1. Worms in Kenya

1.1 Randomising occurred at the school level to allow for the estimation of the intent-to-treat effect, or the overall impact of treating the school on the school and its students.

Randomising treatment at the individual level could have led to biased estimates of the treatment effects due to externalities, within treated schools (for example, re-infection between treated and untreated students) and across treatment to control schools (by infection from interactions between children from treated and untreated schools). Other issues which may arise if randomised at the individual include difficulty in implementing, costliness, selection bias, partial compliance, and attrition.

Randomisation at the school level allowed the researchers to estimate the overall effect of the program (the impact of the deworming treatment) by comparing treatment and comparison schools even in the presence of these externalities. However, it is worth noting that the experimental variation only identifies cross-school externalities, estimating within school externalities required alternative non-experimental methods.

1.2 If I had pre and post treatment school attendance records for all schools, I could do a difference-in-difference (DD) method to estimate the effect of deworming on school attendance. However, it is worth noting that if the randomisation is pure and uncompromised the DD method will yield the same impact estimate as simply comparing the outcomes between the treatment and control group. This is because the randomisation should ensure that the potential different conditions between the groups, including initial outcome positions, observed variables and the expected time paths of different groups, are treated at program inception.

A graph of the estimate of the impact is presented in Figure 1.

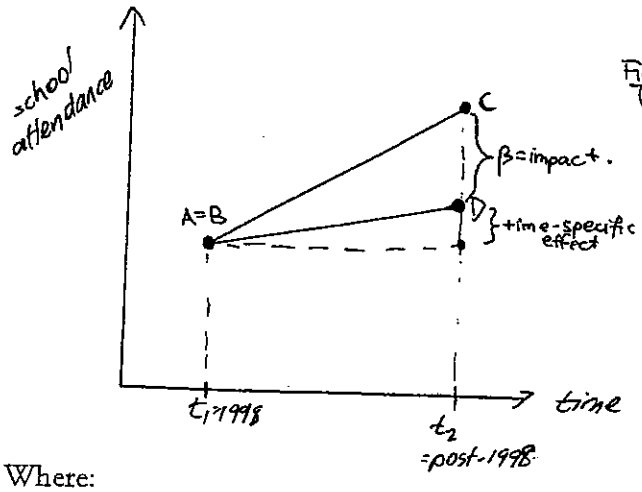


Figure 1. Impact of Deworming on School Attendance.

Where:

A = treatment group, pre-treatment

B = control group, pre-treatment

C = treatment group, post-treatment

D = control group, post-treatment

C – D in Figure 1 represents the impact.

If a DD estimate was used $(D - B) - (C - A)$ would represent the impact.

As shown in Figure 1, $A = C$ due to randomisation and therefore:

$$(D - B) - (C - A) = C - D.$$

However, in this case the DD is nonetheless useful to confirm that the randomisation was pure. While randomisation should ensure that the control group is statistically identical to the treatment group at the start of the experiment, collecting rich baseline data, and even periodic data throughout the experiment, is always good practice, as you can better monitor the integrity of the randomised design.

A few options of estimating impact and respective differences and assumptions are:

a. **Pre-post:** If just the pre and post attendance records for the treatment group alone are compared, the comparison impact will be biased, if there is a time trend to the outcome. This will bias the treatment effects, or estimated impact, because it will be picked up in the estimation. ✓

b. **Post-treatment and control simple difference / OLS:** If I just compare the post-treatment and control outcomes [ignoring pre-treatment outcomes] the impact estimate will be biased if there is a permanent average different in the attendance records between the treatment and control groups and the true treatment effect will be blurred and biased by and permanent differences that might have existed prior to treatment. This is the estimator that I would get from a simple OLS estimation of the form:

$$Y_i = a_2 + \delta_2 t_i + \varepsilon_i$$

where δ_2 is the estimated impact of the treatment [dummy variable t_i].

In the case of a randomised controlled experiment, such as the deworming experiment, the process of randomisation should remove any permanent different between the groups, and both groups should be statistically identical. If this condition holds, this estimator should provide an accurate estimate. ✓

c. **Difference-in-difference (DD) estimator** is the difference in the average outcome in the treatment group before and after treatment and the difference in the average outcome in the control group before and after the treated group is treated. The control group's estimator captures the time trend and the difference estimator for the pre-period is used to estimate the permanent difference. If there is no time trend bias, or permanent difference, the DD estimate should be very equal that of the simple treatment versus control outcome estimator or a simple pre-post estimator.

I would prefer a DD estimator to a straightforward OLS estimate using only post-treatment data if there is any permanent difference in between the groups, as well as if there is a time trend. However, it is worth noting that the DD estimator requires the assumption that if the program did not exist, the two groups would have had identical trajectories over this period. This parallel trend assumption is a common problem in many evaluations and DD estimators can also be biased.

✓ 10/10

1.3 I would run the following regression to estimate the DD described in the previous question:

$$Y_i = \beta_0 + \beta_1 T_i + \beta_2 t_i + \beta_3 T_i * t_i + \varepsilon_i$$

Where:

Y_i = *attendance outcome*

T_i = *treatment dummy, where:*

T= *0 if in control group;* T = *1 if in treatment group*

t_i = *time dummy, where:*

i = *0 if pre; and* i = *1 if post*

β_0 = *intercept / constant term*

β_1 = *treatment group specific effect, which accounts for the differences between treatment and control groups*

β_2 = *time trend effect common to both the treatment and control groups*

β_3 = *the estimated impact of the intervention*

$T_i * t_i$ = *the interaction between the treatment and control, which identifies the group treated, after the treatment*

ε_i = *error term*

✓ 10/10

Alternative

1.1 Randomization at school level captures the overall effect of deworming in the presence of positive spill-over effects across individuals within schools.

If randomization at individual level, with some children within the same school are treated, the rest are in the control group could result in underestimate of the treatment effect, since children in the control group benefit from reduced disease transmission from other children. This tends to underestimate the program effect.

✓ 10/10

1.2 Naive OLS estimate using post-treatment data presupposes the treatment and the control groups have the same baseline characteristics, which may not hold. In addition, it fails to capture the time trend. For example, there may be other programs in place at the same time, also affect school attendance. In this case, school attendance may increase even without the presence of the de-worming program. The third reason for preferring DD to a straightforward OLS estimate using post-treatment data is because DD gives an idea of how big the intervention effect is with a background comparison.

	pre-intervention	post-intervention
control	C_0	C_f
treatment	T_0	T_f

✓ 15/16

For simplicity:

1.3 $Y = \alpha_0 + \alpha_1 \cdot D_{treat} + \alpha_2 \cdot D_{post} + \alpha_3 \cdot D_{treat} \cdot D_{post} + \alpha_4 \cdot X' + e$

D_{treat} is a dummy, $= 1$ if treat $\quad \alpha_3 = (T_f - C_f) - (T_0 - C_0)$: the program eff
$= 0$ if control

D_{post} is a dummy, $= 1$ if post-intervention time
$= 0$ if pre-intervention time

$\alpha_0 = C_0$: average school attendance of the control group at pre-intervention time

$\alpha_1 = T_0 - C_0$: the difference of average school attendance between control & treatm groups at the pre-intervention time

$\alpha_2 = C_f - C_0$: the difference of average school attendance of the control group betw the pre- & post-intervention times.

