

Notes on Econometric and Table Analysis of Financial Aid Impacts

Michael Robinson
Department of Economics
Mount Holyoke College

October 1998

Basic Econometric and Table Analysis approaches.

In both approaches to analyzing the impact of financial aid on yield the goal is to estimate the impact of additional grant on the probability of enrolling, controlling for a variety of factors. In the most simple framework the only control would be for the need of the student. This would produce a regression to be estimated such as that given in equation 1 and a grid analysis similar to that in Table 1.

$$(1) \text{ Yield} = a + b * \text{Total grant} + c * \text{need} + e$$

and

Table 1
Yields

Total Grant	Need			
	0	1000	2000	3000
0	0.10	na	na	na
1000	0.12	0.14	0.16	na
2000	0.14	0.16	0.18	0.20
3000	0.16	0.18	0.20	0.22
4000	na	0.20	0.22	0.24

The estimated coefficients on total grant from equation 1 could then be used to predict the impact on yield from say a \$1000 increase (or decrease) in grant. The estimated impact on yield combined with the revenue from the student (with and without the extra grant) could then be used to estimate the revenue impact of the increase (or decrease in grant).

The exact same approach is used in the table analysis. By reading down the columns of Table 1 you obtain an estimate of the impact of an additional \$1000 in grant on the probability of matriculating holding need constant. Reading across the rows yields an estimate of the impact of need on yield. (Many introductory econometrics text give a table similar to Table 1 to demonstrate this point (Gurjarati, Econometrics 2nd Ed, p. 48).)

The data presented in Table 1 indicate that there is a 2 percent increase in yield for every \$1000 increase in Grants. A regression run on the underlying data would yield the following results. $\text{Yield} = .10 + .02 * \text{grants} + .02 * \text{need}$ (need and grant measured in \$1000s). In this case the econometric and Table analysis approaches are identical and would yield exactly the same policy conclusions.

More complex models

Any analysis of this sort would next want to control for other important factors related yield, for example SAT scores. The econometric approach could be adapted in two ways to do this. An SAT score variable could be added to equation (1). This implicitly assumes that the impact of grant on yield is

identical between students with different SAT scores. Alternatively, equation (1) could be estimated for two or more groups of admits, say high and low SAT scores. This might yield something like the following.

High SAT scores

$$(1a) \text{ Yield} = 0.02 + 0.03 * \text{grant} + 0.01 * \text{need}$$

Low SAT scores

$$(1b) \text{ Yield} = 0.20 + 0.01 * \text{grant} + 0.02 * \text{need}$$

In this case it is clear that the yield is higher for the low SAT score group, but that the impact of grants is more pronounced for the high SAT score group. This econometric analysis is exactly equivalent to the construction of the following two sets of tables.

Table 2

High SAT Scores

Total Grant	Yields			
	0	1000	2000	3000
0	0.02	na	na	na
1000	0.05	0.06	0.07	na
2000	0.08	0.09	0.10	0.11
3000	0.11	0.12	0.13	0.14
4000	na	0.15	0.16	0.17

Low SAT Scores

Total Grant	Yields			
	0	1000	2000	3000
0	0.20	na	na	na
1000	0.21	0.22	0.23	na
2000	0.22	0.23	0.24	0.25
3000	0.23	0.24	0.25	0.26
4000	na	0.25	0.26	0.27

These tables represent the exact equivalent of the separate regressions by SAT group shown above....and as they are constructed yield exactly the same conclusion that a \$1000 increase in grant has a larger impact for high than low SAT scores students 3 and 1 percent per thousand respectively.

The econometric approach and the table approach are exactly equivalent in what they attempt to do and under the assumption that the impact of grant on yield is constant (at least within groups) they would generate identical conclusions and policy recommendations.

Difference Between the Two Approaches (Weaknesses of Econometrics)

While the econometric approach and the table approach are in many respects doing an identical analysis there are several important ways in which they are different and both have advantages and weakness. In fact it is not clear to me which is the superior approach.

Non-linearity's in Response

Table 3 reveals a relationship between yield and total grant that is highly nonlinear. The impact of the grant on yield is very high for the first grants given and then decreases. While this is quite obvious from the table the econometric estimation of equation (1) would never reveal this non-linearity and would give more or less the same results as equation (1) applied to Table 1 data since the average effect of the grant on yield is about 2 percent. In this case the policies suggested by the two analyses would be quite different and the table analysis would provide better results.

