

**Hướng dẫn sử dụng phương pháp Phân tích bao dữ liệu trong Excel:  
Vietnamese DEA add-in for Excel (phiên bản 2.0)**

**Ngô Đăng Thành**

Trường Đại học Kinh tế, ĐHQGHN

Massey Business School, New Zealand

**Tóm tắt:**

*Bài viết trình bày tóm tắt về phương pháp Phân tích bao dữ liệu (Data Envelopment Analysis – DEA) cũng như một số mô hình cơ bản của nó, bao gồm mô hình ước lượng hiệu quả kỹ thuật (sử dụng dữ liệu chéo – cross-sectional data) và mô hình ước lượng năng suất tổng hợp Malmquist TFP theo thời gian (sử dụng dữ liệu bảng – panel data). Phiên bản 2.0 này cũng cung cấp cho người dùng thông số về các bộ trọng số tối ưu (multipliers hay shadow prices). Việc sử dụng phương pháp Phân tích bao dữ liệu giúp các nhà nghiên cứu có thể đánh giá được hiệu quả (tương đối) của các đơn vị/doanh nghiệp/tổ chức (gọi tắt là các đơn vị ra quyết định – Decision making units, DMUs) hoạt động trong cùng một ngành nghề, lĩnh vực như ngân hàng, giáo dục, bệnh viện... Việc xây dựng một tiện ích thực hiện phân tích bao dữ liệu dành cho người Việt được hy vọng sẽ góp phần nhân rộng tính ứng dụng và tính phổ biến của phương pháp này tại Việt Nam. Trong các phiên bản tiếp theo, tác giả hy vọng có thể tích hợp thêm nhiều mô hình như Fisher DEA, SBM DEA, Network DEA... vào tiện ích VDEA nói trên.*

**Abstract:**

*This paper briefly introduces the Data Envelopment Analysis (DEA) in estimating the technical efficiency (using cross-sectional data) and Malmquist total factor productivity changes over time (using balanced panel data), as well as the correspondence multipliers or shadow prices. This technique allows researchers to evaluate the relative efficiency of the decision making units (DMUs) working in the same industry and environment such as banking, education, hospital. The construction of a Vietnamese DEA add-in for Excel is expected to boost up the use of the above technique in the academic area in Vietnam. In the following update, the author aims to include other models such as Fisher DEA, SBM DEA, and Network DEA into the add-in.*

**Key words:** data envelopment analysis, Excel, add-in, Vietnam

## 1. Giới thiệu chung

Bài viết này nhằm mục đích khái quát lại về phương pháp Phân tích bao dữ liệu (Data Envelopment Analysis, viết tắt là DEA) trong việc tính toán và ước lượng hiệu quả (kỹ thuật) của các doanh nghiệp, ngân hàng, trường học,... (gọi chung là các đơn vị ra quyết định – Decision Making Unit, viết tắt là DMU) trong việc sử dụng các nguồn lực đầu vào để tạo ra các kết quả đầu ra. Việc đo lường hiệu quả như vậy dựa trên cơ sở của phương pháp phân tích giới hạn (frontier analysis), theo đó, các DMU có hiệu quả cao nhất sẽ xác lập nên một đường giới hạn khả năng sản xuất (production frontier), và các DMU sẽ được so sánh với đường giới hạn này để xác định xem chúng hoạt động có hiệu quả hay không. Đối với các DMU hiệu quả, vì chúng nằm trên đường giới hạn, nên điểm hiệu quả kỹ thuật (technical efficiency score, gọi tắt là TE) của chúng bằng 1. Đối với các DMU kém hiệu quả (nằm trong đường giới hạn), điểm hiệu quả của chúng sẽ nhỏ hơn 1 (xem thêm trong Mục 2).

Hiện nay, có khá nhiều phần mềm cho phép ước lượng hiệu quả kỹ thuật theo phương pháp DEA, bao gồm cả phần mềm thương mại (phải mua, ví dụ như DEA Frontier, DEA-Excel-Solver Pro) lẫn phần mềm miễn phí (như DEAP, DEAOS,...). Hạn chế lớn nhất của các phần mềm này (trừ vấn đề chi phí đối với các phần mềm thương mại) đối với những người mới bắt đầu nghiên cứu về DEA như học sinh, sinh viên... là có giao diện bằng tiếng Anh. Hạn chế lớn thứ hai là khó khăn trong việc xử lý số liệu, chạy chương trình, và đọc kết quả. Điển hình như DEAP, một phần mềm được sử dụng phổ biến kể cả với những người nghiên cứu phương Tây, sử dụng giao diện trên nền hệ điều hành MS-DOS với các file dữ liệu, thông số, và kết quả đều ở dạng text (\*.txt)<sup>1</sup>. Hay như DEAOS là phần mềm trực tuyến (online solution) đòi hỏi người dùng phải nhập (import) dữ liệu khá phức tạp. Vì vậy, tiện ích VDEA được xây dựng nhằm:

- Tích hợp công việc nhập dữ liệu, đặt tham số, tính toán, và in kết quả trong môi trường Excel. Xuất phát từ thực tế là phần mềm Excel phiên bản 2010 (trở lên) đã và đang được sử dụng rộng rãi tại Việt Nam, tiện ích này có tính tương thích cao với Excel 2010, nhưng vẫn có thể sử dụng được với Excel 2007 và Excel 2013. Với các phiên bản khác, chẳng hạn như Excel 2003, đề nghị liên lạc với tác giả để được hỗ trợ cụ thể.

---

<sup>1</sup> Độc giả có thể tham khảo Hướng dẫn sử dụng DEAP bằng tiếng Việt (cũng do tác giả thực hiện) tại địa chỉ <https://sabapsau.wordpress.com/2011/03/23/h%C6%B0%E1%BB%9Bng-d%E1%BA%ABn-s%E1%BB%AD-d%E1%BB%A5ng-ph%E1%BA%A7n-m%E1%BB%81m-deap-tom-t%E1%BA%Aft/>

- Sử dụng song ngữ (tiếng Việt không dấu và tiếng Anh) để tạo điều kiện cho kể cả những người mới nghiên cứu về DEA cũng có thể dễ dàng sử dụng.

