Dynamical Analysis of the Malmquist Productivity Index

Alexandra Daskovska
Sébastien Van Bellegem   Léopold Simar

Institut de Statistique
Université Catholique de Louvain

September 21, 2007

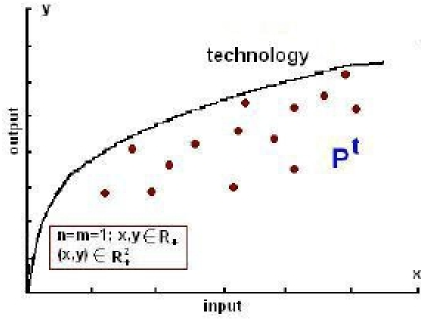
Overview

1. Introduction. Malmquist Productivity Index (MPI)
2. Forecasting the MPI
3. Confidence intervals: bootstrap on correlated pairs vs. bootstrap on correlated triples
4. Finite sample properties
5. Empirical illustration
6. Conclusions and future research

Introduction

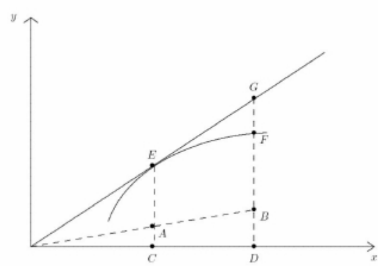
Some facts about productivity

- Let us suppose that a production unit (which can be company, country, etc.) uses n inputs in order to produce m outputs. Denote inputs and outputs as $x \in R_+^n$ and $y \in R_+^m$ correspondingly



- Activity of organization at time t can be described by its production possibilities set $P^t = \{(x, y) \in R_+^{n+m} | x \text{ can produce } y \text{ at time } t\}$
- Upper boundary of P^t is called technology, or production frontier.

Introduction
Distance functions



Production unit moves from the point A at time t_1 to the point B at time t_2, technology remains unchanged. Here, e.g.

$$D^{t_2}(x_i^{t_2}, y_i^{t_2}) = DB/DF$$

$$\Delta^{t_2}(x_i^{t_2}, y_i^{t_2}) = DB/DG$$

- In order to measure the distance from the point to the frontier, the following distance measures are used:
- Output distance function for the i^{th} production unit at time t_j relative to the technology existing at time t_k :

$$D^{t_k}(x_i^{t_j}, y_i^{t_j}) = \inf\{\theta > 0 | (x_i^{t_j}, y_i^{t_j}/\theta) \in P^{t_k}\}$$

- Output distance function for the i^{th} production unit at time t_j relative to the convex cone (with vertex at the origin) V^t spanned by technology existing at time t_k :

$$\Delta^{t_k}(x_i^{t_j}, y_i^{t_j}) = \inf\{\theta > 0 | (x_i^{t_j}, y_i^{t_j}/\theta) \in V^{t_k}\}$$

- Malmquist productivity index (MPI)

$$\Pi_i^{t_1,t_2} = \left(\frac{\Delta^{t_1}(x_i^{t_2}, y_i^{t_2})}{\Delta^{t_1}(x_i^{t_1}, y_i^{t_1})} \cdot \frac{\Delta^{t_2}(x_i^{t_2}, y_i^{t_2})}{\Delta^{t_2}(x_i^{t_1}, y_i^{t_1})} \right)^{1/2}$$

  measures the productivity changes of a given production unit i between two time periods t_1 and t_2
- **Useful feature of MPI:** its ability to be decomposed into the product of efficiency, scale, technology and scale technology changes:

$$\Pi_i^{t_1,t_2} = \left(\frac{D^{t_2}(x_i^{t_2}, y_i^{t_2})}{D^{t_1}(x_i^{t_1}, y_i^{t_1})} \right) \cdot \left(\frac{\Delta^{t_2}(x_i^{t_2}, y_i^{t_2})/D^{t_2}(x_i^{t_2}, y_i^{t_2})}{\Delta^{t_1}(x_i^{t_1}, y_i^{t_1})/D^{t_1}(x_i^{t_1}, y_i^{t_1})} \right)$$

