The Inheritance of Inequality in Urban South India
Data Memo for SFI Workshop on "The Inheritance of Inequality in Pre-Modern Societies"
Mary K. Shenk
Updated 2.14.07

A. The Dataset

Background and Context

My data come from the city of Bangalore (recently changed to Bengaluru) in Karnataka State, South India. They were gathered during a year of fieldwork in 2001-2002. The data were gathered via a survey of adults aged 45-70 which collected detailed retrospective data on three generations of the respondents' families: (a) the respondent's parents and parents-in-law, (b) the respondent's self, siblings, spouse, and spouse's siblings, and (c) the respondents' children. Given the age of the respondents, the individuals covered in this sample span most of the 20th century. The data capture a period in which South India's economy has been progressively moving from a subsistence agricultural base with a limited cash economy to a market economy with a wage-labor base. As such it is not truly a 'pre-modern society' but should help capture the transition from the pre-modern to the modern.

In traditional India, only the sons of upper caste families would have received any significant amount of education, and most of this would have been of a religious nature. Over the course of the 20th century, however, formal secular education became increasingly common, beginning first among those from upper caste backgrounds but eventually spreading to lower castes as well. Similarly, in traditional India occupation would have been determined hereditarily by caste and family membership. However, such systems have been slowly breaking down for more than a century. In modern urban India perhaps only 10-20% of people still follow hereditary occupations, and most of these are merchants or skilled laborers whose occupations have obvious places in a modernizing market economy. The remainder of people, many of whom were farmers or agricultural laborers in rural India, have now adopted education-based or skills-based occupations. Thus in modern Bangalore education and income are usually the clearest signals of social status. While others exist, they are often far more difficult to measure systematically. For example, many families which once owned land have sold it or have allowed other family members who still live in the village to keep it. And while many families own real estate or businesses in Bangalore, assessing the value of such properties with survey data is problematic.

In India, wealth is traditionally divided equally among sons at the death of the father while daughters take their share of their parents' wealth via dowry at marriage. Traditional gender roles dictated that men do most of the labor in the fields and the market, while women do most of the domestic labor. In modern India men are still expected to have primary economic responsibility for their families, and virtually all men who do not own businesses themselves participate in some form of wage labor. While it is becoming more acceptable for women to work outside the home, the prevalence of working women varies a great deal by caste, social class, and father's occupation. Among educated professionals and poor manual laborers women often work whereas among business owners and those involved in low-end white collar jobs women are often expected to be housewives. The education of sons has become ubiquitous, with the level of education based on the child's expected (or hoped for) occupation, and average levels of male educational attainment have risen steadily throughout the 20th century. In the early part of the 20th century the education of daughters lagged well behind that of sons. However, a boom in the education of daughters came in the mid-late 20th century, thus educational rates for women have grown steadily in the past few decades and in contemporary Bangalore many daughters are nearly as well educated as their brothers. However, variation by caste and class is still common with groups that see housewife as the ideal occupation for women lagging behind groups where income-generating occupations are a possibility for daughters.

Data

I have wealth/status data in the form of both education and income for approximately 400 sets of married survey respondents. I also have education and income information on more than 1300 of these parents' children, including around 700 sons (almost all of whom have jobs) and around 600 daughters (of whom around 25% have jobs). Where income information is missing, I have occupation data which might be used along with income data from the rest of my sample to estimate income (though this is not done here). I also have limited data on non-income measures of wealth such as land ownership in acres, and whether the household has a telephone, a house, or a vehicle for all parents in the sample.

In addition to these, I also have education, occupation, and land and house ownership information for both sets of parents of the respondents in my sample. I also have education and occupation information for all of the respondents' siblings and siblings' spouses. I can use these data to measure the inheritance of educational attainment, and if sufficient historical information could be found to reasonably estimate income or wealth from these data I could also estimate the inheritance of income. Land and house ownership could also be used for subsamples of the population. My data on fertility for this grandparental generation is quite good.

Income and capital ownership data are current as of 2002 when most of the data were collected, which presents a concern in that parents' income is terminal, whereas the income of children can be expected to increase to some degree with the age of the child. However, I do have year of birth and/or age data for almost all cases in my sample which should allow some of this bias to be corrected for.

For the purposes of this memo, I will concentrate on the inheritance of education, income, and fertility between survey respondents and their children. The inheritance of education, fertility, land, and possibly income between the respondents' parents and the respondents themselves requires a separate data extraction and set of analytical concerns, so I will leave it until I have sufficient time to complete that work.

Variables

Bangalore is a highly urban area and virtually all men and many women in my sample were engaged in wage labor, and most respondents were able to provide reasonable income figures. The income variable reflects income in Rupees per month as given by the respondent. All incomes were given in 2002 Rupees. Education was collected as highest completed level and then converted into the appropriate number of years. Age and/or year of birth was collected from all respondents, as was fertility information when it was available (i.e. when a particular person had had children). Individuals with zero years of education were arbitrarily assigned one year of education before log conversions were done. This probably does not distort the results as the lowest level of education recorded was literacy, implying two years of education, and thus there were no people in the sample with one year of education who could be confused with those randomly assigned this amount. There were no families with zero income, so a similar simplifying assumption was not necessary.

Sample

Attempting to sample randomly within or across caste or class communities in Bangalore was not realistic. Unbiased enumeration would have been a complicated task well beyond the resources of my research project, and existing lists of people by residence or caste membership are rare, incomplete, biased, or otherwise problematic. Most importantly, without a personal introduction many potential respondents would refuse to participate in a survey or interview. Given these circumstances, recruiting respondents through personal contacts and referrals by snowball sampling was the only feasible way of collecting data. Sampling started through my social network as well as the social networks of three research assistants who came from different caste and class communities within Bangalore. My goal was to survey as broad a group of Bangaloreans as possible, and a comparison of the statistics of my sample population with the population of Bangalore in the Census of India indicates at least moderate success.

B. Estimates

Summary statistics for variables used in the analysis can be found in Table 1 below.

Table 1. Summary Statistics

Variable	Mean	Standard Deviation	Sample Size
Father's Age (years)	61.01438	9.842636	1252
Son's Age (years) [a]	31.95146	7.605609	618
Daughter's Age (years) [a]	30.60714	7.48091	532
Father's Education (years)	8.937984	5.282518	1290
Father's Education (log years)	1.828581	1.037304	1290
Parents' Average Education (years)	7.406977	4.677991	1290
Parents' Average Education (log years)	1.660751	0.967510	1290
Son's Education (years) [a]	12.04479	4.440015	614
Son's Education (log years) [a]	2.32998	0.7263381	614
Daughter's Education (years) [a]	11.30189	4.998903	530
Daughter's Education (log years) [a]	2.19553	0.8652714	530
Father's Income (2002 Rupees)	10872.6	25975.2	1222
Father's Income (log 2002 Rupees)	8.634187	1.101955	1222
Parents' Average Income (2002 Rupees)	10616.6	25841.0	1231
Parents' Average Income (log 2002 Rupees)	8.596908	1.113715	1231
Son's Income (2002 Rupees) [a]	23235.8	226366.0	517
Son's Income (log 2002 Rupees) [a]	8.751115	1.143361	517
Daughter's Income (2002 Rupees) [a]	7901.4	6498.4	141
Daughter's Income (log 2002 Rupees) [a]	8.588522	0.997337	141
Parents' Fertility (number of surviving children)	3.954824	1.660525	1306
Parents' Fertility (log number of surviving children)	1.287754	0.425756	1306
Child's Fertility (number of surviving children) [b]	2.042614	1.186775	352
Child's Fertility (log number of surviving children) [b]	0.672942	0.448776	331

[a] For sons and daughters 21 years of age or older.
[b] For children married before 1993.

Scatterplots of raw and logged data for all measures of wealth can be found in Appendix I.

Estimates of Beta

A summary of all estimates of Beta can be found in Table 2 below.

For the inheritance of education and income, all estimates of Beta were obtained using the regress command in Stata. Each analysis was performed for sons and daughters separately and clustered for family membership using the “robust cluster” command in Stata. Each analysis was limited to children age 21 and older in an attempt to exclude most of those who had not yet finished their educations or obtained jobs. This eliminated approximately 150 out of 1300 cases. Father’s education and father’s income were used as the primary predictors, but parents’ average education and income were also used to incorporate information on the mother’s characteristics. Parents’average variables were constructed by averaging the father’s and mother’s education or the father’s and mother’s income. Where there was no education or income for the mother then the father’s information was used and vice versa.

For the inheritance of fertility, all estimates of Beta were obtained using the regress command in Stata. Each analysis was clustered for family membership using the “robust cluster” command in Stata but was not stratified by child’s gender. Each analysis was limited to children married at least ten years prior to the time of the survey in 2002 so that fertility would be likely to be completed. This limited the sample to 333 children.

All analyses were controlled with child’s age, child’s age squared, father’s age, father’s age squared, and the interaction term (child’s age – 30) * logged father’s/parents’ wealth variable (i.e. education, income, fertility). For complete regression results see Appendix II.

Table 2. Summary of Results for Beta

Predictor	Beta	Standard Deviation	P value	Sample Size
On Son’s Education				
Father’s Education (Logged)	0.4046105	0.0482257	0.000	585
Parents’ Average Education (Logged)	0.4492196	0.0506347	0.000	585
On Daughter’s Education				
Father’s Education (Logged)	0.6106706	0.0597641	0.000	502
Parents’ Average Education (Logged)	0.6824934	0.0603646	0.000	502
On Son’s Income				
Father’s Income (Logged)	0.7764156	0.0500070	0.000	479
Parents’ Average Income (Logged)	0.7729013	0.0500098	0.000	480
On Daughter’s Income				
Father’s Income (Logged)	0.8204771	0.1137438	0.000	135
Parents’ Average Income (Logged)	0.7449284	0.1177571	0.000	136
On Child’s Fertility				
Parents Fertility (Logged)	0.3030856	0.0971886	0.002	313

Estimates of r

A summary of estimates of r (with controls) can be found in Table 3 below.

For education and income, all estimates of Beta were obtained using the regress command in Stata. Each analysis was performed for sons and daughters separately and clustered for family membership using the "robust cluster" command in Stata. Each analysis was also limited to children age 21 and older in an attempt to exclude most of those who had not yet finished their educations or obtained jobs. This eliminated approximately 150 out of 1300 cases. Father's education and father's income were used as the primary predictors, but parents' average education and income were also used to incorporate information on the mother's characteristics. Parents'average variables were constructed by averaging the father's and mother's education or the father's and mother's income. When there was no education or income for the mother then the father's information was used and vice versa.

For fertility, all estimates of Beta were obtained using the regress command in Stata. Each analysis was clustered for family membership using the "robust cluster" command in Stata but was not stratified by child's gender. Each analysis was also limited to children married at least ten years prior to the time of the survey in 2002 so that fertility would be likely to be completed. This limited the sample to 333 children.