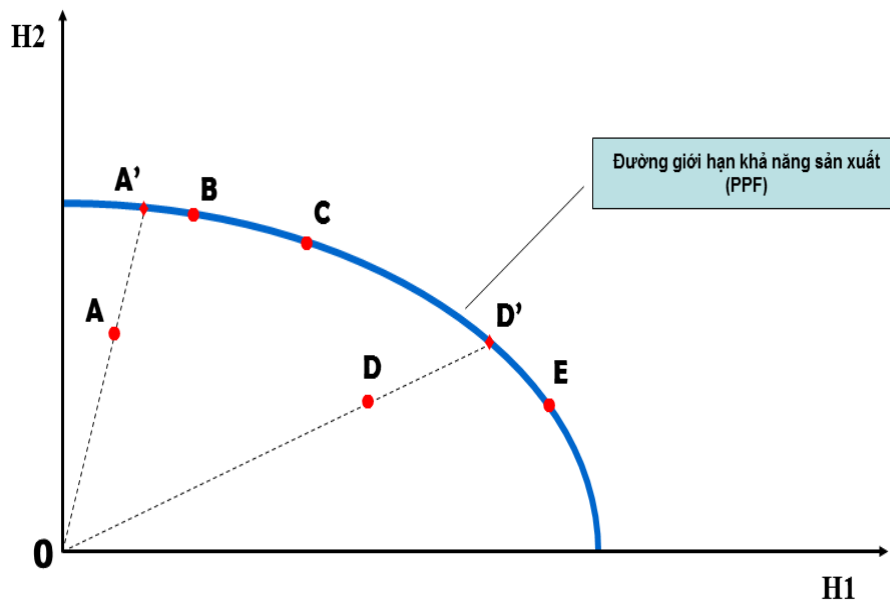
- Phiên bản VDEA 1.1 có khả năng thực hiện một số mô hình cơ bản như mô hình tối thiểu hóa đầu vào (input-minimization), tối đa hóa đầu ra (output-maximization), hiệu quả không đổi theo quy mô (constant-returns-to-scale, CRS), và hiệu quả thay đổi theo quy mô (variable-returns-to-scale, VRS). Phiên bản 1.3 có khả năng tính toán năng suất nhân tố tổng hợp Malmquist TFP. **Trong phiên bản hiện tại (VDEA 2.0), việc tính toán các trọng số tối ưu (multipliers) hay “giá ẩn” (shadow prices) đã được tích hợp trong cửa sổ nhập liệu của VDEA.** Trong thời gian tới, các mô hình khác như SBM, Fisher TFP,... cũng sẽ dần được hoàn thiện. Tác giả rất mong nhận được các ý kiến góp ý, báo lỗi,... để có thể tiếp tục hoàn thiện tiện ích VDEA hơn nữa.

Phần tiếp theo của bài viết được cấu trúc như sau. Trong Mục 2, tác giả khái quát lại một số kiến thức cơ bản về phương pháp phân tích bao dữ liệu (DEA), bao gồm các vấn đề đường giới hạn khả năng sản xuất, bộ trọng số động (dynamic weights hay còn gọi là shadow prices), hiệu quả nhờ quy mô, mô hình hướng theo đầu vào/đầu ra,... Mục 3 tóm lược về việc tính toán chỉ số năng suất tổng hợp (Total Factor Productivity) theo phương pháp Malmquist trong DEA. Mục 4 giới thiệu về tiện ích VDEA cũng như cách sử dụng VDEA trong Excel 2010. Các kết luận, cũng như gợi ý nghiên cứu trong thời gian tới sẽ được trình bày trong Mục 5.

## **2. Phương pháp Phân tích bao dữ liệu (DEA) và đo lường hiệu quả kỹ thuật (tương đối)**

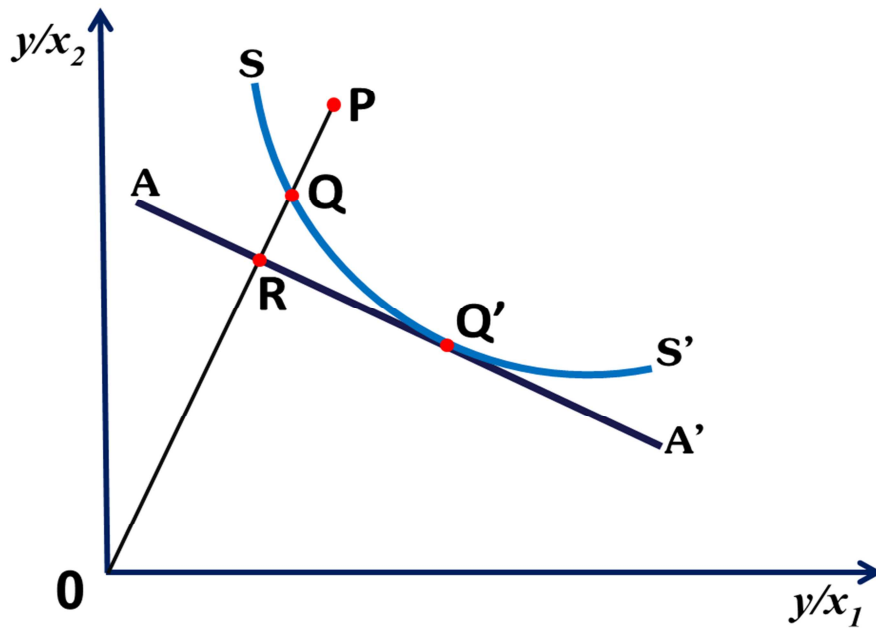
### **Đường giới hạn khả năng sản xuất PPF và phương pháp Phân tích giới hạn (frontier analysis)**

Farrell (1957) đưa ra ý tưởng sử dụng đường giới hạn khả năng sản xuất (production possibilities frontier – PPF) làm tiêu chí đánh giá hiệu quả (tương đối) giữa các công ty trong cùng một ngành; theo đó các công ty đạt đến mức giới hạn sẽ được coi là hiệu quả (hơn) và các công ty không đạt đến đường PPF sẽ bị coi là kém hiệu quả (so với các công ty kia). Cụ thể, trong Hình 1, các DMU B, C và E có  $TE_B = TE_C = TE_E = 1$ ; còn DMU A và D có  $TE_A = OA/OA' < 1$  và  $TE_D = OD/OD' < 1$ .



