Impact of Absent Father on Child's Education

Fragile Families & Child Wellbeing Study
PRINCETON | COLUMBIA

Aarshay, Bindia, Keerti

What is The “Fragile” Families & Wellbeing Study?

- Birth cohort study of ~5000 American families in urban areas over 15 years
  - Collected in 5 waves [1998-2015]
    - As per child’s age - birth, 1st year, 3rd year, 5th year and 9th year

- Addresses 3 areas of great interest to policy makers and community leaders
  - Nonmarital childbearing, welfare reform, and the role of fathers

- Data on unwed fathers is more complete than those from previous surveys
  - Interviewed at least 75% of the unwed fathers
  - Data on the missing fathers completed from mother’s file

Why did we pick this data?

Lets Not Go There!!

- **12943** survey questions over **15 years**
- **4242** children tracked
- Cryptic feature codings
- **~25** files with documentation!
- **~80%** of the time on knowing data!

Children born to unmarried parents have been shown to be more vulnerable to adverse outcomes

- Nearly a third of all children born in the United States today are born to unmarried parents with higher among poor and minority populations
- Some of the existing research is around the effect of:
  - Father involvement on child's behavior
  - Parents' On/Off Relationship on father's involvement
  - Upbringing in disadvantaged families leading to low college performance
- The link between father involvement and child's school education outcomes is not common; thus our primary focus

Our aim is to assess the impact of biological father's involvement upto 9 years on GPA of child at the age of 15

- Interviews questions covered the following variables:
  - Relationships, Parenting behaviour, Employment status, Neighbourho
    Characteristics, Home Environment, Incarceration Histories, Race/Religion etc.

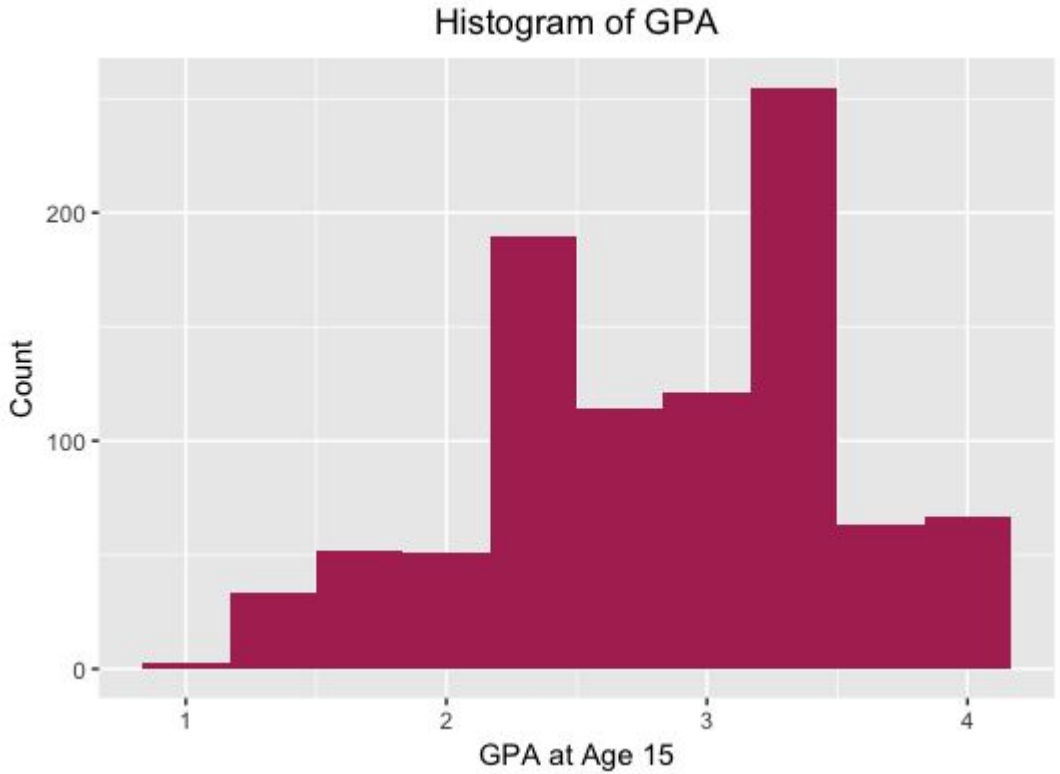
- ~600 constructed variables
  - Created a data dictionary for these variables (66 cells)
  - Can be used to map the variable meaning in the future research

We chose GPA as a measure of educational attainment at the age of 15

Available Outcomes:

- GPA
- Grit
- Material Hardship
- Eviction
- Being laid off

We picked GPA!