✓ 10/1

1.4 Describe the data

```
. desc

Contains data from \\pebble\u4583833\Incentives\worms\worm.dta
  obs:        34,792
 vars:            10                          7 Oct 2012 05:14
 size:       695,840 (93.4% of memory free)
```

variable name	storage type	display format	value label	variable label
pupid	long	%10.0g		Pupil Index#
wgrp	byte	%9.0g		Assigned worm group (1,2,3) 98
sex	byte	%9.0g		Pupil gender, 1=male 0=female
grade98	byte	%8.0g		Pupil grade in early 98
old_girl98	byte	%9.0g		Girl >=13 yrs old in 98
pill98	byte	%9.0g		Child took deworming 98
pill99	byte	%9.0g		Child took deworming 99
treat_sch98	byte	%9.0g		School assigned deworming in 98
infect_early99	byte	%9.0g		Moderate-heavy worm infection, early 99
totpar98	float	%9.0g		Avg school participation 98

```
Sorted by:  wgrp

.

. summarize
```

Variable	Obs	Mean	Std. Dev.	Min	Max
pupid	34792	2146865	699024.2	1071714	9146209
wgrp	34792	1.986175	.8093606	1	3
sex	29470	.5207669	.499577	0	1
grade98	34787	5.676316	10.63853	0	88
old_girl98	33591	.1128278	.3163869	0	1
pill98	30834	.2258221	.4181294	0	1
pill99	30834	.2937018	.4554643	0	1
treat_sch98	34792	.3345309	.471833	0	1
infect_ea~99	2329	.4293688	.4950924	0	1
totpar98	27376	.7642412	.3297393	0	1

There are 34,792 pupils recorded in the sample, each has observations recorded for 9 variables (10 if you include pupil ID number), 52.08% of the students are male, 22.58% of students took the pill in 1998, but that is less than the 33.45% of students were assigned to the treatment schools in that year.

The difference in the proportion of students who were assigned to be treated may be different to the number of students who actually took the pill because some student (or their parents) chose not to take the treatment (parents could tell the school that they preferred their child didn't receive the treatment), absence from school on the day of treatment (this could be due to any number of other factors such as illness, other commitments, truancy), or non-eligibility (girls over the age of 13 were excluded from treatment because of concerns over birth defects if they were pregnant). 10/10

1.5 Summarise the infection outcomes

These students (group 1, below) were in the treatment group in 1998: 26.91% of these students had a moderate to heavy worm infection in the year after their treatment.

```
. summ  infect_early99 if  wgrp==1
```

Variable	Obs	Mean	Std. Dev.	Min	Max
infect_ea~99	862	.2691415	.4437711	0	1

The students in group 2 (shown below) did not receive the treatment in 1998. 52.39% of them had a moderate to heavy worm infection in 1999.

```
. summ  infect_early99 if  wgrp==2
```

Variable	Obs	Mean	Std. Dev.	Min	Max
infect_ea~99	1466	.5238745	.4996001	0	1

To test whether the outcomes of the two groups are statistically significant, we use a two-sample t test. This tests the differences in the means of 2 groups (1 and 2) for the same variable (infect_early99). The P statistic shows whether that difference is statistically significant, here P=0.00 indicates that the difference is significantly different at the 1% level.

```
. ttest  infect_early99 if  wgrp<3, by( wgrp)
```

Two-sample t test with equal variances

Group	Obs	Mean	Std. Err.	Std. Dev.	[95% Conf.	Interval]
1	862	.2691415	.0151149	.4437711	.2394752	.2988079
2	1466	.5238745	.0130483	.4996001	.498279	.5494699
combined	2328	.4295533	.0102617	.4951187	.4094303	.4496762
diff		-.254733	.0205889		-.2951075	-.2143584

```
    diff = mean(1) - mean(2)                                      t = -12.3723
Ho: diff = 0                     degrees of freedom =     2326

    Ha: diff < 0                 Ha: diff != 0                 Ha: diff > 0
 Pr(T < t) = 0.0000         Pr(|T| > |t|) = 0.0000          Pr(T > t) = 1.0000
```

The difference found with this test shows the difference between treatment and control groups after the intervention. Using the notation outlines above, this test is showing:

(A2+C2)-(B2+D2)

Note that it does not exclude the non-eligible students from the sample.

Is it a good estimate of the health effect of de-worming?

The estimate above captures a large part of the health effect of deworming. It includes the intra-school externality effect, but does not capture the externality effect across schools (inter-school externality).

The data available (without pre-treatment information about worm infections) does not allow for A1 and B1 to be recorded and considered. The total health effect of the treatment would need to include the externality effect of the treatment on other schools and exclude any time trend, i.e. changes in school attendance that may have occurred in the absence of the intervention.

There is no reason to suspect that there is any time trend (unless a large scale water and sanitation program took place), so the dominating effect from the above two biases would likely be the inter-school externality. Thus, the estimate provided by the estimation likely understates the total effect of the program.

10/10

1.6 Explain in word, how the authors proceed to resolve the above issue and so to estimate the overall health effect of de-worming program (Hint: explain how they incorporate cross-school externality into the overall estimate).

Randomization of deworming treatment across schools allows estimation of the overall effect of the programme by comparing treatment and comparison schools even if there are within-school externalities. Externalities also occur across schools because children from the same farm often attend di□erent schools.

They estimate cross-school externalities by taking advantage of variation in the local density of treatment schools introduced by randomization.

$$Y_{ijt} = \alpha + \beta_1 \text{Treatment(Year1)}_{it} + \beta_2 \text{Treatment(Year2)}_{it} + X'_{ijt}\delta$$
$$+ \sum_d (\gamma_d N^T_{dit}) + \sum_d (\phi_d N_{dit}) + u_i + e_{ijt}$$

Xijt is a vector of control variables (to increase statistical precision).
N^T_{dit} is the number of pupils randomly assigned to treatment.
N_{dit} is the number of pupils at distance d from school i and year t.
γ_d measures the extent of cross-school externalities
β_1 captures direct e□ect of deworming + within school externalities on untreated children in treated schools for year 1

$\beta_1 + \sum_d (\gamma_d \overline{N}^T_{dit})$ is the average e□ect of deworming treatment on overall infection prevalence in treatment schools in year 1 (including cross school externalities from other treated schools).

If the authors were after the total programme effect in treated schools and cross school externalities they could estimate the equation above.

To get a better estimation of the de-worming health effect, the authors did difference in difference method (the β1 in Point 1.2)

2. School Construction in Indonesia

Describing data:

STATA commands and output

```
. describe

Contains data from \\capstore01\users\redirections\u5255373\Desktop\supas.dta
  obs:'       59,938
 vars:            11                          12 Feb 2000 13:14
 size:     2,877,024 (94.5% of memory free)
-------------------------------------------------------------------------------
              storage  display     value
variable name   type   format      label      variable label
-------------------------------------------------------------------------------
ROB             float  %9.0g                  region of birth
YOB             float  %9.0g                  year of birth
yeduc           float  %9.0g                  years of education
lhwage          float  %9.0g                  log hourly wage
prog_int        float  %9.0g                  number of schools for 1000 children
high            float  %9.0g                  dummy high program intensity region
ch71            float  %9.0g                  number of children in 1971
old             float  %9.0g                  dummy born 1957-1962
young           float  %9.0g                  dummy born 1968-1972
veryold         float  %9.0g                  dummy born 1950-1956
intermed        float  %9.0g                  dummy born 1963-1967
-------------------------------------------------------------------------------
Sorted by:  high
```

2.1 Of the 59,938 observations in the entire sample:

a. the years of schooling range from 0 – 19, with an average of 9.52 years and a standard deviation of 4.03;

b. the ln of the hourly wage ranges from 3.53 to 11.96, with a mean of 6.93 and a standard deviation of 0.69;

c. 24.49 % of people were born between 1968 – 72 (aged 2-6 in 1974);

d. 26.77 % of people were born between 1957 – 62 (aged 12-17 in 1974);

e. 23.39 % of people were born between 1950 – 56 (aged 18-24 in 1974); and

f. 40.56% of people were in the high program intensity region.

As demonstrated by the two sample t-tests below, there is a difference in the number of schools per 1000 children between the high and low program intensity regions and this difference is highly statistically significant, at the 1%.

. sum

Variable	Obs	Mean	Std. Dev.	Min	Max
ROB	59938	3620.584	1733.41	1101	8171
YOB	59938	61.99011	6.317713	50	72
yeduc	59938	9.515599	4.034124	0	19
lhwage	59938	6.933525	.6878968	3.533587	11.95921
prog_int	59938	2.018584	1.097086	.5908243	8.598269
high	59938	.4056191	.4910156	0	1
ch71	59938	164886.6	111863.7	3796	542835
old	59938	.2676599	.442743	0	1
young	59938	.2448697	.4300135	0	1
veryold	59938	.2339417	.4233391	0	1
intermed	59938	.2535286	.4350345	0	1

Average years of schooling for whole sample: 9.5, standard deviation: 4.0. Average ln(hourly wage) for whole sample: 6.9, standard deviation: 0.7. Both are present for all of the 59,938 observations. 24.5% of individuals were "young", 26.8% were "old", and 23.4% were "very old". 40.6% of individuals sampled were in high-intensity regions.