**Hình 1.** Đường giới hạn khả năng sản xuất ứng với hai hàng hóa H1 và H2

Đường PPF cho trường hợp sử dụng 2 yếu tố đầu vào ( $x_1$  và  $x_2$ ) để sản xuất ra 1 yếu tố đầu ra ( $y$ ) có thể được biểu diễn như một đường đẳng lượng (isoquant) trong Hình 2. Theo đó, một DMU sản xuất tại vị trí Q được coi là hiệu quả ( $TE_Q = OQ/OQ = 1$ ), trong khi nếu nó sản xuất tại vị trí P là kém hiệu quả ( $TE_P = OQ/OP < 1$ ). Chú ý là Hình 2 giả thiết là với đầu ra  $y$  xác định,  $SS'$  là đường đẳng lượng thể hiện mức kết hợp tối thiểu của  $x_1$  và  $x_2$  và do đó, Hình 2 áp dụng mô hình tối thiểu hóa đầu vào (input-orientation hoặc input-minimization). Nếu giả thiết giữ nguyên đầu vào mà có thể đạt được mức sản lượng đầu ra cao nhất thì đường PPF sẽ có dạng tương tự như trong Hình 1 và khi đó nó áp dụng mô hình tối đa hóa đầu ra (output-orientation hoặc output-maximization).



**Hình 2.** Đường PPF trong trường hợp tối thiểu hóa đầu vào

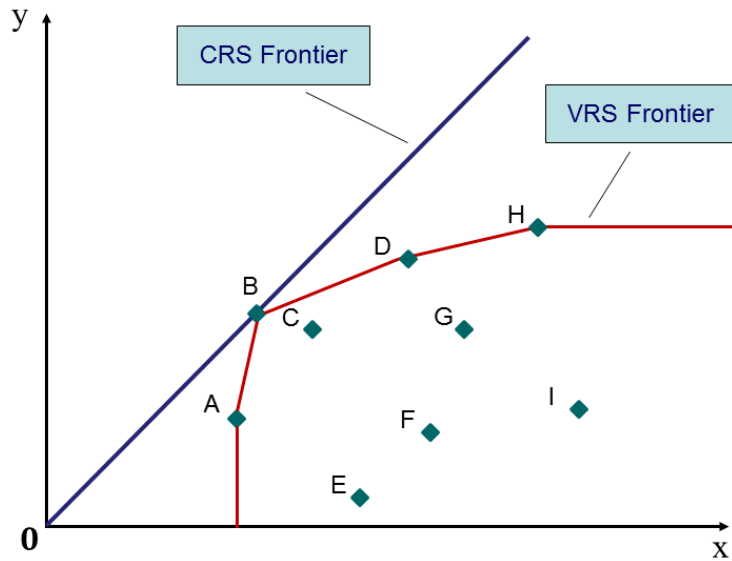
*Nguồn: Farrell (1957)*

Một điểm đáng chú ý khác trong Hình 2 là nếu kết hợp với đường đẳng phí (isocost) AA' thì có thể thấy Q' mới là điểm tối ưu chứ không phải Q. Do đó, nếu DMU sản xuất tại Q thì nó có thể có hiệu quả kỹ thuật (TEQ=1) nhưng lại kém hiệu quả về phân bổ nguồn lực (Q nằm trên đường đẳng phí AA'), do đó QR thể hiện hiệu quả phân bổ đầu vào (allocative efficiency). Một cách ngắn gọn, ta có:

Hiệu quả kinh tế (tổng hợp) = Hiệu quả kỹ thuật x Hiệu quả phân bổ

$$OR/OP = OQ/OP \times OR/OQ$$

Trong phương pháp phân tích giới hạn, một loại hình đồ thị thường gặp khác là đồ thị biểu diễn mối quan hệ giữa tổng đầu vào và tổng đầu ra (Hình 3).



**Hình 3.** Hiệu quả không đổi/thay đổi theo quy mô và đường bao giới hạn PPF

Cách thức để xác định giá trị tổng đầu vào và đầu ra sẽ được trình bày rõ hơn trong mục tiếp theo. Ở đây tác giả muốn lưu ý người đọc về sự khác biệt trong việc xác định đường PPF dưới các điều kiện về hiệu quả không đổi theo quy mô (CRS) hoặc hiệu quả thay đổi theo quy mô (VRS). Theo đó, đường CRS PPF là một đường thẳng nối liền gốc tọa độ và DMU có hiệu quả ( $TE = y/x$ ) cao nhất ( $TE = 1$ ). Đường CRS PPF do đó không tính đến sự khác biệt về quy mô (scale) giữa các DMU mà chỉ đơn giản so sánh các tỷ số hiệu quả ( $TE_i = y_i/x_i$ ) giữa việc sử dụng đầu vào  $x_i$  để tạo ra đầu ra  $y_i$ . Trong khi đó, đường VRS PPF lại tính toán cả đến yếu tố quy mô, vì vậy VRS PPF có hình dạng như một đường bao (envelop) bao quanh các DMU kém hiệu quả khác (CRS PPF cũng là 1 dạng đường bao, nhưng “lỏng lẻo” hơn). Đây chính là nguồn gốc của cái tên phương pháp Phân tích bao dữ liệu.

