

Nature vs. Nurture: Does family background or external factors influence one's early intelligence scores?

Zi Wei Low | April 8, 2022

In this study, I will be utilizing a data study which consists of a nationally representative sample of children born in 2001. The data set includes innate characteristics, family characteristics, and two Bayley Scales of Infant Development (BSID) test scores, which can be used as proxies for an infant's intelligence test scores. The two infant intelligence tests are measured across two different points in time, where the first score is measured when most children are between 8 months to 12 months and the second score is measured when most children are around 24 months. Table I displays the summary statistics of the sample.

The goal of the current study is to test whether there are causal effects of family background inputs on a child's early intelligence scores. Family background inputs, or also known as external characteristics, is roughly defined as factors that are not innately attached to children and can be changed if the child is otherwise adopted. This includes factors such as socioeconomic status (SES), regional location, and number of siblings. A large motivation to understand the causal effect of nurture's influence is due to its high and far-reaching stakes in policy implications. Policies such as anti-poverty and pro-education laws are enacted contingent on the consensus that environment matters. These policies are implemented with the purpose of palliating the environmental differences and to not make those who are already disadvantaged worse off. However, is there truly a significant causal prediction of environmental inputs on intelligence?

To understand the effects of nurture and whether it causally predicts IQ scores, I would have to parse off the effects of family input from internal characteristics, such as race and gender. Internal characteristics are roughly defined as factors an individual is innately born and biologically endowed with. Individual internal characteristics and family characteristics are defined particularly in this study to allow for the clear distinction between genetic and environmental factors. Using OLS linear regression, I aim to first build a strong, parsimonious "nature" model with only internal characteristic regressors while controlling for the intelligence score in the first period. The walkthrough analysis for this portion will be fast tracked and summarized since it is not the meat of the study. However, a better breakdown of the analyses and procedures will be decomposed in the subsequent section. The model construction of the latter model mirrors that of the nature model. Following the construction of the "nature" model, family characteristics will be added to build the "nature and nurture" model. By observing the significance level and the coefficients of the family input variables, I would be able to determine whether nurture truly matters in influencing a child's early intelligence scores.

## **I. Predicting Intelligence Scores with Solely Individual Characteristics**

First, looking at the given dataset, I quickly identified the variables that can be categorized as internal characteristics. This includes birth weight, gender, race, multiple birth status, days premature, and age (when taking the test). Notice that race and birth weight are categorical variables; thus, to allow for further meaningful analysis, I will create dummy variables for each

option individually, and exclude one option from each characteristic to avoid the dummy variable trap. Singleton is also categorical, but I will only create one dummy, singleton, that captures whether an infant is part of a multiple baby birth, which takes value 0, or not, which takes value 1. Additionally, since birth weight stores ordinal categorical information, to allow for better interpretation I created a dummy variable from this indicator as an above or below premature birth weight dummy, which takes value 1 if the infant is born above average and 0 otherwise. This new dummy will allow for valid interpretation of whether the baby was born healthy and with adequate nutrition levels. The average baseline references Stanford's Children Health webpage 2500g value.

With these internal characteristic variables recoded, I've create the following linear regression model that includes all these characteristics as regressors and also controlled for the infant's previous intelligent scores:

$$\begin{aligned} score_2 = & \beta_0 + \beta_1 abv\_avg\_wgt + \beta_2 female + \beta_3 white.bi + \beta_4 black.bi + \beta_5 asian.bi \\ & + \beta_6 native\_hawaii.bi + \beta_7 indian.bi + \beta_8 multi.bi + \beta_9 singleton.bi \\ & + \beta_{10} days\_premature + \beta_{11} age + \beta_{12} score_1 \end{aligned}$$

The OLS results output is presented in table 1 as model 1.

From the output, I can clearly observe that there are several regressors that are not significant in predicting scores. Considering that I want a parsimonious model instead of only valuing fit, I will drop insignificant regressors and only keep regressors that are significant. Thus, to test whether these dropped regressors of p-value of less than 0.1 are truly insignificant, I conducted a anova test with the following joint significance hypothesis:

$$H0: \beta_{4(black.bi)} = \beta_{6(native\_hawaii.bi)} = \beta_{7(indian.bi)} = \beta_{8(multi.bi)} = 0$$

$$H1: \text{At least one of the } \beta_j \neq 0 \text{ where } j= 4, 6, 7, 8$$

The restricted model that captures the hypothesis above is presented in table 1 as model 2. Using the anova test to conduct model comparisons against model 1, I observe that I should not reject the more complex model 1,  $F(-4)=14.275$ ,  $p<0.001$ , relative to the parsimonious model. With this in mind, the anova test results also implies that I should reject the null hypothesis. The omittance of these variables in favor of more parsimony costs too much explanatory power.

