The Impact of Parental Separation on Academic Performance in High School

Joseph Fallait

ECON3306.01

5/12/2020

Abstract

This study focuses on the explanatory variables that determine an individual's likelihood of performing well in high school, and more specifically deals with the inclusion of an explanatory variable for parents' marital status. Using the ELS:2002 dataset, this study uses survey-reported and school-reported information in order to generate several explanatory variables that can be used to predict academic performance. Using an ordered logit regression, this study compares the magnitude of the coefficient on the dummy variable related to separation/divorce and finds that it has almost as large an average effect on cumulative GPA ceteris paribus as gender and race have, and is statistically significant at the 99% level. Furthermore, this effect does not appear to interact with race and gender, being uniform throughout these demographics.

The findings of this study are that parental marital status has a large and statistically significant negative effect on an individual's academic performance, and much of the existing literature attributes these negative effects not to a decline in actual academic ability, but psychological issues. Therefore, it would be be reasonable to experiment with optional counseling or tutoring intervention by schools to raise academic performance in these groups.

Table of Contents

Introduction

Problem Statement and Research Questions

This paper analyzes the explanatory power of parental separation and its effects on a student's cumulative high school GPA. This answer is explored via the use of the Education Longitudinal Study of 2002, or the ELS:2002, and its follow-up collection waves. The ELS:2002 is a national survey following students from the 10^{th} grade up to their postsecondary years and accrues information via student, parent, and teacher submitted surveys along with data submitted by school administrators.

The effects of separation and divorce on a student's academic performance is an interesting area of concern—in large part because it represents irredeemably inequitable elements in a student's conditions. While it is possible to influence aspects of education like school quality or educational attainment of parents, for the most part, government intervention is not very capable of directly altering the relationship of a student's parents. For this reason, it is worthwhile to analyze the effects of parents' marital status on their child's academic performance. If only just to potentially implement academic intervention in the cases of students who are in the middle of such marital statuses, analyzing these effects is productive.

In a sentence, this paper aims to evaluate the impact of parental separation or divorce on the academic performance of high schoolers. Some key objectives are to test the statistical significance of marital status, to determine significant interaction variables on the variable, and to narrow down the most effective model to regress cumulative GPA on marital status.

Literature Review

The first and maybe largest motivator for research into this topic is the growth of divorce rates over the last 50 years. Between 1960 and 2000, the number of divorces per 1000 adults in the US increased from around 3 to almost 9 (Gruber 800). Part of this, as Gruber notes, was caused by the adoption of unilateral divorce by many different states—enabling one spouse to terminate a marriage without the consent of the other. One of the reasons divorce legislation causes such a big splash in the political world is the sheer impact of a split in the marriage on children of divorce. Some of these academic effects are not even strictly intelligence or ability based, with some of the issues stemming from difficulty in coordinating events. Scheduling and having parents attend school activities like parent-teacher conferences or afterschool academic or social enrichment activities is very difficult in a single-parent family (Guidubaldi and Cleminshaw 40). Divorce isn't as simple as a likely drop in income—it has complex interactions with many different advantages students can no longer be afforded.

Charles Murray's exploration into poverty leads him to use NLSY data to sample and create what he calls a utopian sample—733 samples of siblings growing up in an environment without illegitimacy, early divorce, or poverty (Murray 339). The analysis in this paper relates the IQ of siblings in both this utopian subsample and the remaining data points to the percentage of the subjects, and it states that "Being smart is associated with very high levels of marriage among those with children, and being dumb is associated with very big drop-offs" (Murray 342). While not the most delicate way of explaining the concept, Murray's point is clear: nonmarital births strictly increase as the IQ of the individuals involved decreases. This compounds negative impacts on individuals born into those situations, leaving them with not only fewer resources from the get-go often lacking a parent but also maybe leaving them with less academic help than they might need. Simple economies of scale theory explain the economic benefits of being in a marriage. Expenses and savings being shared would necessarily slash costs across the board and leave individuals with more financial resources for their children (Amato and Maynard 118). In fact, the poverty rate of children living with married parents was 8.2%, compared to 31.2% of all other children (Amato and Maynard 128).

As it stands, "nearly half of the children born to married parents in this country go through a divorce experience before they are eighteen" (Clarke-Stewart and Brentano 106). Especially in the case of high schoolers, it is found that adolescents from divorced families generally perform worse and have lower educational aspirations than those from other families (Clarke-Stewart and Brentano 117). For the sake of this exploration, it is also useful to note that Clarke-Stewart and Brentano cite a study stating that high-school dropout rates were equivalent for children with parents divorced within five years of the study as those whose parents divorced more than five years earlier (Clarke-Stewart and Brentano 125). In other words, the year that the separation has occurred is not exceedingly relevant. Another variable of interest, gender, also affects collaboration with separation on the magnitude of the likelihood a student might have academic problems. Boys in households headed by a single mother are a whole "25 percentage points more likely to be suspended in the eighth grade than girls from these households" (Autor and Wasserman 46).

