A Two Part Study: Which Pre-16 and Post-16 Factors Influence Income at 25 years old?

Total word count (including lay report hard copy): 2497

Contents Page

1. Introduction

The aim of this project is to analyse factors influencing income at age 25 (*W8DINCW*) using longitudinal data from 16,000 individuals in England born in 1989-90.
This analysis focuses on life factors to provide clear insights for policy holders; explicitly examining the following research questions:

- **RQ1: Which early life factors collected during mandatory school years 10-11 (Wave 1-4) influence income at age 25?**
- **RQ2: Which later life factors collected post-16 (Wave 5-8) influence income at age 25?**

Waves 1-4 were used for RQ1 to capture pre-16 adolescence, where school environments are broadly similarly structured. RQ2 focuses on later waves, capturing post-16 transitions, where pathways into education, work, or training become more varied.

Our analysis highlights that individuals from single-parent households, routine occupational backgrounds and Black, South Asian or Mixed backgrounds were likely to earn less at age 25. Later life factors suggest that lower income at 25 is associated with having a child by age 17, not being at university at 17 and having a less positive attitude towards debt.

2. Exploratory Data Analysis (EDA)

2.1 Plots

Boxplots assessed categorical predictors against the outcome *W8DINCW* (income). Variables with clear associations to income were noted for likely model inclusion including *W1wrk1aMP*. Scatterplots of continuous predictors revealed skewness and non-linearity. For example, *W1GrssyrHH* was highly right-skewed so applying a log transformation improved linearity, revealing a clear positive correlation against *W8DINCW*. While the transformed variable could've been retained, we instead recoded it into a categorical variable (*hh_income_quintile*) using income band thresholds for interpretability; its boxplot showed stepped increases in median income across quintiles likely for model inclusion. Variables with step-like scatterplots were converted into factors and examined using boxplots - this revealed a more distinguishable association with income for *W6DebtattYP*.

2.2 Missing values, Merging and Relevelling

Missing values, coded as negative numbers, were recoded as N/A for continuous variables and retained for categorical variables by grouping them into a combined 'missing' level. This allowed us to assess whether non-response patterns were informative.

Baselines were set for categorical predictors using the level with most observations; levels were also merged where observations were rare and box plots displayed similar structures with the concordant category for better interpretability and statistical power- Appendix A.

3. Model Fitting

3.1 Method

Our model was developed using backward elimination, refined by removing predictors that were statistically insignificant at the 5% level.

Predictors with high missingness (93% *W6NEETAct*) were dropped. To avoid endogeneity, we also removed Wave 8 predictors that could be simultaneously influenced by the outcome (like *W8DWRK, W8DACTIVITY*). However, *W8CMSEX* (gender) was retained as a key interest and potential confounder.

Conceptually overlapping predictors were removed to avoid redundancy and for better interpretability. For example, parental employment variables (like *W1wrkfullmum*, *W1empsmum*) were excluded in favour of summary variables (*W1wrk1aMP)*. Broader household composition variables (*W1depkids*, *W1NoldBroHS*) were also retained over multiple age-specific child variables (eg. *W1ch0_2HH*) to reduce complexity without losing information.

3.2 Interactions

We tested the interaction between university attendance and attitude to debt, given the conceptual link of university decisions potentially affected by fear of debt and both measured in W6 suggesting their connection. However, the interaction term was not statistically significant (p=0.18) so was excluded.

We also tested if the effect of family social class (NS-SEC) on income differed by ethnicity, given they had the largest standardised coefficients (-0.286 and -0.25 respectively) and linking to forms of socioeconomic disadvantage. While one interaction (routine backgrounds with Black ethnicity) was weakly significant, the overall interaction term was not (p=0.54), thus excluded.

3.3 Outlier Analysis

Outliers were assessed using standardised residual diagnostics, leverage, Cook's Distance, and DFFITS. Observation 2679 stood out across all four diagnostics: a standardised residual of 3.13 (under-estimating income by £116, given an RSE of 37.17), leverage of 0.22 (above the 2p/n = 0.07 threshold), DFFITS of 0.87 (threshold: 0.37), and the highest Cook's Distance (0.018) in the model.

Seven observations were flagged by multiple diagnostics, but even the most extreme (observation 4060) did not distort model fit. The Residuals vs Fitted plot suggested mild heteroscedasticity, with larger errors at higher income levels, though the pattern was weak so no transformations applied. Q-Q plot showed broadly normal residuals, supporting linear model assumptions. Leverage vs Standardised Residuals identified three high-leverage points, and Cook's Distance showed several values above threshold 0.003 but none at problematic levels.

