An Evaluation of Quality and Probity in the Irish Dental Treatment Services Scheme

Hanna Seyoum

02 August, 2017

Declaration

I, Hanna Seyoum, declare that this thesis titled, An Evaluation of Quality and Probity in the Irish Dental Treatment Services Scheme, and the work presented in it are my own. I confirm that the work being submitted is the result of my own analysis except where otherwise stated.

Abstract

Background: A great deal of treatments provided by dentists in the DTSS are amalgam restorations. 90% of Irish adults experience tooth decay, which is commonly corrected through either amalgam or composite fillings. With 21% of total expenditure being from amalgam fillings, there has been a need for a formal probity assurance mechanism. Two mechanisms are used in this paper to analyse quality and probity of dentists in the DTSS.

Methods: Probity and quality in the DTSS were assessed through linear regression analysis. For probity, a normal linear regression was used by first calculating the fraction of total earnings from amalgam restorations (FTEA), by then fitting a normal linear model with FTEA as the response variable with 18 explanatory variables. Quality was analysed by fitting a general linear model with the re-treatment rates (RR) of fillings provided as the response variable with 18 explanatory variables. After validating the data and fitting the models, diagnostic tests were run on the models. The detection of outliers was used as a means of identifying dentists with excessively high or low FTEA and RR.

Results: The identified dentists had either high or low FTEA and RR, with some dentists near the median. When comparing their explanatory variables, the results were less conclusive. Most of their explanatory variables, such as dentist age, patient age, and proportion of amalgam fillings, did not deviate far from the median.

Conclusion: The aim of this study was to identify dentists with excessively high or low FTEA and RR and determine whether there is sufficient reason to investigate them further. A comparative analysis of the identified dentists' explanatory variables showed that, besides their high proportions, there was nothing particularly odd about them. I recommend further study/investigation into the dentists identified to figure out why they have such high proportions.

Contents

List of Tables

List of Figures

1 Introduction

Dental carries is a significant oral health problem in most industrialised countries. Over 90% of Irish adults experience dental caries, excluding the 5% of adults that have no teeth, (compared to 92% of US adults) [1]. Dental caries is a cavity (tooth decay), which is caused by bacteria producing acid that destroys the tooth's enamel surface and its underlying layer, the dentin. The acid is produced when sugars in foods or drinks react with bacteria present on the tooth surface [2]. The most common treatment for tooth decay is a restoration procedure which is done by drilling out the decay and replacing it with a filling made from metal (amalgam), or plastic (composite). Extensive tooth decay may require treatments such as an extraction, root canal (endodontic), or inserting a crown [2].

Dentistry in Ireland is regulated by the Dental Council of Ireland, a statutory body created under the Dentists Act 1985. There are four registers relating to dentistry in the council; one of which is the Register of Dentists. Only dentists listed on the Register of Dentists can legally practice dentistry in Ireland [3]. The dental care system in Ireland is made up of public and private service provision, with most dentists working within the private sector, mainly as general dental practitioners [4].

In Ireland, the Dental Treatment Services Scheme (DTSS), established in November 1994, is a State-run scheme that is managed by the Health Service Executive (HSE). It provides access to dental treatment to medical card holders, aged over 16 years. The scheme entitles medical card holders to a free dental examination, up to two fillings in a calendar year, and extractions if necessary [5]. General dental practitioners in the private sector are the main providers of public dental services in the DTSS, operating on a fee-per-item remuneration [3]. It is important to pay close attention to the provision and remuneration of fillings. Dentists in the scheme provide a great deal of fillings to patients every year. The total expenditure on restorations, for both amalgam and composite, was 15 million in 2003, representing 32% of the total expenditure. Of the 15 million, 65% of the total expenditure on restorations

was from amalgam restorations [6]. This paper aims to evaluate financial probity and performance quality of dentists in the DTSS, and identify dentists with high probity and quality risk in relation to restorations provided.

The provision of fillings is a delicate process that can be difficult to perfect, resulting in variability in their quality. In addition to providing restorations that are up to par, dentists must decide on whether to re-treat fillings. In making this decision, though unusual, some dentists might be more economically motivated; re-treating fillings unnecessarily for additional financial gains. Such decisions are deemed unethical, and are subject to lacking probity.

In 2006, the HSE conducted probity and quality checks in the DTSS. As part of a probity assurance initiative, they established an Examining Dentist scheme where they appointed 20 examining dentists in the DTSS for a period of one year to meet with patients that have been treated by dentists in the DTSS to ensure the quality of treatments provided were up to par, as well as ensuring the procedures provided by the dentists were necessary. This study was conducted on a random sampling basis. Their finding was that most dentists were not a probity risk, but focusing more on dentists in high probity areas rather than randomly selecting them could be more efficient [7].

There is a need for probity and quality assurance mechanisms in the DTSS as there are currently no formal assurance mechanisms in place. Probity can be studied through the mechanism of analysing fractions of total earnings from the provision of amalgam restorations for each dentist in the DTSS. Quality can be studied through the mechanism of analysing the re-treatment rates of all restorations provided the dentists in the DTSS.

The database in this study consists of summary details of treatments provided by a random sample of dentists in the DTSS. It contains some characteristics of dentists and patients in the scheme; such as dentist age, patient age, and health board and practice type in which the dentists work. The two fractions used for the analyses,

fractions of total earnings from amalgam fillings provided as well as the re-treatment rates were calculated in this database.

Probity in the DTSS is assessed by identifying dentists whose fractions of total earnings from the provision of amalgam fillings are excessively high or excessively low. This fraction is calculated by dividing the total earnings dentists have earned from amalgam fillings by total earnings dentists have earned from all treatments provided. If a fraction is excessively high or low, it could indicate that a dentist is overcharging patients, or heavily leaning towards providing amalgam fillings for financial reasons. It could also be a result of other variables affecting it. This paper identifies dentists with high or low proportions and investigates other variables that may have contributed to the excessiveness of the fraction.

On the other hand, quality of restorations placed by dentists in the DTSS is assessed by identifying dentists with excessively high or excessively low re-treatment rates of fillings. The re-treatment rate is calculated by dividing the number of failed fillings by the total number of fillings the dentists have provided. An excessively high or low re-treatment rate could be indicative of negligence on the part of the dentist. It could also result from patients not practicing good dental hygiene, for example. This paper identifies dentists with very high or low re-treatment rates of fillings, and looks at explanatory variables for analysis.

2 Method

2.1 Research Overview

The purpose of this study is to analyse the probity and quality of dentists in Ireland within the DTSS. The goal is to identify dentists who might be providing services with probity and quality that are not up to par by analysing the fractions of total earnings from amalgam restorations provided (FTEA), and re-treatment rates (RR) of fillings provided to patients by the dentists. The analyses were conducted using the statistical software R, and the research design used in this paper is of quantitative nature.

2.2 Setting and Participants

The participants of the study are dentists in Ireland who are registered in the DTSS, observed during the period November 1994 - December 2003. These dentists are from all over the country, some with individual practices, and others with group practices.

2.3 Description of Database

The data comes from the DTSS payments database that has been transferred routinely to the Oral Health Services Research Centre (OHSRC), in University College, Cork where queries have been developed in the OHSRC to provide individual level data on the treatments provided to patients. The following database contains summary details of the fillings provided by a random sample of dentists operating in the DTSS to patients aged 16-34. The original data has been perturbed to produce this database, to preserve the privacy of the original data. It was done so by assigning unique but artificial identification numbers to the dentists.

Dentist Details		
Variable position number in the txt. file		
22	DentistID	Unique but artificial dentist identifier
1	DENT_HB_NUM	Health Board in which the dentist works: 1 = Eastern 2 = Midlands 3 = Mid-Western 4 = North Eastern 5 = North Western 6 = South Eastern 7 = Southern 8 = Western
2	DentAge	Dentist's age at the time first filling under study was provided
3	TotalPatients	Total numbers of patients (all ages) seen since entering DTSS. Non-DTSS patients are not included in this total.
4	PracticeType	Type of practice in which the dentist works: 0 = Single (one dentist) 1 = Group (more than one dentist)
5	Fillings	The number of fillings provided by the dentist between December 1995 and December 2003.
6	Failures	The number of fillings provided by the dentist that failed up to 31st of December 2003.
Summary details of other treatments (some of which have been carried out by other dentists) received by patients during the observation period of their fillings for each dentist		
7	Dentists	Patients do not have to visit the same dentist for all their treatments. Following placement of a filling, a patient is free to visit any other dentist for any other treatments. This variable is the total number of other dentists visited by patients during the observation period of their fillings placed by the dentist.
8	Oral Exam	The total number of oral examinations (some of which have been carried out by other dentists) received by patients during the observation period of their filling placed by this dentist.
9	Prophylaxes	The total number of prophylaxes (scale and polish)
10	Extraction	The total number of extractions
11	ProtractedPerio	The total number of protracted periodontal (gum) treatments
12	Endodontic	The total number of endodontic (root canal) treatments
13	OtherTreatment	The total number of other treatments
Summary details of the fillings provided by each dentist		
14	Amalgam	The number of amalgam (metal) fillings provided by the dentist.
15	Composite	The number of composite (plastic) fillings provided by the dentist.
Summary details of the patients treated with fillings for each dentist		
16	Patients	The number of patients treated with fillings by the dentist.
17	Female	The number of female patients treated with fillings by the dentist
18	MeanAge	The mean age of patients treated with fillings by the dentist
19	MeanBaselineDMFT	The mean number of Decayed, Missing or Filled Teeth (DMFT) at entry to the DTSS of patients treated with fillings by the dentist.
Summary Details of earnings for each dentist		
20	AmalgamEarnings	Total earnings from amalgam fillings provided to all patients for the dentist while in the DTSS.
21	TotalEarnings	Total earnings (Euro) from all treatments provided to all patients for the dentist while in the DTSS. Earnings from non-DTSS patients are not included.

Table 1: Description of Database

2.4 Data Validation

- Minimum dentist age was set to 22 with no upper limit.
- Cases with negative number of total patients were removed.
- Dentists with total earnings less than €5000 were excluded from Model 1.
- Dentists who provided less than 10 fillings were excluded from Model 2.

2.4.1 Cross-Validation

- Dentists that had greater number of failed fillings than total number of fillings provided were removed.
- Number of female patients were checked to ensure there were no cases greater than total number of patients.
- Amount of amalgam earnings were checked to ensure there were no cases greater than total earnings.
- Number of patients treated with fillings were checked to ensure there were no cases greater than total patients.

2.4.2 Missing Values

Three out of 22 of the explanatory variables had cases with missing values, resulting in 635 missing values combined. The variables were age, total patients and practice type. To check if the three variables were significant, a dataframe without the missing values was created and two models, full and reduced, were fitted from that dataframe. The reduced model was without dentist age, total patients, and practice type. After conducting the F-test and analysing the F and p values, the three variables were not removed.

In addition to the F-test two groups were created, with and without missing values, and were compared with regards to their explanatory variables and significant differences were observed, indicating that the removal of the variables would distort the analyses. Hence supporting the decision to keep the variables.

2.5 Model 1: Probity

Two models were used to analyse probity in the DTSS. The first was a full model with all the explanatory variables, and the second was a reduced model excluding dentist age, total patients, and practice type.

2.5.1 Description of Model

The full model, model1a, is a linear regression model used to analyse the probity of dentists in the DTSS. The model was arrived at after a running a few diagnostic checks on the initial full model, model1.1i, resulting in the application of square root transformations to some of the variables. The response variable in this model is the fraction of total earnings from the provision of amalgam restorations (FTEA) (transformed), calculated by dividing amalgam earnings by total earnings. It is regressed on the following explanatory variables: dentist health board (categorical: 8 dummy variables), dentist age (transformed), total number of patients (transformed), practice type (categorical: 2 dummy variables), fillings, failures, dentists (transformed), oral exam, prophylaxes, extraction, protracted periodontal, endodontic, other treatment, amalgam fillings, composite fillings, patients, female patients, mean age of patients, and mean baseline DMFT. The variables amalgam earnings and total earnings were excluded from the model because they were already contained in the response variable, FTEA. The variable, composite, was also excluded because composite and amalgam fillings are collinear variables.