As a child of divorce, I do not have the same perspective on this problem as many of these authors, who seem to view divorce strictly as a problem that can and should be solved. I am of the vein of thought that separation will always exist, as it is human nature to make mistakes in choosing partners, and there are always problems in marriage that cannot be predicted. However, what economics can do is effectively gauge divorce's effects on academic performance and motivate greater intervention in aiding those who are affected.

Methods

The Education Longitudinal Study brings with it several different caveats in its use, but for the intents and purposes of this paper, it had a lot of key information that was, importantly, available for unrestricted use. I posit the way this was managed was by representing personal information not as numeric responses, but in ranges. For example, the variable relating to the cumulative GPA of the student throughout their entire high school career is not a distinct number but a range relating to a specific GPA level. Because this is the variable to be explained, it makes sense to leave this as an ordered logit or probit variable. But, because income and *freelunch* are essentially ranked ranges as well I selected the median value of the ranges and replaced the variable to be equal to those. For example, where *freelunch* == 1, 0-5% is the range of students on the *freelunch* plan, so *freelunch* == 1 was replaced with a value of 2.5%. While this makes the *freelunch* values a little bit inaccurate, the general trends remain the same and when regressed they become much more comprehensible than the rank format.

freelunch and *income* are rather strongly negatively correlated because individuals with higher incomes tend not to attend poorer schools, so the choice between these two variables (to avoid multicollinearity) ended up being a decision in which reported measure I trusted more. Because *freelunch* is reported by school administrators instead of parents, there is less motive to misreport data so I eventually chose it as the superior measure of school quality to include in the model. Parental ability was then explained through *pardegree,* which tracks whether the parent has received a college degree yet. It takes on the value 1 when the parent responding to the survey has a 2-year or 4-year degree, or a Master's or Ph.D., and 0 otherwise. Finally, the *separated* variable was created such that it equals 1 when the individual's parents are either separated or divorced and 0 when they are still married. While not pinpoint accurate due to the ranged or string datatypes of many of the explanatory variables, probit and logit-modeling using these altered median values is the best we can do with the unrestricted information provided by ELS and should still offer statistically significant insight into the importance of the separation variable.

The decision to use either an ordered probit/logit model also derives from this study's purpose to calculate probabilities of students falling into each grade level via the regression—so no other model would make sense in that case. The *cumGPA* variable is discrete, which further lends itself to either ordered probit or logit. The two models are rather similar, but because I am working with a healthy but not overly large number of observations or independent variables I opted for logit.

The method to analyze this data was therefore to alter the data points presented in string format into readable integers as explanatory variables. As discussed in the literature review, it was also necessary to attempt incorporating interaction terms for the binary variables white and female—as many of the journals on the topic speak about this disproportionate effect on performance between gender and race. These models were further refined by implementing new variables provided they were statistically significant, all the while testing the correlation of all the explanatory variables to assure there was no multicollinearity. As it turned out in this process, the separated interaction terms were insignificant when included alongside the original separated explanatory variable—rendering those interaction effects statistically insignificant. This is why they were ultimately excluded from the model. The final necessary explanatory factor was a variable controlling for student ability. Much of the standardized test data offered up by the ELS dataset was restricted, but as a proxy for ability, we do have access to a variable called F18113 (renamed to *writingabil*)—a teacher-reported measure of the students' writing ability. While this might miss situations where the individual's writing ability is far less or more developed than their other skills, *writingabil* has a .5836 correlation to cumGPA, so it is very usable as a proxy. It is extremely important to use some sort of proxy for ability in the regression to explain cumGPA because omitting it would severely limit the model's ability to predict an individual's academic performance.

The regression equation I finalized was:

$$cumGPA = \beta_0 + \beta_1 white + \beta_2 female + \beta_3 pardegree + \beta_4 separated + \beta_5 writingabil + \beta_6 freelunch + e$$

Results

Table 1 is the summarized statistics for all the explanatory variables used in this model. Four of the variables are dummy variables: *white*, *female*, *pardegree*, and *separated*. *writingabil*, by design, takes values between -2 and 2, and *freelunch* tracks the percentage of individuals the school reports as receiving free lunch, so it could take values between 0 and 100.