These outliers likely reflected natural variation in income prediction so were not removed.

3.4 Cross Validation

A subset of our final model predictors was created and unused factor levels dropped to prevent estimation errors. Model performance was evaluated using 100 iterations of random sub-sampling (80/20 split). The average in-sample MSE was 1229.25 and out-of-sample MSE was 1324.91. A 70/30 split produced similar results, decreasing in-sample MSE to 1221.56 and increasing out-of-sample MSE to 1336.25 suggesting minor overfitting and reasonable generalisability.

3.5 Final Model Comparison

We evaluated eight models using AIC and adjusted R^2. The inclusion of *hh_income_quintile*, (49% missing) in Model_6, was compared with Model_7 (without it). Despite missingness, Model_6 performed better, with a lower AIC (12,165.22 vs 12,170.01) and higher adjusted R^2 (0.6469 vs 0.6443) suggesting a better model at explaining income variation.

Multicollinearity was assessed using variance inflation factors (VIF). *W1marstatmum* displayed extreme collinearity with *W1hiqualmum* (GVIF~700) so was removed very early on. Model_8 reintroduced it, as it was statistically significant. Although AIC decreased slightly (12,128.38), VIF again exceeded thresholds. We instead retained *W1famtyp2*, which improved adjusted R^2 and resolved multicollinearity issues (the highest being 3.37 across this Model_6). It also offers clearer understanding as a straightforward binary indicator of family structure.

Model_6 was therefore selected as the final model, explaining 64.7% of the variation in income at age 25. An F-statistic of 56.09 confirms model significance. This represents a decent model fit given the complexity of income trajectories. Assumptions of normality were checked using a Q-Q plot with no major concerns. However, the residuals vs fitted plot observed mild heteroscedasticity.

4. Results for Main Model

Predictor	**Estimate**	**P-value**	**Signif.**
$(Intercept)$	361.2289	$< 2e{-}16$	$* * *$
$W1wrk1aMPParttime$	7.9226	0.000659	$* * *$
$W1wrk1aMPEconomicallyinactive/Unemployed/Other$	-29.8002	0.172334	
$W1wrk1aMPLookingafterhome/family$	-45.6353	0.007100	$**$
$W1hiqualmumDegree$	-12.2621	0.000807	$* * *$
$W1hiqualmumHighereducation(predegree)$	-7.2659	0.028565	$*$
$W1hiqualmumFurthereducation(16-18)$	-2.3066	0.466818	
$W1hiqualmumNoqualification$	-28.4066	$1.90e{-}12$	$* * *$
$W1famtyp2Yes$	-34.7867	$< 2e{-}16$	$* * *$
$W1nssecfamMissing$	-13.6717	0.024850	$*$
$W1nssecfamIntermediate$	-5.1661	0.220247	
$W1nssecfamSelf-employed$	7.8196	0.091167	.
$W1nssecfamSupervisory/technical$	-3.5307	0.357905	
$W1nssecfamRoutine/Semi-routine$	-39.4779	$< 2e{-}16$	$* * *$
$W1nssecfamNeverWorked/unemployed$	12.4573	0.744513	
$W1ethgrpYPMissing$	-84.7001	0.024219	$*$
$W1ethgrpYPOther/Mixed$	-79.1094	$< 2e{-}16$	$* * *$
$W1ethgrpYPSouthAsian$	-69.7603	$< 2e{-}16$	$* * *$
$W1ethgrpYPBlack-African/Caribbean$	-78.2691	$< 2e{-}16$	$* * *$
$W1heposs9YPDon'tknow$	-5.3389	0.455022	
$W1heposs9YPFairlylikely$	-2.1040	0.412527	
$W1heposs9YPNotverylikely$	-4.9381	0.175611	
$W1heposs9YPNotatalllikely$	-17.9275	0.000685	$* * *$
$W1disabYPHasdisability, schoolingaffected$	-31.2265	$4.59e{-}08$	$* * *$
$W1disabYPHasdisability, schoolingnotaffected$	-5.0433	0.184658	
$W2disc1YPMissing$	-10.4322	0.075874	.
$W2disc1YPYes$	-11.4767	0.008119	$**$
$W4CannTryYPYes$	16.5639	$2.34e{-}12$	$* * *$
$W4schatYP$	0.9423	0.005784	$**$
$W5EducYPMissing$	-72.6742	0.054067	.
$W5EducYPNo$	-7.7301	0.003226	$**$
$W5Apprent1YPYes$	-17.6955	0.000365	$* * *$
$W6JobYPNo$	-7.1723	0.003141	$**$
$W6UnivYPYes$	13.4015	$8.10e{-}07$	$* * *$
$W6OwnchiDVYes$	-29.7219	$9.03e{-}05$	$* * *$
$W6DebtatYP$	1.1774	0.000771	$* * *$
$W8CMSEXMale$	-30.9429	$< 2e{-}16$	$* * *$
$hh\_income\_quintileBottomQuintile(< 13,057)$	-14.7447	0.002216	$**$
$hh\_income\_quintile4thQuintile(13,057-19,204)$	-9.9028	0.028928	$*$
$hh\_income\_quintile3rdQuintile(19,204-29,939)$	-6.6767	0.038211	$*$
$hh\_income\_quintileTopQuintile(>= 49,400)$	1.8188	0.573437	

