



DATA ENVELOPMENT ANALYSIS

МЕТОД ОБОЛОЧЕЧНОГО АНАЛИЗА ДАННЫХ

Автор презентации:

Соколова Т.В., к.ф.-м.н., аналитик ЛАФР

Москва, 2018

Литература

Учебник:

1. Cooper W., Seiford L., Tone K. (2007). Data Envelopment Analysis. A Comprehensive Text with Models, Applications, References and DEA-Solver Software. Springer, 2nd Edition.

Примеры применения DEA для решения задач финансовой экономики

1. Выгон Г.В. (2001). Методы оценки нефтяных компаний в условиях неопределенности. *Аудит и финансовый анализ*. №1.
2. Arshinova T. (2011). Construction of Equity Portfolio on the Basis of Data Envelopment Analysis Approach. *Scientific Journal of Riga Technical University*. Vol. 49.
3. Karimzadeh M. (2012). Efficiency Analysis by Using Data Envelopment Analysis Model: Evidence from Indian Banks. *Int. J Latest Trends Fin. Econ*. Vol. 2. №.3. P. 228-237.
4. Теплова Т.В., Соколова Т.В. (2017). Непараметрический метод оболочечного анализа для портфельных построений на российском рынке облигаций. *Экономика и математические методы*. Т. 53. №3. С.110-128.

Назначение DEA

Назначение DEA – оценка и сопоставление по эффективности единиц принятия решения (в задачах экономики - экономических агентов).

Мера эффективности – коэффициент вида $\frac{Output}{Input}$

В простейшем случае:

Output – выходной параметр (например, выпуск продукции)

Input – входной параметр (например, затраты)

Как сопоставить экономических агентов по нескольким входным и выходным параметрам?

