Assignment 1: What is the impact of VSL on school absenteeism?

a. Pre-treatment effect analysis
i.

In this dataset, we consider absenteeism for children in school. On average, the children of the households in this dataset missed 16.45763 days of school every month before the implementation of the program and its promotion.

To counter this, a government implemented a Village Savings and Loan Program, accompanied by a randomized control trial where they led a campaign promoting the VSL program in randomly selected villages. In this case, 51.21% of the households were part of the treatment group (in villages where the VSL program was promoted) and 48.79% were in the control group (where the program was not promoted).

Overall, by the end of the campaign 12.75% of the individuals were participating in the program and it was found that all of them were part of the treatment group. This suggested that none of the households in the control group participated in the microcredit program at the endline.

In addition, when considering whether the households participating in the program were pro-government or anti-government, we find that all the households who decided to participate in the program were pro-government. This helped conclude that the effect of the program on absenteeism might also be influenced by the effect of being pro-government on absenteeism.

ii.

EXPERIMENTAL CONDITIONS	CONTROL	TREATMENT	DIFFERENCE
age_hh	46.46562 [14.53959]	45.86179 [14.36723]	-0.60383 [0.21982]
age_sp	40.83357 [12.72111]	40.55736 [12.80636]	-0.27621 [0.19413]
educ_hh	2.887169 [2.803986]	3.06289 [2.87855]	0.175721 [0.04325]
educ_sp	2.620673 [2.549188]	2.70551 [2.674839]	0.084837 [0.03976]
female_hh (binary)	0.077762 [0.2678602]	0.0745257 [0.2626839]	-0.0032363 [0.00403]
indigenous (binary)	0.350403 [0.4772093]	0.3690154 [0.4826471]	0.0186124 [0.00730]
dirtfloor (binary)	0.5865339 [0.4925717]	0.6124661 [0.4872973]	0.0259322 [0.00745]
bathroom (binary)	0.6230441 [0.4847387]	0.6350497 [0.481525]	0.0120056 [0.00735]
land (binary)	2.095306 [2.957626]	2.057362 [3.345568]	-0.037944 [0.04809]
progov (binary)	0.4495021 [0.4975614]	0.5045167 [0.5000926]	0.0550146 [0.00760]

b. Regression analysis and presentation of results
i.

Source	SS	df	MS			
				Number of obs	=	2,161
				F(1, 2159)	=	161.51
Model	4632.61404	1	4632.61404	Prob > F	=	0.0000
Residual	61925.978	2,159	28.6827133	R-squared	=	0.0696
				Adj R-squared	=	0.0692
Total	66558.592	2,160	30.814163	Root MSE	=	5.3556

days	Coef.	Std. Err.	t	P>|t|	[95% Conf. Interval]	
treatment	-2.92873	.2304498	-12.71	0.000	-3.380657	-2.476803
_cons	15.73978	.1643417	95.77	0.000	15.4175	16.06207

$$\widehat{D}_i = 15.74 - \underset{(0.23)}{2.92}\, treatment_i$$
$$\phantom{\widehat{D}_i = }{(0.16)}$$

(1)

Note: this table shows the results of regressions examining the change in average number of days children in households missed schools in the past 30 days, depending on whether the child was in the treatment group. D^ is the predicted number of days absent.

ii.

Source	SS	df	MS			
				Number of obs	=	1,032
				F(1, 1030)	=	357.98
Model	9371.9357	1	9371.9357	Prob > F	=	0.0000
Residual	26965.507	1,030	26.1801039	R-squared	=	0.2579
				Adj R-squared	=	0.2572
Total	36337.4427	1,031	35.2448523	Root MSE	=	5.1166

days	Coef.	Std. Err.	t	P>|t|	[95% Conf. Interval]	
treatment	-6.040963	.319284	-18.92	0.000	-6.667484	-5.414441
_cons	15.63983	.2332991	67.04	0.000	15.18203	16.09762

(2)
$$\widehat{D}_i = 15.64 - \underset{(0.32)}{6.04}\, treatment_i$$
$$\phantom{\widehat{D}_i = }{(0.23)}$$

Note: this table shows the results of regressions examining the change in average numbers of days children in households missed schools in the past 30 days among households that are pro-government. It calculates the effect on days if the child was part of the treatment or not .

iii.

Source	SS	df	MS			
Model	10777.1312	2	5388.56559			
Residual	55781.4609	2,158	25.8486844			
Total	66558.592	2,160	30.814163			

			Number of obs	=	2,161
			F(2, 2158)	=	208.47
			Prob > F	=	0.0000
			R-squared	=	0.1619
			Adj R-squared	=	0.1611
			Root MSE	=	5.0842

days	Coef.	Std. Err.	t	P>|t|	[95% Conf. Interval]	
treatment	-2.764991	.2190265	-12.62	0.000	-3.194516	-2.335466
progov	-3.379836	.2192153	-15.42	0.000	-3.809731	-2.94994
_cons	17.27057	.1849256	93.39	0.000	16.90792	17.63322

(3) $\widehat{D}_i = 17.27 - \underset{(0.18)}{} \underset{(0.22)}{2.76}\, treatment_i - \underset{(0.22)}{3.38}\, progov_i$

Source	SS	df	MS			
Model	17175.2862	12	1431.27385			
Residual	49383.3059	2,148	22.9903659			
Total	66558.592	2,160	30.814163			

