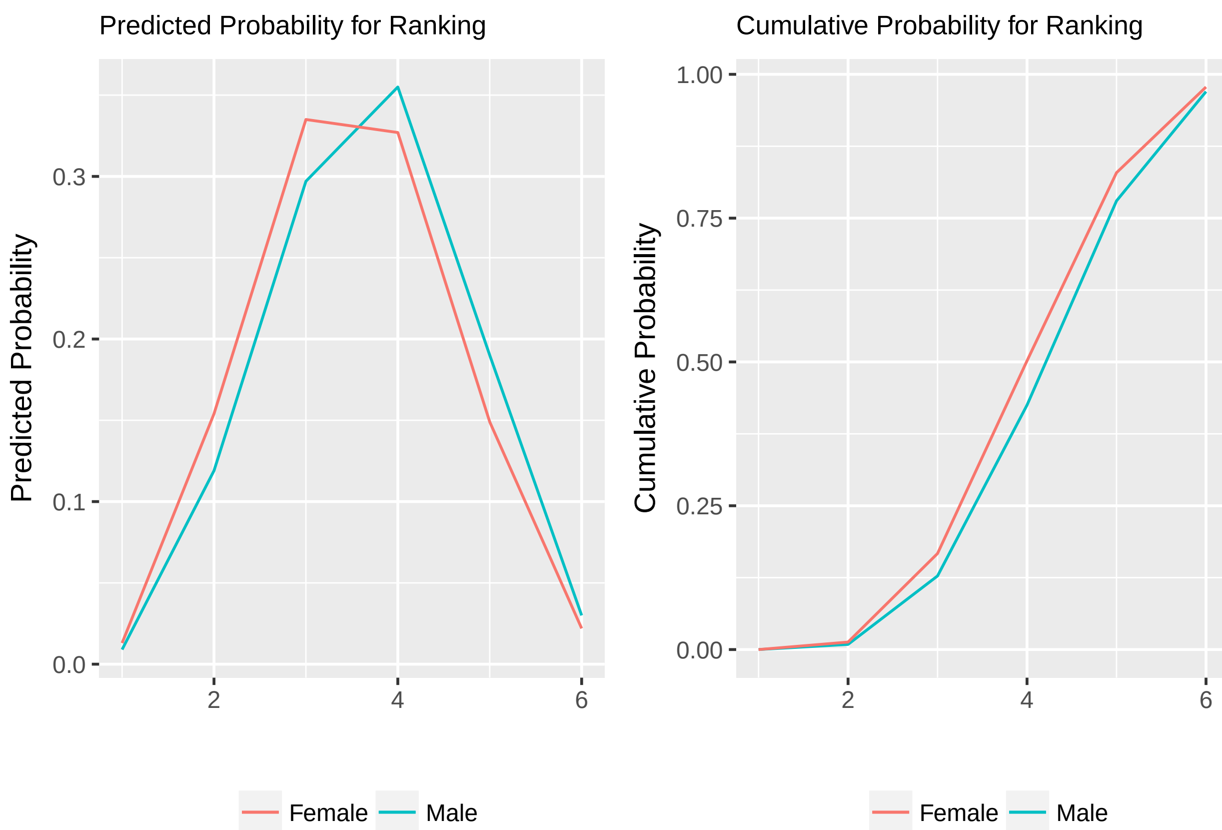


- **Ranking**: This last model has Ranking as a response and Gender, Division, PercentFemale, IsContinuation, InstType, PreviousRequest and logAmount as predictors. We are not considering here the grades given to the applicant track record and to the project, as we just want to see the influence of the demographic data and the project information in each grade. A comparison of this model with the same one without gender may suggest that gender is significant to the model: p.value of 0.014531.

	Estimate	Std. Error	z value	Pr(>|z|)
1|2	2.5331	1.1114	2.2793	0.0227
2|3	5.2874	1.0852	4.8723	0.0000
3|4	6.9031	1.0879	6.3453	0.0000
4|5	8.4717	1.0948	7.7381	0.0000
5|6	10.6965	1.1091	9.6439	0.0000
Genderf	-0.3100	0.1101	-2.8155	0.0049
PercentFemale	0.2667	0.1187	2.2466	0.0247
DivisionDiv 2	0.2367	0.1326	1.7842	0.0744
DivisionDiv 3	-0.2166	0.1258	-1.7221	0.0851
IsContinuation1	0.8434	0.1148	7.3488	0.0000
PreviousRequest1	0.2065	0.1321	1.5626	0.1182
InstTypeOther	-0.7457	0.2270	-3.2844	0.0010
InstTypeUAS/UTE	-0.9866	0.2004	-4.9223	0.0000
InstTypeUni	-0.4358	0.1262	-3.4545	0.0006
logAmount	0.5707	0.0842	6.7783	0.0000

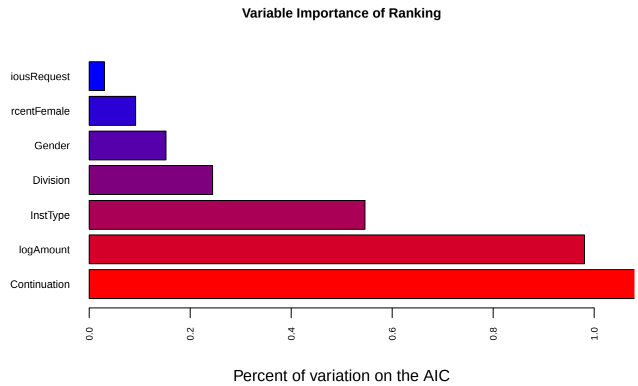
```
            Male  Female Difference
poor        0.009  0.013     -0.003
average     0.119  0.154     -0.035
good        0.297  0.335     -0.038
very good   0.355  0.327      0.028
excellent   0.190  0.149      0.041
outstanding 0.030  0.022      0.008
```

The predicted probabilities of achieving certain grade for male and female is shown in the next table. The average difference of the cumulative probability is here as well close to zero (0.0255). Notice that the only difference is that female applicants are mole likely to get a “good” grade rather than a “very good”, compared to male applicants. From the cumulative probability plot below, we see that the trend is the same for both genders and that the difference is not relevant.



Once more, though gender is significant to the model, when estimating its relative importance in comparison with the other variables, it turns out that gender is one of the least important variable in the model, in average it increments the AIC 0.15%.

	feature	Diff.AIC	Percent.AIC
4	IsContinuation	54.061531	1.1642844
7	logAmount	45.544583	0.9808610
6	InstType	25.354826	0.5460487
2	Division	11.344618	0.2443209
1	Gender	7.057893	0.1520008
3	PercentFemale	4.265595	0.0918651
5	PreviousRequest	1.405770	0.0302751



Results

External Step

The external step model is not a good explanation of the variation in the approval of applications. We nevertheless look into the influence of gender at this stage, but couldn't find evidence of its effect on the final decision.

When looking at the different grades in this step, we found that gender has a small influence in the Application track assessment, but as this variable is far less important than the project assessment in the determination of the Overall Grade, the influence vanishes out.

Internal Step

The internal model is a quite good explanation of the variation in the approval of applications (around 70% of the variability is explained). When looking at each criteria of evaluation, we found that gender seems to have a small influence in the grade given, however when we checked the variable importance it always turned out to be one of the least important variable in the model.

Considering all the analysis that we have done so far, we can say that there is no evidence of gender bias in the funding decision at the Swiss National Science Foundation. Gender is never an important variable in all the regressions we performed and the difference in predicted probability between male and female applicants is not relevant.

Relative Importance of the Different Steps

Our second research question was to assess the relative importance of each step in the process, and the relative importance of each criteria within each step.