The reduced model, model1b, is a linear regression model used to analyse the probity of dentists in the DTSS. The model was arrived at after running a few diag-

nostic checks on the initial reduced model, model1.2i, resulting in the application of square root transformations to some of the variables. The response variable FTEA is regressed on the following explanatory variables: dentist health board (categorical: 8 dummy variables), fillings (transformed), failures, dentists (transformed), oral exam (transformed), prophylaxes (transformed), extraction, protracted periodontal, endodontic, other treatment, amalgam fillings, composite fillings, patients, female patients, mean age of patients, and mean baseline DMFT. The variables composite, amalgam earnings, and total earnings were excluded from the model for the same reasons as model1a.

See figure 3 for a summary of the response and explanatory variables in model 1

2.5.2 Diagnostics

The following assumptions were made about the initial models prior to the transformations:

- The relationship between the response variables, FTEA, and explanatory variables is linear
- The errors e_i, and hence the responses have constant variance
- The errors e_i are normally distributed

The following regression diagnostic plots were used to test the above assumptions.

Test 1: Response variables vs. explanatory variables plot to assess the assumptions of linearity and constant variance

Test 2: Residuals vs. fitted values of the response variables plot to also assess the assumption of linearity and constant variance

Test 3: Histogram of the residuals and the normal probability of residuals plot to test the assumption of normality

Both initial models failed all three tests. Therefore, square root transformations were applied to linearize the models, stabilize their variances, and give normal residuals, resulting in the final full and reduced models, model1a and model1b respectively.
See figures 7-10 in the appendices chapter for diagnostic residual plots for model 1

Case Identification

The method used for identifying unusual cases was the detection and analysis of outliers. Cut-off points (-3,3) were used to identify outliers. Since the residuals are normally distributed, 5% of the cases were expected to lie outside of the (-2,2) range, resulting in 42 cases, which would have been too many.

Assumptions made

- The errors are normally and independently distributed; $e_i \sim$ NID(0, σ^2).

Studentized residuals were first calculated to identify outliers. Through a plot of studentized residuals vs. observation numbers, it was possible to visually locate cases outside of the (-3,3) range and manually count them. After locating the cases, their dentist ID numbers were identified.
See chapter 3.2 for the list of identified dentists.

2.6 Model 2: Quality

Two models were used to analyse quality in the DTSS. The first was a full model with all the explanatory variables, and the second was a reduced model excluding dentist age, total patients, and practice type.

2.6.1 Description of Model

The full model, model2a, is a general linear regression model used to analyse the quality of dentists in the DTSS with respect to the re-treatment rates of fillings they have provided. This model is made up of Bernoulli trials with independent cases, making it a binomial model. It is binomial and not merely binary since the number

of success per dentist is being counted as opposed to having just success or failure at an individual filling level.

Model2a was arrived at after running a few diagnostic checks on the initial full model, model2.1i, resulting in the application of square root transformations to some of the variables. The response variable in this model is the re-treatment rates of fillings (RR) which was calculated by dividing the fillings that failed by the total number of fillings the dentist provided. RR is a vector of successes and failures, the successes being the failed fillings, regressed on the following explanatory variables: dentist health board (categorical: 8 dummy variables), dentist age (transformed), total number of patients (transformed), practice type (categorical: 2 dummy variables), dentists (transformed), oral exam, prophylaxes, extraction, protracted periodontal, endodontic, other treatment, amalgam fillings, patients (transformed), female patients, mean age of patients, mean baseline DMFT, amalgam earnings (transformed), and total earnings (transformed). The variables fillings, failures, and composite fillings were excluded from the model because they were already contained in the response variable RR.

The reduced model, model2b, is also a general linear regression model used to analyse quality of dentists in the DTSS with respect to the re-treatment rates of fillings provided by them. This model is also binomial and was arrived at by applying square root transformations to some of the variables in the initial reduced model, model2.2i. The response variable RR was regressed on the following explanatory variables: dentist health board (categorical: 8 dummy variables), dentists (transformed), oral exam, prophylaxes, extraction, protracted periodontal, endodontic, other treatment, amalgam fillings, patients (transformed), female patients, mean age of patients, mean baseline DMFT, amalgam earnings (transformed), and total earnings (transformed). The variables fillings, failures, and composite fillings were excluded from the model for the same reason as model2a.

See figure 4 for a summary of the response and explanatory variables in model 2

2.6.2 Diagnostics

The following assumption was made about model 2:

- Each filling is an independent Bernoulli trial

The independence assumption is likely to be valid on the entire model with few violations expected. Such violations may be due to fillings that are placed within the same patient, hence counting the same patient more than once.

Checking the systemic component

Index plots of Pearson and deviance residuals and a plot of residuals vs. linear predictor were plotted for both initial full and reduced models using the standardized forms of the residuals. The standardized forms were calculated using the leverage values. Both index and residual plots had a pattern, indicating an issue with the systematic component. Therefore, square root transformations were applied on some of the variables ending up with model2a and model2b.

See figures 11-14 in the appendices chapter for diagnostic plots for model 2 **Case Identification**

Outliers were used to identify unusual dentists. The index plot of Pearson and deviance residuals with standardized residuals were used to detect outliers. Cutoff points (-5,5) were used for the outliers instead of (-2,2) because the former resulted in too many outliers. Since the residuals are only approximately normally distributed, outliers outside of the (-2,2) range can be expected to be more that 5% of the cases. Even with a cutoff of (-2,2) in a perfect normal distribution would result in too many outliers since 5% of 702 cases is 35. Therefore, 28 dentists outside of the (-5,5) range were identified as outliers for model2a and model2b combined.

See chapter 3.3 for a list of identified dentists in model 2.

See figures 15-16 in the appendices chapter for model 2a and model 2b outliers plots

3 Results

There are dentists that are possibly placing restorations with lacking quality and providing unnecessary treatments. DTSS expenditure on restorations is high and it is important to study it closely. The aim was to identify potential dentists that were a high quality and probity risk by analysing their FTEA and RR.

3.1 Data Validation

- There were no dentists outside of the set range, so no cases were deleted.
- One case with a negative number of total patients was identified and removed.
- 76 dentists were identified as having a RR of fillings greater than one, i.e., number of failed fillings greater than total number of fillings, and were removed.
- 13 dentists with total earnings less than €5000 were identified and excluded from model 1.
- 207 dentists were identified as having a provision of less than 10 fillings and were excluded from model 2.

Note: there was some overlap between total earnings and fillings provided. Eight out of ten dentists with total earnings less than €5000 also provided less than 10 fillings, automatically excluding dentists with total earning less than €5000 from model 2.

Missing Values

Three out of 22 of the variables had cases with missing values, resulting in 641 missing values combined prior to data validation. After data validation, there were 536 cases with missing values for model 1 and 302 cases with missing values for model 2.

Model 1 had 192 cases with missing values from dentist age, 176 cases from total patients, and 168 cases from practice type. 22.7%, 20.8%, and 19.9% respectively.

Model 2 had 87 cases with missing values from dentist age, 70 cases from total patients, and 64 cases from practice type. 13.2%, 10.6%, and 9.7% respectively.

Partial F-test to determine the significance of the three variables

Null Hypothesis: Variables dentist age, total patients, and practice type have no effect on the FTEA. ANOVA was used to compare two models, full and reduced, to determine if the three variables with missing values were significant. The following results were gathered from the ANOVA. F = 28.439 with critical value 1.14 ¡ 28.439, and the p-value = 2.2e-16 ¡ 0.05; therefore, the null hypothesis was rejected with 95% confidence that the three variables dentist age, total patients, and practice type have little to no effect on FTEA.

See figure 1 for ANOVA results.

```
Model 1: FTEA ~ factor(DENT_HB_NUM) + Fillings + Failures + Dentists +
    OralExam + Propylaxes + Extraction + ProtractedPerio + Endodontic +
    OtherTreatment + Amalgam + Patients + Female + MeanAge +
    MeanBaselineDMFT
Model 2: FTEA ~ factor(DENT_HB_NUM) + DentAge + TotalPatients +
factor(PracticeType) +
    Fillings + Failures + Dentists + OralExam + Propylaxes +
    Extraction + ProtractedPerio + Endodontic + OtherTreatment +
    Amalgam + Patients + Female + MeanAge + MeanBaselineDMFT
  Res.Df    RSS Df Sum of Sq      F    Pr(>F)
1    623 2.4308
2    620 2.1367  3   0.29403 28.439 < 2.2e-16 ***
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1
```

Figure 1: ANOVA results

When comparing the two groups, with and without missing values, significant differences in the values of explanatory variables were observed. This indicated the removal of the three variables would distort the analyses. Thus, supporting the results of the partial F-test.

See figure 5 in the appendices chapter for a summary of the explanatory variables in a dataframe with missing values.

See figure 6 in the appendices chapter for a summary of the explanatory variables in

```
   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.    NA's
0.02996 0.16817 0.20704 0.21516 0.25842 0.50781       6
```

Figure 2: Summary of FTEA

a dataframe without the missing values

3.2 Model 1 Identified Dentists

The outliers identified between model1a and model1b were dentist ID numbers 43, 338, 517, 585, 905, 982, and 1086. Three dentists were from model1a and the other four from mode1b. All dentists identified had FTEA greater than the median fraction of 20.7%. Dentist ID numbers 43, 585, and 982 had a very high FTEA; with 44.9%, 49.8%, and 43% respectively. The highest FTEA came from dentists in health boards one and eight. All seven dentists provided fillings to patients aged near the median age of 25 years. Dentist ID number 43 provided six amalgam fillings out of eight fillings in total and earned €11,755.

See figure 2 for the minimum, maximum, and median FTEA values in model 1. See table 2 for the full list of identified dentists.

3.3 Model 2 Identified Dentists

The outliers identified between model2a and model2b were dentist ID numbers 2, 24, 152, 173, 181, 188, 209, 221, 378, 425, 463, 503, 528, 566, 567, 647, 653, 804, 805, 941, 960, 1022, 1059, 1092, 1096, 1129, 1132, 1136, 1138, 1265, 1357, 1359, 1385, 1432. With 34 dentists identified in total, 13 of them came from model2a and 23 of them from model2b, with two dentists in common. See table below for identified dentists' details. The median of RR in this model was 15%, with a minimum of 0% and a maximum of 72.73%. Three dentists had very low RR with rates below 5%; 7 dentists had excessively large RR with rates greater than 35%. Most of the

	FTEA	RR	DENT_HB_NUM	DentAge	TotalPatients	PracticeType
30	0.4490755	0.37500000	8	NA	NA	NA
210	0.3572633	0.50000000	2	NA	NA	NA
321	0.2651090	0.07692308	7	38.56188	548	0
362	0.4982603	0.14285714	1	28.96936	198	1
544	0.2393036	0.09497207	4	66.22343	1740	0
593	0.4300882	0.25000000	1	41.96729	51	1
652	0.3169463	0.10714286	8	30.54807	122	1

	Fillings	Failures	Dentists	OralExam	Propylaxes	Extraction	ProtractedPerio
30	8	3	14	16	9	5	8
210	2	1	8	8	12	2	0
321	117	9	127	149	89	0	64
362	56	8	44	44	41	4	0
544	179	17	79	73	65	146	7
593	28	7	23	22	22	71	0
652	224	24	340	382	205	68	285

	Endodontic	OtherTreatment	Amalgam	Composite	Patients	Female	MeanAge
30	0	7	6	2	4	4	32.52167
210	0	8	0	2	1	0	23.38703
321	20	33	99	18	18	13	24.72993
362	0	14	52	4	6	5	26.74493
544	7	10	159	20	26	17	23.12673
593	0	1	25	3	5	5	23.27157
652	69	293	188	36	29	18	23.99055

	MeanBaselineDMFT	AmalgamEarnings	TotalEarnings	DentistID
30	10	11755	26176	43
210	12	4449	12453	338
321	11	68567	258637	517
362	13	36088	72428	585
544	13	97202	406187	905
593	16	4537	10549	982
652	8	79480	250768	1086

Table 2: Identified Dentists in Model 1

dentists identified were from health boards one and seven; and health boards with least number of identified dentists were two and four.

