

Internationally, microcredit has become one of the most important tools used to combat poverty and to enhance families' wellbeing. Microcredit--more generally micro-financing--is a term that indicates "the offer of financial services to customers with low income levels". Theory suggests that microcredit is an important a tool to boost economic growth, to counteract the effects of economic instability and to empower women.

This research aims at testing the following hypothesis: microcredit is positively linked to women's socio-economic wellbeing in Cairo.<sup>1</sup> Nader and her colleagues interviewed 100 women. 50 of them had been receiving loans for at least 3 years. The other 50 had just joined the program and had not received loans yet.

To understand the impact of microcredit, the author chose 6 dependent variables to represent the outcome socio-economic wellbeing of women and estimated 6 separate models using the same independent variables in all of them.

These dependent variables are: family income, their possession of assets, the schooling of the girls, the schooling of the boys, the perception of women's health and the perception of the harmony in the family. The author provides no further details on any of these variables.

The independent variables are:

- Age (in years)
- Microcredit: a dummy that it is equal to 1 if the woman has been receiving microcredit for at least 3 years and equal to 0 if the woman has just joined the program.
- Education (in years)
- Marital status: a dummy that equals 1 if married and 0 otherwise.
- Number of children

N=100.

The results of the analysis are summarized in the table below:

	Income		Assets		Health		Harmony in Family		Schooling of girls		Schooling of boys	
	Coeff	t-stat	Coeff	t-stat	Coeff	t-stat	Coeff	t-stat	Coeff	t-stat	Coeff	t-stat
Age	0.036	-0.024	-2.901	-0.296	-0.032	-2.173	0.020	-1.371	0.003	-0.600	0.001	-0.214
Microcredit	253.504	-10.307	363.788	-2.279	-0.016	-0.069	0.377	-1.608	0.284	-3.679	0.306	-4.163
Education	19.707	-0.486	246.227	-1.347	0.211	-0.776	0.225	-0.836	0.106	-1.190	0.055	-0.644
Marital status	30.812	-0.969	271.136	-1.313	0.262	-0.852	2.043	-6.732	-0.084	-0.855	0.117	-1.199
Number of children	9.585	-1.374	99.011	-2.187	-0.094	-1.396	-0.058	-0.874	-0.022	-0.968	-0.009	-0.373
Constant	153.078	-2.680	1049.207	-2.830	7.731	-13.991	5.503	-10.096	0.333	-1.727	0.365	-2.174
Adjusted R2	0.542		0.119		0.140		0.330		0.130		0.141	
F	24.387		3.676		4.221		10.773		3.490		3.922	

<sup>1</sup> Nader, Yasmine (2008). "Microcredit and the socio-economic wellbeing of women and their families in Cairo", *Journal of Socio-Economics* 37: 644-656

1) Please explain in **non-technical terms** what you can say with a **.01 level of significance** about the impact of MICROCREDIT on FAMILY WELLBEING? (show your calculations; look at all the models) (15 points)

To answer this question, we need to know what the critical value of  $t$  is at a .01 level of significance. To figure that out, we need to know the degrees of freedom:  $df=100-5-1=94$ . Using the table, I estimate the critical value of  $t$  to be around 2.64 (the midpoint between  $t$  at 60  $df$  and 120  $df$ ). Using  $t$  at where  $df=60$  (rather than where  $df=120$ , where  $t=6.17$ ) would get you a slightly more conservative (larger values) of 2.66. Using the formula  $=tinv(.01,94)$  in Excel I get a value of 2.63. I'm going to use that, but all but  $t=6.17$  would be ok.

$t_c=2.63$  I formally can set up hypothesis tests for each model (but you did not have to). For each model:

$H_0: \beta_{\text{microcredit}}=0$

$H_a: \beta_{\text{microcredit}} \neq 0$  If  $|t| > t_c$  I can reject the null hypothesis (for a 2-sided test).

For income:  $|-10.307| > 2.63$  so I **can** reject the null hypothesis at a .01 level

For Assets:  $|-2.279| < 2.63$  so I cannot reject the null hypothesis at a .01 level

For Health:  $|-0.069| < 2.63$  so I cannot reject the null hypothesis at a .01 level

For Harmony:  $|-1.608| < 2.63$  so I cannot reject the null hypothesis at a .01 level

For Schooling Girls:  $|-3.679| > 2.63$  so I **can** reject the null at a .01 level

For Schooling Boys:  $|-4.163| > 2.63$  so I **can** reject the null at a .01 level

**In non-technical terms:** The receipt microcredit for at least 3 years is associated with higher incomes and more education for both boys and girls, in comparison to the incomes and education for boys and girls of those who just started the program.<sup>1</sup> In contrast, microcredit has no impact on assets, health, or harmony in the family.<sup>2</sup>

Because we don't know how some of the dependent variables are coded, it is difficult to make specific non-technical interpretations of the coefficients. I guess you could say that those having had microcredit had 253 more units of income, .28 more units of schooling for girls, and .306 units of schooling for boys but that is not so easy to understand.

---

<sup>1</sup>These differences are significant at a .01 level, and holding constant education, marital status, age, and number of children.

<sup>2</sup>Again, assessed at a .01 level of significance, holding constant education, marital status, age, and number of children.