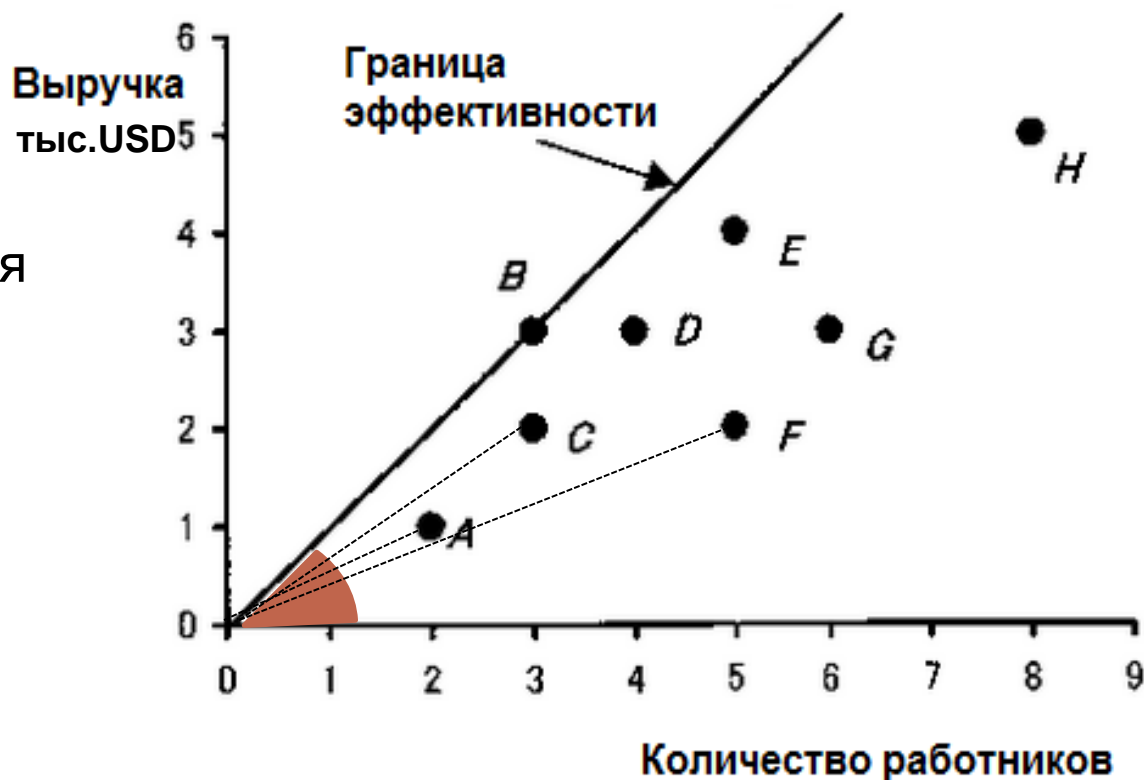
Преимущества DEA

- ✓ **Не требует задания** пользователем **весов** входных и выходных параметров
 -
- ✓ **Не требует** формулирования и проверки **гипотез о функциональных связях** между входными и выходными параметрами (в отличие от регрессионного анализа)
- ✓ Предоставляет возможность работы с **большим количеством** входных и выходных **параметров** (в отличие от регрессионного анализа)

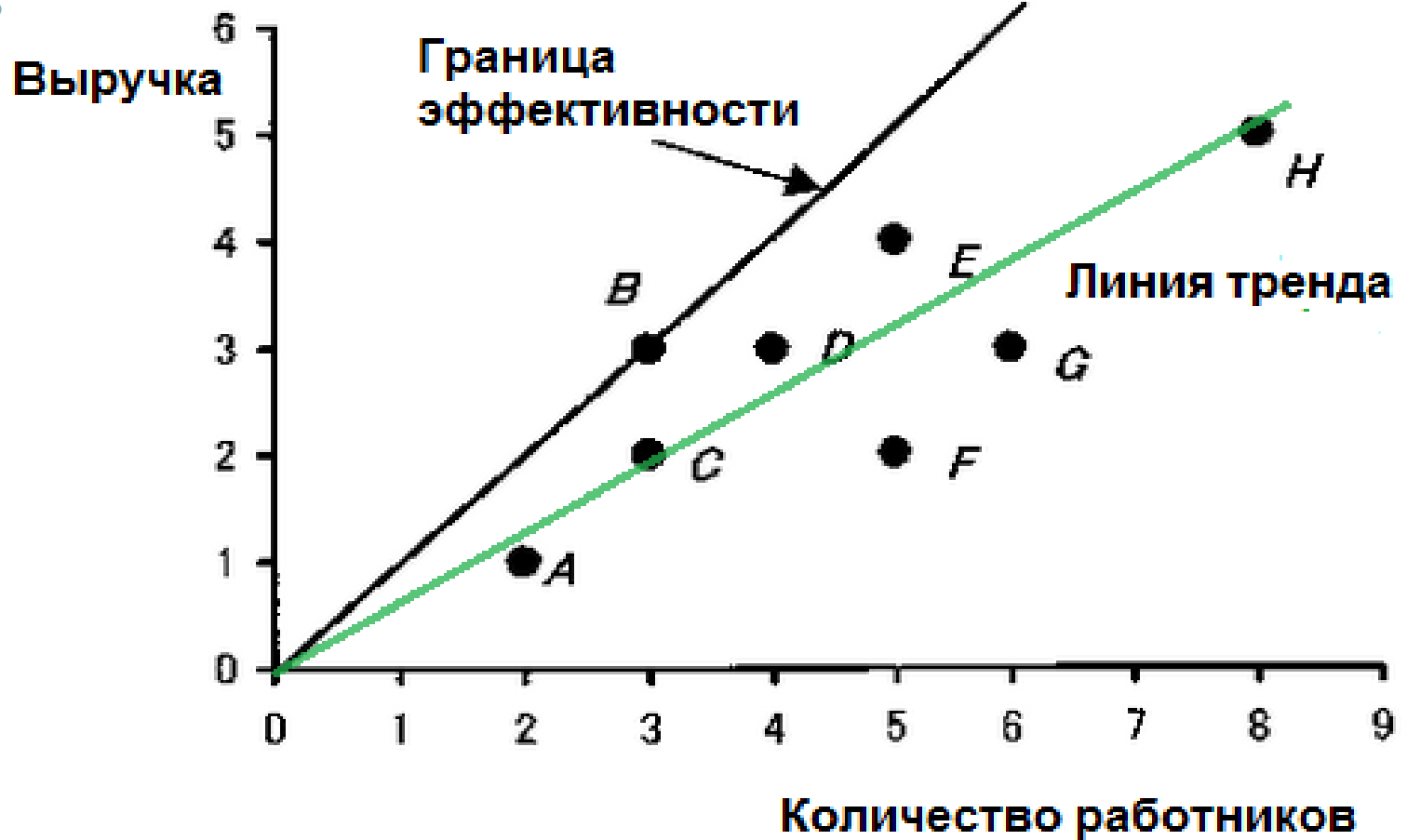
Пример 1. Модель DEA с единственным входным и единственным выходным параметром

