05/10/2020

**OVERVIEW REPORT**

AN ANALYSIS ON LSE SURVEY SCORES

AUTHORS:

Hasan Younis

Yunxiao Chen

Ellen Austin

**TABLE OF CONTENTS**

|  |  |
| --- | --- |
| Part | Page Number |
| **Introduction** | 3 |
| **Exploratory Analysis and Methodology** | 3 |
| **Results and Interpretations** | 6 |
| *Regression Analysis* | 6 |
| *Dependence Analysis* | 20 |
| **Conclusion** | 26 |
| **Appendix A** | 27 |
| **Appendix B** | 28 |
| **Appendix C** | 28 |

Introduction

Each year the Office for Students sponsors the National Student Survey, gathering final year undergraduate students’ opinions on their courses for all universities in the UK. This survey consists of 27 core questions across eight categories and results are used to assess institutional performance and inform future student choice. Institutions do not have access to student-level data from the NSS but can see overall results broken down according to programme, department and a range of demographic characteristics.

Like many institutions, LSE runs a ‘mirror’ NSS survey for first- and second-year undergraduates so that students can reflect and share their views on their programme as it progresses. As well as allowing LSE to improve and develop programmes ‘in real time’, these surveys provide student-level data about satisfaction on a range of metrics, using the same questions and groupings as the NSS.

This report aims to explore results from LSE’s internal year 1 and year 2 programme surveys for 2018-9 to establish whether student characteristics affect their satisfaction levels, and to identify whether responses to one question are strongly correlated with responses to another.

Exploratory Analysis and Methodology

The original dataset is made up of 3421 LSE students, with 51 given variables (27 of which represent the students answer to the LSE’s adaptation of the NSS questionnaire, and the remaining 24 being covariates describing different aspects about the student). Three students are removed at this stage, with the commonality being that they have the Student sex “Other” as the lack of sample size does not allow us to make any reasonable conclusions. This leaves us with 3418 students. It is decided that besides the 27 questions which will serve as the main focus behind our regressions, only 10 of the 24 covariates listed in the original document are useful to us. These 10 covariates are displayed in the table below:

|  |  |  |  |
| --- | --- | --- | --- |
| **Covariates** | **Code** | **Options** | **Baseline** |
| Student sex | Studentsex | “M” for Male, “F” for Female | “M” |
| Maturity | Maturity | “Young” for Students who were younger than 21 at the commencement of studies, “Mature” for Students who were older than 21 at commencement of studies. | “Young” |
| OECD | OECD | “1” for if a student’s country of domicile is an OECD country, “0” for if a student’s country of domicile is not an OECD country. | “1” |
| Bursary | Bursary | “No Bursary” if a student did not hold a bursary at any point during their undergraduate study, “Bursary” if a student did hold a bursary at some point | “No Bursary” |
| English Status | englishstatus | “Native” if student was recorded as a native English speaker on application, “Non-Native” if student is recorded as a non-native English speaker on application. | “Native” |
| Fee Status | fee | “HUK” if student is a UK student, “HEU” if student is an EU Student, “OS” if student is an Overseas student. | “HUK” |
| Average mark in year | avgpassinyr | Average of recorded marks for modules taken in academic year | Considered as “0” |
| Ethnicity HESA values | ethnicityHESA | “White” if student is from White ethnic background, “Asian” if student is from Asian ethnic background, “Otherincmixed” if student from a Mixed or Other ethnic background, “Missingpnts” if a student declined information, “Black” if student is from a Black ethnic background. | “White” |
| Disability Status | disability | “No Disability” if the student has not declared a disability, “Disability” if the student declared a disability. | “No Disability” |
| Student Year | `Student Year` | “Year 1” if student is currently in their first academic year, “Year 2” if a student is current in their second academic year. | “Year 1” |

The following regressions use all of the above covariates as factors to consider. The categorical variates in particular are calculated with respect to a baseline, which is displayed in the third column of the table. These baselines are generally speaking the most common option for each of the covariates and therefore dictate the most common student by which the regressions are calculated. For the numerical covariate “Average mark in year”, the regression assumes the student has “0”, and then for every average mark above 0 the student gets, the coefficient is then multiplied by that number and then added to the intercept value.

The unused 14 covariates are then removed. Leaving us with 3418 students and these 37 key variates to work with. Due to the missingness of values within these covariates, a complete case dataset is constructed; selecting only students who have values for each of the above 37 key variates. The complete case data set has 1801 students and forms the basis for which all the regressions and indeed dependence analyses are run. The following tables represent the distribution of students for each of the categorical factors.

|  |  |  |
| --- | --- | --- |
| Student Sex Options | Male | Female |
| Number of Students | 920 | 881 |

|  |  |  |
| --- | --- | --- |
| Maturity Options | Young | Mature |
| Number of Students | 1730 | 71 |
| OECD Options | “1” | “0” |
| Number of Students | 1133 | 668 |

|  |  |  |
| --- | --- | --- |
| Bursary Options | No Bursary | Bursary |
| Number of Students | 1481 | 320 |

|  |  |  |
| --- | --- | --- |
| English Status Options | Native | Non-Native |
| Number of Students | 1205 | 596 |

|  |  |  |  |
| --- | --- | --- | --- |
| Fee Status Options | HUK | HEU | OS |
| Number of Students | 801 | 282 | 718 |

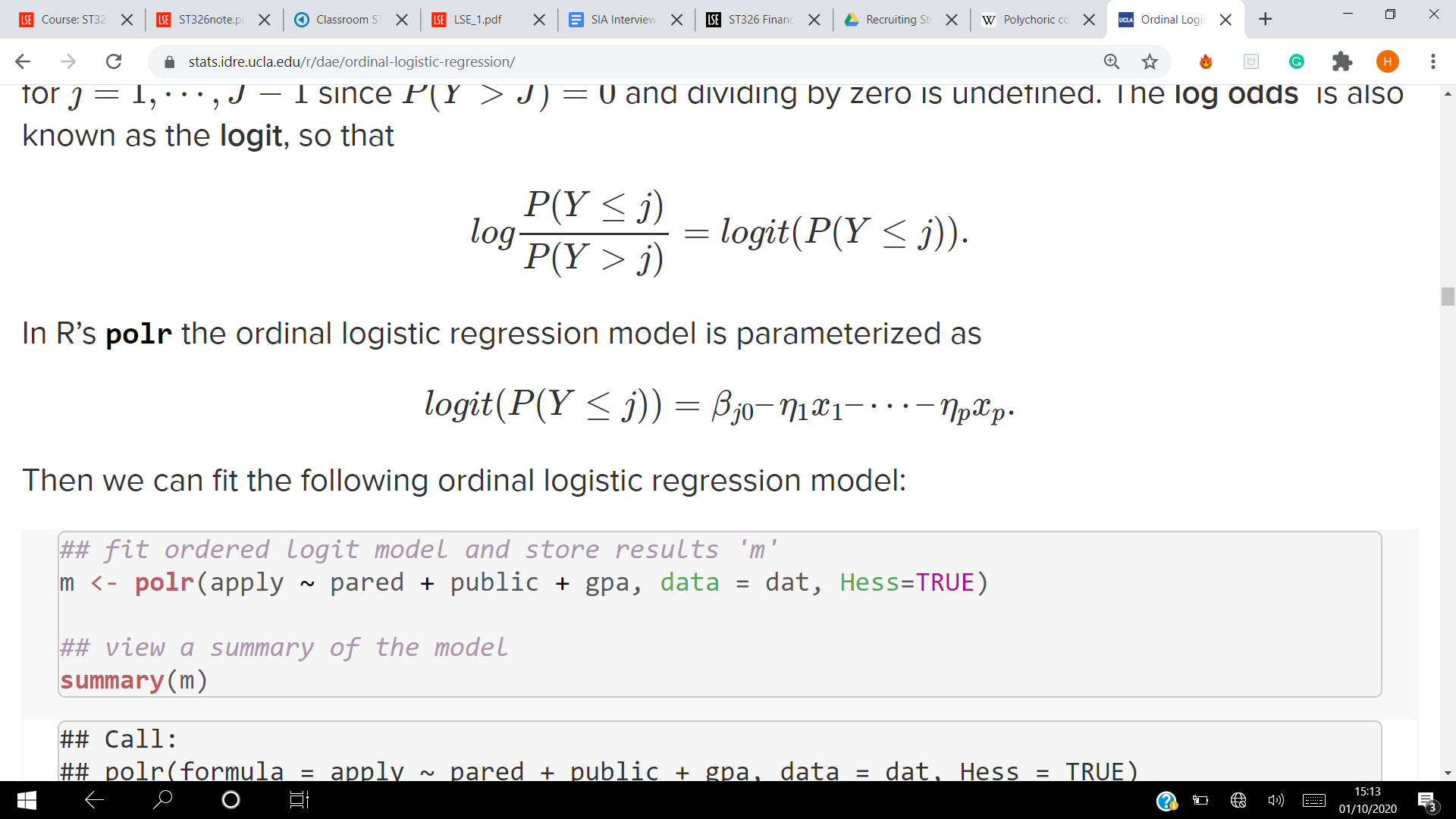
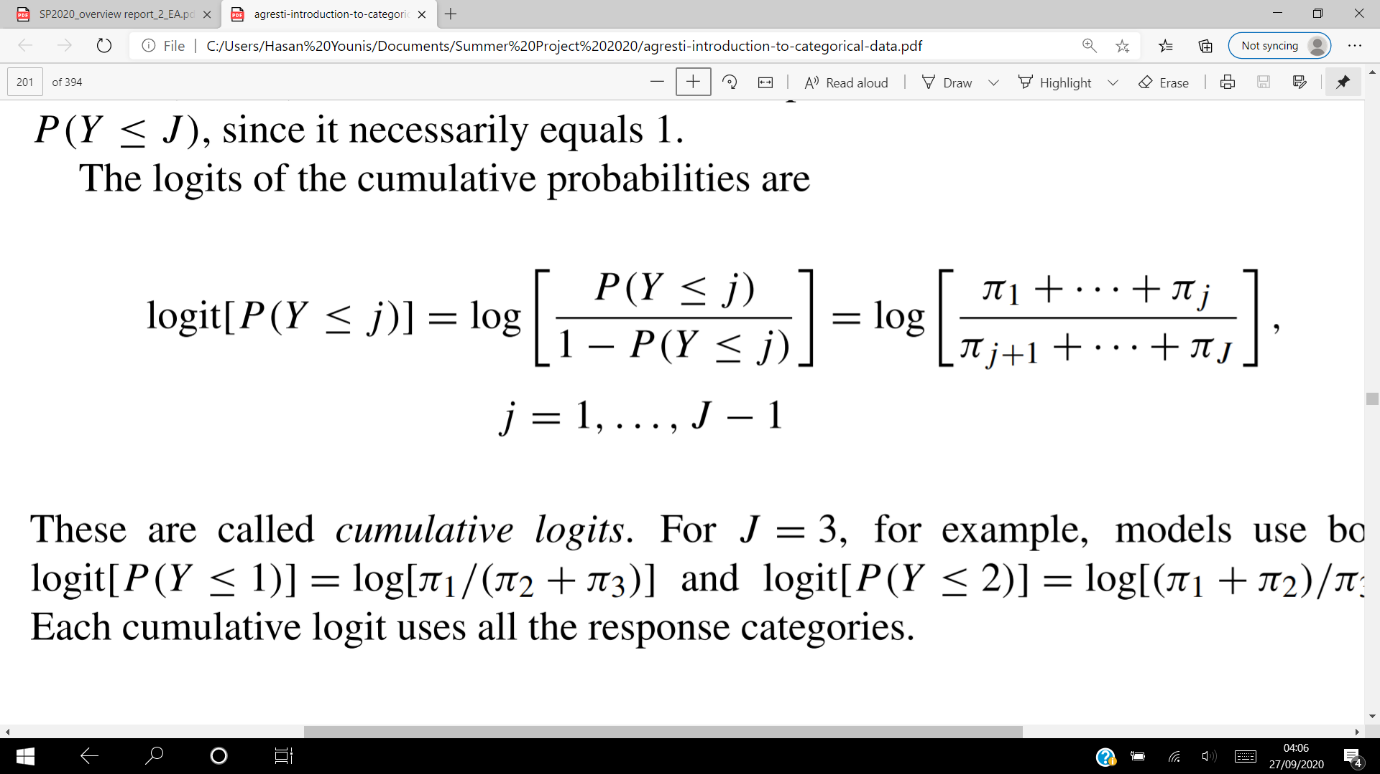
|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| Ethnicity HESA Options | White | Asian | Other including Mixed | Missing | Black |
| Number of Students | 638 | 928 | 127 | 60 | 48 |

|  |  |  |
| --- | --- | --- |
| Disability Options | No Disability | Disability |
| Number of Students | 1634 | 167 |

|  |  |  |
| --- | --- | --- |
| Student Year Options | Year 1 | Year 2 |
| Number of Students | 991 | 810 |

The potential answers to each of the questions are made up of a Likert scale of 1 to 5, with “1” being definitely disagree, and “5” being definitely agree. There was an additional answer “0” for students who believe the question was not applicable them. We cannot infer any information from the “0” answers, and therefore before each regression is run for the respective questions, we removed the “0” answers. This allows us to retain the maximal data for each of the regressions, rather than having to compromise for the sake of having a similar dataset. For the rest of this report, each question will be referred to as their numerical counterpart. Refer to Appendix A, for a key which describes what a certain questions definition is, and which group of questions it is part of.

These 10 key covariates are then regressed across each question using a fixed-effect cumulative logit model for ordinal responses[[1]](#footnote-1), where the answer to each question is defined as the ordinal response variable (being categorical and ordered). The formula that is used to calculate coefficients is shown below:



The function we used to emulate this was the polr() function from the “MASS” R package. Significances will then be determined to establish any understanding for which certain covariates affect question scoring the most. For each of the questions, we will also be fitting a cumulative link mixed model[[2]](#footnote-2) with one random effect (being the programme code) via the Laplace approximation to attempt to further understand the covariate effects. Now, in the mixed-effects model, the programme code is accounted for by a “programme-specific” intercept. We used the clmm() function from the “ordinal” R package to do so. The above fixed-effect model assumes the aforementioned programme-specific intercepts to be a fixed parameter, however, the mixed-effect model assumes these intercepts to be random variables which are independent and identically distributed, with its own variance.

As well as this, a dependence analysis based on the polychoric correlation[[3]](#footnote-3) is run on the same dataset, between all of the pairs of questions available. This was achieved using the polychor() function from the “polycor” R package. Commonly referred to as the Pearson correlation for ordinal data, this will tell us how much the scoring of each question is dependent on the other. This is calculated by supposing each of the question responses is obtained by “categorizing” a normally distributed core variate, and that those two unobserved variates also adhere to a bivariate normal distribution. The maximum likelihood estimate of the previously mentioned correlation is then referred to as the polychoric correlation.

To check if the trend is consistent from Year 1 to Year 2 (and therefore, reasonably transferable to Year 3 students) we will also partition the dataset into a Year 1 only and Year 2 only form, and then re-run the dependence analysis. We shall simply compare the differences between these partitional correlations to the correlations from the full dataset, as well as a quantitative analysis to determine whether any changes from the Year 1 and Year 2 dataset are significant. Therefore, if the difference is small comparing each of the correlation values, we can conclude that the trend is consistent from Year 1 to Year 2. We also will run a hierarchical clustering analysis to better visualise the dependence of each question with each other, and which groups of questions are most dependent on the others.

Results and Interpretation

Regression Analysis

To briefly summarise the following section on the Regression Analysis, we will generate significance tables based on the relative strength of the covariates as the model changes from question to question. This will be completed for an initial fixed effects model, with a secondary look at a mixed effects model where a programme-specific intercept is used, and each covariate is treated as its own random variable. The advantage of this is to reduce the effects of difference in opinion in accordance with the course programme. By adjusting for this variable using the mixed effects model, we believe that it is reasonable to suggest that this effect will be less influential and will give us a purer version of our results.

We will then analyse the possible differences and explanations between each of the significance tables. Significance at the 5% level is denoted by a bold enumeration whereas Significance at the 10% level and blank boxes denote generally considered as non-bold or blank boxes. The direction of the significance is given by the sign inside the brackets (+ or -), with \*, \*\* and \*\*\* denoting significance at the 5%, 1% and 0.1% level, respectively.

Teaching: Fixed Effects Model

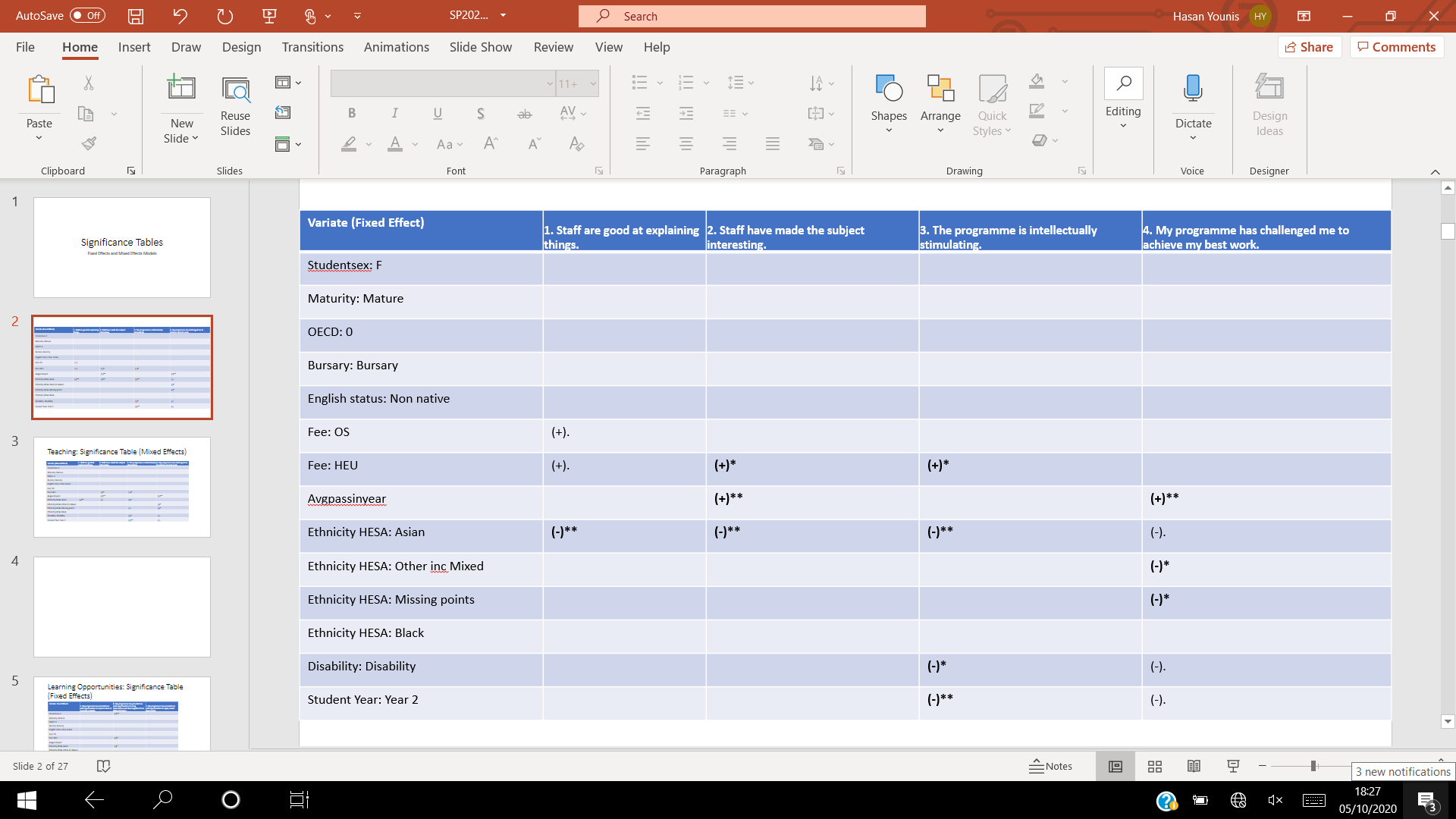


Table 1: Significance table for the Teaching questions, fixed effects model.