$$\cdot \left(\frac{D^{t_1}(x_i^{t_2}, y_i^{t_2})}{D^{t_2}(x_i^{t_2}, y_i^{t_2})} \cdot \frac{D^{t_1}(x_i^{t_1}, y_i^{t_1})}{D^{t_2}(x_i^{t_1}, y_i^{t_1})} \right)^{1/2}$$

$$\cdot \left(\frac{\Delta^{t_1}(x_i^{t_2}, y_i^{t_2})/D^{t_1}(x_i^{t_2}, y_i^{t_2})}{\Delta^{t_2}(x_i^{t_2}, y_i^{t_2})/D^{t_2}(x_i^{t_2}, y_i^{t_2})} \cdot \frac{\Delta^{t_1}(x_i^{t_1}, y_i^{t_1})/D^{t_1}(x_i^{t_1}, y_i^{t_1})}{\Delta^{t_2}(x_i^{t_1}, y_i^{t_1})/D^{t_2}(x_i^{t_1}, y_i^{t_1})} \right)^{1/2}$$

$$= \Delta\text{PureEff}_i^{t_1,t_2} \cdot \Delta\text{Scale}_i^{t_1,t_2} \cdot \Delta\text{PureTech}_i^{t_1,t_2} \cdot \Delta\text{ScaleTech}_i^{t_1,t_2}$$

Forecasting the MPI
Objective

- **Problem:** How can the performance of a production unit be forecasted in terms of productivity?
  - 'Naive' method (static): forecast productivity change of interest with a help of geometrical mean of (estimated) productivity indices for all previous years

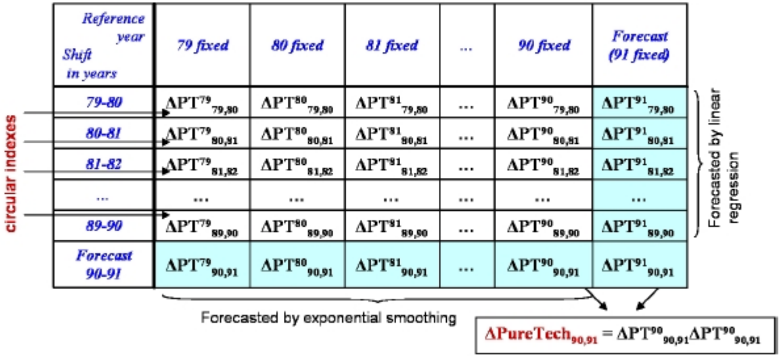
$$\Pi_i^{t,t+1} = \sqrt[t-1]{\Pi_i^{1,2} \cdot \ldots \cdot \Pi_i^{t-1,t}}$$

  - Drawback: this approach hides a potentially valuable information given by the evolution of productivity over time
  - Goal: to develop a dynamic approach for forecasting MPI, taking into acount its behaviour over time.

- **Why circularity of MPI is important?**
  - In axiomatic index theory circularity is regarded as one of the fundamental properties an index should obey
  - Circularity is necessary reqiurement to build a meaningful forecasting theory for an index
- By definition, an index I is called circular if

$$I_i^{t_1,t_3} = I_i^{t_1,t_2} \cdot I_i^{t_2,t_3}$$

  Unfortunately, MPI is (in general) not circular
- **Decompose it into circular components!**