Филиал	A	B	C	D	E	F	G	H
Работников	2	3	3	4	5	5	6	8
Выручка, тыс. USD	1	3	2	3	4	2	3	5
Выручка на работника	0,5	1	0,667	0,75	0,8	0,4	0,5	0,625

Граница эффективности – линия, соединяющая начало координат и точки филиалов, с наибольшим углом наклона



Регрессионный анализ и DEA

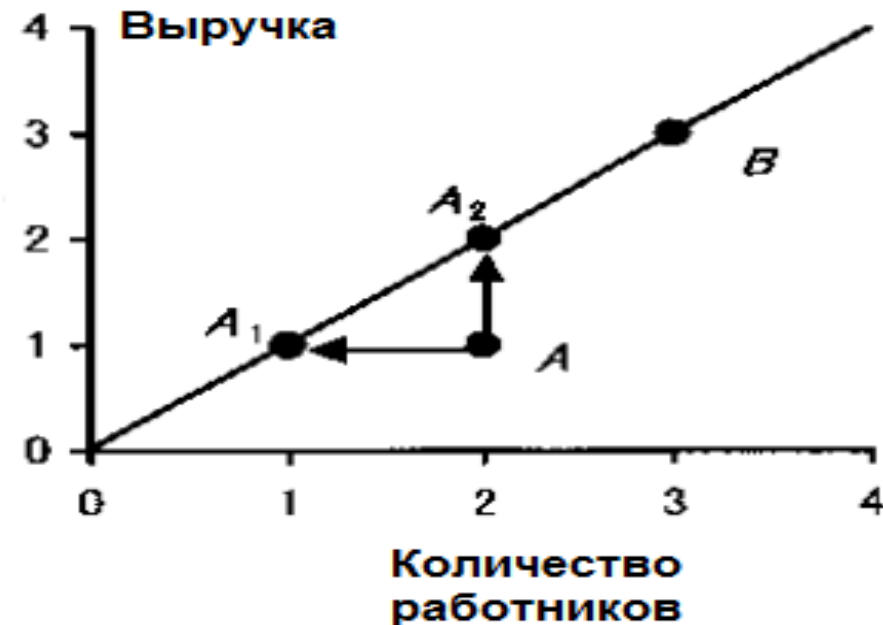


Оценка эффективности предприятий

$$0 \leq \frac{\text{Выручка на работника для предприятия } i}{\text{Выручка на работника для предприятия – лидера}} \leq 1$$

Филиал	A	B	C	D	E	F	G	H
Выручка на работника (тыс.\$)	0,5	1	0,667	0,75	0,8	0,4	0,5	0,625
Эффективность (% от эффективности лидера)	0,5	1	0,667	0,75	0,8	0,4	0,5	0,625

Как повысить эффективность предприятия A путем изменения входных и выходных параметров?