### Phương pháp phân tích bao dữ liệu DEA

Một cách đơn giản, hiệu quả (mang tính kỹ thuật) của việc sử dụng yếu tố đầu vào  $x$  để thu được yếu tố đầu ra  $y$  có thể được đo lường theo công thức:

$$TE = \frac{\text{Đầu ra}}{\text{Đầu vào}} = \frac{y}{x} \quad (1)$$

Công thức (1) chỉ có thể được áp dụng trong trường hợp chỉ có 1 biến đầu vào (input) và 1 biến đầu ra (output), ví dụ như hiệu quả sử dụng vốn (Doanh thu/Vốn) hay hiệu quả sử dụng lao động (Thu nhập/Lao động). Khi áp dụng cho một doanh nghiệp

(hay gọi chung là DMU) có  $k$  yếu tố đầu vào và sản xuất ra  $m$  kết quả đầu ra, thì cần phải dựa trên giá cả  $p_i$  và  $w_j$  của các yếu tố đầu vào/đầu ra đó để tính toán:

$$TE = \frac{\text{Tổng đầu ra}}{\text{Tổng đầu vào}} = \frac{p_1 y_1 + \dots + p_m y_m}{w_1 x_1 + \dots + w_k x_k} \quad (2)$$

Tuy nhiên, việc xác định giá cả của từng yếu tố đầu vào/đầu ra thường rất phức tạp, nhất là trong những lĩnh vực như tài chính ngân hàng, giáo dục đào tạo,... Trong trường hợp này, có thể giả thiết là mỗi DMU sẽ sử dụng những trọng số nhất định  $u_m$  và  $v_k$  sao cho điểm hiệu quả TE của nó là cao nhất, nói cách khác,  $u_m$  và  $v_k$  là những trọng số giúp cho DMU đó tiến gần đến đường giới hạn khả năng sản xuất PPF nhất. Vì vậy, chúng còn được gọi là “giá ẩn” (shadow prices) vì mặc dù chúng không phải là giá cả thực (true prices) nhưng lại đóng vai trò như giá cả trong việc tính toán hiệu quả kỹ thuật TE.

Một cách tổng quát, với bài toán có  $n$  DMU, mỗi DMU sử dụng  $k$  yếu tố đầu vào  $x_k$  để tạo ra  $m$  yếu tố đầu ra  $y_m$ , việc xác định hiệu quả  $TE_0$  của một DMU<sub>0</sub> bất kỳ sẽ được tính toán như sau:

$$\max_{u,v} TE_0 \quad (3)$$

Trong điều kiện:

$$TE_0 = \frac{\sum u_{0m} y_{0m}}{\sum v_{0k} x_{0k}} \quad (\text{Điểm hiệu quả của DMU}_0)$$

$$TE_j = \frac{\sum u_{jm} y_{jm}}{\sum v_{jk} x_{jk}} \leq 1, \quad j = 1, \dots, n \quad (\text{Điểm hiệu quả của tất cả các DMU không vượt quá 1, tức là không vượt quá khỏi đường PPF})$$

$$u_m, v_k \geq 0 \quad (\text{Các “giá ẩn” là không âm})$$

Charnes và đồng sự (1978) đã áp dụng phương pháp tối ưu hóa tuyến tính phi tham số (non-parametric linear optimization) vào việc giải quyết công thức (3), ứng với giả thiết hiệu quả không đổi theo quy mô (CRS). Sau đó, Banker và đồng sự (1984) đã phát triển bài toán này cho trường hợp hiệu quả thay đổi theo quy mô (VRS). Đến nay, đã có khá nhiều mô hình DEA được phát triển như Malmquist DEA, network DEA, SBM DEA,... (xem thêm Cooper *et al.*, 2006; Cook & Seiford, 2009; Paradi *et al.*, 2011), nhưng bản chất vẫn dựa trên mô hình cơ bản của công thức (3). Công thức (3) và các mô hình DEA cơ bản dựa trên nó (CRS I, VRS I, CRS O, và VRS O – xem thêm Mục 3) vẫn là một chuẩn mực trong nghiên cứu DEA: dù nghiên cứu có phức tạp thế nào thì đầu tiên vẫn

phải phân tích các mô hình cơ bản nói trên. Do đó, trong phiên bản đầu tiên của VDEA, tác giả chỉ tập trung giải quyết các mô hình DEA cơ bản nói trên. Các phiên bản tiếp theo sẽ tiếp tục mở rộng tới các mô hình khác.

### **3. Ước lượng năng suất tổng hợp bằng chỉ số Malmquist trong DEA**

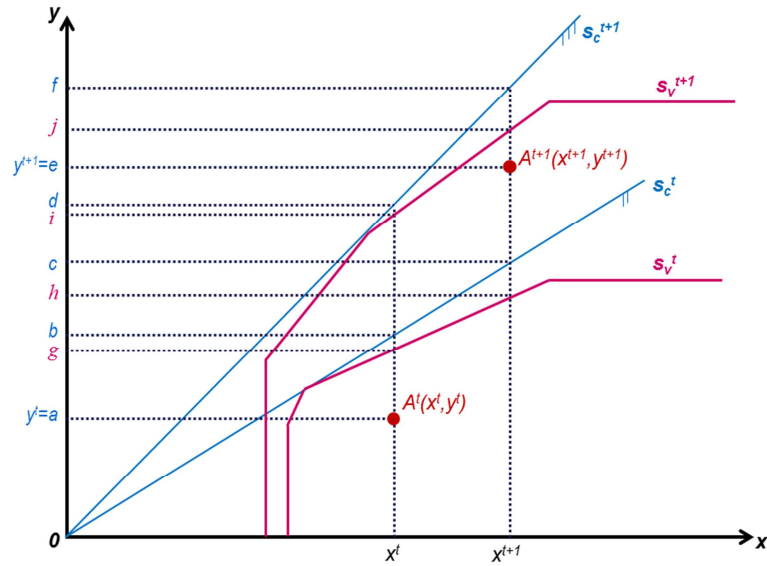
Bên cạnh việc tính toán hiệu quả kỹ thuật tại một thời điểm nhất định, việc tính toán hiệu quả theo thời gian cũng quan trọng không kém. Sự so sánh các mức hiệu quả giữa các giai đoạn khác nhau giúp các nhà nghiên cứu có cái nhìn rõ nét hơn về sự thay đổi của hiệu quả theo thời gian, từ đó có thể đánh giá về những thay đổi trong các giai đoạn đó có tác động thế nào tới hiệu quả, cũng như có thể phần nào dự báo được biến động của hiệu quả trong tương lai. Trong phiên bản 1.3 lần này, VDEA đã được tích hợp để sử dụng trong tính toán sự thay đổi của năng suất tổng hợp theo thời gian theo chỉ số Malmquist nhằm đáp ứng yêu cầu nghiên cứu nói trên.