See Table 3 for the full list of identified dentists.

	FTEA	RR	DENT_HB_NUM	DentAge	TotalPatients	PracticeType	Fillings	Failures
1	0.1674804	0.08622398	1	26.42389	5866	1	1009	87
15	0.2340569	0.02739726	1	29.64432	730	1	73	2
98	0.2260437	0.26368159	8	60.81054	1542	1	201	53
112	0.1761524	0.36842105	1	NA	NA	NA	19	7
116	0.4156194	0.03125000	7	NA	NA	NA	96	3
119	0.2786462	0.10303030	8	29.87558	93	1	165	17
132	0.1718772	0.68000000	6	52.46209	696	0	25	17
139	0.2195781	0.07973856	5	32.38649	2672	0	765	61
232	0.2213873	0.31106870	3	NA	3012	0	524	163
284	0.3148072	0.09848485	8	24.49265	477	1	396	39
263	0.1860472	0.20360825	1	NA	2945	0	388	79
311	0.1490079	0.18766067	5	63.89758	2699	0	389	73
328	0.2278599	0.32413793	7	31.92351	1323	1	145	47
350	0.1701140	0.08966862	3	28.49301	2488	0	513	46
351	0.1532780	0.36363636	1	72.51430	1983	0	88	32
399	0.1696008	0.15923567	3	28.40099	2712	0	314	50
405	0.3421879	0.22751323	5	35.65332	1171	1	378	86
487	0.1919284	0.19642857	6	NA	3858	0	336	66
488	0.3362069	0.07692308	1	NA	NA	NA	13	1
569	0.2098756	0.31137725	7	32.40467	186	1	167	52
581	0.2686256	0.23003195	6	57.19073	2731	0	313	72
616	0.2214045	0.13209393	8	31.78963	3530	0	1022	135
635	0.1980634	0.12109375	8	28.85913	950	0	256	31
655	0.1277917	0.26666667	7	NA	NA	NA	15	4
658	0.2538984	0.20696517	4	47.75865	3219	0	1005	208
680	0.2817100	0.26751592	7	29.14387	1004	1	314	84
682	0.1996558	0.08888889	1	33.47044	668	0	45	4
685	0.1412460	0.12643678	1	56.32171	2139	0	87	11
686	0.1154220	0.60465116	7	60.91444	494	0	43	26
761	0.1666999	0.04464286	7	35.72392	1377	1	336	15
814	0.3023654	0.23013699	1	70.14498	1720	0	365	84
816	0.1971206	0.36781609	1	27.68221	1115	1	87	32
831	0.2861746	0.38181818	7	53.56169	628	0	110	42
866	0.1207473	0.68965517	7	42.20334	506	0	29	20

Table 3: Identified Dentists in Model 2

4 Discussion

4.1 Data Validation Justification

- In Ireland, on average, students sit the leaving certification exams around age 17-18, and complete dentistry school after five years following the leaving certification; therefore, it is reasonable to set the minimum dentist age to 22. The average retirement age of dentists listed on the Irish Register of Dentists is 65, with nearly 10% of the dentists being older than 65 [3]. Hence, an upper limit was not set.

- Cases with total earnings less than €5000 were excluded from Model 1 because

they could indicate dentists who have not treated many patients, or dentists who entered the DTSS much later than the start of the observational period, resulting in unrepresentative data. Additionally, with such dentists, a small unusual pattern could have a high impact on the FTEA. For instance, it could be the case that all the treatments provided by a dentist were amalgam fillings. Such a case would result in 100% FTEA. A 100% FTEA would generally be of interest in this study, but since the dentist only has total earnings less than €5000, it might cost the system more to investigate the dentist than what the saving might be. Cases with €0 were also removed since they indicate dentists who are paid a salary, instead of being paid on a treatment-by-treatment basis. Such cases do not contribute to this study, particularly when analysing FTEA, since a salary paid dentist receives the same amount regardless of the treatment provided. They also would not have sufficient motivation to commit fraud and behave indecently since they are not affected in the variability of prices the patients are charged.

- Dentists who provided less than 10 fillings were excluded from Model 2 for similar reasons as the former variable. Low number of fillings could be due to censoring, i.e. dentist that entered the scheme after the start of observational period, dentist that entered at the start but exited before end date, or dentists that entered late and exited early. It could also be due to low activity and experience. These dentists have not provided enough fillings for reasonable conclusions to be reached about their re-treatment rates, and would have similar effects on the proportion as the former variable would on FTEA. So, it might not be worth investigating dentists with very small sample size of fillings.

4.1.1 Missing Values

There are cases in the dataset with missing values. These missing values could have occurred due to data entry error, or due to nonresponse. Missing values can raise an

issue when analysing data because they reduce the representativeness of the sample; however, removing them would have resulted in removing too much data making the analysis less valid.

4.2 Overdispersion

Overdispersion occurs in binomial data when the variability in the data is greater than what would be expected if the responses were truly binomial. Model 2 was checked for overdispersion since it is a binomial data and resulted in an apparent overdispersion. The mean deviance of the model was checked because it is indicative of overdispersion; i.e. a mean deviance greater than one may indicate overdispersion; as was the case in this model. This may be explained by an incorrectly specified systematic component or presence of outliers. However, the systematic component of the model was checked through plots and corrected through applying transformations; so, the issue was not with the systematic component. On the other hand, 34 outliers were detected in the model, which might explain the apparent overdispersion.

The overdispersion could be explained by the correlation between the responses. There might exist a correlation between the responses in the database used since there might be cases where patients appeared twice because of getting more than one treatment.

While the overdispersion problem exists in this model, standard error corrections were not made because the standard errors of the model parameters were not used; instead, only the estimates of the model parameters were used to calculate the predictive values and residuals. Since this logistic regression model does not take the overdispersion into account, it will underestimate the standard errors of the model parameters, but the estimates of the model parameters remain consistent.

4.3 Identified Dentists

The aim of this study was to identify dentists with extreme FTEA and RR values. The identified dentists from Model 1 had large FTEA. Three out of seven of the dentists had excessively high FTEA, with 49.8%, 43%, and 44.9%. The first two were from health board 2, and the third one was from health board 8. The other four values were not necessary extreme, but they were higher than the median, ranging from 23.9% to 35.7%. Two of the dentists had fillings less than 10, which might explain their high percentage. Dentists with lower number of fillings provided tend to have higher FTEA. Model 2 had many outliers, totaling 34 dentists. 29% of the identified dentists in Model 2 had either excessively large or excessively low RR. The model had more identified dentists in health board numbers one and seven, whereas health board numbers two and four had the least number of dentists identified. This might have been due to the potential effects of water fluoridation, or lack thereof in some regions in Ireland.

4.4 Relevant Studies

There have been a few studies regarding probity assurance in the DTSS. In 2009, Oral Care Consulting produced a report on probity assurance in the DTSS stating that there had been improvement in probity assurance in the DTSS between the first report conducted by the HSE in 2002 and the second report in 2007. However, they reported a decline in commitment to probity assurance by the HSE since 2007, and estimated at least 10% of payments were likely to be inappropriate [7]. The Public Dental Service (PDS) also produced a report highlighting the role they could play in probity assurance, specifically in estimating the financial aspect of the scheme.

5 Conclusion

Tooth decay is prevalent in Ireland, with more than 90% of adults experiencing it. It is commonly treated with a restoration procedure, with either an amalgam filling or a composite filling. The DTSS has been providing access to dental treatments to medical card holders who are over the age of 16 since November 1994. With 32% of total expenditures being from restorations, and 65% of the restorations being from amalgam fillings; the need for probity and quality assurance in the DTSS has been identified. Many patients are affected by the provision of restorations, as well as the government, as they provide funding for these procedures. It is imperative that the provision of amalgam fillings is given more attention.

Probity and quality of dentists in the DTSS were analysed by initially calculating the fraction of total earnings from amalgam fillings, as well as the re-treatment rates of fillings provided. Linear regression analysis was used to identify 3 dentists with excessively high FTEA; and 10 dentists were identified as having excessively high or low re-treatment rates. A comparative analysis of the identified dentists' explanatory variables showed that, besides their high proportions, there was nothing particularly odd about them. I recommend further study/investigation into the dentists identified to figure out why they have such high proportions.

6 Appendices

```
      FTEA                RR            DENT_HB_NUM       DentAge
 Min.   :0.02996   Min.   :0.0000   Min.   :1.000   Min.   :22.58
 1st Qu.:0.16907   1st Qu.:0.1098   1st Qu.:1.000   1st Qu.:29.11
 Median :0.20730   Median :0.1613   Median :4.000   Median :33.86
 Mean   :0.21533   Mean   :0.2123   Mean   :4.125   Mean   :38.22
 3rd Qu.:0.25833   3rd Qu.:0.2667   3rd Qu.:7.000   3rd Qu.:44.66
 Max.   :0.50781   Max.   :0.8571   Max.   :8.000   Max.   :72.51
                                                    NA's   :176
 TotalPatients     PracticeType       Fillings        Failures
 Min.   :   4.0   Min.   :0.0000   Min.   :   2.0   Min.   :  0.00
 1st Qu.: 280.8   1st Qu.:0.0000   1st Qu.:  20.0   1st Qu.:  4.00
 Median : 660.5   Median :1.0000   Median :  60.0   Median : 10.00
 Mean   : 930.1   Mean   :0.5149   Mean   : 118.6   Mean   : 16.95
 3rd Qu.:1303.2   3rd Qu.:1.0000   3rd Qu.: 157.0   3rd Qu.: 20.00
 Max.   :5866.0   Max.   :1.0000   Max.   :1439.0   Max.   :208.00
 NA's   :161      NA's   :153
    Dentists         OralExam        Propylaxes       Extraction
 Min.   :   0.0   Min.   :   0.0   Min.   :   0.0   Min.   :  0.00
 1st Qu.:  20.0   1st Qu.:  21.0   1st Qu.:  17.0   1st Qu.:  5.00
 Median :  57.0   Median :  64.0   Median :  50.0   Median : 25.00
 Mean   : 104.4   Mean   : 125.7   Mean   : 102.8   Mean   : 59.25
 3rd Qu.: 132.0   3rd Qu.: 161.0   3rd Qu.: 134.0   3rd Qu.: 73.00
 Max.   :1376.0   Max.   :1511.0   Max.   :1146.0   Max.   :823.00

 ProtractedPerio    Endodontic     OtherTreatment       Amalgam
 Min.   :  0.00   Min.   :  0.00   Min.   :   0.00   Min.   :   0.00
 1st Qu.:  0.00   1st Qu.:  0.00   1st Qu.:   5.00   1st Qu.:  15.00
 Median :  6.00   Median :  0.00   Median :  19.00   Median :  50.00
 Mean   : 24.18   Mean   : 10.31   Mean   :  53.42   Mean   :  93.64
 3rd Qu.: 23.00   3rd Qu.: 10.00   3rd Qu.:  61.00   3rd Qu.: 117.00
 Max.   :800.00   Max.   :207.00   Max.   :1021.00   Max.   :1229.00

   Composite         Patients         Female         MeanAge
 Min.   :  0.00   Min.   :  1.00   Min.   :  0.00   Min.   :16.67
 1st Qu.:  4.00   1st Qu.:  6.00   1st Qu.:  4.00   1st Qu.:23.67
 Median : 12.00   Median : 15.00   Median : 10.00   Median :25.08
 Mean   : 24.91   Mean   : 26.52   Mean   : 17.06   Mean   :25.33
 3rd Qu.: 33.00   3rd Qu.: 35.00   3rd Qu.: 23.00   3rd Qu.:26.61
 Max.   :317.00   Max.   :254.00   Max.   :174.00   Max.   :36.65

 MeanBaselineDMFT AmalgamEarnings  TotalEarnings      DentistID
 Min.   : 0.000   Min.   :   575   Min.   :   5223   Min.   :   2.0
 1st Qu.: 8.000   1st Qu.: 15883   1st Qu.:  76630   1st Qu.: 386.0
 Median : 9.000   Median : 36637   Median : 186501   Median : 759.0
 Mean   : 9.661   Mean   : 60961   Mean   : 304098   Mean   : 754.2
 3rd Qu.:11.000   3rd Qu.: 81737   3rd Qu.: 406290   3rd Qu.:1135.0
 Max.   :27.000   Max.   :527257   Max.   :2233700   Max.   :1505.0
```