In order to assess whether there is a difference between high vs. low intensity regions in the number of schools for every 1000 children, we cannot simply run "ttest prog_int, by(high)" in Stat since each region appears many times in the individual-level data. This command would therefore overestimate the sample size due to the duplication of regional data. It is instead necessary to create a new dataset containing one observation per region – the first few observations of this dataset (with obs = 280) are as follows:

ROB	prog_int	high	ch71
1101	2.729536	1	66678
1102	2.673724	0	36653
1103	2.367049	1	89563
1104	2.062135	0	29581
1105	2.455087	1	60283
1106	2.386285	0	47773
1107	2.64525	1	73339
1108	2.400766	1	129542

ch71 values have also been included in this data set as this field is also region-specific.

This produces the following results for the comparison of means t-test:

```
. ttest prog_int,by(high)

Two-sample t test with equal variances
```

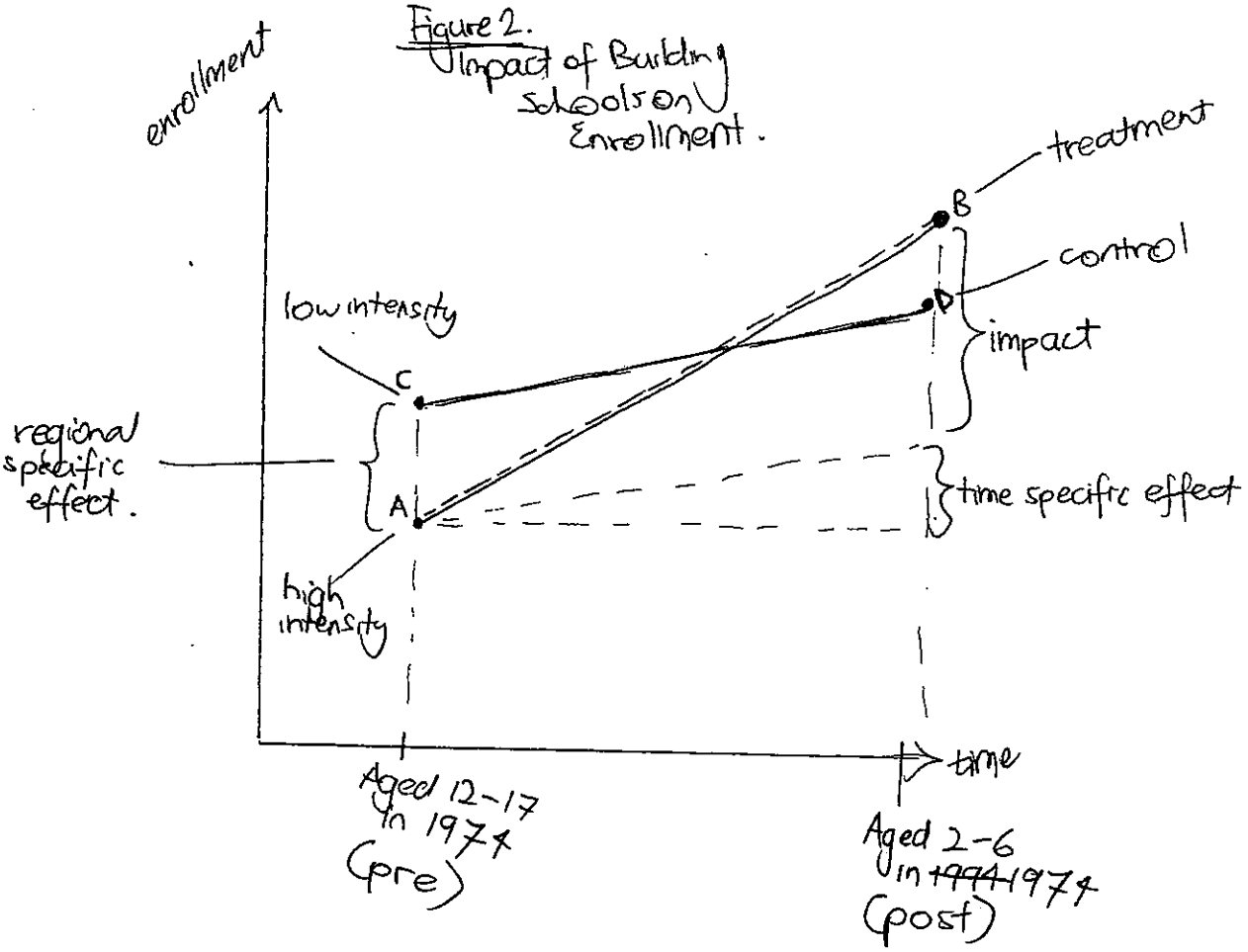
Group	Obs	Mean	Std. Err.	Std. Dev.	[95% Conf.	Interval]
0	150	1.884218	.0978176	1.198016	1.690929	2.077507
1	130	2.867817	.0999917	1.140081	2.669981	3.065653
combined	280	2.340889	.0758043	1.268449	2.191668	2.49011
diff		-.9835989	.1403783		-1.259938	-.7072595

```
    diff = mean(0) - mean(1)                                  t =  -7.0068
Ho: diff = 0                     degrees of freedom =      278

    Ha: diff < 0                 Ha: diff != 0                 Ha: diff > 0
 Pr(T < t) = 0.0000         Pr(|T| > |t|) = 0.0000          Pr(T > t) = 1.0000
```

The results are statistically significant (p = 0.0000), and indicate that there are more schools per 1000 children in high program intensity areas.

2.2 As the school-building program was targeted at areas of lower education, and we know that education and many other development indicators move together, the high treatment groups are expected to have very different observable characteristics to the low treatment (control) groups.

The impact is illustrated below in Figure 2:



Where:

A = Aged 12 to 17 in 1974 (old/pre-) and high intensity (treatment group)

B = Aged 2 – 6 in 1974 (young/post-) and high intensity (treatment group)

C = Aged 12 to 17 in 1974 (old) and low intensity (control group)

D = Aged 2 – 6 in 1974 (young) and low intensity (control group)

β = *Impact*
$= (B - A) - (D - C)$

2.2 First, estimate the impact of INPRES on schooling using simple difference-in-differences (DD) with cross-cohort (young versus old) and cross-regional (high-intensity versus low-intensity) variation.

(i) What are the mean years of education of the young cohort in the high- and low-program intensity regions; are these significantly different? What are the mean years of education of the old cohort in the high- and low-program intensity regions; are these significantly different? Interpret these results.

```
. ttest  yeduc if  young ==1, by( high )

Two-sample t test with equal variances
```

Group	Obs	Mean	Std. Err.	Std. Dev.	[95% Conf.	Interval]
0	8476	10.11892	.0363381	3.345479	10.04769	10.19016
1	6201	8.914207	.0442166	3.481897	8.827528	9.000887
combined	14677	9.609934	.0285211	3.455297	9.554029	9.665839
diff		1.204717	.0568793		1.093226	1.316207

```
    diff = mean(0) - mean(1)                                      t =  21.1802
Ho: diff = 0                     degrees of freedom =    14675

    Ha: diff < 0                 Ha: diff != 0                 Ha: diff > 0
 Pr(T < t) = 1.0000         Pr(|T| > |t|) = 0.0000          Pr(T > t) = 0.0000
```

The mean years of education of the young cohort in the high- and low-program intensity regions are 8.914 and 10.119. They are significantly different. The young cohort in high intensity region has lower mean years of education than that in low intensity region. This reflects the allocation rule that more schools were to be built in the region where enrollment rates were low.

```
. ttest  yeduc if  old==1, by( high )

Two-sample t test with equal variances
```

Group	Obs	Mean	Std. Err.	Std. Dev.	[95% Conf.	Interval]
0	9569	9.861114	.0421165	4.11989	9.778557	9.943671
1	6474	8.539234	.0541712	4.358678	8.43304	8.645427
combined	16043	9.327682	.0336908	4.267314	9.261644	9.39372
diff		1.32188	.0678761		1.188835	1.454925

```
    diff = mean(0) - mean(1)                                      t =  19.4749
Ho: diff = 0                     degrees of freedom =    16041

    Ha: diff < 0                 Ha: diff != 0                 Ha: diff > 0
 Pr(T < t) = 1.0000         Pr(|T| > |t|) = 0.0000          Pr(T > t) = 0.0000
```