Table 3

Total Grant	Non-linear Impact of Grant on Yield			
	Yields			
	Need			
	0	1000	2000	3000
0	0.05	na	na	na
1000	0.12	0.10	0.12	na
2000	0.13	0.16	0.18	0.18
3000	0.14	0.18	0.20	0.22
4000	na	0.18	0.20	0.24

Of course it would be possible to write down an econometric specification that would capture this non-linearity. For instance you could separate the grant variable into levels and estimate a model like the following:

$$\text{Yield} = a + b_1 * (\text{grants from } 0\text{-}1000) + b_2 * (\text{grants from } 1000\text{-}2000) \dots$$

This is similar to the model that Stephen Brooks reports in his article in the Journal of Student Financial Aid, Fall 1996. However, the econometric analysis cannot suggest where to put the break points and so if one were going to construct an econometric model it would be important to conduct the table analysis first.

Differences in effect of grant across groups

Table 4 shows another problematic situation for the regression analysis. In this case the effect of grant on yield is highest for the group with need of 0 (4 percent) and lowest for the group with need of 3000 (0 percent). Again this is obvious from the table analysis, but estimation of equation (1) would give misleading results. Since the average effect of grant on yield is about 2 percent that is the estimate that would be produced by the econometric estimation. This would yield policy recommendations that would be substantially inferior to those from the table analysis.

Table 4

Total Grant	Non-linear Impact of Grant on Yield			
	Yields			
	Need			
	0	1000	2000	3000
0	0.10	na	na	na
1000	0.14	0.16	0.18	na
2000	0.18	0.19	0.19	0.21
3000	0.22	0.21	0.20	0.21

4000	na	0.23	0.21	0.21
------	----	------	------	------

Of course, as before, it would be possible to estimate an econometric model that captures this effect. To do this you would need to add interaction between need levels and grant. The appropriate model in this case would be:

$$\text{Yield} = a + b_1 * (\text{grant for need}=0) + b_2 * (\text{grant for need} = 100) \dots$$

this would produce four different estimates of the grant coefficient ranging from 4 to 0 percent. The econometrics would not tell you what interactions to include and even if one wanted to do the econometric analysis it would be advisable to complete the table analysis first.

Of course in reality there might be many combinations of non-linearity's and differential effects scattered throughout the data. An econometric model could be constructed to capture these effects once it was known what they were, but it would much easier and more effective to find these effects through use a table analysis.

Difference Between the Two Approaches (Weakness of Table Analysis)

The table analysis also has a weakness. Above I demonstrated that it is possible to control for SAT scores as well as need and grant in both the econometric and table analyses. While in principle this could be done for any number of variables, there will generally not be sufficient data to control for say 15 different variables in the table analysis. The econometric model mentioned above by Stephen Brooks uses 12 variables. Models I typically use at Mount Holyoke might have as many as 15 to 20 variables. The difficulty for the table analysis can be demonstrated by reconsidering Table 3. It might be the case that the apparent non-linearity's in the response to grant are caused by other variables omitted from the analysis that are not controlled for in Table 3. Estimation of a complete econometric model would control for these other factors and come up with a correct and linear estimate of the effect of grant on yield.

Of course the econometric analysis still must assume the effect of grant on yield is linear. If the impact is different for each sub group in the population than the econometric model would quickly run into degrees of freedom problems as well. Really what you are left with is the necessity of making some simplifying assumptions. The econometric model controls for a variety of other factors but must assume the effect of grant on yield is linear and constant across subgroups. The table analysis allows the effects to be non-linear and different across groups, but cannot control for many other factors.

Given sufficient data (that may not be available) it would be possible to make all the tables necessary to control for all factors and estimate an econometric model that allows for all possible interactions between grant and the various subgroups.

Summary

1. Table analysis and econometrics are basically trying to do the same thing in ways that are really not that dissimilar.
2. Under the assumption of a small number of factors affecting yield and linear grant effects, table analysis and econometrics will yield identical conclusions about the effect of grant on aid.
3. If there are substantial non-linearity's in the effects or different grant effects in different groups table analysis is probably superior to the econometric analysis. (Especially if there are not many other factors affecting yield.)

4. If there are a very large number of variables that effect yield and if the grant effects are linear and similar across groups, than the econometric approach is probably better.