Two sets of analyses were run, one with controls and one without controls. Controlled analyses were run with child's age, child's age squared, father's age, father's age squared, and the interaction term (child's age – 30) * father's/parents' wealth variable (education, income, or fertility). For complete regression results for both controlled and uncontrolled analyses see Appendix III. Only controlled estimates are presented in Table 3 below.

Table 3. Summary of Results for r

Predictor	R	Standard Deviation	P value	Sample Size
On Son's Education				
Father's Education	0.5519064	0.0385413	0.000	585
Parents' Average Education	0.6331615	0.0413057	0.000	585
On Daughter's Education				
Father's Education	0.6929836	0.0505511	0.000	502
Parents' Average Education	0.7968004	0.0527539	0.000	502
On Son's Income				
Father's Income	2.092783	0.5441475	0.000	479
Parents' Average Income	2.094237	0.5434221	0.000	480
On Daughter's Income				
Father's Income	0.4197135	0.0805855	0.000	135
Parents' Average Income	0.4259540	0.0806285	0.000	136
On Child's Fertility				
Parents' Fertility	0.1257296	0.0572460	0.029	333

C. Interpretation

At .40 for sons and .61 for daughters the Beta coefficients for education are moderately high as well as highly significant. Results of this sort are to be expected since educated parents are likely to (a) value education more than less educated parents, and (b) be more efficient at producing educated children through the quality of their interactions with children and their ability to model and/or help children with education-related knowledge. The coefficients for parents' education are higher than those for father's education, most likely because the parents' education term incorporates information about the mother's education. Mother's education is expected to have an influence on the level of her child's education since the mother is likely to spend substantial amounts of time interacting with her children and helping them with schoolwork, typically much more so than the father in Indian culture.

In the case of education, the Beta coefficients for daughters vis-à-vis their fathers/parents are higher than the coefficients for sons. There are a variety of possible reasons for this. First, sons will customarily be educated to the best of the parents' ability and will always be expected to have market-related occupations, whereas parents have greater flexibility in how well they educate daughters or whether they allow them to obtain market-related work. In less educated poor families, scarce resources may cause a greater degree of sex-biased investment as parents sacrifice education for their daughters in order to fund better opportunities for their sons. This would cause the intergenerational correlation to drop for sons. Conversely, in wealthier families the cost of education is not likely to be limiting and thus daughters can be educated to a level that closer approaches that of their fathers, causing an increase in the inheritance of education by daughters. Second, a similar effect might be caused if sons have begun to hit a ceiling in the utility of education on the job market. In this case, the difference between the educational level of the sons of educated parents and the educational level of sons of less well educated parents might converge and cause a muting in the intergenerational inheritance of education. If daughters have not yet begun to hit this ceiling, then parental influence would remain higher.

The Beta coefficients for the inheritance of income are high (.77 for sons, .81 for daughters) and highly statistically significant in all cases. Since estimates come from income data they are unlikely to be due to simple inheritance of wealth. Some fathers may allow sons to run and take the income from businesses that the father owns, but this would be a small minority of cases.

There are several potential reasons for these high coefficients on income. First, the high coefficients may be the result of the childrens' occupations. In India, parents often exert a strong influence on, if not total control over, the occupations their children adopt. It is also quite common for sons (or even daughters) to adopt an occupation similar to their father's. Even when this is not the case, however, wealthier parents value earning potential highly and often encourage their children to enter occupations associated with high incomes. High coefficients might also be the results of the educational and networking opportunities available to children because of their parents' wealth. Getting into selective, high status high schools, colleges, or professional schools often entails the giving of large 'donations', or unofficial admission fees (effectively bribes). While talented students often do not have to give donations, more average students can often only obtain admission into highly selective schools by paying such amounts. In this way wealthy parents can influence their child's quality of education, social connections, and likely employment opportunities in a way that parents of similarly-talented but poorer children cannot. Finally, wealthy parents can influence the income of their children by expending capital in ways that are likely to directly increase a child's earning potential after his or her education is complete. For example, a son educated as a doctor might expect to make a good income.

However, the doctor son of a wealthy man might be able to start his own clinic thereby increasing his income potential substantially.

In the case of income, the Beta coefficients for daughters vis-à-vis their fathers/parents are also higher than the coefficients for sons. A similar argument to that discussed for education holds in this case as sons are far more likely to receive the levels of unusual investment necessary to allow them to increase their income and social class in comparison to their parents. While daughters' marriages might also show such a social climbing trend, this cannot be seen without looking at wedding expenditures or husband's characteristics. Additionally, however, some social class or caste groups have norms about daughters working outside the house that may limit the incomes of daughters from middle-income households in comparison to those of higher-income households or even their presence in the analysis. Daughters of highly educated parents, if they work, are likely to take on high-status, high-pay occupations (such as medicine or engineering) similar to those of their fathers or, alternatively, not to work at all. Daughters of poor manual laborers are also likely to take on manual labor type jobs. Daughters of middle-ranking, middle income groups are much less likely to work outside the home at all but if they do are likely to take moderate-income jobs. Thus the category of 'daughters with incomes' is biased towards high and low rank occupations, with parents similarly situated. With few social-climbers and few daughters of middle income families in the analysis a strong relationship between the income of fathers/parents and daughters is to be expected.

The differences in the results for father's versus parents' education and income are not very great. This is most likely because the two sets of numbers are not terribly different. Averaging mother's with father's education does tend to lead to consistently smaller numbers than one gets using just father's education, but in many cases the differences are not very large. Averaging mother's with father's income usually leads to no change in the value at all because mothers usually do not earn any income and thus father's income is used instead. In the cases where mothers do have incomes those values are usually smaller than the father's income, and so those averages will be lower. However, this affects only 232 of 1306 children in the sample, which may help account for its modest effects on the estimates of Beta.

The high levels of status and wealth transmission that I find in my sample may be related to, or at least influenced by, two large-scale socioeconomic trends in mid-late 20th century Indian society. The first is a demographic transition to lower fertility; this may aid the process of status transmission by allowing family resources to be strategically concentrated in a relatively small (2-4) number of children rather than dispersed over a larger number of children. Second is the rapid rate of economic growth experienced in India over the last several decades of the 20th century, and especially noticeable in Bangalore. To the degree that this growth has meant a growth in jobs or other types of economic opportunities, this trend may have contributed to high rates of transmission by decreasing competition for resources both between families and between children within families.

As we might expect, parents' fertility does significantly predict child's fertility. On average fertility is linked to educational and occupational status and to the degree that parents and children resemble each other in these ways, we should expect similar fertility behavior. Other unobserved characteristics such as caste and religiosity should also increase our estimates of Beta.

However, the Beta coefficients on fertility are not as high or as statistically significant as those for education or income. Perhaps the most likely explanation for this is that India is in the middle of a demographic transition which is not taking place at the same rate in all families. While better educated families are likely to have a smaller than average number of children (usually 2), wealthier families especially in the business castes still often have larger families of 3 or 4 children. And while poorer laboring families still often have several offspring to aid them with their work, those with moderate amounts of education who aspire to middle class status often restrict their fertility in order that their children can achieve it. Thus despite prevailing fertility differentials in some groups, many people in urban India have adopted or are adopting the norm of two-three child families which might obscure the more noticeable differences in fertility that existed among the parents one generation ago as well as lead to lower estimates of Beta.

On the other hand, it is also possible that since this sample is truncated using the relatively recent year of marriage of 1993 some portion of the fertility I am measuring is incomplete. I expect that I will get more definitive results for this variable when I compare the parental generation in this sample to their own parents in the next phase of my work.

D. Next Steps

The next step for me will be to extract the data and construct the variables I will need to run intergenerational wealth comparisons on the grandparental versus the parental generation in my dataset. With this first generation dataset I will be able to examine at least three wealth proxies: (a) education for the entire sample (N = approx. 3000 parent-child pairs in nearly 800 families), (b) land for the portion of the sample which owns land in both generations (N = approx. 200), and (c) income if I can find a mutually satisfactory way to impute income from occupation (N = approx. 1500 father-son pairs).

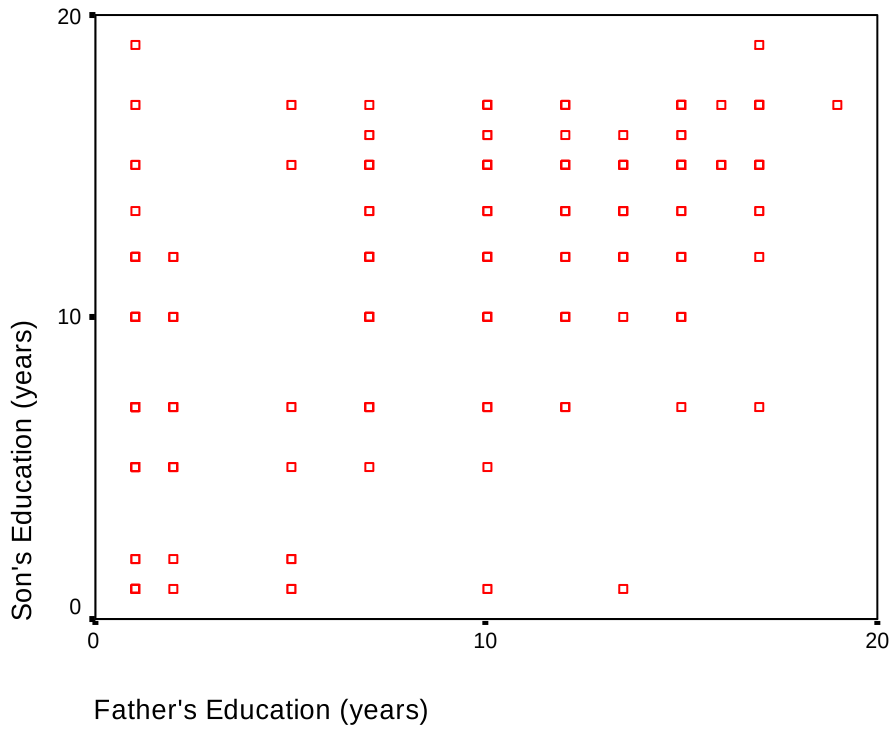
As for general next steps, as a group I think that we ought to consider the following questions and topics:

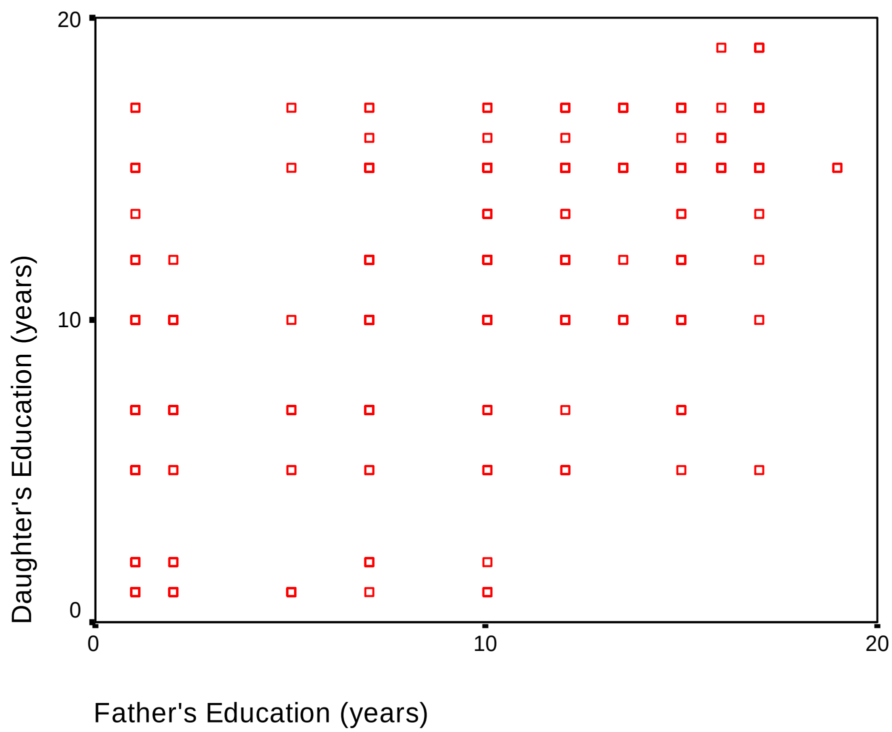
1. In my study population, and most likely in the study populations of several others, men and women do not inherit either wealth or status in the same ways. In India, women typically do not inherit property on the death of the parents but rather receive dowry on marriage. Likewise, many women even in modern urban India do not work in the labor market, and so their social status and/or wealth is largely determined by the occupation, wealth, or social status of their husband. Parents, however, influence the lifelong social status of daughters through arranged marriages with negotiated dowries and other marriage costs. Thus, in some circumstances appropriate measures of a daughter's status would include measures of her husband's status.
2. We should discuss guidelines and data requirements for imputing income from educational and occupational data, especially for recent historical populations. I could have income estimates for around 1500 additional men going back to the early 20th century if this can be worked out; in many cases this would give me three generations of income data.
3. We should consider whether it is necessary or useful to attempt to estimate wealth using price data and catalogues of consumer items.
4. We should consider the question of how to deal with the demographic transition in our interpretation of the intergenerational correlation of fertility.
5. We should consider whether using the average of parents' characteristics (i.e. wealth, income, etc.) is an appropriate way of incorporating information from two parents. For example, using father's income to predict son's or daughter's income does not allow for any effects of mother's income or lack thereof. And averaging mother's and father's income allows us some information on mother's income, but only if all or most mothers work. In my sample most mothers do not work, so for them the 'parents' average income' is the father's income. This means that if fathers in two families earn the same income, the one with a working wife will actually have a lower average income than the

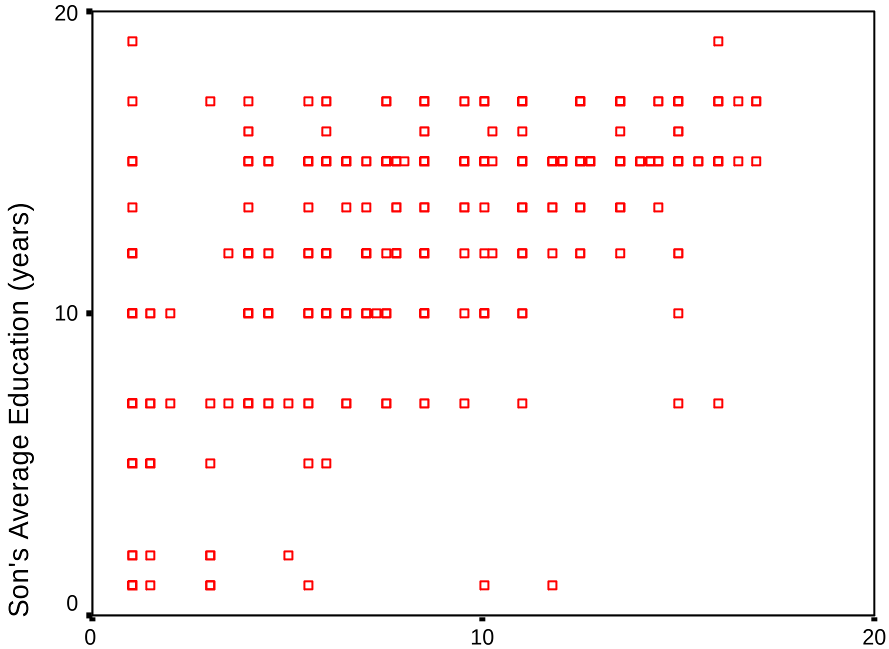
one with just the father working—as if, absurdly, a working mother's income actually decreased the income available to the family. A better solution should be found for constructing a husband-wife composite variable. Even simply adding the mother's and father's incomes would be more realistic than the current construction in the case of income in India.

Appendix I. Scatterplots of Raw and Logged Data for all Measures of Wealth

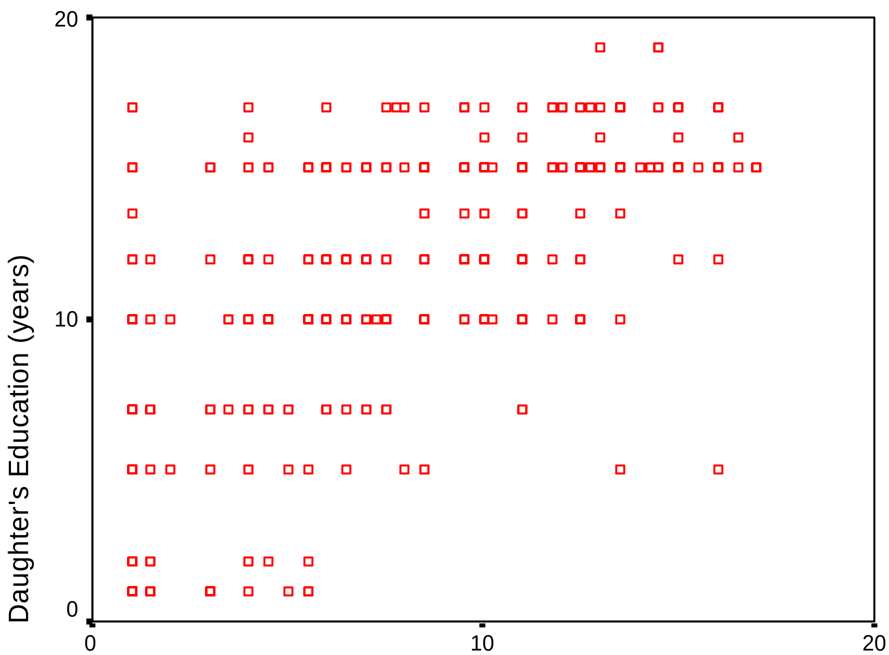
Raw Education





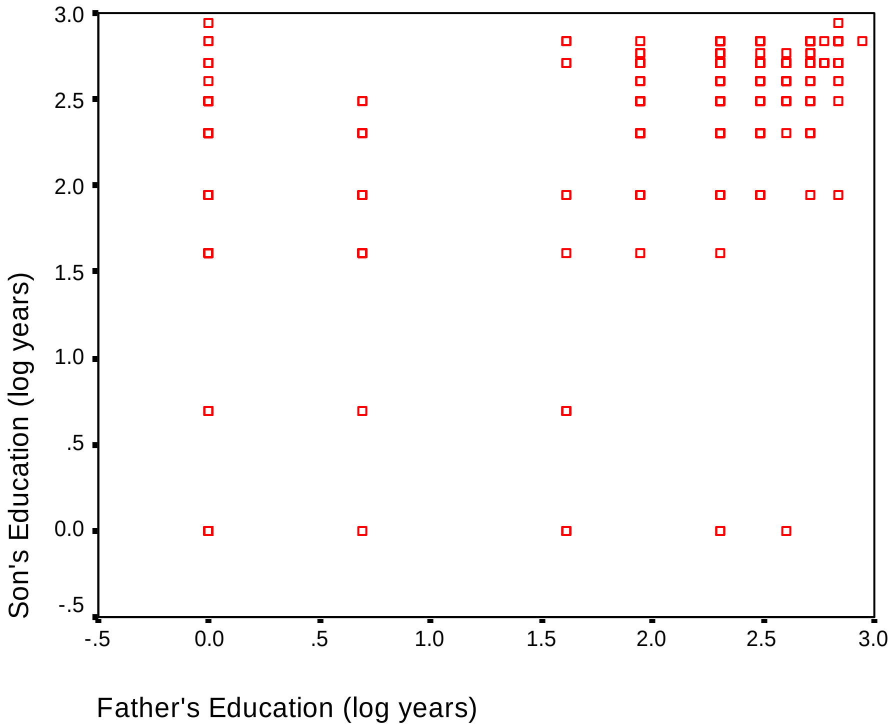


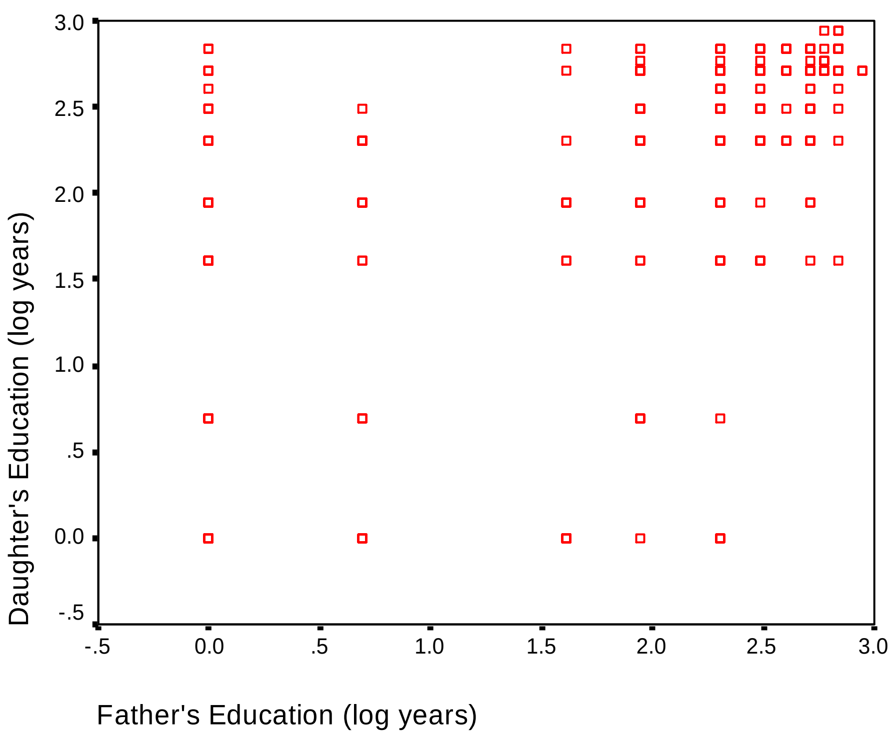
Parents' Average Education (years)