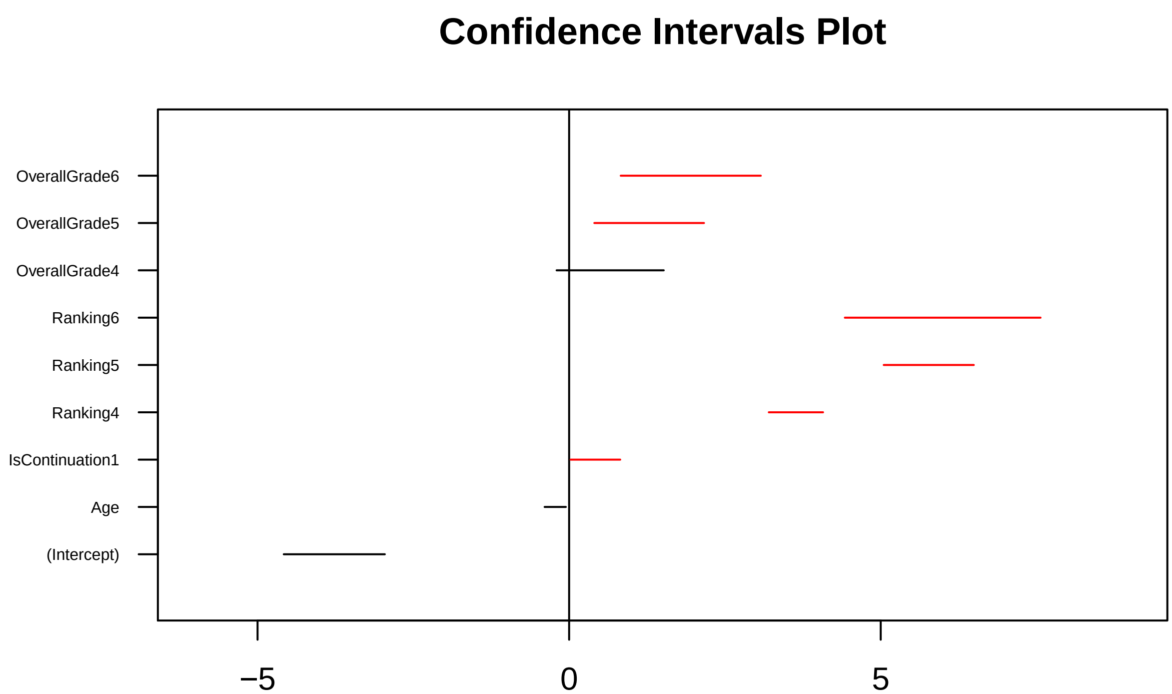
Analysis

Most Important Step

To approach the question of which step in the process is most important, we first fit a logistic regression with IsApproved as our binary response variable using the glm function in R. We fit a full model with all potential demographic predictors and interactions across the different steps of the process, and the summary grade given to an application in the external (OverallGrade) and internal (Ranking) step. To address the first part of the question (relative importance of each step in the process), we used only the summary grade in each step due to the correlation between the individual grades given within each step and the summary grade given. With the full model (predictors: Gender, Division, Age, IsContinuation, InstType, log(AmountRequested), PercentFemale, Ranking, OverallGrade, Gender:Division, PercentFemale:Gender), we achieved a pseudo-R^2 value of 0.7251, indicating that percent of the variation in Y can be explained by the model.

As our goal was to explain the most important factors, we then did backwards variable selection using the AIC. This left us with a model with only 4 predictors: Ranking, OverallGrade, Age, and IsContinuation. The pseudo R^2 measure of this model is 0.7234, which indicates that this simplified model nearly explains exactly as much variance in the data as the full model, and so we can be content to use just the small model.

Now that we've reduced our model to 4 predictors, we wanted to understand exactly how important each of those predictors are to the final funding decision. To do this, we looked looked at the confidence intervals of the coefficients to see which had the largest impact. To do this, we needed to first standardize our continuous variables (Age). When we plot the confidence intervals of our coefficients, we can see that the Ranking has by far the largest coefficinet, and thus the biggest impact on the final funding decision.



Most Important Criteria Within Each Step

The second aspect of this question was to identify what was the most important criteria within each step. To understand this, we again did a permutation test of the different predictors determining the summary grade given in each the external and the internal review step. We used the Ordinal Regressions from earlier: one for the external OverallGrade using the demographic data, Scientific Proposal grade, and Applicant Track grade as predictors, and a second one predicting the Ranking using the demographic data, Scientific Proposal grade, and Applicant Track grade as predictors.

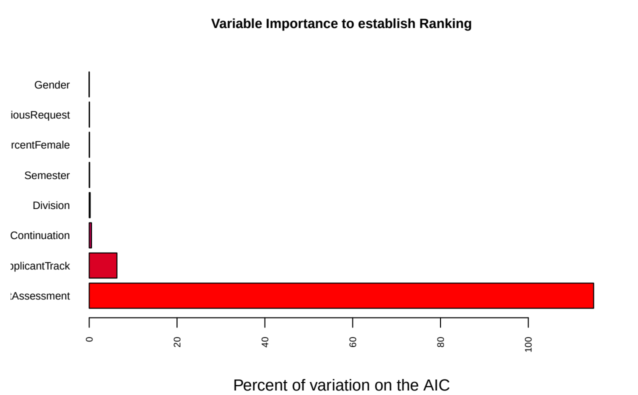
In each of these two regressions, we again computed the variable importance by permuting the values of the predictors one at a time. Since it is an Ordinal Regression, and there is no R^2 equivalent to measure the goodness of fit, we assessed goodness of fit based on the percent of variation in the AIC, another measure of goodness of fit. For both the external and the internal step, we found permuting the grade given to the Scientific Proposal by far had the biggest impact on the quality of the regression. This led us to conclude that the grade given to the Scientific Proposal far outweighs the grade given to the Applicant Track Record, or any of the demographic predictors, in explaining the overall grade given to an application.

Internal step

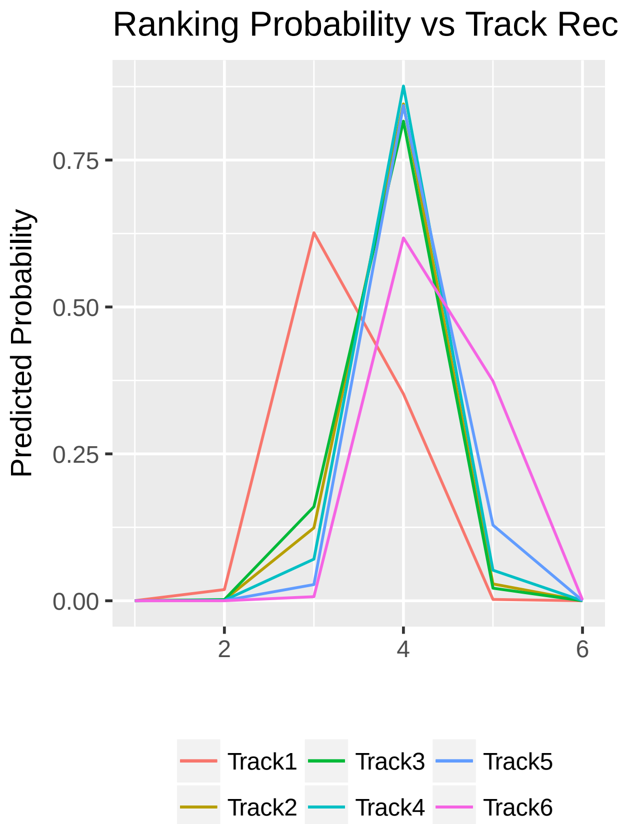
In order to check which variable has the biggest influence on the final grade given by the internal referees, we fitted an ordinal regression using the "Ranking" grade as multinomial response and the Applicant Track grade, the Project Assessment grade and some demographic data as predictors. We started as always fitting the full model and we did variable selection using the AIC. The explanatory variables included in the final model are: ApplicantTrack, ProjectAssessment, Gender, Division, PercentFemale, IsContinuation, PreviousRequest and Semester.