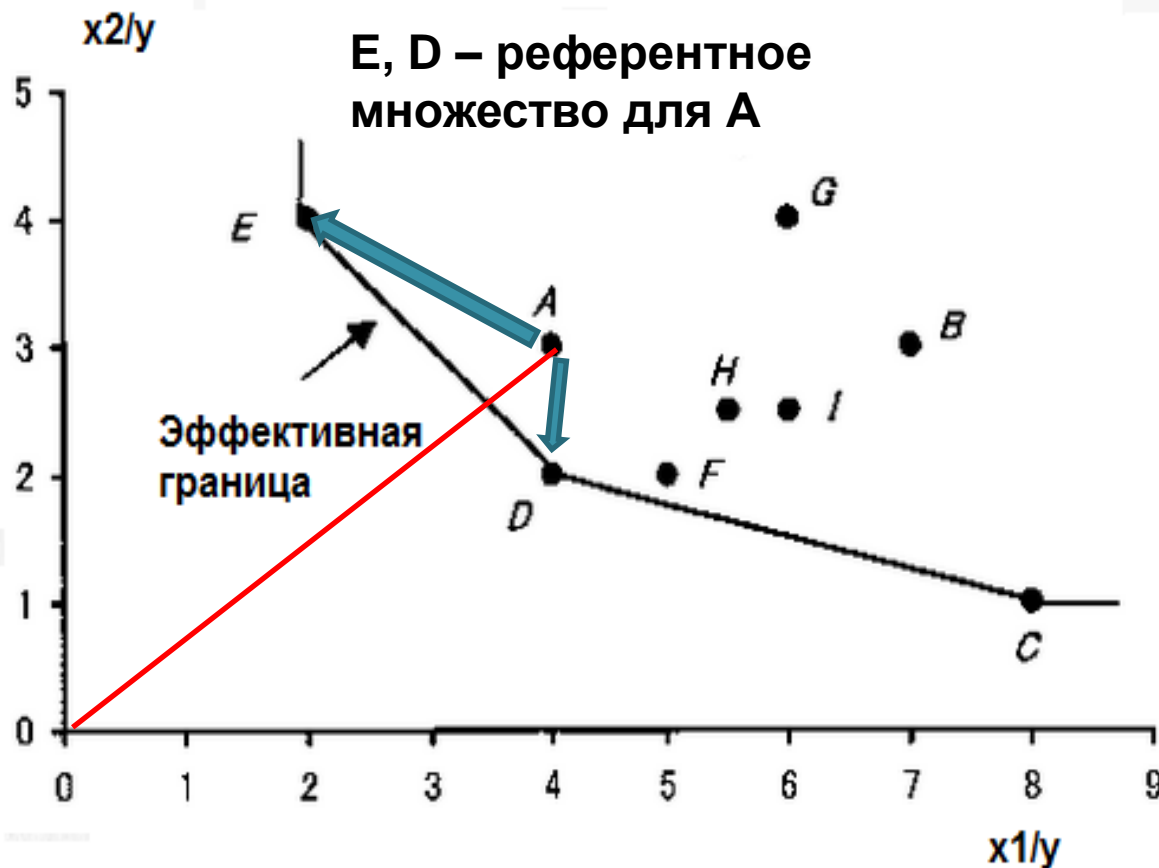


Пример 2. Модель DEA с двумя входными и единственным выходным параметром

Филиал	A	B	C	D	E	F	G	H	I
Работников (x1)	4	7	8	4	2	5	6	5.5	6
Площадь торг. зала (x2)	3	3	1	2	4	2	4	2.5	2.5
Выручка (y)	1	1	1	1	1	1	1	1	1

Более эффективны те филиалы, которые используют меньше входов на единицу выхода.

