Report

Zhiyi Jin

Part 1

A. Assumptions

Parental education can be a good predictor of children's educational achievement. Children may have better educational attainment if their parents have higher levels of education degrees. This correlation could be due to the income effect and motivation effect. Parents with excellent educational backgrounds are more likely to have a high level of wages to provide their children with good educational opportunities. Moreover, parents are role models to their children. If they have extraordinary education attainments, their children are more likely to be motivated to pursue similar achievements.

I don't think there will be significant differences in the influence of father's or mother's schooling on their children's educational attainment. However, I would like to assume that when considering the sex of the child, the same-sex parent can have a significant effect on the schooling attainment of the child because the model effect tends to be much bigger when it's the same sex.

B. Estimate a multinomial logistic regression model

Table: Multinomial logistic regression

	Children Educational Attainment(Ref = < High School)			
	High School	Junior College	Bachelor	Graduate
	(1)	(2)	(3)	(4)
FatherHigh School(Ref = < High School)	1.617***	2.001***	2.202***	1.998***
	(0.180)	(0.254)	(0.219)	(0.243)
FatherJunior College	1.115*	2.806***	2.458***	2.379***
	(0.623)	(0.665)	(0.646)	(0.679)
FatherBachelor	1.636***	2.188***	3.755***	3.511***
	(0.436)	(0.520)	(0.443)	(0.460)
FatherGraduate	2.735***	3.287***	4.845***	5.201***
	(1.018)	(1.084)	(1.019)	(1.023)
Constant	0.877***	-1.340***	-0.784***	-1.175***
	(0.097)	(0.180)	(0.146)	(0.169)
Akaike Inf. Crit.	5,480.379	5,480.379	5,480.379	5,480.379