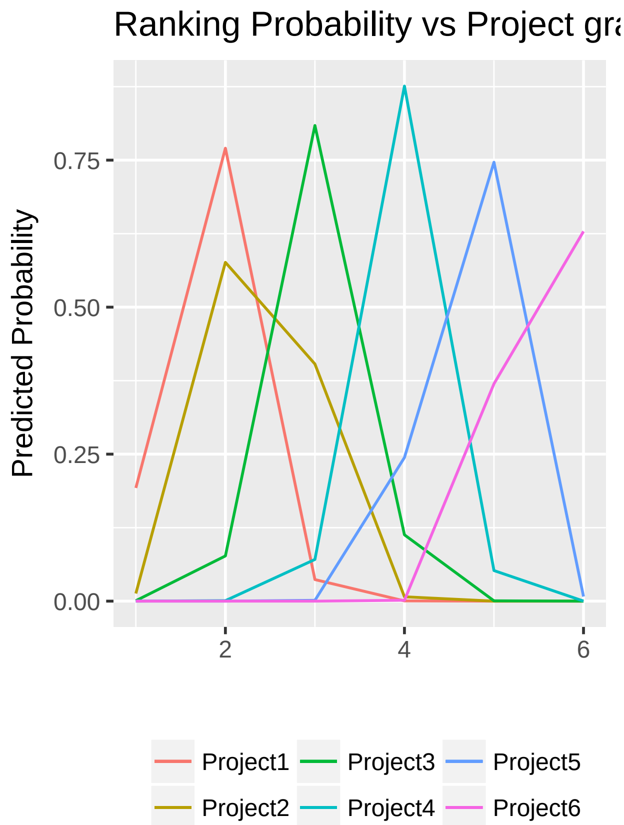
To assess the importance of each variable we performed a permutation test, shuffling the value of one variable at a time and computing the average variation in AIC. As you can see from the plot below, the Project Asessment is by far the most important variable to determine the internal Ranking.

	Feature	Diff.AIC	Percent.AIC
4	ProjectAssessment	2126.3085747	114.8798985
5	ApplicantTrack	116.7845310	6.3096181
6	IsContinuation	9.6902958	0.5235459
2	Division	4.0776721	0.2203079
8	Semester	1.8560758	0.1002798
3	PercentFemale	1.3841074	0.0747804
7	PreviousRequest	1.2664100	0.0684214
1	Gender	-0.4434171	-0.0239569



When we permutRanking Probability vs Track Recorde the Project Assessment grades, the AIC increases by -0.02%. The second most important variable is the Applicant Track, but shuffling its values the AIC increases by "only" 0.03%. We also computed the predicted probabilities for each of the Ranking grade, varying the Applicant Track grade (plot on the left) and the Project Assessment grade (plot on the right). It is clear that the same variation in grade implies different changes in probabilities: when the project grade changes, the probability variation is much bigger. This confirms what we found before: the quality of the project has a greater influence on the final ranking, compared to the track record.





Results

Budget Cuts

Analysis

Results

Conclusion

Appendix

Detailed Data Description

Applications

- **AmountRequested**: Rounded to the next 10k CHF
- **AmountGranted**: Rounded to the next 10k CHF

- **IsApproved**: 1 if the application was approved, 0 if it was rejected
- **GradeFinal**: Comparative ranking of the application as determined by the evaluation body (the division of the National Research Council). A: "belongs to the 10% best percent"; AB: "10% are worse, 75% are better"; B: "50% are worse, 25% are better"; BC: " 25% are worse, 50% are better"; C"10% are worse, 75% are better"; D:"90% of the applications are better"
- **Division**: Evaluation Body in which the application was evaluated. Division 1 evaluates Social Sciences and Humanities; Division 2 Mathematics, Natural Sciences and Engineering; Division 3 Biology and Medicine
- **MainDiscipline**: as chosen by the applicant from the SNSF discipline list
- **MainDisciplineLevel2**: category in the SNF discipline list grouping disciplines into fields of research
- **CallTitle**: Call for proposals under which the application was submitted. Applications from the same Call are evaluated together, i.e. in competition to each other
- **CallEndDate**: Submission deadline of the Call
- **ResponsibleApplicantAcademicAgeAtSubmission**: Years since the applicant's PhD at time of submission; data only available since mid 2016
- **ResponsibleApplicantAgeAtSubmission**: Biological age of the applicant at time of submission; data only available since mid 2016
- **ResponsibleApplicantProfessorshipType**: employment situation of the applicant at time of submission; data only available since mid 2016
- **Gender**: of the main applicant
- **NationalityIsoCode**: Nationality of the main applicant
- **IsHasPreviousProjectRequested**: 0 if it is the applicant's first application at the SNSF, 1 if not
- **InstType**: Type of institution where the applicant is employed
- **IsContinuation**: 1 if the project is a thematic continuation of a previously approved project, 0 if not
- **ProjectID**: Anonymized identifier of the application

Referee Grades

- **Question**: Evaluation criterion
- **QuestionRating**: The (co-)referee's assessment of the evaluation criterion
- **OverallRanking**: The (co-)referee's overall comparative ranking of the application. A: "belongs to the 10% best percent"; same scale as the GradeFinal
- **RefereeRole**: Some applications have one referee evaluation, some have two. The role indicates who was the primary and who was the secondary referee (also called co-referee)
- **RefereeGender**
- **IDs**: Anonymized identifiers of the application, the referee and the evaluation by the referee