Figure 3: Summary of explanatory variables in the first dataframe

```
     FTEA               RR          DENT_HB_NUM       DentAge
Min.   :0.04212   Min.   :0.0000   Min.   :1.000   Min.   :22.58
1st Qu.:0.16818   1st Qu.:0.1053   1st Qu.:1.000   1st Qu.:29.09
Median :0.20621   Median :0.1500   Median :4.000   Median :33.85
Mean   :0.21553   Mean   :0.1803   Mean   :4.206   Mean   :38.04
3rd Qu.:0.25872   3rd Qu.:0.2222   3rd Qu.:7.000   3rd Qu.:44.47
Max.   :0.50781   Max.   :0.7273   Max.   :8.000   Max.   :72.51
NA's   :2                                          NA's   :114
TotalPatients     PracticeType       Fillings        Failures
Min.   :   4.0   Min.   :0.0000   Min.   :  10.0   Min.   :  0.00
1st Qu.: 332.0   1st Qu.:0.0000   1st Qu.:  31.0   1st Qu.:  6.00
Median : 716.0   Median :1.0000   Median :  83.0   Median : 12.00
Mean   : 980.2   Mean   :0.5104   Mean   : 135.6   Mean   : 19.16
3rd Qu.:1348.0   3rd Qu.:1.0000   3rd Qu.: 179.0   3rd Qu.: 23.00
Max.   :5866.0   Max.   :1.0000   Max.   :1439.0   Max.   :208.00
NA's   :96       NA's   :90
   Dentists         OralExam       Propylaxes      Extraction
Min.   :   0.0   Min.   :   0.0   Min.   :   0.0   Min.   :  0.0
1st Qu.:  32.0   1st Qu.:  35.0   1st Qu.:  26.0   1st Qu.:  9.0
Median :  73.0   Median :  82.0   Median :  64.0   Median : 33.0
Mean   : 118.8   Mean   : 143.1   Mean   : 117.1   Mean   : 67.6
3rd Qu.: 156.0   3rd Qu.: 182.0   3rd Qu.: 151.0   3rd Qu.: 83.0
Max.   :1376.0   Max.   :1511.0   Max.   :1146.0   Max.   :823.0

ProtractedPerio    Endodontic     OtherTreatment       Amalgam
Min.   :  0.00   Min.   :  0.00   Min.   :   0.00   Min.   :   4.0
1st Qu.:  1.00   1st Qu.:  0.00   1st Qu.:   8.00   1st Qu.:  23.0
Median :  8.00   Median :  1.00   Median :  26.00   Median :  61.0
Mean   : 27.62   Mean   : 11.72   Mean   :  60.91   Mean   : 107.1
3rd Qu.: 27.00   3rd Qu.: 14.00   3rd Qu.:  72.00   3rd Qu.: 137.0
Max.   :800.00   Max.   :207.00   Max.   :1021.00   Max.   :1229.0

  Composite         Patients         Female         MeanAge
Min.   :  0.00   Min.   :  1.00   Min.   :  0.00   Min.   :17.22
1st Qu.:  6.00   1st Qu.:  9.00   1st Qu.:  6.00   1st Qu.:23.71
Median : 16.00   Median : 19.00   Median : 12.00   Median :25.02
Mean   : 28.48   Mean   : 30.16   Mean   : 19.39   Mean   :25.15
3rd Qu.: 37.00   3rd Qu.: 40.00   3rd Qu.: 25.00   3rd Qu.:26.29
Max.   :317.00   Max.   :254.00   Max.   :174.00   Max.   :34.95

MeanBaselineDMFT AmalgamEarnings  TotalEarnings       DentistID
Min.   : 1.000   Min.   :     0   Min.   :      0   Min.   :   2.0
1st Qu.: 8.000   1st Qu.: 21633   1st Qu.: 106998   1st Qu.: 378.0
Median : 9.000   Median : 45566   Median : 228566   Median : 751.0
Mean   : 9.665   Mean   : 68571   Mean   : 341952   Mean   : 755.8
3rd Qu.:11.000   3rd Qu.: 91558   3rd Qu.: 449933   3rd Qu.:1145.0
Max.   :22.000   Max.   :527257   Max.   :2233700   Max.   :1505.0
```

Figure 4: Summary of explanatory variables in the second dataframe

6.1 Figures 3-16

```
 DENT_HB_NUM       DentAge    TotalPatients  PracticeType    Fillings
Min.   :1.000   Min.   : NA   Min.   : NA   Min.   : NA   Min.   :  1.00
1st Qu.:1.000   1st Qu.: NA   1st Qu.: NA   1st Qu.: NA   1st Qu.:  5.00
Median :4.000   Median : NA   Median : NA   Median : NA   Median : 10.00
Mean   :4.173   Mean   :NaN   Mean   :NaN   Mean   :NaN   Mean   : 15.99
3rd Qu.:7.000   3rd Qu.: NA   3rd Qu.: NA   3rd Qu.: NA   3rd Qu.: 20.25
Max.   :8.000   Max.   : NA   Max.   : NA   Max.   : NA   Max.   :105.00
                NA's   :168   NA's   :168   NA's   :168
   Failures         Dentists        OralExam       Propylaxes
Min.   : 0.000   Min.   :  1.00   Min.   :  1.00   Min.   :  0.00
1st Qu.: 2.000   1st Qu.: 11.00   1st Qu.: 14.00   1st Qu.:  9.00
Median : 3.000   Median : 20.50   Median : 23.50   Median : 18.00
Mean   : 3.399   Mean   : 33.62   Mean   : 37.38   Mean   : 28.82
3rd Qu.: 5.000   3rd Qu.: 44.25   3rd Qu.: 47.00   3rd Qu.: 38.50
Max.   :13.000   Max.   :276.00   Max.   :311.00   Max.   :258.00

  Extraction     ProtractedPerio     Endodontic    OtherTreatment
Min.   :  0.00   Min.   :  0.000   Min.   : 0.00   Min.   :  0.00
1st Qu.:  2.00   1st Qu.:  0.000   1st Qu.: 0.00   1st Qu.:  3.00
Median :  6.00   Median :  2.000   Median : 0.00   Median :  8.00
Mean   : 15.82   Mean   :  7.845   Mean   : 2.97   Mean   : 18.16
3rd Qu.: 18.75   3rd Qu.:  9.000   3rd Qu.: 4.00   3rd Qu.: 21.25
Max.   :138.00   Max.   :165.000   Max.   :35.00   Max.   :189.00

   Amalgam        Composite         Patients         Female          MeanAge
Min.   : 0.00   Min.   : 0.000   Min.   : 1.00   Min.   : 0.000   Min.   :18.51
1st Qu.: 4.00   1st Qu.: 1.000   1st Qu.: 2.00   1st Qu.: 1.000   1st Qu.:25.21
Median : 8.00   Median : 2.000   Median : 4.00   Median : 2.500   Median :27.72
Mean   :12.11   Mean   : 3.887   Mean   : 5.31   Mean   : 3.744   Mean   :27.63
3rd Qu.:16.00   3rd Qu.: 6.000   3rd Qu.: 7.00   3rd Qu.: 5.000   3rd Qu.:30.13
Max.   :88.00   Max.   :20.000   Max.   :23.00   Max.   :21.000   Max.   :38.67

MeanBaselineDMFT AmalgamEarnings TotalEarnings      DentistID
Min.   : 0.000   Min.   : 1288   Min.   :  5223   Min.   :  22.0
1st Qu.: 7.000   1st Qu.: 9228   1st Qu.: 42359   1st Qu.: 349.5
Median : 9.000   Median :17950   Median : 80556   Median : 721.0
Mean   : 9.327   Mean   :21044   Mean   : 99280   Mean   : 710.7
3rd Qu.:11.000   3rd Qu.:28845   3rd Qu.:139260   3rd Qu.:1055.8
Max.   :32.000   Max.   :90924   Max.   :547690   Max.   :1503.0

     FTEA              RR
Min.   :0.04027   Min.   :0.0000
1st Qu.:0.18261   1st Qu.:0.1333
Median :0.21877   Median :0.2679
Mean   :0.22752   Mean   :0.3591
3rd Qu.:0.26681   3rd Qu.:0.5000
Max.   :0.47327   Max.   :1.0000
```

Figure 5: Summary of explanatory variables in a dataframe with missing values

```
 DENT_HB_NUM       DentAge      TotalPatients     PracticeType
Min.   :1.000   Min.   :22.58   Min.   :   4.0   Min.   :0.0000
1st Qu.:1.000   1st Qu.:29.14   1st Qu.: 271.0   1st Qu.:0.0000
Median :4.000   Median :33.92   Median : 645.0   Median :1.0000
Mean   :4.132   Mean   :38.31   Mean   : 905.5   Mean   :0.5271
3rd Qu.:7.000   3rd Qu.:44.66   3rd Qu.:1283.0   3rd Qu.:1.0000
Max.   :8.000   Max.   :72.51   Max.   :5866.0   Max.   :1.0000
   Fillings         Failures         Dentists         OralExam
Min.   :   1.0   Min.   :  0.00   Min.   :   0.0   Min.   :   0.0
1st Qu.:  35.0   1st Qu.:  6.00   1st Qu.:  28.0   1st Qu.:  29.0
Median :  89.0   Median : 13.00   Median :  74.0   Median :  86.0
Mean   : 142.3   Mean   : 19.89   Mean   : 120.2   Mean   : 145.5
3rd Qu.: 192.0   3rd Qu.: 24.00   3rd Qu.: 162.0   3rd Qu.: 200.0
Max.   :1439.0   Max.   :208.00   Max.   :1376.0   Max.   :1511.0
  Propylaxes       Extraction     ProtractedPerio    Endodontic
Min.   :   0.0   Min.   :  0.00   Min.   :  0.00   Min.   :  0.00
1st Qu.:  22.0   1st Qu.:  8.00   1st Qu.:  0.00   1st Qu.:  0.00
Median :  67.0   Median : 32.00   Median :  7.00   Median :  0.00
Mean   : 119.4   Mean   : 68.33   Mean   : 28.23   Mean   : 11.94
3rd Qu.: 164.0   3rd Qu.: 90.00   3rd Qu.: 27.00   3rd Qu.: 14.00
Max.   :1146.0   Max.   :823.00   Max.   :800.00   Max.   :207.00
OtherTreatment       Amalgam         Composite         Patients
Min.   :   0.00   Min.   :   0.0   Min.   :  0.00   Min.   :  1.0
1st Qu.:   6.00   1st Qu.:  26.0   1st Qu.:  6.00   1st Qu.:  9.0
Median :  24.00   Median :  67.0   Median : 18.00   Median : 21.0
Mean   :  60.37   Mean   : 112.4   Mean   : 29.82   Mean   : 31.4
3rd Qu.:  75.00   3rd Qu.: 153.0   3rd Qu.: 38.00   3rd Qu.: 42.0
Max.   :1021.00   Max.   :1229.0   Max.   :317.00   Max.   :254.0
    Female          MeanAge      MeanBaselineDMFT AmalgamEarnings
Min.   :  0.00   Min.   :16.67   Min.   : 2.000   Min.   :   575
1st Qu.:  5.00   1st Qu.:23.43   1st Qu.: 8.000   1st Qu.: 18427
Median : 13.00   Median :24.82   Median : 9.000   Median : 48811
Mean   : 20.12   Mean   :24.82   Mean   : 9.757   Mean   : 69899
3rd Qu.: 27.00   3rd Qu.:26.05   3rd Qu.:11.000   3rd Qu.: 95827
Max.   :174.00   Max.   :34.70   Max.   :27.000   Max.   :527257
TotalEarnings       DentistID           FTEA                RR
Min.   :   5424   Min.   :   2.0   Min.   :0.02996   Min.   :0.0000
1st Qu.:  91583   1st Qu.: 370.0   1st Qu.:0.16731   1st Qu.:0.1071
Median : 244680   Median : 739.0   Median :0.20439   Median :0.1522
Mean   : 347654   Mean   : 752.1   Mean   :0.21294   Mean   :0.1964
3rd Qu.: 483164   3rd Qu.:1138.0   3rd Qu.:0.25390   3rd Qu.:0.2300
Max.   :2233700   Max.   :1505.0   Max.   :0.50781   Max.   :1.0000
```