The mean years of education of the old cohort in the high- and low-program intensity regions are 8.539 and 9.861. They are significantly different. The old cohort in high intensity region has lower mean years of education than that in low intensity region. This satisfied the allocation rule of the program

(ii) Are there significant differences between mean schooling of the young and old cohorts in each of the program intensity region? Are these increases in education between cohorts at the same rate in both program regions? What might be accounted for these differences across regions? (Hint: as the young cohort was exposed to the program, the larger increase in education between cohorts in the high-program intensity region should be a result of the program)

```
. gen young_old=young if young==1
(45261 missing values generated)

. replace young_old=2 if old==1
(16043 real changes made)

. ttest  yeduc if high==1, by( young_old)

Two-sample t test with equal variances
```

Group	Obs	Mean	Std. Err.	Std. Dev.	[95% Conf.	Interval]
1	6201	8.914207	.0442166	3.481897	8.827528	9.000887
2	6474	8.539234	.0541712	4.358678	8.43304	8.645427
combined	12675	8.722682	.0351596	3.958382	8.653764	8.791601
diff		.3749735	.0702593		.2372546	.5126925

```
    diff = mean(1) - mean(2)                                      t =   5.3370
Ho: diff = 0                     degrees of freedom =    12673

    Ha: diff < 0                 Ha: diff != 0                 Ha: diff > 0
 Pr(T < t) = 1.0000         Pr(|T| > |t|) = 0.0000          Pr(T > t) = 0.0000
```

The significant difference between cohorts in high intensity region is 0.37.

```
. ttest  yeduc if  high==0, by( young_old)

Two-sample t test with equal variances
```

Group	Obs	Mean	Std. Err.	Std. Dev.	[95% Conf.	Interval]
1	8476	10.11892	.0363381	3.345479	10.04769	10.19016
2	9569	9.861114	.0421165	4.11989	9.778557	9.943671
combined	18045	9.982211	.0281249	3.77806	9.927084	10.03734
diff		.25781	.0563221		.1474134	.3682066

```
    diff = mean(1) - mean(2)                                      t =   4.5774
Ho: diff = 0                     degrees of freedom =    18043

    Ha: diff < 0                 Ha: diff != 0                 Ha: diff > 0
 Pr(T < t) = 1.0000         Pr(|T| > |t|) = 0.0000          Pr(T > t) = 0.0000
```

In both high and low program intensity regions, mean number of years of education increased over time. It increased more in high intensity regions (0.37 years) than in low intensity regions (0.26 years). This is expected because the young cohort in the high intensity region could benefit much more from the greater number of schools built there (due to the program) than those in the low intensity region.

(iii) Bring these results together to construct a DD two-by-two grid (similar to panel A in Table 3 of the paper). What might be the effect of INPRES on education according to DD? Interpret the result, how large would years of schooling increase because of the program. You also learn from (ii) that the difference between the numbers of schools constructed per 1,000 children in high and low program regions is 1.04). What does this imply about an increase in years of schooling because of one additional school constructed for 1,000 children?

	Years of education		
	Level of program in region of birth		
	High	Low	Difference

Panel A:			
Age 2-6 in 1974	8.914	10.119	1.205
	(0.044)	(0.036)	(0.057)
Age 12-17 in 1974	8.539	9.861	1.322
	(0.054)	(0.042)	(0.068)
Difference	0.375	0.258	0.117
	(0.070)	(0.056)	(0.089)

In both cohorts, the average educational attainment in regions that received fewer schools is higher than in regions that received more schools. This reflects the fact that more schools were to be built in regions where enrollment rates were low. These statistic figures increased over time in both types of regions. However, it increased more in regions that received more schools. The difference in these differences can be interpreted as the casual effect of the program. An individual young enough, born in high program region, received on average 0.117 more years of education.

In addition, we have the difference between the number of schools constructed per 1000 children in high and low program regions is 1.048 (see the statistical table below). This suggests that one school per 1000 children contributed to an increase in education by 0.11 years (0.117 divided by 1.048).

(iv) Under what assumption would this estimate be an unbiased estimate of the INPRES impact? Redo the above estimates and construct a DD grid for the "control experiment" among old and veryold cohorts (who both were not exposed to the program). What do you find? Can these results confirm the identification assumption above?

"The difference in these differences can be interpreted as the casual effect of the program under the assumption that in the absence of the program, the increase in educational attainment would not have been systematically different in low and high program regions" (Duflo, 2001).

```
. ttest  yeduc if  veryold==1, by( high)

Two-sample t test with equal variances
```

Group	Obs	Mean	Std. Err.	Std. Dev.	[95% Conf.	Interval]
0	8431	9.59981	.0443433	4.071627	9.512886	9.686734
1	5591	8.288678	.0590032	4.411849	8.173009	8.404348
combined	14022	9.077022	.0359676	4.25909	9.00652	9.147523
diff		1.311132	.072621		1.168785	1.453479

```
    diff = mean(0) - mean(1)                                      t =  18.0544
Ho: diff = 0                     degrees of freedom =    14020

    Ha: diff < 0                 Ha: diff != 0                 Ha: diff > 0
 Pr(T < t) = 1.0000         Pr(|T| > |t|) = 0.0000          Pr(T > t) = 0.0000

. gen old_veryold=old if old==1
(43895 missing values generated)

. replace old_veryold=2 if  veryold==1
(14022 real changes made)

. ttest  yeduc if  high==1,by( old_veryold)

Two-sample t test with equal variances
```

Group	Obs	Mean	Std. Err.	Std. Dev.	[95% Conf.	Interval]
1	6474	8.539234	.0541712	4.358678	8.43304	8.645427
2	5591	8.288678	.0590032	4.411849	8.173009	8.404348
combined	12065	8.423125	.0399214	4.384997	8.344872	8.501377
diff		.2505556	.0800283		.0936873	.4074239

```
    diff = mean(1) - mean(2)                                      t =   3.1308
Ho: diff = 0                     degrees of freedom =    12063

    Ha: diff < 0                 Ha: diff != 0                 Ha: diff > 0
 Pr(T < t) = 0.9991         Pr(|T| > |t|) = 0.0017          Pr(T > t) = 0.0009

. ttest  yeduc if  high==0,by( old_veryold)

Two-sample t test with equal variances
```

Group	Obs	Mean	Std. Err.	Std. Dev.	[95% Conf.	Interval]
1	9569	9.861114	.0421165	4.11989	9.778557	9.943671
2	8431	9.59981	.0443433	4.071627	9.512886	9.686734
combined	18000	9.738722	.0305545	4.099316	9.678832	9.798612
diff		.2613038	.0612022		.1413416	.381266

```
    diff = mean(1) - mean(2)                                      t =   4.2695
Ho: diff = 0                     degrees of freedom =    17998

    Ha: diff < 0                 Ha: diff != 0                 Ha: diff > 0
 Pr(T < t) = 1.0000         Pr(|T| > |t|) = 0.0000          Pr(T > t) = 0.0000
```

	Years of education		
	Level of program in region of birth		
	High	Low	Difference
Panel B:			
Age 12-17 in 1974	8.539	9.861	1.322
	(0.054)	(0.042)	(0.068)
Age 18-24 in 1974	8.289	9.600	1.311
	(0.059)	(0.044)	(0.073)
Difference	0.250	0.261	0.011
	(0.080)	(0.061)	(0.099)

The difference in education between cohorts across the two regions is not significant. This implies that the program is the only systemic factor driving these cross-region differences and confirms the above assumption.

