Malmquist Productivity Index using DEA frontier in Stata

Kyong-Rock Lee
Korea National Defense University
Seoul/Republic of Korea
rock6371@naver.com

Byung-Ihn Leem
Nextreesoft
Seoul/Republic of Korea
byleem@nextree.co.kr

Choon Woo Lee
Samsung Thales
Seoul/Republic of Korea
cw0403.lee@samsung.com

Choonjoo Lee
Korea National Defense University
Seoul/Republic of Korea
bloom.rampike@gmail.com

Abstract.

This paper introduces a user written Malmquist Productivity Index(MPI) using Data Envelopment Analysis(DEA) frontier in Stata. MPI measures the productivity changes for units between time periods. The MPI using DEA frontier suggested by Färe, Grosskopf, Norris, and Zhang (1994) is the most prevalent method to measure the productivity changes among the various methods developed(Coelli, Rao, O´Donnel, and Battese 2005). The user-written MPI approach in Stata will provide some possible future extensions of Stata programming in the productivity analysis.

Keywords: st0001, Stata, Malmquist Productivity Index, Data Envelopment Analysis, Panel data, non-parametric, Efficiency Change, Technology Change

1 Introduction

This paper introduces a new application in Stata system for Malmquist Index of productivity changes of Decision Making Units(DMUs) using DEA Frontiers. Ji and Lee (2010) introduced the user written DEA codes and explains about the basic concept of DEA frontier. The DEA provides the efficiency estimates relatve to the best practices under the assumption that the technology is fixed at current level. However, considering the notion that it is better to have more data to analyze, if panel data are available we are apt to use them. In this situation, we need to consider the changes in technology as well as the changes in efficiency according to the definition of efficiency and productivity.

2 The Basics of Malmquist Productivity Index using DEA Frontier

Malmquist Productivity Index(MPI) measures the productivity changes along with time variations and can be decomposed into changes in efficiency and technology with DEA-like nonparametric approach. Productivity decomposition into technical change and

efficiency catch-up necessitates the use of a contemporaneous version of the data and the time variants of technology in the study period. The MPI can be expressed in terms of distance function(E) as Equation (1) and Equation (2) using the observations at time t and t+1.

$$MPI_I^t = \frac{E_I^t(x^{t+1}, y^{t+1})}{E_I^t(x^t, y^t)} \tag{1}$$

$$MPI_I^{t+1} = \frac{E_I^{t+1}(x^{t+1}, y^{t+1})}{E_I^{t+1}(x^t, y^t)} \tag{2}$$

where I denotes the orientation of MPI model.

The geometric mean of two MPI in Equation (1) and Equation (2) gives the Equation (3).

$$MPI_I^G = (MPI_I^t MPI_I^{t+1})^{1/2} = [(\frac{E_I^t(x^{t+1}, y^{t+1})}{E_I^t(x^t, y^t)}) \cdot (\frac{E_I^{t+1}(x^{t+1}, y^{t+1})}{E_I^{t+1}(x^t, y^t)})]^{1/2} \tag{3}$$

The input oriented geometric mean of MPI can be decomposed using the concept of input oriented technical change(TECHCH) and input oriented efficiency change(EFFCH) as given in Equation (4).

$$MPI_I^G = (EFFCH_I) \cdot (TECHCH_I^G) = (\frac{E_I^{t+1}(x^{t+1}, y^{t+1})}{E_I^t(x^t, y^t)}) \cdot [(\frac{E_I^t(x^t, y^t)}{E_I^{t+1}(x^t, y^t)}) \cdot (\frac{E_I^t(x^{t+1}, y^{t+1})}{E_I^{t+1}(x^{t+1}, y^{t+1})})]^{1/2} \tag{4}$$

The first and second terms represent the efficiency change and the technology change respectively. MPI given by Equation (3) and Equation (4) can be defined using DEA like distance function. That is, the components of MPI can be deriven from the estimation of distance functions defined on a frontier technology. Färe, Grosskopf, Norris, and Zhang (1994) provided the formal derivation of MPI and it is the most popular method among the various methods that have been developed to estimate a production technology (Coelli, Rao, O´Donnel, and Battese 2005; Thanassoulis 2001). By utilizing both CRS and VRS DEA frontiers to estimate the distance functions in Equation (4), the technical efficiency can be decomposed into scale efficiency and pure technical efficiency components. A scale efficiency change(SECH) is given in (5).