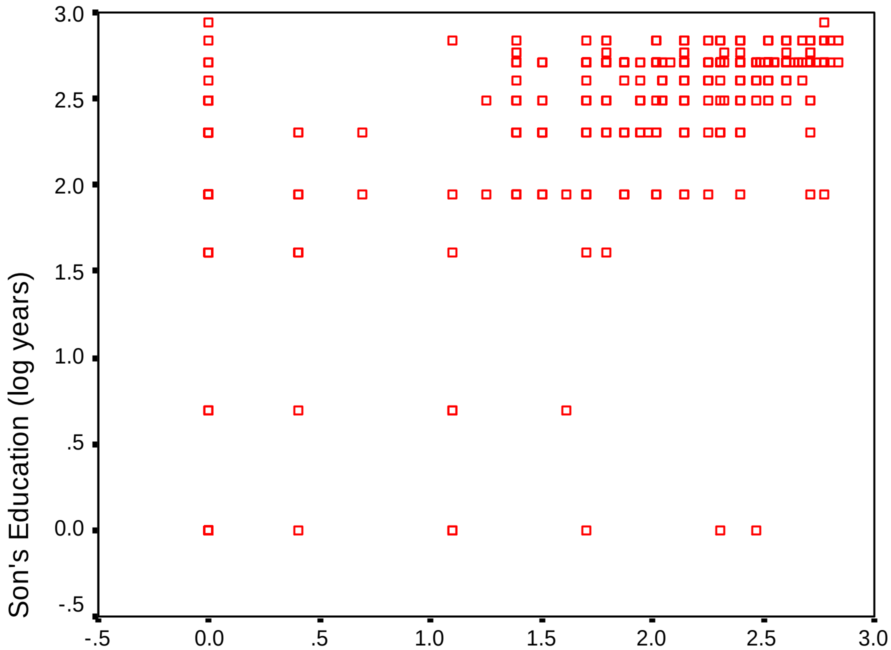


Parents' Average Education (years)

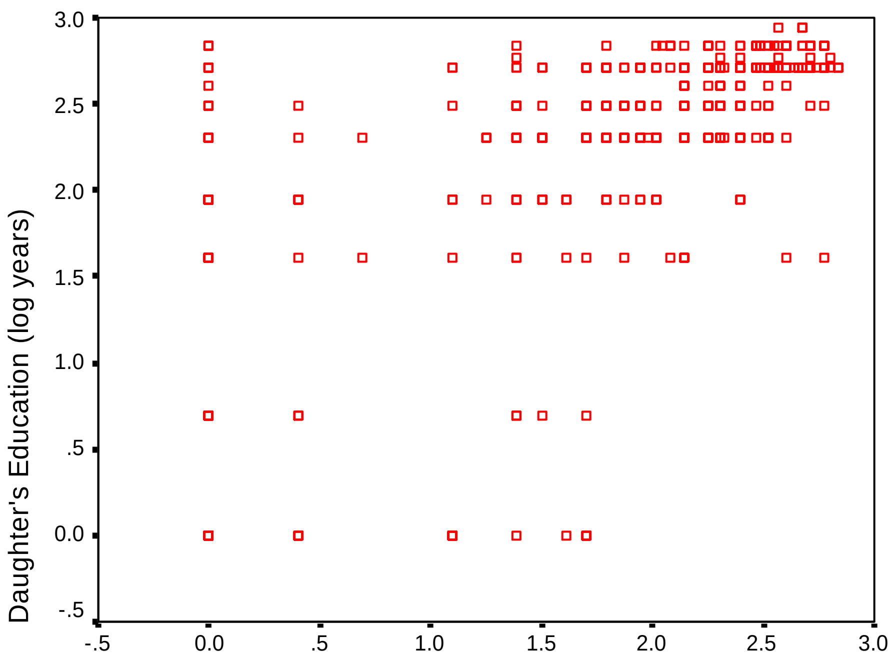
Log Education





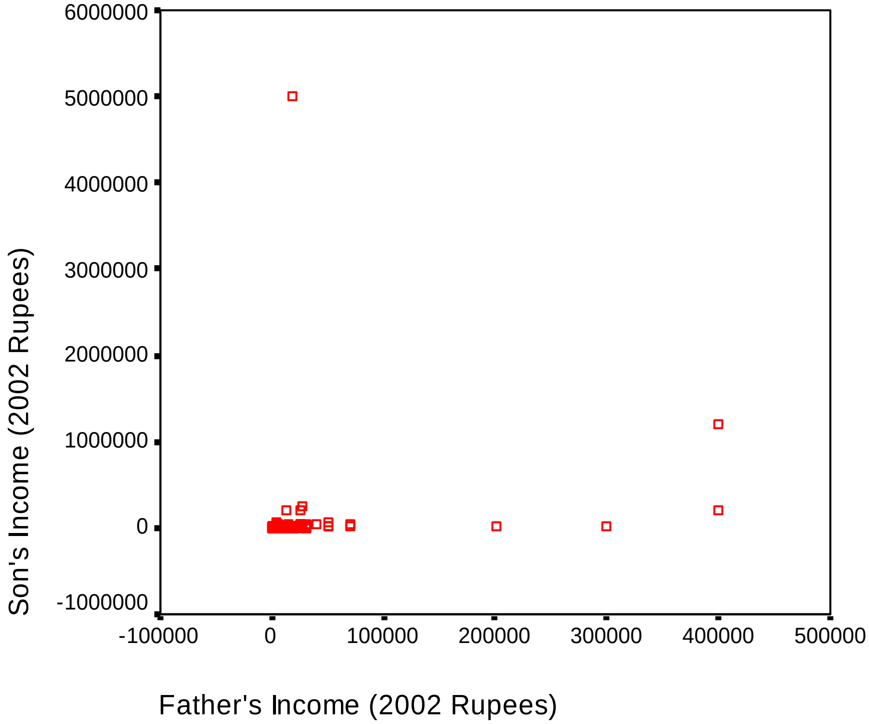


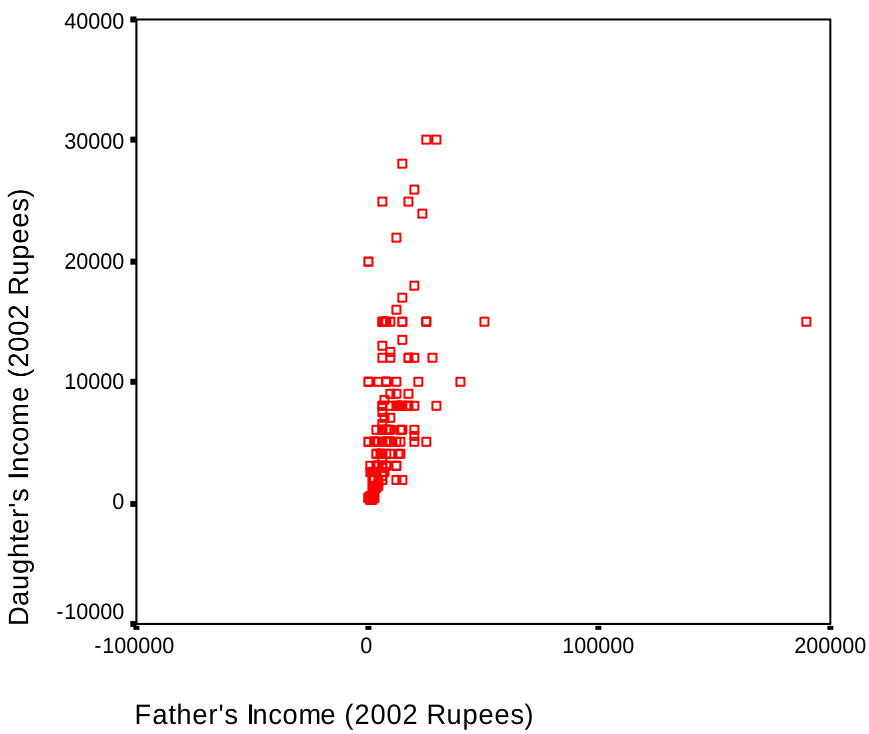
Parents' Average Education (log years)