Next I tried adding quadratic terms to see whether there are any nonlinear relationships between period 2 intelligence scores and the continuous variables period 2 age and period 1 score. Typically, I would utilize a scatterplot to see if there is an observable quadratic function, but due to the large amount of observations, it was difficult to visualize and observe with the naked eye. Thus, to check whether there exists a significant quadratic effect, I directly ran a linear regression instead with the addition of the "score 1" and "period 2 age" quadratic term first followed by the deletion of the "period 2 age" quadratic term (due to insignificance). The OLS results are presented in table 1 as model 3 and 4, respectively. Through an anova test for model 1 against both 3 or 4, I observe that I should reject both model 3,  $F(2)=5.922$ ,  $p<0.01$ , and model 4,  $F(1)=8.942$ ,  $p<0.01$ , and stick with model 1. Thus, from these comparisons, I will temporarily

conclude that the best parsimonious linear regression model 1 with small explanatory power, adjusted  $R^2=0.2146$ , is my “nature” model. Interestingly, all the regressors chosen for the initial “nature” model are significant.

## II. Integrating Family Background Input Variables

Once I’ve determined my “nature” model and the internal characteristic regressors that significantly predict scores in period 2, I began to build the comprehensive “nature and nurture” models that will include environmental characteristics that will help me answer my research questions.

Similar to the construction of the “nature” model, prior to including all the family characteristics into the linear regression model, I reviewed the input variables to rewrite categorical variables, such as region and SES quintile, into dummies for better analysis and interpretation. Similar to before, I will create dummy variables for each option individually, and exclude one option from each characteristic to avoid the dummy variable trap.

With the family background input variables recoded, I ran the following regression that incorporates the “nature” model concluded in the previous section along with the family input characteristics:

$$\begin{aligned} score_2 = & \beta_0 + \beta_1 abv\_avg\_wgt + \beta_2 female + \beta_3 white.bi + \beta_4 black.bi + \beta_5 asian.bi \\ & + \beta_6 native\_hawaii.bi + \beta_7 indian.bi + \beta_8 multi.bi + \beta_9 singletion.bi \\ & + \beta_{10} days\_premature + \beta_{11} age + \beta_{12} score_1 + \beta_{13} parenting\_score \\ & + \beta_{14} northeast.bi + \beta_{15} midwest.bi + \beta_{16} south.bi + \beta_{17} num\_sibings \\ & + \beta_{18} 2\_bio\_parents + \beta_{19} 1\_bio\_parent + \beta_{20} 1\_bio\_1\_step + \beta_{21} SESqu\_1.bi \\ & + \beta_{22} SESqu\_2.bi + \beta_{23} SESqu\_3.bi + \beta_{24} SESqu\_4.bi \end{aligned}$$

The OLS results output is presented in table 2 as model 1.

From the output, I can clearly observe that there are several environmental input regressors that are not significant in predicting scores. Considering that I want a parsimonious model instead of only valuing fit, I tried to drop the insignificant family structure input. To test whether the dropped regressors are truly insignificant, I conducted a anova test with the following joint significance hypothesis:

$$H0: \beta_{14(2bioparents.bi)} = \beta_{15(1bioparent.bi)} = \beta_{16(1bio1step.bi)} = 0$$

$$H1: \text{At least one of the } \beta_j \neq 0 \text{ where } j= 14, 15, 16$$

The restricted model that captures the hypothesis above is presented in table 2 as model 2. Using the anova test to conduct model comparisons against model 1, I observe that I should reject the more complex model 1,  $F(-3)=0.1821$ ,  $p=0.9086$ , in favor of the complex, unrestricted model. With this in mind, I tried checking whether the family structure is truly completely insignificant, and conducted a joint F-test, which only reinforced the results of the anova test,  $F(3)=0.1821$ ,  $p=0.9086$ .

