Power of significance test of dummies in Simar-Wilson two-stage efficiency analysis model

Zelenyuk, Valentin

Institut de Statistique, Université Catholique de Louvain, Belgium, EERC, UPEG

24 May 2005

I N S T I T U T   D E
S T A T I S T I Q U E

UNIVERSITÉ CATHOLIQUE DE LOUVAIN

D I S C U S S I O N
P   A   P   E   R

0522

POWER OF SIGNIFICANCE TEST OF DUMMIES IN SIMAR-WILSON TWO-STAGE EFFICIENCY ANALYSIS MODEL

V. ZELENYUK

http://www.stat.ucl.ac.be

Power of Significance Test of Dummies in Simar-Wilson Two-Stage Efficiency Analysis Model

Valentin Zelenyuk*

Abstract

In this note we investigate the power of significance test for dummy-variables in the context of Simar and Wilson (2003) two-stage efficiency analysis model.

Key Words: truncated regression, production efficiency, dummy variables.

JEL Code: C24, D24

May 24, 2005

* Institut de Statistique, Université Catholique de Louvain, Belgium and UPEG (Ukrainian Productivity and Efficiency Group) at EROC/EERC of "Kyiv-Mohyla Academy", Kyiv, Ukraine. *Mailing address:* 20 voie du Roman Pays, B 1348, Louvain-la-Neuve, Belgium.
Tel. : +32-10-474311, fax : +32-10-473032. zelenyuk@stat.ucl.ac.be

Acknowledgement: Research support from the "Interuniversity Attraction Pole", Phase V (No.P5/24) from the Belgian Government (Belgian Science Policy) is acknowledged. Any views and remaining errors are our responsibility.

Introduction

A recent paper of Simar and Wilson (2003) (hereafter SW) gives a simple but revolutionary to the current common practice view on analyzing how explanatory variables determine the efficiency scores (measured via Farrell/Debreu-type index), when the latter is estimated via DEA.

Our comment is motivated by our application of SW procedure to various *real data* sets and observing that we often *could not* reject a hypothesis that a coefficient of a dummy variable (modeled to represent a difference in levels between some two groups) is significantly different from zero. Our immediate reaction was to investigate the *power* of such tests and we present some interesting evidence here in this comment.

Importance of our note is motivated by the fact that much of economic data involves categorical variables. In (production) efficiency analysis, in particular, researchers often categorize firms into groups (public vs. private, local vs. foreign, or/and by different regions, etc.), desiring to infer whether belonging to a particular group induces different levels of efficiency. Such inference is most commonly done by including dummy variables (indicator functions) into regression equations and testing whether their coefficients are significantly different from zero or not.

Our main conclusion is that even for reasonably large sample sizes, it might be impossible to identify statistical significance of dummy-variable coefficients that are *significant* from economics perspective, unless the true coefficient is very large. For practitioners, this means that in the so-called Two-Stage DEA analyses, one shall be very careful with policy conclusions (such as, say, state ownership is as efficient as private one and thus privatization did not improve performance, etc.) that are just based on not finding statistical significance of the corresponding coefficients of dummy variables. However, if significance is statistically identified for a particular sample then the difference in efficiency levels between the groups modelled by the dummy variable is likely to be quite large and this should add for the researcher in the corresponding policy conclusions.

Design of Monte Carlo Experiment

Our true model is assumed to be

$$TE_j = Z_j\delta + \varepsilon_j, \qquad j = 1, \ldots, n$$

where, TE_j is the true technical efficiency score of firm j, Z_j is a (row) vector of observation-specific variables for firm j that effect firms' efficiency score through the vector of parameters δ to be estimated, and where ε_j is a statistical noise.

Here, we would only present the case when TE_j is actually observed. In this case, no additional DEA-estimation is needed. Also, the percentile bootstrap method SW used is not expected to improve on the conventional inference based on the Fisher Information matrix, so we will use the latter for our inference. (As a result we save a lot of computational time, as well make the point more general, not related to the DEA or the bootstrap estimation.)

Same as in SW, we assume that $\varepsilon_j \sim N(\mu_\varepsilon, \sigma_\varepsilon)$, s.t. $\varepsilon_j \geq 1 - Z_j \delta$. The first regressor, Z_{j1}, is a vector of ones, the second regressor is a continuous variable, and we assume $Z_{j2} \sim N(\mu_Z, \sigma_Z)$. The third regressor is a dummy variable that divides the sample into two equal groups A and B, $Z_{j3} = I(j \text{ is in group } A)$, where I is the indicator function. For the sake of presentation, we choose parameters as in SW: $\mu_\varepsilon = 0, \sigma_\varepsilon = 1$, $\mu_Z = 2, \sigma_Z = 2$.