FEATURE ENGINEERING

We started with some primary data cleaning

- Considering Ids where GPA is present (~54%)
- Considered only the mothers present in all waves (~82%)
- Drop children under father’s custody (~0.2%)
- Missing value imputation

We have engineered 5 features to access the impact of biological father's involvement upto 9 years on GPA of child at the age of 15

1. Father's Involvement
2. Father's Presence
3. Cohabiting with biological father
4. Cohabiting with any partner
5. Number of partner's changed

Presence of biological father was determined based on the time lived in the same house.

- Full presence - father lives in the same house
- Partial presence - father has seen the child in the last 2 years
- No presence or None - father has not seen the child in the last 2 years

Full = 5
Partial = 2.5
None = 0

Categories

Year 1	Year 3	Year 5	Year 9
x			
0.1	0.2	0.3	0.4

Interview years

High (>=3.5)
Medium (1.75-2.5)
Low (<=1.75)

Conclusion

Father's involvement with child was determined based on frequency of spending time with the child.

- Full involvement - 1+ hour everyday.
- Partial involvement - a few times a week or a few times a month
- No involvement or None - very little time or not at all

Full = 5
Partial = 2.5
None = 0

Categories

Year 1	Year 3	Year 5	Year 9
x			
0.1	0.2	0.3	0.4

Interview years

High (>=3.5)
Medium (1.75-2.5)
Low (<=1.75)

Conclusion

Presence of a fatherly figure is based on the mother living with any partner

Cohabiting with biological father

- Number of times mother answered “yes” to living with the biological father

Cohabiting with any partner

- Number of times mother answered “yes” to living with the father or any other partner

The number of romantic partners of the mother, apart from biological father, were also captured.

- Number of romantic partners the mother had
- Expected to have a negative connotation to the child's academic performance
- Data only available for the last two interviews (span of 6 years)

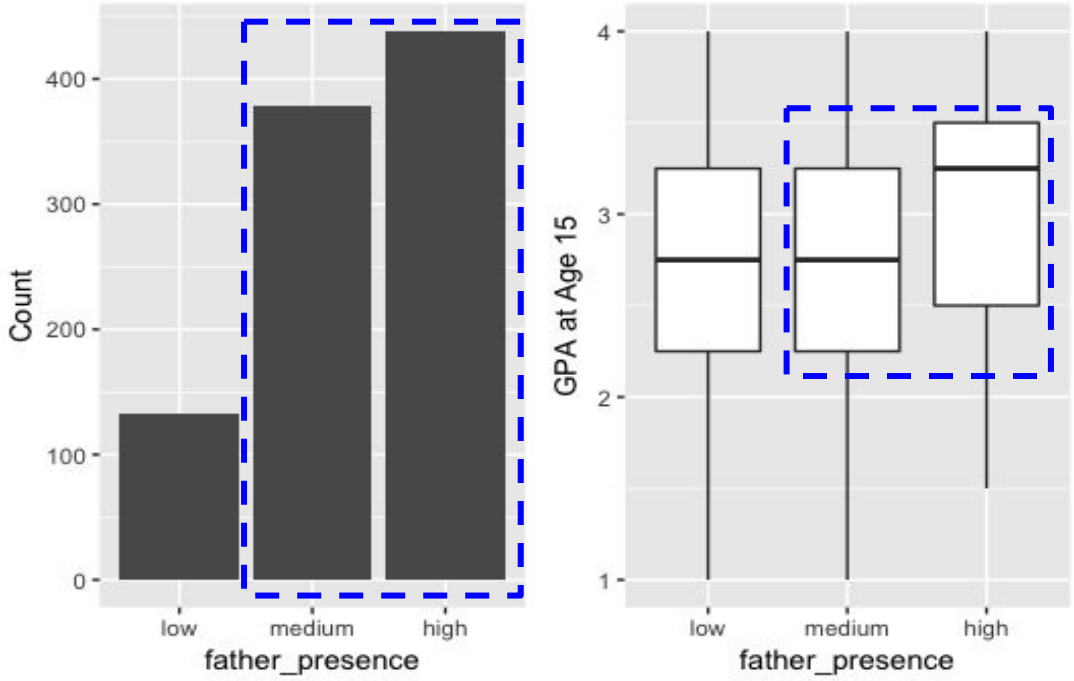
We had some control variables to account for other factors/confounders.