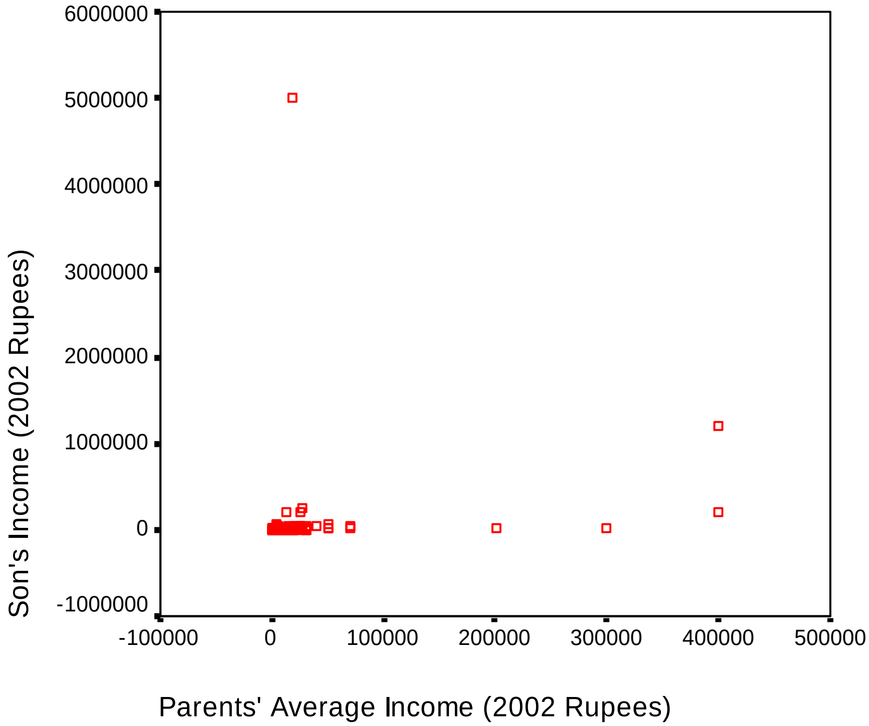


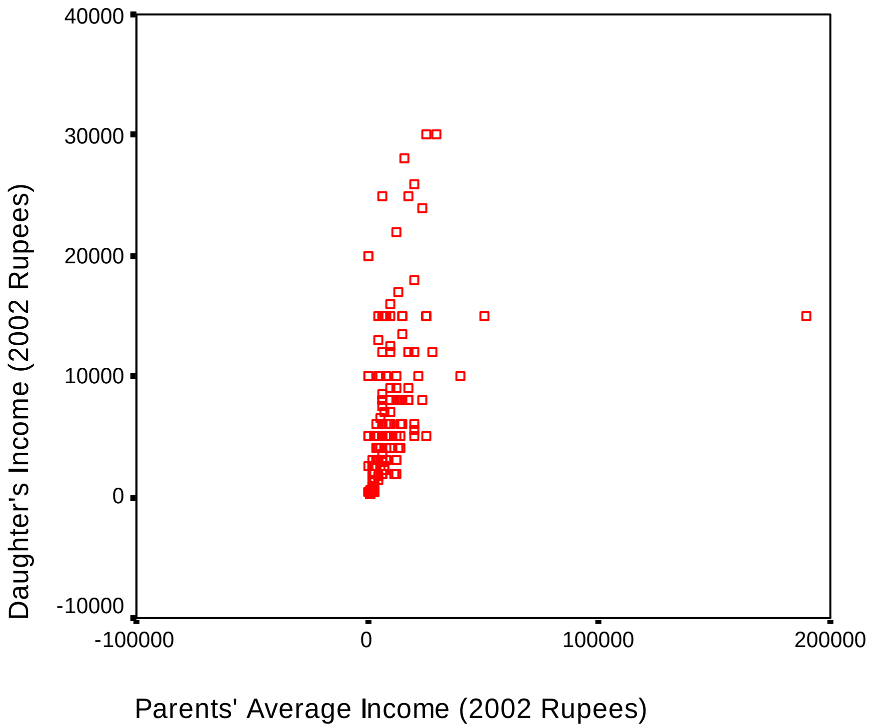
Parents' Average Education (log years)

Raw Income

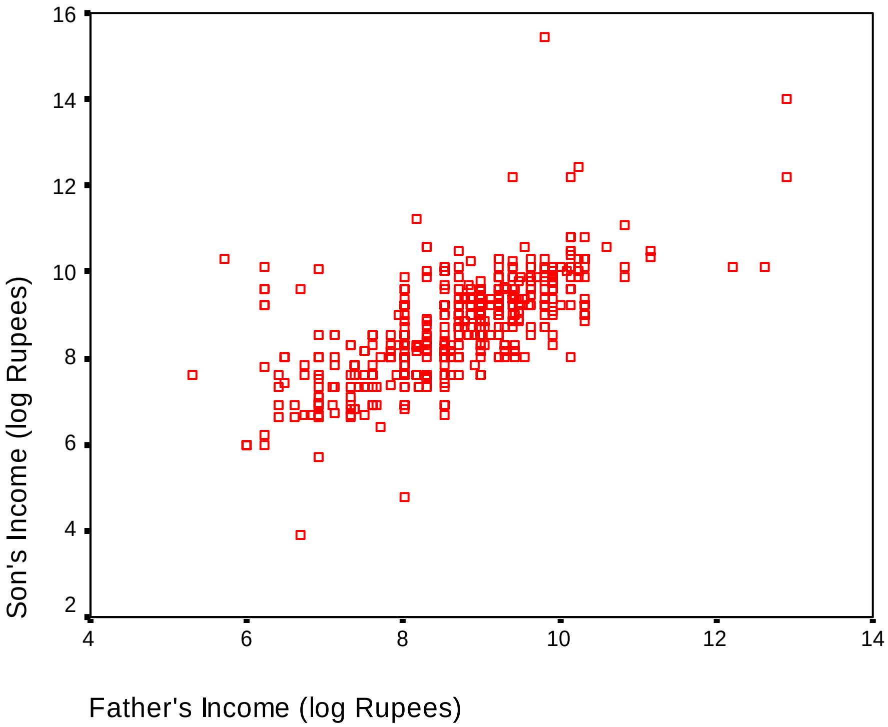

6000000
5000000
4000000
3000000
2000000
1000000
0
-1000000
Son's Income (2002 Rupees)
-100000
0
100000
200000
300000
400000
500000
Father's Income (2002 Rupees)

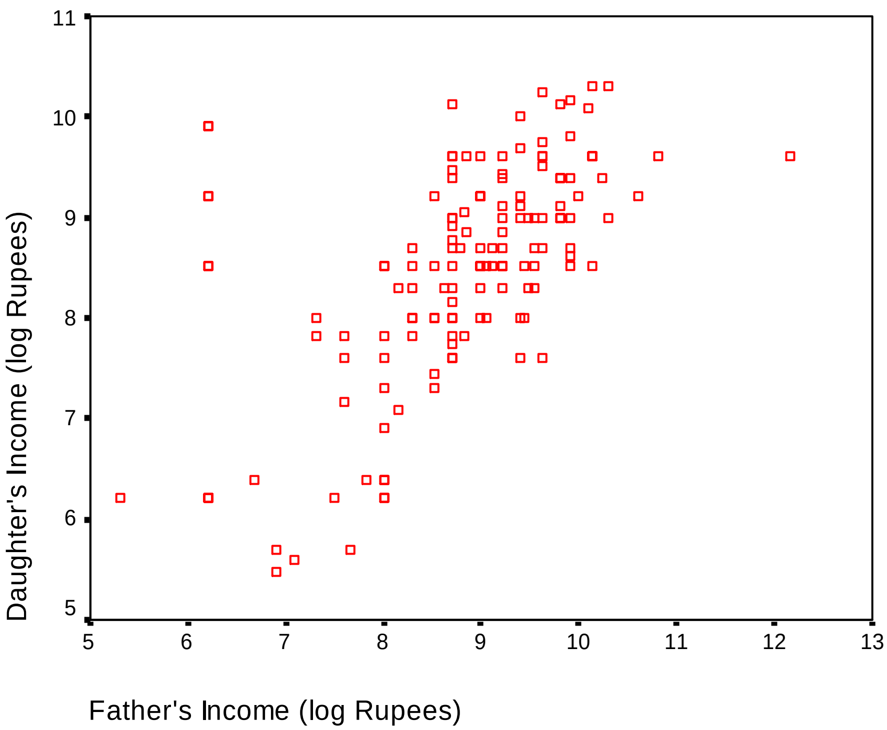

40000
30000
20000
10000
0
-10000
Daughter's Income (2002 Rupees)
-100000
0
100000
200000
Father's Income (2002 Rupees)

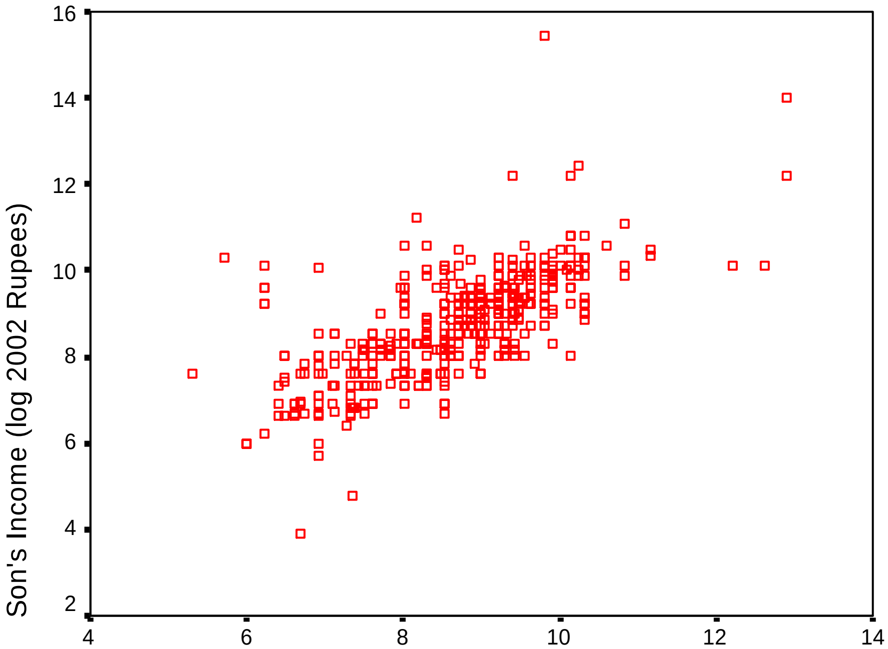

6000000
5000000
4000000
3000000
2000000
1000000
0
-1000000
Son's Income (2002 Rupees)
-100000
0
100000
200000
300000
400000
500000
Parents' Average Income (2002 Rupees)


40000
30000
20000
10000
0
-10000
Daughter's Income (2002 Rupees)
-100000
0
100000
200000
Parents' Average Income (2002 Rupees)

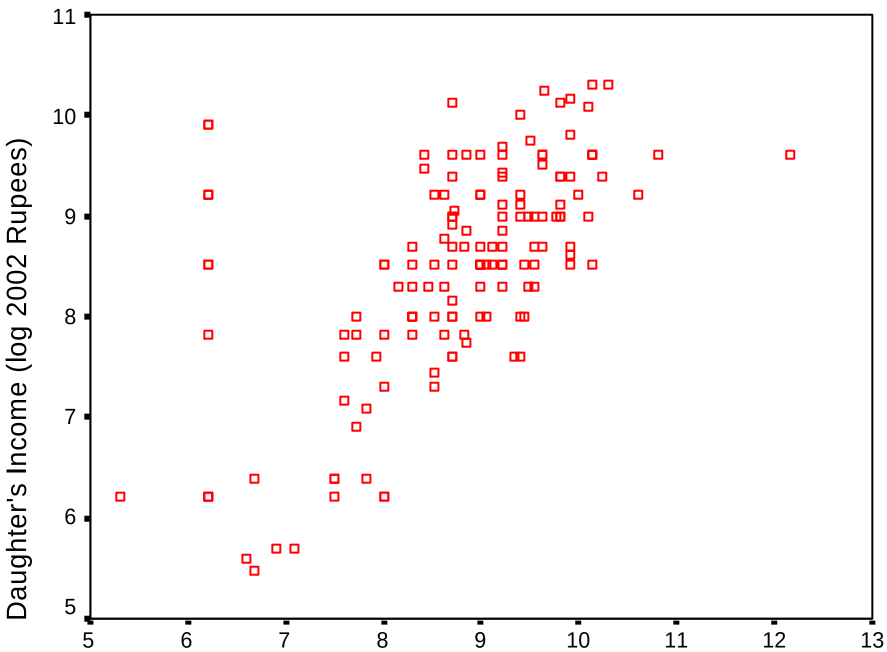
Log Income


16
14
12
10
8
6
4
2
Son's Income (log Rupees)
4
6
8
10
12
14
Father's Income (log Rupees)


