CIRJE-F-354

Observational Equivalence between the Malmquist Index and the Solow Residual for the G-7 Countries

Jeong-Joon Lee
Towson University

July 2005

Observational Equivalence between the Malmquist Index and the Solow Residual for the G-7 Countries

Jeong-Joon Lee*
Department of Economics
Towson University
Towson, MD 21252
E-mail: jjlee@towson.edu

July 2005

Abstract

This study examines the empirical validity of the equivalence proposition on two productivity indexes, studied by Caves, Christensen, and Diewert (1982). Based on permutation tests, this study shows that the measured Malmquist index and the Solow residual are observationally equivalent for the U.S., Canada, France, Germany, and Italy. While the Solow residual is predominantly used in the existing macroeconomic studies, findings suggest that the Malmquist index is empirically compatible to the Solow residual.

JEL classification: O47; D24

Keywords: Observational Equivalence; Solow Residual; Malmquist index; Permutation Tests; G-7 Countries

*I would like to thank Michael Boskin and Lawrence Lau for their support. I have benefited from discussions with Matthew Shum and Kangoh Lee. All remaining errors are mine.

1 Introduction

By focusing on the Group-of-Seven (G-7) countries, this study examines the empirical validity of the equivalence proposition on two productivity indexes studied by Cave, Christensen, and Diewert (1982, henceforth CCD).[1]

Aggregate productivity growth is important to understand the short run business cycles and the long run economic growth. Most growth models and the standard Real Business Cycle (RBC) models predominantly rely on the Solow residual as a proxy for measured aggregate productivity growth. While an alternative aggregate productivity measure, the Malmquist index, has been proved equivalent to the Solow residual by CCD, it has not been widely used in the existing macroeconomic studies.

This paper contributes to the literature by investigating the empirical relevance of the equivalence between the Malmquist index and the Solow residual, primarily based on permutation tests. Despite its importance of aggregate productivity measures in macroeconomic studies, few studies have examined the empirical relevance of the equivalence proposition. The study finds that the measured Malmquist index and the Solow residual are *observationally equivalent*.[2] The results are robust for the U.S., Canada, France, Germany, and Italy.

What is the intuition behind these results? The Malmquist index and the Solow residual are *observationally equivalent* to an empirical researcher with data on aggregate output and aggregate input because two indexes are constructed using only a subset of the equivalence conditions given by CDD. While the *equivalence* holds when all the conditions are used, the findings suggest that even with the less than full conditions often used in the existing studies, the measured Malmquist index and the Solow residual are indistinguishable for the empirical researcher.

2 Equivalence: Theory

Define a output distance function in period t as T_t,

$$T_t(y,x) \equiv \min_\omega \{\omega : G_t(\frac{y}{\omega}, \hat{x}) \leq x^1\} \quad (1)$$

where $\hat{x}$ is input vector excluding the first input x^1, y is output, and $x^1 = G_t(y, \hat{x})$ is the input requirements function.

Then the Malmquist productivity index between period t and s becomes,

$$M_t(x_s, x_t, y_s, y_t) = T_t(y_s, x_s) \equiv \min_\omega \{\omega : G_t(\frac{y_s}{\omega}, \hat{x}_s) \leq x_s^1\} \quad (2)$$

[1] This study uses both *the Solow residual* and *the Malmquist index* terms to refer to aggregate productivity *growth*, not levels. The Solow residual is defined as changes in the Tönquivst indexes and the Malmquist index as the geometric mean of two Malmquist productivity indexes.

[2] This study focuses on *observational equivalent*, which roughly implies *similar*. A slightly different concept, *behavioral equivalent* means *equal*.

$M_t(x_s, x_t, y_s, y_t)$ measures the minimal output deflation factor needed to deflate output at time s to be on the production surface at time t, given the time s input vector. As in CCD, assume that the output distance function is translog (lng_t) and linearly homogenous in input vector x and output y. Then, the Malmquist input indexes is equal to the Solow residual.[3]

$$\textbf{Malmquist Index, } P_t^*$$

$$= \overbrace{\frac{1}{2}lnM_t(x_s, x_t, y_s, y_t) + \frac{1}{2}lnM_s(x_s, x_t, y_s, y_t)}^{\text{Geometric Mean of Malmquist Productivity Indexes}} \quad (3)$$

$$= \frac{1}{2}[lnT_t(y_s, x_s) - lnT_t(y_t, x_t)] + \frac{1}{2}[lnT_s(y_s, x_s) - lnT_s(y_t, x_t)] \quad (4)$$

$$= \frac{1}{2}[lng_t(y_s, x_s) - lng_t(y_t, x_t)] + \frac{1}{2}[lng_s(y_s, x_s) - lng_s(y_t, x_t)] \quad (5)$$