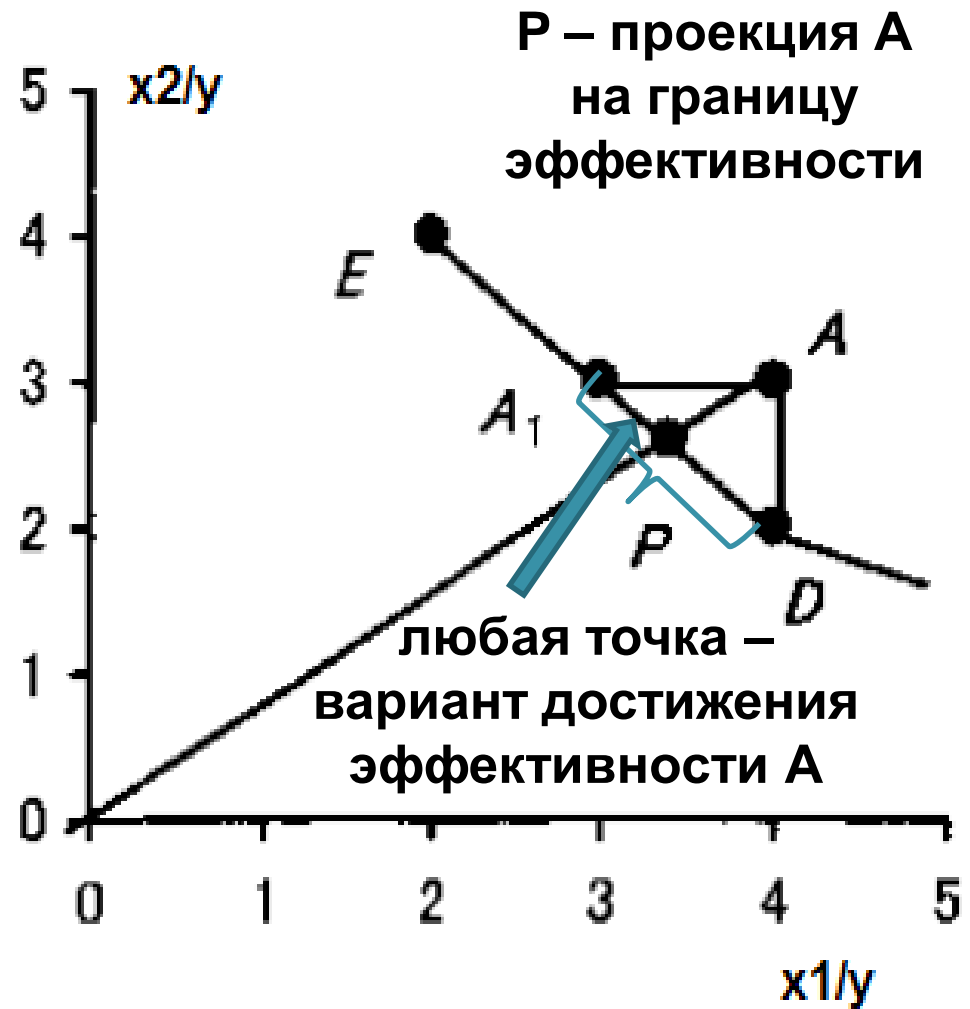
Предприятия на границе эффективности могут уменьшить один входной параметр только за счет увеличения другого входного параметра



Оценка эффективности предприятий

$$\text{Эффективность } A = \frac{OP}{OA}$$

Как повысить эффективность предприятия А?



Пример 3. Модель DEA с единственным входным и двумя выходными параметрами

Филиал	A	B	C	D	E	F	G
Работников (x1)	1	1	1	1	1	1	1
Покупателей (y1)	1	2	3	4	4	5	6
Выручка (y2)	5	7	4	3	6	5	2

Эффектив. $D = OD / OP$
 Координаты $P (16/3; 4)$ –
 это пересечение прямых

$$\begin{cases} y_2 = 3/4 \cdot y_1 & (OD) \\ y_2 = 20 - 3 \cdot y_1 & (FG) \end{cases}$$

$$OD = \sqrt{4^2 + 3^2} = 5$$

$$OP = \sqrt{(16/3)^2 + 4^2} = 20/3$$

$$\text{Эфф}(D) = 0.75 = 5 / (20/3)$$

$$D \dashrightarrow P$$

$$(4/3 * 4, 4/3 * 3) = (16/3, 4)$$



Переход к большому количеству параметров

Подходы:

- ✓ Присвоить **фиксированные, заранее определенные** веса каждому параметру и рассматривать взвешенную сумму параметров
- ✓ DEA – использовать **переменные веса параметров**
 - ✓ **веса присваиваются автоматически**
 - ✓ **выбираются наилучшие веса:** максимизируется эффективность каждого предприятия при условии, что для измерения эффективности всех остальных предприятий применяются те же веса параметров (в рамках одной задачи DEA!)

Модель CCR (Charnes, Coopers, Rhodes, 1978)

x_{1j}, \dots, x_{mj} – входные параметры для DMU_j

y_{1j}, \dots, y_{sj} – выходные параметры для DMU_j

v_1, \dots, v_m – веса входных параметров

u_1, \dots, u_s – веса выходных параметров

Для каждой DMU

$$\text{Virtual input} = v_1 x_{10} + \dots + v_m x_{m0}$$

$$\text{Virtual output} = u_1 y_{10} + \dots + u_s y_{s0}$$

$$\frac{\text{Virtual output}}{\text{Virtual input}} \rightarrow \max \quad \begin{array}{l} x_{10}, x_{20}, \dots, x_{m0} > 0 \\ y_{10}, y_{20}, \dots, y_{s0} > 0 \end{array}$$

при ряде ограничений

DMU =
единица
принятия
решений
(Decision
Making
Unit)

Оптимальные веса различны для различных DMU!

**Выбор параметров: какие значения предпочтительнее –
большие или меньшие?**

Модель ССР – постановка задачи

m – количество входных параметров,

s – количество выходных параметров,

n – количество DMU

DMU_j input : $(x_{1j}, x_{2j}, \dots, x_{mj})^T$, output $(y_{1j}, y_{2j}, \dots, y_{sj})^T$

X – матрица
входных
параметров

$$X = \begin{pmatrix} x_{11} & x_{12} & \dots & x_{1n} \\ x_{21} & x_{22} & \dots & x_{2n} \\ \dots & \dots & \dots & \dots \\ x_{m1} & x_{m2} & \dots & x_{mn} \end{pmatrix}$$

Y – матрица
выходных
параметров

$$Y = \begin{pmatrix} y_{11} & y_{12} & \dots & y_{1n} \\ y_{21} & y_{22} & \dots & y_{2n} \\ \dots & \dots & \dots & \dots \\ y_{s1} & y_{s2} & \dots & y_{sn} \end{pmatrix}$$

Модель ССР – постановка задачи

n задач оптимизации (для каждой DMU в отдельности)

$$\max_{v,u} \theta = \frac{u_1 y_{1o} + u_2 y_{2o} + \dots + u_s y_{so}}{v_1 x_{1o} + v_2 x_{2o} + \dots + v_m x_{mo}} \quad (o = 1, \dots, n)$$

$$\frac{u_1 y_{1j} + u_2 y_{2j} + \dots + u_s y_{sj}}{v_1 x_{1j} + v_2 x_{2j} + \dots + v_m x_{mj}} \leq 1 \quad (j = 1, \dots, n)$$

$$v_1, v_2, \dots, v_m \geq 0$$

$$u_1, u_2, \dots, u_s \geq 0$$

что эквивалентно задаче линейного программирования:

$$\max_{\mu,v} \theta = \mu_1 y_{1o} + \mu_2 y_{2o} + \dots + \mu_s y_{so}$$

$$v_1 x_{1o} + \dots + v_s x_{mo} = 1$$

$$\mu_1 y_{1j} + \dots + \mu_s y_{sj} \leq v_1 x_{1j} + \dots + v_s x_{mj} \quad (j = 1, \dots, n)$$

$$v_1, v_2, \dots, v_m \geq 0$$

$$\mu_1, \mu_2, \dots, \mu_s \geq 0$$

Модель ССR – определение эффективности

Определение ССR-эффективности

DMU_0 является **ССР-эффективным**, если $\theta^* = 1$ и существует по крайней мере один набор оптимальных весов (v^*, u^*) со всеми положительными компонентами