2.3

i. The DD effect of INPRES on years of schooling is then estimated using a simple linear regression as follows:

$$Yeduc_{ijt} = c + a_j + \beta_t + (P_j \times T_i)\,\varphi + (C_j \times T_i)\,\delta + \varepsilon_{ijt}$$

Where:

$Yeduc_{ijt}$ = *outcome variable; years of schooling*

C = *constant*

a_j = *region of birth fixed effect*

β_t = *cohort of birth fixed effect*

P_j = *intensity of exposure of the program region*

T_i = *dummy indicating whether the individual belongs to the exposed cohort*

φ = this coefficient could represent the impact of exposure to the program on an exposed cohort

C_j = *other control regional-specific variables (such as the allocation of the water and sanitation program, and the enrolment rate of the population)*

δ = this coefficient could capture the impact of regional specific variables on the exposed cohort, to improve the precision of φ.

ε_{ijt} = *error term*

(ii) . regress yeduc high young young_high if young==1|old==1

Source	SS	df	MS	Number of obs =	30720
				F(3, 30716) =	282.28
Model	12555.4369	3	4185.14564	Prob > F =	0.0000
Residual	455397.438	30716	14.8260658	R-squared =	0.0268
				Adj R-squared =	0.0267
Total	467952.875	30719	15.2333369	Root MSE =	3.8505

yeduc	Coef.	Std. Err.	t	P>|t|	[95% Conf.	Interval]
high	-1.32188	.0619635	-21.33	0.000	-1.443331	-1.200429
young	.25781	.0574332	4.49	0.000	.1452387	.3703814
young_high	.1171635	.0893285	1.31	0.190	-.057924	.292251
_cons	9.861114	.0393622	250.52	0.000	9.783962	9.938266

A young enough individual whose region of birth experienced high program intensity, received on average 0.12 more years of education on average due to the program. This is the causal impact of the program. However the coefficient is not statistically significant (small t statistic 1.31), even at the 10% significance level.

An old individual in a high program intensity region received 1.32 fewer years of education on average than in a low program intensity region. This is the pre-program difference between high and low intensity regions. The coefficient is statistically significant at the 1% significance level.

In a low intensity region, a young individual receives 0.26 more years of education on average than an old individual. This is the increase in education over time without high exposure to the program. The coefficient is statistically significant at the 1% significance level.

It should be noted that these estimates are only unbiased if there are no omitted time-varying and region-specific effects correlated with the program.

(iii)

```
. xtreg yeduc young young_high if young==1|old==1, i(ROB) fe

Fixed-effects (within) regression               Number of obs      =     30720
Group variable: ROB                             Number of groups   =       280

R-sq:  within  = 0.0020                         Obs per group: min =        16
       between = 0.1467                                        avg =     109.7
       overall = 0.0000                                        max =       635

                                                F(2,30438)         =     30.95
corr(u_i, Xb)  = -0.1228                        Prob > F           =    0.0000
```

yeduc	Coef.	Std. Err.	t	P>|t|	[95% Conf.	Interval]
young	.246888	.0549691	4.49	0.000	.1391462	.3546297
young_high	.173821	.0852243	2.04	0.041	.0067778	.3408643
_cons	9.309491	.0288556	322.62	0.000	9.252933	9.366049
sigma_u	1.3860061					
sigma_e	3.6336204					
rho	.12701575	(fraction of variance due to u_i)				

F test that all u_i=0: F(279, 30438) = 16.36 Prob > F = 0.0000

Including region of birth fixed effects rather than just two types of regions (high, low) allows for time-invariant region-specific factors to be better controlled for. Therefore, standard errors of all coefficients decrease compared to (ii). The estimate of the program impact (0.17) increases relative to (ii) and becomes statistically significant at the 5% significance level. This is an unbiased estimate of the program effect as long as there are no omitted time-varying effects or other region-specific effects correlated with the program.

```
. xtreg yeduc young_intensity young young_ch71 if young==1|old==1, i(ROB) fe

Fixed-effects (within) regression               Number of obs      =     30720
Group variable: ROB                             Number of groups   =       280

R-sq:  within  = 0.0035                         Obs per group: min =        16
       between = 0.1151                                        avg =     109.7
       overall = 0.0003                                        max =       635

                                                F(3,30437)         =     35.21
corr(u_i, Xb)  = -0.2178                        Prob > F           =    0.0000
```

yeduc	Coef.	Std. Err.	t	P>|t|	[95% Conf.	Interval]
young_inten~y	.1964597	.0426272	4.61	0.000	.1129087	.2800108
young	-.5455719	.1397112	-3.90	0.000	-.8194118	-.2717321
young_ch71	2.82e-06	4.23e-07	6.67	0.000	1.99e-06	3.65e-06
_cons	9.305287	.0288426	322.62	0.000	9.248754	9.361819
sigma_u	1.4206334					
sigma_e	3.6310775					
rho	.13275074	(fraction of variance due to u_i)				

F test that all u_i=0: F(279, 30437) = 17.57 Prob > F = 0.0000

Including region of birth fixed effects, the number of schools per 1000 children and a region control of the number of children in 1971 allows for better control of region-specific effects. Therefore, standard errors of all coefficients further decrease compared to (ii) and (i). The estimate of the program impact (0.20) further increases relative to (ii) and (i) and increases in statistical significance. It is statistically significant at the 1% significance level. This is an unbiased estimate of the program effect as long as there are no omitted time-varying effects or other region-specific effects correlated with the program.

2.4 Following the generation of dummy variables for each age group based on their year of birth, the effect of INPRES on years of schooling by age cohorts is estimated using a more generalised linear regression of the form:

$$Yeduc_{ijt} = c + a_j + \beta_t + \Sigma\,(P_j \times d_{il})\,\varphi + \Sigma\,(C_j \times d_{il})\,\delta + \varepsilon_{ijt}$$

Where:

Σ and d_{il} = dummy variables for age-specific cohort, with d_{il} representing the dummy variable whether individual I is age l in 1974, where l = 2....12 (in Σ)

and, as before:

$Yeduc_{ijt}$ = *outcome variable; years of schooling*
C = *constant*
a_j = *region of birth fixed effect*
β_t = *cohort of birth fixed effect*
P_j = *intensity of exposure of the program region*
T_i = *dummy indicating whether the individual belongs to the exposed cohort*
C_j = *other control regional-specific variables*
ε_{ijt} = *error term*

STATA commands

Generating dummies

```
. gen d62 = (YOB==62)
. gen d63 = (YOB==63)
. gen d64 = (YOB==64)
. gen d65 = (YOB==65)
. gen d66 = (YOB==66)
. gen d67 = (YOB==67)
. gen d68 = (YOB==68)
. gen d69 = (YOB==69)
. gen d70 = (YOB==70)
. gen d71 = (YOB==71)
. gen d72 = (YOB==72)
. gen d62_intensity=d62*prog_int
. gen d63_intensity=d63*prog_int
. gen d64_intensity=d64*prog_int
. gen d65_intensity=d65*prog_int
. gen d66_intensity=d66*prog_int
. gen d67_intensity=d67*prog_int
. gen d68_intensity=d68*prog_int
. gen d69_intensity=d69*prog_int
. gen d70_intensity=d70*prog_int
. gen d71_intensity=d71*prog_int
. gen d72_intensity=d72*prog_int
. gen d62_ch71=d62*ch71
. gen d63_ch71=d63*ch71
. gen d64_ch71=d64*ch71
. gen d65_ch71=d65*ch71
. gen d66_ch71=d66*ch71
. gen d67_ch71=d67*ch71
. gen d68_ch71=d68*ch71
. gen d69_ch71=d69*ch71
. gen d70_ch71=d70*ch71
. gen d71_ch71=d71*ch71
. gen d72_ch71=d72*ch71
```