11
10
9
8
7
6
5
Daughter's Income (log Rupees)
5
6
7
8
9
10
11
12
13
Father's Income (log Rupees)

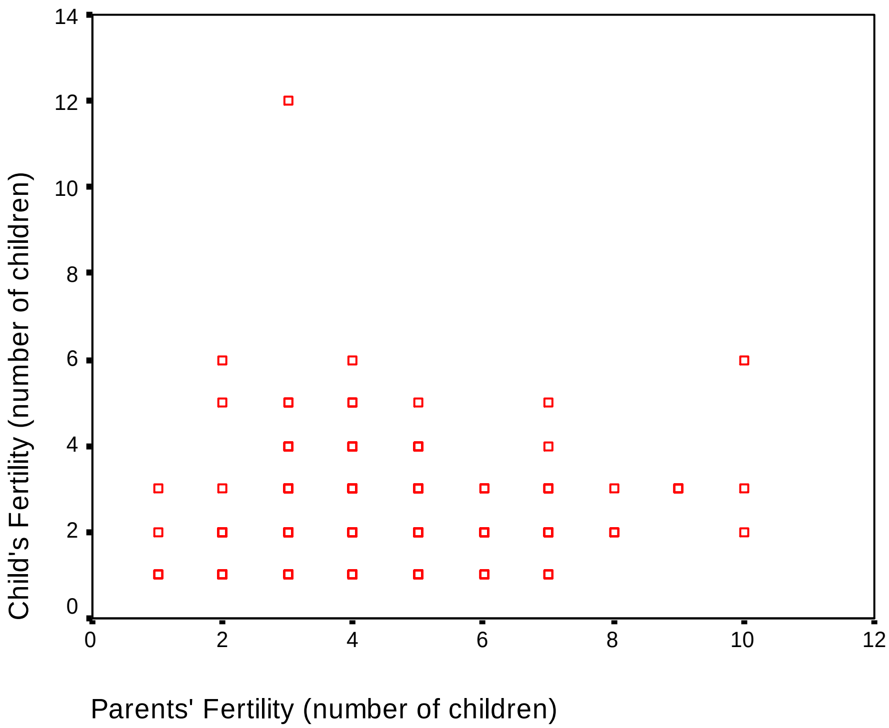


Parents' Average Income (log 2002 Rupees)

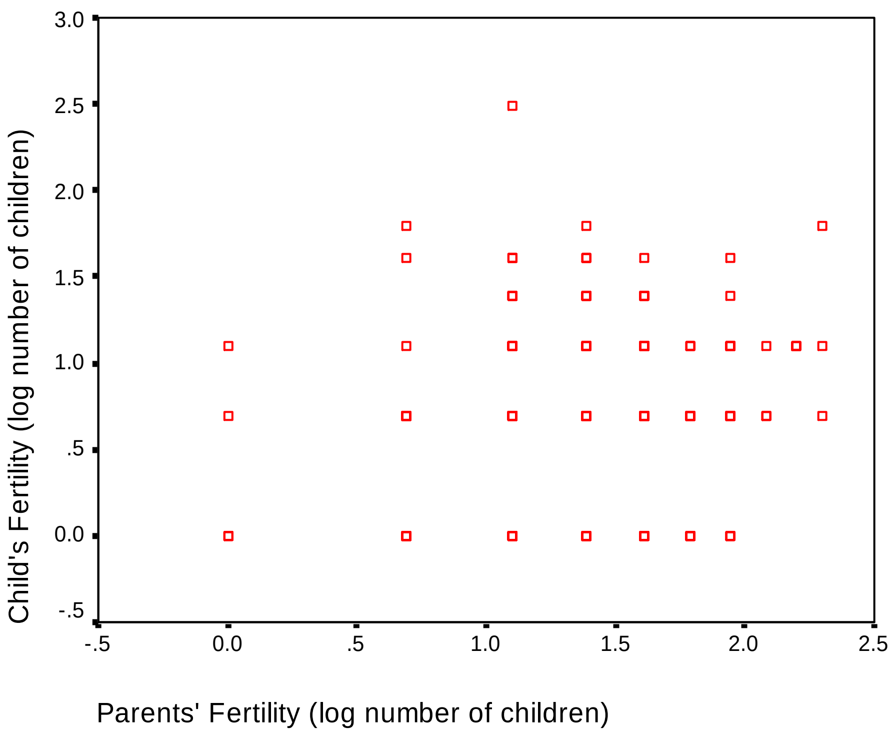


Parents' Average Income (log 2002 Rupees)

Fertility



Log Fertility



Appendix II. Complete Regression Results for Estimates of Beta
For each analysis, the dependent variable and predictor of interest are highlighted in yellow as are the coefficient (Beta), standard error, t value, and p value for the predictor of interest in each set of results.

1. Dependent Variable = Child's Education (Logged)
Independent Variables = Father's Education (Logged), Child's Age, Child's Age Squared, Father's Age, Father's Age Squared, and the interaction (Child's Age – 30) * Father's Education (Logged)

by var11b, sort: regress chedyrln faedyrln chage chagesqd faage faagesqd faedintl if chage>20 , robust cluster(respno)

```
------------------------------------------------------------------------------
SONS
Linear regression                          Number of obs =     585
                                           F(  6,   303) =   13.06
                                           Prob > F      =  0.0000
                                           R-squared     =  0.3502
Number of clusters (respno) = 304          Root MSE      =  .56759
------------------------------------------------------------------------------
             |               Robust
    chedyrln |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
    faedyrln |   .4046105   .0482257     8.39   0.000     .3097109    .4995101
       chage |   .0078262   .0322696     0.24   0.809    -.0556747     .071327
    chagesqd |   .0000709   .0003543     0.20   0.842    -.0006263    .0007681
       faage |   .0200787   .0313518     0.64   0.522    -.0416162    .0817735
    faagesqd |   -.000118   .0002289    -0.52   0.607    -.0005686    .0003325
    faedintl |  -.0051219    .004939    -1.04   0.301    -.0148411    .0045973
       _cons |    .507568   1.149342     0.44   0.659    -1.754135    2.769271
------------------------------------------------------------------------------
------------------------------------------------------------------------------
DAUGHTERS
Linear regression                          Number of obs =     502
                                           F(  6,   291) =   18.46
                                           Prob > F      =  0.0000
                                           R-squared     =  0.5095
Number of clusters (respno) = 292          Root MSE      =  .56885
------------------------------------------------------------------------------
             |               Robust
    chedyrln |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
    faedyrln |   .6106706   .0597641    10.22   0.000     .4930459    .7282953
       chage |  -.0493003   .0329651    -1.50   0.136    -.1141805      .01558
    chagesqd |   .0003632   .0003292     1.10   0.271    -.0002849    .0010112
       faage |   -.034805   .0227731    -1.53   0.128    -.0796259    .0100159
    faagesqd |   .0002368   .0001846     1.28   0.201    -.0001266    .0006002
    faedintl |   .0102215   .0073812     1.38   0.167    -.0043057    .0247488
       _cons |   3.383903   .9363967     3.61   0.000     1.540934    5.226872
------------------------------------------------------------------------------
```

2. Dependent Variable = Child's Education (Logged)
Independent Variables = Parents' Average Education (Logged), Child's Age, Child's Age Squared, Father's Age, Father's Age Squared, and the interaction (Child's Age – 30) * Parents' Average Education (Logged)

```
. by var11b, sort: regress chedyrln paravedl chage chagesqd faage faagesqd paedintl if chage>20 ,
robust cluster(respno)

-------------------------------------------------------------------------------
SONS

Linear regression                          Number of obs =     585
                                       F(  6,   303) =   14.77
                                       Prob > F      =  0.0000
                                       R-squared     =  0.3741
Number of clusters (respno) = 304          Root MSE      =  .55704

------------------------------------------------------------------------------
             |               Robust
    chedyrln |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
    paravedl |   .4492196   .0506347     8.87   0.000     .3495795    .5488597
       chage |    .010014   .0332857     0.30   0.764    -.0554865    .0755144
    chagesqd |   .0000388   .0003658     0.11   0.916     -.000681    .0007586
       faage |   .0045197   .0324032     0.14   0.889    -.0592441    .0682834
    faagesqd |   9.53e-06   .0002363     0.04   0.968    -.0004555    .0004745
    paedintl |  -.0052888   .0051678    -1.02   0.307     -.015458    .0048805
       _cons |   .9304342   1.174417     0.79   0.429    -1.380612    3.241481
------------------------------------------------------------------------------

-------------------------------------------------------------------------------
DAUGHTERS

Linear regression                          Number of obs =     502
                                       F(  6,   291) =   21.86
                                       Prob > F      =  0.0000
                                       R-squared     =  0.5517
Number of clusters (respno) = 292          Root MSE      =  .54379

------------------------------------------------------------------------------
             |               Robust
    chedyrln |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
    paravedl |   .6824934   .0603646    11.31   0.000     .5636869    .8012999
       chage |  -.0462416   .0336052    -1.38   0.170    -.1123815    .0198984
    chagesqd |   .0002952   .0003428     0.86   0.390    -.0003796    .0009699
       faage |   -.054749   .0209856    -2.61   0.010    -.0960517   -.0134463
    faagesqd |   .0004162   .0001675     2.49   0.014     .0000866    .0007459
    paedintl |   .0119781   .0075625     1.58   0.114     -.002906    .0268623
       _cons |   3.860553   .9460983     4.08   0.000      1.99849    5.722616
------------------------------------------------------------------------------
```

3. Dependent Variable = Child's Income (Logged)
Independent Variables = Father's Income (Logged), Child's Age, Child's Age Squared, Father's Age, Father's Age Squared, and the interaction (Child's Age – 30) * Father's Income (Logged)

```
. by var11b, sort: regress chincln faincmln chage chagesqd faage faagesqd faintl if chage>20,
robust cluster(respno)

-------------------------------------------------------------------------------
SONS

Linear regression                                  Number of obs =     479
                                                   F(  6,   270) =   52.72
                                                   Prob > F      =  0.0000
                                                   R-squared     =  0.5212
Number of clusters (respno) = 271                  Root MSE      =  .78381

------------------------------------------------------------------------------
             |               Robust
     chincln |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
    faincmln |   .7764156    .050007    15.53   0.000     .6779623    .8748689
       chage |   .2449935   .0796955     3.07   0.002     .0880899     .401897
    chagesqd |  -.0005247   .0007221    -0.73   0.468    -.0019464    .0008969
       faage |  -.0062957   .0357838    -0.18   0.860    -.0767465    .0641551
    faagesqd |   .0001778   .0003072     0.58   0.563    -.0004271    .0007827
     faintl  |   -.022344   .0081562    -2.74   0.007    -.0384019   -.0062861
       _cons |  -5.169575   2.420929    -2.14   0.034    -9.935874   -.4032762
------------------------------------------------------------------------------

-------------------------------------------------------------------------------
DAUGHTERS

Linear regression                                  Number of obs =     135
                                                   F(  6,   102) =   13.41
                                                   Prob > F      =  0.0000
                                                   R-squared     =  0.4914
Number of clusters (respno) = 103                  Root MSE      =  .70749

------------------------------------------------------------------------------
             |               Robust
     chincln |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
    faincmln |   .8204771   .1137438     7.21   0.000     .5948667    1.046087
       chage |   .4555934   .1264226     3.60   0.000     .2048347    .7063521
    chagesqd |  -.0015776   .0006914    -2.28   0.025     -.002949   -.0002062
       faage |   -.086214    .052269    -1.65   0.102    -.1898893    .0174614
    faagesqd |    .000801    .000428     1.87   0.064    -.0000479    .0016499
     faintl  |  -.0368119   .0106153    -3.47   0.001    -.0578672   -.0157565
       _cons |  -8.766879   3.948661    -2.22   0.029    -16.59903   -.9347292
------------------------------------------------------------------------------
```