Trong phương pháp DEA, việc ước lượng hiệu quả kỹ thuật được thực hiện dựa trên một đường giới hạn (frontier) xác định, và do đó, so sánh hiệu quả giữa hai giai đoạn dựa trên hai đường giới hạn khác nhau là rất phức tạp. Tuy nhiên, nếu quy về cùng một gốc tọa độ thì vấn đề trở nên đơn giản hơn với sự giúp đỡ của các hàm khoảng cách (distance functions)<sup>2</sup>. Fare và đồng sự (1994) đưa ra mô hình xác định mức thay đổi của năng suất tổng hợp theo thời gian trong đó một DMU bất kỳ sẽ được nghiên cứu tại hai thời điểm khác nhau  $t$  và  $t+1$  (tương ứng với hai đường frontier khác nhau tại hai thời điểm  $t$  và  $t+1$ ) rồi so sánh sự thay đổi về năng suất tổng hợp của DMU đó (Hình 4).

---

<sup>2</sup> Xem thêm Caves, Christensen, and Diewert (1982) và Shephard (1970).





**Hình 4.** Chỉ số Malmquist TFP đầu ra (output-based)

*Nguồn: Fare và đồng sự (1994)*

Cụ thể, trong điều kiện hiệu quả không đổi theo quy mô, chỉ số Malmquist TFP của DMU A tại thời điểm  $t$  (điểm  $A_t$ ) so với thời điểm  $t+1$  ( $A_{t+1}$ ) có thể được tính toán theo trung bình nhân (geometric mean) của hai chỉ số Malmquist đầu ra (output-based Malmquist index): chỉ số thứ nhất lấy đường giới hạn tại thời điểm  $t$  làm cơ sở tính toán, chỉ số thứ hai lấy đường giới hạn thời điểm  $t+1$  làm cơ sở tính toán.

$$\begin{aligned}
 m_0(x^{t+1}, y^{t+1}, x^t, y^t) &= \left[ \frac{d_c^t(x^{t+1}, y^{t+1})}{d_c^t(x^t, y^t)} \times \frac{d_c^{t+1}(x^{t+1}, y^{t+1})}{d_c^{t+1}(x^t, y^t)} \right]^{\frac{1}{2}} \\
 &= \frac{d_c^{t+1}(x^{t+1}, y^{t+1})}{d_c^t(x^t, y^t)} \times \left[ \frac{d_c^t(x^{t+1}, y^{t+1})}{d_c^{t+1}(x^{t+1}, y^{t+1})} \times \frac{d_c^t(x^t, y^t)}{d_c^{t+1}(x^t, y^t)} \right]^{\frac{1}{2}} \\
 &= \left( \frac{0e}{0f} \frac{0b}{0a} \right) \times \left[ \frac{0f}{0c} \frac{0d}{0b} \right]^{\frac{1}{2}}
 \end{aligned} \tag{1}$$

$$\text{TFPCH} = (\text{EFCH}) \times [\text{TECHCH}]$$

Nếu áp dụng cho trường hợp hiệu quả thay đổi theo quy mô (VRS), ta có:

$$\begin{aligned}
EFCH &= PECH \times SECH \\
&= \frac{d_v^{t+1}(x^{t+1}, y^{t+1})}{d_v^t(x^t, y^t)} \times \frac{s^{t+1}(x^{t+1}, y^{t+1})}{s^t(x^t, y^t)} \\
&= \frac{0e \ 0g}{0j \ 0a} \times \frac{0j \ 0b}{0f \ 0g}
\end{aligned} \tag{2}$$

Một cách tổng quát:

$$TFPCH = (PECH \times SECH) \times [TECHCH]$$

Trong đó:

- TFPCH: Mức thay đổi của năng suất tổng hợp (Chỉ số Malmquist TFP)
- EFCH: Mức thay đổi của hiệu quả kỹ thuật (trong điều kiện CRS)
- PECH: Mức thay đổi của hiệu quả kỹ thuật thuần (trong điều kiện VRS)
- TECHCH: Mức thay đổi của công nghệ hay đường giới hạn (frontier)
- SECH: Mức thay đổi của hiệu quả nhờ quy mô (trong điều kiện VRS)

Như vậy, khi  $m_0$  hay  $TFPCH > 1$ , năng suất tổng hợp của DMU A đã có sự gia tăng tại thời điểm  $t+1$  so với tại thời điểm  $t$ . Nếu  $m_0 < 1$ , ta nói năng suất tổng hợp của DMU A bị suy giảm trong gian đoạn từ  $t$  đến  $t+1$ . Cách sử dụng cũng như đọc kết quả phân tích chỉ số Malmquist cũng như các kết quả khác được trình bày trong mục tiếp theo.

## 4. VDEA phiên bản 2.0

### 4.1. Hướng dẫn cài đặt

- Download add-in VDEA cho Excel phiên bản 2.0 (file “VDEA ver 2.0.xlam”) tại đây: <https://www.dropbox.com/s/gcyh2d85r4hmqzo/VDEA%20ver%202.0.xlam?dl=0>

- Chạy VDEA ver 2.0.xlam, Excel sẽ hỏi xem có cho phép sử dụng tiện ích này không, chọn “Enable Macro”.

- VDEA sẽ thiết lập một menu điều khiển có tên “Vietnamese DEA” trong mục Add-Ins của Excel.

### 4.2. Hướng dẫn chuẩn bị file dữ liệu:

- Đối với phân tích hiệu quả kỹ thuật TE: Mô hình này sử dụng dữ liệu chéo (cross-sectional data) trong đó tại thời điểm cần nghiên cứu, các DMUs cùng hoạt động

trong một lĩnh vực sử dụng các yếu tố đầu vào (inputs) giống nhau để tạo ra các yếu tố đầu ra (outputs) giống nhau (tham khảo Hình 5a).

- Đối với tính toán năng suất tổng hợp Malmquist TFP: Mô hình này sử dụng dữ liệu bảng (panel data) theo đó dữ liệu của các năm được sắp xếp theo trình tự của từng DMUs và được bố trí từ trên xuống dưới theo các năm (tham khảo Hình 5b). Lưu ý là dữ liệu bảng này phải cân xứng (balanced).