Estimation

```
. xtreg yeduc d62_intensity d63_intensity d64_intensity d65_intensity d66_intensity
d67_intensity d68_intensity d69
> _intensity d70_intensity d71_intensity d72_intensity d62 d63 d64 d65 d66 d67 d68 d69
d70 d71 d72 d62_ch71 d63_ch7
> 1 d64_ch71 d65_ch71 d66_ch71 d67_ch71 d68_ch71 d69_ch71 d70_ch71 d71_ch71 d72_ch71,
i(ROB)fe

Fixed-effects (within) regression               Number of obs      =     59938
Group variable: ROB                             Number of groups   =       280

R-sq:  within  = 0.0116                         Obs per group: min =        35
       between = 0.0860                                        avg =     214.1
       overall = 0.0058                                        max =      1219

                                                F(33,59625)        =     21.26
corr(u_i, Xb)  = -0.0764                        Prob > F           =    0.0000

------------------------------------------------------------------------------
       yeduc |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
d62_intens~y |  -.0293967   .0771595    -0.38   0.703    -.1806296    .1218361
d63_intens~y |   .0135931   .0733083     0.19   0.853    -.1300913    .1572776
d64_intens~y |   .1004009   .0752573     1.33   0.182    -.0471036    .2479054
d65_intens~y |   .0675011    .065491     1.03   0.303    -.0608616    .1958638
d66_intens~y |    .179483   .0785186     2.29   0.022     .0255862    .3333798
d67_intens~y |   .1140695   .0720687     1.58   0.113    -.0271854    .2553244
d68_intens~y |   .2271345   .0698847     3.25   0.001     .0901602    .3641088
d69_intens~y |   .1304583   .0748622     1.74   0.081    -.0162718    .2771884
d70_intens~y |   .1933287    .068685     2.81   0.005     .0587058    .3279516
d71_intens~y |   .1501673   .0788419     1.90   0.057    -.0043631    .3046977
d72_intens~y |   .2073191   .0731715     2.83   0.005     .0639026    .3507356
         d62 |   .6057801   .2428372     2.49   0.013     .1298184    1.081742
         d63 |   .6278222   .2362194     2.66   0.008     .1648312    1.090813
         d64 |   .5976244   .2409197     2.48   0.013     .1254209    1.069828
         d65 |   .1064284   .2135537     0.50   0.618    -.3121377    .5249945
         d66 |   .4521036   .2501605     1.81   0.071    -.0382119     .942419
         d67 |   .4283851   .2389643     1.79   0.073    -.0399859    .8967561
         d68 |  -.0608854   .2369797    -0.26   0.797    -.5253665    .4035957
         d69 |   .0278174   .2471091     0.11   0.910    -.4565173    .5121521
         d70 |  -.5476832   .2324437    -2.36   0.018    -1.003274   -.0920927
         d71 |  -.2363503   .2620425    -0.90   0.367    -.7499546    .2772541
         d72 |  -.8587165   .2500345    -3.43   0.001    -1.348785   -.3686479
    d62_ch71 |   3.66e-07   7.57e-07     0.48   0.629    -1.12e-06    1.85e-06
    d63_ch71 |   4.40e-07   7.19e-07     0.61   0.541    -9.69e-07    1.85e-06
    d64_ch71 |   3.49e-08   7.23e-07     0.05   0.961    -1.38e-06    1.45e-06
    d65_ch71 |   1.47e-06   6.45e-07     2.28   0.023     2.07e-07    2.73e-06
    d66_ch71 |   1.67e-06   7.44e-07     2.24   0.025     2.08e-07    3.12e-06
    d67_ch71 |   2.77e-06   7.26e-07     3.81   0.000     1.34e-06    4.19e-06
    d68_ch71 |   2.68e-06   7.31e-07     3.66   0.000     1.24e-06    4.11e-06
    d69_ch71 |   2.23e-06   7.38e-07     3.02   0.003     7.79e-07    3.67e-06
    d70_ch71 |   2.21e-06   6.90e-07     3.20   0.001     8.55e-07    3.56e-06
    d71_ch71 |   2.34e-06   7.89e-07     2.96   0.003     7.89e-07    3.88e-06
    d72_ch71 |   4.13e-06   7.77e-07     5.31   0.000     2.60e-06    5.65e-06
       _cons |    9.15919   .0227839   402.00   0.000     9.114534    9.203847
-------------+----------------------------------------------------------------
     sigma_u |  1.4326517
     sigma_e |   3.746644
         rho |  .12756448   (fraction of variance due to u_i)
------------------------------------------------------------------------------
       F test that all u_i=0:     F(279, 59625) =    31.20          Prob > F = 0.0000
```

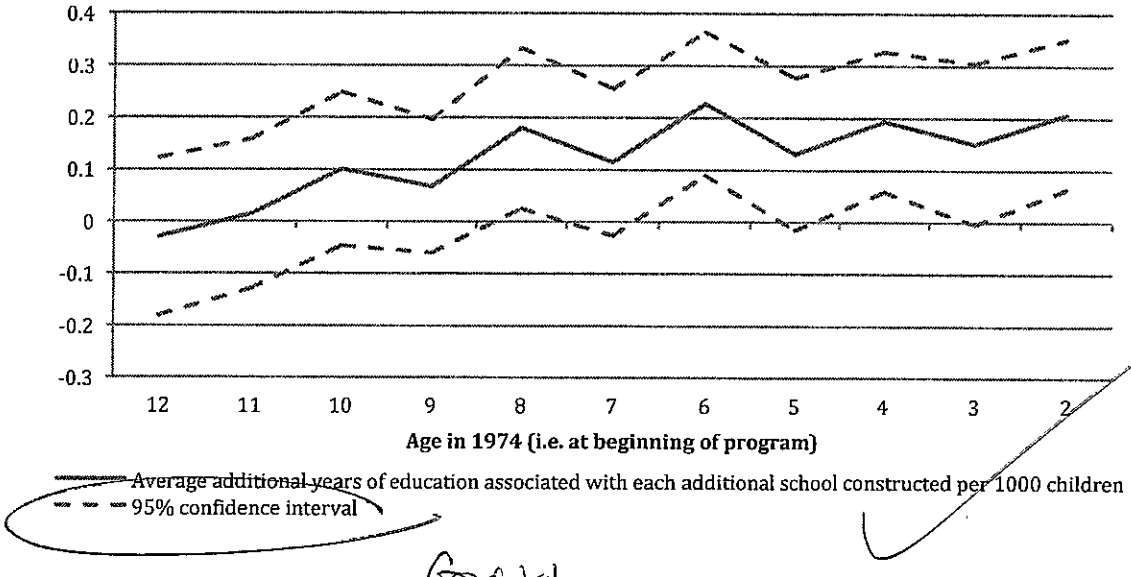
```
. test d62_intensity d63_intensity d64_intensity d65_intensity d66_intensity d67_intensity d68_intensi
> ty d69_intensity d70_intensity d71_intensity d72_intensity

 ( 1)  d62_intensity = 0
 ( 2)  d63_intensity = 0
 ( 3)  d64_intensity = 0
 ( 4)  d65_intensity = 0
 ( 5)  d66_intensity = 0
 ( 6)  d67_intensity = 0
 ( 7)  d68_intensity = 0
 ( 8)  d69_intensity = 0
 ( 9)  d70_intensity = 0
 (10)  d71_intensity = 0
 (11)  d72_intensity = 0

       F( 11, 59625) =    2.63
            Prob > F =    0.0023
```

The F-test of the joint significance of the interactions between program intensity and age cohort shows that they are in fact jointly significant, with p = 0.0023. This indicates that the program did in fact have an impact on years of education. For all cohorts except those born in 1962 this impact is estimated to be positive, although it is only statistically significant at the 5% level for a few.

One of these few is the 1972 cohort, who were 2 years old when the program started and thus stand to benefit the most from the program among the cohorts considered. With an average of 1.98 schools built per 1000 children as per the question, and since the co-efficient of d72_intensity is estimated to be 0.2073, we can estimate the total effect of the INPRES school-construction program on this cohort as increasing the years of schooling by 0.41 on average.

The below graph illustrates the values of the co-efficients of the dxx_intensity interaction variables across the cohorts.

Impact of INPRES school construction program
Average additional years of education associated with each additional school constructed per 1000 children



Good job...