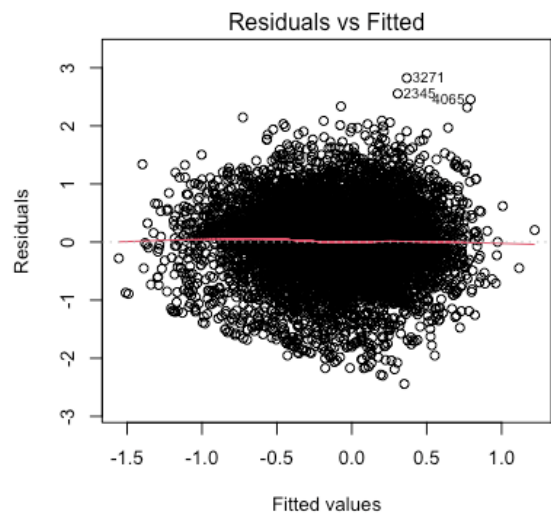
Next I tried adding quadratic terms to see whether there are any nonlinear relationships between period 2 intelligence scores and the parenting score and number of siblings. The anova test tests yielded to reject the more complex model (3),  $F(2)=1.4631$ ,  $p=0.2316$ , in favor of the more parsimonious model. Thus, I will stick with model 2 as the best “nature and nurture” model.

### ***II.1. Inspecting Multiple Linear Regression Assumptions***

Prior to interpreting the coefficients of the variables of interest, I tested the multiple linear regression assumptions to ensure that my estimators are unbiased and inferences are valid. Specifically, I tested the no perfect multicollinearity condition (MLR3), zero mean condition (4), homoskedasticity (5), and the notion that residuals are normally distributed (6).

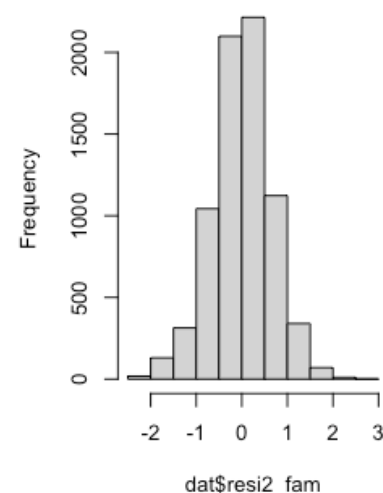
*MLR3 No perfect multicollinearity.* To check whether there are perfect multicollinearity issues, I regressed each individual regressor onto all other regressors to determine whether there is too much information overlap by observing the  $R^2$  of each regression output. As noted in the R file, there is no perfect multicollinearity issue and the variance inflation factor is small (at most 2).

*MLR4 Zero conditional mean.* To visually check whether there are potential violations of the zero mean condition, I indirectly inspected by checking whether the residuals have constant means by looking into the residuals vs fitted graph. With the naked eye, I can discern that the points plotted are roughly mirroring each other with no distinct pattern of too many sitting above or below the zero line. Thus, I would infer that there is minimal, if any, zero mean condition violation.



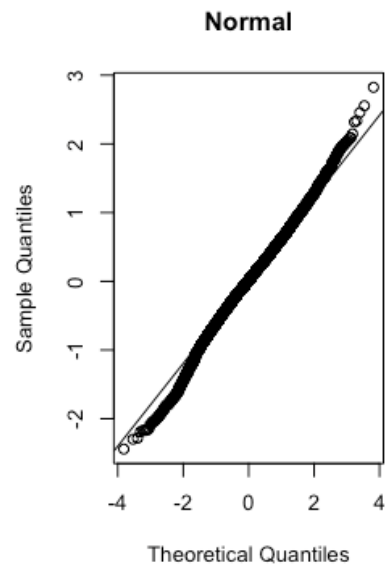
*MLR5 Homoskedastic.* To test whether there exists heteroskedasticity, I conducted a Breusch-Pagan Test, which has a null hypothesis that assumes homoskedasticity. The test yielded  $BP=93.48$  (21),  $p<0.0001$ . This implies that I should reject the notion that the variance is constant. Thus, to correct for this, I conducted a robust SE to my final “nature and nurture” regression to derive the linear regression output available in table 2 model 2 robust. Comparing the significance values of the robust SE and the non-robust SE regressions, I can observe that the significance values are exactly the same. The phenomenon is likely attributable to the fact that the data set is very large; hence, heteroskedasticity would not be as big of an issue.