Figure 6: Summary of explanatory variables in a dataframe without missing values

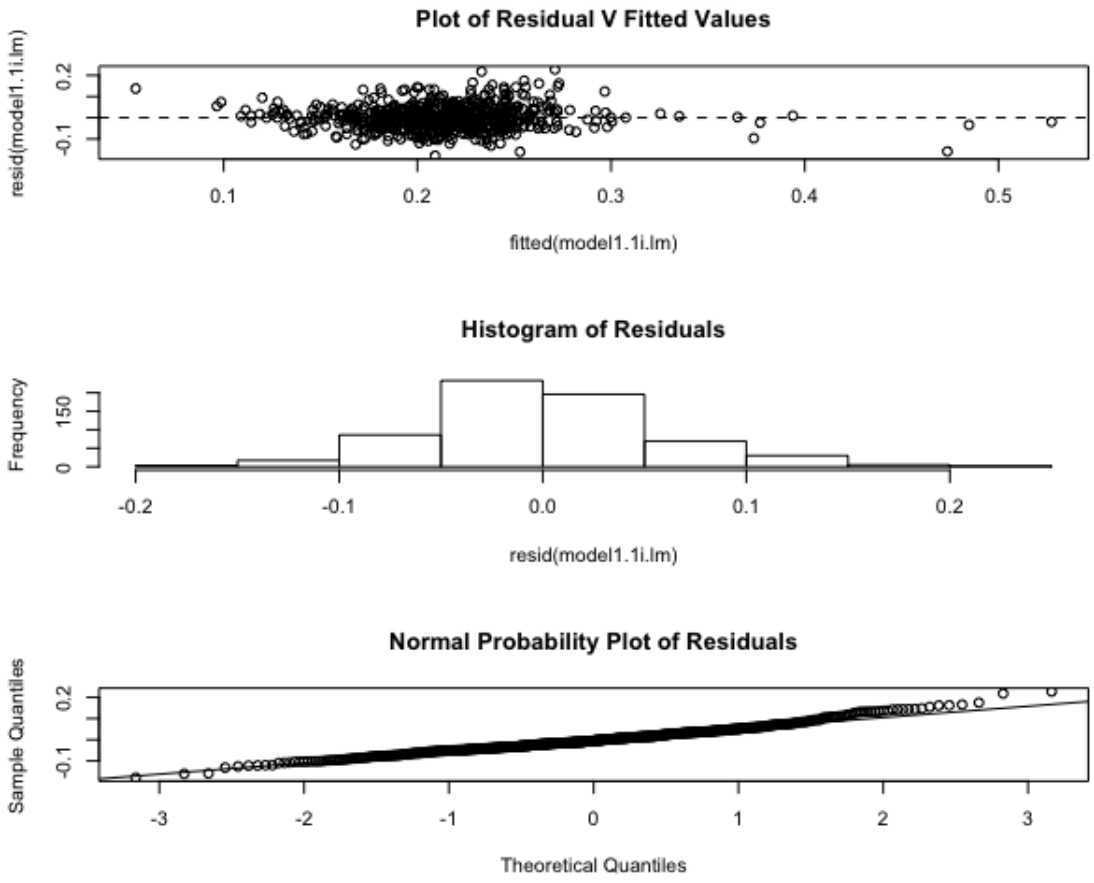

Figure 7: Diagnostic residual plots of the initial full model; model 1.1i

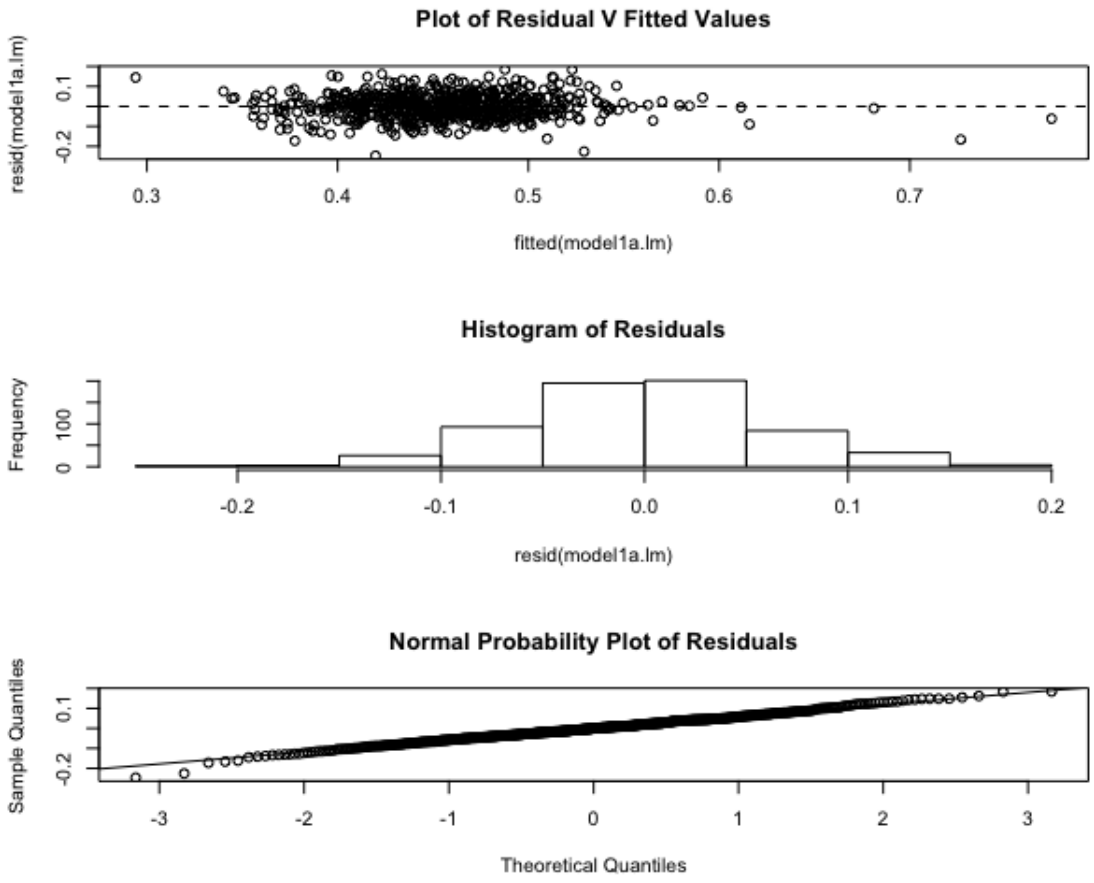

Figure 8: Diagnostic residual plots of the initial full model; model 1a

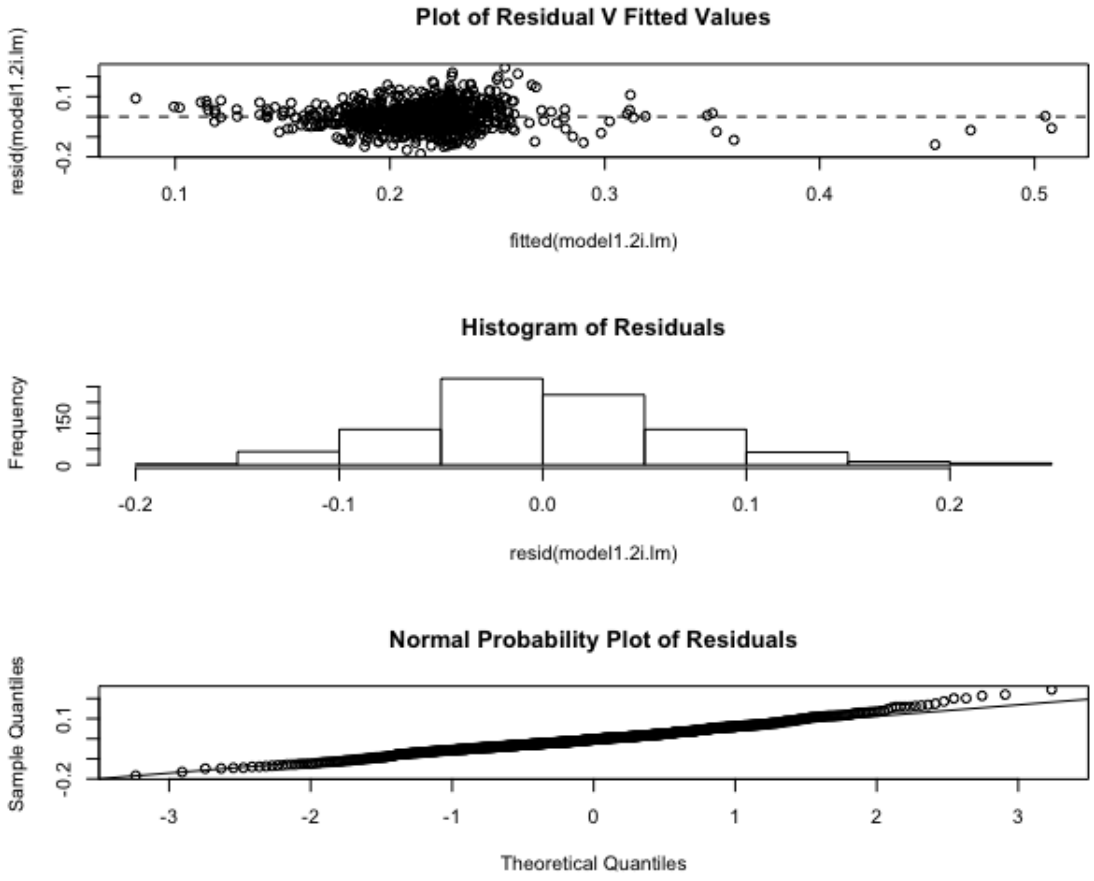

Figure 9: Diagnostic residual plots of the initial full model; model 1.2i

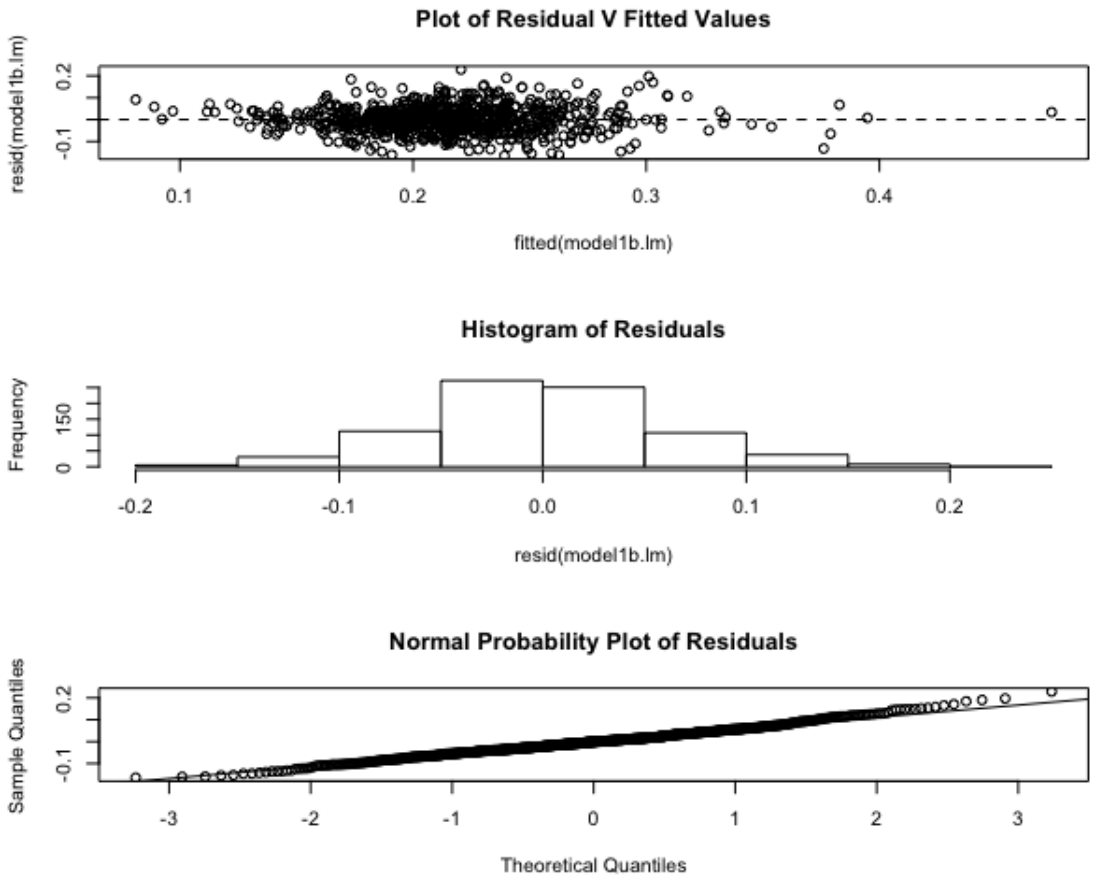

