ECOM90003: Applied Microeconometric Modelling – Assignment 1

Q	Answer
1	The author uses five measures of public goods in villages for this paper. They are the presence of: a primary health centre (a *Puskesmas*), at least one doctor, access to safe drinking water, community based healthcare facilities (*Posyandus*) and a garbage disposal system. I would expect the provision of all these public goods to be positively impacted by the mass school construction program, but for different reasons. For the health-related public goods, I would expect the impact to be driven significantly by both the general impact of labour market augmentation as well as improved public governance. This is because the labour required to provide to provide these public goods is quite specialised, where labour supply would greatly benefit from improved education. Similarly, improving the education of village leaders should also make public governance more aware of the developmental benefits of these public goods. Given I would not expect the specialisation of labour to be a limit to administering access to safe drinking water and a garbage disposal system, I assume their increased provision would be largely driven by improving local governance.
2	By simply comparing average public good provisions before and after the first election in areas with a positive amount of INPRES schools we are lacking a control group for our analysis. Without knowledge of how those villages without any INPRES schools change their public good provision we lack an effective counterfactual of the policy, that is we can't truly estimate what would have occurred without the treatment.
3	Because, to derive accurate causal impacts, we need information about outcomes pre- and post-treatment for each village. If we simply compare outcomes for villages that did and did not receive the treatment and estimate causal effects we are effectively assuming that all villages are completely homogenous, and there is no reason apart from the intervention that different outcomes may arise. This is clearly an infeasible assumption.
4	A mass education intervention improves the level of education in the entire labour force as well as the education level of village heads. Therefore, the key challenge is disentangling to impact of changes in local governance (village head education levels) from this general labour market augmenting effect.
5	
5a	The author's overall strategy is to first show if increases in public good provisions corresponds with when wholly INPRES educated cohorts start contesting elections. She then seeks quantify the overall impact of the INPRES program on public good provision. Equation 1 seeks to establish if the increase in public good provision corresponds with the first village election post-1992, which is the first year wholly INPRES educated cohorts can contest elections. It is estimated for three different groups of villages, where grouping is determined by when after 1992 the election occurs (groups are limited by the fact census data is published triennially). Equation 2 the second equation seeks to establish the magnitude of the overall effect of this intervention on public good provision in each village.
5b	In Equation 1: γ_s. In Equation 2: λ_2.

Joshua Copeland (SID:1444772)

5c	In Equation 1, γ_s measures the impact of improved public governance on public good provision for each village group. In Equation 2, λ_2 measures the impact of improved public governance on overall public good provision. The purpose of the prior parameter is to demonstrate a descriptive relationship between public good provision and improved public governance. The latter's purpose is to estimate the overall effect of improved public governance on public good provision.
6	It needs to be shown that the first village election post-1992 and its interaction with the school construction program as quasi-random (i.e. as good as random). She considers this assumption holds given that after testing exogeneity with 50 different pairwise correlations, only 3 of them are significant at the 10 per cent level.
7	No, given this is panel data set, it is highly unlikely errors will be iid. This is primarily because of the unobserved characteristics in each village independent of the intervention which would impact the provision of public goods over time. These unobserved characteristics are likely to reflect common geographical and cultural difference common across Java. Therefore, it makes sense to cluster the standard errors at the group level. In this dataset, that corresponds with the district level variable.
8	
8a	The unit of observation are villages. There are 9855 unique villages in this dataset observed over 6 different years: 1986, 1990 1993, 1996, 2000, 2003.
8b	61000 schools were constructed between 1974 and 1978. On average, 0.89 schools were constructed in each village. Children between the ages of 7 and 12 could attend these new schools.
8c	32.14% had zero, 46.33% had 1 and 21.53% had two schools.
8d	The timing of village elections is the other key source of variation used by the author to isolate the impact of improved public governance on public good provision. Because village elections are not synchronised, their staggering means their timings are quasi-random.
8e	Village heads need to be at least 25 years old in Indonesia. The first cohort educated entirely within the INPRES program schools reached age 25 in 1992.
9	
9a	Yes, there is variation. Around three quarters of villages had their first election between 1997 and 1999. 1996 had the lower proportion of villages with their first post-1992 election (1.08%), whereas 1999 had the highest (36.00%). There is no constant/linear pattern in the data: the proportion of each year is pretty flat beforeincreasing significantly from 1997 and 1999 and then crashes again in 2000.
9b	This variation is not significantly correlated with changes in the number of doctors, health centres or health posts in the pre-treatment period up to the 10% significance level. If we could demonstrated significant correlation, it would imply the timing of village election was not random. This would be a probably as it would suggest the timing of an election was influenced endogenously by village characteristics and therefore could not be used to estimate causal impacts of the INPRES program.
10	The first row of coefficients in Table 2 reflects the difference in probability the first election post 1992 has on public good provision. Depending on the public good, this election increases the odds of having these public goods by 5-6 percentage points. This impact is significant at the 1% significance level in all cases.