**Histogram of dat\$resi2\_fam**



*MLR6 Residuals are normally distributed.* To check whether this assumption is true, I began by looking at whether the residuals hold a bell shape using a histogram (shown on right).

It was difficult to say for certain but it does hold a silhouette of a bell shaped distribution. To substantiate this observation, I conducted a Q-Q plot as another approach to check for normality. From the figure on the right, I would infer that there are hints of normality but the distribution is also heavy tailed. However, luckily, considering the sample size for this data set is very large, I am also able to infer that the residuals for this linear regression model enjoys an approximately normal distribution. With this assumption being satisfied, this implies that the probability distribution of the beta estimators are also normally distributed. This observation is crucial for hypothesis testing and interpretation since the probability distribution of OLS estimators is necessary to draw inferences about their true population values.



## II.II. Interpreting the “Nature and Nurture” Model

Seeing that the assumptions are not violated, I can safely move on to interpreting the estimated coefficients and interpret whether nurture or family inputs have an effect on children’s early intelligence controlling for nature characteristics and how they influence intelligence. From the final “nature and nurture” model (2), we can broadly classify the different environmental inputs as internal family characteristics, such as number of siblings and parenting score, and external family characteristics, such as region located and SES quintile level.

*Internal family characteristics (parenting score, number of siblings, family structure).*

*I. Parenting score.* The coefficient for parenting score is 0.010(SD=0.002). This implies that for every 1 unit increase in parenting score, which evaluates how good parent(s) act as a teacher for their child, I would observe, on average, a 0.01 increase in standard deviation of early intelligence for the child. This observation can potentially have strong policy implications such that a government can mandate all parents to take classes that teaches them how to become good parent teachers to potentially improve a child’s early intelligence scores and future learning outcomes.

*II. Number of Siblings.* The coefficient for parenting score is -0.048(SD=0.007). This implies that for every 1 extra sibling a child, an infant would test, on average, 0.048 standard deviation behind. Intuitively, we can rationalize this observation as parents who decide to have more children are usually children from poorer families and are most likely going to have less attention allocated to each individual child. From this observation, policy makers may potentially enact birthing limitation statutes, such as the one-child policy in China, to converge resources onto one child instead of deviating financial and attentional resources to many children. The intentions behind the policies would, again, be to improve children’s early intelligence scores and future learning outcomes.

*II. Family Structure.* Family structure is found to be insignificant in influencing a child’s early intelligence scores as shown from the model and joint significance hypothesis test. Put

differently, whether a child is raised by one biological parent, one step parent and one biological parent, or any other form, it does not negatively or positively influence an infant's intelligence scores.

*External family characteristics (region, SES quintile level).* Both of these characteristics are analyzed as dummies to help for interpretation.

*I. Region.* As we can see from the regression output (table 2 model 2 robust), there seems to be some variations where living in one region will hurt an infant's intelligence scores more than others. Namely, on average, those infants living in the midwest score 0.118 standard deviations lower than the scores of comparable infants in other regions. This observation calls for awareness of what is going on to account for these differences between the midwest and other regions. For example, are other regions outside the midwest implementing programs or policies giving themselves a comparative advantage?

*II. SES quintile levels.* As we can see from the regression output (table 2 model 2 robust), there seems to be a consistent variation on how lower levels of SES would hurt a child's intelligence scores more. Just from comparing the estimated values, each increase in SES quintile levels would improve 0.4-0.9 standard deviation of an average infant's intelligence scores. This has strong policy implications - it highlights how the society structure and welfare system is not supporting new families with adequate resources. This observation underscores the notion that children do not enlist in kindergarten or elementary on the same ground, and that a family's wealth background can potentially provide enough resources to give their child a significant comparative advantage over others.