Reference year / Shift in years	79 fixed	80 fixed	81 fixed	...	90 fixed	Forecast (91 fixed)
79-80	$\Delta PT^{79}_{79,80}$	$\Delta PT^{80}_{79,80}$	$\Delta PT^{81}_{79,80}$	...	$\Delta PT^{90}_{79,80}$	$\Delta PT^{91}_{79,80}$
80-81	$\Delta PT^{79}_{80,81}$	$\Delta PT^{80}_{80,81}$	$\Delta PT^{81}_{80,81}$	...	$\Delta PT^{90}_{80,81}$	$\Delta PT^{91}_{80,81}$
81-82	$\Delta PT^{79}_{81,82}$	$\Delta PT^{80}_{81,82}$	$\Delta PT^{81}_{81,82}$	...	$\Delta PT^{90}_{81,82}$	$\Delta PT^{91}_{81,82}$
...	...	...	...	...	...	...
89-90	$\Delta PT^{79}_{89,90}$	$\Delta PT^{80}_{89,90}$	$\Delta PT^{81}_{89,90}$	...	$\Delta PT^{90}_{89,90}$	$\Delta PT^{91}_{89,90}$
Forecast 90-91	$\Delta PT^{79}_{90,91}$	$\Delta PT^{80}_{90,91}$	$\Delta PT^{81}_{90,91}$	...	$\Delta PT^{90}_{90,91}$	$\Delta PT^{91}_{90,91}$

The corresponding table for $\Delta ScaleTech_{90,91}$ is obtained in the similar way.

Confidence Intervals

Bootstrap on correlated pairs

- Want to make inference on MPI $\rightarrow$ BOOTSTRAP
- Bootstrapping the MPI = generating B pseudo-samples $X_b^* = \{(x_{it}^*, y_{it}^*) | i = 1, ..., N; t = 1, ..., T\}; b = 1, ..., B \rightarrow$ applying the original forecasting procedure $\rightarrow$ bootstrap empirical distributions for MPI
- Simple resampling from the original input/output set is not valid! (inconsistent estimation of confidence intervals)
- **Solution:** smooth bootstrap, including possible time-dependent structure of the data.
- Using bivariate kernel instead of the univariate one:

$$\hat{f}(z) = \frac{1}{Nh^2} \sum_{i=1}^{N} K\left(\frac{z - Z_i}{h}\right), \; Z_i = [\hat{D}_i^t \; \hat{D}_i^{t+1}]$$

- $\hat{D}_i^t$ and $\hat{D}_i^{t+1}$ (output distance functions) are bounded by unity from below $\rightarrow$ $\hat{f}(\cdot)$ **inconsistent and asymptotically biased on the boundary.**
- $\Rightarrow$ Adapt the reflection method (Silverman (1986)): reflect the data about the boundaries in two-dimensional space:

$$\Delta = \begin{bmatrix} A & B \\ 2-A & B \\ 2-A & 2-B \\ A & 2-B \end{bmatrix}$$

where $A = [\hat{D}_1^t, ..., \hat{D}_N^t]'$ and $B = [\hat{D}_1^{t+1}, ..., \hat{D}_N^{t+1}]'$.

- The temporal correlation of the original data $[A\ B]$ (and, $[2-A\ B]$ resp.) is measured by the estimated covariance matrices

$$\hat{\Sigma} = \begin{bmatrix} \hat{\sigma}_1^2 & \hat{\sigma}_{12} \\ \hat{\sigma}_{12} & \hat{\sigma}_2^2 \end{bmatrix} \text{ and } \hat{\Sigma}_R = \begin{bmatrix} \hat{\sigma}_1^2 & -\hat{\sigma}_{12} \\ -\hat{\sigma}_{12} & \hat{\sigma}_2^2 \end{bmatrix}$$

- Then

$$\hat{g}(z) = \frac{1}{4Nh^4} \sum_{j=1}^{4N} K_j\left(\frac{z-\delta_j}{h}\right)$$

  is a kernel estimator of the density of the $4N$ reflected points represented by the rows of Δ.
- Consistent estimate of the density of the original data $[A\ B]$ with bounded support is

$$\hat{g}^*(z) = \begin{cases} 4\hat{g}(z), & \textit{for } z_1 \geq 1, z_2 \geq 1 \\ 0, & \text{otherwise} \end{cases}$$