$$\begin{aligned} &= \{\nabla_{lny} lng_t(y_t, x_t) + \nabla_{lny} lng_s(y_s, x_s)\} \cdot [lny_s - lny_t] \\ &+ \{\nabla_{lnx} lng_t(y_t, x_t) + \nabla_{lnx} lng_s(y_s, x_s)\} \cdot [lnx_s - lnx_t] \end{aligned} \quad (6)$$

$$= \underbrace{(lnY_s + \alpha_* lnL_s + (1 - \alpha_*)lnK_s) - (lnY_t + \alpha_* lnL_t + (1 - \alpha_*)lnK_t)}_{\text{changes in Törnquiqust indexes}} \quad (7)$$

$$= \textbf{Solow Residual, } P_t^*$$

where $\nabla_{lny} lng_t(y_t, x_t)$ and $\nabla_{lnx} lng_t(y_t, x_t)$ are column vectors of the partial derivatives of lng_t with respect to lny and lnx. α_* is the average of labor share and $x \equiv (L, K)$ for a two input case.

Equation (4) is obtained base on the definition and Equation (5) assumes that the distance function T is translog. Equation (6) is obtained by Translog Identity and Quadratic Identity (See Lau 1979). Equation (7) is for the two-input case with assumptions of the cost-minimization, the revenue-maximization, and constant returns to scale.[4]

3 Date and Measured Productivity Indexes

3.1 Data

The data set is taken from Jorgensen and Yip (2001). This study focuses on the G-7 countries because they are relatively homogeneous in terms of size and other economic environments, which provide a suitable setting for the Data Envelopment Analysis (DEA) to construct the Malmquist index. In addition, the

[3] For a more detailed proof, see CDD.

[4] Boskin and Lau (2000) showed that estimated local returns to scale were close to one for the G-7 countries.

results of the study are comparable to existing aggregate productivity studies.[5]

This study considers a single-output, two-input production technology. Output is measured by real GDP and inputs are labor hours and physical capital. The sample period runs from 1960 to 1995.

3.2 Malmquist Index

The Malmquist index is constructed primarily relying on Färe, Grosskopf, Norris, and Zhang (1994)'s approach. Their method is deviated from the approach given by CDD. Based on the DEA, the world technology frontier is constructed first. Then, the Malmquist index ($P_{m,t}$) is obtained as in Equation (3).[6]

3.3 Solow residual

The Solow residual is constructed based on the standard growth accounting framework. This approach is also deviated from the framework given by CDD. Assume Hicks neutral technical progress and constant returns to scale. Then, the Solow residual ($P_{s,t}$) is constructed as in Equation (7).

4 Observational Equivalence

The equivalence conditions are, indeed, quite restrictive for an empirical researcher. Assume that measured aggregate productivity growth contains a random noise term, ε.

$$P_{i,t} = P_t^* + \varepsilon_{i,t} \qquad i = m, s. \tag{8}$$

where $P_{i,t}$ represents a measured aggregate productivity growth for i, P_t^* the true aggravate productivity growth, $\varepsilon_{i,t}$ a random noise term. m stands for the Malmquist index and s for the Solow residual.

The random disturbance term captures possible specification errors and other unobserved factors. Introducing $\varepsilon_{i,t}$ can be justified because this study deviates from the framework suggested by CDD. If $\varepsilon_{m,t} = \varepsilon_{s,t}$, the equivalence between the two indexes holds.

Assume that ε_i follows some probability distribution $f(\varepsilon_i; \theta_i)$ with a parameter θ_i, where $i = m, s$. And define U as the probability distribution for m and W be the probability distribution for s. Finally, consider the *observational equivalence* as in Rothenberg (1971):

Definition: Two parameter θ_m and θ_s said to be *observationally equivalent*, if $U(P; \theta_m) = W(P; \theta_s) \quad for\ all\ P \in \mathbb{R}^n$

[5] See Jorgenson (1995) for studies on international comparisons.

[6] The Malmquist indexes are computed using the software, Data Envelopment Analysis Program, developed by Tim Coelli.

4.1 Permutation Tests

Observational equivalence is tested by permutation tests, which need no distributional assumptions and the validity of the analysis only depends on the randomization. One appealing feature is that mean, median, or other test statistics can be used to obtain exact calculations for significance levels.

This study primarily follows Effron and Tibshirani's (1993, henceforth ET) algorithm. Let $P_{i,m}$ and $P_{i,s}$ be the ith observation. An empirical researcher observes $P_m = (P_{1,m}, \cdots, P_{j,m})$ and $P_s = (P_{1,s}, \cdots, P_{k,s})$ drawn from possibly different probability distributions, U and W. j and k are numbers of observations. She wishes to test the null hypothesis H_o of no difference between U and W,

$$H_o : U = W$$

If H_o is true, there is no difference between the probabilistic behavior of P_m or P_s. Thus, the two indexes can be randomized. Under the null hypothesis, the conditional distribution of the observations given their combined ordered statistics is permutation invariant.[7]

To take advantage of using permutation tests, this study chooses four different test statistics, $\phi = \theta_m - \theta_s$: the mean difference, the median difference, the first quartile difference, and the third quartile difference.