$$SECH = [\frac{E_{vrs}^{t+1}(x^{t+1}, y^{t+1})/E_{crs}^{t+1}(x^{t+1}, y^{t+1})}{E_{vrs}^{t+1}(x^t, y^t)/E_{crs}^{t+1}(x^t, y^t)} \cdot \frac{E_{vrs}^t(x^{t+1}, y^{t+1})/E_{crs}^t(x^{t+1}, y^{t+1})}{E_{vrs}^t(x^t, y^t)/E_{crs}^t(x^t, y^t)}]^{1/2} \tag{5}$$

And a pure efficiency change(PECH) is given in (6).

$$PECH = \frac{E_{vrs}^{t+1}(x^{t+1}, y^{t+1})}{E_{crs}^{t}(x^{t}, y^{t})} \qquad (6)$$

In Figure 1, five observations and frontiers at time t and t+1 determined by constant returns to scale are presented considering one input and one output. Consider observations at time t, A(0.5, 0.5), B(1,1), C(1,2), D(2,2), E(2,3.5) and corresponding observations at time t+1 A_1(1,0.5), B_1(1,2.5), C_1(1,3), D_1(2,3), E_1(4,3.5), we can calculate the MPI using DEA like distance function. For example, $E_I^t(x^t, y^t)$ for observation D in Equation (4) can be obtained from the following process; firstly find distance measures from the point D as well as the frontier line to the reference point depending on types of distance measures, secondly find the ratio of two distance measures, which will be the distance measure of D relative to CRS Frontier at $Year_t$) devided by the distance measure of D_1 relative to CRS Frontier at $Year_{t+1}$=(1/2)/(1/2)=1 for the example. This means that there is no efficiency gains on D in between time t and t+1. When we use the DEA distance measures, we can obtain all the values in the Equations(4 ,5 ,6).
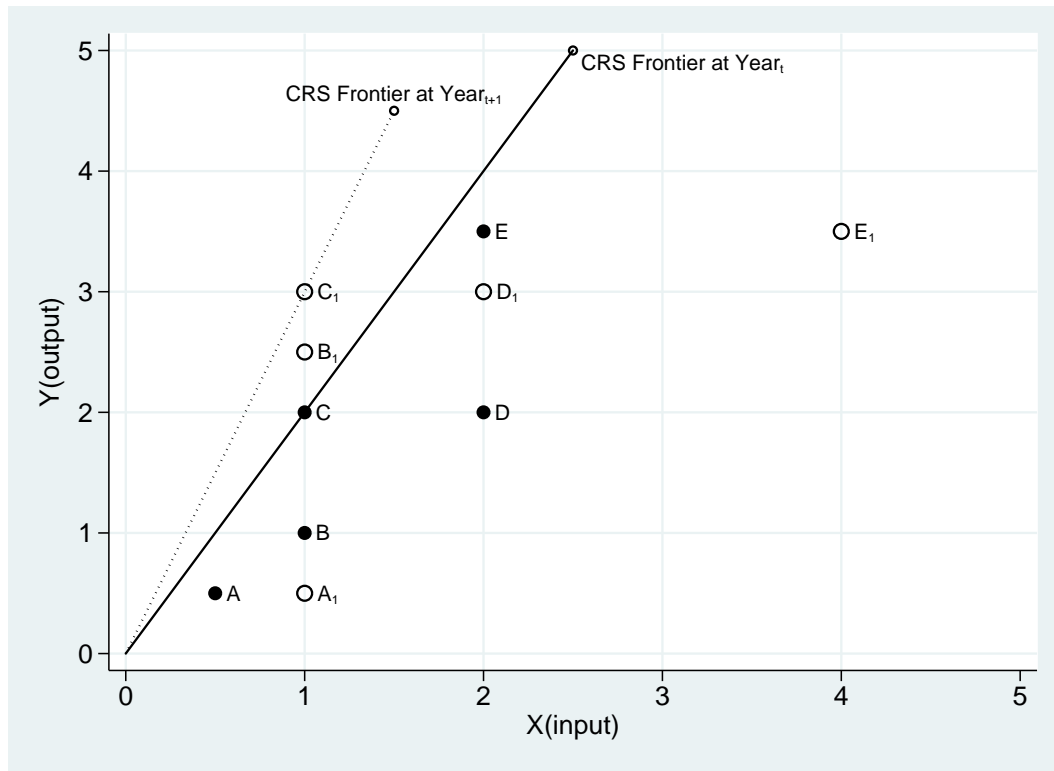