	The second row of coefficients reflects the difference in the probability of public good provision owing to the school construction program after the first election post 1992. This impact is only significant positive for having a primary health centre in the village, increasing the probability by around 1 percentage point. For doctors and safe drinking water, this program increases the probability by 0.4 and 0.5 percentage points respectively. However, these two coefficients are insignificant up to the 10% level.
11	After making this amendment, there is no impact on our coefficient estimates (except increasing doctors in the village by 0.001) or their significance levels. Therefore, the specification from Question 10 cannot be used to obtain the causal impact of the school construction program. We need to account for the two dimensions of fixed effects (time and village) to sufficiently estimate the causal impact of this intervention.
12	In general, all interaction coefficients increase in size and become significant at least to the 10% significance level. Therefore, at this significance level, they are all significantly positive. All interaction coefficient terms increase. The chance of having a primary health centre, doctor or access to safe drinking water in the village increases by 1.2, 0.6 and 1.8 percentage points respectively.
13	Yes, they are large. For primary health centre, doctor and access to safe drinking water, these estimates show a 7%, 6% and 1.7% increase in outcomes relative to the sample mean respectively.
14	The author finds the interaction effects larger for those villages with bad quality of service to begin with, meaning the effects are a function of villagers' demand for these public goods. This suggests the more intense the school construction program was in a village, the more receptive public governance became to what public goods were needed most.
15	These results tell us that, for most public goods, the more intense the school construction program the higher the chance there is of having good public good provision in a village.
16	The parallel trends assumption is satisfied by exploiting the unsynchronised timing of village election and showing that these elections are not correlated with public goods provisions. This implies that without the mass school construction, villages would have adhered to similar trends in public good production over time.

Word count: 1400 words

Appendix A: Requested tables

Table 1: Year of first election post 1992

Year	Proportion	Frequency
1992	4.09	403
1993	6.89	679
1994	5.72	564
1995	2.34	231
1996	1.08	106
1997	12.62	1244
1998	28.88	2846
1999	36.00	3548
2000	2.37	234

Table 2: The effects of school construction in public good provision (Q10-12)

	Primary health centre in the village			Doctors in the village			Access to safe drinking water		
	(1)	(2)	(3)	(1)	(2)	(3)	(1)	(2)	(3)
Post first 1992 election	0.049***	0.005	-0.005*	0.064***	0.021**	0.001	0.052***	0.027*	-0.005
	(0.003)	(0.007)	(0.003)	(0.004)	(0.010)	(0.003)	(0.015)	(0.015)	(0.005)
Post x num INPREs schools	0.010***	0.010***	0.012***	0.004	0.005	0.006*	0.005	0.005	0.018**
	(0.003)	(0.003)	(0.003)	(0.004)	(0.004)	(0.003)	(0.014)	(0.014)	(0.007)
Observations	59130	59130	59130	59130	59130	59130	39420	39420	39420
Mean of dependent variable	0.10	0.10	0.10	0.12	0.12	0.12	0.77	0.77	0.77
Year fixed effects	No	Yes	Yes	No	Yes	Yes	No	Yes	Yes
Village fixed effects	No	No	Yes	No	No	Yes	No	No	Yes