Question 2.5.

```
. xtreg lhwage d62_intensity d63_intensity d64_intensity d65_intensity d66_intensity d67_intensity d68_in
> nsity d69_intensity d70_intensity d71_intensity d72_intensity d62 d63 d64 d65 d66 d67 d68 d69 d70 d71 d
>  d62_ch71 d63_ch71 d64_ch71 d65_ch71 d66_ch71 d67_ch71 d68_ch71 d69_ch71 d70_ch71 d71_ch71 d72_ch71, i(
> OB ) fe

Fixed-effects (within) regression               Number of obs      =     59938
Group variable: ROB                             Number of groups   =       280

R-sq:  within  = 0.0527                         Obs per group: min =        35
       between = 0.0278                                        avg =     214.1
       overall = 0.0417                                        max =      1219

corr(u_i, Xb)  = -0.0668                        F(33,59625)        =    100.52
                                                Prob > F           =    0.0000
```

lhwage	Coef.	Std. Err.	t	P>|t|	[95% Conf.	Interval]
d62_intensity	.0167573	.0131777	1.27	0.204	-.0090709	.0425856
d63_intensity	-.0109646	.0125199	-0.88	0.381	-.0355037	.0135745
d64_intensity	.0025206	.0128528	0.20	0.845	-.0226709	.0277121
d65_intensity	.0078741	.0111849	0.70	0.481	-.0140482	.0297965
d66_intensity	.0184183	.0134098	1.37	0.170	-.0078649	.0447015
d67_intensity	-.011231	.0123082	-0.91	0.362	-.0353551	.0128932
d68_intensity	.0105849	.0119352	0.89	0.375	-.0128082	.033978
d69_intensity	.0200868	.0127853	1.57	0.116	-.0049725	.045146
d70_intensity	.0183333	.0117303	1.56	0.118	-.0046582	.0413249
d71_intensity	.0088371	.013465	0.66	0.512	-.0175543	.0352286
d72_intensity	.0158472	.0124966	1.27	0.205	-.0086462	.0403405
d62	-.1541639	.0414729	-3.72	0.000	-.2354508	-.0728769
d63	-.1096652	.0403427	-2.72	0.007	-.188737	-.0305935
d64	-.193306	.0411454	-4.70	0.000	-.2739511	-.1126608
d65	-.2999463	.0364717	-8.22	0.000	-.371431	-.2284616
d66	-.3248363	.0427236	-7.60	0.000	-.4085747	-.241098
d67	-.2973862	.0408114	-7.29	0.000	-.3773768	-.2173956
d68	-.4471452	.0404725	-11.05	0.000	-.5264714	-.3678189
d69	-.4735488	.0422024	-11.22	0.000	-.5562658	-.3908319
d70	-.4818379	.0396978	-12.14	0.000	-.5596457	-.40403
d71	-.5375945	.0447528	-12.01	0.000	-.6253102	-.4498787
d72	-.606798	.0427021	-14.21	0.000	-.6904942	-.5231018
d62_ch71	3.87e-07	1.29e-07	2.99	0.003	1.34e-07	6.41e-07
d63_ch71	2.46e-07	1.23e-07	2.00	0.045	5.17e-09	4.86e-07
d64_ch71	4.26e-07	1.23e-07	3.45	0.001	1.84e-07	6.68e-07
d65_ch71	5.63e-07	1.10e-07	5.11	0.000	3.47e-07	7.78e-07
d66_ch71	6.24e-07	1.27e-07	4.92	0.000	3.75e-07	8.73e-07
d67_ch71	5.99e-07	1.24e-07	4.83	0.000	3.56e-07	8.42e-07
d68_ch71	8.69e-07	1.25e-07	6.96	0.000	6.24e-07	1.11e-06
d69_ch71	7.64e-07	1.26e-07	6.06	0.000	5.17e-07	1.01e-06
d70_ch71	5.30e-07	1.18e-07	4.49	0.000	2.99e-07	7.61e-07
d71_ch71	6.96e-07	1.35e-07	5.17	0.000	4.32e-07	9.61e-07
d72_ch71	8.24e-07	1.33e-07	6.21	0.000	5.64e-07	1.08e-06
_cons	7.063692	.0038911	1815.32	0.000	7.056065	7.071318

sigma_u	.20868514	
sigma_e	.6398693	
rho	.09613951	(fraction of variance due to u_i)

F test that all u_i=0: F(279, 59625) = 23.36 Prob > F = 0.0000

The estimates above for the effects of the program on wage are varied; some are negative, and none are individually statistically significant at the 5% level. This begs the question of whether the program's overall impact on wage was statistically significant, so I ran the same F-test as in the previous question, i.e.

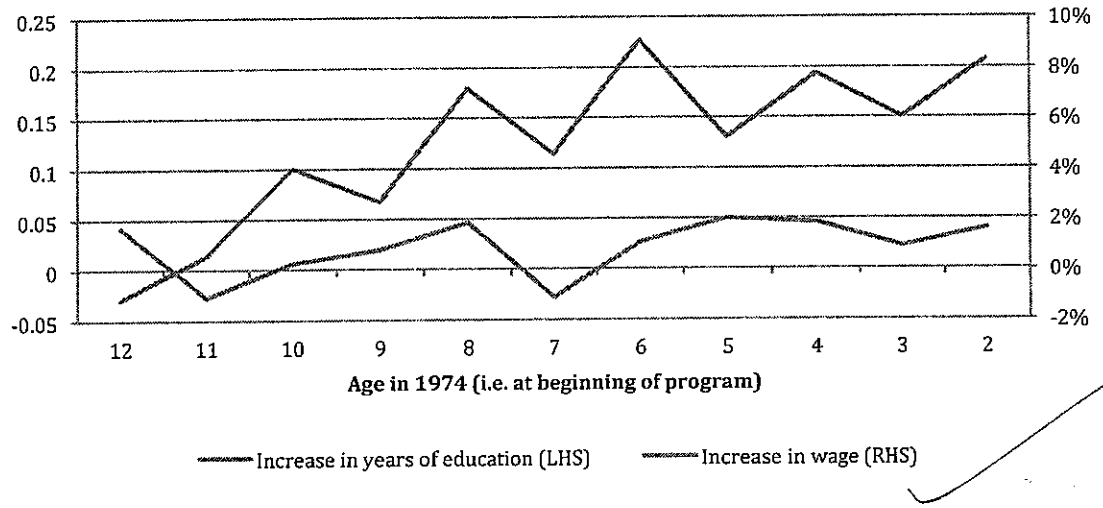
testing the restriction that the co-efficients of the dxx_intensity terms were all equal to zero. The results are below.

```
. test d62_intensity d63_intensity d64_intensity d65_intensity d66_intensity d67_intensity d68_intensity
> 9_intensity d70_intensity d71_intensity d72_intensity

 ( 1)  d62_intensity = 0
 ( 2)  d63_intensity = 0
 ( 3)  d64_intensity = 0
 ( 4)  d65_intensity = 0
 ( 5)  d66_intensity = 0
 ( 6)  d67_intensity = 0
 ( 7)  d68_intensity = 0
 ( 8)  d69_intensity = 0
 ( 9)  d70_intensity = 0
 (10)  d71_intensity = 0
 (11)  d72_intensity = 0

       F( 11, 59625) =    1.05
            Prob > F =    0.3973
```

The results of this test indicate that the dxx_intensity terms are not jointly significant at the 5% level (p = 0.3973). In other words, the program did not have a statistically significant impact on wage.

The values of the dxx_intensity co-efficients are plotted below for both (yeduc and lhwage) regressions.

Impact of INPRES school construction program
Average increase in years of education and wage associated with each additional school constructed per 1000 children



10/10

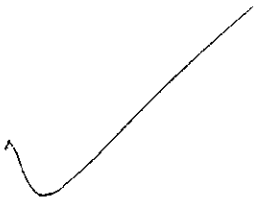
2.6 (i)

$$lnw_i = \alpha_i + bS_i$$

A one-year increase in years of schooling is associated with a b% increase in wages on average. This OLS estimate is not a good estimate of the returns to education because there are several problems are likely to cause an overestimation:

- Endogeneity bias
    - Omitted variable bias: For example, unobserved child ability is correlated with a child's years of schooling because smarter students may find school less difficult and choose to obtain more schooling to signal their high ability, while ability is also likely to have a direct effect on income. Parents' intellectual capacities can also be an omitted variable.
    - Reverse causality: Higher income is likely to result in more schooling
    - Measurement errors
- Data quality: Difficulties in measuring schooling
- Specification errors: A linear return to education may not be realistic

(ii) From 2.4, the program exposure instruments are jointly significantly different from 0 at the 1% significance level (F-statistic is 2.63). They are also mostly individually significant. This indicates that the instruments have high correlation with schooling. From 2.5, there is no evidence of joint significance of the exposure variables even at the 10% significance level (small F-statistic 1.05). In addition, the graph from 2.5 suggests that the program impacts on wages and education across cohorts track each other. This suggests that the instruments have little or no direct effect on log wages, but affect log wages via education. Therefore program exposure is likely to be a good instrument.