Hình 5a.

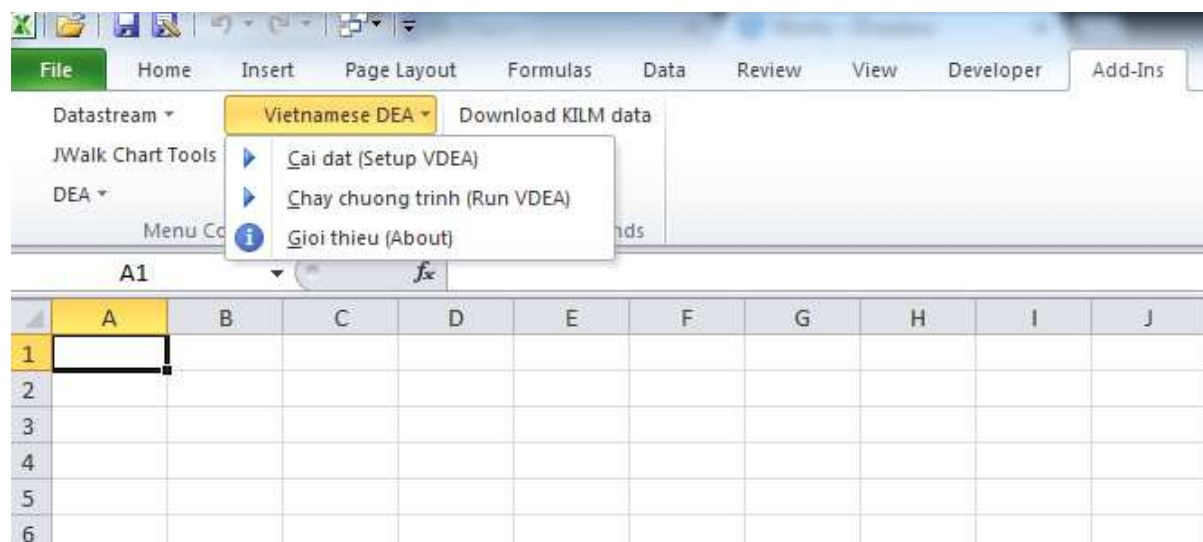
	A	B	C	D	E	F	G
1	DMU	X1	X2	Y1	Y2		
2	a	1	1	2	2		
3	b	2	2	1	3		
4	c	3	3	2	3		
5	d	3	2	1	4		
6	e	4	1	1	2		
7							
8							
9							
10							
11							

Hình 5b.

	A	B	C	D	E	F	G
1	DMU	X1	X2	Y1	Y2		
2	a	1	1	2	2		
3	b	2	2	1	3		<b>Năm 1</b>
4	c	3	3	2	3		
5	d	3	2	1	4		
6	e	4	1	1	2		
7		5	5	6	6		
8		6	6	5	7		<b>Năm 2</b>
9		7	7	6	7		
10		7	6	5	8		
11		8	5	5	6		
12							
13							
14							
15							

### 4.3. Hướng dẫn sử dụng

- Trên thanh công cụ của Excel, chọn Add-Ins, chọn tiếp Vietnamese DEA (như Hình 6) để chạy VDEA.



**Hình 6.** Trình điều khiển của VDEA trong menu Add-Ins của Excel 2010

- Nếu chọn “Cài đặt (Setup VDEA)”: VDEA sẽ hiển thị cửa sổ Cài đặt để thiết lập Ngôn ngữ (Tiếng Anh hoặc Tiếng Việt); Độ chính xác (của việc tính toán chỉ số hiệu quả) có giá trị từ 0 đến 1, trong đó giá trị càng nhỏ thì mức độ chính xác càng cao; và lựa chọn Chuẩn hóa số liệu (quy đổi các biến đầu vào/đầu ra theo cùng một mức tỷ lệ - scale - để dễ tính toán) theo đó tỷ lệ giữa giá trị lớn nhất và giá trị nhỏ nhất của 1 biến số bất kỳ được khuyến nghị không vượt quá 1:1000 nếu thiết lập này được chọn (xem thêm SAITECH Inc., 2012, p. 29).

- Nếu chọn “Giới thiệu (About)”: VDEA sẽ hiển thị thông tin về tác giả cũng như số hiệu phiên bản (hiện tại là bản 2.0).

- Nếu chọn “Chạy chương trình (Run VDEA)”: VDEA sẽ hiển thị cửa sổ nhập số liệu bao gồm các nội dung sau:

**Hình 7.** Cửa sổ Nhập số liệu của VDEA

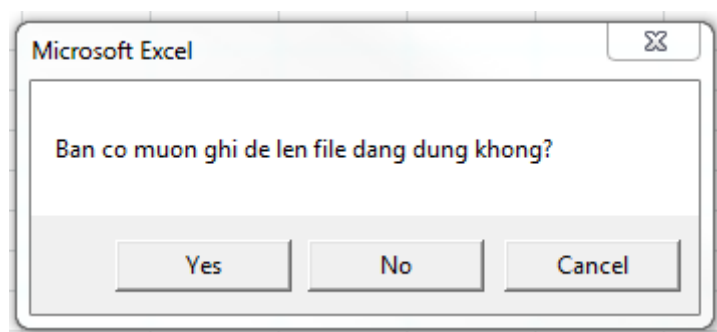
1. Nhập số liệu: Sử dụng con chuột (mouse) để chọn/bôi đen các hàng/cột Excel có chứa thông tin về tên của các DMUs, tên các biến đầu vào/đầu ra, và số liệu biến đầu vào/đầu ra tương ứng.

2. Chọn mô hình VDEA: Chọn mô hình “Tinh hieu qua ky thuat (cross-section)” nếu chỉ cần tính hiệu quả trong 1 năm/thời điểm xác định. Nếu muốn ước lượng năng suất nhân tố tổng hợp thay đổi theo thời gian, mời chọn “Tinh nang suat tong hop Malmquist TFP (balanced panel)”. Trong trường hợp này, số liệu phải cân xứng (balanced) và được trình bày theo dạng bảng (panel) – xem thêm mục **Hướng dẫn chuẩn bị số liệu**, đồng thời số liệu về giai đoạn nghiên cứu cũng cần phải được nhập vào trong mục “Co bao nhieu giai doan (periods)?” (giá trị mặc định là 1 nếu là cross-section hoặc 2 nếu là Malmquist TFP).