It seems with respect to the questions based on Teaching, the student’s sex, maturity, OECD origins, bursary status or English status seem to have little influence on whether the score goes significantly up or down. However, the other covariates seem to influence the scoring but often in different directions. An example is that a student with an Asian ethnicity consistently thinks that the teaching is of a worse quality across these questions in comparison to the baseline with the exception being the fourth question. A more in depth look into the confidence intervals that indeed the fourth question is very close to being significant but ended up falling just short and in the same direction. This seems pretty consistent across the rest of the ethnicities as well, with most of them falling just short of significance outside of the fourth question, but the key is that they are all leaning in the negative direction, meaning all the ethnicities (to varying degrees) feel negatively about the teaching in comparison to the baseline. Other important things to note are that students with a HEU fee status tend to find the teaching better in comparison to the HUK fee status, as well as students that achieve higher average marks. Finally, it seems that students in second year or who are disabled tend to find the courses less intellectually stimulating in comparison to their baselines, but this is not a consistent trend across the rest of the teaching questions.

Teaching: Mixed Effects Model

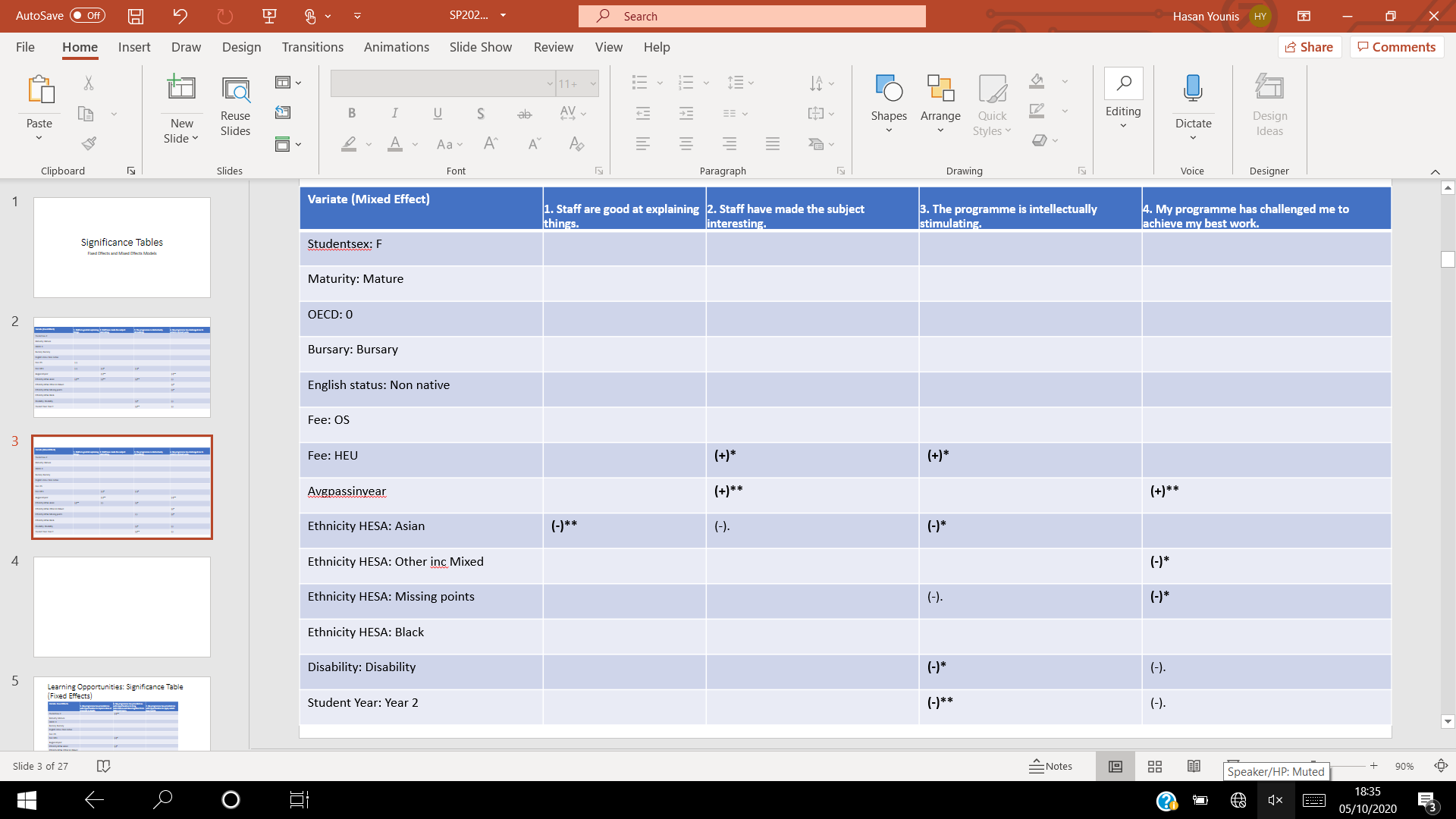


Table 2: Significance table for the Teaching questions, mixed effects model.

Once we account for the programme-specific intercept, as shown above, we can see that largely the same significance structures remain intact. Question 1 sees little impact to the change in model but an important thing to note are for question 2 is that the Asian factor for ethnicity becomes non-significant. This suggests that when the programme code is accounted for, Asians tend to score more happily. For question 3, the missing factor for ethnicity becomes significant at the 10% level (non-significant by the 5% standard we are using) and suggests more so that other ethnicities also feel negatively towards this question in particular. It is also worth noting that the missing points value has a small sample size, which could be a large contributing factor as to why its only significant at the 10% level. Finally, for question 4, both the disability factor and student year 2 become significant at the aforementioned 10% level (non-significant at 5%). This implies that, very similarly to question 3, that both of these factors are contributory towards the negative opinions of students. This is intuitive to understand as the polychoric correlations for these two questions in particular is quite high (second highest polychoric correlation, only bested by the question 1 to question 2 correlation) relative to the rest of the questions.

Learning Opportunities: Fixed Effects Model

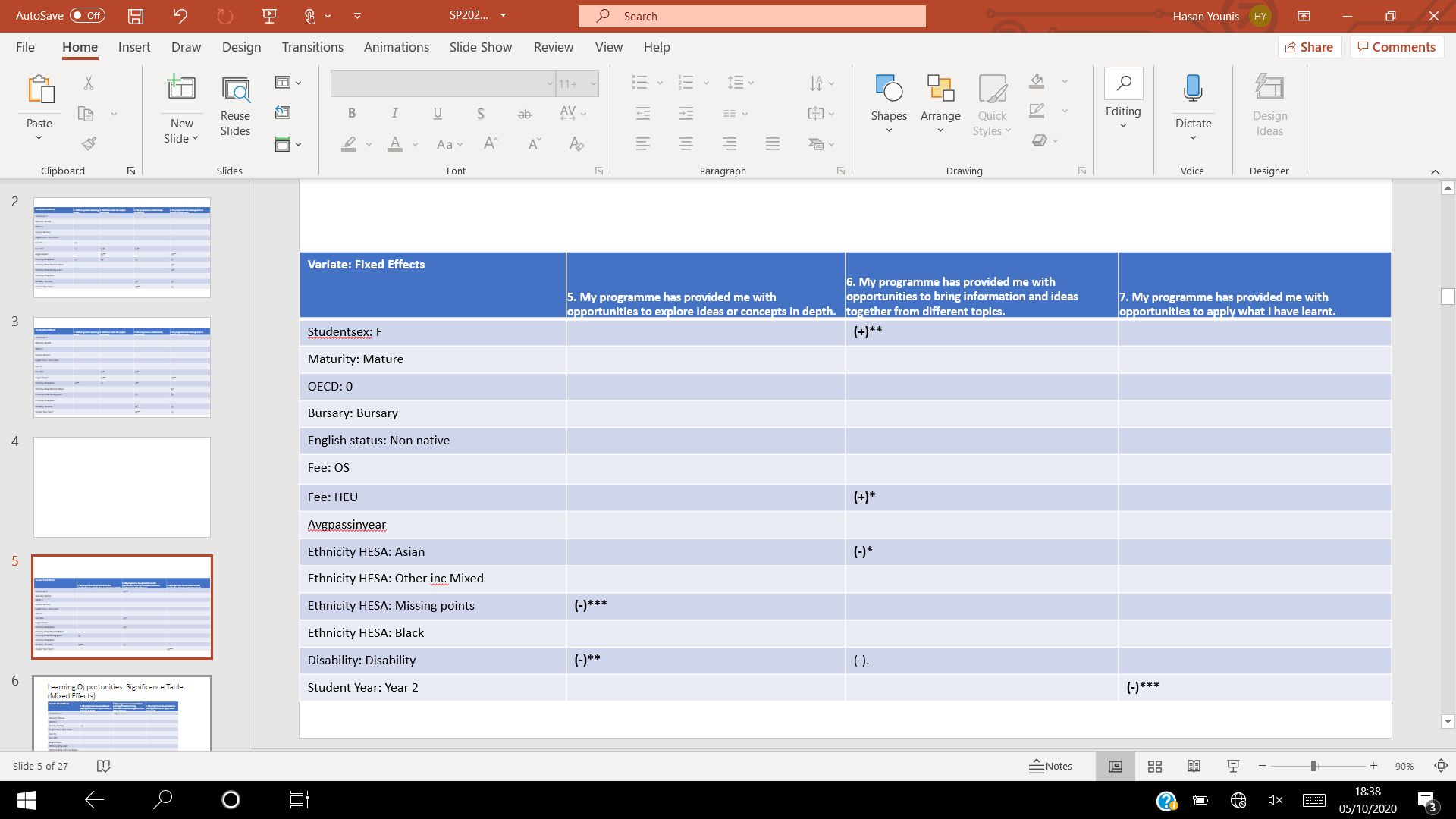


Table 3: Significance table for Learning Opportunities questions, fixed effects model.

With respect to the questions on Learning Opportunities, there is very little in terms of consistent significance in more than one question for one of these covariate factors as shown above. Important things to note are that women feel good about question 6 in comparison to the males, as well as students of the HEU fee status. For question 7, it appears as though students from Year 2 feel particularly unhappy with their scoring. This could be due to being close to the end of their degree having gained less applicable knowledge in comparison to the baseline Year 1 students. Finally, disability seems to play the most consistent impactful variate, with significance in question 5 and near significance in question 6.

Learning Opportunities: Mixed Effects Model

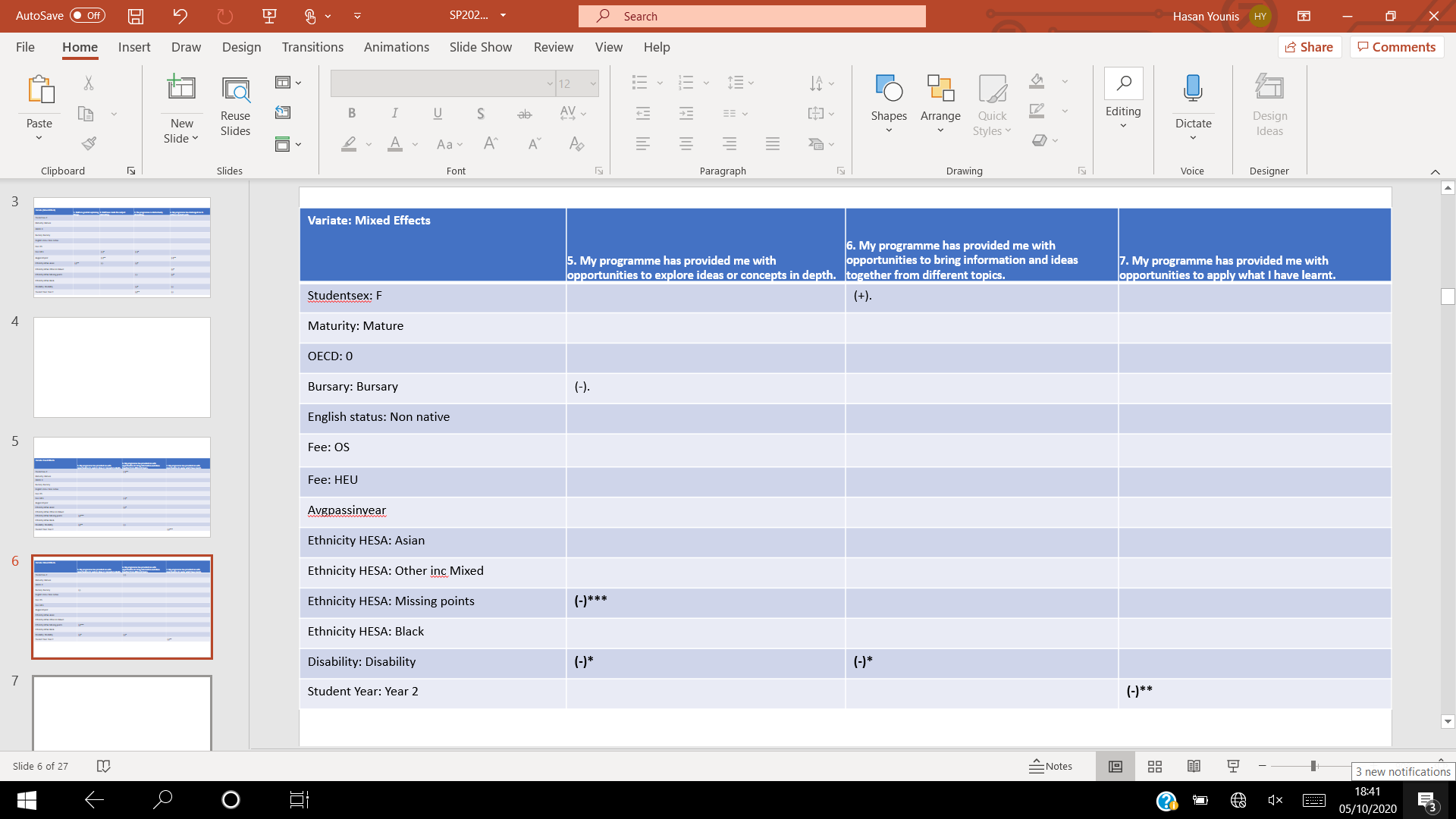


Table 4: Significance table for Learning Opportunities questions, mixed effects model.

In the mixed effects models for learning opportunity, little change in significance can be observed compared to their fixed effect counterparts can be seen for the most part. Particularly in question 5 and question 7, the main observable differences come in the form of some significance at the 10% level (still non-significant overall). In question 5, the bursary factor becomes significant at this 10% level, which is intuitive to suggest given the fact that disabled students are a significant factor in this question consistently, who are more likely to be the recipients of bursaries. This trend follows onto question 7, where the mixed effect model mirrors the significance for fixed effects model. This implies that even when the programme is accounted for, year 2 students still consistently think more negatively towards question 7. This trend breaks for question 6 however, with large changes in significance. EU fee status, Asian ethnicity, and Female student sex all become non-significant (Female students are an exception at the 10% significance level). This is in favour of significance towards the disability factor. The implications from this are that, once the programme codes are accounted for, that the disability factor swallows a lot of the explanatory power for the aforementioned fixed effect factors.

Assessment & Feedback: Fixed Effects Model

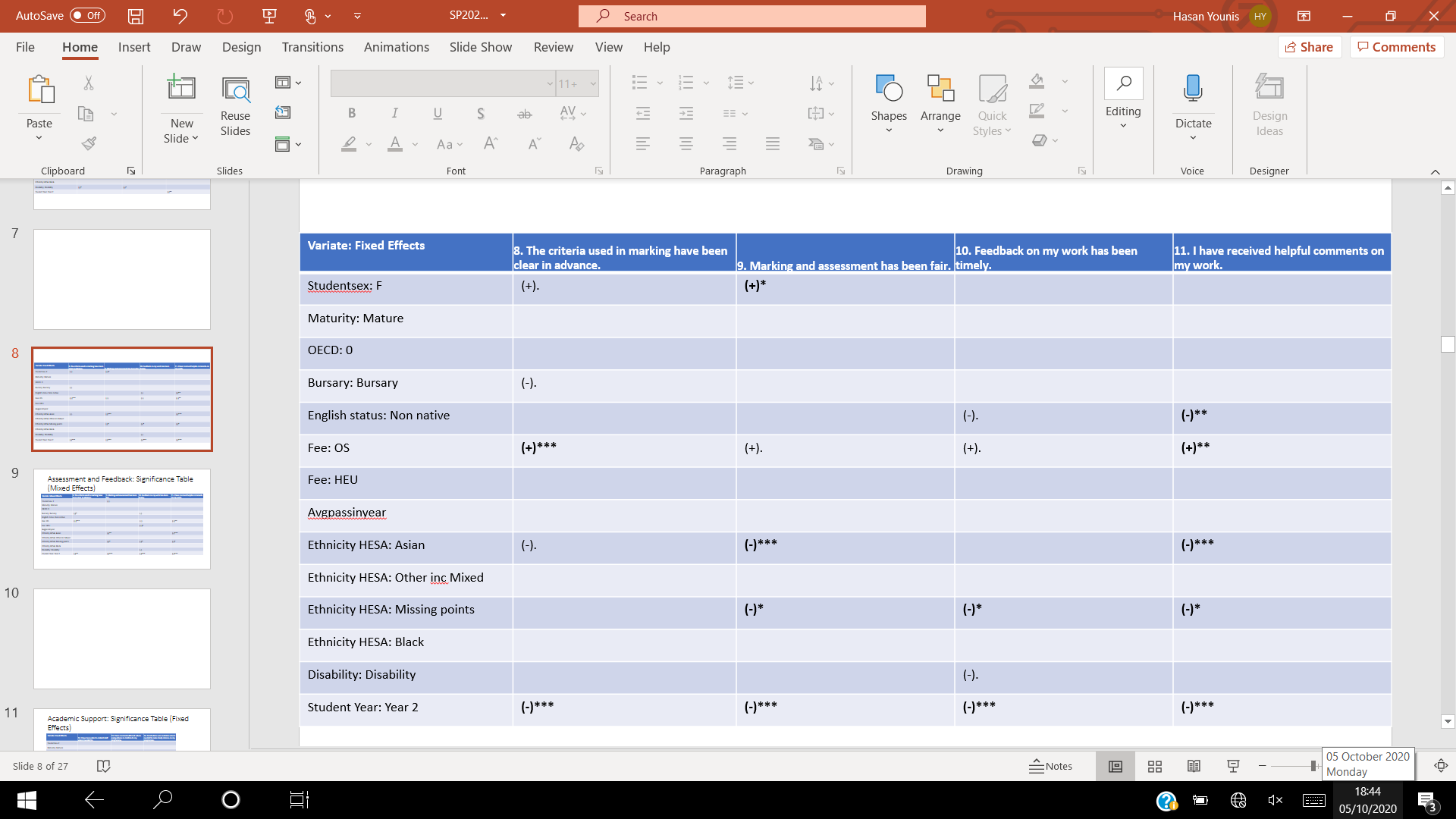


Table 5 Significance table for Assessment & Feedback questions, fixed effects model.