			Number of obs	=	2,161
			F(12, 2148)	=	62.26
			Prob > F	=	0.0000
			R-squared	=	0.2580
			Adj R-squared	=	0.2539
			Root MSE	=	4.7948

days	Coefficient	Std. err.	t	P>|t|	[95% conf. interval]	
treatment	.279567	.2857831	0.98	0.328	-.2808734	.8400074
participation	-6.354367	.414164	-15.34	0.000	-7.166571	-5.542162
progov	.1128626	.3119205	0.36	0.718	-.498835	.7245602
age_hh	.0195598	.0134372	1.46	0.146	-.0067915	.0459111
age_sp	-.0112076	.0148881	-0.75	0.452	-.0404042	.017989
educ_hh	-.0118987	.0455727	-0.26	0.794	-.1012699	.0774724
educ_sp	.0237019	.0498443	0.48	0.634	-.0740462	.12145
female_hh	1.069567	.403428	2.65	0.008	.2784168	1.860717
indigenous	-.9457889	.2266839	-4.17	0.000	-1.390332	-.5012461
dirtfloor	-.5459723	.2284214	-2.39	0.017	-.9939224	-.0980221
bathroom	-.0855203	.2176394	-0.39	0.694	-.5123261	.3412856
land	-.0756914	.0328616	-2.30	0.021	-.1401353	-.0112475
_cons	15.99262	.6278872	25.47	0.000	14.76129	17.22395

.

Note: this table provides the results of regressions examining the changes in average numbers of days children in households missed schools in the past 30 days whether the child is part of the treatment group and whether they are in a pro-government household.

(4)

$\hat{D}_i = \underset{(0.62)}{15.99} + \underset{(0.29)}{0.28}\, treatment_i - \underset{(0.41)}{6.35}\, participation_i + \underset{(0.31)}{0.11}\, progov_i + \underset{(0.01)}{0.02}\, age\_hh_i - \underset{(0.01)}{0.01}\, age\_sp_i - \underset{(0.05)}{0.01}\, educ\_hh_i + \underset{(0.05)}{0.02}\, educ\_sp_i$

$+ \underset{(0.40)}{1.06}\, female\_hh_i - \underset{(0.23)}{0.95}\, indigenous_i - \underset{(0.22)}{0.55}\, dirtfloor_i - \underset{(0.22)}{0.09}\, bathroom_i - \underset{(0.03)}{0.08}\, land_i$

Note: this table provides the results of regressions examining the changes in average numbers of days children in households missed schools in the past 30 days, controlling for whether the child was part of the treatment, whether they participated in the VSL program, and controlling for other variables like whether the household is pro government, the age of head of household and spouse, education of the head of household and spouse, whether the head of the household is a woman or if they speak an indigenous language, wealth (as shown by 'dirtfloor', 'bathroom' and 'land').

iv.

Dependent Variable	(1) Days	(2) Days	(3) Days	(4) Days
Treatment (Endline)	-2.93873 [0.2304498]		-2.764991 [0.2190265]	0.279567 [0.2857831]
Treatment* Endline* Progov		-6.040963 [0.319284]		
Progov			-3.379836 [0.2192153]	0.1128626 [0.3119205]
Participation				-6.354367 [0.414164]
Age_hh				0.0195598 [0.0134372]
Age_sp				-0.0112076 [0.0148881]
Educ_hh				-0.0118987 [0.0455727]
Educ_sp				0.0237019 [0.0498443]
female_hh				1.069567 [0.403428]
Indigenous				-0.9457889 [0.2266839]
Dirtfloor				-0.5459723 [0.2284214]
Bathroom				-0.0855203 [0.2176394]
Land				-0.0756914 [0.0328616]
Mean of Dependent Variable	16.44976	16.44976	16.44976	16.44976
Observations	4323	4323	4323	4323
R-squared	0.0696	0.2579	0.1619	0.2580

c. Discussion of results and further analysis

i.

Firstly, we found that the VSL program had a causal impact on absenteeism. This was evident in the regressions run in which we analyzed the relationship between days absent and the independent variables in both treatment and control groups.

When we look at absenteeism more precisely, we see that at the endline, households who did not participate in the program still missed 15.84225 days of school per month on average. On the other hand,we found that the households that participated in the program then missed only 9.5988865 days of school per month on average.

The first regression shows that the treatment has an impact on absenteeism, whilst the second and third regression indicate that this effect is increased when the households are pro-government. However, when we control for participation (evident in the fourth regression), those effects disappear. As a result, we found that only participation in the program impacts absenteeism, by reducing the days missed by 6 on average.

As participation in the program can be impacted by other variables such as the households being pro-government or other, we used the treatment as an instrumental variable to isolate the effect of the program, outside of other variables.

ii.

We would like to have a variable showing the number of children of schooling age per household. We believe there will be a positive relationship between the number of children in a household of school age and school absenteeism. We also believe there to be greater program participation for larger households. Given income is controlled for, larger Households with more children of schooling age will face a greater burden. A larger number of children of schooling age will mean that the costs associated with schooling will represent a larger percentage of household income. This assumption was highlighted by (Romero & Lee, 2008), with larger families having "[chronically absent]" children in school. Therefore we expect that larger households will be more likely to participate in the program. In addition, we anticipate that the program effect may be even larger for large households as opposed to small households on school absenteeism. This is due to our hypothesis that children in larger families may be more prone to school absenteeism due to a lower amount of income available per child; holding the assumption that income is spread evenly per child.

Furthermore, for children whose Family participated in the Programme we would also like to measure years of completed education and adult wages. These variables would help extend the analysis, as we

expect there to be an inverse relationship between school absenteeism and years of completed education. + adult wages. Did the children of families in the program continue school for a longer time? Did the program have an impact on their wages? If true, the program may also prove viable in being a solution to increasing years of completed education, and increasing adult wages.

Reference:

- Romero, M., & Lee, Y.-S. (2008). *The Influence of Maternal and Family Risk on Chronic Absenteeism in Early Schooling.* New York, NY: Columbia Univ. National Center for Children in Poverty.