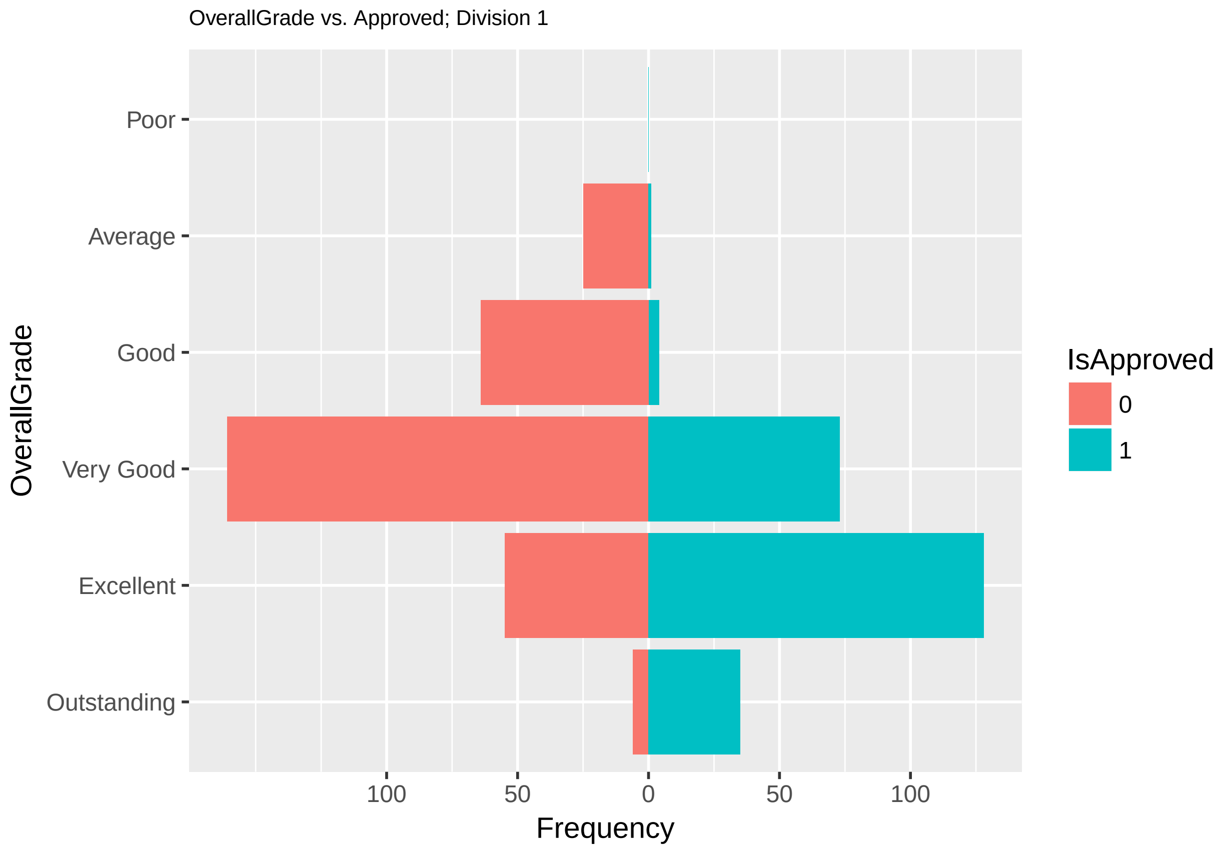
Reviews

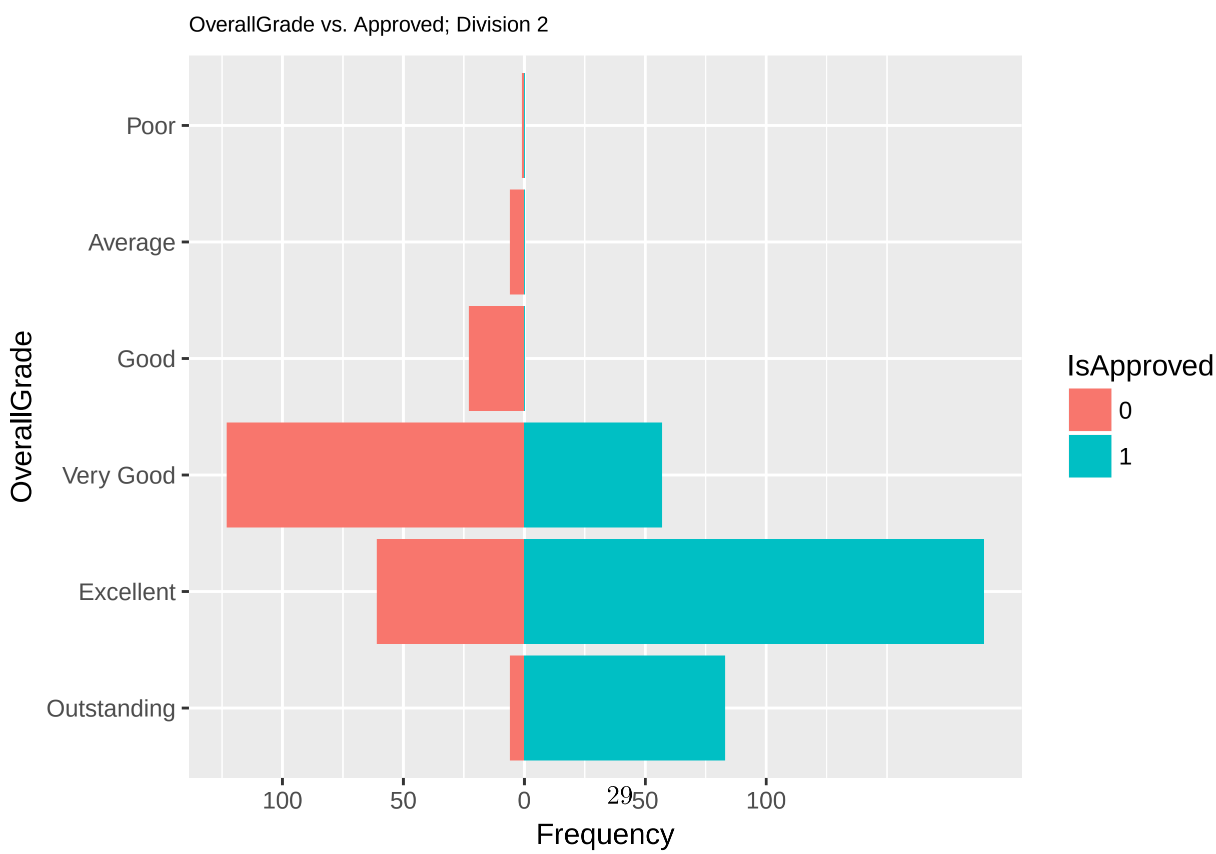
- **Question**: Evaluation criterion
- **QuestionRating**: The external reviewer's assessment of the evaluation criterion
- **OverallGrade**: The external reviewer's overall assessment of the application
- **SourcePerson**: Who suggested the reviewer?
- **Gender**
- **Country**: where the reviewer is located. Not always known

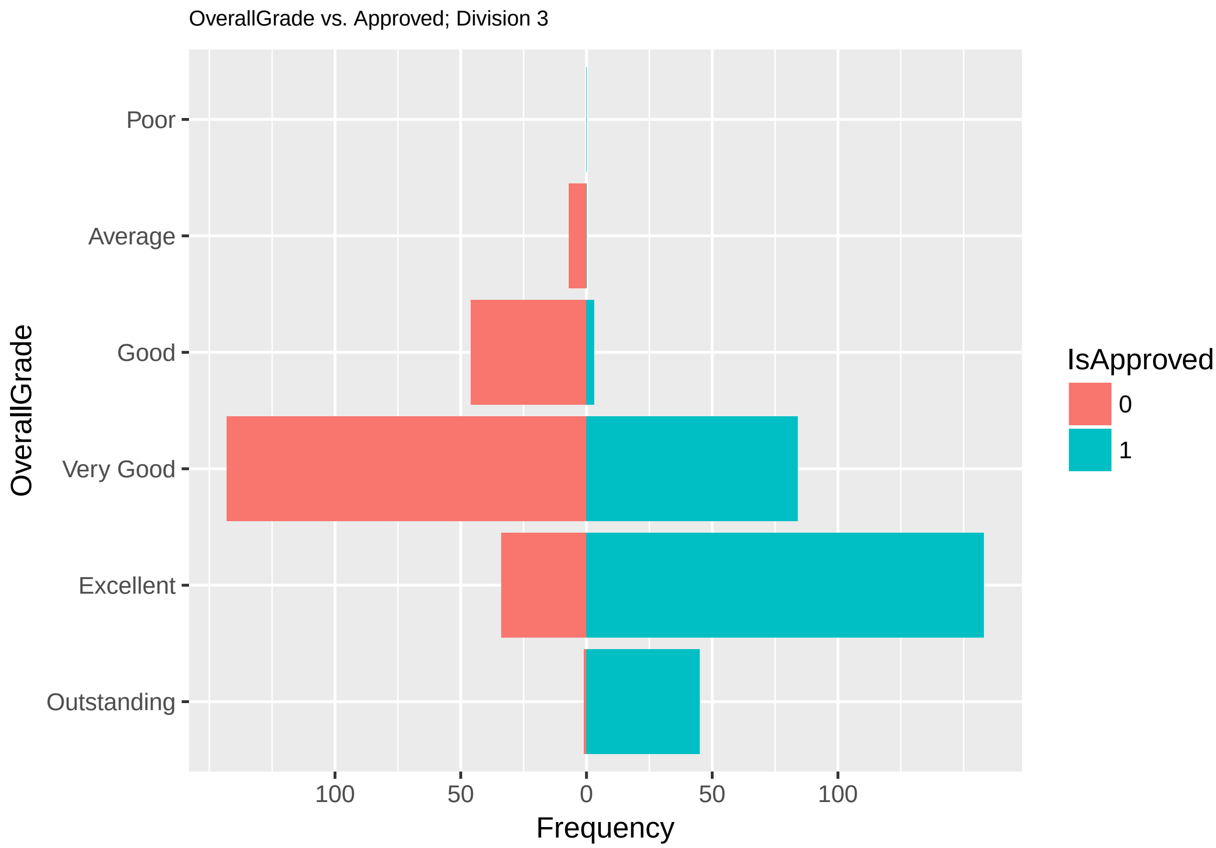
- **EmailEnding**: ending of the reviewer's email address. Might be used as an approximation of the country where the reviewer is located in cases where this data is missing
- **IDs**: Anonymized identifiers of the application, the reviewer and the review