With respect to the Assessment and Feedback questions, more defined trends seem to emerge from the analysis. As is evident from the above, Year 2 students in comparison to the baseline Year 1 students think far more negatively towards assessment and feedback at LSE across the board. The significance here is large and consistent, implying a large change in student attitude amongst second years in comparison to first years. A similar logic can be applied to the overseas fee status, but to a lesser extent, with questions 8 and 11 being significant, with 9 and 10 being near significant. It also consistent that students who have a missing ethnicity also feel negatively towards assessment and feedback. In particular questions 9 and 11, all ethnicities seem to exhibit some type of negative behaviour and in these two particular questions, this can be largely attributed to the lack of sample size especially with respect to the Black ethnicity value. It can also be noted that students who have overseas fee statuses tend to think positively in comparison to the baseline UK fee status.

Of particular importance to us is question 10, which represents the timeliness of feedback to the student. In this particular question, only the missing ethnicity value and year 2 student year value are significant, both of which think negatively of the timeliness of feedback in comparison to their respective baselines. Taking a more in depth look into the confidence intervals suggests that all ethnicities tend to think negatively of the timeliness of feedback in comparison to the baseline, to varying degrees however (with all apart from missing ethnicity and year 2 factors being non-significant). Another important thing to note are that students who are disabled also tend to think more negatively about the timeliness of feedback in comparison to the baseline non-disabled students. This is also somewhat reflected in the bursary factor, which also demonstrates a similar phenomenon to the disability factor but has no significance even at the 10% level. Therefore, it can be reasonably imputed that students with disabilities tend to find the timeliness of feedback worse in comparison to their baselines, despite the overall non-significance in these areas.

Assessment & Feedback: Mixed Effect Models

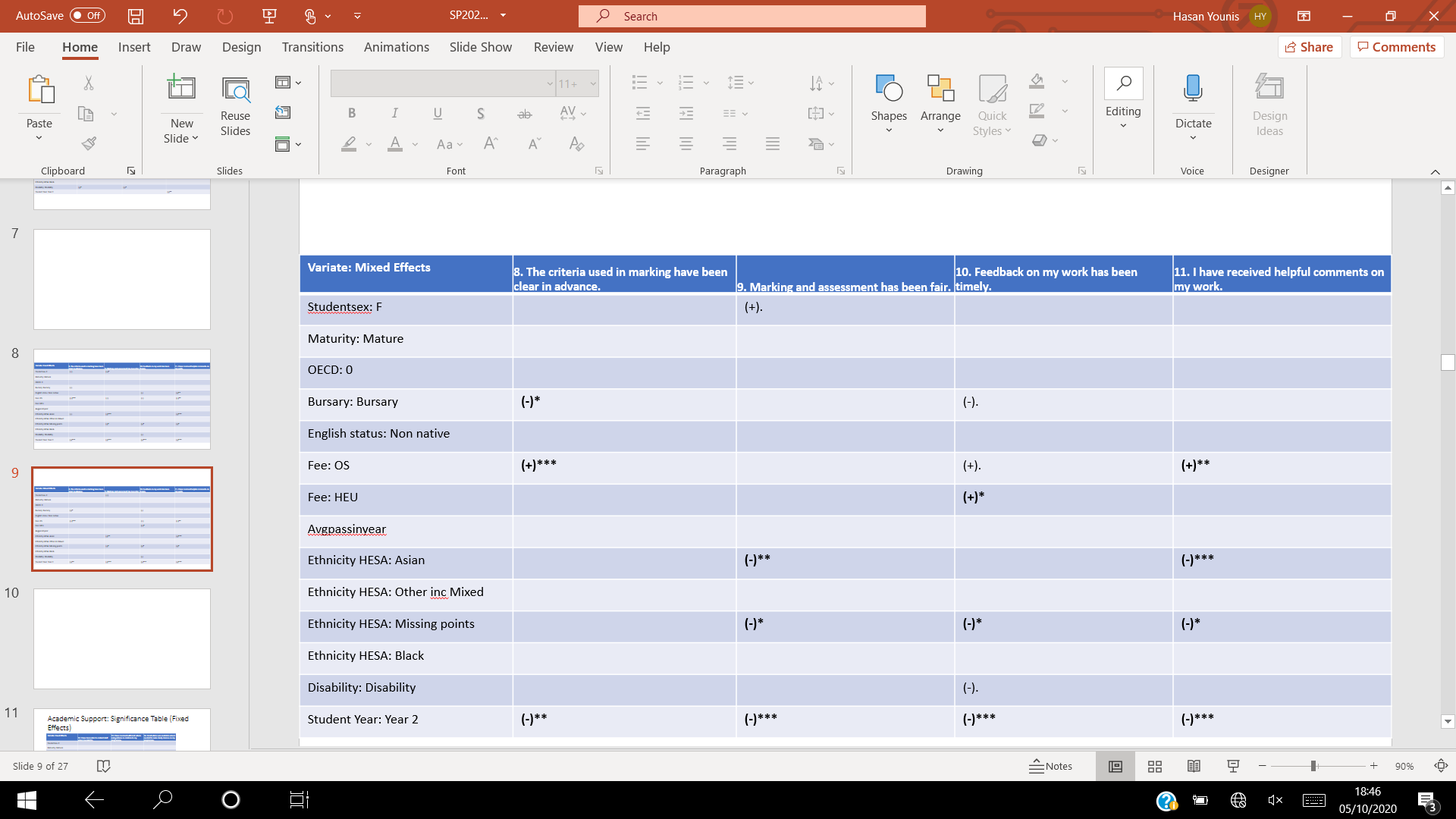


Table 6: Significance table for Assessment & Feedback questions, mixed effects model.

For the assessment and feedback questions, the same significance structures remain intact for the most part. Year 2 students consistently feel negatively towards all of the questions (even up to the 0.1% significance level for the last three questions, 1% level for the first question), and the missing ethnicity factor also consistently feel negatively with respect to the last three questions even when factoring in the programme code. Changes that are important to note are that in question 8, the bursary factor becomes significant, as well as nearly significant (10% level) in question 10. The female factor becomes non-significant overall, but significant at the 10% level in question 9. This implies that student sex plays a large factor in the scoring of assessment and feedback depending on the course. Similarly, in question 11, the non-native English status factor becomes non-significant even including the 10% level. This can partly be explained by the fact that very similar factors (particularly fee status overseas, Asian ethnicity) are also significant. Finally, for question 10, the current significance structure at the 5% level remains intact, but other factors seem to play more of a factor. Bursary, overseas fees, and disability factors all become significant at the 10% level (not significant overall) and EU fees does become significant at the 5% level. Programme code therefore seems to play a large part in the factors of students that score more or less.

Academic Support: Fixed Effects Model

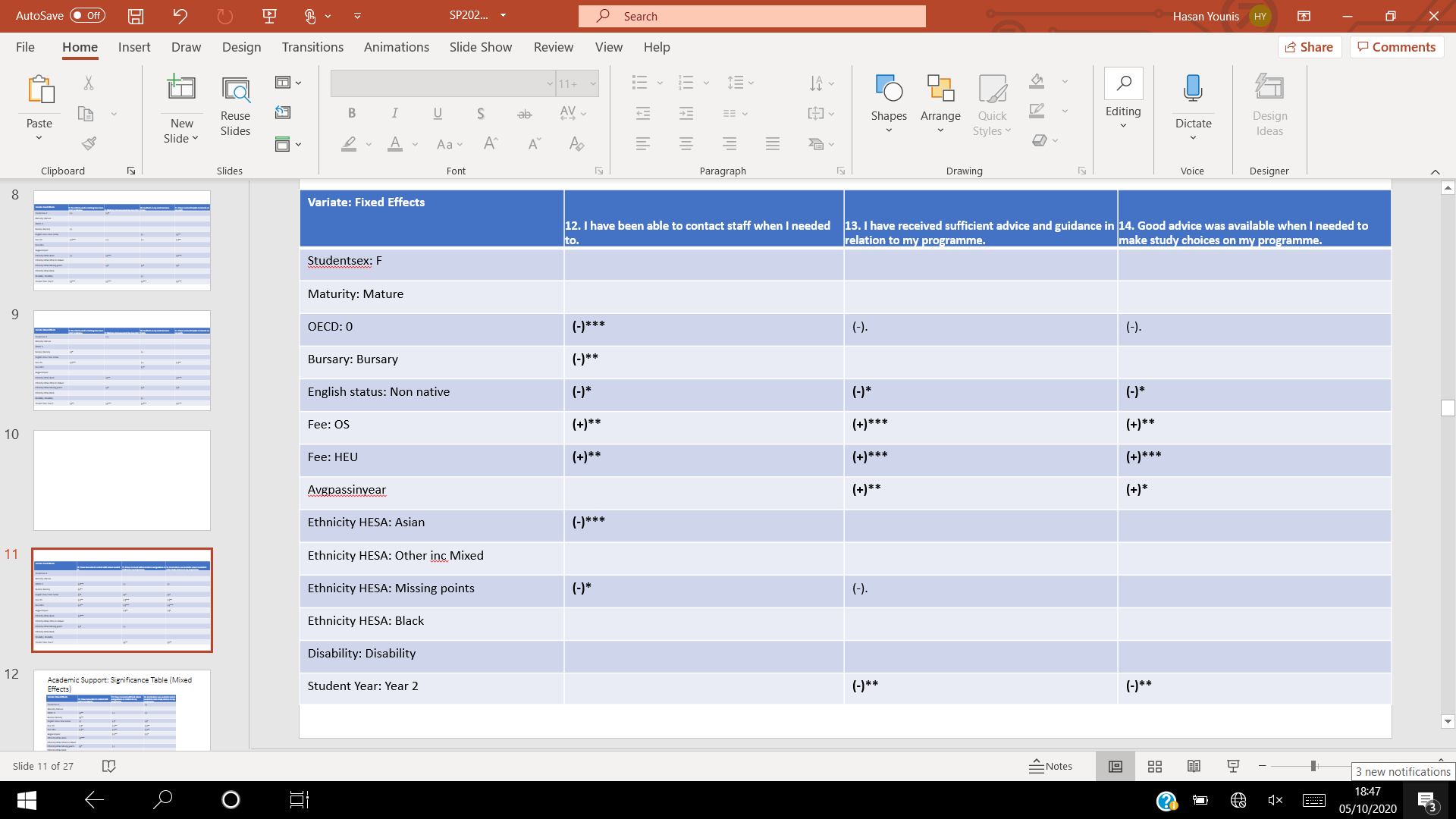


Table 7: Significance table for Academic Support questions, fixed effects model.

In the academic support questions, there seems to be a lot more in terms of significant trends in comparison to the other groups. A key area that should be noted are that students who do not have English as their native language think negatively of the academic support on offer in comparison to the baseline native English speakers. This is an area of concern, since the common language at LSE located in the UK will be English and that can often mean students who may be less comfortable with English are alienated as a result. This is in contrast however to students who pay an EU or Overseas fee. These students seem to feel positively about the academic support consistently, which does not an intuitive proposition since these same students are more likely to not have English as their native language. This implies that a lot of the foreign students tend to either definitely agree or definitely disagree with respect to academic support, with few opting for the in-between options, creating this interesting dichotomy.

Other important things to note are that students who average higher marks tend to think more positively about the academic support structures whereas students who are in their second year tend to think much more negatively. It can also be noted that maturity, ethnicity, and disability seem to have little effect on the scoring of these questions (besides question 12 with ethnicity).

Academic Support: Mixed Effects Model

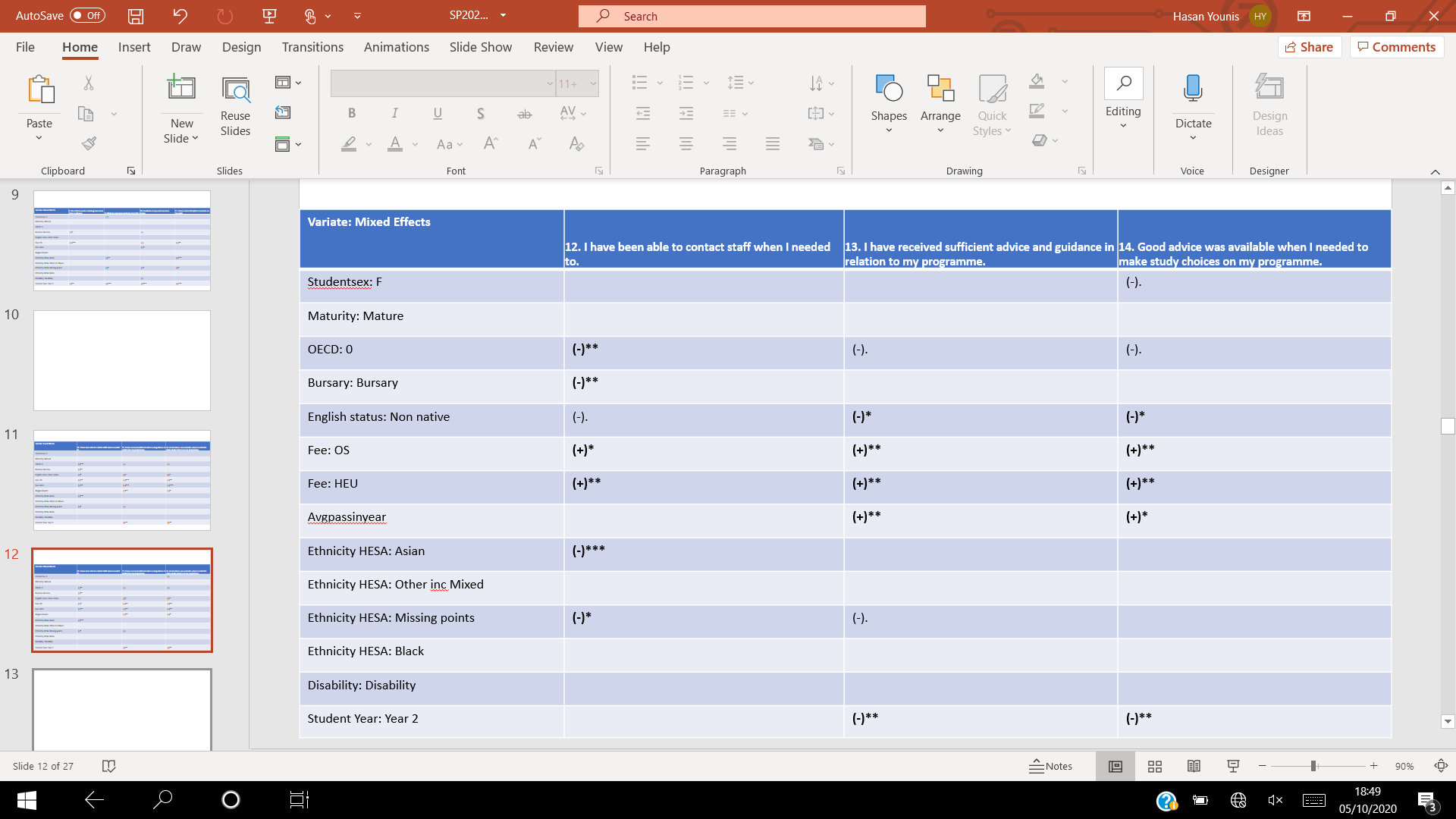


Table 8: Significance table for Academic Support questions, mixed effects model.

The mixed effect model seems to have little effect on the overall significance of the fixed effect models but seems to highlight certain underlying traits that are more difficult to see. An example of this is the OECD factor, which is significant at the 10% level in both questions 13 and 14 and is not highlighted as such in the original fixed effect models. This seems to suggest that if the student comes from a non-OECD country that it tends to have effects that are only captured when the programme specific intercept is introduced, explaining that its mostly captured by other factors such as an overseas fee status in the fixed effects model.

Organisation & Feedback: Fixed Effects Model

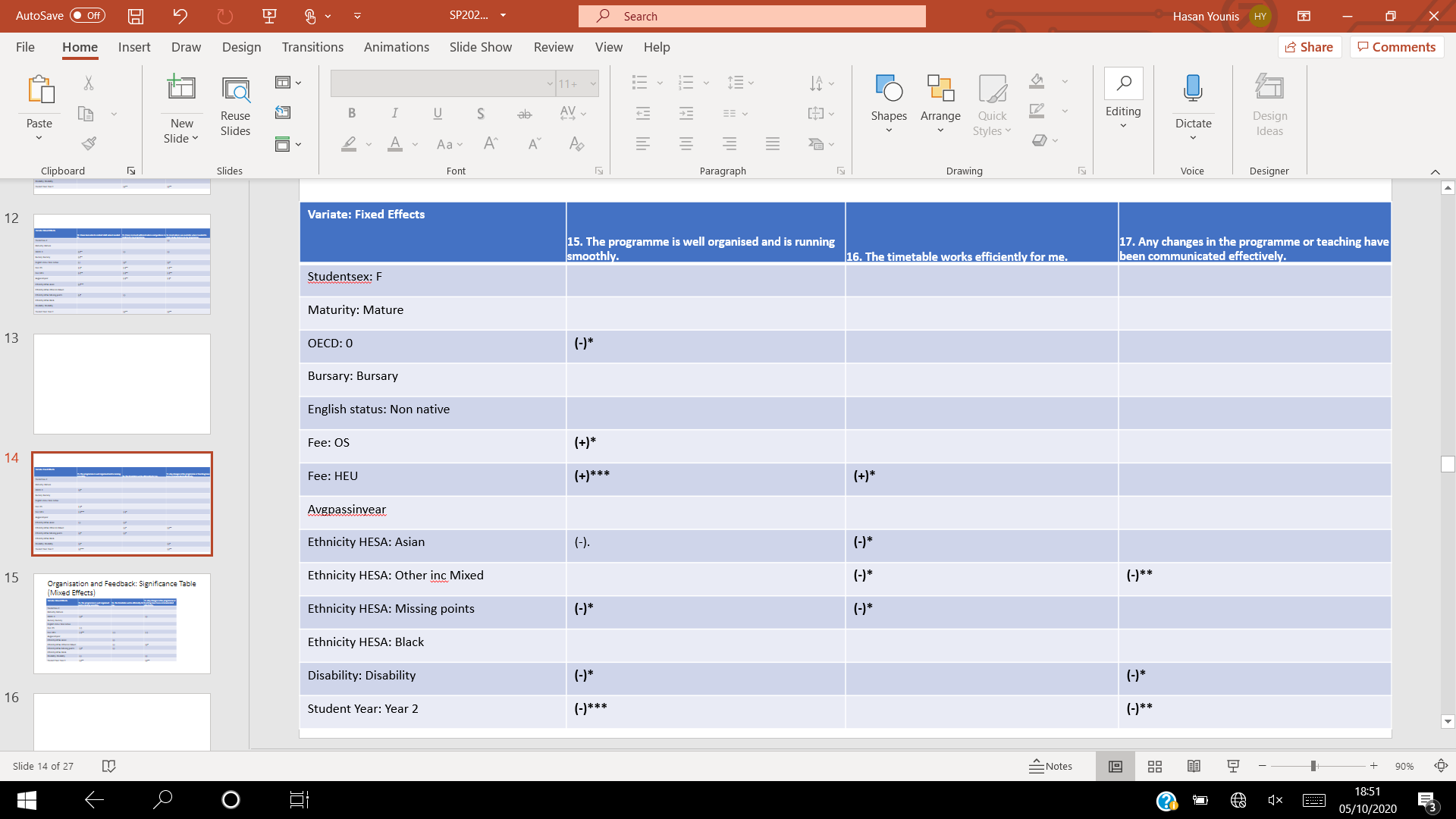


Table 9: Significance table for Organisation & Feedback questions, fixed effects model.

