Week 5. Regression Analysis: Poverty and Food Security in Southern Madagascar

Abdumuttolib Salako

June 25, 2022

Descriptive Statistics

Table 1: Descriptive Statistics

Measure	Mean	SD
Daily Per Capita Expenditure	1.09	0.75
Food Consumtion Score	25.38	12.55
Adults	2.20	1.20
Age Under 6	1.53	1.11
Age 6-15	1.68	1.55
Female Adults	0.28	0.45
Treatment	0.42	0.49
Poor	0.86	0.35

Table 1 above presents the descriptive statistics of the variables under consideration. The average daily consumption score is 25.38. The table also shows that the average daily per capital spending is 1.09. On average, the population of adults, population under 6, population aged between 6 -15 and female adults are 2.20,1.53, 1.68 and 0.28, respectively. The variable "Poor" is a dummy, which categorizes those above or below the international $1.90/day poverty line. From the results, it appears that 86% of the southern Madagascar. This poverty rate is higher than the 2020 poverty rate of 81% in Madagascar and the 2018 poverty rate of 40% in sub-saharan Africa.

Regression Analysis

Table 2: OLS estimates of effect of treatment on poverty in southern Madagascar

	Dependent variable:		
	exp_dollars		poor
	(1)	(2)	(3)
Number of adults	-0.097^{***}	-0.091^{***}	0.035^{***}
	(0.011)	(0.010)	(0.005)
Number 6-15	-0.085^{***}	-0.076^{***}	0.027^{***}
	(0.007)	(0.007)	(0.003)
Number <6	-0.029^{***}	-0.027^{***}	0.012^{***}
	(0.010)	(0.009)	(0.004)
Female adult only present	0.064^{**}	0.065^{**}	-0.033^{***}
	(0.028)	(0.027)	(0.013)
Treatment		0.365^{***}	-0.128^{***}
		(0.021)	(0.010)
Constant	1.472^{***}	1.288^{***}	0.781^{***}
	(0.033)	(0.034)	(0.016)
Observations	4,595	4,595	4,595
R^2	0.080	0.136	0.088
Adjusted R^2	0.079	0.135	0.087
Residual Std. Error	0.724 (df = 4590)	0.701 (df = 4589)	0.332 (df = 4589)
F Statistic	99.674^{***} (df = 4; 4590)	144.973^{***} (df = 5; 4589)	88.040^{***} (df = 5; 4589)
Note:			*p<0.1; **p<0.05; ***p<0.01

In Table 2 above, I present the OLS estimates of the effect of the treatment on daily per capita expenditure and the dummy variable "poor", which measures the proportion of the population that are below the poverty line. The first regression result shows that an increase by one adult in the household reduces the daily expenditure by 9cents. Similarly, an increase in the number of children aged between 6-15 by one, will cause the daily expenditure to decrease by 8 cents. Also, when there is unit increase in the number of children below age 6, the household experiences a reduction in daily expenditure by 2 cents. However, when the number female adult only household increases by 1, there is an increase in the daily expenditure. Generally, this implies that, as the population of adults, children between 6-15 and children under 6 increases, spending reduces. However, the reverse is the case for a female adult only household. This interpretation also holds true for the second regression output, but the reduction in expenditure seems to have subsided a bit, largely because of the treatment effect. Similar reaction is seen in the female only household: the expenditure increases by a dime relative to the first regression output. The treatment effect shows that for those who received cash assistance, their expenditure grew by 36 cents per day, which close to one-fifth of the poverty line.

In the case of the third regression result, the signs of the coefficient estimates have reversed. Now, the dummy variable "Poor" is now the dependent variable. An increase in the number of adults, number of children aged between 6-15, and the number of children under 6 will increase the household probability of getting poor by 0.035, 0.027, and 0.012, respectively. Meanwhile, a unit increase in the female adult only household, will reduce their probability of getting poor by 0.033. Finally, the household who received cash will have their probability of getting poor reduced by 0.128. All the coefficient estimates are within the 5% significance level.

Table 3: OLS estimates of effect of treatment on food consumption in southern Madagascar

	Dependent variable:		
	FCS		
	(1)	(2)	(3)
Number of adults	0.484***	0.541***	1.393***
	(0.188)	(0.187)	(0.160)
Number 6-15	0.063	0.139	0.860***
	(0.125)	(0.124)	(0.107)
Number <6	0.117	0.129	0.392***
	(0.170)	(0.169)	(0.144)
Female adult only present	0.717	0.740	0.176
	(0.488)	(0.484)	(0.412)
Treatment		3.010***	−0.365
		(0.380)	(0.333)
Expenditure			9.329***
			(0.226)
Constant	23.823***	22.283***	10.169***
	(0.585)	(0.612)	(0.598)
Observations	4,464	4,464	4,464
R^2	0.002	0.016	0.287
Adjusted R^2	0.001	0.015	0.287
Residual Std. Error	12.542 (df = 4459)	12.456 (df = 4458)	10.599 (df = 4457)
F Statistic	2.157* (df = 4; 4459)	14.322*** (df = 5; 4458)	299.703*** (df = 6; 4457)
Note:			*p<0.1; **p<0.05; ***p<0.01

Table 3 above presents three regression models, which shows the effect of covariates on the food consumption score. The food consumption score ranges from 0-100, with mean 25.38. The first regression output shows that a unit increase in the number of adults, increases the consumption score by 0.484 point. A unit increase in the number of children aged between 6-15 will increase the consumption score by 0.063 point. A unit increase in the number of children below age 6, will increase the food consumption score by 0.117 point. And a unit increase in the number of female adults only household will increase the food consumption point 0.717. Generally, it is observed that increase in household number will increase the food consumption score, but this increase is not sufficient as it is way below the average consumption score. Moreover, female adult household seem to have the highest increase in food consumption score. It also observed that all the coefficient estimates are not significant except number of adults.