Figure 1: Concepts of Malmquist Productivity Index using CRS DEA Frontier

3   The malmq command

3.1   Syntax

The syntax of the `malmq` command is:

`malmq ivars = ovars` $\big[\, if \,\big]$ $\big[\, in \,\big]$ $\big[$`using/`*filename* $\big]$ $\big[$ `, ort(`*string*`) period(`*string*`)`
     `trace sav`*ing*`(`*filename*`)` $\big]$

3.2   Options

`ort(`*string*`)` specifies the orientation. The default is `ort(i)` or `ort(in)`, meaning the input oriented MPI. `ort(o)` or `ort(out)` means the output oriented MPI.

`period` identifies the time variable.

`trace` lets all the sequences displayed in the result window and also saved in the "malmq.log" file. The default is to save the final results in the "malmq.log" file.

`sav`*ing*`(`*filename*`)` specifies that the results be saved in *filename*.dta.

3.3   Description

`malmq` selects the input and output variables from the user designated data file or in the opened data set and solves Malmquist Productivity Index models by options specified.

The `malmq` program requires initial panel data set that contains the input and output variables for observed units.

Variable names must be identified by ivars for input variable, by ovars for output variable, and by period for panel period of time to allow that `malmq` program can identify and handle the multiple input-output data set. And the variable name of Decision Making Units(DMUs) is to be specified by `"dmu"`. `sav(`*filename*`)` option creates a *filename*.dta file that contains the results of malmq including the informaions of the DMU, input and output data used, efficiency change, technical change, total factor productivity change, scale efficiency change, and pure technical change. The log file `"malmq.log"` will be created in the working directory.

3.4   Saved Results

`malmq` saves the following in `r( )`:

Matrix

   `r(prodidxrslt)`

n x 7 matrix of the results of `malmq` command where rows show dmu corresponding to the interested period of time. Columns correspond to the variables including period(time period), dmu, tfpch(total factor productivity change),effch (efficiency change), techch(technology change),pech(pure efficiency change), and sech(scale efficiency change).

`r(effrslt)`

n x 3 matrix of the results of `malmq` command where rows show dmu corresponding to the time variable and columns correspond to the variables including dmu, CRS_eff(efficiency score under the assumption of constant returns to scale), VRS_eff(efficiency score under the assumption of variable returns to scale).

4 Applications of `malmq`

4.1 Data

This section provides examples taken using data from Coelli et al.(2005, p.296, Table 11.1) for illustration of the malmq program. The data of (Coelli, Rao, O´Donnel, and Battese 2005) consist of five firms that use one input i_x to produce one output o_q for three years period.