Figure 10: Diagnostic residual plots of the initial full model; model 1b

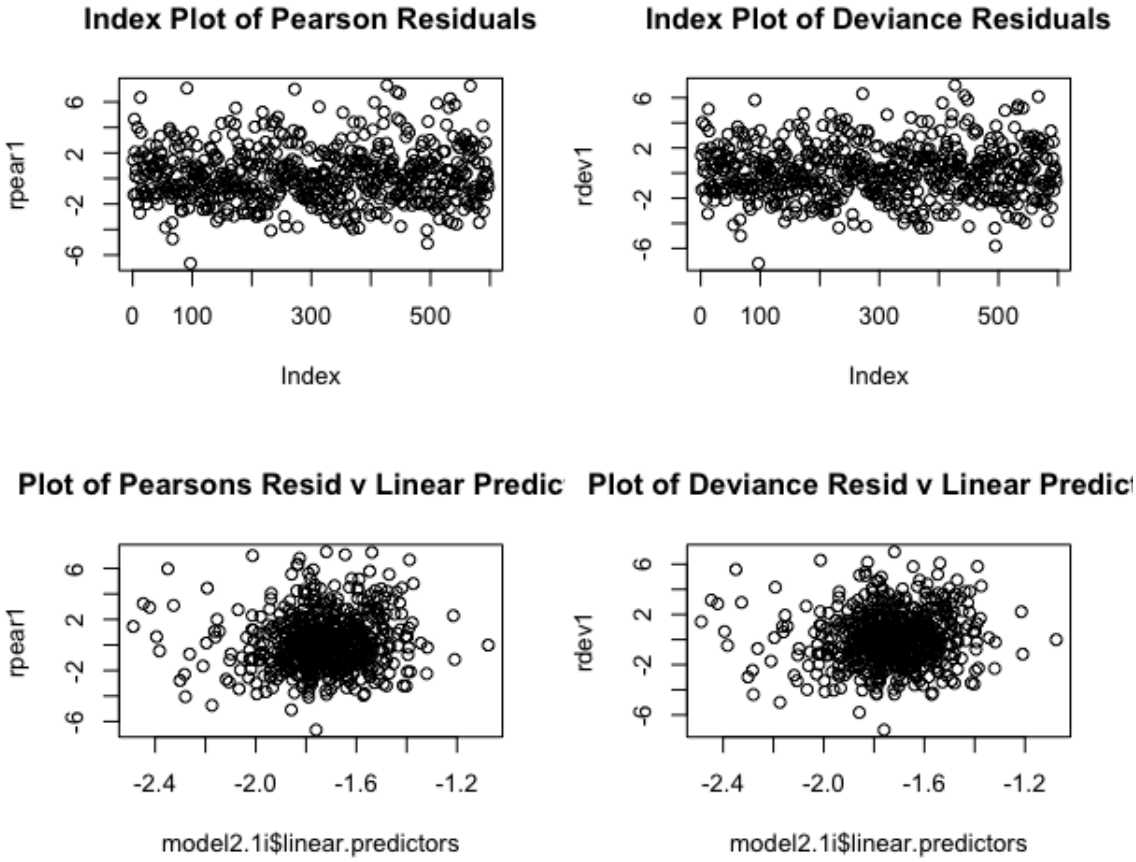

Figure 11: Checking the systematic component of the initial full model; model 2.1i

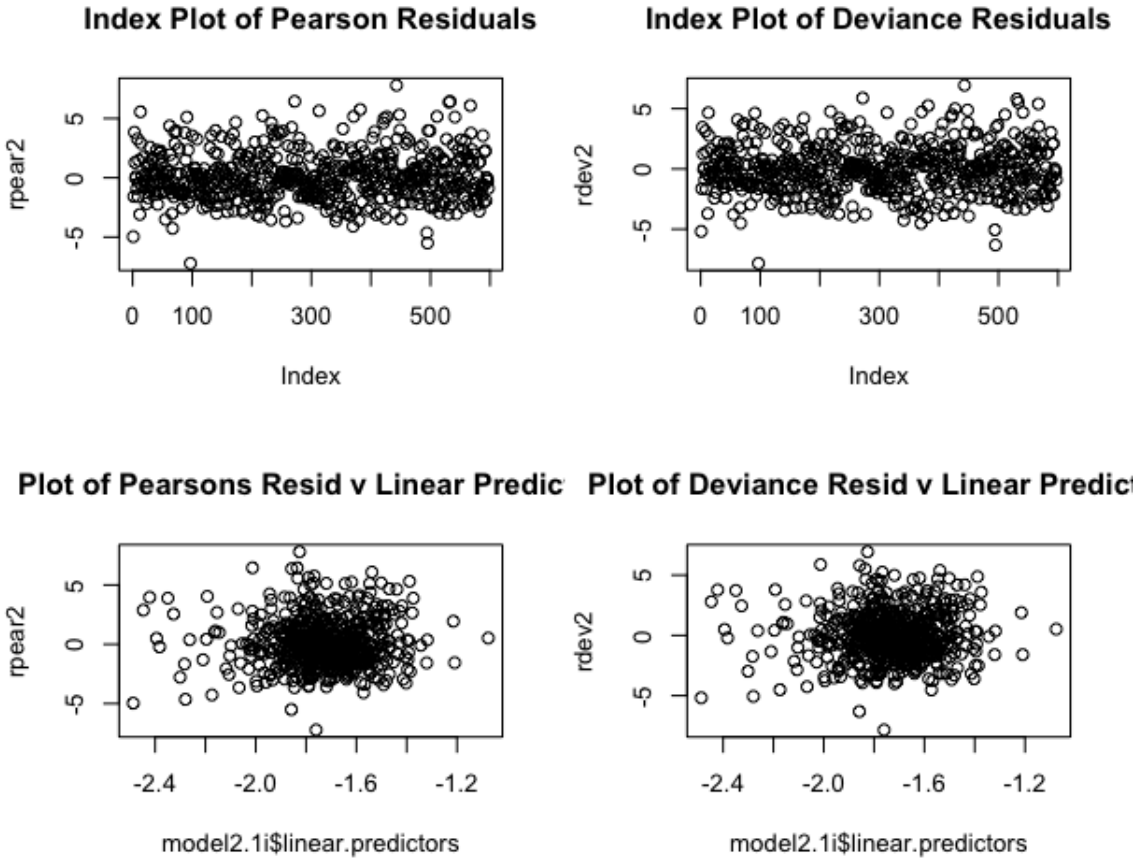

Figure 12: Checking the systematic component of the initial full model; model 2a

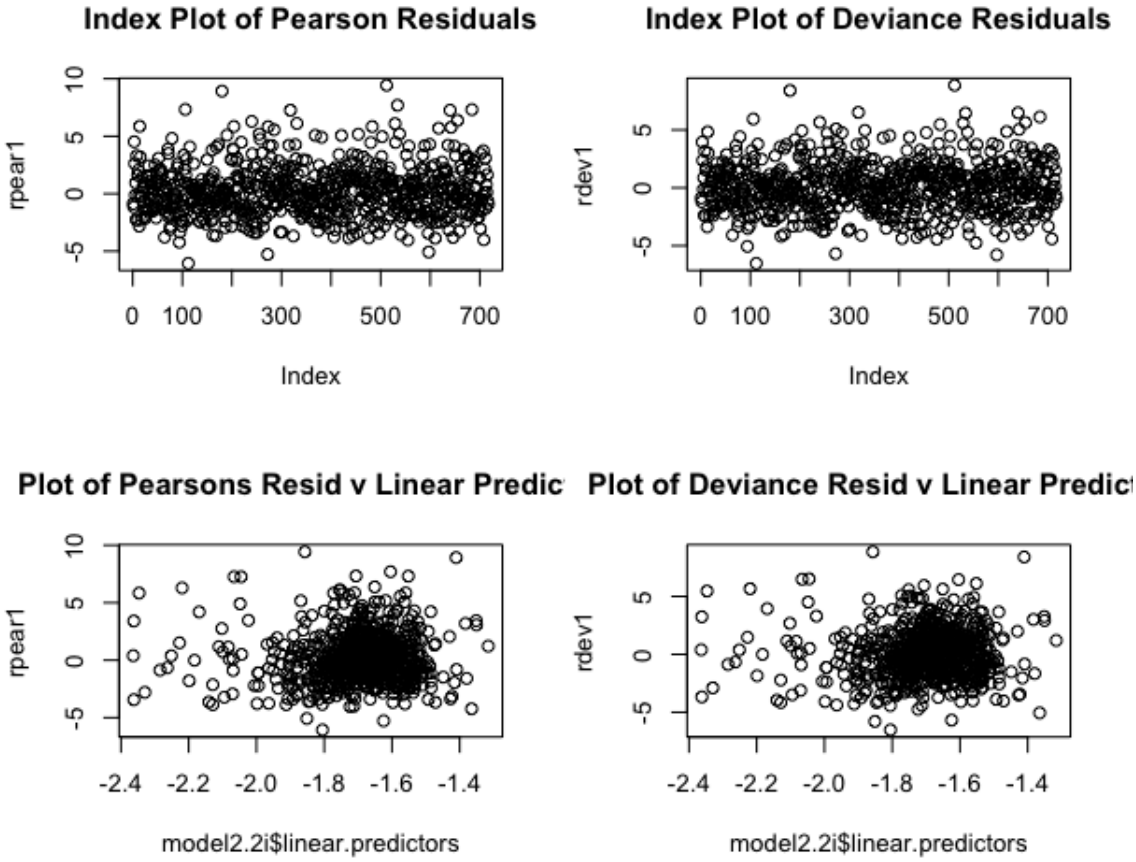

Figure 13: Checking the systematic component of the initial reduced model; model 2.2i