2) Does the education of a woman (a) have an impact on her children's education? (b) Explain how you arrive at your answer. (c) What factors may explain this result? (15 points)

a) No, the education of a woman does not have an impact on her children's education.

b) If we set up a formal hypothesis test,

$$H_0: \beta_{\text{microcredit}}=0$$

$$H_a: \beta_{\text{microcredit}} \neq 0 \quad \text{If } |t| > t_c, \text{ reject the null hypothesis (for a 2-sided test).}$$

The critical value of  $t$  at .05 level of significance:  $df=94$ ,  $t_c=1.99$ . In for both models,  $t_c$  is not greater than the calculated  $t$ -statistic. Therefore, we cannot reject the null hypothesis at a .05 level. In short, the absolute values of the calculated  $t$  is not larger than the critical value of  $t$  at a .05 level of significance. (You could also have done this test at a .01 level of significance).

c) Here are some possible explanations:

- (1) It may be there is very little variation in education among women in this culture, and so education may not vary with children's education.
- (2) It may be that other variables in the model account for the differences in outcomes that might be attributed to education in a bivariate analysis. For example, those with higher levels of education may have been receiving education loans for 3 years, while those with less education just got them.
- (3) It may be that the sample size is too small to detect a significant difference.
- (4) It may be that these models don't meet one of the assumptions of OLS and that the standard errors are not efficient, unbiased, and minimum.

3) If you could add a slope dummy to the regression on ASSETS, (a) which slope dummy variable would you add? (b) Why do you think that would make sense? (c) What sign do you expect it to have? (d) What change do you expect to see in the coefficients of the two related variables? (20 points)

A slope dummy is the interaction of a continuous variable with a dummy variable. Given that only Microcredit and number of children are significant in the model, these are the only variables that really make sense to interact. It is unlikely that other variables would suddenly be significant. At the same time, one could argue other interactions as well. Some examples:

(1) (a) add number of children\*microcredit (b) because these are both significant and to interact them would allow for an analysis of their combined impact on assets and allow an understanding of the impact of children for those with microcredit and those just starting the program. (c) I expect a positive sign on the interaction because I theorize that the more children one has the more microcredit, and the number of children increase so do assets. (d) Both the coefficients on microcredit and number of children will decrease.

(2) (a) add marital status\*number of children (b) because those who are married likely have more kids. (c) I expect a positive sign on the interaction and (c) expect the coefficient on marital status to be smaller than it is currently and the coefficient on number of children to be smaller as well. It may be that there will

be no change on any of the related coefficients, since marital status is not significant.

(3) If one interacts 2 variables that are not significant, it is unlikely that they will have any impact on the model in combination, and none of the coefficients would change at all.

4) If you could change the regression on INCOME and make it non-linear, (a) how would you specify it? (b) Why? (c) How would you test whether that specification is better than the current specification? (15 points)

- (1) (a) Add age-squared (b) because it may be that income increases with age up to a point--as a person ages out of childhood to adulthood--and then becomes negative as a person enters old age. (c) To test if age should be nonlinear, I would look at the t-test on the squared term. If it were significant, then this nonlinear form is appropriate.
- (2) (a) Add number of children-squared (b) because one might argue that a few children are associated with increased income but a lot cause income to fall (at least, that is what some of you argued!). (c) To test if number of children should be nonlinear, I would look at the t-test on the squared term. If it were significant, then this nonlinear form is appropriate.

One could continue in the same way, adding squared terms for any of the continuous variables, if one had similar theories of the relationship between each of those variables and the dependent variable (that is, that income increases to a point with increasing values of the independent variable and then at some point income decreases as the independent variable increases).

- (3) (a) One could have logged the dependent variables, assuming one thought they were continuous. (b) Logging the dependent variable implies that as each continuous independent variable increases, the income also increases at an increasing rate. (c) To check whether this new specification is better one could see if more of the coefficients are significant, is F significantly different than zero, and if the regression residuals follow more of the assumptions for OLS. If theory suggested the relationship was correct, you might keep the transformed dependent variable regardless of these statistical issues.
- (4) (a) One could have logged any of the continuous independent variables, as long as one had an appropriate theory about functional form. (b) That theory would be that as the continuous independent variable increases, income increases at a decreasing rate (the slope of the curve becomes flatter). (c) If the t on logged term is significant and the regression residuals are improved--and theory about functional form was correct--they you could retain the transformation.
- (5) (a) One could have created a double-log model, logging both the dependent variable and the continuous independent variables. (b) This transformation implies that as the independent variable increases, income either is increasing at a decreasing rate or decreasing at an increasing rate. (c) To check whether this specification was better, you'd look at the t values on the transformed independent variables. If

they are significant, retain the transformation if theory suggests the non-linear form is the correct one. You also would check the regression residuals.

**5A) What type of research design did the authors of this research employ? (5 points)**

This is a quasi-experimental design with a non-equivalent comparison group.

**5B): What feature of the research design encourages causal inference? (5 points)**

The use of a similarly motivated control group eliminates selection bias, as those who had not received microcredit had just entered the program. That is, everyone was similarly motivated but only half had actually received microcredit. This means that differences between the groups are likely not associated with selection but with a program impact.

**5C) Discuss a specific problem for causal inference with the design. (5 points)**

The lack of a pre-test means that one does not know if all changes over time are actually attributable to the program or some non-random differences between women who were in the program over the last 3 years and those who were not.