4. Dependent Variable = Child's Income (Logged)
Independent Variables = Parents' Average Income (Logged), Child's Age, Child's Age Squared, Father's Age, Father's Age Squared, and the interaction (Child's Age – 30) * Parents' Average Income (Logged)

. by var11b, sort: regress chincln paravinl chage chagesqd faage faagesqd painintl if chage>20, robust cluster(respno)

SONS

```
Linear regression                              Number of obs =     480
                                               F(  6,   271) =   50.55
                                               Prob > F      =  0.0000
                                               R-squared     =  0.5204
Number of clusters (respno) = 272              Root MSE      =  .78387
```

chincln	Coef.	Robust Std. Err.	t	P>|t|	[95% Conf. Interval]	
paravinl	.7729013	.050098	15.43	0.000	.6742705	.871532
chage	.2336234	.079293	2.95	0.003	.0775149	.389732
chagesqd	-.0004935	.0007189	-0.69	0.493	-.0019089	.0009219
faage	-.0057379	.0362251	-0.16	0.874	-.0770563	.0655805
faagesqd	.0001631	.000308	0.53	0.597	-.0004433	.0007694
painintl	-.0212393	.0082425	-2.58	0.011	-.0374668	-.0050118
_cons	-4.79392	2.450403	-1.96	0.051	-9.618166	.030327

DAUGHTERS

```
Linear regression                              Number of obs =     136
                                               F(  6,   103) =   14.97
                                               Prob > F      =  0.0000
                                               R-squared     =  0.4781
Number of clusters (respno) = 104              Root MSE      =  .71567
```

chincln	Coef.	Robust Std. Err.	t	P>|t|	[95% Conf. Interval]	
paravinl	.7449284	.1177571	6.33	0.000	.5113849	.9784718
chage	.3870194	.1340712	2.89	0.005	.1211208	.6529181
chagesqd	-.0013758	.0006989	-1.97	0.052	-.0027619	.0000103
faage	-.1197025	.0615235	-1.95	0.054	-.2417198	.0023149
faagesqd	.0010461	.0004974	2.10	0.038	.0000597	.0020325
painintl	-.0309372	.011548	-2.68	0.009	-.0538399	-.0080345
_cons	-5.061162	4.844418	-1.04	0.299	-14.66892	4.546598

5. Dependent Variable = Child's Fertility (Logged)
Independent Variables = Parents' Fertility (Logged), Child's Age, Child's Age Squared, Father's Age, Father's Age Squared, and the interaction (Child's Age – 30) * Parents' Fertility (Logged)

```
. regress chnochln parnochln chage chagesqd faage faagesqd parnochintln if yrmg<1993, robust
cluster(respno)

Linear regression                              Number of obs =     313
                                               F(  6,   158) =    4.97
                                               Prob > F      =  0.0001
                                               R-squared     =  0.0969
Number of clusters (respno) = 159              Root MSE      =  .42715

------------------------------------------------------------------------------
             |               Robust
    chnochln |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
   parnochln |   .3030856   .0971886     3.12   0.002     .1111292     .495042
       chage |  -.0946847   .0340585    -2.78   0.006    -.1619534   -.0274161
    chagesqd |   .0012704    .000411     3.09   0.002     .0004586    .0020822
       faage |  -.0063367   .0260879    -0.24   0.808    -.0578626    .0451892
    faagesqd |   .0000798   .0001879     0.42   0.672    -.0002914    .0004509
parnochintln |  -.0128164   .0075789    -1.69   0.093    -.0277855    .0021527
       _cons |   2.163426   1.063404     2.03   0.044     .0631048    4.263748
------------------------------------------------------------------------------
```

Appendix III. Complete Regression Results for Estimates of r (both with and without controls)

1. Dependent Variable = Child's Education (Not Logged)
Independent Variables = Father's Education (Not Logged), Child's Age, Child's Age Squared, Father's Age, Father's Age Squared, and the interaction (Child's Age – 30) * Father's Education (Not Logged)

```
by var11b, sort: regress chedyrs faedyrs chage chagesqd faage faagesqd faedintn if chage>20,
robust cluster(respno)
-------------------------------------------------------------------------------
SONS

Linear regression                              Number of obs =     585
                                          F(  6,   303) =   38.25
                                          Prob > F      =  0.0000
                                          R-squared     =  0.4533
Number of clusters (respno) = 304                Root MSE      =  3.2087

------------------------------------------------------------------------------
             |               Robust
     chedyrs |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
     faedyrs |   .5519064   .0385413    14.32   0.000     .4760639    .6277489
       chage |   .0963417    .174028     0.55   0.580    -.2461147    .4387982
    chagesqd |  -.0004023   .0021595    -0.19   0.852    -.0046518    .0038472
       faage |   .1264594    .182367     0.69   0.489    -.2324067    .4853255
    faagesqd |  -.0007952   .0013724    -0.58   0.563    -.0034958    .0019054
    faedintn |  -.0037194   .0040913    -0.91   0.364    -.0117704    .0043317
       _cons |  -.1603459   6.025384    -0.03   0.979    -12.01724    11.69655
------------------------------------------------------------------------------
-------------------------------------------------------------------------------
DAUGHTERS

Linear regression                              Number of obs =     502
                                          F(  6,   291) =   35.34
                                          Prob > F      =  0.0000
                                          R-squared     =  0.5144
Number of clusters (respno) = 292                Root MSE      =  3.3495

------------------------------------------------------------------------------
             |               Robust
     chedyrs |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
     faedyrs |   .6929836   .0505511    13.71   0.000     .5934914    .7924758
       chage |    -.24332   .1742608    -1.40   0.164    -.5862914    .0996513
    chagesqd |   .0016172    .002183     0.74   0.459    -.0026793    .0059137
       faage |  -.1661284   .1437664    -1.16   0.249    -.4490821    .1168253
    faagesqd |   .0012875   .0011772     1.09   0.275    -.0010293    .0036043
    faedintn |   .0099021   .0054434     1.82   0.070    -.0008114    .0206156
       _cons |   15.78541   4.973639     3.17   0.002      5.99654    25.57427
------------------------------------------------------------------------------
```

1b. Dependent Variable = Child's Education (Not Logged)
Independent Variable = Father's Education (Not Logged)

- Raw correlation; no controls

```
. by var11b, sort: regress chedyrs faedyrs if chage>20, robust cluster(respno)

-------------------------------------------------------------------------------
SONS

Linear regression                              Number of obs =     607
                                        F(  1,   314) =  241.26
                                        Prob > F       =  0.0000
                                        R-squared      =  0.4644
Number of clusters (respno) = 315                 Root MSE       =  3.2546

------------------------------------------------------------------------------
             |               Robust
     chedyrs |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
     faedyrs |  .5718163   .0368139    15.53   0.000     .4993831    .6442494
       _cons |  6.885877   .4580885    15.03   0.000     5.984566    7.787188
------------------------------------------------------------------------------

-------------------------------------------------------------------------------
DAUGHTERS

Linear regression                              Number of obs =     526
                                        F(  1,   301) =  235.73
                                        Prob > F       =  0.0000
                                        R-squared      =  0.5463
Number of clusters (respno) = 302                 Root MSE       =  3.3794

------------------------------------------------------------------------------
             |               Robust
     chedyrs |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
     faedyrs |  .7262548   .0473027    15.35   0.000      .633169    .8193406
       _cons |  4.367461   .5906281     7.39   0.000     3.205178    5.529745
------------------------------------------------------------------------------
```

2. Dependent Variable = Child's Education (Not Logged)
Independent Variables = Parents' Average Education (Not Logged), Child's Age, Child's Age Squared, Father's Age, Father's Age Squared, and the interaction (Child's Age – 30) * Parents' Average Education (Not Logged)

. by var11b, sort: regress chedyrs paraved chage chagesqd faage faagesqd paedintn if chage>20, robust cluster(respno)

```
------------------------------------------------------------------------------
SONS

Linear regression                               Number of obs =     585
                                                F(  6,   303) =   43.65
                                                Prob > F      =  0.0000
                                                R-squared     =  0.4684
Number of clusters (respno) = 304               Root MSE      =  3.1643

------------------------------------------------------------------------------
             |               Robust
     chedyrs |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
     paraved |   .6331615   .0413057    15.33   0.000     .5518791    .7144438
       chage |   .1008827   .1791528     0.56   0.574    -.2516585    .4534239
    chagesqd |  -.0005715   .0022228    -0.26   0.797    -.0049456    .0038026
       faage |   .0280814   .2006228     0.14   0.889     -.366709    .4228718
    faagesqd |   .0000703   .0014904     0.05   0.962    -.0028625    .0030031
    paedintn |  -.0025717   .0046802    -0.55   0.583    -.0117815     .006638
       _cons |   2.828574   6.508221     0.43   0.664     -9.97846    15.63561
------------------------------------------------------------------------------

------------------------------------------------------------------------------
DAUGHTERS

Linear regression                               Number of obs =     502
                                                F(  6,   291) =   40.41
                                                Prob > F      =  0.0000
                                                R-squared     =  0.5455
Number of clusters (respno) = 292               Root MSE      =  3.2405

------------------------------------------------------------------------------
             |               Robust
     chedyrs |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
     paraved |   .7968004   .0527539    15.10   0.000     .6929729     .900628
       chage |  -.1709361   .1776452    -0.96   0.337    -.5205683    .1786962
    chagesqd |   .0002292   .0023299     0.10   0.922    -.0043564    .0048149
       faage |  -.2941452    .135768    -2.17   0.031     -.561357   -.0269335
    faagesqd |   .0025353   .0011037     2.30   0.022     .0003632    .0047075
    paedintn |   .0145999     .00609     2.40   0.017     .0026139    .0265859
       _cons |   18.26047   5.101146     3.58   0.000     8.220648    28.30028
------------------------------------------------------------------------------
```