Table 1: SUMMARY STATISTICS OF EXPLANATORY VARIABLES

Variable	Obs	Mean	Std. Dev.	Min	Max
white	5,883	.6610573	.4733906	0	1
female	5,883	.5072242	.4999903	0	1
pardegree	5,883	.6068332	.4884949	0	1
separated	5,883	.1436342	.3507482	0	1
writingabil	5,883	.1952696	.9770652	-1.847	1.951
freelunch	5,883	20.89903	21.511	2.5	88

Of interest as well is the distribution of cumulative GPAs in the *cumGPA* variable that we are predicting. Table 2 displays this distribution, with the important information that 50% of the GPAs in this dataset are 3 and lower.

Table 2: DISTRIBUTION OF CUMGPA

GPA for all courses taken in the 9th - 12th grades	Freq.	Percent	Cum.
0.00 - 1.00	16	0.27	0.27
1.01 - 1.50	97	1.65	1.92
1.51 - 2.00	462	7.85	9.77
2.01 - 2.50	960	16.32	26.09
2.51 - 3.00	1,410	23.97	50.06
3.01 - 3.50	1,511	25.68	75.74
3.51 - 4.00	1,427	24.26	100.00
Total	5,883	100.00	

The last key table to look at is Table 3, which displays the correlation between every explanatory variable used in the model. Looking at this, none of the variables are strongly correlated, with the strongest correlation being a negative correlation of -.2714 between *freelunch* and *white*, which makes sense because the white individuals in this sample report higher incomes than non-white.

Table 3: Correlation of Explanatory Variables

	white	female	pardeg~e	separa~d	writin~l	freelu~h
white	1.0000					
female	-0.0069	1.0000				
pardegree	0.0721	-0.0138	1.0000			
separated	-0.0641	-0.0025	0.0240	1.0000		
writingabil	0.1095	0.1534	0.1111	-0.0686	1.0000	
freelunch	-0.2714	0.0383	-0.1684	0.0761	-0.1140	1.0000

Table 4: Ordered Logit vs Ordered Probit

VARIABLES	(1) OrderedLogit	(2) OrderedProbit
white	0.537***	0.314***
	(0.0532)	(0.0307)
female	0.430***	0.253***
	(0.0489)	(0.0284)
pardegree	0.174***	0.112***
	(0.0503)	(0.0291)
separated	-0.399***	-0.229***
	(0.0695)	(0.0397)
writingabil	1.316***	0.754***
	(0.0299)	(0.0165)
freelunch	-0.00335***	-0.00205***
	(0.00118)	(0.000683)
/cut1	-6.036***	-2.962***
	(0.259)	(0.0960)
/cut2	-4.024***	-2.155***
	(0.115)	(0.0561)
/cut3	-2.138***	-1.207***
	(0.0780)	(0.0437)
/cut4	-0.651***	-0.374***
	(0.0709)	(0.0411)
/cut5	0.787***	0.467***
	(0.0708)	(0.0412)
/cut6	2.368***	1.394***
	(0.0762)	(0.0434)
Observations	5,883	5,883
Pseudo R2	0.147	0.146

Standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

As evidenced by Table 4, the ordered logit and probit models perform very similarly in this dataset, with nearly identical Pseudo R2's and every single explanatory variable and cutoff value being statistically significant at the 99% level. Likewise, we can tell that *separated* has a large effect on *cumGPA*, as a change to the *separated* status has nearly the same impact as a change in the *female* dummy variable. This confirms suspicions that parents' marital status has a distinct and statistically significant large negative impact on an individual's *cumGPA*. Furthermore, Table 5 demonstrates that the predicted probabilities of this ordered logit model are very similar to the actual distribution of *cumGPA*. Note that GPAs below a 3 still consist of about 50% of the data points at the means of these predicted probabilities. The individual predicted probabilities are not identical, but typically only differ by a percent or so.

Table 5: Ordered Logit Predicted Probabilities

Variable	Obs	Mean	Std. Dev.	Min	Max
0.00 - 1.00	5,883	.0037326	.0099774	1.63e-07	.1233023
1.01 - 1.50	5,883	.0175466	.0319806	8.36e-06	.239208
1.51 - 2.00	5,883	.0774491	.0885883	.0003915	.3643069
2.01 - 2.50	5,883	.1586545	.1040573	.0054699	.3229238
2.51 - 3.00	5,883	.2381757	.0823943	.040751	.3259995
3.01 - 3.50	5,883	.2623009	.0916166	.0109035	.3568873
3.51 - 4.00	5,883	.2421407	.2081207	.0006946	.7739248

Discussion

The ordered logit model demonstrated that the separation of parents has a statistically significant negative impact on the probability that an individual would have a high cumulative GPA. The model itself is capable of explaining much of the variation because its predicted breakdown of the actual grades was very similar to the actual distribution of the *cumGPA* variable. What this paper complicates, however, were some of the ideas presented in the literature existing on the topic already.