### **III. Discussion**

From our "nature and nurture" model, I've identified that there are family input characteristics that influence early intelligence scores even when controlling for individual characteristics and previous intelligence scores. Seeing that nurture is a crucial aspect in influencing early intelligence scores, These results suggest and highlight the importance of putting emphasis on society's current welfare system in hopes of bettering posterity and future outlooks. Specifically, for example, as highlighted in the previous sections, there can be certain mandates for parent classes before the family has a baby to ensure that they have the appropriate knowledge of how to become a good teacher, or evaluations on whether certain regions are implementing good policies or programs that can be emulated to improve an infant's intelligence scores. These policies are inferred (from the linear regression model) to have substantial influence on a child's intelligence scores even as early as 2 year old. Considering the far-reaching effects of good education, these policy or program implementations may even ideally improve human productivity levels. The intuition behind this postulation stems from the notion that if these two year olds can score higher on early intelligence test, relative to their previous generation counterparts, as a result of these mandates, the two year olds can also begin building their human capital at a higher starting point too.

#### ***III.I. Limitations***

However, it is important to recognize the limitations of the current study in interpreting the effects of nurture. This study only uses linear regression which implies that many assumptions and limitations of a linear regression model will also be superimposed to this study. For example, by implementing linear regression, I am only looking at “on average” or the mean of the dependent variable, which is not exactly the complete description of the variable on its own. Other types of regression, such as quantile regression, may help palliate this limitation. Additionally, linear regressions are prone to underfitting the dependent variable since it assumes that there exists a linear relationship between the regressand and regressor; while in real life, the relationship between the two may not be exactly linear. Thus, the data will not fit the straight line assumption as well. As a result, the models will have lower fit which translates to less accuracy. Furthermore, there are concerns of omitted variable bias where there are some variables in the error term that may be correlated with the “nurture,” family input characteristic regressors and thus may bias the estimated coefficient seen in the model. An example would be the temperament of a child - a scale from easy to difficult. This omitted variable may have a negative relationship with parent score ( $\text{cov}(\text{temperament}, \text{parenting score}) < 0$ ) where the more difficult a child’s temperament is, the more likely the parent would be viewed as a bad teacher whereas if a child has an extremely easy temperament, the parent may be viewed as a better teacher. In regards to the estimated coefficient if I were able to regress temperament against intelligent scores in period 2, I predict that I will observe a negative estimated coefficient considering that children with more difficult temperament may find it more challenging to respond to the novelty of an intelligence test and may lack patience to solve harder questions. If both the covariance and the estimated coefficient is predicted to be negative, I can reasonably infer that this implies that there is a positive bias with the parenting score estimate.

### ***III. II. Future Directions***

Considering the results derived from the “nature and nurture” model, the natural question to ask is whether there are certain family input characteristics that hold more significance in some periods more than others. Put differently, are there certain environmental inputs that only have significant effects in particular critical periods and are not as influential in predicting early intelligence scores outside of the critical period? Additionally, are there certain family input characteristics that are not sensitive in significantly predicting a child’s early intelligence scores? These questions can be answered with more monitoring and data of early children’s intelligent scores. In the future, I hope to be able to check whether there are observable patterns in how different nurture inputs influence early children’s intelligence scores over time.

**Table 1. Ordinary Least Squares Regression Output (“Nature” Model Construction)**

	Model 1	Model 2	Model 3	Model 4
Input Variable	Estimate	Estimate	Estimate	Estimate
Intercept	-3.810(0.209)***	-3.742(0.209)***	4.857(5.096)	-3.812(0.209)***
abv_avg_wgt	-0.121(0.025)***	-0.125(0.025)***	-0.117(0.025)***	-0.117(0.025)***
female	0.226(0.016)***	0.227(0.016)***	0.225(0.016)***	0.225(0.016)***
white.bi	0.414(0.021)***	0.359(0.017)***	0.416(0.021)***	0.414(0.021)***
black.bi	0.043(0.027)	-	0.044(0.027) .	0.044(0.021) .
asian.bi	0.289(0.031)***	0.233(0.028)***	-0.290(0.031)***	0.288(0.031)***
native_hawaii.bi	-0.063(0.128)	-	-0.061(0.127)	-0.066(0.128)
indian.bi	-0.007(0.054)	-	-0.001(0.054)	-0.003(0.053)
multi.bi	0.242(0.033)	-	0.244(0.033)***	0.243(0.033)***
singleton.bi	0.138(0.023)***	0.139(0.023)***	0.136(0.023)***	0.137(0.023)***
days_premature	-0.007(0.0006)***	-0.007(0.0006)***	-0.007(0.0006)***	-0.007(0.0006)***
age_2	0.141(0.009)***	0.140(0.009)***	-0.574(0.420)	0.141(0.009)***
age_2^2	-	-	0.015(0.009) .	-
score_1	0.147(0.012)***	0.147(0.012)***	0.155(0.012)***	0.154(0.012)***
score_1^2	-	-	-0.026(0.009)**	-0.026(0.009)**
F-stat (df)	168.7(7350)***	244.1(7354)***	145.6(7348)***	156.6(7349)***
R^2	0.216	0.210	0.217	0.217
Adjusted R^2	0.215	0.209	0.216	0.216