Exploratory Analysis

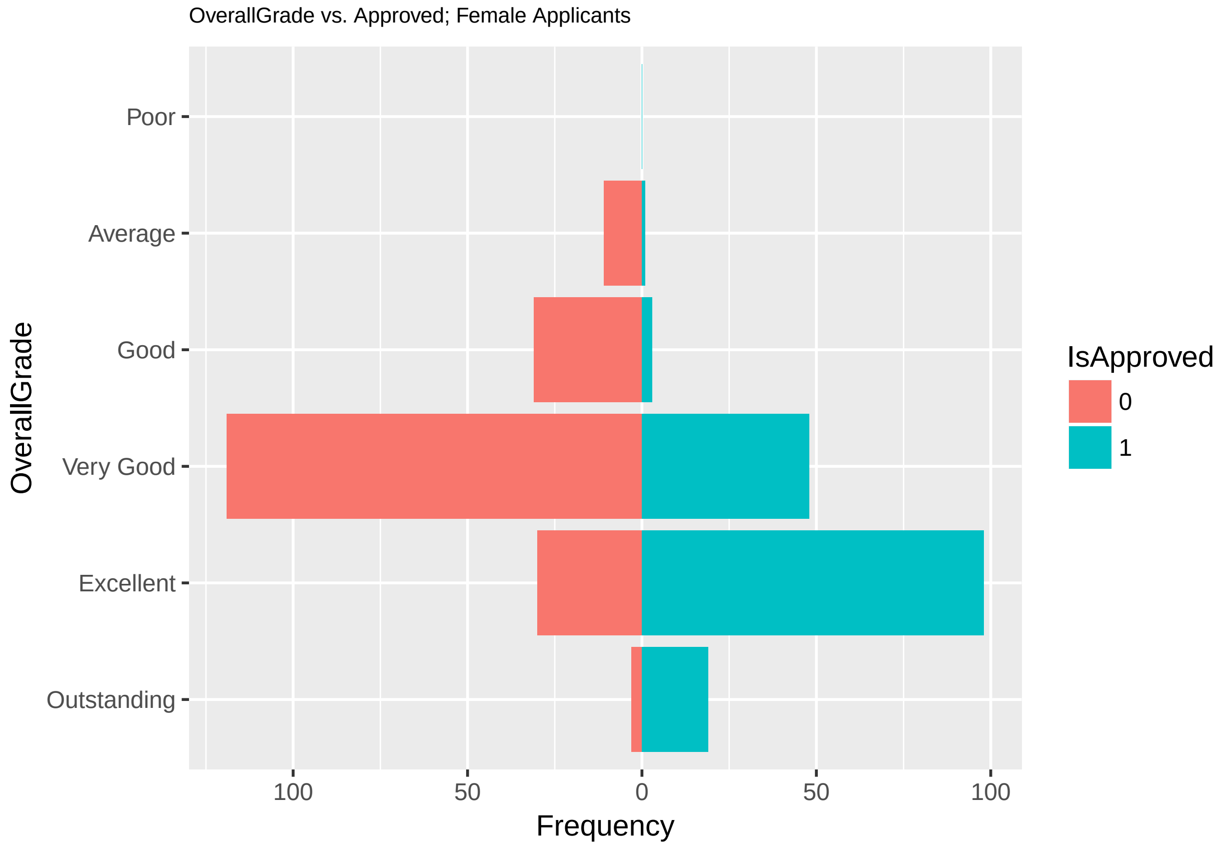
OverallGrade vs. IsApproved, by Division

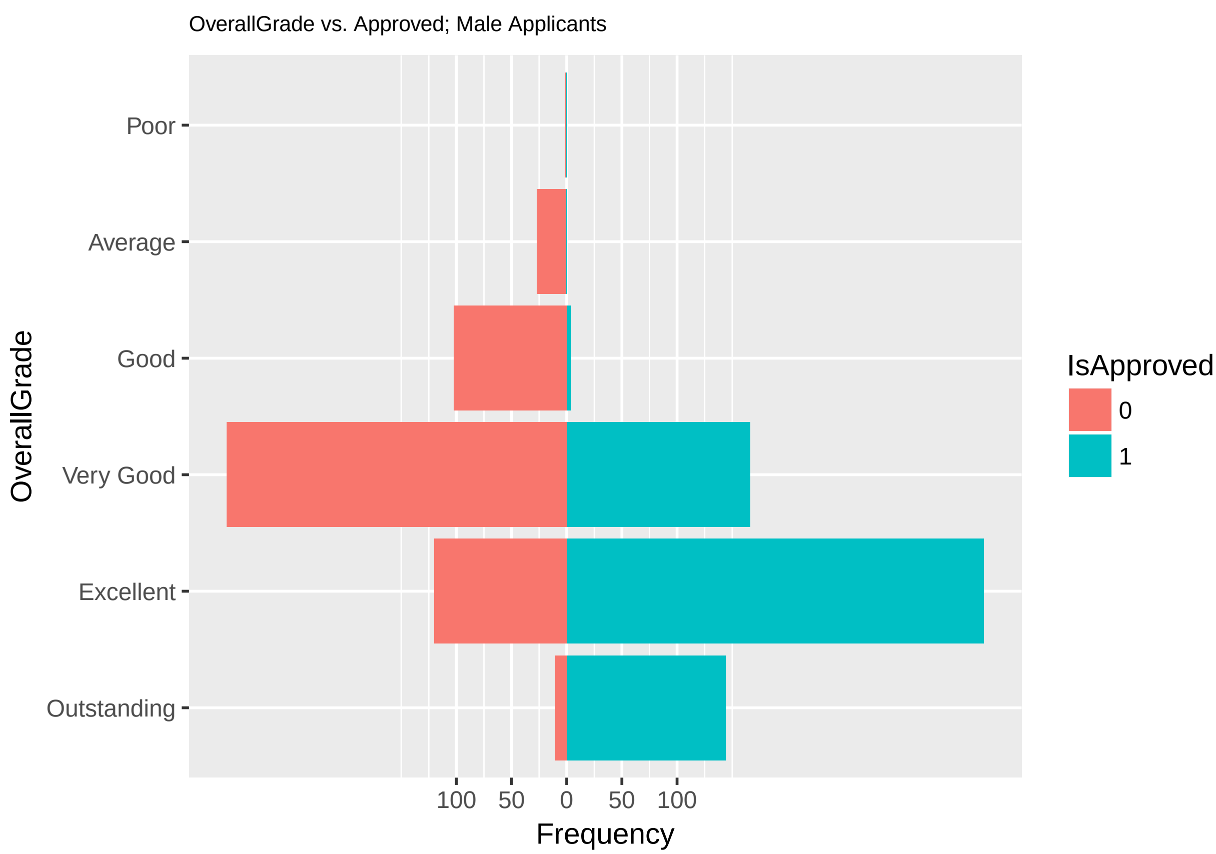





OverallGrade vs. Approved; Division 3
OverallGrade
Poor
Average
Good
Very Good
Excellent
Outstanding
100
50
0
50
100
Frequency
IsApproved
0
1