*Note: standard errors in parentheses. * = p < 0.1, ** = p < 0.05, *** = p < 0.01*

Table 3: Intensity of school construction (Q15)

	Primary health centre in the village (1)	Doctors in the village (2)	Access to safe drinking water (3)
Post first 1992 election	-0.017***	-0.006	-0.017**
	(0.004)	(0.005)	(0.007)
Post x INPRES schools = 1	0.013***	0.009*	0.010
	(0.005)	(0.005)	(0.009)
Post x INPRES schools = 2	0.024***	0.011	0.036**
	(0.006)	(0.007)	(0.014)
Observations	59130	59130	39420
Mean of dependent variable	0.10	0.12	0.77
Year fixed effects	Yes	Yes	Yes
Village fixed effects	Yes	Yes	Yes

*Note: standard errors in parentheses. * = p < 0.1, ** = p < 0.05, *** = p < 0.01*

Appendix B: Stata code

```
*******************************************************************************

********** ECOM90003 - Applied Microeconometric Modelling  **********

****************************** ASSIGNMENT 1 ****************************

*******************************************************************************



*******************************************************************************

******************************** SET-UP ********************************

*******************************************************************************


clear // clear all data from memory

log close


set more off // pause more message


cd "C:\Users\joshc\OneDrive\Desktop\git\mae_unimelb\2024\S2\ECOM90003 - Applied Microeconometric Modelling\Assignments\Assignment 1\"


global tablefile "C:\Users\joshc\OneDrive\Desktop\git\mae_unimelb\2024\S2\ECOM90003 - Applied Microeconometric Modelling\Assignments\Assignment 1\"



log using "assignment_1.log", replace // if you get log file open error message, insert: "log close" into command


* Import Data


use Educ_PolEcon_Data, clear
```

****

********** Q8 **********

****

//a) # of unique years and villages?

levelsof year // 6 different years

distinct v_id // There are 9855 unique villages

//b)

// # schools constructed b/w 1974 & 1978?

summarize num_PSINPRES1980 if year == 1986, detail // 8810 schools were constructed

// Average number of schools on the villages?

mean num_PSINPRES1980 if year == 1986 // 0.894 is the average

//c) What prop. of villages had 0, 1 or 2 schools?

tabulate num_PSINPRES1980 // 0 == 32.14, 1 == 46.33, 2 == 21.53.

****

********** Q9 **********

****

// Creating table 1

preserve

duplicates drop v_id, force

tabulate ele1v_post92

restore

****

********** TABLE 2 **********

****

// Creating table 2

// Step 1

```
generate post92 = 0

replace post92 = 1 if year >= ele1v_post92

// Step 2

generate interaction = post92 * num_dev



// Creating actual table

local outcome_var dum doc safe

eststo clear

foreach y in `outcome_var' {

        eststo: quietly reg `y' post92 num_dev interaction, cl(idkab_num)
        eststo: quietly areg `y' post92 num_dev interaction, absorb(year) cl(idkab_num)
        eststo: quietly areg `y' post92 num_dev interaction, absorb(year v_id) cl(idkab_num)

}

esttab using "table_2.csv", b(%5.3f) se obslast r2 star(* 0.1 ** 0.05 *** 0.01) label replace
keep(post92 interaction) title(Table 2. The effect of school construction on public good provision)
```

****

Joshua Copeland (SID:1444772)

```
******************************** TABLE 3 ********************************
*************************************************************************

eststo clear

foreach y in `outcome_var' {

        eststo: quietly areg `y' post92 i.post92##i.num_PSINPRES1980, absorb(year v_id)
cl(idkab_num)

}

esttab using "table_3.csv", b(%5.3f) se obslast r2 star(* 0.1 ** 0.05 *** 0.01) label replace title (Table
3. Intensity of School Construction) keep(post92 1.post92#1.num_PSINPRES1980
1.post92#2.num_PSINPRES1980)
```