- **Parent's education (Integer)**
  - Both father's and mother's education
- **CIDI (Integer)**
  - Cases of depression, anxiety, alcohol and drug abuse involving mothers
- **Household Income (log scale)**
  - Average of the income throughout the interview years
- **Cases of Spanking (Integer)**
  - Taken from the child's interview at age 9

EXPLORATORY ANALYSIS

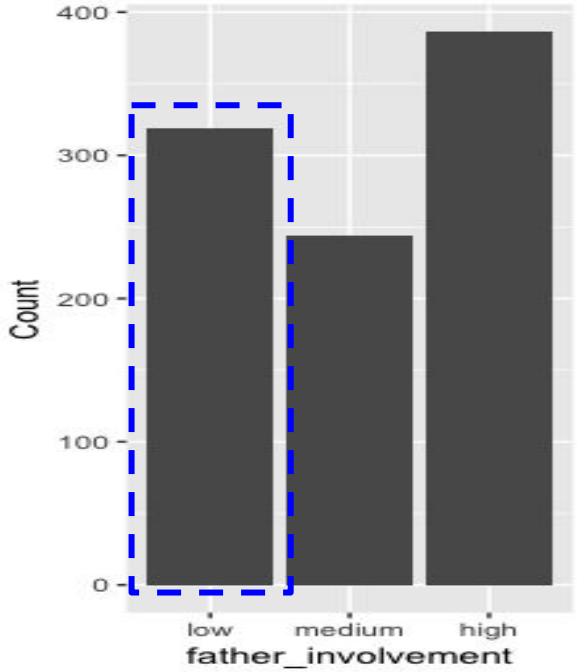
1. Univariate Analysis
2. Interaction Modeling Using OLS

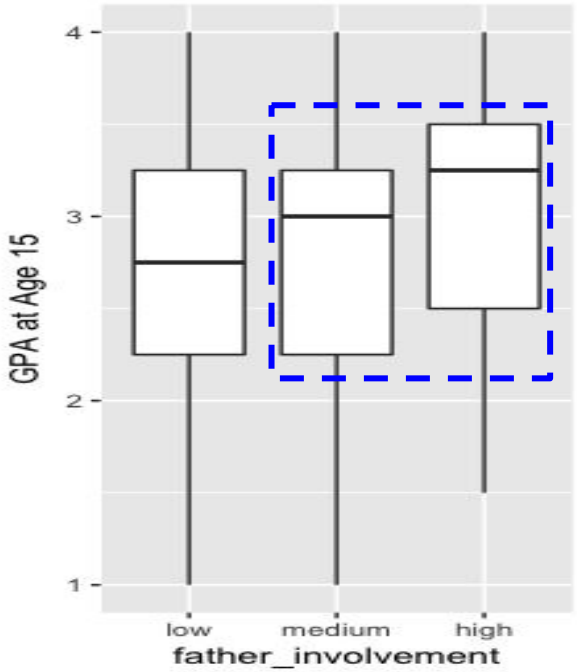
Children who experience high presence of biological father have higher GPA on average