Moving to the second regression result, which has the treatment as one of the independent variables. The result shows that with the treatment present in the model, the previous result in the first regression holds, with a subtle increase in the consumption score of adults, children between 6-15, children under 6 and female adult only household. The effect for adult remains the only significant variable. Similarly, female adult only household maintains the highest consumption score increase. The treatment shows that a $1 increase in cash assistance, will increase food consumption score by 3.010 points. This coefficient estimate is significant at the 1% level, implying that the cash assistance contributes significantly to the improvement of the food consumption score in southern Madagascar. It should, however, be noted that the effect of the cash treatment, although high, is still below the average food consumption score.

The third regression result has an additional variable included, which is daily expenditure. With the inclusion of this variable, the food consumption score of adults, children between 6-15 and children under 6, after a unit increase, has increased by over 50% with relative to the second regression and have also become significant. Meanwhile, female adult only consumption score, after a unit increase, reduced by over 50% compared to the second regression, but remains not significant. the addition of daily expenditure has reversed the sign of the coefficient estimate of the treatment effect. The treatment now has a negative effect on food consumption score, which means that the cash assistance reduces the food consumption score by 0.365 points. The new variable, daily expenditure, has a positive effect on food consumption score. A one dollar increase in daily expenditure will increase the food consumption score by 9.329 points. And this effect is significant at the 1% level.

It is important to know whether the expenditure data is from the baseline study or the post-treatment period because only then can we know and comment of the true effect of the treatment. If the data is from the pre-treatment period, and the coefficient of treatment is negative, then there nothing to worry about because no cash assistance has been given during that period. So, we expect that the treatment would not have any good impact on the food consumption score. If the data is from the post-study period, and the treatment is negative, then it is glaring that expenditure is capturing the effect of the treatment and is a mediator in this case and should not be included in the model. Because at the post-treatment period, expenditure is the aggregate of the household expenditure during the treatment period, of which the cash assistance is also included. Another interpretation, from my perspective, could also be that the cash assistance did not contribute to the increase in their food consumption score (which is highly doubtful anyway). Hence, I think that the data for the expenditure variable in this regression is from the pre-treatment variable.

Descrepancy in the number of observations in the merged data.

The anthro data set only includes children under 60 months, while the hh data set was collected based on household: adults, female only, under 6, between 6-15. So, it is possible that a household has more than one child below 60months or none, but they are all registered as one household in the hh data.

Table 4: OLS estimates of effect of treatment on child WAZ score in southern Madagascar

	Dependent variable:		
	waz06		sev_underweight
	(1)	(2)	(3)
Number of adults	0.049^{***}	0.049^{***}	-0.010^{**}
	(0.017)	(0.017)	(0.005)
Number 6-15	-0.046^{***}	-0.044^{***}	0.008^{***}
	(0.012)	(0.012)	(0.003)
Number <6	0.025	0.025	-0.001
	(0.017)	(0.017)	(0.004)
Female adult only present	0.073	0.071	-0.018
	(0.047)	(0.047)	(0.012)
Female Child	0.159^{***}	0.160^{***}	-0.028^{***}
	(0.036)	(0.036)	(0.009)
Age of Child in Months	-0.016^{***}	-0.016^{***}	0.001^{***}
	(0.001)	(0.001)	(0.0003)
Treatment		0.074^{**}	-0.012
		(0.037)	(0.010)
Constant	-1.316^{***}	-1.350^{***}	0.160^{***}
	(0.070)	(0.072)	(0.019)
Observations	5,748	5,748	5,748
R^2	0.045	0.046	0.005
Adjusted R^2	0.044	0.044	0.004
Residual Std. Error	1.361 (df = 5741)	1.361 (df = 5740)	0.354 (df = 5740)
F Statistic	45.049^{***} (df = 6; 5741)	39.208^{***} (df = 7; 5740)	4.235^{***} (df = 7; 5740)
Note:			*p<0.1; **p<0.05; ***p<0.01

Table 3 presents the regression results of the effect of treatment on weight-for-age Z-score (hereinafter referred to as WAZ score). The WAZ score according to the codebook is -1.61. In the first regression result, adults are 0.049 standard deviation above the average weight, and is significant at the 1% level. Furthermore, children aged between 6-15 are 0.0046 standard deviation below the average weight. This effect is also significant at the 1% level. Children below 6years are 0.025 standard deviation above the average weight. Similarly, female adults are 0.073 above the average weight. Also, the female child, with the highest positive coefficient estimate, 0.159 above the average wight. This coefficient estimate is significant at the 1% level. For children below 60months, they are 0.0016 standard deviation below the average weight, which is significant at the 1% level.

In regression 2, the treatment has been introduced. The other variables have similar effects as in regression 1. The treatment shows that an increase of $1 in cash assistance will make the treatment group 0.074 standard deviation above the average weight. This is significant at the 5% level. The third regression shows severely underweight as the dependent variable. The independent variables now have opposite signs relative to the second regression. Adults now weigh below the average severe underweight level. This holds true for children below 6, female adult, female child, as well as the treatment. The children between 6-15years are above the average severe underweight level and as well as the children below 60months.

The WAZ score only tells us those who are below or above the average weight standard, which does not necessarily mean that those below the average weight score are underweight. So, for policy makers who wants to make policy decisions surrounding children who are underweight (i.e malnourished) the underweight variables is the most appropriate in this regard. Because it shows how far above or below the standard benchmark underweight score the children are.