OverallGrade vs. IsApproved, by Gender





External Logistic Regression

- Summary of the final model

Ordinal external Regression

Project grades (ProposalCombined)

Summary of the final model:

```
formula:
ProposalCombined ~ Gender + Division + PercentFemale + IsContinuation + InstType + logAmount
data:    external_regression_data

 link  threshold nobs logLik   AIC     niter max.grad cond.H
 logit flexible  1623 -1949.71 3927.42 7(0)  8.06e-10 4.9e+05

Coefficients:
                Estimate Std. Error z value Pr(>|z|)
Genderf         -0.18423    0.11413  -1.614  0.10649
DivisionDiv 2    0.36126    0.14190   2.546  0.01090 *
DivisionDiv 3   -0.28708    0.13153  -2.183  0.02907 *
PercentFemale   -0.43375    0.20023  -2.166  0.03029 *
IsContinuation1  0.57973    0.11661   4.971 6.65e-07 ***
InstTypeOther   -0.63461    0.23714  -2.676  0.00745 **
InstTypeUAS/UTE -0.83721    0.20535  -4.077 4.56e-05 ***
InstTypeUni     -0.27226    0.13099  -2.078  0.03767 *
logAmount        0.44651    0.08627   5.176 2.27e-07 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Threshold coefficients:
    Estimate Std. Error z value
1|2   -1.280      1.310  -0.977
2|3    1.798      1.114   1.613
3|4    3.416      1.108   3.084
4|5    5.709      1.112   5.132
5|6    8.174      1.123   7.276
```

Odd Ratios and Confidence intervals:

	OR	2.5 %	97.5 %
Genderf	0.83	0.66	1.04
DivisionDiv 2	1.44	1.09	1.90
DivisionDiv 3	0.75	0.58	0.97
PercentFemale	0.65	0.44	0.96
IsContinuation1	1.79	1.42	2.25
InstTypeOther	0.53	0.33	0.84
InstTypeUAS/UTE	0.43	0.29	0.65
InstTypeUni	0.76	0.59	0.98
logAmount	1.56	1.32	1.85

Applicant Track

Summary of the final model:

```
formula:
ApplicantTrack ~ Gender + Division + PercentFemale + IsContinuation + InstType + logAmount + Gender
data:    external_regression_data

 link  threshold nobs logLik   AIC     niter max.grad cond.H
 logit flexible  1623 -1895.84 3821.69 7(0)  7.99e-12 5.1e+05

Coefficients:
                      Estimate Std. Error z value Pr(>|z|)
Genderf               -0.51334    0.16102  -3.188 0.001433 **
DivisionDiv 2          0.51317    0.14066   3.648 0.000264 ***
DivisionDiv 3         -0.19141    0.12973  -1.475 0.140095
PercentFemale         -0.79751    0.23535  -3.389 0.000703 ***
IsContinuation1        0.47308    0.11730   4.033 5.51e-05 ***
InstTypeOther         -0.55931    0.23642  -2.366 0.017996 *
InstTypeUAS/UTE       -1.26531    0.20288  -6.237 4.46e-10 ***
InstTypeUni           -0.23820    0.13069  -1.823 0.068343 .
logAmount              0.55195    0.08643   6.386 1.70e-10 ***
Genderf:PercentFemale  1.04403    0.41418   2.521 0.011711 *
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Threshold coefficients:
    Estimate Std. Error z value
1|2  -0.7523     1.4897  -0.505
2|3   2.0350     1.1325   1.797
3|4   3.8149     1.1106   3.435
4|5   6.1707     1.1115   5.551
5|6   8.2987     1.1208   7.404
```

Odd Ratios and Confidence intervals:

	OR	2.5 %	97.5 %
Genderf	0.60	0.44	0.82
DivisionDiv 2	1.67	1.27	2.20
DivisionDiv 3	0.83	0.64	1.06
PercentFemale	0.45	0.28	0.71
IsContinuation1	1.60	1.28	2.02
InstTypeOther	0.57	0.36	0.91
InstTypeUAS/UTE	0.28	0.19	0.42
InstTypeUni	0.79	0.61	1.02
logAmount	1.74	1.47	2.06
Genderf:PercentFemale	2.84	1.26	6.40

Overall Grade

Summary of the final model:

```
formula:
```

```
OverallGrade ~ ApplicantTrack + ProposalCombined + PercentFemale + Gender + IsContinuation + Percent
data:    external_regression_data

 link  threshold nobs logLik  AIC     niter max.grad cond.H
 logit flexible  1623 -833.31 1704.61 8(0)  6.33e-08 2.9e+03

Coefficients:
                      Estimate Std. Error z value Pr(>|z|)
ApplicantTrack.L        6.2284     2.2048   2.825  0.00473 **
ApplicantTrack.Q        1.0113     1.9494   0.519  0.60392
ApplicantTrack.C       -0.8432     1.3748  -0.613  0.53969
ApplicantTrack^4        0.6535     0.8014   0.816  0.41477
ApplicantTrack^5       -0.2462     0.3864  -0.637  0.52407
ProposalCombined.L     13.1111     1.6336   8.026 1.01e-15 ***
ProposalCombined.Q      2.2458     1.4162   1.586  0.11278
ProposalCombined.C     -1.7896     0.9897  -1.808  0.07057 .
ProposalCombined^4      0.6906     0.5696   1.212  0.22534
ProposalCombined^5     -0.4098     0.2677  -1.531  0.12588
PercentFemale          -0.7295     0.3218  -2.267  0.02337 *
Genderf                -0.4804     0.2296  -2.092  0.03641 *
IsContinuation1         0.3165     0.1554   2.036  0.04171 *
PercentFemale:Genderf   1.6716     0.5947   2.811  0.00494 **
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Threshold coefficients:
    Estimate Std. Error z value
1|2 -12.2445     1.3270  -9.228
2|3  -6.0098     0.8169  -7.357
3|4  -1.0010     0.7898  -1.267
4|5   4.9117     0.8014   6.129
5|6  10.3953     0.8246  12.607
```

Odd Ratios and Confidence intervals:

	OR	2.5 %	97.5 %
ApplicantTrack.L	506.94	21.01	19755.74
ApplicantTrack.Q	2.75	0.12	44.50
ApplicantTrack.C	0.43	0.05	3.76
ApplicantTrack^4	1.92	0.50	7.92
ApplicantTrack^5	0.78	0.37	1.59
ProposalCombined.L	494405.74	34781.51	10341935.85
ProposalCombined.Q	9.45	0.72	90.71
ProposalCombined.C	0.17	0.03	0.99
ProposalCombined^4	1.99	0.72	5.71
ProposalCombined^5	0.66	0.39	1.09
PercentFemale	0.48	0.26	0.90
Genderf	0.62	0.39	0.97
IsContinuation1	1.37	1.01	1.86
PercentFemale:Genderf	5.32	1.66	17.11