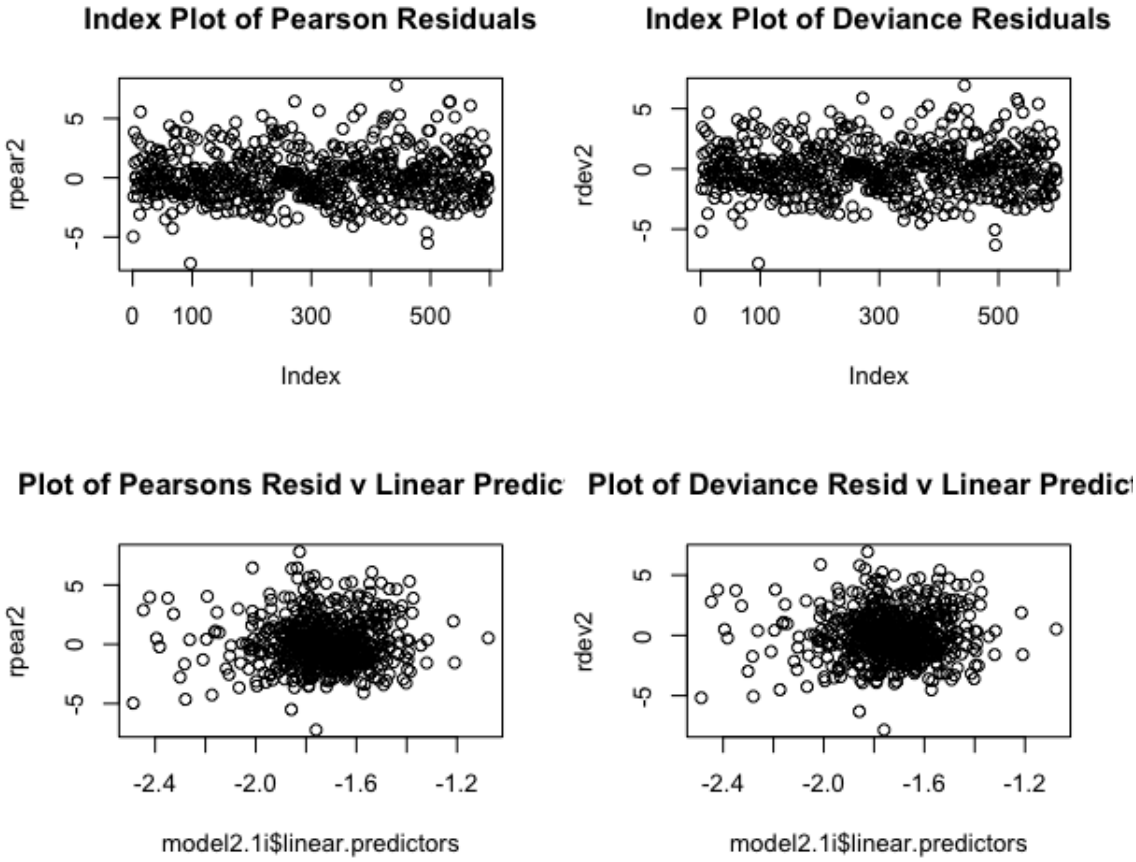

Figure 14: Checking the systematic component of the initial full model; model 2b

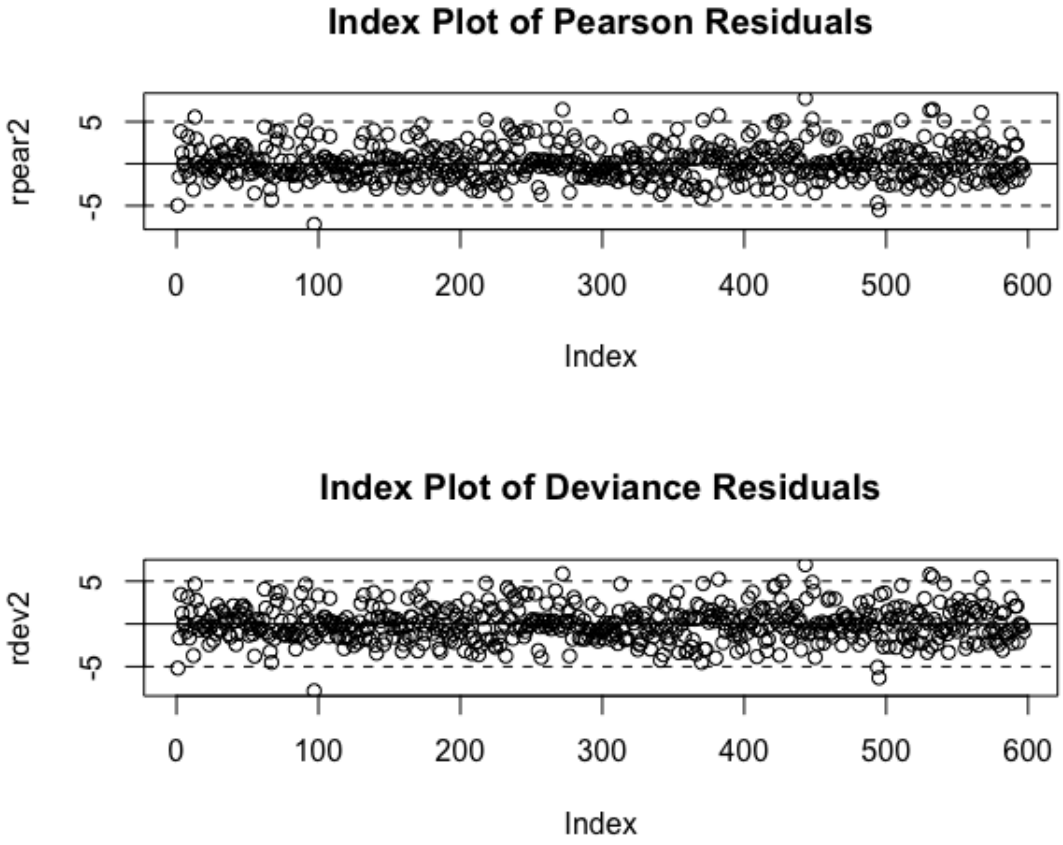

Figure 15: Outliers in model 2a

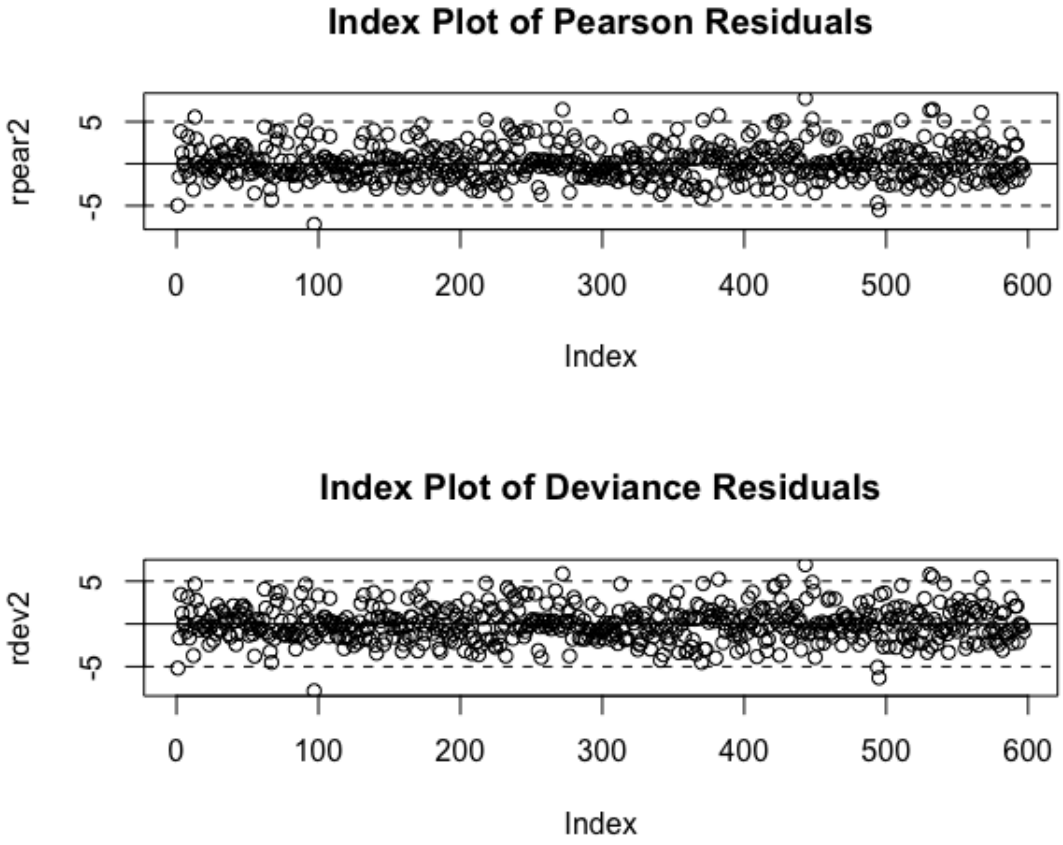

Figure 16: Outliers in Model 2b

6.2 R Code

```
data.df <-read.csv("~/Documents/Modules/ST6009/data.csv", header=T)

#remove negative total patient value
data.df <- data.df[-602,]

attach(data.df)

#create new column for retreatment rate : failures/fillings
RR <- Failures/Fillings
data.df <- data.frame(RR, data.df)

#create new column for amalgam earning ratio
FTEA <- AmalgamEarnings/TotalEarnings
data.df <- data.frame(FTEA, data.df)

# MODEL 1

#remove rows where total earnings < 5000
dtss1 <- data.df[data.df[,23]>4999,]

#remove rows where retreatment rates are > 1
dtss1 <- dtss1[dtss1[,2]<1,]

#Model 1: assessing missing values
#with NAs
#with=dtss1[is.na(dtss1["DentAge"]) & is.na(dtss1["TotalPatients"])
#& is.na(dtss1["PracticeType"]) , ]

#w/o NAs
#without=dtss1[!is.na(dtss1["DentAge"]) & !is.na(dtss1["TotalPatients"])
#& !is.na(dtss1["PracticeType"]) , ]