For the organisation and feedback questions, there tend to be more in the way of general trends. Student sex, maturity, bursary status, English status and marks on the whole seem to have little effect on the outcome of these questions, in comparison to the other variates fee status, ethnicity, disability and student year which tend to have more of an impact. It can be seen that students which have an overseas or EU fee status tend to think more positively towards the organisation and feedback of the university, but this is contrary to the ethnicity values which on the whole demonstrate a more negative perception of the organisation and feedback. It can also be witnessed that having a disability or being in year 2 is also linked with negative opinions towards the organisation and feedback in comparison to their respective baselines.

Organisation & Feedback: Mixed Effects Model

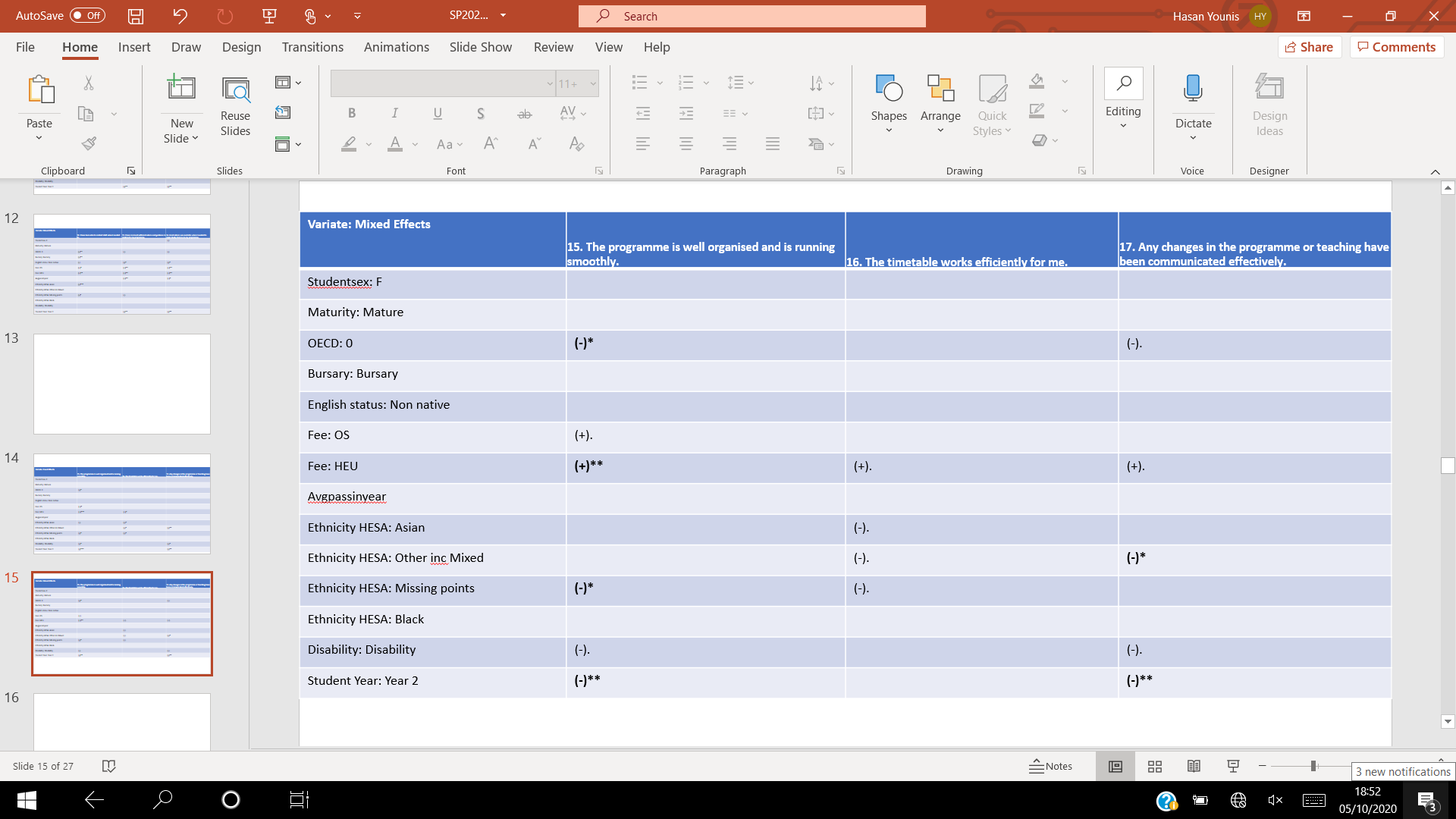


Table 10: Significance table for Organisation & Feedback questions, mixed effects model.

A consistent trend we can see in the mixed effects model is that when the programme is accounted for, a lot of the significant variates at the 5% level become non-significant. It seems as though the programme specific intercept subtracts from the severity of the scoring, be it positive or negative. Starting with question 15, both the overseas fee status and disability factors become non-significant at the 5% level (significant at the 10% level). The remaining factors sustained their significance. An interesting effect is that for question 16 (which we will find out later, tends be the least dependent question in terms of correlation) loses all of its significant factors. None of the factors that are significant at the 5% level in the fixed effects model remain significant in the mixed effects model. Therefore, it is reasonable to suggest the explanatory power of our variates do not properly explain the results of question 16 that well. Finally, for question 17, we lose significance at the 5% level (still significant at the 10% level) for the disability factor. This gives rise to some other near significance in both the EU fee status and OECD factors.

Learning Resources: Fixed Effects Models

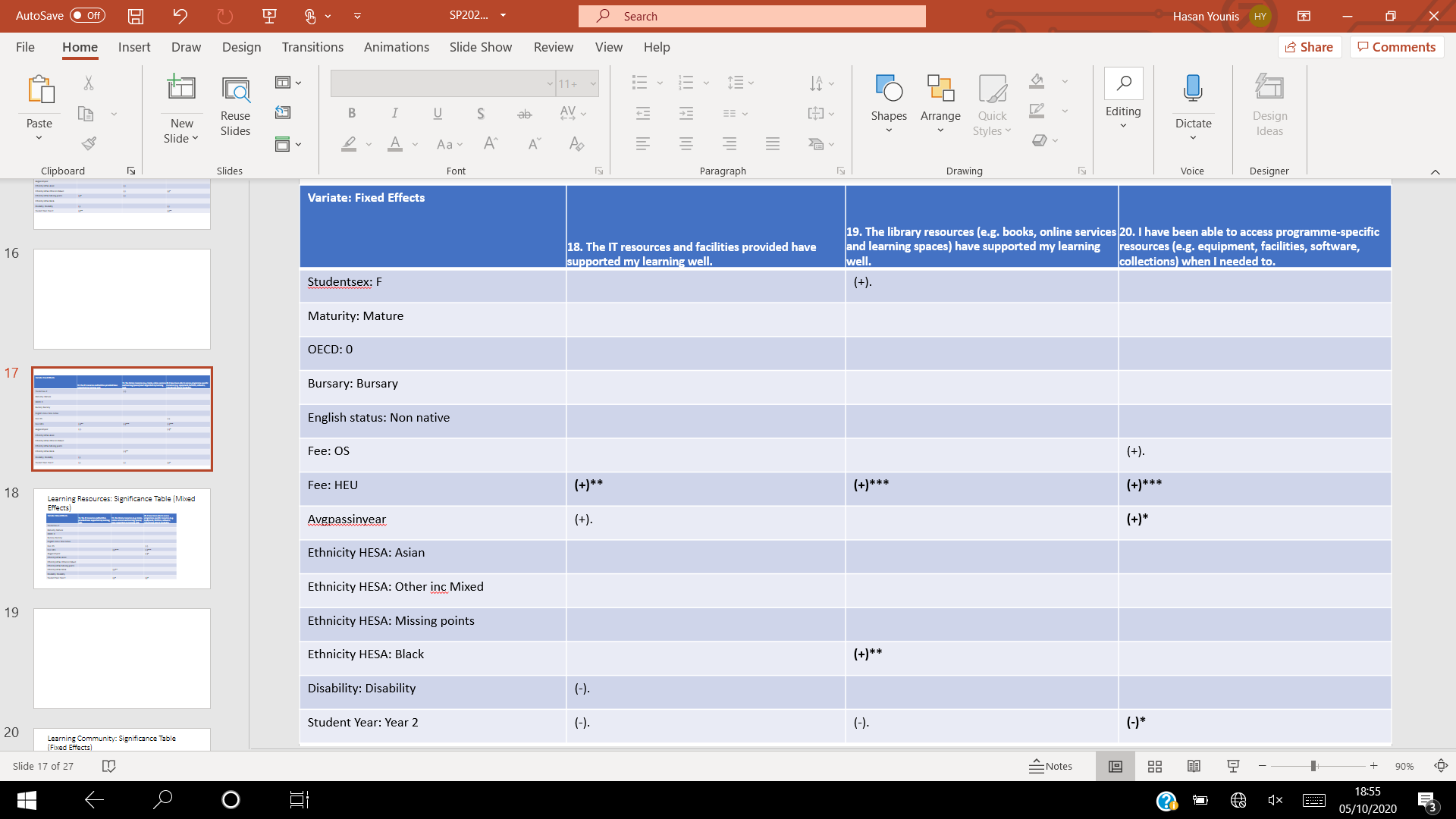


Table 11: Significance table for Learning Resources questions, fixed effects model.

For the learning resources questions, there is not much to analyse in terms of overall significance. Despite this, it is evident that students with an EU fee status tend to think consistently and largely in favour of the learning resources in comparison to the baseline. An important thing to note is that for question 19, students of black ethnicity think positively of the library resources, despite the general non significance of other ethnicities. Since there was little data for black students in comparison to the rest of the ethnicities listed, this could be down to a lack of sample size rather than demonstrating any real effect.

Learning Resources: Mixed Effects Model

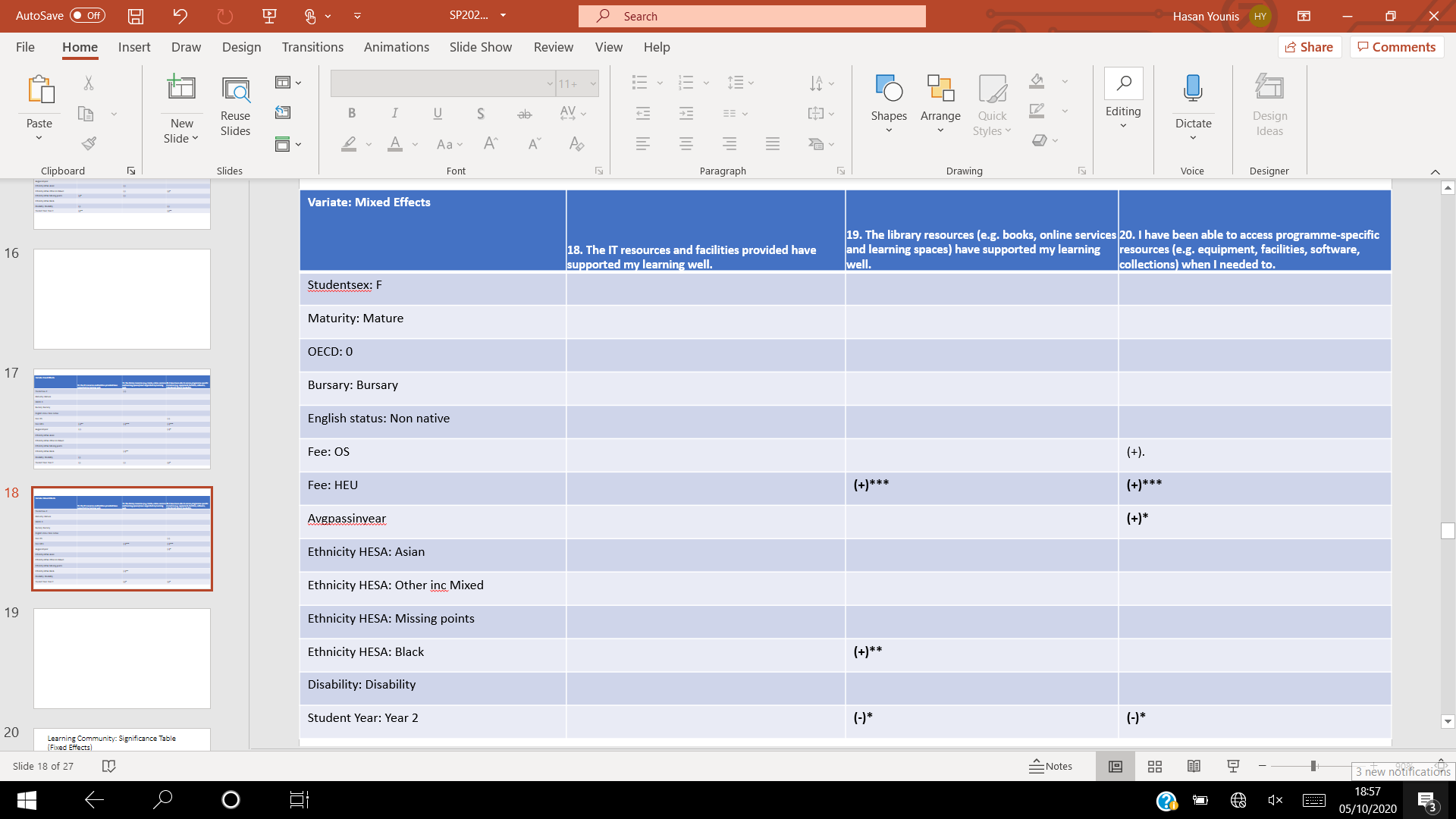


Table 12: Significance table for Learning Resources questions, mixed effects model.

For the mixed effects models in learning resources, accounting for the programmes has increased the effect of certain variates whilst maintaining the impact of the variates which were significant in the fixed models. In question 19, for example, the EU fee status and Black ethnicity factor remained significant, but the year 2 student factor also became significant in the mixed effect model. This trend remains consistent for question 20 as well, with EU fee status, average pass in year, and year 2 student factors all remaining significant. In addition to these items, the overseas fee status also achieves near significance (significant at the 10% level). No significance was recorded for question 18.

Learning Community: Fixed Effects Model

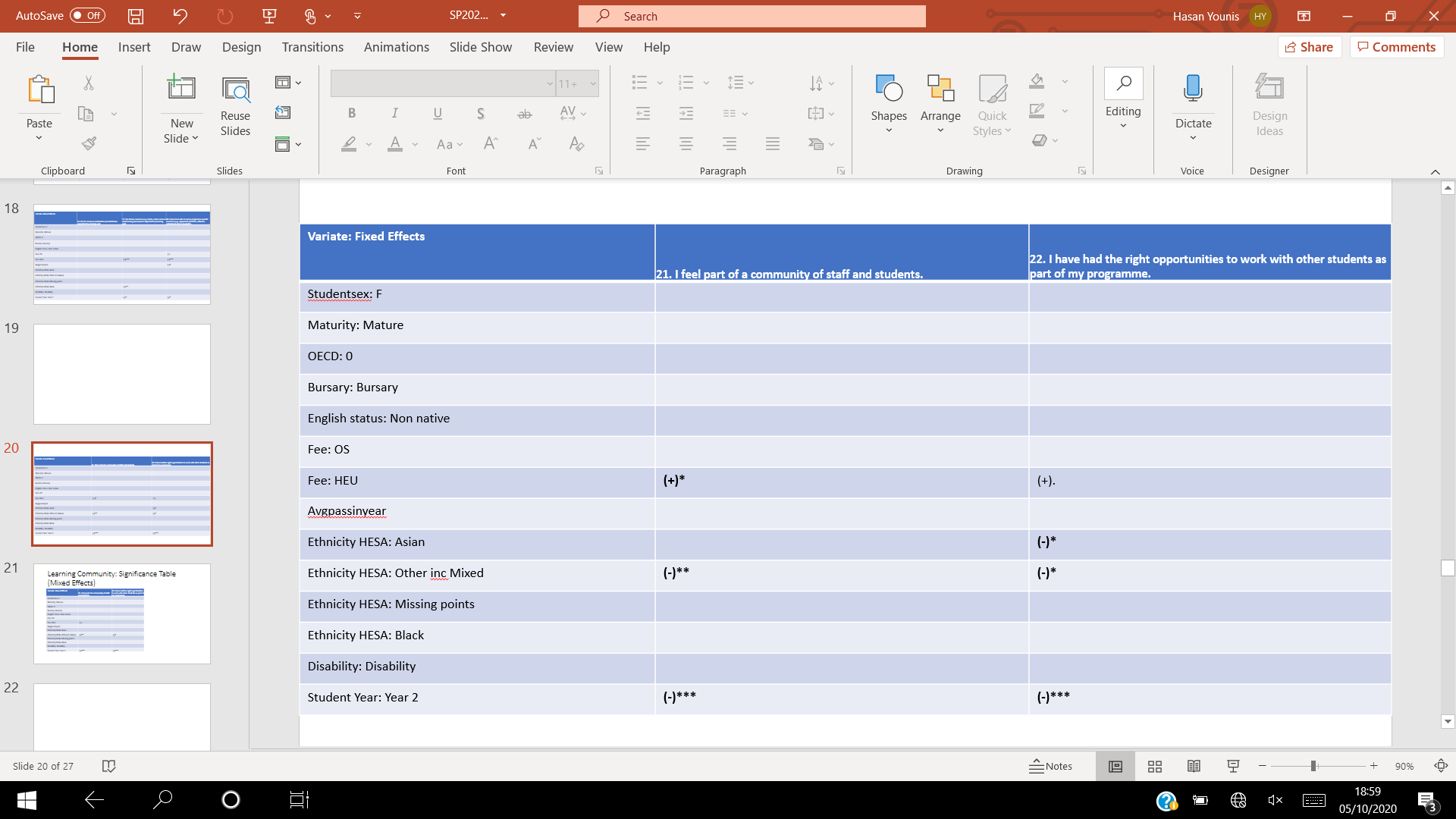


Table 13: Significance table for Learning Community questions, fixed effects model.