Note: *p<0.1; **p<0.05;** p<0.01

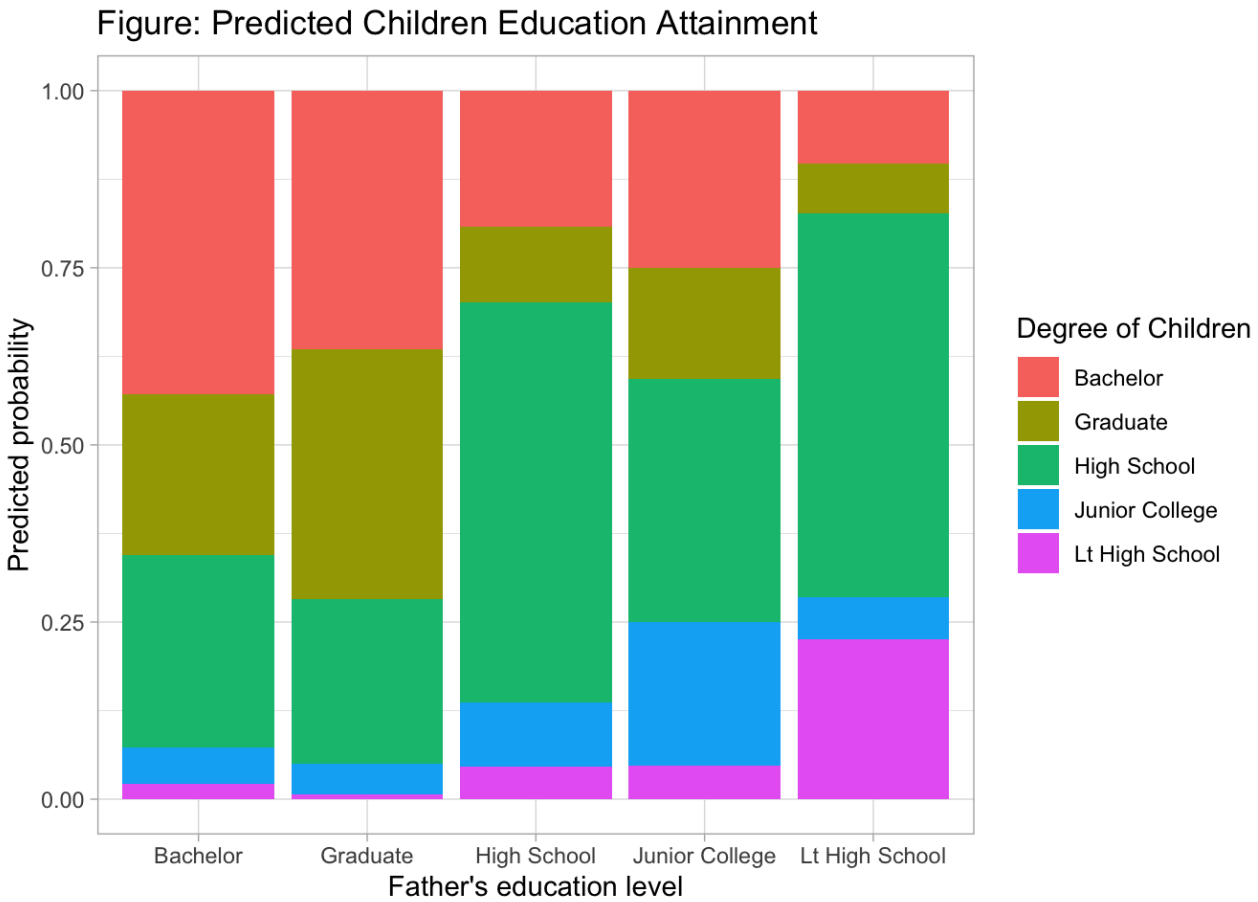
Data from General Social Survey

C. Interpret the model

It can be seen from the table that all coefficients are prominently significant and positive. This evidence confirms a positive influence of a father's educational background on children's educational attainment. The coefficient of a father's graduate degree on children's graduate achievement is exceptionally high. Children whose father has a graduate degree have exp(5.21)=183 times greater

odds of finishing a graduate education compared to those whose father has an education background lower than high school level. The other interesting coefficient is the one of the father's junior college background on children's high school attainment. It indicates that compared to children whose father has an education background lower than high school level, children whose father has a junior college background have exp(1.115)=3 times greater odds of completing high school, but this impact is relatively lower and less significant than others. Overall, the impact of the father's educational categories seems to be strongest when it's predicted for corresponding categories of children's education.

D. Produce a plot of predicted probabilities



The predicted probability plot does not fully support my former interpretation, but it does show that if the father's educational background is at high school, bachelor, or graduate level, children are more likely to complete education at the same level. However, even if the father has a junior college degree, the children could not be guaranteed to have the same level of degree. Instead, they are more likely to only complete a high school degree. For children whose father has less than a high school level degree, they are still very likely to complete high school education. This could be due to the widespread growth of high school education in this generation.

E. Estimate a new model with mother's education

I added the mother's educational background as an explanatory variable in model 2. And the result of model 1 and model 2 are summarised in the following table. After adding new variables, coefficients are still significant and positive, which implies the similar effect of mother's education on children's

educational attainment as the father's.

Table: Multinomial logistic regression

	Children Educational Attainment(Ref = < High School)							
	High School (1)	Junior College (2)	Bachelor (3)	Graduate (4)	High School (5)	Junior College (6)	Bachelor (7)	Graduate (8)
FatherHigh School(Ref = < High School)	1.617***	2.001***	2.202***	1.998***	0.921***	1.270***	1.289***	1.097***
	(0.180)	(0.254)	(0.219)	(0.243)	(0.225)	(0.315)	(0.269)	(0.294)
FatherJunior College	1.115*	2.806***	2.458***	2.379***	0.321	1.937***	1.312*	1.160
	(0.623)	(0.665)	(0.646)	(0.679)	(0.649)	(0.701)	(0.680)	(0.720)
FatherBachelor	1.636***	2.188***	3.755***	3.511***	1.167**	1.506**	2.964***	2.691***
	(0.436)	(0.520)	(0.443)	(0.460)	(0.561)	(0.656)	(0.572)	(0.589)
FatherGraduate	2.735***	3.287***	4.845***	5.201***	1.779*	2.037*	3.411***	3.550***
	(1.018)	(1.084)	(1.019)	(1.023)	(1.058)	(1.139)	(1.063)	(1.071)
MotherHigh School(Ref = < High School)					1.287***	1.109***	1.651***	1.607***
					(0.220)	(0.313)	(0.269)	(0.298)
MotherJunior College					13.387***	13.691***	14.015***	13.757***
					(0.193)	(0.304)	(0.235)	(0.287)
MotherBachelor					0.884*	1.229**	1.732***	1.596***
					(0.456)	(0.544)	(0.483)	(0.510)
MotherGraduate					1.768*	2.459**	2.716**	3.262***
					(1.048)	(1.107)	(1.059)	(1.064)
Constant	0.877***	-1.340***	-0.784***	-1.175***	0.556***	-1.606***	-1.278***	-1.599***
	(0.097)	(0.180)	(0.146)	(0.169)	(0.109)	(0.203)	(0.175)	(0.200)
Akaike Inf. Crit.	5,480.379	5,480.379	5,480.379	5,480.379	5,096.569	5,096.569	5,096.569	5,096.569

Note: *p<0.1;* ***p<0.05;*** p<0.01

Data from General Social Survey

To test whether the new model is sufficient, I use log-ratio tests and BIC statistics to compare the two models. The comparison shows that the likelihood ratio test prefers the second model in which the mother's education is included as an explanatory variable ($p < 0.001$). Also, the BIC decreases from model 1 to model 2, indicating that the improvement of the model does lead to a parsimonious improvement of model fit. This is evidence in favor of the role of mother's education over father's education, which does not support my original assumption.