*Table 1: *** for p < 0.001, ** for p < 0.01, * for p < 0.05, . for p < 0.1, and no symbol for p ≥ 0.1*

4.1 RQ1 Discussion

The intercept (361.23) represents a predicted weekly income of £361.23 for an individual with all continuous predictors at 0 and belonging to their reference categories. The key ones being a female, White ethnicity, no disability, full-time parental employment and medium household income. All estimates represent differences in predicted weekly earnings (£ per week)

Parental employment showed that having a parent working part time vs full time (reference level) predicted a higher income of £7.92 per week, while parents not in paid work and looking after the home predicted earning £45.64 less per week. Having a mother with 'No qualification' predicted £28.41 less per week. Interestingly, individuals whose mothers had a degree resulted in £12.26 lower compared to those whose mothers had general secondary education.

Single parent households were associated with £34.78 less. Routine/ semi-routine households predicted £39.48 lower income and being in a family class of self-employed, over managerial/professional backgrounds, was predictive of higher income of £7.82.

Ethnicity had one of the strongest effects, with all minority groups earning substantially less than their White peers. Black and Mixed/Other groups experienced the largest reductions (£78.27 and £79.11). Those experiencing unfair treatment due to skin colour or ethnicity were predicted to earn £11.48 less per week.

Disability affecting schooling predicted £31.23 lower income. A positive school attitude predicted £0.94 higher income per week, and early cannabis usage (by age 15) predicted £16.56 higher income per week.

Finally, households earning less than £13,057 predicted earning £14.74 less than those from middle quintile households [£29,939–£49,400). The second lowest quintile [£13,057–£19,204) earned £9.90 less and [£19,204–£29,939) earned £6.68 less.

4.2 RQ2 Discussion

Post-16 predictors (Wave 5+) mostly related to education and employment. Not being in education at age 16 earned £7.73 less per week at age 25. Interestingly, those doing an apprenticeship at age 16 earned £17.70 less while those attending university predicted earning £13.40 more.

Not working or having a child at age 17 predicted a lower income by £7.17 and £29.72 per week respectively. Finally, 17-year-olds reporting more positive attitudes towards debt earned £1.18 more. This may reflect greater willingness to take out a student loan to attend university, potentially being the real driver behind higher income.

4.3 Missing Level Analysis

Significant missingness predicted lower weekly income. NS-SEC class missingness led to earning £13.67 less per week, while being treated unfairly due to ethnicity earned £10.43 less. Missing responses for education aged 16 also earned £76.67 less. On inspection, missingness was due to 'refused' based on 10 observations so it might be an unstable, inflated estimate. Consequently, we conducted a complete case analysis.

5. Complete Case Analysis

5.1 Model Comparison

We examined whether predictor effects changed without any missing data. Backward elimination was reapplied and the only variable no longer significant was *hh_income_quintile,* thus removed.

The model's adjusted R-squared reduced from 0.6468 to 0.6348 - slightly lower explanatory power. This suggests missingness may have been informative. AIC might not be meaningful given a different data size but it was considerably lower.

5.2 Outlier Analysis

Plots showed similar patterns but all influence scores were higher, likely due to reduced sample size increasing the relative impact of outliers. For example, observation 2679 remained the most influential across all diagnostics, particularly its standardised residual (3.15) now implying an under-prediction of £118.

Ten observations now violated more than one diagnostic criterion. The previous influential observation 4060 was no longer present. New ones emerged such as case 5118 under-predicting income by £111.