Learning Community is of particular importance at LSE, so despite the general non-significance, we will focus in detail on the significance shown. In particular, both the other ethnicity value and year 2 students think negatively of the learning community in comparison to their respective baselines, with an EU fee status showing positively related significance for question 21 only and the Asian ethnicity value picking up a negative significance for question 22. After diving deeper into the confidence tables, it seems that for all of the ethnicity values for both questions 21 and 22 that they tend to think more negatively than positively towards the learning community, just not enough to be significant. This can in large part due to limited sample sizes, but the general trend in the non-significance is closer to negative and can be observed in the significance of the other ethnicity value. Another aspect is that the year 2 students also think negatively of the learning community in comparison to the baseline. This can be attributed to an often-larger workload and less free time in general.

Learning Community: Mixed Effects Model

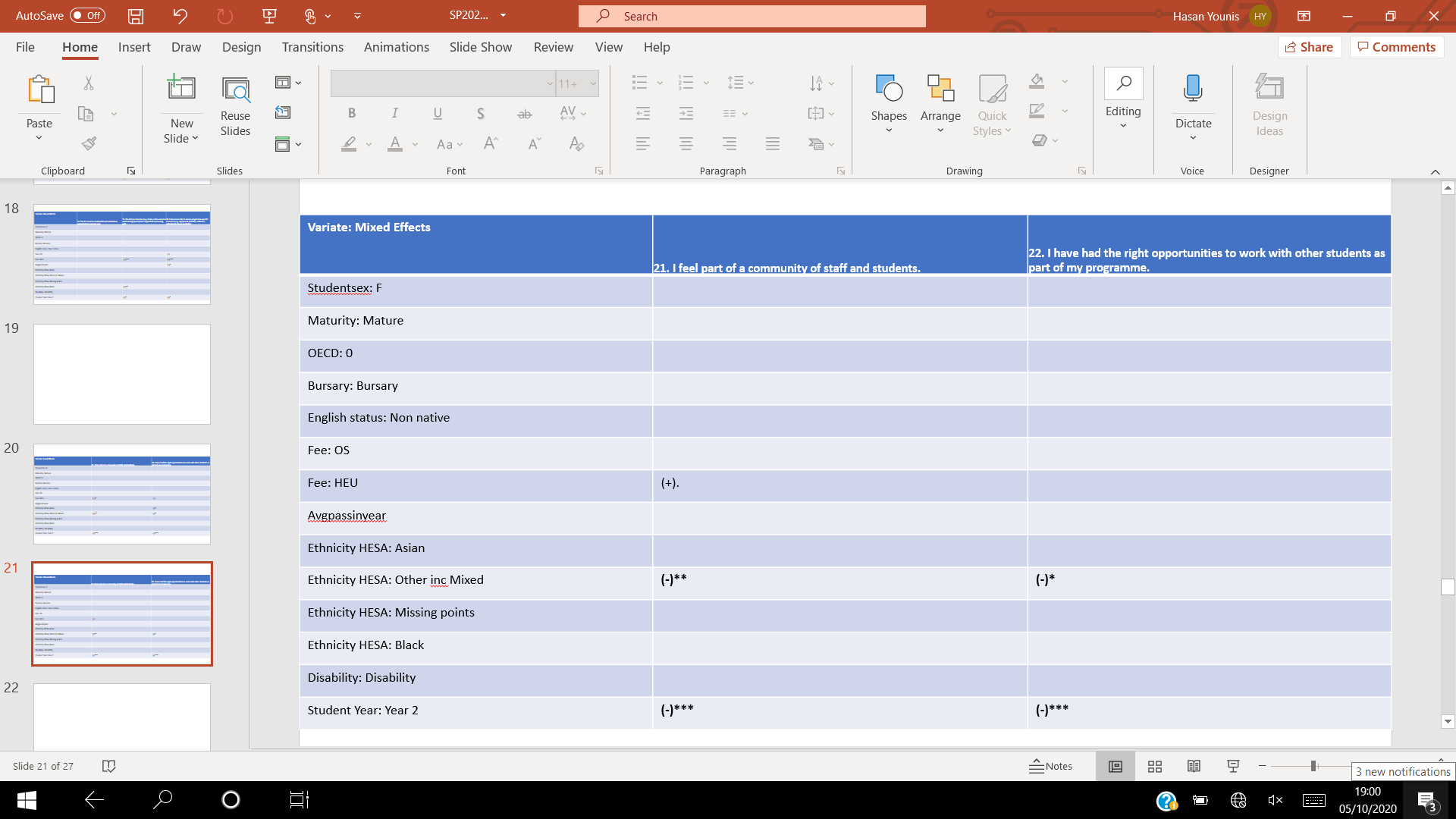


Table 14: Significance table for Learning Community questions, mixed effects model.

The mixed effects models for learning community seem to consistently subtract from the importance of significant predictors in the fixed effects model. This seen in both question 21 and question 22. In the former, the other ethnicity factor and year 2 student factors both remain significant, however the EU fee status is only significant at the 10% level (non-significant overall). This is consistent with question 22 as well, with the Asian ethnicity factor losing complete significance in the mixed effects model, with the other two factors (other ethnicity and year 2 students) remaining significant

Student Voice: Fixed Effects Model

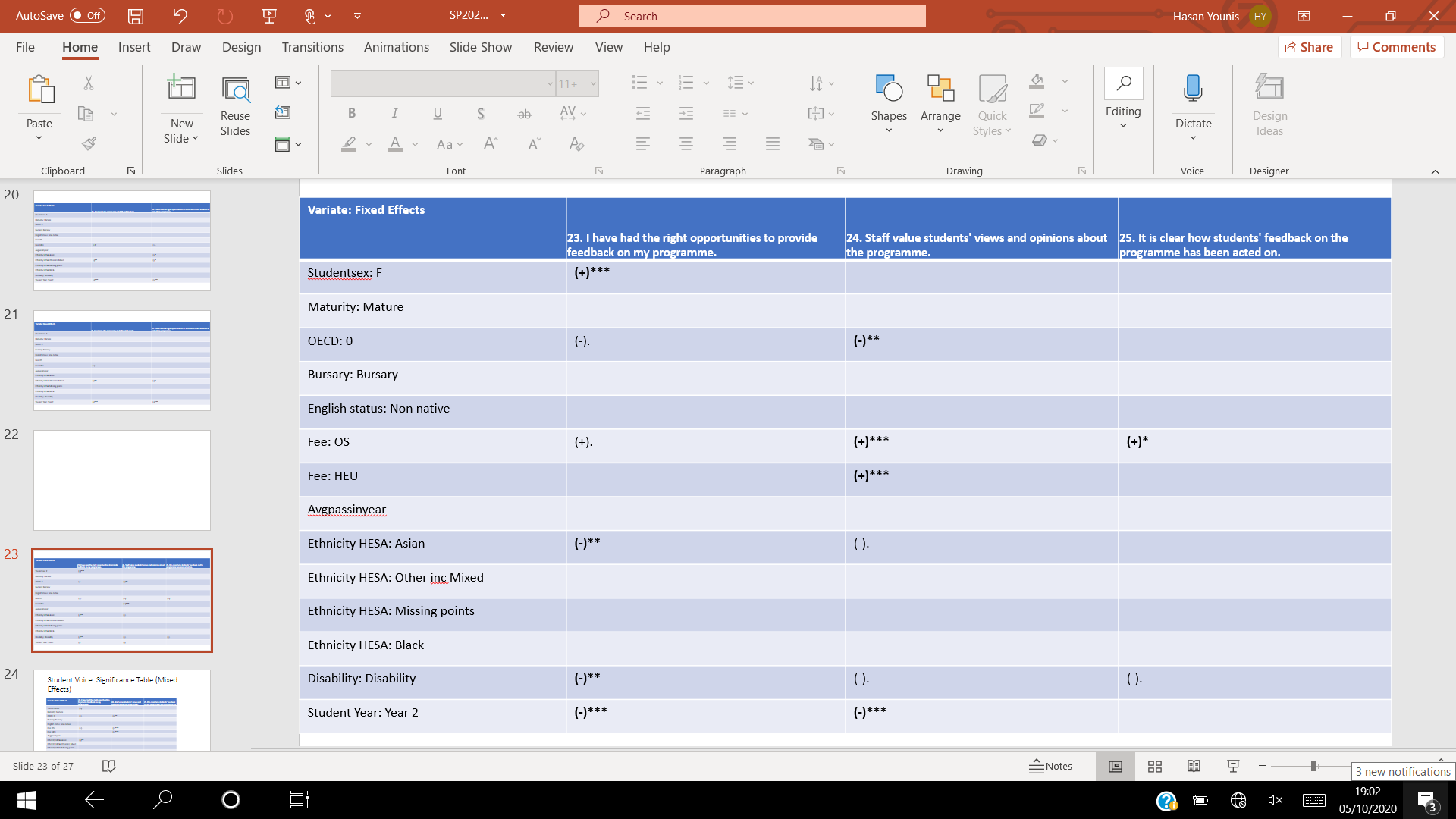


Table 15: Significance table for Student Voice questions, fixed effects model.

For the questions regarding student voice, it seems as though trends differ by question, with a couple of factors affecting all of the questions with different factors in each question. It is consistent that having an overseas fee status means that you think positively about the student voice questions, with question 23 being the only exception (but was very close to being positively significant as well). The same logic can be applied to year 2 students, but on the whole, they thought negatively of the student voice in comparison to the baseline year 1 students. It is also important to note that disability despite only having significance in question 23, was very close to being negatively significant for the rest of the questions as well.

If we take particular focus on Question 23, we can see that females tend to think positively with respect to providing feedback about the programme. Generally speaking, as well, despite a large section of non-significance, having a different ethnicity to the baseline also tended to score negatively about opportunities to provide feedback. It is also evident that having a disability or being in year 2 also significantly effects the scoring negatively in comparison to the baselines.

Student Voice: Mixed Effects Model

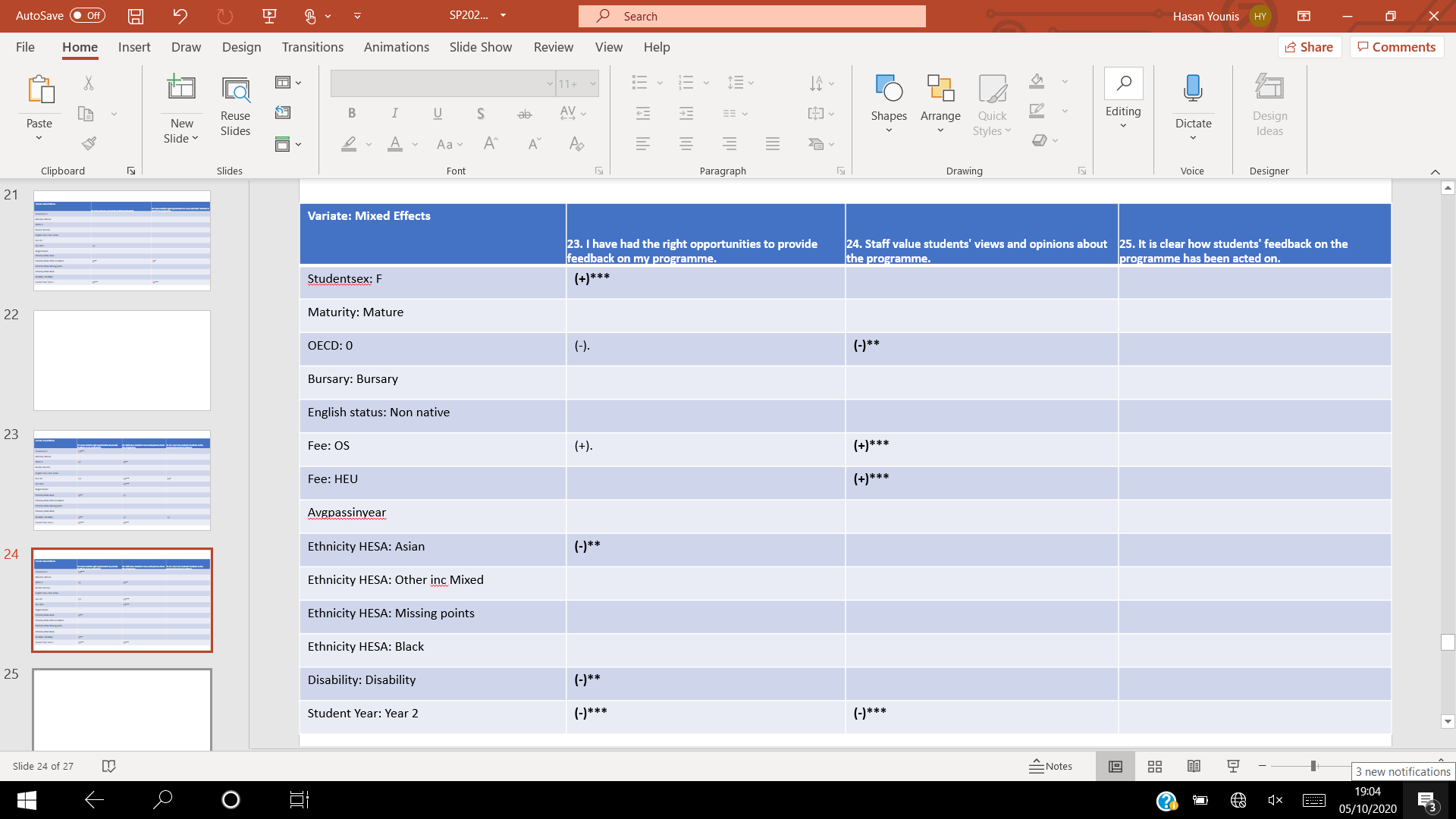


Table 16: Significance table for Student Voice questions, mixed effects model.

For the mixed effects models for student voice, the effect of the programme specific intercept seems to change depending on the question. Question 23 for example, see rises in the significance of the female student sex, Asian ethnicity, disability, and year 2 student factors as well as the near significance of the OECD factor and overseas fee status. This remains consistent for question 24, with all of the factors that are significant in the fixed effect model also significant in the mixed effects model except there significance level seems to be more influential. This trend breaks for question 25. In the mixed effects model, no factor is significant in determining the scoring for question 25.

Final Two: Fixed Effects Model

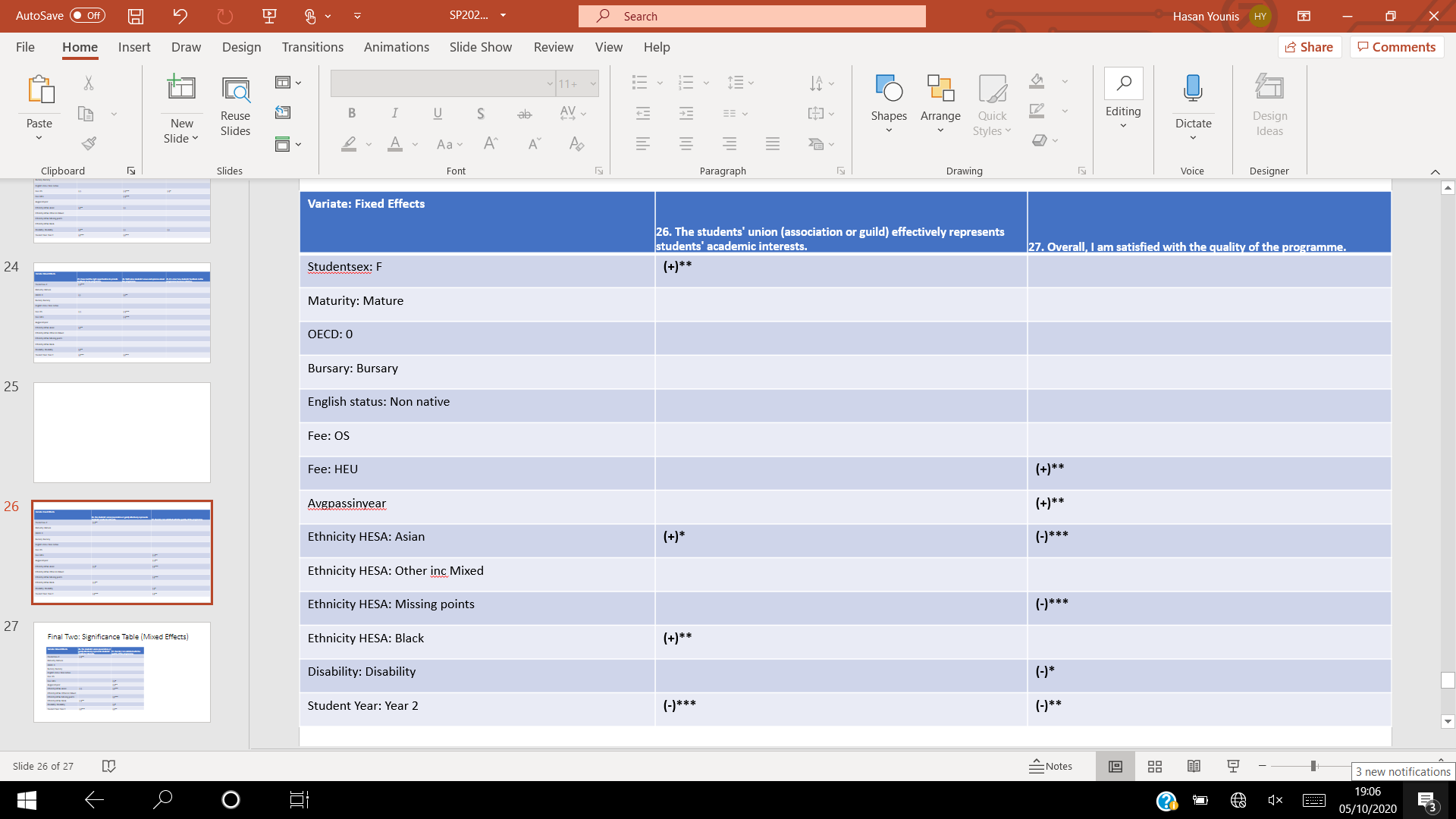


Table 17: Significance table for the Final Two questions, fixed effects model.

For question 26 which represents the student union, it appears as though Females and for the most part all of the Ethnicity variates seem to think positively of how they represent or portray their feelings. This is in contrast to year 2 students who tend to think negatively of how the student union represents their academic interests. For question 27 which represents overall satisfaction, it seems as though consistently, students who have an EU fee status score more highly as well as people who are score higher in their exams which in particular makes intuitive sense. Despite this, it is evident that for the ethnicity factors, disability factor and if you are in student year 2, you are consistently less satisfied overall in comparison to their respective baselines.