(iii) First stage from 2.4:

```
. xtreg yeduc d62_intensity d63_intensity d64_intensity d65_intensity d66_intens
> ity d67_intensity d68_intensity d69_intensity d70_intensity d71_intensity d72_
> intensity d62 d63 d64 d65 d66 d67 d68 d69 d70 d71 d72 d62_ch71 d63_ch71 d64_ch
> 71 d65_ch71 d66_ch71 d67_ch71 d68_ch71 d69_ch71 d70_ch71 d71_ch71 d72_ch71, i(
>  ROB ) fe

Fixed-effects (within) regression               Number of obs      =     59938
Group variable: ROB                             Number of groups   =       280

R-sq:  within  = 0.0116                         Obs per group: min =        35
       between = 0.0860                                        avg =     214.1
       overall = 0.0058                                        max =      1219

                                                F(33,59625)        =     21.26
corr(u_i, Xb)  = -0.0764                        Prob > F           =    0.0000
```

yeduc	Coef.	Std. Err.	t	P>|t|	[95% Conf.	Interval]
d62_intensity	-.0293967	.0771595	-0.38	0.703	-.1806296	.1218361
d63_intensity	.0135931	.0733083	0.19	0.853	-.1300913	.1572776
d64_intensity	.1004009	.0752573	1.33	0.182	-.0471036	.2479054
d65_intensity	.0675011	.065491	1.03	0.303	-.0608616	.1958638
d66_intensity	.179483	.0785186	2.29	0.022	.0255862	.3333798
d67_intensity	.1140695	.0720687	1.58	0.113	-.0271854	.2553244
d68_intensity	.2271345	.0698847	3.25	0.001	.0901602	.3641088
d69_intensity	.1304583	.0748622	1.74	0.081	-.0162718	.2771884
d70_intensity	.1933287	.068685	2.81	0.005	.0587058	.3279516
d71_intensity	.1501673	.0788419	1.90	0.057	-.0043631	.3046977
d72_intensity	.2073191	.0731715	2.83	0.005	.0639026	.3507356
d62	.6057801	.2428372	2.49	0.013	.1298184	1.081742
d63	.6278222	.2362194	2.66	0.008	.1648312	1.090813
d64	.5976244	.2409197	2.48	0.013	.1254209	1.069828
d65	.1064284	.2135537	0.50	0.618	-.3121377	.5249945
d66	.4521036	.2501605	1.81	0.071	-.0382119	.942419
d67	.4283851	.2389643	1.79	0.073	-.0399859	.8967561
d68	-.0608854	.2369797	-0.26	0.797	-.5253665	.4035957
d69	.0278174	.2471091	0.11	0.910	-.4565173	.5121521
d70	-.5476832	.2324437	-2.36	0.018	-1.003274	-.0920927
d71	-.2363503	.2620425	-0.90	0.367	-.7499546	.2772541
d72	-.8587165	.2500345	-3.43	0.001	-1.348785	-.3686479
d62_ch71	3.66e-07	7.57e-07	0.48	0.629	-1.12e-06	1.85e-06
d63_ch71	4.40e-07	7.19e-07	0.61	0.541	-9.69e-07	1.85e-06
d64_ch71	3.49e-08	7.23e-07	0.05	0.961	-1.38e-06	1.45e-06
d65_ch71	1.47e-06	6.45e-07	2.28	0.023	2.07e-07	2.73e-06
d66_ch71	1.67e-06	7.44e-07	2.24	0.025	2.08e-07	3.12e-06
d67_ch71	2.77e-06	7.26e-07	3.81	0.000	1.34e-06	4.19e-06
d68_ch71	2.68e-06	7.31e-07	3.66	0.000	1.24e-06	4.11e-06
d69_ch71	2.23e-06	7.38e-07	3.02	0.003	7.79e-07	3.67e-06
d70_ch71	2.21e-06	6.90e-07	3.20	0.001	8.55e-07	3.56e-06
d71_ch71	2.34e-06	7.89e-07	2.96	0.003	7.89e-07	3.88e-06
d72_ch71	4.13e-06	7.77e-07	5.31	0.000	2.60e-06	5.65e-06
_cons	9.15919	.0227839	402.00	0.000	9.114534	9.203847
sigma_u	1.4326517					
sigma_e	3.746644					
rho	.12756448	(fraction of variance due to u_i)				

F test that all u_i=0: F(279, 59625) = 31.20 Prob > F = 0.0000

Second stage:

```
. xtreg lhwage pdt_yeduc d62 d63 d64 d65 d66 d67 d68 d69 d70 d71 d72 d62_ch71 d63_ch
> 71 d64_ch71 d65_ch71 d66_ch71 d67_ch71 d68_ch71 d69_ch71 d70_ch71 d71_ch71 d72_ch7
> 1, i( ROB ) fe

Fixed-effects (within) regression               Number of obs      =     59938
Group variable: ROB                             Number of groups   =       280

R-sq:  within  = 0.0526                         Obs per group: min =        35
       between = 0.0270                                        avg =     214.1
       overall = 0.0417                                        max =      1219

corr(u_i, Xb)  = -0.0659                        F(23,59635)        =    143.89
                                                Prob > F           =    0.0000
```

lhwage	Coef.	Std. Err.	t	P>|t|	[95% Conf.	Interval]
pdt_yeduc	.0630352	.0317407	1.99	0.047	.0008233	.1252472
d62	-.1434002	.0282569	-5.07	0.000	-.1987839	-.0880165
d63	-.1813136	.0302164	-6.00	0.000	-.240538	-.1220893
d64	-.2411569	.0356571	-6.76	0.000	-.311045	-.1712689
d65	-.2967864	.02198	-13.50	0.000	-.3398673	-.2537056
d66	-.3341656	.037427	-8.93	0.000	-.4075226	-.2608085
d67	-.3758223	.0323265	-11.63	0.000	-.4391825	-.3124622
d68	-.4540055	.0286753	-15.83	0.000	-.5102092	-.3978019
d69	-.4419755	.025486	-17.34	0.000	-.4919282	-.3920227
d70	-.4296395	.0212049	-20.26	0.000	-.4712011	-.3880779
d71	-.5245893	.0249035	-21.06	0.000	-.5734003	-.4757784
d72	-.5447683	.0250211	-21.77	0.000	-.5938098	-.4957268
d62_ch71	2.89e-07	1.19e-07	2.43	0.015	5.62e-08	5.22e-07
d63_ch71	2.68e-07	1.11e-07	2.41	0.016	5.00e-08	4.86e-07
d64_ch71	4.39e-07	1.13e-07	3.89	0.000	2.18e-07	6.60e-07
d65_ch71	4.54e-07	1.06e-07	4.29	0.000	2.47e-07	6.62e-07
d66_ch71	4.89e-07	1.18e-07	4.16	0.000	2.59e-07	7.20e-07
d67_ch71	5.11e-07	1.32e-07	3.87	0.000	2.52e-07	7.69e-07
d68_ch71	7.19e-07	1.22e-07	5.89	0.000	4.80e-07	9.58e-07
d69_ch71	5.68e-07	1.23e-07	4.61	0.000	3.26e-07	8.09e-07
d70_ch71	3.61e-07	1.12e-07	3.21	0.001	1.41e-07	5.82e-07
d71_ch71	5.53e-07	1.31e-07	4.21	0.000	2.95e-07	8.10e-07
d72_ch71	5.51e-07	1.55e-07	3.55	0.000	2.46e-07	8.56e-07
_cons	6.486339	.2907645	22.31	0.000	5.916439	7.056238
sigma_u	.20863272					
sigma_e	.63985654					
rho	.09609933	(fraction of variance due to u_i)				

F test that all u_i=0: F(279, 59635) = 23.47 Prob > F = 0.0000

A one-year increase in the number of years of schooling results in a 6.30% increase in hourly wages on average. This represents the economic returns to each year of schooling. The coefficient is statistically significant and depends on the validity of the program exposure instruments used.