5.3 Cross Validation

A slightly higher in-sample MSE of 1266.92 (from 1221.56) indicates a looser fit to the training data, but a lower out-of-sample MSE of 1296.10 (from 1336.25) suggests improved generalisation. The narrower gap between them also indicates reduced overfitting.

6. Results for Complete Case

Predictor	Estimate	P-value	Signif.
(Intercept)	361.2899	< 2e−16	***
W1wrk1aMPPart time	6.0727	0.010533	*
W1wrk1aMPEconomically inactive/Unemployed/Other	−32.5047	0.138682	
W1wrk1aMPLooking after home/ family	−47.7817	0.004987	**
W1hiqualmumDegree	−10.9897	0.003269	**
W1hiqualmumHigher education (pre degree)	−5.3227	0.124758	
W1hiqualmumFurther education (16–18)	0.2367	0.943332	
W1hiqualmumNo qualification	−27.6637	1.47e−10	***
W1famtyp2Yes	−42.7757	< 2e−16	***
W1nssecfamIntermediate	−8.6250	0.045884	*
W1nssecfamSelf − employed	3.6846	0.436078	
W1nssecfamSupervisory/technical	−7.7383	0.043988	*
W1nssecfamRoutine/Semi − routine	−44.5800	< 2e−16	***
W1nssecfamNever Worked/unemployed	1.8020	0.962472	
W1ethgrpYPOther/Mixed	−81.9791	< 2e−16	***
W1ethgrpYPSouth Asian	−72.3710	< 2e−16	***
W1ethgrpYPBlack − African/Caribbean	−80.5245	< 2e−16	***
W1heposs9YPFairly likely	−2.6125	0.328304	
W1heposs9YPNot very likely	−5.6927	0.137366	
W1heposs9YPNot at all likely	−15.3142	0.006529	**
W1disabYPHas disability, schooling affected	−33.0315	5.75e−08	***
W1disabYPHas disability, schooling not affected	−5.0689	0.195832	
W2disc1YPYes	−11.9519	0.007133	**
W4CannTryYPYes	17.7592	1.01e−12	***
W4schatYP	0.9807	0.006787	**
W5EducYPNo	−8.1893	0.003112	**
W5Apprent1YPYes	−19.8813	0.000116	***
W6JobYPNo	−8.7383	0.000717	***
W6UnivYPYes	14.3680	7.45e−07	***
W6OwnchiDVYes	−26.9933	0.000690	***
W6DebtattYP	1.1092	0.002464	**
W8CMSEXMale	−30.8311	< 2e−16	***

Table 2: Complete Case Model Output

*** for p < 0.001, ** for p < 0.01, * for p < 0.05, . for p < 0.1, and no symbol for p ≥ 0.1*

6.1 RQ1 Discussion

Results were highly consistent with the full model with most predictors retaining the same significance and relative effect.

There were three notable changes. Maternal pre-degree education was no longer significant at predicting lower income. Families with supervisory/ technical occupations became significant, predicting £7.74 lower weekly income than professional backgrounds. Finally, household income quintiles was removed as a predictor entirely at the 10% level, suggesting no impact of household income on an individual's future earnings. However, we believe its removal shouldn't be interpreted as evidence of no effect, instead its 49% missing values reduced its power to detect one. Policymakers could further research this.

6.2 RQ2 Discussion

For RQ2, all key later-life predictors retained their significance and relative effect, suggesting observed relationships of post-16 choices and later income are stable.

7. Reflection & Limitations

7.1 Reflection

Our results suggest areas for policy consideration. Young people in Wave 1 ‘not likely at all’ to attend university were predicted to earn £17.93 less per week, while by Wave 6, those attending university earned £13.40 more. This highlights the long-term income impacts of educational aspiration and access. We believe further research is needed to understand why some students, particularly around Year 10 do not aspire to attend university.

An unexpected result showed individuals whose mother has general secondary education (14-16) earned more than those who held degrees. While this may reflect complex labour market or intergenerational factors, we propose policymakers to examine maternal educational pathways and employment patterns more closely.

Finally, managing missing data presented challenges and valuable insights. In *W1hiqualdad,* 25% missingness reflected ‘father not present’ rather than non-response so we retained *W1famtyp2* to prevent aliased errors. However, removing *W1hiqualdad* risked losing contextually important information. *W6NEETACT* (93%) and *W4Childck1YP* (99%), were initially excluded for simplicity. However, missingness reflected ‘Not Applicable’ potentially due to survey design or the individual’s age. When recoded as binary predictors, *W6is_NEET* became statistically significant (Appendix B). This reinforces that missingness could reflect meaningful outcomes rather than random loss. As such, we propose policymakers to carefully consider patterns of missingness in future surveys, potentially indicating overlooked subgroups that warrant targeted sampling efforts.