- Draw randomly with replacement N rows from Δ to form the $(N \times 2)$ matrix $\Delta^* = [\delta_{ij}], i = 1, ..., N, j = 1, 2$
- Generate an $(N \times 2)$ matrix ϵ^* containing N independent draws ϵ_i^* from a normal density with shape $\hat{\Sigma}$ ($\hat{\Sigma}_R$ resp.)
- Compute the $N \times 2$ matrix

$$\Gamma = (1 + h^2)^{-1/2}(\Delta^* + h\epsilon^* - C \begin{bmatrix} \bar{\delta}_{.1} & 0 \\ 0 & \bar{\delta}_{.2} \end{bmatrix}) + C \begin{bmatrix} \bar{\delta}_{.1} & 0 \\ 0 & \bar{\delta}_{.2} \end{bmatrix}$$

- For each element γ_{ij} of Γ, set

$$\gamma_{ij}^{*} = \begin{cases} \gamma_{ij}, & \textit{if} \;\; \gamma_{ij} \geq 1, \\ 2 - \gamma_{ij}, & \text{otherwise} \end{cases}$$

  The resulting $(N \times 2)$ matrix $\Gamma^{*} = [\gamma_{ij}^{*}]$ consists of two column-vectors of simulated distance function values.

- Obtain pseudo samples $X^{*,t}$, $X^{*,t+1}$ where

$$x_{it}^{*} = x_{it}, \;\; y_{it}^{*} = (y_{it}/\gamma_{ij}^{*}) \cdot \hat{D}_{i}^{t}$$

  for $i = 1, ..., N$.

1. As a result we obtain a bootstrap pseudo sample

$$X^* = \{(x_{it}^*, y_{it}^*)|i = 1, ..., N; t = 1, ..., T\}.$$

   Basing on this pseudo sample as on reference set, forecast by dynamical method the Malmquist productivity index $\hat{\Pi}_i^{*T,T+1}$

2. Loop, for the selected unit k, the previous steps B times, yielding the bootstrap empirical distributions of Malmquist indices and its components, for the unit k.

Remark: it is possible to extend the method allowing for three or more subsequent time periods to be correlated => **bootstrap on correlated triples**

- Compare the performance of confidence intervals: bootstrap on correlated pairs vs. bootstrap on correlated triples
- $\rightarrow$ MC simulations
- Input-output set is generated according to the **Cobb-Douglas model**

$$y_{it} = \alpha_t + \beta_t' x_{it} - \eta_{it}, \quad i = 1, ..., N, \quad t = 1, ..., T$$

where $x_{it} \in R^d$ is a d-dimensional vector of logs of inputs X_{it}, y_{it} is the log of outputs Y_{it}, $\alpha_t \in R^d$ and $\beta_t \in R^d$.

Finite Sample Properties

MC Simulations: the Algorithm

1. Choose the sample size N (number of productivity units), time span T, and the number of inputs d.
2. Generate the regressors according to the bivariate VAR model

$$x_{it} = Rx_{i,t-1} + \nu_{it}, \quad \text{where } \nu_{it} \sim I\mathcal{N}(0, \sigma_X^2 I_2)$$

3. Generate α_t, β_t according to the formulae

$$\alpha_t = \alpha + \gamma_1 \frac{t}{T}, \;\; \beta_t = \beta + \gamma_2 \frac{t}{T}$$

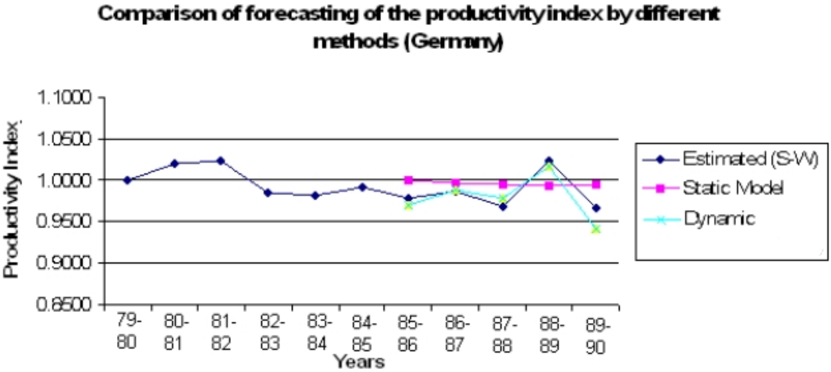
4. The error term η_{it} of the model is generated as