4.2 Input-oriented Malmquist Model

The default of malmq program specifies input-oriented CRS Frontier. If you want to use this specification for your analysis, just use the command `malmq` as below using the Coelli et al.(2005, p.296, Table 11.1) data.

```
. malmq i_x=o_q,ort(i) period(year)
-------------------------------------------------------------------------------
Input Data:


     +--------------------------+
     |   dmu   year   o_q   i_x |
     |--------------------------|
  1. | firm1      1     1     2 |
  2. | firm2      1     2     4 |
  3. | firm3      1     3     3 |
  4. | firm4      1     4     5 |
  5. | firm5      1     5     6 |
     |--------------------------|
  6. | firm1      2     1     2 |
  7. | firm2      2     3     4 |
  8. | firm3      2     4     3 |
  9. | firm4      2     3     5 |
 10. | firm5      2     5     5 |
     |--------------------------|
 11. | firm1      3     1     2 |
 12. | firm2      3     3     4 |
```

```
13. | firm3     3     4     3 |
14. | firm4     3     3     5 |
15. | firm5     3     5     5 |
     +------------------------+


Cross CRS-DEA Result:
              from     thru        t       t1
dmu:firm1        1        2       .5     .375
dmu:firm2        1        2      .75     .375
dmu:firm3        1        2  1.33333      .75
dmu:firm4        1        2       .6       .6
dmu:firm5        1        2        1     .625
dmu:firm1        2        3     .375     .375
dmu:firm2        2        3    .5625    .5625
dmu:firm3        2        3        1        1
dmu:firm4        2        3      .45      .45
dmu:firm5        2        3      .75      .75


Malmquist efficiency INPUT Oriented DEA Results:


     +---------------------------------+
     | year     dmu    CRS_eff   VRS_eff |
     |---------------------------------|
 1. |    1    firm1        .5         1 |
 2. |    1    firm2        .5      .625 |
 3. |    1    firm3         1         1 |
 4. |    1    firm4        .8        .9 |
 5. |    1    firm5  .833333         1 |
     |---------------------------------|
 6. |    2    firm1     .375         1 |
 7. |    2    firm2    .5625   .666667 |
 8. |    2    firm3        1         1 |
 9. |    2    firm4      .45   .533333 |
10. |    2    firm5      .75         1 |
     |---------------------------------|
11. |    3    firm1     .375         1 |
12. |    3    firm2    .5625   .666667 |
13. |    3    firm3        1         1 |
14. |    3    firm4      .45   .533333 |
15. |    3    firm5      .75         1 |
     +---------------------------------+


Malmquist productivity index INPUT Oriented DEA Results:


     +------------------------------------------------------------------+
     | period    dmu     tfpch    effch    techch      pech       sech |
     |------------------------------------------------------------------|
 1. |   1~2    firm1        1      .75  1.33333        1       .75 |
 2. |   1~2    firm2      1.5    1.125  1.33333  1.06667  1.05469 |
 3. |   1~2    firm3  1.33333        1  1.33333        1        1 |
 4. |   1~2    firm4      .75    .5625  1.33333  .592593  .949219 |
 5. |   1~2    firm5      1.2       .9  1.33333        1       .9 |
     |------------------------------------------------------------------|
 6. |   2~3    firm1        1        1        1        1        1 |
 7. |   2~3    firm2        1        1        1        1        1 |
 8. |   2~3    firm3        1        1        1        1        1 |
 9. |   2~3    firm4        1        1        1        1        1 |
10. |   2~3    firm5        1        1        1        1        1 |
     +------------------------------------------------------------------+
```

The results display the data set used, CRS and VRS efficiency scores for each time period, and five indices including total factor productivity change(tfpch), technical efficiency change relative to a CRS technology(effch), technical change(techch), pure technical efficiency change relative to a VRS technology(pech), and scale efficiency change(sech).

4.3 Output-oriented Malmquist Model

If you want to use CRS output-oriented Malmquist Model, you can specify `malmq` program as below using the "coelli_table11.2a.dta" file.

```
. malmq i_x=o_q,ort(o) period(year)
-----------------------------------------------------------------------------
Input Data:

     +-------------------------+
     |   dmu   year   o_q   i_x |
     |-------------------------|
 1.  | firm1      1     1     2 |
 2.  | firm2      1     2     4 |
 3.  | firm3      1     3     3 |
 4.  | firm4      1     4     5 |
 5.  | firm5      1     5     6 |
     |-------------------------|
 6.  | firm1      2     1     2 |
 7.  | firm2      2     3     4 |
 8.  | firm3      2     4     3 |
 9.  | firm4      2     3     5 |
10.  | firm5      2     5     5 |
     |-------------------------|
11.  | firm1      3     1     2 |
12.  | firm2      3     3     4 |
13.  | firm3      3     4     3 |
14.  | firm4      3     3     5 |
15.  | firm5      3     5     5 |
     +-------------------------+


Cross CRS-DEA Result:
              from     thru        t        t1
dmu:firm1        1        2        2   2.66667
dmu:firm2        1        2  1.33333   2.66667
dmu:firm3        1        2      .75   1.33333
dmu:firm4        1        2  1.66667   1.66667
dmu:firm5        1        2        1       1.6
dmu:firm1        2        3  2.66667   2.66667
dmu:firm2        2        3  1.77778   1.77778
dmu:firm3        2        3        1         1
dmu:firm4        2        3  2.22222   2.22222
dmu:firm5        2        3  1.33333   1.33333

Malmquist efficiency OUTPUT Oriented DEA Results:

     +----------------------------------+
     | year     dmu   CRS_eff   VRS_eff |
     |----------------------------------|
```

```
 1. |   1   firm1        2          1 |
 2. |   1   firm2        2    1.83333 |
 3. |   1   firm3        1          1 |
 4. |   1   firm4     1.25    1.08333 |
 5. |   1   firm5      1.2          1 |
    |---------------------------------|
 6. |   2   firm1  2.66667          1 |
 7. |   2   firm2  1.77778        1.5 |
 8. |   2   firm3        1          1 |
 9. |   2   firm4  2.22222    1.66667 |
10. |   2   firm5  1.33333          1 |
    |---------------------------------|
11. |   3   firm1  2.66667          1 |
12. |   3   firm2  1.77778        1.5 |
13. |   3   firm3        1          1 |
14. |   3   firm4  2.22222    1.66667 |
15. |   3   firm5  1.33333          1 |
    +---------------------------------+
```