3-way ANOVA also shows significant results

But children whose fathers were even mildly involved have higher GPA on average

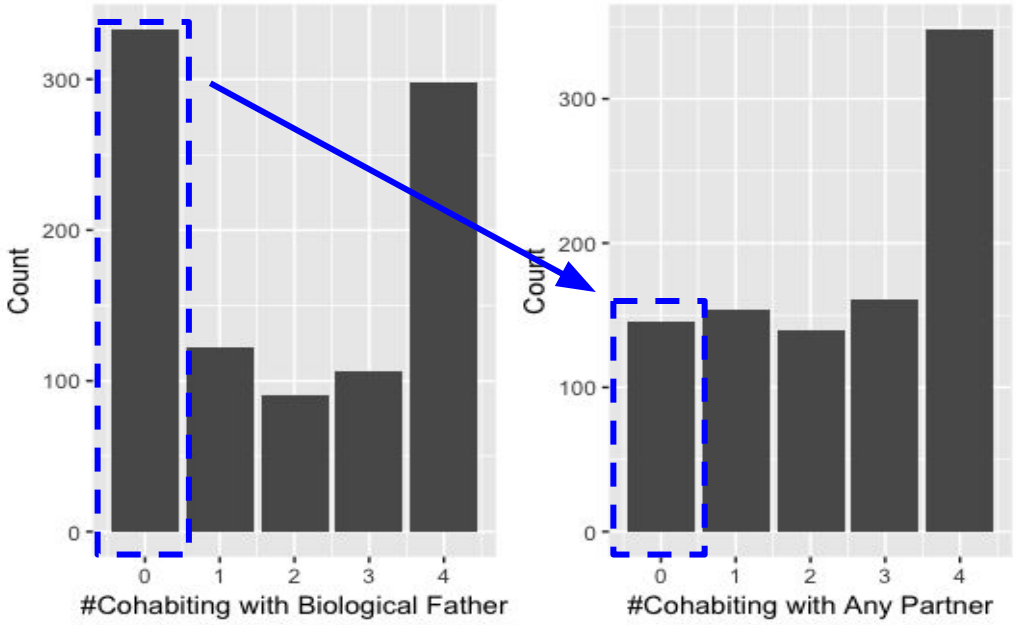




3-way ANOVA also shows significant results

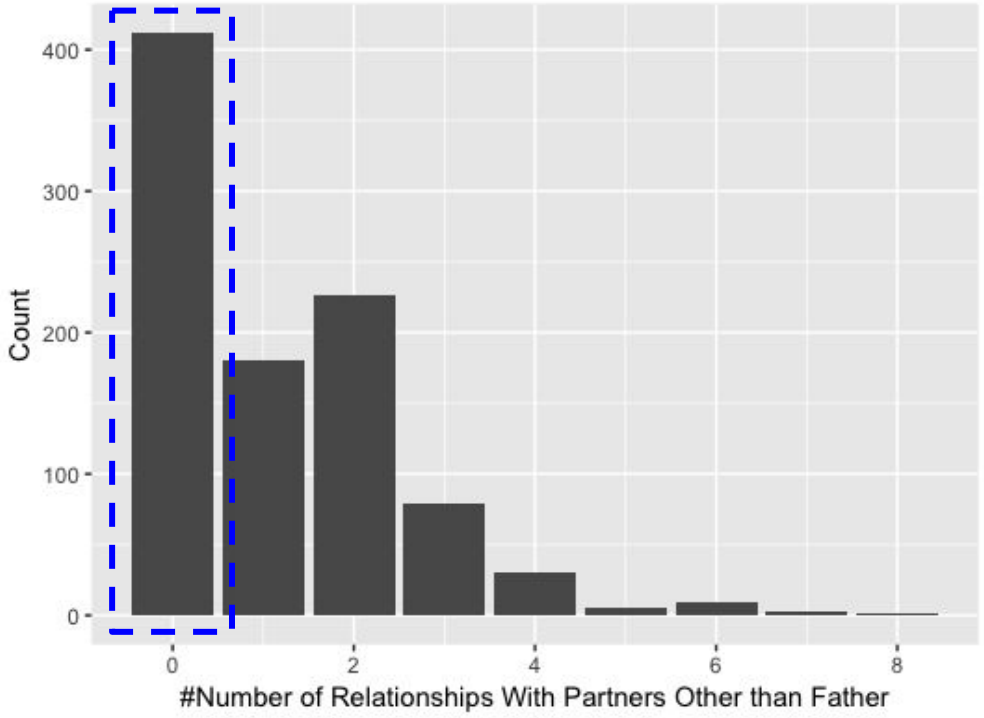
Cohabitation with biological or even other partners show positive correlations with GPA

Correlation Coeff: **0.24** **0.22**



More than half of mothers had non-marital relationships, which are negative correlated with GPA

Corr = **-0.18**



REGRESSION AS AN EXPLORATORY TOOL

Cohabiting biological fathers seem to have a significant impact on GPA

	Dependent variable:	
	gpa	
	(1)	(2)
Constant	2.714	2.707
	p = 0.000***	p = 0.000***
num_cohab_biof		0.083
		p = 0.006***
father_presencemedium	0.037	-0.008
	p = 0.576	p = 0.903
father_presencehigh	0.328	0.040
	p = 0.00000***	p = 0.743
Observations	949	949
R2	0.049	0.057
Note:	*p<0.1; **p<0.05; ***p<0.01	

But biological father's high involvement with the child is more important even if not living together.