Internal Logistic Regression

- Summary of the final model

```
Call:
glm(formula = IsApproved ~ Gender + Age + Semester + IsContinuation +
    PercentFemale + ApplicantTrack + ProjectAssessment + logAmount,
    family = "binomial", data = data)

Deviance Residuals:
    Min       1Q   Median       3Q      Max
-2.6763  -0.4069   0.2040   0.5752   2.6648

Coefficients:
                    Estimate Std. Error z value Pr(>|z|)
(Intercept)          3.84166    2.08032   1.847 0.064796 .
Genderf              0.14332    0.19051   0.752 0.451869
Age                 -0.01377    0.01017  -1.354 0.175648
SemesterOktober      0.19430    0.16517   1.176 0.239434
IsContinuation1      0.58642    0.19803   2.961 0.003063 **
PercentFemale       -0.38897    0.20067  -1.938 0.052584 .
ApplicantTrack.L     0.88660    0.28885   3.069 0.002145 **
ApplicantTrack.Q    -0.44577    0.22437  -1.987 0.046952 *
ApplicantTrack.C    -0.04565    0.14807  -0.308 0.757844
ProjectAssessment.L  4.29435    0.52753   8.140 3.94e-16 ***
ProjectAssessment.Q -1.52049    0.40461  -3.758 0.000171 ***
ProjectAssessment.C -0.22306    0.25653  -0.870 0.384557
logAmount           -0.17391    0.15548  -1.118 0.263354
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

(Dispersion parameter for binomial family taken to be 1)

    Null deviance: 2243.9  on 1622  degrees of freedom
Residual deviance: 1073.2  on 1610  degrees of freedom
AIC: 1099.2

Number of Fisher Scoring iterations: 6
```

- Effect plots

Ordinal internal Regressions

Project Assessment

Summary of the final model:

```

    formula:
    ProjectAssessment ~ Gender + Division + PercentFemale + Age + IsContinuation + InstType + logAmount
    data:    internal_regression_data

     link  threshold nobs logLik   AIC     niter max.grad cond.H
```

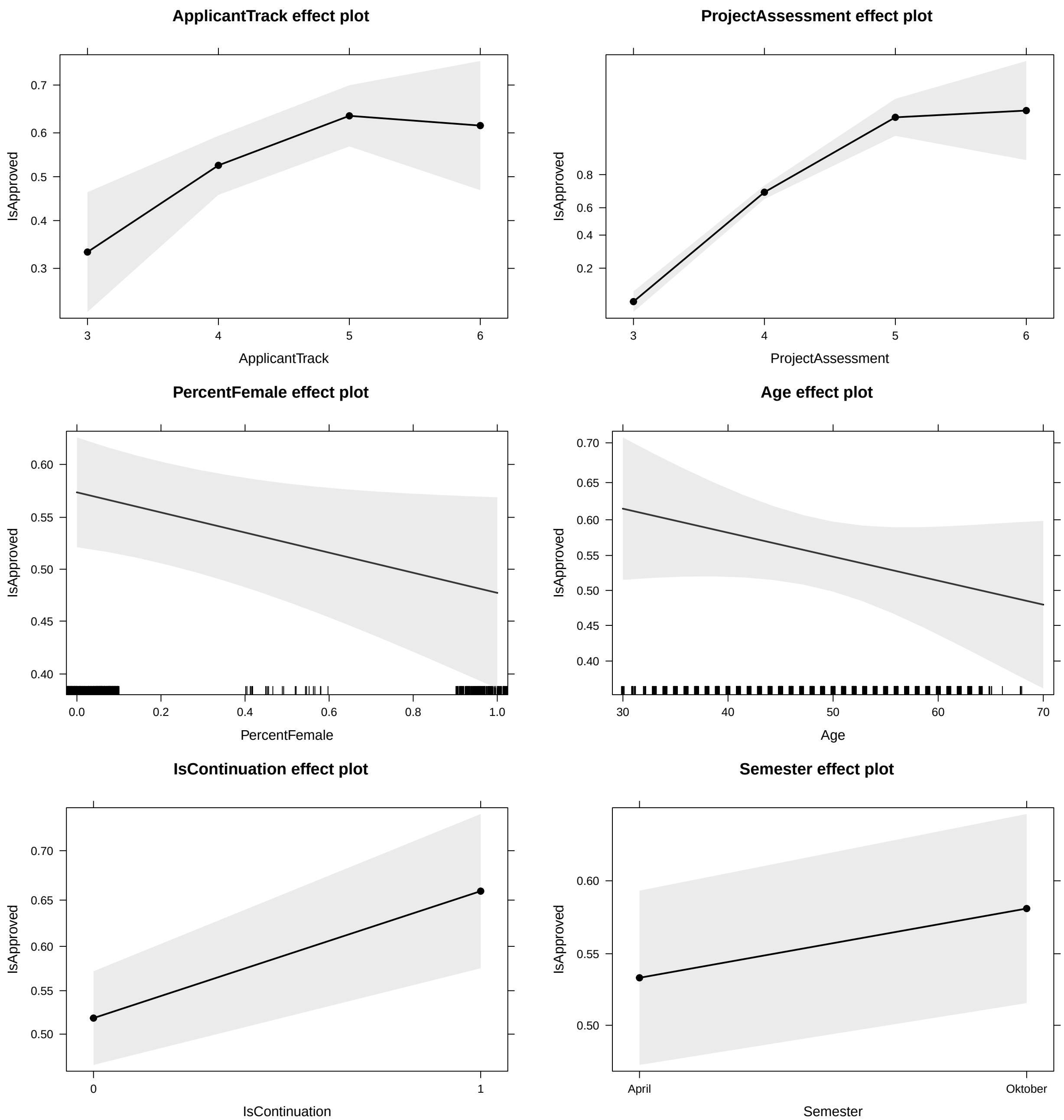


Figure 1: Effects of the external logistic regression

```
 logit flexible  1623 -2352.59 4735.17 6(0)  3.00e-09 7.9e+06

Coefficients:
                 Estimate Std. Error z value Pr(>|z|)
Genderf         -0.247031   0.110412  -2.237 0.025263 *
DivisionDiv 2    0.473118   0.133485   3.544 0.000394 ***
DivisionDiv 3    0.007914   0.125690   0.063 0.949793
PercentFemale    0.328773   0.118157   2.783 0.005394 **
Age              0.008426   0.005777   1.459 0.144662
IsContinuation1  0.737721   0.115973   6.361 2.00e-10 ***
InstTypeOther   -0.625207   0.227598  -2.747 0.006015 **
InstTypeUAS/UTE -0.813258   0.200263  -4.061 4.89e-05 ***
InstTypeUni     -0.434411   0.126555  -3.433 0.000598 ***
logAmount        0.537325   0.084040   6.394 1.62e-10 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Threshold coefficients:
    Estimate Std. Error z value
1|2    3.113      1.124   2.770
2|3    5.358      1.111   4.823
3|4    6.916      1.114   6.211
4|5    8.394      1.120   7.496
5|6   10.708      1.134   9.441
```

Odd Ratios and Confidence intervals:

	OR	2.5 %	97.5 %
Genderf	0.78	0.63	0.97
DivisionDiv 2	1.60	1.24	2.09
DivisionDiv 3	1.01	0.79	1.29
PercentFemale	1.39	1.10	1.75
Age	1.01	1.00	1.02
IsContinuation1	2.09	1.67	2.63
InstTypeOther	0.54	0.34	0.84
InstTypeUAS/UTE	0.44	0.30	0.66
InstTypeUni	0.65	0.51	0.83
logAmount	1.71	1.45	2.02