Final Two: Mixed Effects Model

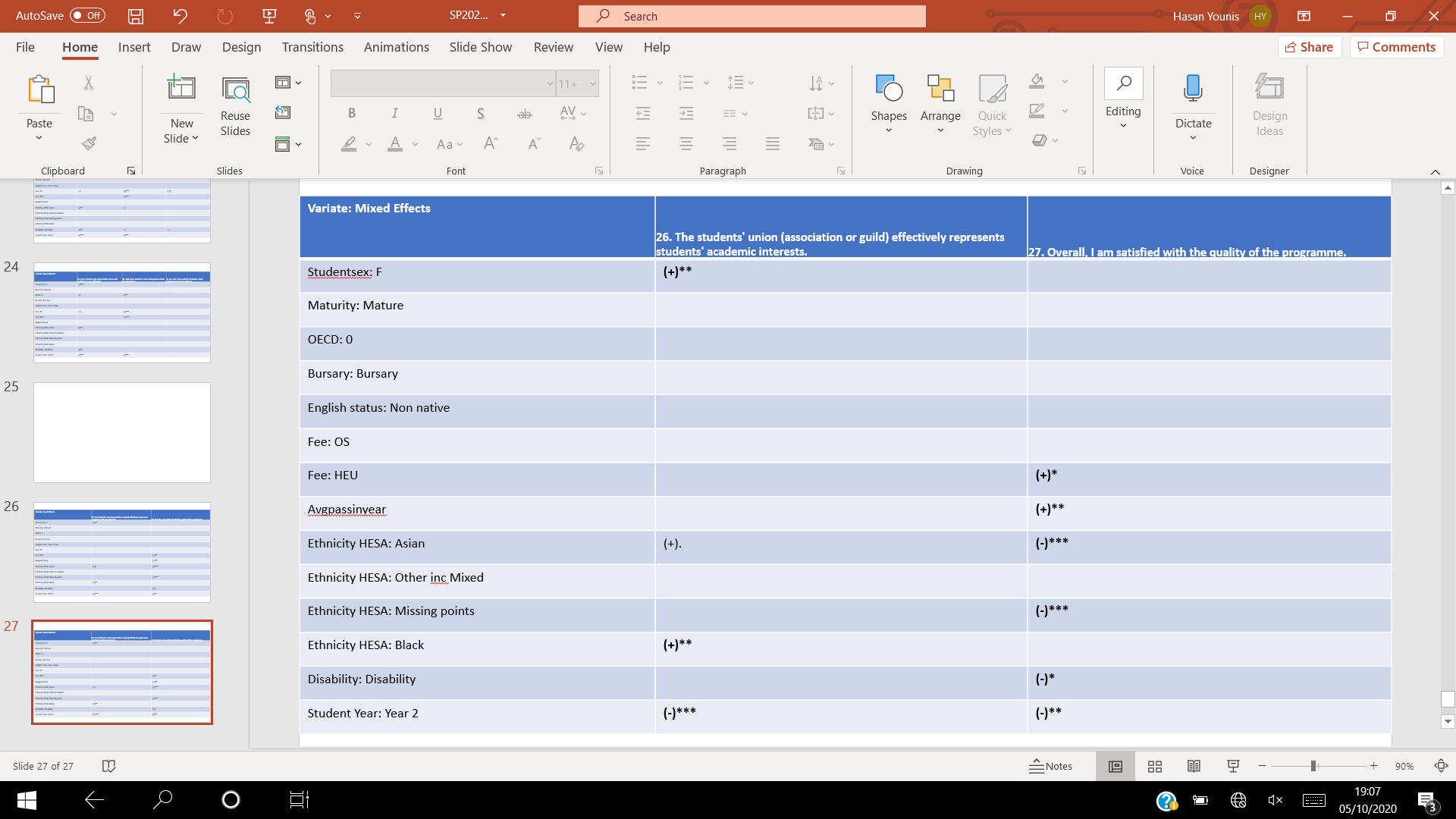


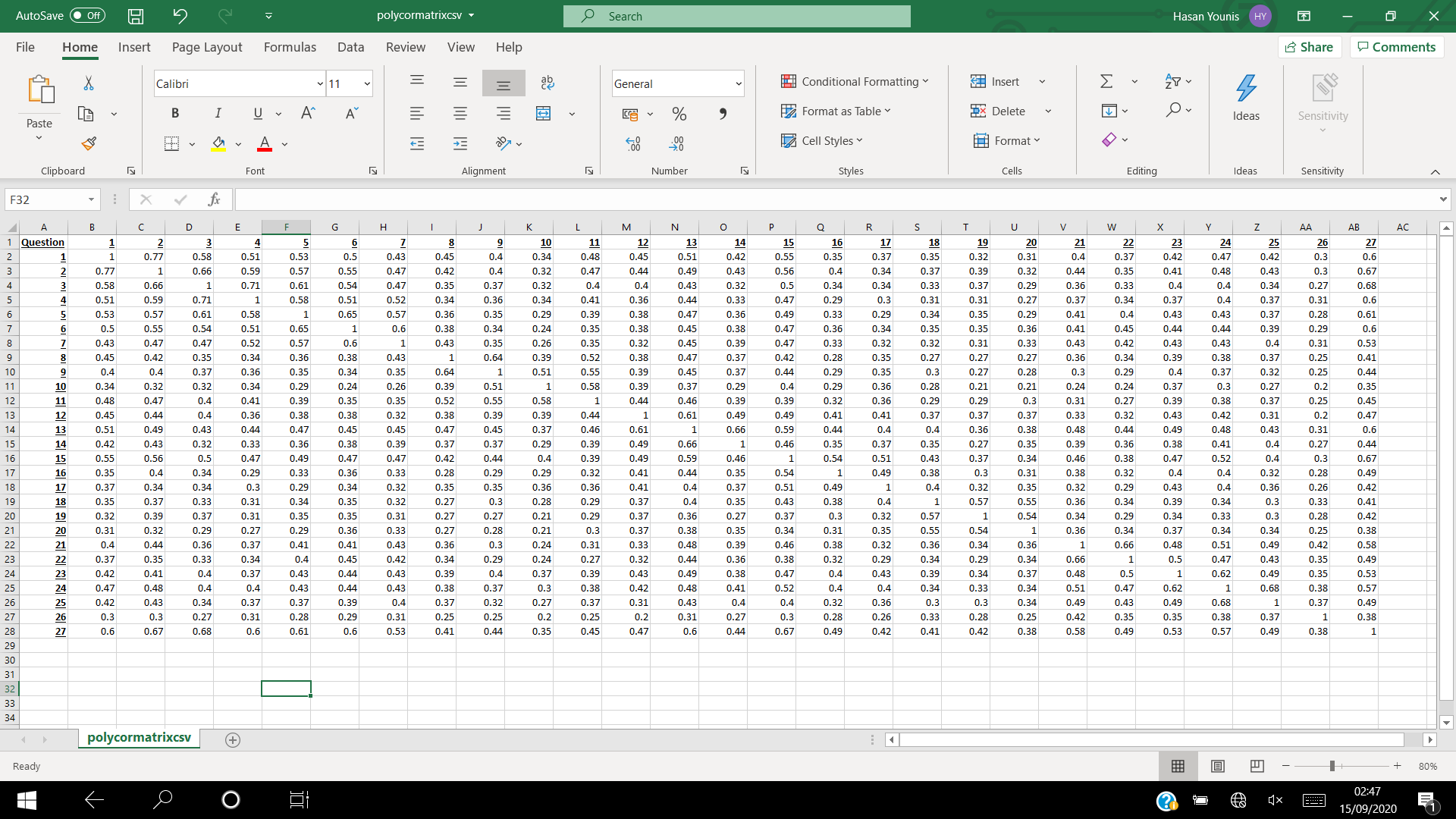
Table 18: Significance table for Final Two questions, mixed effects model.

For the final two questions, the significance structures in place via the fixed effects model seem to have not been affected too heavily by introducing the programme specific intercept. Question 27 for example has exactly the same factors that are significant in both the fixed effects and mixed effects model. This is also mostly consistent with question 26 as well, with the only difference being the near significance of the Asian ethnicity factor when it was previously significant.

Dependence Analysis

In this dependence analysis, we will explore via the polychoric correlation the dependence of each individual question on the others, and also explore this in more depth using hierarchical clustering. Using this unsupervised machine learning algorithm, we will group together questions based on their dependence on the other, allowing us to create a more useful, general approach to improving scores in other areas, with an example being an efficient way of improving scores for the teaching questions by improving the assessment & feedback scores, all things even. Finally, we will analyse the differences between a partitional Year 1 and Year 2 dependence analysis, to see if results from year 1 and year 2 can be generalised to a dataset which includes Year 3 students.

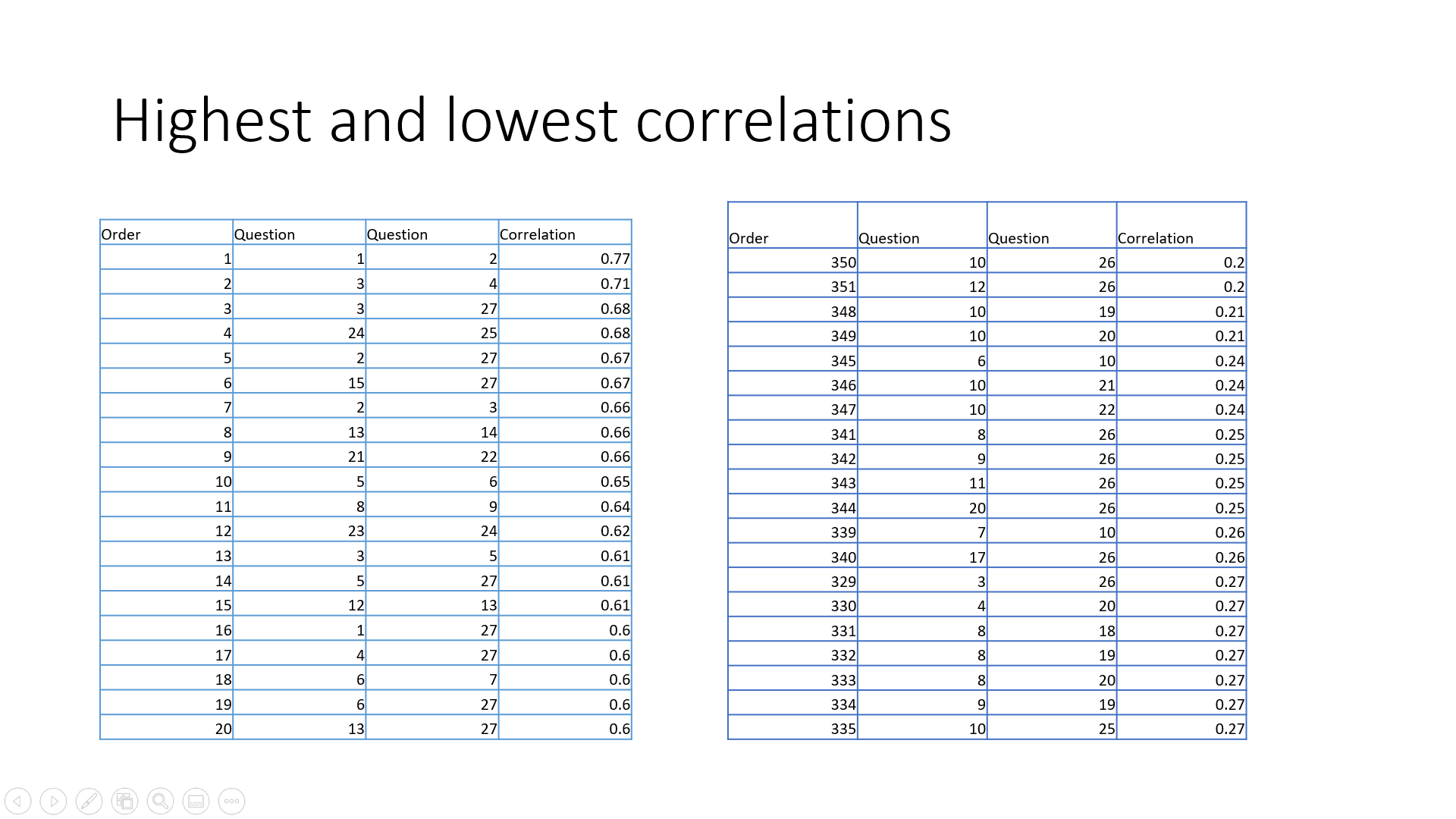
Full Polychoric Correlation Matrix



The above matrix also represented in Appendix B demonstrates the polychoric correlation between each of the questions and the other. The polychoric correlation is a measure of “bivariate association” when both the variates are ordered and categoric. The higher the polychoric correlation, it can be perceived that there is a higher dependence between the two items.

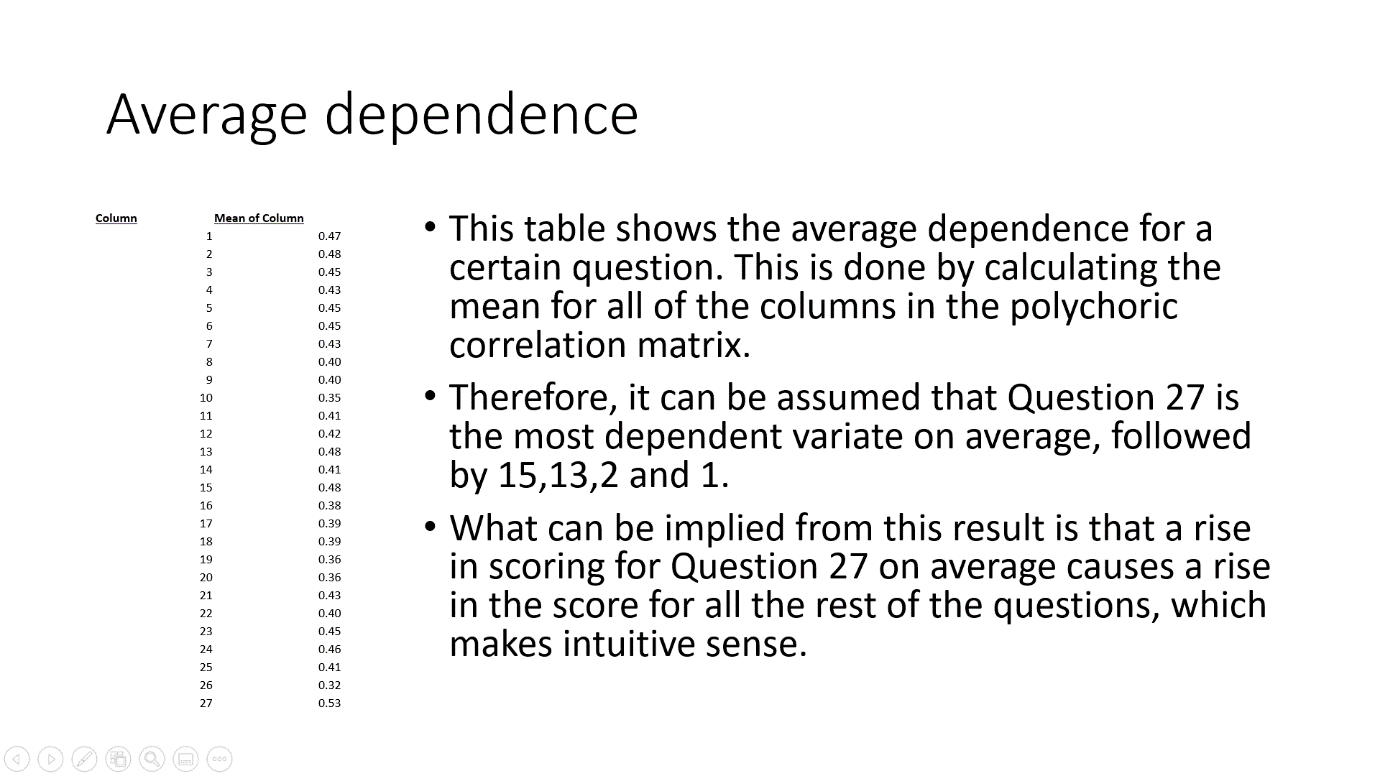
It appears that the within group dependence (i.e. for teaching, questions 1 to 4) have higher values for their polychoric correlation than pairs of questions that are not part of the same group. This makes intuitive sense as the factors that are likely to affect a question, are likely to affect the group of questions that are similar.

This is demonstrated by the tables below:



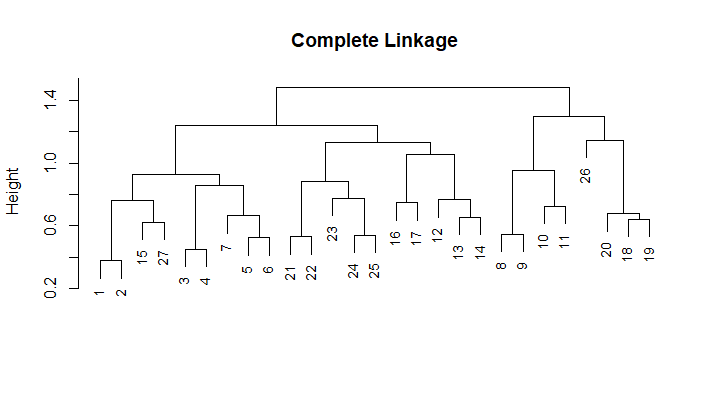
Question 1 and Question 2 appear to be the two items that depend on each other the most, with 3 & 4, 3 & 27, 24 & 25, and 2 & 27 being the next highest respectively. As stated above, it is evident that the items with the highest dependence correlations tend to be within the same group of questions. It can also be observed that Question 27 is the most common presence with these item pairs. With the teaching questions also prominent. This is also intuitive since it is likely that if a student is satisfied with a certain question, that they are also satisfied overall.

Question 10 and 26 appear to be the two least dependent items. Questions 10 and Question 26 individually are also the most commonly occurring in this list of least dependent questions. This implies that these questions are generally speaking the ones that associate the least with the rest of the questions.

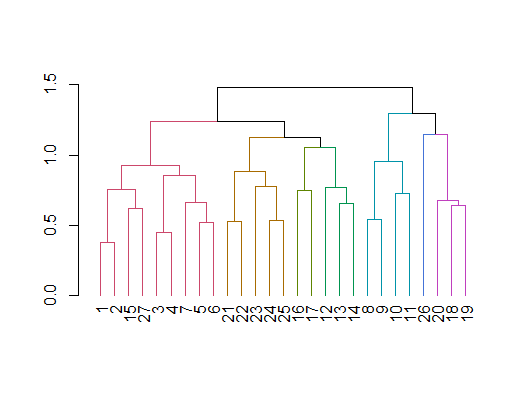
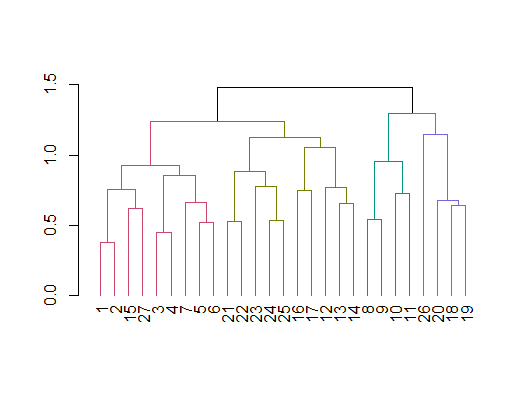
Another important aspect of dependence for each of the questions, is the average dependence which can be derived from the column means of the full polychoric correlation matrix. This shown below:

What can be implied from this result is that a rise in scoring for Question 27 on average causes a rise in the score for all the rest of the questions, which makes intuitive sense. Therefore, it can be assumed that Question 27 is the most dependent variate on average, followed by 15,13,2 and 1. Therefore it can be reasonably assumed that by increasing the scores for questions 27, 15, 13, 2 and 1 respectively, that respective increases in the rest of the questions should also occur. In this case as well the converse is also true. An increase in the scoring of the rest of the questions should also lead to an increase in the scoring for these questions too.