For estimating the power-function, we vary the coefficient δ_3 in the range [-0.5, 0.5] and fix the intercept at $\delta_1 = 1$ and $\delta_2 = -0.5$. Note that since $(1/TE_j) \in (0,1]$, the coefficient of categorical variable of 0.1 and 0.5 corresponds to about 9% and 33% difference in efficiency levels, respectively. Results (which were not sensitive to the change in parameters) are presented and discussed in the next section.

Results of the Monte Carlo Experiment

We repeat the Monte Carlo exercise described above for 1000 replications for the sample sizes of *n* = 100, 200, 400, 600, 800, 1000. To save on space we only present the *plots* of estimated power-functions for testing the hypothesis that the true coefficient equals zero, with the size of the test fixed at 5% (tables are available upon request).

<Insert Figure 1 about here>

Figure 1 presents power function for the sample sizes considered. We see that the power functions for the case of 100 and 200 observations are quite flat. For example, for the dummy variable case, the estimated probability of identifying difference (at 5%-significance level) of 20 percentage points in efficiency levels between two groups is only about 10% for the sample size of 100 and 20% for the sample size of 200. As it is expected for the MLE, power functions improve for the sample of size 400 and more. However, *even for the sample size of 1000, the estimated power-functions suggest that it is virtually impossible to identify the difference in efficiency between groups of about 10% percentage* points, e*.g., 90% efficient vs. 80% efficient, etc*. This is perhaps a good example when statistical (in)significance is quite different from economic significance of the estimates. A technical reason for this is that the variation on the regressand is fairly small (often between 1 and 2 or even 1.5) and might be poorly identified by a dummy variable whose variation is also small, unless its impact (coefficient) is quite large (e.g., reflecting 25 percentage points difference in efficiency for more than 400 observations).

For practitioners, our findings suggest that in the so-called Two-Stage DEA analysis, one shall be very careful with policy conclusions (e.g., conclusing that foreign ownership is as efficient as private one, etc.) that are just based on reporting *statistical* insignificance of the corresponding coefficients of dummy variables. This does not imply that categorical variables shall not be included in the truncated regression model—if so is dictated by the model then they should be included and might happen to be statistically significant. But if significance is not found—it still might be substantial from economic point of view—but might be not empirically identified due to low power of the test on coefficients of dummy-variables. On the other hand, if such significance is identified, it is likely to be quite large in reality.

References

Simar L. and P. Wilson, 2003, "Estimation and inference in two-stage, semi-parametric models of production processes," Discussion Paper (0307) of Institute of Statistics, University Catholique de Louvain. (forthcoming in *Journal of Econometrics*).

FIGURES

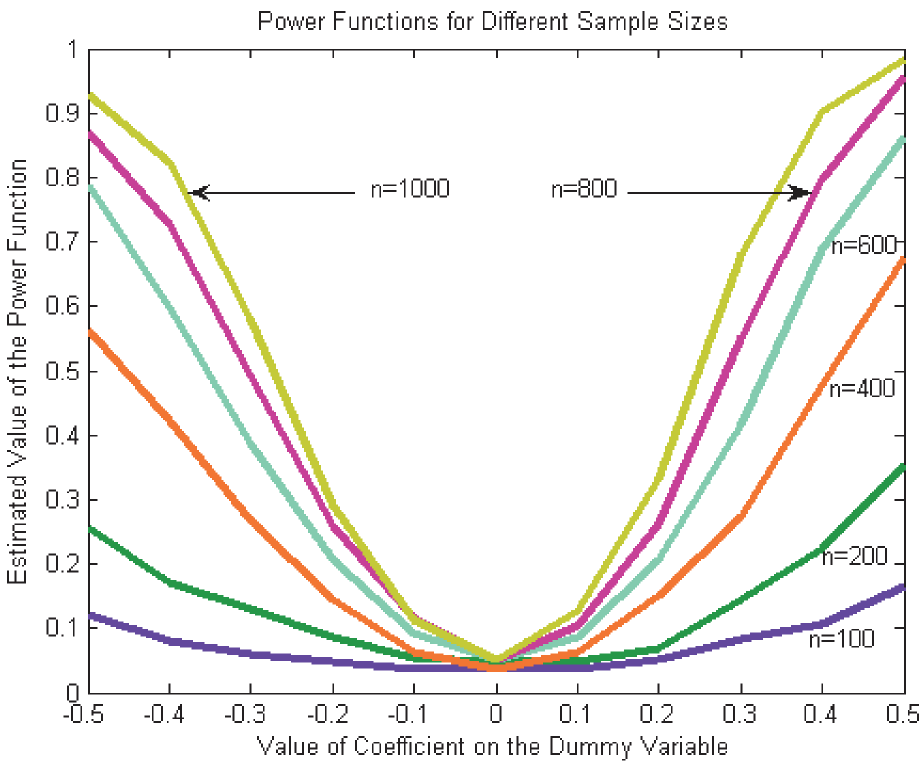

Figure 1. Estimated power functions for testing hypotheses that coefficients on continuous (dotted curve) and dummy (solid curve) variables are equal to zero, n=100 and n=200.