Note: The standard errors are in parentheses; \*\*\*p<0.001, \*\*p<0.01, \*p<0.05, . p<0.1; the final model used as the “nature” model is highlighted in grey.



**Table 1. Ordinary Least Squares Regression Output (“Nature and Nurture” Model Construction)**

Input Variable	Model 1	Model 2		Model 3
	Estimate	Estimate	Estimate (Robust SE)	Estimate
Intercept	-3.945(0.234)***	-3.880(0.214)***	-3.880(0.227)***	-3.528(0.360)***
abv_avg_wgt	-0.110(0.024)***	-0.110(0.024)***	-0.110(0.025)***	-0.108(0.024)***
female	0.228(0.015)***	0.228(0.015)***	0.228(0.015)***	0.228(0.015)***
white.bi	0.313(0.023)***	0.313(0.023)***	0.313(0.023)***	0.312(0.023)***
black.bi	0.076(0.028)*	0.078(0.027)**	0.078(0.027)**	0.078(0.027)**
asian.bi	0.087(0.032)***	0.087(0.032)**	0.087(0.035)*	0.087(0.032)**
native_hawaii.bi	-0.052(0.125)	-0.051(0.124)	-0.051(0.133)	-0.055(0.125)
indian.bi	0.013(0.052)	0.012(0.052)	0.012(0.047)	0.013(0.052)
multi.bi	0.162(0.033)***	0.161(0.033)***	0.161(0.032)***	0.161(0.033)***
singleton.bi	0.127(0.024)***	0.127(0.024)***	0.127(0.024)***	0.131(0.024)***
days_premature	-0.007(0.001)***	-0.007(0.001)***	-0.007(0.001)***	-0.007(0.001)***
age_2	0.147(0.008)***	0.147(0.008)***	0.147(0.009)***	0.147(0.008)***
score_1	0.132(0.011)***	0.132(0.011)***	0.132(0.012)***	0.133(0.011)***
parenting score	0.009(0.002)***	0.009(0.002)***	0.009(0.002)***	-0.012(0.017)
parenting score^2	-	-	-	0.0003(0.0003)
# of siblings	-0.048(0.007)***	-0.048(0.007)***	-0.048(0.007)***	-0.031(0.016) .
# of siblings^2	-	-	-	-0.005(0.004)
2 bio_parents	0.067(0.097)	-	-	-
1 bio_parents	0.072(0.098)	-	-	-
1 bio 1 step	0.066(0.121)	-	-	-
northeast.bi	-0.061(0.026)*	-0.060(0.026)*	-0.060(0.028)*	-0.060(0.026)*
midwest.bi	-0.112(0.023)***	-0.118(0.023)***	-0.118(0.022)***	-0.118(0.023)***
south.bi	-0.089(0.021)***	-0.088(0.021)***	-0.088(0.021)***	-0.089(0.021)***
SESqu1.bi	-0.400(0.028)***	-0.399(0.027)***	-0.399(0.027)***	-0.398(0.027)***
SESqu2.bi	-0.348(0.026)***	-0.347(0.025)***	-0.347(0.025)***	-0.347(0.025)***
SESqu3.bi	-0.268(0.025)***	-0.267(0.024)***	-0.267(0.024)***	-0.267(0.024)***

SESqu4.bi	-0.171(0.024)***	-0.171(0.024)***	-0.171(0.024)***	-0.171(0.024)***
F-stat (df)	108.4(7338)	124(7341)***		113.3(7339)
R <sup>2</sup>	0.262	0.262		0.262
Adjusted R <sup>2</sup>	0.259	0.260		0.260

Note: The standard errors are in parentheses; \*\*\*p<0.001, \*\*p<0.01, \*p<0.05, . p<0.1; the final model used as the “nature and nurture” model is highlighted in grey.