2b. Dependent Variable = Child's Education (Not Logged)
Independent Variable = Parents' Average Education (Not Logged)

- Raw correlation; no controls

```
. by var11b, sort: regress chedyrs paraved if chage>20, robust cluster(respno)

------------------------------------------------------------------------------
SONS

Linear regression                               Number of obs =     607
                                                F(  1,   314) =  264.90
                                                Prob > F      =  0.0000
                                                R-squared     =  0.4649
Number of clusters (respno) = 315               Root MSE      =  3.2529

------------------------------------------------------------------------------
             |               Robust
     chedyrs |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
     paraved |   .6448177   .0396183    16.28   0.000     .5668668    .7227685
       _cons |   7.259879   .4291716    16.92   0.000     6.415464    8.104295
------------------------------------------------------------------------------

------------------------------------------------------------------------------
DAUGHTERS

Linear regression                               Number of obs =     526
                                                F(  1,   301) =  281.11
                                                Prob > F      =  0.0000
                                                R-squared     =  0.5658
Number of clusters (respno) = 302               Root MSE      =   3.306

------------------------------------------------------------------------------
             |               Robust
     chedyrs |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
     paraved |   .8276232   .0493621    16.77   0.000     .7304846    .9247618
       _cons |    4.71165   .5403116     8.72   0.000     3.648384    5.774917
------------------------------------------------------------------------------
```

3. Dependent Variable = Child's Income (Not Logged)
Independent Variables = Father's Income (Not Logged), Child's Age, Child's Age Squared, Father's Age, Father's Age Squared, and the interaction (Child's Age – 30) * Father's Income (Not Logged)

. by var11b, sort: regress var19 hvar19 chage chagesqd faage faagesqd fainintn if chage>20 , robust cluster(respno)

SONS

Linear regression

Number of obs	= 479
F(6, 270)	= 4.80
Prob > F	= 0.0001
R-squared	= 0.0466
Root MSE	= 2.3e+05

Number of clusters (respno) = 271

var19	Coef.	Robust Std. Err.	t	P>|t|	[95% Conf.	Interval]
hvar19	2.092783	.5441475	3.85	0.000	1.021471	3.164095
chage	1171.789	6153.747	0.19	0.849	-10943.64	13287.22
chagesqd	-54.24778	88.17951	-0.62	0.539	-227.8546	119.3591
faage	3998.173	5290.927	0.76	0.451	-6418.546	14414.89
faagesqd	-28.51637	43.58819	-0.65	0.514	-114.3323	57.29958
fainintn	.3465705	.1413323	2.45	0.015	.068317	.624824
_cons	-119662.1	235143.7	-0.51	0.611	-582610.5	343286.4

DAUGHTERS

Linear regression

Number of obs	= 135
F(6, 102)	= 9.66
Prob > F	= 0.0000
R-squared	= 0.3273
Root MSE	= 5479.5

Number of clusters (respno) = 103

var19	Coef.	Robust Std. Err.	t	P>|t|	[95% Conf.	Interval]
hvar19	.4197135	.0805855	5.21	0.000	.2598725	.5795545
chage	1175.827	424.8102	2.77	0.007	333.2181	2018.436
chagesqd	-13.67399	4.785286	-2.86	0.005	-23.16558	-4.182396
faage	-835.6212	689.9656	-1.21	0.229	-2204.165	532.9222
faagesqd	7.183949	5.381364	1.33	0.185	-3.489962	17.85786
fainintn	-.005617	.0142233	-0.39	0.694	-.0338288	.0225947
_cons	4555.967	22328.26	0.20	0.839	-39732.03	48843.96

3b. Dependent Variable = Child's Income (Not Logged)
Independent Variable = Father's Income (Not Logged)

- Raw correlation; no controls

```
. by var11b, sort: regress var19 hvar19 if chage>20, robust cluster(respno)

------------------------------------------------------------------------------
SONS

Linear regression                                 Number of obs =     495
                                               F(  1,   280) =   11.51
                                               Prob > F      =  0.0008
                                               R-squared     =  0.0301
Number of clusters (respno) = 281              Root MSE      =  2.3e+05

------------------------------------------------------------------------------
             |               Robust
       var19 |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
      hvar19 |   1.310833   .3863074     3.39   0.001      .550398    2.071269
       _cons |   8730.945   9892.513     0.88   0.378     -10742.2    28204.09
------------------------------------------------------------------------------

------------------------------------------------------------------------------
DAUGHTERS

Linear regression                                 Number of obs =     137
                                               F(  1,   104) =    3.45
                                               Prob > F      =  0.0662
                                               R-squared     =  0.0936
Number of clusters (respno) = 105              Root MSE      =  6251.6

------------------------------------------------------------------------------
             |               Robust
       var19 |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
      hvar19 |   .1150379   .0619619     1.86   0.066    -.0078349    .2379106
       _cons |   6674.396   824.2825     8.10   0.000     5039.813     8308.98
------------------------------------------------------------------------------
```

4. Dependent Variable = Child's Income (Not Logged)
Independent Variables = Parents' Average Income (Not Logged), Child's Age, Child's Age Squared, Father's Age, Father's Age Squared, and the interaction (Child's Age – 30) * Parents' Average Income (Not Logged)

```
. by var11b, sort: regress var19 paravinc chage chagesqd faage faagesqd paintn if chage>20 ,
robust cluster(respno)

------------------------------------------------------------------------------
SONS

Linear regression                              Number of obs =     480
                                               F(  6,   271) =    4.82
                                               Prob > F      =  0.0001
                                               R-squared     =  0.0467
Number of clusters (respno) = 272              Root MSE      =  2.3e+05

------------------------------------------------------------------------------
             |               Robust
       var19 |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
    paravinc |   2.094237   .5434221     3.85   0.000     1.024371    3.164103
       chage |    1293.55   6097.076     0.21   0.832    -10710.11    13297.21
    chagesqd |  -55.64907   87.71878    -0.63   0.526     -228.346    117.0478
       faage |   4011.169   5296.984     0.76   0.450    -6417.302    14439.64
    faagesqd |  -28.87288   43.58645    -0.66   0.508     -114.684    56.93822
    painintn |   .3467852   .1413479     2.45   0.015     .0685055    .6250648
       _cons |  -121272.7   235003.5    -0.52   0.606    -583937.2    341391.9
------------------------------------------------------------------------------

------------------------------------------------------------------------------
DAUGHTERS

Linear regression                              Number of obs =     136
                                               F(  6,   103) =   10.05
                                               Prob > F      =  0.0000
                                               R-squared     =  0.3344
Number of clusters (respno) = 104              Root MSE      =  5443.7

------------------------------------------------------------------------------
             |               Robust
       var19 |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
    paravinc |    .425954   .0806285     5.28   0.000     .2660464    .5858617
       chage |   1185.732     423.88     2.80   0.006     345.0665    2026.398
    chagesqd |  -13.69528   4.774606    -2.87   0.005    -23.16458   -4.225974
       faage |  -789.7822   662.9166    -1.19   0.236    -2104.521    524.9566
    faagesqd |   6.814865   5.187346     1.31   0.192    -3.473013    17.10274
    painintn |  -.0072612   .0141453    -0.51   0.609    -.0353152    .0207927
       _cons |   2904.153   21194.53     0.14   0.891    -39130.21    44938.51
------------------------------------------------------------------------------
```

4b. Dependent Variable = Child's Income (Not Logged)
Independent Variable = Parents' Average Income (Not Logged)

- Raw correlation; no controls

```
. by var11b, sort: regress var19 paravinc if chage>20 , robust cluster(respno)

------------------------------------------------------------------------------
SONS

Linear regression                                  Number of obs =     496
                                                   F(  1,   281) =   11.56
                                                   Prob > F      =  0.0008
                                                   R-squared     =  0.0302
Number of clusters (respno) = 282                  Root MSE      =  2.3e+05

------------------------------------------------------------------------------
             |               Robust
       var19 |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
    paravinc |   1.312873   .3860798     3.40   0.001     .5528968    2.072848
       _cons |    8827.35    9850.03     0.90   0.371    -10561.86    28216.56
------------------------------------------------------------------------------

------------------------------------------------------------------------------
DAUGHTERS

Linear regression                                  Number of obs =     138
                                                   F(  1,   105) =    3.40
                                                   Prob > F      =  0.0679
                                                   R-squared     =  0.0961
Number of clusters (respno) = 106                  Root MSE      =  6236.1

------------------------------------------------------------------------------
             |               Robust
       var19 |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
    paravinc |   .1167839   .0633174     1.84   0.068    -.0087628    .2423306
       _cons |   6646.813   817.8722     8.13   0.000     5025.123    8268.502
------------------------------------------------------------------------------
```

5. Dependent Variable = Child's Fertility (Not Logged)
Independent Variables = Parents' Fertility (Not Logged), Child's Age, Child's Age Squared, Father's Age, Father's Age Squared, and the interaction (Child's Age – 30) * Parents' Fertility (Not Logged)

. regress chnoch var8 chage chagesqd faage faagesqd parnochintn if yrmg<1993, robust cluster(respno)

```
Linear regression                              Number of obs =     333
                                       F(  6,   163) =    4.08
                                       Prob > F      =  0.0008
                                       R-squared     =  0.0771
Number of clusters (respno) = 164              Root MSE      =  1.1386

------------------------------------------------------------------------------
             |               Robust
      chnoch |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
        var8 |   .1257296    .057246     2.20   0.029     .0126903    .2387689
       chage |  -.2773176   .0940368    -2.95   0.004     -.463005   -.0916301
    chagesqd |    .003533   .0011876     2.97   0.003     .0011879     .005878
       faage |  -.0520893   .0695129    -0.75   0.455    -.1893512    .0851726
    faagesqd |   .0004844    .000497     0.97   0.331    -.0004969    .0014657
 parnochintn |  -.0072792   .0044541    -1.63   0.104    -.0160744     .001516
       _cons |   8.299845   2.588535     3.21   0.002      3.18846    13.41123
------------------------------------------------------------------------------
```

5b. Dependent Variable = Child's Fertility (Not Logged)
Independent Variable = Parents' Fertility (Not Logged)

- Raw correlation; no controls

. regress chnoch var8 if yrmg<1993, robust cluster(respno)

```
Linear regression                              Number of obs =     352
                                       F(  1,   171) =    2.08
                                       Prob > F      =  0.1510
                                       R-squared     =  0.0080
Number of clusters (respno) = 172              Root MSE      =  1.1837

------------------------------------------------------------------------------
             |               Robust
      chnoch |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
        var8 |   .0628822    .043592     1.44   0.151    -.0231655    .1489299
       _cons |   1.781259    .194688     9.15   0.000     1.396958    2.165561
------------------------------------------------------------------------------
```