3. Chọn thông số: Hiện nay VDEA chấp nhận 4 mô hình ước lượng hiệu quả kỹ thuật tối thiểu hóa đầu vào không đổi theo quy mô (CRS I), tối thiểu hóa đầu vào thay đổi theo quy mô (VRS I), tối đa hóa đầu ra không đổi theo quy mô (CRS O), và tối đa hóa đầu ra thay đổi theo quy mô (VRS O).

4. Thiết lập lựa chọn để tính toán trọng số tối ưu hay “giá ẩn” (shadow prices): Nếu bạn muốn VDEA đồng thời thực hiện tính toán bộ trọng số tối ưu  $u$  và  $v$  (xem công thức 3 ở trên) thì chọn “Có chứ”. Nếu không thì chọn “Không cần”. VDEA thiết lập “Không cần” là lựa chọn mặc định.

5. Chọn OK để chạy chương trình và xuất kết quả. Chương trình sẽ hỏi xem bạn muốn lưu kết quả vào đâu. Chọn “Yes” để ghi đè (overwrite) lên file đang sử dụng. Chọn “No” để mở cửa sổ Save As và lưu file mới. Chọn “Cancel” để xuất kết quả ra file hiện tại nhưng không lưu lại.



**Hình 8.** Lựa chọn ghi lại kết quả của VDEA

#### **4.4. Hướng dẫn đọc kết quả Tính hiệu quả kỹ thuật (cross-section)**

Sau khi chạy chương trình, VDEA sẽ xuất kết quả ra sheet “VDEA results”. Các kết quả được trình bày trong sheet này bao gồm Số liệu gốc, Mức độ cải thiện của các biến (slacks), DMU nào là mục tiêu phấn đấu của các DMU kém hiệu quả (peers), điểm hiệu quả kỹ thuật (TE) của từng DMU, và trạng thái quy mô sản xuất mà DMU đang gặp phải (CRS, IRS hoặc DRS tương ứng với quy mô tối ưu, lợi thế nhờ quy mô, hoặc bất lợi thế nhờ quy mô) nếu lựa chọn thông số Hiệu quả thay đổi theo quy mô ở Hình 7.

	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T
1	So lieu goc						Muc do cai thien					Muc tieu phan dau (				Ket qua tinh toan theo mo hinh DEA VRS I				
2	DMUs	X1	X2	Y1	Y2		X1 slacks	X2 slacks	Y1 slacks	Y2 slacks		DMU(trong so)				CRS TE	VRS TE	SE	NIRS TE	Trang thai
3	a	1	1	2	2		0.000	0.000	0.000	0.000		a(1)				1.000	1.000	1.000	1.000	CRS
4	b	2	2	1	3		0.000	0.500	-0.500	0.000		a(0.5)		d(0.5)		0.750	1.000	0.750	1.000	DRS
5	c	3	3	2	3		0.000	0.000	0.000	0.000		c(1)				0.500	1.000	0.500	1.000	DRS
6	d	3	2	1	4		0.000	0.000	0.000	0.000		d(1)				1.000	1.000	1.000	1.000	CRS
7	e	4	1	1	2		3.000	0.000	-1.000	0.000		a(1)				1.000	1.000	1.000	1.000	CRS
8																				

**Hình 9.** Sheet “VDEA results”: Ước lượng hiệu quả kỹ thuật

Nếu VDEA được thiết lập để tính toán “giá ẩn”, kết quả tính toán sẽ có dạng như trong Hình 10. Điểm khác biệt giữa Hình 10 và Hình 9 là chúng ta có thêm thông tin về các “giá ẩn” tương ứng cho từng biến đầu vào và đầu ra của từng DMU, ví dụ như v1 là “giá ẩn” của X1, u2 là “giá ẩn” của Y2. Các “giá ẩn” u và v này là tối ưu cho từng DMU, theo đó TE của DMU này nếu tính theo u và v sẽ có giá trị lớn nhất:

$$TE = \frac{\sum u_m y_m}{\sum v_k x_k} \text{ maximum} \quad (\text{Điểm hiệu quả của DMU là tối đa})$$

	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U	
1	So lieu goc						Muc do cai thien					Muc tieu phan dau (				Ket qua tinh toan theo mo hinh DEA VRS I						
2	DMUs	X1	X2	Y1	Y2		X1 slacks	X2 slacks	Y1 slacks	Y2 slacks		DMU(trong so)				CRS TE			v1	v2	u1	u2
3	a	1	1	2	2		0.000	0.000	0.000	0.000		a(1)				1.000			0.000	4.036	0.000	2.018
4	b	2	2	1	3		0.000	0.000	-2.000	0.000		a(1.5)			0.750			0.000	4.036	0.000	2.018	
5	c	3	3	2	3		0.000	0.000	-1.000	0.000		a(1.5)			0.500			0.000	4.036	0.000	2.018	
6	d	3	2	1	4		0.000	0.000	0.000	0.000		d(1)				1.000			0.000	4.036	0.000	2.018
7	e	4	1	1	2		2.667	0.000	0.000	0.000		a(0.333)		d(0.333)		1.000			0.000	3.511	0.000	1.756
8																						

**Hình 10.** Sheet “VDEA results”: Ước lượng hiệu quả kỹ thuật có tính “giá ẩn”

#### 4.5. Hướng dẫn đọc kết quả Tính năng suất tổng hợp Malmquist TFP

Sau khi chạy chương trình, VDEA sẽ xuất kết quả ra sheet “VDEA results”. Các kết quả được trình bày bao gồm Số liệu gốc, Mức thay đổi của hiệu quả kỹ thuật (EFCH) ứng với điều kiện CRS, Mức thay đổi của công nghệ (TECHCH) hay còn gọi là thay đổi của đường giới hạn (frontier shifts), và Mức thay đổi năng suất tổng hợp (TFPCH) – xem Hình 11. Nếu lựa chọn thông số Hiệu quả thay đổi theo quy mô (ở Hình 7) thì có thêm kết quả về Mức thay đổi của hiệu quả kỹ thuật thuần túy (PECH) và Mức thay đổi của hiệu quả theo quy mô (SECH).