Applicant Track

Summary of the final model:

```
formula:
ApplicantTrack ~ Gender + Division + PercentFemale + IsContinuation + InstType + Semester + logAmour
data:    internal_regression_data

 link  threshold nobs logLik   AIC     niter max.grad cond.H
 logit flexible  1623 -2017.29 4072.59 6(0)  9.12e-07 5.3e+05

Coefficients:
                            Estimate Std. Error z value Pr(>|z|)
Genderf                     -0.11356    0.17347  -0.655 0.512717
DivisionDiv 2                0.32544    0.15799   2.060 0.039406 *
DivisionDiv 3               -0.41435    0.15915  -2.604 0.009226 **
PercentFemale                0.72319    0.17787   4.066 4.78e-05 ***
IsContinuation1              0.45958    0.11657   3.942 8.06e-05 ***
InstTypeOther               -0.83234    0.23740  -3.506 0.000455 ***
InstTypeUAS/UTE             -1.34969    0.20382  -6.622 3.55e-11 ***
InstTypeUni                 -0.38629    0.12918  -2.990 0.002788 **
SemesterOct                 -0.15398    0.09791  -1.573 0.115780
logAmount                    0.77252    0.08994   8.589  < 2e-16 ***
Genderf:DivisionDiv 2       -0.60542    0.29360  -2.062 0.039204 *
Genderf:DivisionDiv 3       -0.48321    0.26212  -1.844 0.065256 .
DivisionDiv 2:PercentFemale -0.73185    0.29184  -2.508 0.012153 *
DivisionDiv 3:PercentFemale -0.28183    0.29618  -0.952 0.341324
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Threshold coefficients:
    Estimate Std. Error z value
1|2    3.159      1.273   2.483
2|3    5.786      1.146   5.051
3|4    7.550      1.138   6.632
4|5    9.706      1.145   8.475
5|6   11.906      1.160  10.265
```

Odd Ratios and Confidence intervals:

	OR	2.5 %	97.5 %
Genderf	0.89	0.64	1.25
DivisionDiv 2	1.38	1.02	1.89
DivisionDiv 3	0.66	0.48	0.90
PercentFemale	2.06	1.45	2.92
IsContinuation1	1.58	1.26	1.99
InstTypeOther	0.44	0.27	0.69
InstTypeUAS/UTE	0.26	0.17	0.39
InstTypeUni	0.68	0.53	0.88
SemesterOct	0.86	0.71	1.04
logAmount	2.17	1.82	2.58
Genderf:DivisionDiv 2	0.55	0.31	0.97
Genderf:DivisionDiv 3	0.62	0.37	1.03
DivisionDiv 2:PercentFemale	0.48	0.27	0.85
DivisionDiv 3:PercentFemale	0.75	0.42	1.35

Ranking

Summary of the final model:

```
formula:
Ranking ~ Gender + PercentFemale + Division + IsContinuation + PreviousRequest + InstType + logAmou
data:    internal_regression_data

 link  threshold nobs logLik   AIC     niter max.grad cond.H
 logit flexible  1623 -2306.66 4643.33 7(0)  4.93e-12 5.1e+05

Coefficients:
                 Estimate Std. Error z value Pr(>|z|)
Genderf           -0.3100     0.1101  -2.816 0.004870 **
PercentFemale      0.2667     0.1187   2.247 0.024668 *
DivisionDiv 2      0.2367     0.1326   1.784 0.074392 .
DivisionDiv 3     -0.2166     0.1258  -1.722 0.085054 .
IsContinuation1    0.8434     0.1148   7.349 2.00e-13 ***
PreviousRequest1   0.2065     0.1321   1.563 0.118156
InstTypeOther     -0.7457     0.2270  -3.284 0.001022 **
InstTypeUAS/UTE   -0.9866     0.2004  -4.922 8.55e-07 ***
InstTypeUni       -0.4358     0.1262  -3.455 0.000551 ***
logAmount          0.5707     0.0842   6.778 1.22e-11 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Threshold coefficients:
    Estimate Std. Error z value
1|2    2.533      1.111   2.279
2|3    5.287      1.085   4.872
3|4    6.903      1.088   6.345
4|5    8.472      1.095   7.738
5|6   10.696      1.109   9.644
```

Odd Ratios and Confidence intervals:

	OR	2.5 %	97.5 %
Genderf	0.73	0.59	0.91
PercentFemale	1.31	1.03	1.65
DivisionDiv 2	1.27	0.98	1.64
DivisionDiv 3	0.81	0.63	1.03
IsContinuation1	2.32	1.86	2.91
PreviousRequest1	1.23	0.95	1.59
InstTypeOther	0.47	0.30	0.74
InstTypeUAS/UTE	0.37	0.25	0.55
InstTypeUni	0.65	0.50	0.83
logAmount	1.77	1.50	2.09

Relative importance within Internal step

Summary of the final model:

```
formula:
Ranking ~ Gender + Division + PercentFemale + ProjectAssessment + ApplicantTrack + IsContinuation +
data:    data

 link  threshold nobs logLik  AIC     niter max.grad cond.H
 logit flexible  1623 -903.45 1850.90 8(0)  9.60e-10 1.2e+03

Coefficients:
                    Estimate Std. Error z value Pr(>|z|)
Genderf             -0.13394    0.15488  -0.865 0.387149
DivisionDiv 2       -0.38782    0.16595  -2.337 0.019443 *
DivisionDiv 3       -0.32140    0.17128  -1.876 0.060602 .
PercentFemale       -0.25690    0.16889  -1.521 0.128234
ProjectAssessment.L 16.50814    0.59118  27.924  < 2e-16 ***
ProjectAssessment.Q  1.85162    0.43349   4.271 1.94e-05 ***
ProjectAssessment.C -0.07799    0.30991  -0.252 0.801317
ProjectAssessment^4 -0.05701    0.19537  -0.292 0.770435
ProjectAssessment^5  0.49586    0.13401   3.700 0.000216 ***
ApplicantTrack.L     3.98863    1.01703   3.922 8.79e-05 ***
ApplicantTrack.Q    -0.11820    0.90649  -0.130 0.896252
ApplicantTrack.C     0.99379    0.64982   1.529 0.126183
ApplicantTrack^4    -0.65197    0.39483  -1.651 0.098685 .
ApplicantTrack^5     0.33621    0.21250   1.582 0.113608
IsContinuation1      0.55149    0.17047   3.235 0.001216 **
PreviousRequest1     0.28230    0.17894   1.578 0.114663
SemesterOct         -0.22606    0.13088  -1.727 0.084126 .
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Threshold coefficients:
    Estimate Std. Error z value
1|2  -9.8221     0.5077 -19.348
2|3  -5.1330     0.4055 -12.659
3|4  -0.6016     0.3926  -1.532
4|5   4.8578     0.4100  11.850
5|6  10.8194     0.4996  21.655
```

Odd Ratios and Confidence intervals:

	OR	2.5 %	97.5 %
Genderf	0.87	0.65	1.19
DivisionDiv 2	0.68	0.49	0.94
DivisionDiv 3	0.73	0.52	1.01
PercentFemale	0.77	0.56	1.08
ProjectAssessment.L	14770410.53	4762009.57	48526691.84
ProjectAssessment.Q	6.37	2.70	14.93
ProjectAssessment.C	0.92	0.51	1.73
ProjectAssessment^4	0.94	0.64	1.39
ProjectAssessment^5	1.64	1.26	2.14
ApplicantTrack.L	53.98	8.36	483.71
ApplicantTrack.Q	0.89	0.12	4.66
ApplicantTrack.C	2.70	0.82	10.85
ApplicantTrack^4	0.52	0.23	1.10
ApplicantTrack^5	1.40	0.93	2.13
IsContinuation1	1.74	1.24	2.43
PreviousRequest1	1.33	0.93	1.88
SemesterOct	0.80	0.62	1.03