Hierarchical Clustering

****

Dendrogram 1: Initial Complete Linkage Dendrogram



Dendrogram 2: 7 clusters, h=1.0 Dendrogram 3: 4 clusters, h=1.2

In this section, we will further analyse the dependencies between questions using Hierarchical Clustering. Hierarchical Clustering is based on an agglomerative machine learning algorithm, where each observation (in this case, dependence correlation) begins as its own defined cluster, and is successively merged with the other clusters until there are none left to merge. The goal is to merge items which are similar as possible within the clusters, and as different as possible outside of the clusters. This provides value to us in this case, since we can draw conclusions on which questions are most likely to improve the scores of other questions based on this algorithm. Indeed, it can be acknowledged that “clusters” are already “built-in” in the sense that questions 1 to 27 are split into 8 distinct subgroups already such as teaching, assessment & feedback, etc. Therefore, we should pay particular attention to which questions that leave their pre-determined subgroup of questions, and indeed which groups of questions are similar to each other. This would allow for departments to focus on which pre-determined group of questions are most likely linked to another group, and therefore allowing a different avenue of approach in order to improve scores.

There are several different aspects on which the described algorithm’s merging is based. First off, we have to describe a dissimilarity measure. In this case, we opt for a Euclidean distance dissimilarity measure, linking clusters of questions together based on their distance from each other. To interpret this for our specific case, the clusters are linked to each other based on how close their dependencies on each other are. Therefore, similarly dependent questions will be clustered together showing which are most likely to affect the others. Another aspect on which we have to decide is the linkage. This plays a large part in deciding which clusters merge together as it decides which clusters are to be merged with the other. We opt for a Complete linkage. This is because it results in the most balanced dendrogram out of the options given. Complete linkage calculates all of the pairwise distances from each of the clusters and uses the largest of them.

Using the dendextend package and hclust() function in R, the three initial plots are produced. The initial plot gives us the most information, as it displays the various heights of each of the branches displayed by question numbers. The heights represented on the y-axis of each of the dendrograms shows us the distance between the clusters as is shown. An example is that, questions 1 and 2 are drastically different to question 16, but not too distant from question 3 and 4. As we move from the bottom of the tree upwards, we begin to fuse our branches, which corresponds to observations that are similar to each other, however, the higher up the tree we move the less similar the questions become. Verticality plays a large role in how similar two items are before fusing. Therefore, it is understandable how questions 1 and 2 are very similar to questions 3 and 4, but drastically dissimilar to question 26. This cannot be concluded from the final two plots.

What the final two plots do give us is a simple interpretation of results. By cutting the tree at different heights, we can create itemized lists of clusters full of questions that are most similar to each other. This is colour coded into the above plot. The results are summarised below:

|  |  |
| --- | --- |
| Cluster | Questions |
| 1 | 1,2,15,27,3,4,7,5,6 |
| 2 | 21,22,23,24,25 |
| 3 | 16,17 |
| 4 | 12,13,14 |
| 5 | 8,9,10,11 |
| 6 | 26 |
| 7 | 20,18,19 |

Dendrogram 2 Summary

By cutting Dendrogram 1 at height 1 labelled on the y-axis, we generate the above clusters. These can be interpreted as questions that are most similar to each other. Since we have some pre-determined groupings specified, we pay particular attention to questions that leave their group, our groups themselves that have joined together. It should also be noted that this level of clustering was chosen as it imitates the numbering of pre-determined groups as well. It is clear that from cluster 1, that questions 15 and 27 have a lot of co-dependence around the teaching questions. It is also interesting to note that the teaching and learning opportunities questions are also grouped together in cluster 1. This can also be observed in cluster 2, which represents a grouping up of the learning community and student voice questions. The rest of the questions seem to remain within their pre-determined groups, which is an intuitive result since the questions themselves tend to be the most similar within their group.

Another question which arises are which groups are most dependent on each other. Therefore, we can decide to make a higher cut in the dendrogram, and observe which groups come together as the most similar. This can also provide some useful information to learning management, as a more general focus if a certain group of scores need to be improved can be more practical than an item-level suggestion.

This new cut is shown below:

|  |  |
| --- | --- |
| Cluster | Questions |
| 1 | 1,2,15,27,3,4,7,5,6 |
| 2 | 21,22,23,24,25,16,17,12,13,14 |
| 3 | 8,9,10,11 |
| 4 | 26,20,18,19 |

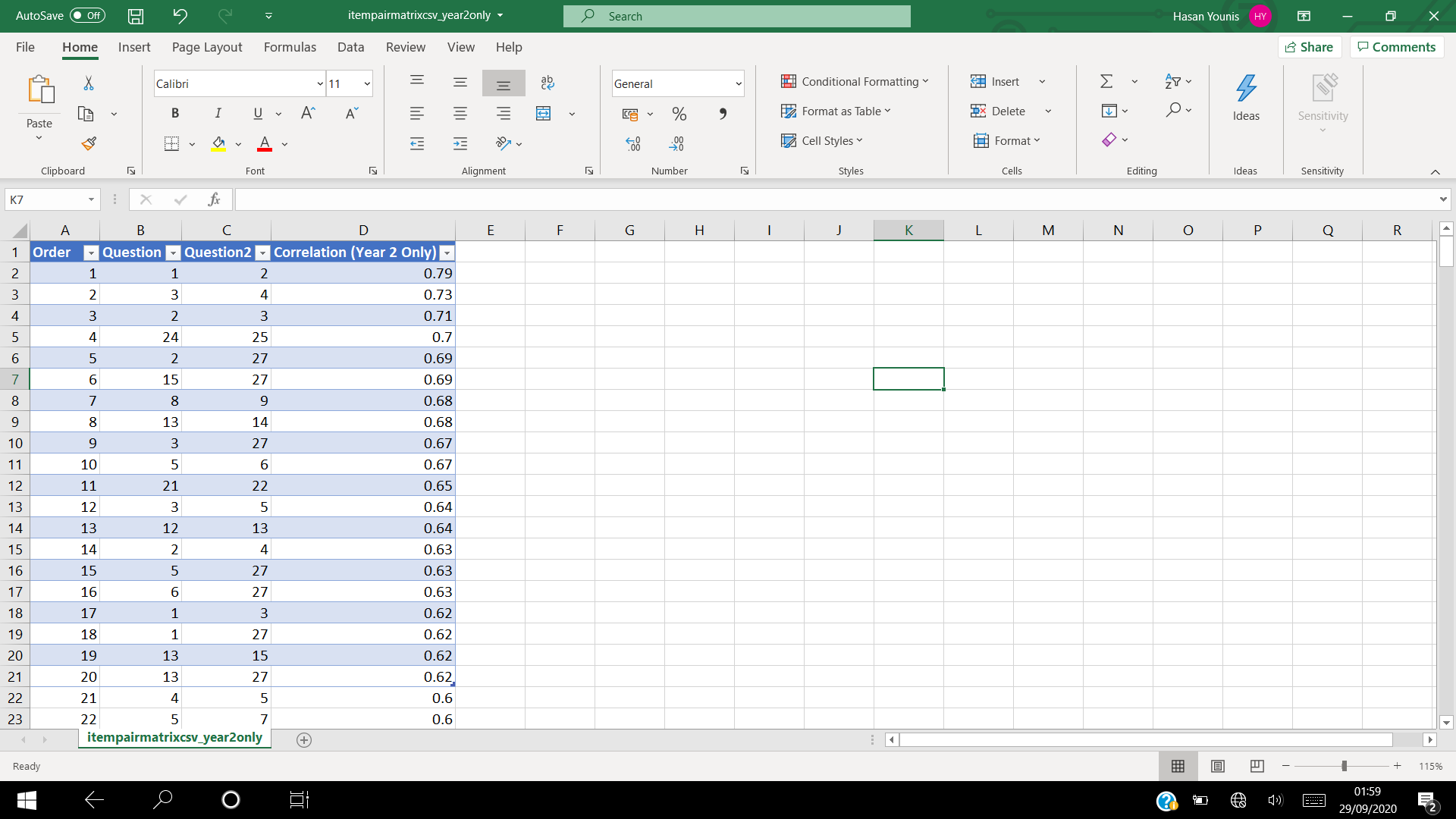
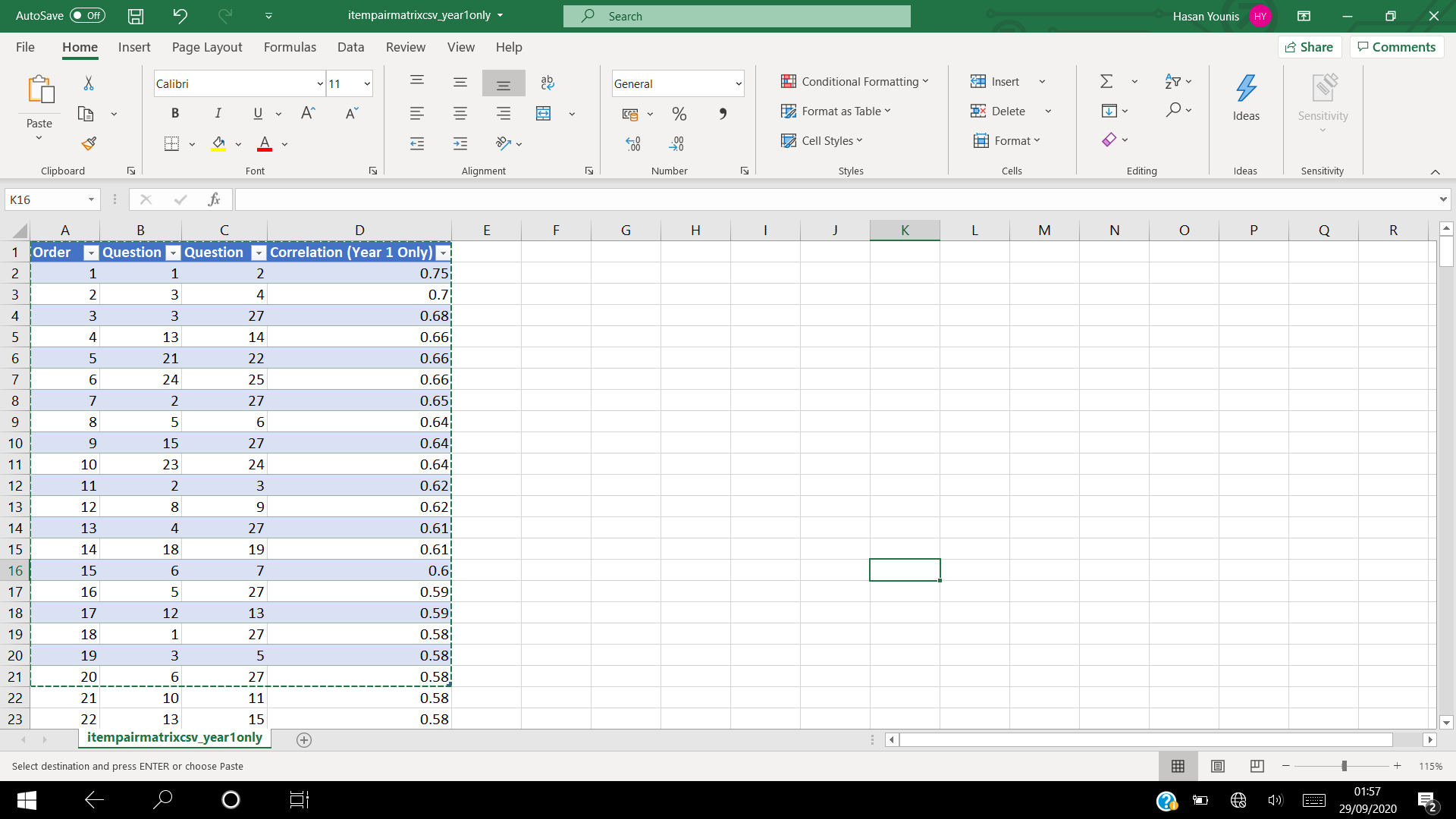
Dendrogram 3 Summary

From the above, it is clear to show that the groups of questions for teaching and learning opportunities are linked, as well as the odd question 15 and 27 which we determined from the last dendrogram, were more dependent on these features. It should also be noted that the groups of questions for student voice, learning community, organisation & feedback, and academic support are tightly linked as well. It can also be found that the questions for assessment & feedback are generally more dependent on each other, with the learning resources questions being most tightly linked with question 26, which, as we have seen, tends to correlate the least with the other questions anyway.

3rd Year Generalisation

Another important aspect of the dependence analysis is whether this information is consistent as a student passes through both the first year, and second year of University. We therefore want to find out whether the correlations that exist for year one exists to a similar extent in year 2, and from this we can impute the result for year 3 students as well.

This was achieved by partitioning the dataset into a Year 1 form, and Year 2 form, and running a dependence analysis individually.



The table on the left displays the top 20 correlation values for the Year 1 dataset only, whereas the table on the right displays the top 20 correlation values for the students in Year 2. An important thing to note is that once compared, 16 out of 20 of the item pairs exist in both lists. This implies that the same factors that cause dependence in the year 1 dataset also exist in the year 2 dataset, and therefore by extension, the year 3 dataset. It is therefore reasonable to suggest that the high correlation values we see from the teaching questions and the overall satisfaction for example extend to a year 3 student population. The individual matrices for the Year 1 and Year 2 polychoric correlations can be found in Appendix C.

A more quantitative method of determining whether this data is truly generalisable from Year 1 to Year 2 and then by extension to Year 3 would be to see if the differences between the polychoric correlations for the Year 1 and Year 2 dependence matrices are significant. If a pair of questions have significantly different polychoric correlation, then we can determine (at least for that question) that the change from Year 1 to Year 2 seems to have impacted the results and therefore, a generalisation is no longer appropriate if too many of these occurrences happen.

In order to generate significances, we have to first have a test objective. What we want to test is whether the difference between the correlations for the year 1 and year 2 polychoric correlations is therefore equivalent to 0. Because each of the polychoric correlations comes with a variance built into the function, we input those variances into a matrix. The variance for the test function is then simply the addition of the variances for the year 1 correlation matrix and the year 2 correlation matrix. We square root this in order to get the, more relevant, standard error. Using this standard error, we construct a 95% confidence interval for each of the correlations. If the interval contains 0, then we judge the correlation to be non-significant overall at the 5% level.

Of the 351 observable differences from the year 1 and year 2 polychoric correlations, 19 were adjudged as significant. This was separated between 8 lower bounds above 0, and 11 upper bounds below 0. This summarised in the tables below:

|  |  |
| --- | --- |
| Question Combination | Confidence Interval |
| 8,10 | (0.00982,0.15218) |
| 10,18 | (0.01269, 0.20770) |
| 13,18 | (0.00557,0.20228) |
| 7,20 | (0.02117,0.19961) |
| 10,23 | (0.06956,0.24725) |
| 2,26 | (0.00342,0.12400) |
| 15,26 | (0.00351,0.16066) |
| 20,26 | (0.03566,0.17098) |

Table 19: Lower bounds above zero.

|  |  |
| --- | --- |
| Question Combination | Confidence Interval |
| 1,3 | (-0.0068, -0.1417) |
| 2,3 | (-0.0322, -0.1391) |
| 1,5 | (-0.0116, -0.1592) |
| 1,6 | (-0.0555, -0.2162) |
| 3,6 | (-0.0176, -0.1678) |
| 4,6 | (-0.0058, -0.144) |
| 4,11 | (-0.0495, -0.2399) |
| 1,12 | (-0.053, -0.2083) |
| 3,14 | (-0.0032, -0.169) |
| 9,16 | (-0.006, -0.1931) |
| 1,19 | (-0.0051, -0.1978) |

Table 20: Upper bounds below zero.

The implications that can be drawn from the above tables is that in Table 19, we tend to find polychoric values of the questions that tend to be on the lower end. We have previously found that questions 8, 10, and 26 which come up prominently in this table are also prominent in terms of the questions that tend to have the least polychoric correlation. It is therefore intuitive as to why these question combinations are significant. The same can be applied for the values in Table 20. We consistently see question numbers here that are towards the top of the polychoric correlation item pair table, such as questions 1, 2 and 3. So it is similarly intuitive as in Table 19, but for the other end. The logic that can be implied is that the variance towards the polychoric correlation values towards the higher and lower ends of the spectrum tend to be more variable overall which means that they become significant in these tests.

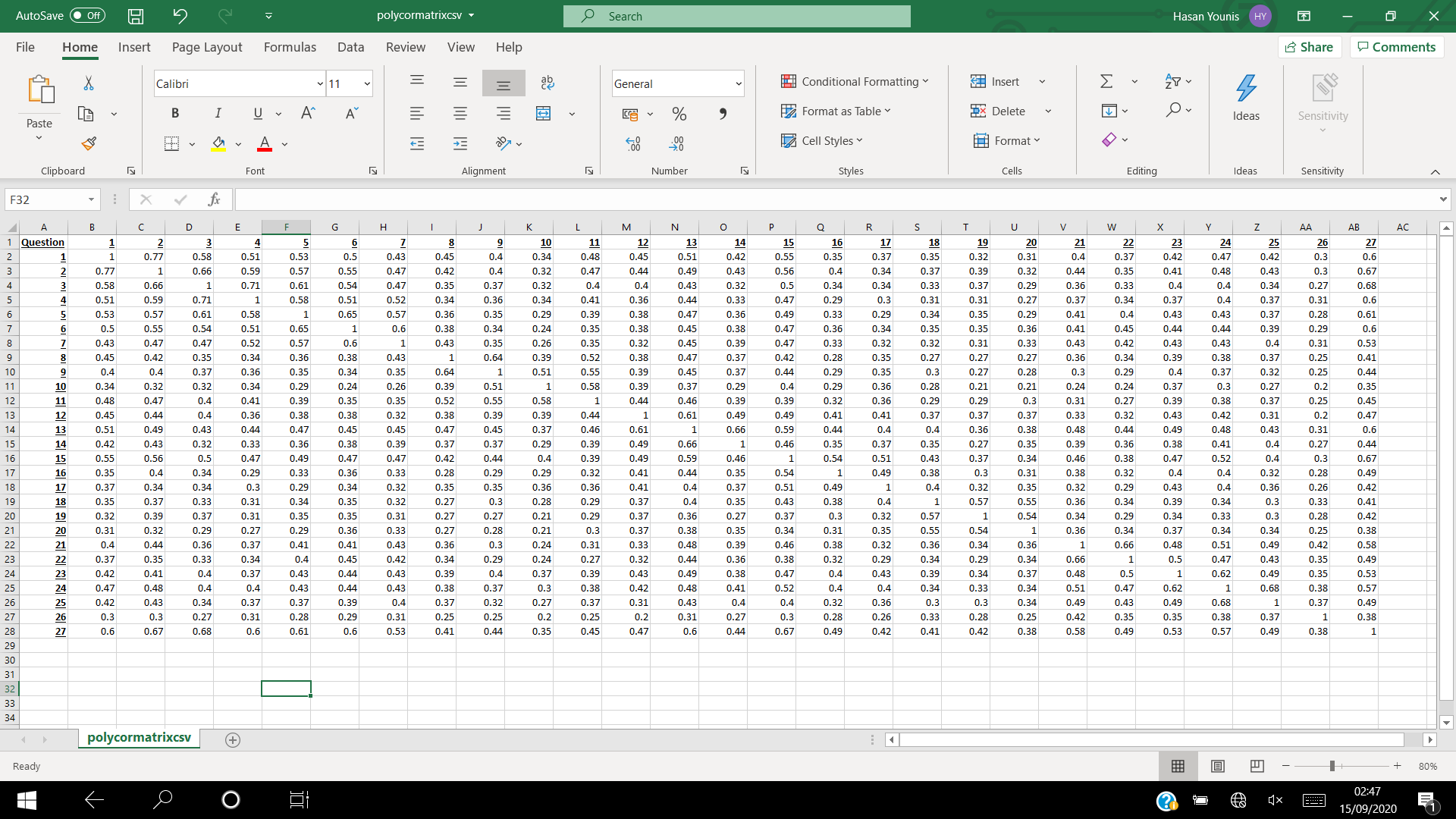
Therefore, overall, I conclude that the Year 1 and Year 2 findings should be generalisable to the Year 3 students.