A number of the papers written on this topic seem to assert that specific demographics suffer more from divorce than others do, but because no models using the interaction terms *female*separation* or *white*separation* were statistically significant, I needed to discard those variables in the regression. However, there were nuances that the data provided by ELS did not allow this regression to account for. For example, Autor and Wasserman attribute some of the negative interaction between gender and divorce to the fact that boys raised by single mothers are more likely to act out than girls in the same scenario (Autor and Wasserman 46). This would be interesting to examine, but datasets generally restrict personal identifying information—which would make it difficult to find one that freely offers information on which parent has custody and the nature of the parents' post-separation relationships. These interactions might be significant working with another dataset, but with what we have available we cannot say with any certainty at all that different demographics suffer more from divorce than others.

The *separation* variable I defined was also a potentially unorthodox interpretation of divorce, but I think it makes sense in the context of the dataset. Some individuals separate and never formally divorce, but the effects should logically be the same as divorce in these situations, and what we are interested in is the effects of these altered family structures—not their exact logistics.

Overall, the conclusions made by this study with the available data was that separation has a large negative effect on individuals' high school performance, almost to the extent of gender's effect on GPA, and barely lower than race's effect. Divorce has clear economic costs on the nation as a whole that are much more widespread than even a couple decades ago, so counseling or even academic intervention for individuals' pre and post-divorce could have sizable positive effects on much of the population (Guidubaldi and Cleminshaw 40). Whether these interventions should be prioritized in the current socio-economic climate of the United

States is another question entirely, as racial and financial inequalities are very much widespread and deserving of intervention themselves. What we can say with certainty based on this study is that academic performance is worsened by divorce across the board, and based off the data available from ELS:2002, it affects all demographics similarly.

Recommendations

This study could be expanded by implementing a test group that receives this sort of academic or psychiatric intervention. My recommendation would be that high schools reach out to students whose parents report separation or divorce to offer optional academic or psychiatric help. The data found that these negative divorce effects on academic performance were generally uniform throughout demographics, and as mentioned in the literature review these effects of divorce do not tend to diminish over time. As a result, the data supports equally offering this counseling to anyone who has experienced a divorce effect. However, because the effects of institutionalized racism are persisting and have an even larger negative coefficient on cumulative GPA in this model, it would make more sense for legislation to prioritize intervening in these racial inequalities.

Because the study in this paper is focused more on identifying and confirming the existence of the problem, the next logical step would be to intervene and determine both the costs and benefits of doing so. School funding has been in a constant state of decline especially in the public sector, so these expanded counseling options would not normally be implemented. The best way to make a case for these interventions is to demonstrate that their benefits outweigh their costs, which would be the focus of this follow-up study.

References

Amato, Paul R., and Rebecca A. Maynard. “Decreasing Nonmarital Births and Strengthening Marriage to Reduce Poverty.” *The Future of Children*, vol. 17, no. 2, 2007, pp. 117–141., doi:10.1353/foc.2007.0012.

Autor, David, and Melanie Wasserman. “Wayward Sons - The Emerging Gender Gap in Labor Markets and Education.” *Third Way*, 20 Mar. 2013.

Brand, Jennie, et al. “Why Does Parental Divorce Lower Children’s Educational Attainment? A Causal Mediation Analysis.” *Sociological Science*, vol. 6, 2019, pp. 264–292., doi:10.15195/v6.a11.

Clarke-Stewart, Alison, and Cornelia Brentano. *Divorce Causes and Consequences*. Yale University Press, 2006.

Gruber, Jonathan. “Is Making Divorce Easier Bad for Children? The Long Run Implications of Unilateral Divorce.” 2000, doi:10.3386/w7968.

Guidubaldi, John, and Helen Cleminshaw. “Divorce, Family Health, and Child Adjustment.” *Family Relations*, vol. 34, no. 1, 1985, p. 35., doi:10.2307/583755.

Murray, Charles. “IQ and Income Inequality in a Sample of Sibling Pairs from Advantaged Family Backgrounds.” *American Economic Review*, vol. 92, no. 2, 2002, pp. 339–343., doi:10.1257/000282802320191570.

Sander, William. “Unobserved Variables and Marital Status The Schooling Connection.” *Journal of Population Economics*, vol. 5, no. 3, 1992, doi:10.1007/bf00172094.