Malmquist productivity index OUTPUT Oriented DEA Results:

```
     +------------------------------------------------------------------+
     | period   dmu     tfpch     effch   techch      pech      sech |
     |------------------------------------------------------------------|
 1. |   1~2   firm1        1   1.33333      .75         1   1.33333 |
 2. |   1~2   firm2  .666667   .888889      .75   .818182   1.08642 |
 3. |   1~2   firm3      .75         1      .75         1         1 |
 4. |   1~2   firm4  1.33333   1.77778      .75   1.53846   1.15556 |
 5. |   1~2   firm5  .833333   1.11111      .75         1   1.11111 |
     |------------------------------------------------------------------|
 6. |   2~3   firm1        1         1        1         1         1 |
 7. |   2~3   firm2        1         1        1         1         1 |
 8. |   2~3   firm3        1         1        1         1         1 |
 9. |   2~3   firm4        1         1        1         1         1 |
10. |   2~3   firm5        1         1        1         1         1 |
     +------------------------------------------------------------------+
```

We get the same results of Coelli et al.(2005, pp.298-300). Note that `malmq` generates the results in the reversed number of the Coelli et al. for the case of output oriented `malmq` analysis. If you are interested in the efficiency catch-up and technology progress, we can analyze the patterns of efficiency and technology changes using some summarizing statistics of the `malmq` results.

5   Conclusion

This paper introduces an author-written Malmquist Productivity Index command `malmq` among those efficiency and productivity analysis codes released in the Stata conference since 2009. The program introduced in this paper is a new application in Stata and enables users measure the productivity changes over time.

6 Acknowledgments

7 References

Coelli, T., D. S. P. Rao, C. J. O´Donnel, and G. E. Battese. 2005. *An Introduction to Efficiency and Productivity Analysis, 2nd ed.* 233 Spring Street, New York, NY 10013, USA: Springer.

Färe, R., S. Grosskopf, B. Lindgren, and P. Ross. 1992. Productivity Changes in Swedish Pharmacies 1980-1989: a non-parametric Malmquist approach. *Journal of Productivity Analysis* 3: 85–102.

Färe, R., S. Grosskopf, M. Norris, and Z. Zhang. 1994. Productivity Growth, Technical Progress, and Efficiency Changes in Industrialised Countries. *American Economic Review* 84: 66–83.

Ji, Y., and C. Lee. 2010. Data Envelopment Analysis. *Stata Journal* 10: 267–280.

Thanassoulis, E. 2001. *Introduction to the Theory and Application of Data Envelopment Analysis: A Foundation Text with Integrated Software.* Boston, USA: Kluwer Academic Publishers.

About the authors

Kyong-Rock Lee received his Ph.D in the Defense Science and Technology Department at Korea National Defense University in Seoul, Republic of Korea.

Byung-in Leem is a Principal Research Engineer in Nextreesoft(Inc.).

Choon Woo Lee is a Principal Research Engineer in Samsung Thales(Inc.).

Choonjoo Lee(Corresponing author) is an associate professor in the Defense Science and Technology Department at Korea National Defense University in Seoul, Republic of Korea.