Conclusion

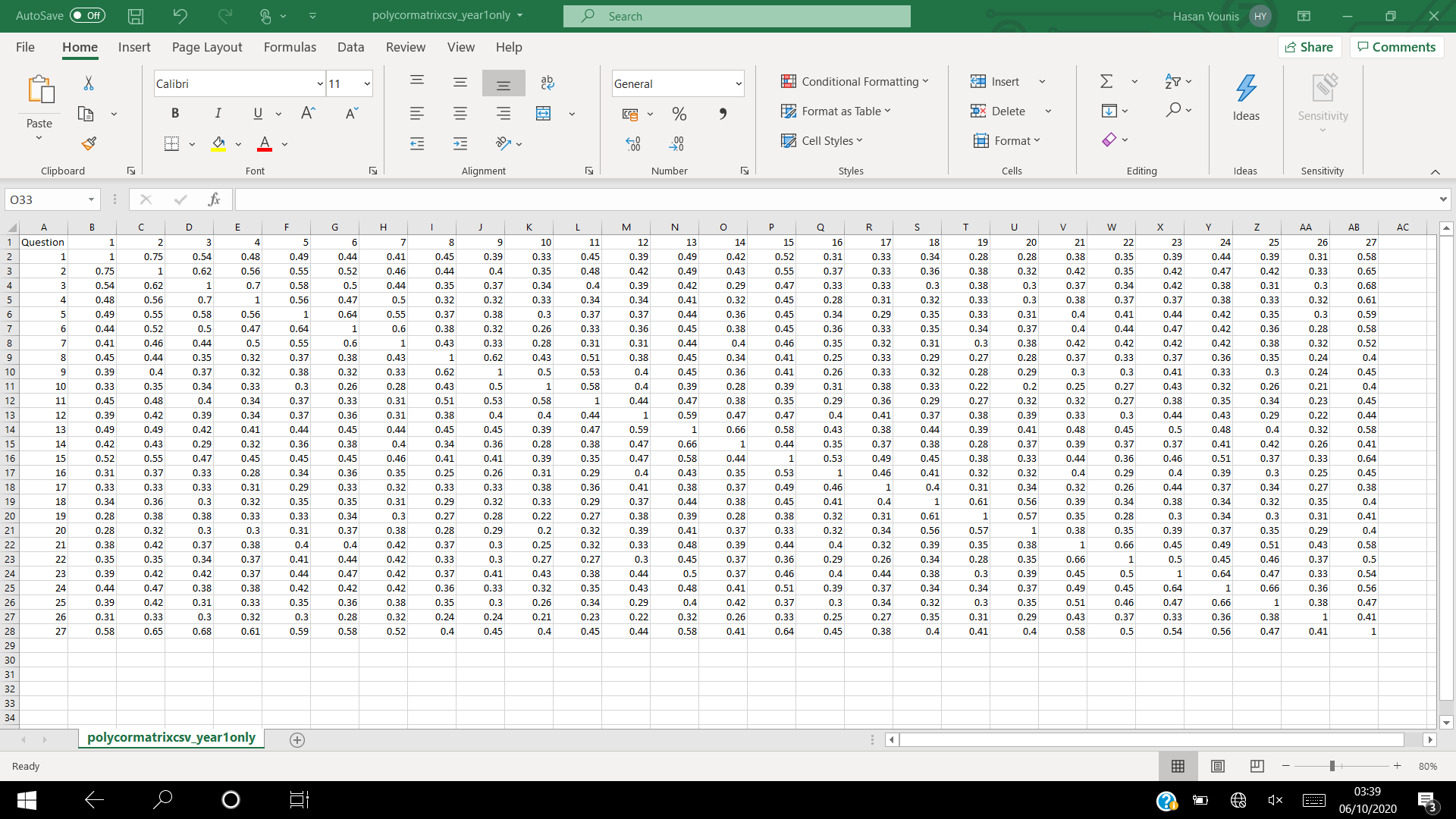
In Conclusion, I believe over the course of this report we have effectively understood and explained the variables that affect the changes in scoring across the students at the LSE. Not only this, but I believe we can also provide useful feedback to departments across the University with areas they can focus on in order to improve their student satisfaction, and indeed improve the student experience overall that is within their control.

Appendix A

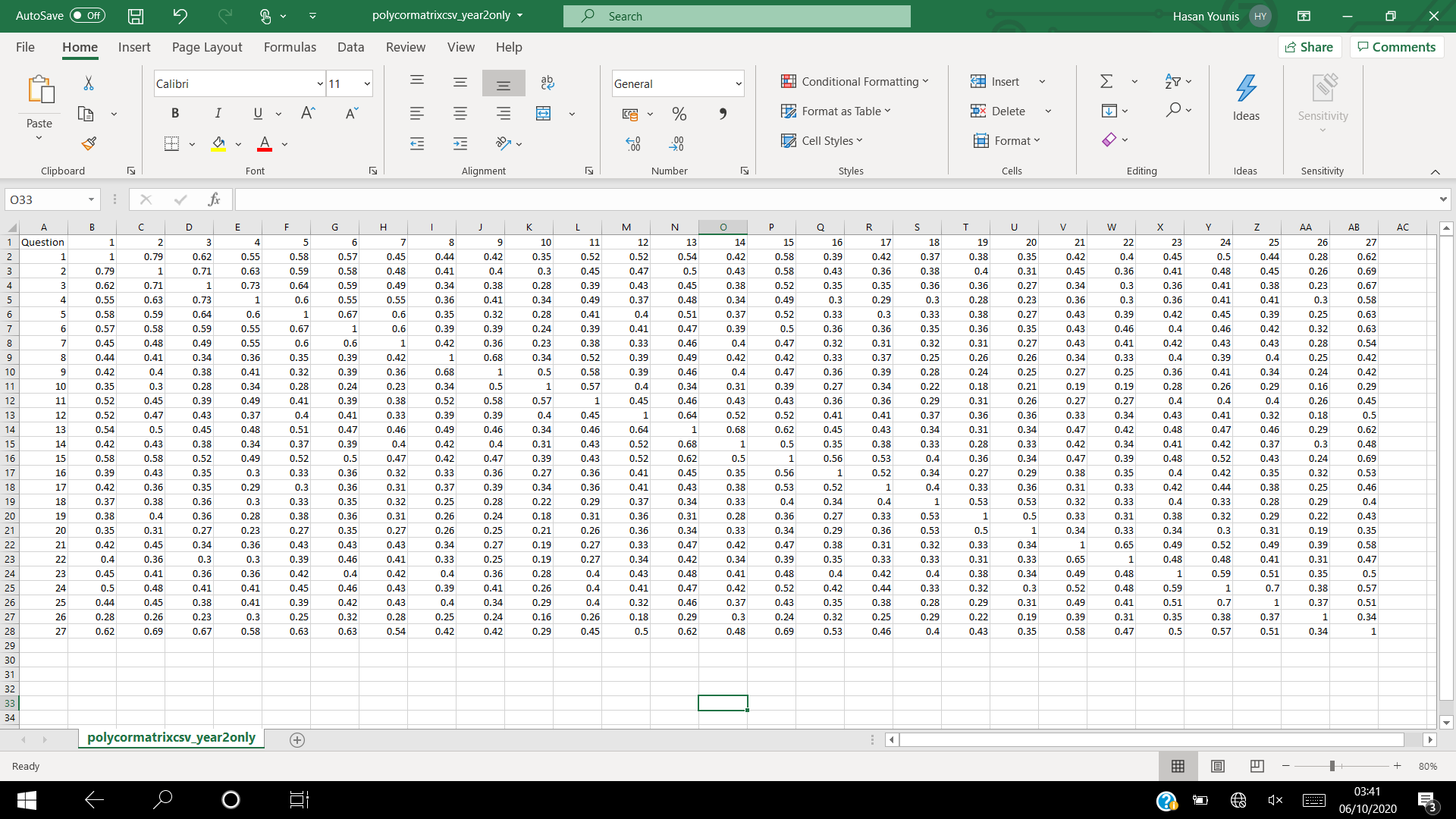
* Question 1 - 1. Staff are good at explaining things.
* Question 2 - 2. Staff have made the subject interesting.
* Question 3 - 3. The programme is intellectually stimulating.
* Question 4 - 4. My programme has challenged me to achieve my best work.
* Question 5 - 5. My programme has provided me with opportunities to explore ideas or concepts in depth.
* Question 6 - 6. My programme has provided me with opportunities to bring information and ideas together from different topics.
* Question 7 - 7. My programme has provided me with opportunities to apply what I have learnt.
* Question 8 - 8. The criteria used in marking have been clear in advance.
* Question 9 - 9. Marking and assessment has been fair.
* Question 10 - 10. Feedback on my work has been timely.
* Question 11 - 11. I have received helpful comments on my work.
* Question 12 - 12. I have been able to contact staff when I needed to.
* Question 13 - 13. I have received sufficient advice and guidance in relation to my programme.
* Question 14 - 14. Good advice was available when I needed to make study choices on my programme.
* Question 15 - 15. The programme is well organised and is running smoothly.
* Question 16 - 16. The timetable works efficiently for me.
* Question 17 - 17. Any changes in the programme or teaching have been communicated effectively.
* Question 18 - 18. The IT resources and facilities provided have supported my learning well.
* Question 19 - 19. The library resources (e.g. books, online services and learning spaces) have supported my learning well.
* Question 20 - 20. I have been able to access programme-specific resources (e.g. equipment, facilities, software, collections) when I needed to.
* Question 21 - 21. I feel part of a community of staff and students.
* Question 22 - 22. I have had the right opportunities to work with other students as part of my programme.
* Question 23 - 23. I have had the right opportunities to provide feedback on my programme.
* Question 24 - 24. Staff value students' views and opinions about the programme.
* Question 25 - 25. It is clear how students' feedback on the programme has been acted on.
* Question 26 - 26. The students' union (association or guild) effectively represents students' academic interests.
* **Question 27 - 27. Overall, I am satisfied with the quality of the programme**.
* Teaching
* Learning Opportunities
* Assessment and Feedback
* Academic Support
* Organisation and Feedback
* Learning Resources
* Learning Community
* Student Voice

Appendix B

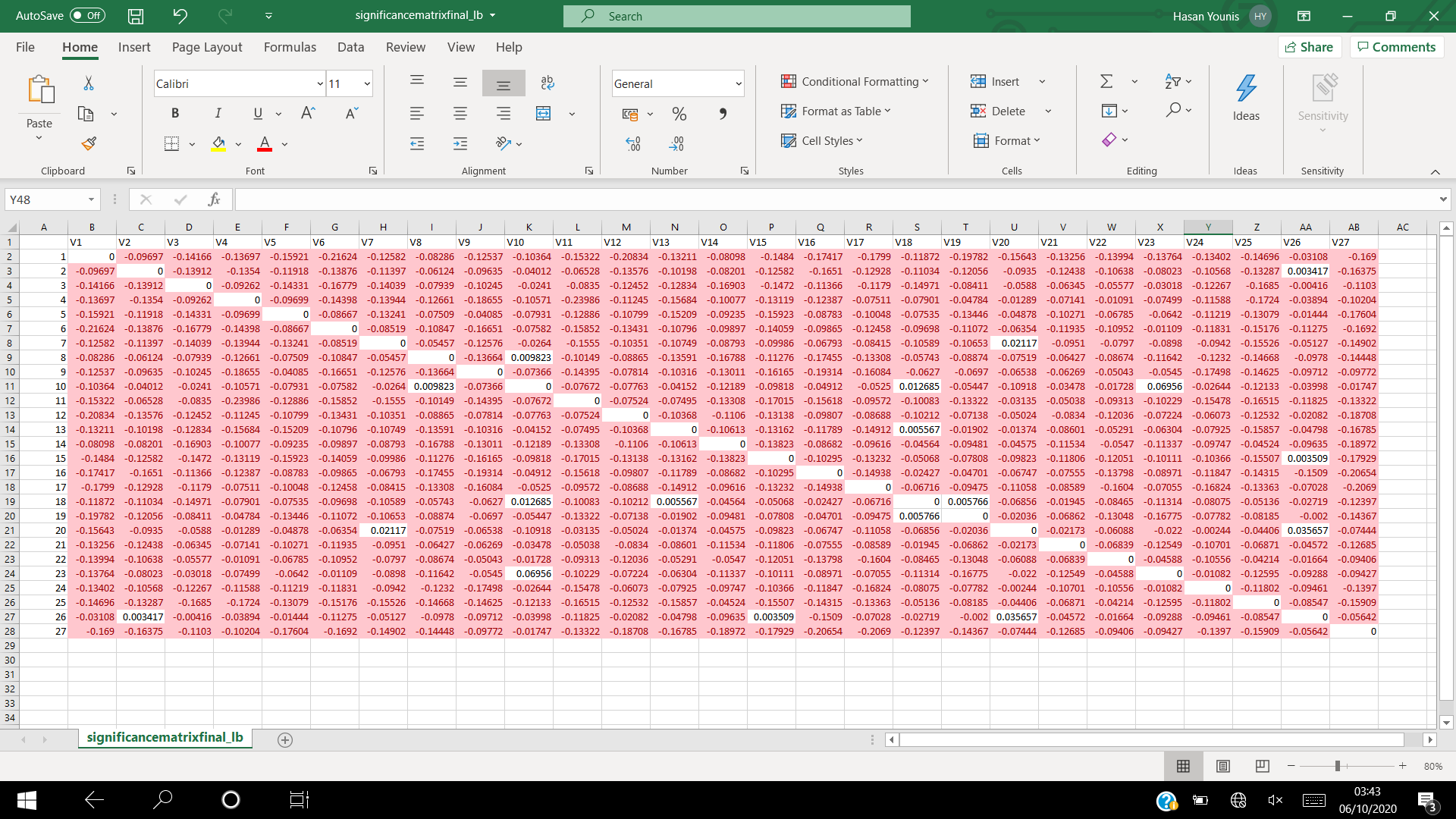
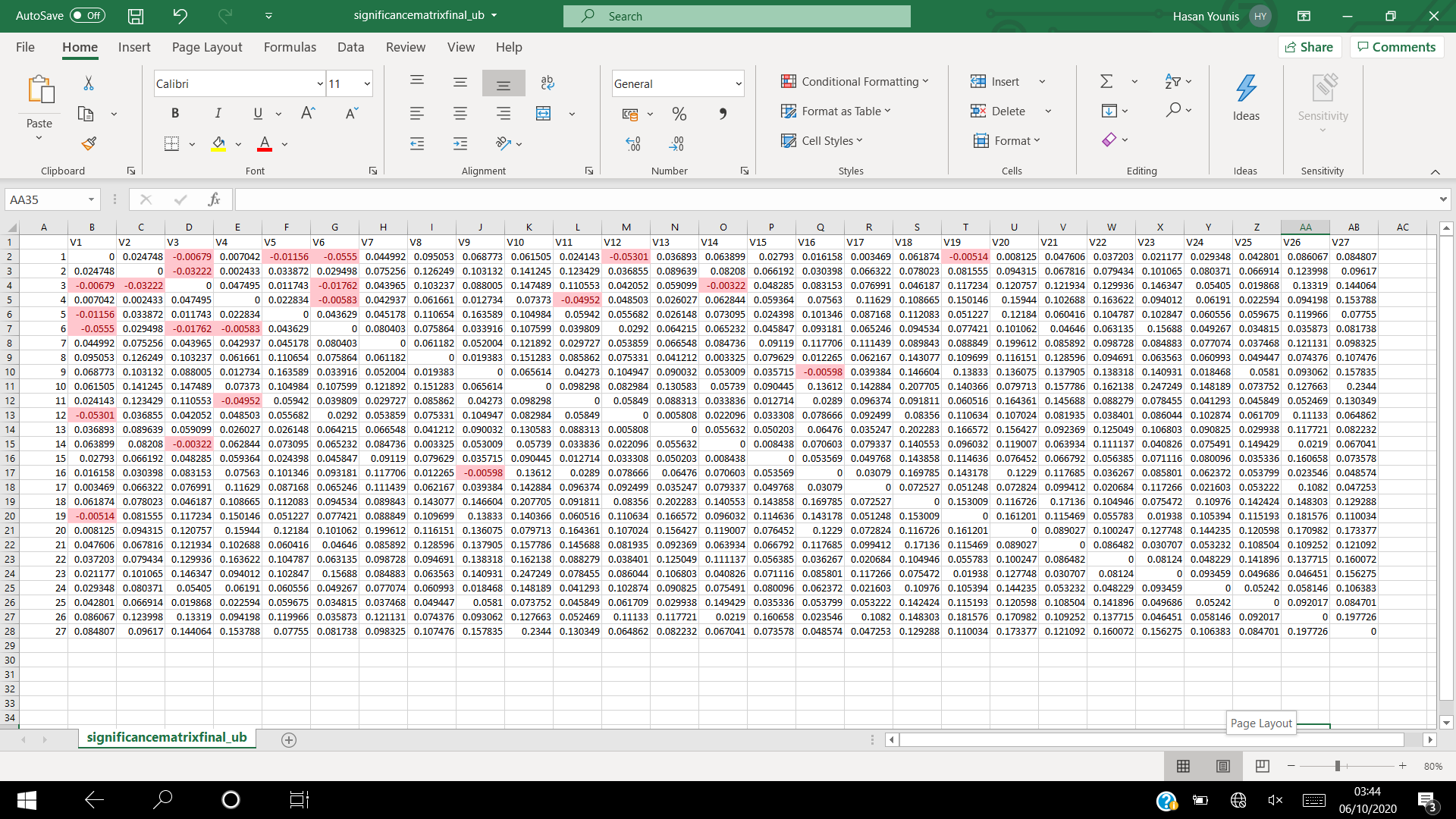
Full Polychoric Correlation Matrix

Appendix C

Year 1 only Polychoric Correlation Matrix



Year 2 only Polychoric Correlation Matrix

 Lower Bound Significance Matrix

Upper Bound Significance Matrix

1. Agresti (2007), An Introduction to Categorical Data Analysis, Pg180-189 [↑](#footnote-ref-1)
2. Agresti (2007), An Introduction to Categorical Data Analysis, Pg 297-317 [↑](#footnote-ref-2)
3. Drasgow (1986), Polychoric and Polyserial Correlations, Pg 68-73 [↑](#footnote-ref-3)