	Dependent variable:		
	gpa		
	(1)	(2)	(3)
Constant	2.707	2.704	2.678
	p = 0.000***	p = 0.000***	p = 0.000***
num_cohab_biof	0.083	0.063	0.058
	p = 0.006***	p = 0.043**	p = 0.006***
father_presencemedium	-0.008	-0.046	
	p = 0.903	p = 0.531	
father_presencehigh	0.040	-0.058	
	p = 0.743	p = 0.653	
father_involvementmedium		0.086	0.071
		p = 0.182	p = 0.235
father_involvementhigh		0.195	0.181
		p = 0.024**	p = 0.026**
Observations	949	949	949
R2	0.057	0.062	0.062
Note:			*p<0.1; **p<0.05; ***p<0.01

Non-biological fathers don't appear to have a significant impact on child's GPA

	Dependent variable:	
	gpa	
	(1)	(2)
Constant	2.678	2.641
	p = 0.000***	p = 0.000***
num_cohab_biof	0.058	0.030
	p = 0.006***	p = 0.333
father_involvementmedium	0.071	0.081
	p = 0.235	p = 0.183
father_involvementhigh	0.181	0.198
	p = 0.026**	p = 0.017**
num_cohab_anyp		0.033
		p = 0.220
Observations	949	949
R2	0.062	0.063
Note:	*p<0.1; **p<0.05; ***p<0.01	

The number of non-marital relationships of mother has a slight negative impact

	Dependent variable:	
	gpa	
	(1)	(2)
Constant	2.641	2.680
	p = 0.000***	p = 0.000***
num_cohab_biof	0.030	0.020
	p = 0.333	p = 0.539
father_involvementmedium	0.081	0.081
	p = 0.183	p = 0.180
father_involvementhigh	0.198	0.190
	p = 0.017**	p = 0.022**
num_cohab_anyp	0.033	0.035
	p = 0.220	p = 0.195
num_partners		-0.019
		p = 0.376
Observations	949	949
R2	0.063	0.064
Note:	*p<0.1; **p<0.05; ***p<0.01	

Control Variables seem to overpower the explanatory variables, but the interpretation is still the same

	Dependent variable:	
	gpa	
	(1)	(2)
Constant	2.680	1.642
	p = 0.000***	p = 0.00000***
num_cohab_biof	0.020	-0.007
	p = 0.539	p = 0.830
father_involvementmedium	0.081	0.074
	p = 0.180	p = 0.202
father_involvementhigh	0.190	0.123
	p = 0.022**	p = 0.125
num_cohab_anyp	0.035	0.029
	p = 0.195	p = 0.276
num_partners	-0.019	-0.011
	p = 0.376	p = 0.586
mothers_education		0.078
		p = 0.007***
fathers_education		0.105
		p = 0.0002***
num_cidi_cases		-0.045
		p = 0.110
hh_income		0.074
		p = 0.042**
kid_punished		-0.012
		p = 0.267
Observations	949	949
R2	0.064	0.146
Note:	*p<0.1; **p<0.05; ***p<0.01	

In conclusion, we found 3 factors to correlate* well with high GPA outcomes

* No causal claims!!

In order of priority:

1. Biological father's high involvement (~95% conf)
2. Biological father's mild involvement (~80% conf)
3. Presence of a fatherly figure (~80% conf)

Challenges/Shortcomings

- 80% of the time spent on:
  - Understanding the questionnaire hierarchy
  - Data understanding and cleaning
  - Tracking the changes in question and response codes in-between waves
- We used information up to 9 years of age to predict outcomes at 15 yrs, missing out crucial
- Assumptions and proxy measures used to account for the factors taken in consideration

Future work

- Analyze the data geographically - get contract data (secured)
- Understanding the effect of social desirability bias and response bias
- Find natural experiments in the data and make causal claims
- Attempt fancy imputation techniques instead of dropping (though not sure if that would really help)
- Incorporate more factors and try to get more significant results

THANK YOU!!