This study considers 1,000 resamples, each of which is divided into two groups and then it computes test statistics, $\hat{\phi}$.[8] Under the null hypothesis, any statistics from the two groups should exhibit no differences, $\phi = 0$. Having observed $\hat{\phi}^*$ from the original data, the Achieved Significance Level (ASL) of each permutation test (for $\hat{\phi}^* > 0$ case) is computed:[9]

$$A\hat{S}L = Prob(\hat{\phi} \geq \hat{\phi}^*) = \frac{\#(\hat{\phi} \geq \hat{\phi}^*)}{\frac{(j+k)!}{j!k!}}, \tag{9}$$

where # indicates the number of times.

4.2 Results

Table 1 shows the test results. For the U.S. Canada, France, Germany, and Italy, this study finds no evidence against the null hypotheses based on the four different test statistics. The conventional t-tests also show the same results. For the U.K. and Japan, the test results are ambiguous.

The findings suggest that the measured Malmquist index and the Solow residual are observationally equivalent for the U.S., Canada, France, Germany, and Italy in the post-World War II period. For U.K and Japan, it is likely that the two indexes are not observationally equivalent.

[7]For details about randomization tests, see Kennedy (1995) or Good (2000).

[8]The minimum number of replications recommended by ET is 1,000.

[9]Due to a large number of possible randomizations, Monte Carlo methods are used to approximate the ASL.

5 Conclusion

This study investigates the empirical validity of the equivalence proposition on two productivity indexes studied by CDD. The results suggest that two measured indexes are observationally equivalent for the U.S., Canada, France, and Germany, and Italy.

The observational equivalence between the two aggregate productivity measures could have non-trivial implications on quantitative studies on economic growth and business cycles. While the Solow residual is predominantly used in the literature as a proxy for measured aggregate productivity growth, the study shows that the Malmquist index is empirically compatible to the Solow residual.

References

[1] Boskin, M. and L. Lau, 2000, Generalized Solow-Neutral Technical Progress and Postwar Economic Growth. *NBER* No. 8023.

[2] Caves, D., L. Christensen and W. Diewert, 1982, The Economic Theory of Index Numbers and the Measurement of Input, Output and Productivity, *Econometrica*, 50:6, 1393-1414.

[3] Coelli, T., 1996, A Guide to DEAP: A Data Envelopment Analysis (Computer) Program, University of New England.

[4] Effron, B. and R. Tibshirani, 1993, *An introduction to the Bootstrap* (Chapman & Hall, New York).

[5] Färe, R., S. Grosskopf, M. Norris, and Z. Zhang, 1994, Productivity Growth, Technical Progress, and Efficiency Change in Industrial Countries, *American Economic Review*, 84, 66-83.

[6] Good, P., 2000, *Permutation tests*: A Practical Guide to Resampling Methods for Testing Hypothesis (Springer-Verlag, Heidelberg).

[7] Jorgenson, D., 1995, Productivity, *International Comparisons of Economic Growth* (MIT Press, Cambridge).

[8] Jorgenson, D. and E. Yip, 2001, Whatever Happened to Productivity Growth? in : E. Dean, M. Harper, and C. Hulten, eds., *New Developments in Productivity Analysis* (University of Chicago Press, Chicago) 205-246.

[9] Kennedy, P., 1995, Randomization Tests in Econometrics, J*ournal of Business and Economic Statistics*, 13,85-94.

[10] Lau, L., 1979, On Exact Index Numbers, *Review of Economics and Statistics*, 61, 73-82.

[11] Rothenberg, T., 1971 Identification in Parametric Models, *Econometrica*, 39, 577-591.

Table 1: Test Results: two-tailed

	Permutation		Tests	(ASLs)	t-tests (P-values)
$\phi = \theta_m - \theta_s$	Mean	Median	1st Q	3rd Q	Mean
U.S.	0.38	0.43	0.65	0.62	0.74
Canada	0.42	0.51	0.52	0.56	0.80
U.K.	0.06	0.21	0.12	0.02**	0.12
France	0.32	0.28	0.07	0.74	0.61
Germany	0.34	0.23	0.38	0.48	0.68
Italy	0.24	0.37	0.32	0.42	0.49
Japan	0.05*	0.02**	0.01**	0.31	0.09*

Significant test statistics at 5% and 10%levels are indicated with ** and *, respectively. ASL stands for the achieved significance level.