Table: Fit statistics

				Likelihood Ratio Tests			
Model	Log likelihood	Parameters	BIC	vs. Model	Chi sq.	df	p value

				Likelihood Ratio Tests			
Model	Log likelihood	Parameters	BIC	vs. Model	Chi sq.	df	p value
1	-2720	20	5594	•	•	•	•
2	-2512	36	5299	1	416	16	<.001

Part 2

A. Assumptions

Both religion and marital status could be predictors for the number of sexual partners a person has had. In terms of religion, I guess some religious people would have fewer sex partners than non-religious people. This might be due to the sexual conservatism that is promoted by certain religions. Some religions also strengthen the commitment to family and long-term marriage. This could also lead to a conservative sexual attitude. The marital status could be an even more prominent predictor than religion. People who are not in a stable marriage are likely to have more sex partners than married people. It could be because this group of people does not have the obligation for one sex partner. Also, those who don't choose to marry are usually more open to sex life.

B. Estimate a binomial logistic regression model

Multinomial logistic regression

	Sexual Partners(Ref = < 5)
Protestant(Ref = None)	-1.078***
	(0.368)
Catholic	-0.652
	(0.407)
Jewish	-7.568
	(35.938)
Other	0.104
	(0.483)
Widowed(Ref = Married)	0.726
	(1.160)
Divorced	2.167***
	(0.657)
Separated	-7.677
	(100.426)
Never Married	2.722***
	(0.608)
Constant	-4.878***
	(0.612)
Akaike Inf. Crit.	406.109
Note:	*p<0.1;* ***p<0.05;*** p<0.01
	Data from General Social Survey

C. Potential separation problems

After fitting the model, I found that the estimates in the result are either strongly positive or insignificant at all, which indicates a sign of separation problem. Specifically, the coefficient of Jewish is extremely small at -7.568, implying that Jewish has 0.00052 times smaller odds of having over 5 sex partners when comparing to non-religious people. This extreme situation also happens in the coefficient of Separated. It indicates that separated people are particularly unlikely to have sex partners over 5. And both of these two coefficients have very large standard errors, 35.9 and 100.4 respectively. This problem might be caused by the small size of the sample.

D. Diagnose potential separation problems

To further diagnose the separation problem, I made two tables below. It can be seen that 33 Jewish people in the sample while none of them have sex partners over 5. Likewise, there are 53 separated people, while none of them have sex partners over 5. Considering the relatively smaller size of Jewish than other religions, I recode Jewish to Other. It makes sense to me that this group of people are the minority and does not necessarily determine the outcome, or have any testing values under this context of the study. For the separated category of the marital status variable, I aggregated it to the divorced. Note that among 126 widows, only 1 person has sex partners over 5, which is also very small. So, I aggregated it to the married. In the theoretical sense, people who initiatively choose to divorce or separate from a marriage could share certain similarities in terms of sexual attitude. Likewise, widows do not choose the end of a marriage by themselves. Their values on sexual behavior might be still similar to married people.

Table by Religion

	None	Protestant	Catholic	Jewish	Other
Patners < 5	376	807	387	33	92
Patners >= 5	23	12	9	0	6

Table by Marital

	Married	Widowed	Divorced	Separated	Never Married
Patners < 5	746	126	295	53	481
Patners >= 5	3	1	11	0	35

E. Re-estimate your model

Re-estimating the model, the coefficient of protestant is negative and significant. It implies that compared to having less than 5 sexual partners, protestant has 0.33 times smaller odds of having more than 5 sexual partners than non-religious people. Although the impact of other religion on the number of sex partners is not significant, the result still supports my original hypothesis to some extent that some religion may be a prominent predictor. In terms of marital status, both divorced and

never married people are more likely to have over 5 sexual partners than married people. This result confirms my assumption that people who are not in a marriage are likely to have more sex partners than married people.

Table: Multinomial logistic regression

	Sexual Partners(Ref = < 5)
Other Religion(Ref = None)	-0.137
	(0.479)
Protestant	-1.085***
	(0.368)
Catholic	-0.657
	(0.407)
Divorced(Ref = Married)	1.864***
	(0.589)
Never Married	2.603***
	(0.534)
Constant	-4.748***
	(0.540)
Akaike Inf. Crit.	406.588
Note:	*p<0.1;* ***p<0.05;*** p<0.01
	Data from General Social Survey