7.2 Limitations

While our predictors covered key domains, gaps remained. More variables measuring early financial hardship, mental health and community environment could’ve strengthened the models. Furthermore, unobserved variables such as resilience, motivation and social capital were not captured in the available data. These unmeasured influences may have impacted the accuracy of our model.

8. Lay report

8.1 Link To Creative Lay Report

https://drive.google.com/file/d/1C7-glGrmeNpb6-OOkV0IM72Lmq0M459G/view?usp=sharing

8.2. Hard Copy of Lay Report

8.2.1 Key Findings from RQ1

Family Background

Young people whose mothers didn’t have formal qualifications, or who grew up in single-parent households, tended to earn less by age 25. Similarly, young people from households in routine occupations tended to earn less compared to those whose parents held managerial, professional, supervisory, technical, or self-employed positions. Young people whose parents were not in paid work or were looking after the home/family earned significantly less than those whose parents were working. Interestingly, young people whose parents worked part-time were predicted to earn slightly more than those whose parents worked full-time.

Education

Students who were fairly/very likely to attend university and/or had a more positive attitude towards school tended to earn more by age 25. This was in comparison to those showing a less positive attitude towards school and/or were unlikely or unsure about attending university.

Ethnicity and Disabilities

Ethnic minority young people, especially those treated unfairly at school due to skin colour, tended to earn less at age 25 than their White peers. Reporting a disability, particularly when it had affected schooling, was also strongly linked to a reduced income. Finally, we found that females generally earned more than males.

One unexpected finding was that young people who had tried cannabis by the age of 15 were predicted to earn more by age 25.

8.2.2 Key findings from RQ2

Later-life predictors, gathered after age 16, revealed additional factors influencing income. Individuals who were not in school or college and instead pursued an apprenticeship tended to earn less. In contrast, attending university was linked to higher earnings.

Other important factors included early parenthood where having a child had a strong negative effect on income. Lastly, a more positive attitude towards debt was also linked to higher income.

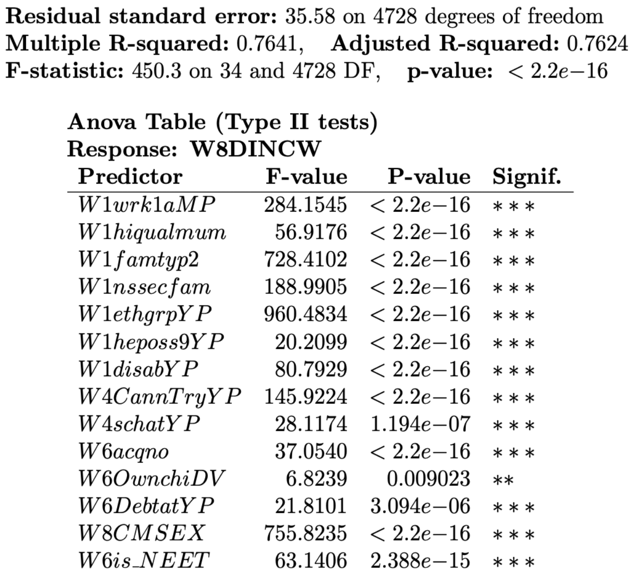
Appendix A: Merging and Levelling

Original Predictor Name	**New Predictor Name**	**Original Levels**	**New Level**	**Baseline Level**	**Justification**
W1GrssyrHH	hh_income_ quintile	Continuous Predictor	Bottom Quintile (<£13,057), 4th Quintile [£13,057–£19,204), 3rd Quintile [£19,204–£29,939), 2nd Quintile [£29,939–£49,400), Top Quintile (≥£49,400)	2nd Quintile [£29,939–£49,400)	Scatterplot looked like it had a positive correlation, turned it into a categorical variable to have better use in the model, more interpretable.
W1wrk1aMP	-	1,3 2,4 5 10 6,7,8,11,12 -91,-99	Full time Part time Unemployed Looking after home/family Economically inactive/Other Missing	Full time	Reducing paramters. Missing values were later dropped after fitting a few models and finding them insignificant.
W1hiqualmum	-	1,2 3,4,5,6 7,8,9,10,11,12,13 14,15,16,17,18 20 -999,-99,-94 -98	Degree Higher education (pre degree) Further education (16-18) General secondary education (14-16) No qualification Missing Mother not present		Reducing parameters based on intuition and similar boxplot structures.
W1nssecfam	-	1,2 3 4 5 6,7 8 -99,-94,-91,-999	Managerial/Professional Intermediate Self-employed Supervisory/technical Routine/Semi-routine Never Worked/unemployed Missing	Managerial/Professiona	Reducing levels for better statistical power. Merged rare levels based on boxplots.
W1ethgrpYP	-	1 2,8 3,4,5 6,7 -999, -92	White Other/Mixed South Asian Black - African/Caribbean Missing	White	Due to small sample sizes in some levels.
W1heposs9YP	-	-1	Don’t know		For interpretability reasons, we wanted to retain responses of ‘don’t know’ for whether the YP wanted to go to university, as we believe this is a valid answer and