$$\eta_{it} = \lambda_i + \epsilon_{it},$$

   where $\lambda_i \sim Exp(1)$, and $\epsilon_{it} \sim 0.5\epsilon_{i,t-1} + e_{it}$, where $e_{it} \sim Unif[-\lambda_i/2, \lambda_i/2]$. Note that the resulting random variable is such that $\eta_{it} > 0$.
5. Finally, obtain X_{it} and Y_{it} by taking exponents.

Table: **Smooth bootstrap on correlated pairs vs. triples.** Performance of confidence intervals: average length of interval and estimation of coverage probabilities, $N = 30$, $T = 10, 15, 30$. Number of Monte-Carlo experiments $M = 500$, bootstrap $B = 1000$.

	Corr. Pairs		**Corr. Triples**	
T	**Length**	**Coverage**	**Length**	**Coverage**
10	0.76	0.83	0.68	0.88
15	0.54	0.91	0.61	0.94
30	0.6	0.88	0.45	0.93
45	0.64	0.90	0.48	0.93



The annual data collected for 17 OECD countries for years 1979-90.

Input: labor, capital. **Output:** gross domestic product (GDP)

Empirical Illustration

OECD dataset

Years	Estimated	Forecasted	
		Static	*Dynamic*
79-80	1.0158		
80-81	1.0182		
81-82	1.0210		
82-83	1.0179		
83-84	0.9981		
84-85	0.9911		
85-86	1.0187	1.0103	1.0065
86-87	1.0000	1.0115	1.0037
87-88	0.9726	1.0100	0.9896
88-89	0.9755	1.0058	0.9705
89-90	0.9626	1.0027	0.9395
Median abs.err.		0.0303	0.0121
Median SE		0.0009	0.0001
Ind. function		0.4	1

Table: Forecasted vs. estimated productivity index for Japan.

Empirical Illustration

The quality of forecasting

	Static			Dynamic		
	MAE	MSE	F(t)	MAE	MSE	F(t)
Australia	0.0212	0.0004	0.6	0.0077	0.0001	0.8
Austria	0.0251	0.0006	0.8	0.0088	0.0001	0.6
Belgium	0.0168	0.0003	1	0.0078	0.0001	1
Canada	0.0218	0.0005	0.8	0.0081	0.0001	0.8
Denmark	0.0140	0.0002	0.4	0.0119	0.0001	0.8
Finland	0.0234	0.0005	0.8	0.0073	0.0001	0.8
France	0.0145	0.0002	1	0.0080	0.0001	1
Germany	0.0270	0.0007	0.6	0.0072	0.0001	1
Greece	0.0067	0.0000	0.6	0.0120	0.0001	0.8
Italy	0.0180	0.0003	0.2	0.0114	0.0001	0.8
Japan	0.0303	0.0009	0.4	0.0121	0.0001	1
Norway	0.0239	0.0006	0.6	0.0072	0.0001	1
Spain	0.0324	0.0010	0.2	0.0110	0.0001	0.8
Sweden	0.0032	0.0000	1	0.0105	0.0001	1
UK	0.0232	0.0005	0.2	0.0037	0.0000	0.8
USA	0.0107	0.0001	0.4	0.0092	0.0001	0.8

Quality of forecasting by 2 methods for OECD countries, where MAE is **median absolute error**, MSE is **median squared error** and $F(t)$ is the **indicator function** showing the estimated probability to predict the right direction of productivity change.

Conclusions and Future Work

- Dynamical method seems to perform better than a static one
- Bootstrap on correlated triples seems to perform better than bootstrap on correlated pairs
- Incorporate the "best" bandwidth h into the bootstraping procedure
- Model the MPI: Include environmental (economic) variables, conditional MPI ...