```

```
33 #summary(with)
34 #summary(without)
35
36 detach(data.df)
37 attach(dtss1)
38
39 #full model scatter smooth plots
40 par(mfrow=c(2,2))
41 scatter.smooth(factor(DENT_HB_NUM), FTEA, main="Scatter.smooth Plot (HB
       #)")
42 scatter.smooth(DentAge, FTEA, main="Scatter.smooth Plot (Age)")
43 scatter.smooth(TotalPatients, FTEA, main="Scatter.smooth Plot (TP)")
44 scatter.smooth(factor(PracticeType), FTEA, main="Scatter.smooth Plot (PT
       )")
45
46 #residual plots
47 model1.1i.lm <- lm(FTEA ~ factor(DENT_HB_NUM) + DentAge + TotalPatients
48                    + factor(PracticeType) + Fillings + Failures +
       Dentists
49                    + OralExam + Propylaxes + Extraction +
       ProtractedPerio
50                    + Endodontic + OtherTreatment + Amalgam + Patients +
       Female
51                    + MeanAge + MeanBaselineDMFT, data=dtss1)
52
53 par(mfrow=c(3,1))
54 plot(fitted(model1.1i.lm),resid(model1.1i.lm),main="Plot of Residual V
       Fitted Values")
55 abline(h=0, lty=2)
56 hist(resid(model1.1i.lm),main="Histogram of Residuals")
57 qqnorm(resid(model1.1i.lm),main="Normal Probability Plot of Residuals")
58 qqline(resid(model1.1i.lm))
59
60 # linear model1a post transformation residual plots
```

```
model1a.lm <- lm(sqrt(FTEA) ~ factor(DENT_HB_NUM) + sqrt(DentAge)
                 + sqrt(TotalPatients) + factor(PracticeType) + Fillings
                 + Failures + sqrt(Dentists) + OralExam + Propylaxes
                 + Extraction + ProtractedPerio + Endodontic +
    OtherTreatment
                 + Amalgam + Patients + Female + MeanAge +
    MeanBaselineDMFT, data=dtss1)

par(mfrow=c(3,1))
plot(fitted(model1a.lm),resid(model1a.lm),main="Plot of Residual V
    Fitted Values")
abline(h=0, lty=2)
hist(resid(model1a.lm),main="Histogram of Residuals")
qqnorm(resid(model1a.lm),main="Normal Probability Plot of Residuals")
qqline(resid(model1a.lm))

par(mfrow=c(2,1))
#model1a high leverage cases
#plot of leverage vs. observation number &
h <- lm.influence(model1a.lm)$hat
plot(h, type="h", main="Plot of Leverage vs. observation number", ylab="
    Leverage")
abline(h=0.2, lty=2)
#dtss1[c(1, 70, 77, 151, 458, 492),]

#calculate studentized residuals
e <- resid(model1a.lm)
s <- summary(model1a.lm)$sigma
r <- e / (s*(1-h)^0.5)

#plot of studentized residuals vs. observation number
plot(r, type="h", main="Plot of Studentized Residuals vs. observation
    number",
      ylab="Studentized Residuals", ylim=c(-4,4))
```

```
90 abline(h=0, lty=1)
91 abline(h=c(-3,3), lty=2)
92 #model1a outliers
93 #dtss1[c(287, 492, 591),]
94
95 # model1.2i residual plots
96 model1.2i.lm <- lm(FTEA ~ factor(DENT_HB_NUM) + Fillings + Failures +
       Dentists
97                     + OralExam + Propylaxes + Extraction +
       ProtractedPerio
98                     + Endodontic + OtherTreatment + Amalgam + Patients
99                     + Female + MeanAge + MeanBaselineDMFT, data=dtss1)
100
101 par(mfrow=c(3,1))
102 plot(fitted(model1.2i.lm),resid(model1.2i.lm),main="Plot of Residual V
       Fitted Values")
103 abline(h=0, lty=2)
104 hist(resid(model1.2i.lm),main="Histogram of Residuals")
105 qqnorm(resid(model1.2i.lm),main="Normal Probability Plot of Residuals")
106 qqline(resid(model1.2i.lm))
107
108
109 # model1b post transformatiom residual plots
110 model1b.lm <- lm(FTEA ~ factor(DENT_HB_NUM) + sqrt(Fillings) + Failures
111                     + sqrt(Dentists) + sqrt(OralExam) + sqrt(Propylaxes)
112                     + sqrt(Extraction) + ProtractedPerio + Endodontic
113                     + OtherTreatment + sqrt(Amalgam) + sqrt(Patients)
114                     + sqrt(Female) + MeanAge + MeanBaselineDMFT, data=dtss1
       )
115
116 par(mfrow=c(3,1))
117 plot(fitted(model1b.lm),resid(model1b.lm),main="Plot of Residual V
       Fitted Values")
118 abline(h=0, lty=2)
```

```
119 hist(resid(model1b.lm),main="Histogram of Residuals")
120 qqnorm(resid(model1b.lm),main="Normal Probability Plot of Residuals")
121 qqline(resid(model1b.lm))
122
123 par(mfrow=c(2,1))
124 #calculate the leverages h
125 #plot of leverage vs. observation number
126 h <- lm.influence(model1b.lm)$hat
127 plot(h, type="h", main="Plot of Leverage vs. observation number", ylab="
        Leverage")
128 abline(h=0.2, lty=2)
129 #dtss1[c(86, 189, 204, 596, 640),]
130
131 #calculate studentized residuals
132 h <- lm.influence(model1b.lm)$hat
133 e <- resid(model1b.lm)
134 s <- summary(model1b.lm)$sigma
135 r <- e / (s*(1-h)^0.5)
136
137 #plot of studentized residuals vs. observation number
138 plot(r, type="h", main="Plot of Studentized Residuals vs. observation
        number",
139       ylab="Studentized Residuals", ylim=c(-4,4))
140 abline(h=0, lty=1)
141 abline(h=c(-3,3), lty=2)
142 #dtss1[c(26, 185, 326, 539),]
143
144 #combiend outliers; model1a & model1b
145 #dtss1[c(26, 185, 287, 326, 492, 539, 591),]
146
147 detach(dtss1)
148 #-------------------------------------------------------------------
149 attach(data.df)
150 # MODEL 2
```

```
151
152 #remove rows where fillings < 10
153 dtss2<-data.df[data.df[,7]>=10,]
154
155 #remove rows where retreatment rates are > 1
156 dtss2 <- dtss2[dtss2[,2]<1,]
157
158 detach(data.df)
159 attach(dtss2)
160
161 RR <- cbind(Failures,Fillings-Failures)
162
163 # model2.1i initial full model
164 model2.1i <- glm(RR ~ factor(DENT_HB_NUM) + DentAge + TotalPatients
165                  + factor(PracticeType) + Dentists + OralExam +
      Propylaxes
166                  + Extraction + ProtractedPerio + Endodontic +
      OtherTreatment
167                  + Amalgam + Patients + Female + MeanAge +
      MeanBaselineDMFT
168                  + AmalgamEarnings + TotalEarnings, family=binomial(link
      =logit))
169
170 #standardizing the residuals
171 h1<-lm.influence(model2.1i)$hat
172 #pearson and deviance residuals
173 rpear1 <- residuals(model2.1i, "pearson")/sqrt(1-h1)
174 rdev1 <- residuals(model2.1i, "deviance")/sqrt(1-h1)
175
176 # Model 2 Diagnostics
177 par(mfrow=c(2,2))
178 #1.1 Index Plot
179 plot(rpear1,main="Index Plot of Pearson Residuals")
180 plot(rdev1,main="Index Plot of Deviance Residuals")
```

```
#1.2 Plot of Residuals vs. Linear Predictor
plot(model2.1i$linear.predictors, rpear1, main="Plot of Pearsons Resid v Linear Predictor")
plot(model2.1i$linear.predictors, rdev1, main="Plot of Deviance Resid v Linear Predictor")
#pattern present in plots. Apply transformations

#model2a post transformations
model2a <- glm(RR ~ factor(DENT_HB_NUM) + sqrt(DentAge) + sqrt(TotalPatients)
                + factor(PracticeType) + sqrt(Dentists) + OralExam + Propylaxes
                + Extraction + ProtractedPerio + Endodontic + OtherTreatment
                + Amalgam + sqrt(Patients) + Female + MeanAge + MeanBaselineDMFT
                + sqrt(AmalgamEarnings) + sqrt(TotalEarnings), family=binomial(link=logit))

#model2a systematic component
h2<-lm.influence(model2a)$hat
rpear2 <- residuals(model2a, "pearson")/sqrt(1-h2)
rdev2 <- residuals(model2a, "deviance")/sqrt(1-h2)
par(mfrow=c(2,2))
plot(rpear2,main="Index Plot of Pearson Residuals")
plot(rdev2,main="Index Plot of Deviance Residuals")
#1.2 Plot of Residuals vs. Linear Predictor
plot(model2.1i$linear.predictors, rpear2, main="Plot of Pearsons Resid v Linear Predictor")
plot(model2.1i$linear.predictors, rdev2, main="Plot of Deviance Resid v Linear Predictor")

#check for outliers
par(mfrow=c(2,1))
```

```
206 plot(rpear2,main="Index Plot of Pearson Residuals")
207 abline(h=0, lty=1)
208 abline(h=c(-5,5), lty=2)
209 #dtss2[c("13", "91", "97", "218", "272", "313", "382", "443", "495",
      "511", "531", "533"),]
210
211 plot(rdev2,main="Index Plot of Deviance Residuals")
212 abline(h=0, lty=1)
213 abline(h=c(-5,5), lty=2)
214 #dtss2[c("1", "97", "272", "382", "443", "495", "531", "533"),]
215
216 #COMBINED OUTLIERS rpear2 & rdev2 for FULL MODEL model2a
217 #dtss2[c(1, 13, 91, 97, 218, 272, 313, 382, 443, 495, 511, 531, 533 ),]
218
219 # REDUCED MODEL
220
221 #initial model
222 model2.2i <- glm(RR ~ factor(DENT_HB_NUM) + Dentists + OralExam +
      Propylaxes
223                  + Extraction + ProtractedPerio + Endodontic +
      OtherTreatment
224                  + Amalgam + Patients + Female + MeanAge +
      MeanBaselineDMFT
225                  + AmalgamEarnings + TotalEarnings, family=binomial(link
      =logit))
226
227 h1<-lm.influence(model2.2i)$hat
228 rpear1 <- residuals(model2.2i, "pearson")/sqrt(1-h1)
229 rdev1 <- residuals(model2.2i, "deviance")/sqrt(1-h1)
230
231 # Diagnostic #1; check the systematic component
232 par(mfrow=c(2,2))
233 #1.1 Index Plot
234 plot(rpear1,main="Index Plot of Pearson Residuals")
```

```
235 plot(rdev1,main="Index Plot of Deviance Residuals")
236 #1.2 Plot of Residuals vs. Linear Predictor
237 plot(model2.2i$linear.predictors, rpear1, main="Plot of Pearsons Resid v
      Linear Predictor")
238 plot(model2.2i$linear.predictors, rdev1, main="Plot of Deviance Resid v
    Linear Predictor")
239
240 #Model2b post transformations
241 model2b <- glm(RR ~ factor(DENT_HB_NUM) + sqrt(Dentists) + OralExam +
    Propylaxes
242                 + Extraction + ProtractedPerio + Endodontic +
    OtherTreatment
243                 + Amalgam + sqrt(Patients) + Female + MeanAge +
    MeanBaselineDMFT
244                 + sqrt(AmalgamEarnings) + sqrt(TotalEarnings), family=
    binomial(link=logit))
245
246 #model2b systematic component
247 h2<-lm.influence(model2b)$hat
248 rpear2 <- residuals(model2b, "pearson")/sqrt(1-h2)
249 rdev2 <- residuals(model2b, "deviance")/sqrt(1-h2)
250
251 par(mfrow=c(2,2))
252 plot(rpear2,main="Index Plot of Pearson Residuals")
253 plot(rdev2,main="Index Plot of Deviance Residuals")
254 plot(model2b$linear.predictors, rpear2, main="Plot of Pearsons Residuals
      v Linear Predictor")
255 plot(model2b$linear.predictors, rdev2, main="Plot of Deviance Residuals
    v Linear Predictor")
256
257 #Model2b Outliers
258 par(mfrow=c(2,1))
259 plot(rpear2,main="Index Plot of Pearson Residuals")
260 abline(h=0, lty=1)
```

```
261 abline(h=c(-5,5), lty=2)
262 #dtss2[c(1, 79, 106, 112, 180, 202, 240, 255, 272, 273, 318,
263 #381, 455, 483, 512, 529, 534, 597, 640, 642, 655, 684),]
264
265 plot(rdev2,main="Index Plot of Deviance Residuals")
266 abline(h=0, lty=1)
267 abline(h=c(-5,5), lty=2)
268 #dtss2[c(1, 94, 106, 112, 180, 272, 273, 318, 455,
269 #512, 529, 534, 597, 640, 642, 655, 684),]
270
271 #COMBINED OUTLIERS rpear2 & rdev2 for REDUCED MODEL model2b
272 #dtss2[c(1, 79, 94, 106, 112, 180, 202, 240, 255, 272,
273 #273, 318, 381, 455, 483, 512, 529, 534, 597, 640, 642, 655, 684 ),]
274
275 #COMBINED OUTLIERS for model2a and model2b
276 #dtss2[c(1, 13, 79, 91, 94, 97, 106, 112, 180, 202, 218,
277 #240, 255, 272, 273, 313, 318, 381, 382, 443, 455, 483, 495, 511,
278 #512, 529, 531, 533, 534, 597, 640, 642, 655, 684 ),]
279
280 #
```

7 Bibliography

"Dental Caries (Tooth Decay) in Adults (Age 20 to 64)."National Institute of Dental and Craniofacial Research, U.S. Department of Health and Human Services, www.nidcr.nih.gov/Data Statistics/FindDataByTopic/DentalCaries/DentalCariesAdults20to64.htm. [1]

Foundation, Dental Health. "Dental Caries (Tooth Decay)."Dental Health Foundation, 20 Nov. 2014, www.dentalhealth.ie/dentalheal th/causes/dentalcaries.html. [2]

Foundation, Dental Health. "Dental Services."Dental Health Foundation, 30 Sept. 2010, www.dentalhealth.ie/dentalhealth/services.html. [3]

Woods, N., et al. "The Healthcare System and the Provision of Oral Healthcare in European Union Member States. Part 7: Republic of Ireland." Vol. 222, no. 7, 7 Apr. 2017, pp. 541–548. [4]

"Dental Services."Citizens Information, www.citizensinformation.ie/en/health/healthservices/dentalauralandopticalservices/dentalservices.html. [5]

Whelton, Helen, et al.Dental Treatment Services Scheme – Service: Utilisation and Treatment Need – Lot 9. health.gov.ie/wp-content/uploads/2014/03/lot9.pdf+. [6]

Lynch, Liam. "Results of a peer review process: the distribution of codes by examining dentists in the Republic of Ireland 2006-2007." Irish Dental Association, 02 Aug. 2017, http://www.lenus.ie/hse/handle/10147/236471 [7]