					should be treated as a level.
W1disabYP	-	1 2 3 -97, -94, -1, -99	Has disability, schooling affected Has disability, schooling not affected No disability Missing	No disability	We wanted to retain missing as a level.
W2ghq12scr	GHQ_bin	0 1, 2, 3 4, 5, 6, 7, 8, 9, 10, 11, 12	No distress Mild distress High distress	No distress	Turned it into categorical given standard GHQ thresholds.
W4CannTryYP	-	1 2 -91, -99, -97, -96, -92, -1	Yes No Missing	No	Retained missing values, later dropped as they were rare and not significant.
W6NEETAct	W6is_NEET	1,2,3 -91 (Not applicable)	NEET not NEET	not NEET	This variable had 93% missing values from ‘not applicable’. We assumed that this was individuals ‘not in employment or education’ so made it binary.
W6acqno	-	1,2 3,4 5,6 7,8 9	Degree/ HE A/AS level GCSE Other/unknown No academic study aim	No academic study aim	Due to rare populated levels, merges reduced parameters. We made merges intuitively based on meaning.
W6OwnchiDV	-	1 2 -94, -92,-1	Yes No Missing	No	Wanted to retain missing values as a level. These were later dropped as they were rare and not significant in the model.

Table 1: Merging and Relevelling

Appendix B: Our best model



Residual standard error: 35.58 on 4728 degrees of freedom
Multiple R-squared: 0.7641, **Adjusted R-squared:** 0.7624
F-statistic: 450.3 on 34 and 4728 DF, **p-value:** $< 2.2e{-16}$

Anova Table (Type II tests)
Response: W8DINCW

Predictor	F-value	P-value	Signif.
$W1wrk1aMP$	284.1545	$< 2.2e{-16}$	$***$
$W1hiqualmum$	56.9176	$< 2.2e{-16}$	$***$
$W1famtyp2$	728.4102	$< 2.2e{-16}$	$***$
$W1nssecfam$	188.9905	$< 2.2e{-16}$	$***$
$W1ethgrpYP$	960.4834	$< 2.2e{-16}$	$***$
$W1heposs9YP$	20.2099	$< 2.2e{-16}$	$***$
$W1disabYP$	80.7929	$< 2.2e{-16}$	$***$
$W4CannTryYP$	145.9224	$< 2.2e{-16}$	$***$
$W4schatYP$	28.1174	$1.194e{-07}$	$***$
$W6acqno$	37.0540	$< 2.2e{-16}$	$***$
$W6OwnchiDV$	6.8239	0.009023	$**$
$W6DebtatYP$	21.8101	$3.094e{-06}$	$***$
$W8CMSEX$	755.8235	$< 2.2e{-16}$	$***$
$W6is\_NEET$	63.1406	$2.388e{-15}$	$***$

Table 2: Our Best Model

Upon reflection, W6NeetAct (whether the young person was in employment or education at 17 years old) consisted of 93% missing values (coded as negative numbers). Thus, we removed it from the main model. On further inspection, these missing values were accredited to '-91 = Not applicable'. We believed this variable was important for interpretability as we want to know whether being in education or employment impacts future income. Therefore, we converted this into a binary variable, with levels of 'not applicable' and 'yes'. Given the structure of the survey, those who did not identify as being in education or employment may have answered 'not applicable' so we felt that this assumption was justified. Before turning it into a binary variable, 1 observation of '-92 = refused' was removed, as we cannot interpret this to mean NEET/not NEET. After conversion, the boxplot showed a lower median for NEETs, suggesting those not in education or employment at age 17 had a lower income at 25. We fitted this model and found that W6is_NEET was in fact significant and improved all of the models diagnostics and residuals. However, these results remain speculative due to the assumptions made. Consequently, we suggest that future studies should research NEET status more thoroug