	A	B	C	D	E	F	G	H	I	J	K	L	M
1	So lieu goc						Ket qua tinh toan theo mo hinh DEA VRS I						
2	DMUs	X1	X2	Y1	Y2		DMUs(t1->EFCH	TECHCH	TFPCH	PECH	SECH		
3	a	1	1	2	2								
4	b	2	2	1	3								
5	c	3	3	2	3								
6	d	3	2	1	4								
7	e	4	1	1	2								
8		5	5	6	6		a(t0->t1)	1.000	0.600	0.600	1.000	1.000	
9		6	6	5	7		b(t0->t1)	1.296	0.600	0.778	1.000	1.296	
10		7	7	6	7		c(t0->t1)	1.667	0.600	1.000	1.000	1.667	
11		7	6	5	8		d(t0->t1)	1.000	0.667	0.667	1.000	1.000	
12		8	5	5	6		e(t0->t1)	0.956	0.647	0.618	1.000	0.956	
13													

**Hình 11.** Sheet “VDEA results”: Ước lượng năng suất tổng hợp Malmquist

## 5. Kết luận và hướng phát triển

Nhìn chung, VDEA chạy ổn định trong môi trường Excel 2010. Do sử dụng hàm Solver có sẵn của Excel để thực hiện giải thuật tối ưu hóa tuyến tính, một hạn chế của VDEA là chỉ có thể được sử dụng tối đa cho 200 DMU và 100 biến (đầu vào + đầu ra)<sup>3</sup>. Phiên bản tiếp theo của VDEA dự kiến sẽ được mở rộng thêm cho mô hình Fisher DEA (sử dụng với bảng số liệu theo thời gian – time series data)<sup>4</sup>, cũng như tích hợp tính toán theo mô hình slacks-based (SBM) hoặc network DEA. Mọi góp ý, báo lỗi,... xin gửi về [T.Ngo@massey.ac.nz](mailto:T.Ngo@massey.ac.nz). Trân trọng.

<sup>3</sup> Tuy nhiên, đối với hầu hết nghiên cứu về DEA thì như vậy là đủ dùng.

<sup>4</sup> Xem thêm Ngo và Tripe (2014).



## Tài liệu tham khảo

- Aigner, D. J., & Chu, S. F. (1968). On estimating the industry production function. *American Economic Review*, 58(4), 826-839.
- Banker, R. D., Charnes, A., & Cooper, W. W. (1984). Some models for estimating technical and scale inefficiencies in data envelopment analysis. *Management Science*, 30(9), 1078-1092.
- Caves, D. W., Christensen, L. R., & Diewert, W. E. (1982). The economic theory of index numbers and the measurement of input, output and productivity. *Econometrica*, 50, 1393-1414.
- Charnes, A., Cooper, W. W., & Rhodes, E. (1978). Measuring the efficiency of decision making units. *European Journal of Operational Research*, 2, 429-444.
- Coelli, T. J. (1996). *A Guide To DEAP Version 2.1: A Data Envelopment Analysis (Computer) Program*. CEPA Working Paper No. 8/96. Department of Econometrics. University of New England.
- Cook, W. D., & Seiford, L. M. (2009). Data envelopment analysis (DEA) – Thirty years on. *European Journal of Operational Research*, 192(1), 1-17.
- Cooper, W. W., Seiford, L. M., & Tone, K. (2006). *Data Envelopment Analysis: A Comprehensive Text with Models, Applications, References, And DEA-Solver Software* (2nd ed.): Springer
- Färe, R., Grosskopf, S., Norris, M., & Zhang, Z. (1994). Productivity growth, technical progress, and efficiency change in industrialized countries. *American Economic Review* 84(1), 66-83.
- Farrell, M. J. (1957). The measurement of productive efficiency. *Journal of the Royal Statistical Society*, 120(3), 253-281.
- Fuentes, H. J., Grifell-Tatje, E., & Perelman, S. (2001). A parametric distance function approach for Malmquist productivity index estimation. *Journal of Productivity Analysis*, 15, 79-94.
- Gong, B.-H., & Sickles, R. C. (1992). Finite sample evidence on the performance of stochastic frontiers and data envelopment analysis using panel data. *Journal of Econometrics*, 51(1-2), 259-284.
- Malmquist, S. (1953). Index numbers and indifference surfaces. *Trabajos de Estadística*, 4(2), 209-242.
- Ngo, D. T., & Tripe, D. (2014). *Nonparametric Fisher Total Factor Productivity Index using Shadow Prices: Panel vs. Time Series Data*. Paper presented at the PMAA 2015, The University of Auckland, Auckland, NZ.
- Nishimizu, M., & Page, J. M., Jr. (1982). Total factor productivity growth, technological progress and technical efficiency change: Dimensions of productivity change in Yugoslavia, 1965-78. *Economic Journal*, 92(368), 920-936.
- Paradi, J., Yang, Z., & Zhu, H. (2011). Assessing Bank and Bank Branch Performance: Modeling Considerations and Approaches. In W. W. Cooper, L. M. Seiford & J. Zhu (Eds.), *Handbook on Data Envelopment Analysis* (pp. 315-361): Springer US.

- SAITECH Inc. (2012). *Introduction to DEA-Solver-Pro Professional Version 9.0*. SAITECH, Inc., Hazlet, New Jersey.
- Schmidt, P., & Sickles, R. C. (1984). Production frontiers and panel data. *Journal of Business & Economic Statistics*, 2(4), 367-374.
- Shephard, R. W. (1970). *Theory of cost and production functions*. Princeton, NJ: Princeton University Press.
- Solow, R. M. (1957). Technical change and the aggregate production function. *Review of Economics and Statistics*, 39(3), 312-320.
- Tulkens, H., & vanden Eeckaut, P. (1995). Non-parametric efficiency, progress and regress measures for panel data: Methodological and aspects. *European Journal of Operational Research*, 80, 474-499.