Можно показать, что для неэффективного DMU_0 с $\theta^* < 1$ существует по крайней мере одно DMU , для которого выполняется:

$$\sum_{r=1}^s u_r^* y_{rj} = \sum_{i=1}^m v_i^* x_{ij}$$

Подмножество E_o , состоящее из ССR-эффективных DMU , будем называть **референтным множеством** для DMU_0 .
Это **эффективная граница** для DMU_0

Оптимальные веса (v^*, u^*) отражают вклад входных и выходных параметров в формирование показателя эффективности DMU_0

Пример 1 - Модель CCR с единственным входным и единственным выходным параметром

DMU	A	B	C	D	E	F	G	H
Input	2	3	3	4	5	5	6	8
Output	1	3	2	3	4	2	3	5
CCR-эфф	0.5	1	0.667	0.75	0.8	0.4	0.5	0.625

Чтобы найти эффективность DMU, необходимо решить следующие задачи линейного программирования:

DMU₀ = A

$$\max_{u,v} \theta = u \quad \text{при усл. } 2v = 1$$

$$u \leq 2v \quad (A) \quad 3u \leq 3v \quad (B)$$

$$2u \leq 3v \quad (C) \quad 3u \leq 4v \quad (D)$$

$$4u \leq 5v \quad (E) \quad 2u \leq 5v \quad (F)$$

$$3u \leq 6v \quad (G) \quad 5u \leq 8v \quad (H)$$

$$v^* = u^* = \theta^* = 0.5$$

DMU₀ = B

$$\max_{u,v} \theta = 3u \quad \text{при усл. } 3v = 1$$

$$v^* = u^* = 1/3 \quad \theta^* = 1$$

Референт. множ-во: B

Пример 2. Модель ССР с двумя входными и единственным выходным параметром

DMU	A	B	C	D	E	F
Input 1	4	7	8	4	2	10
Input 2	3	3	1	2	4	1
Output 1	1	1	1	1	1	1

DMU₀ = A

$$\max_{u,v} \theta = u$$

при усл. $4v_1 + 3v_2 = 1$

$$u \leq 4v_1 + 3v_2 \text{ (A)} \quad u \leq 7v_1 + 3v_2 \text{ (B)}$$

$$u \leq 8v_1 + v_2 \text{ (C)} \quad u \leq 4v_1 + 2v_2 \text{ (D)}$$

$$u \leq 2v_1 + 4v_2 \text{ (E)} \quad u \leq 10v_1 + v_2 \text{ (F)}$$

$$v_1^* = v_2^* = 0.143$$

$$u^* = \theta^* = 0.857$$

Рефер. мн-во: D, E

DMU₀ = B

$$\max_{u,v} \theta = u$$

при усл. $7v_1 + v_2 = 1$

$$v_1^* = 0.053 \quad v_2^* = 0.211$$

$$u^* = \theta^* = 0.632$$

Рефер. мн-во: D, C

Пример 2. Модель ССР с двумя входными и единственным выходным параметром

DMU	A	B	C	D	E	F
Input 1	4	7	8	4	2	10
Input 2	3	3	1	2	4	1
Output 1	1	1	1	1	1	1

DMU₀ = F

$$\max_{u,v} \theta = u \text{ при усл. } 10v_1 + v_2 = 1$$

$$u \leq 4v_1 + 3v_2 (A) \quad u \leq 7v_1 + 3v_2 (B)$$

$$u \leq 8v_1 + v_2 (C) \quad u \leq 4v_1 + 2v_2 (D)$$

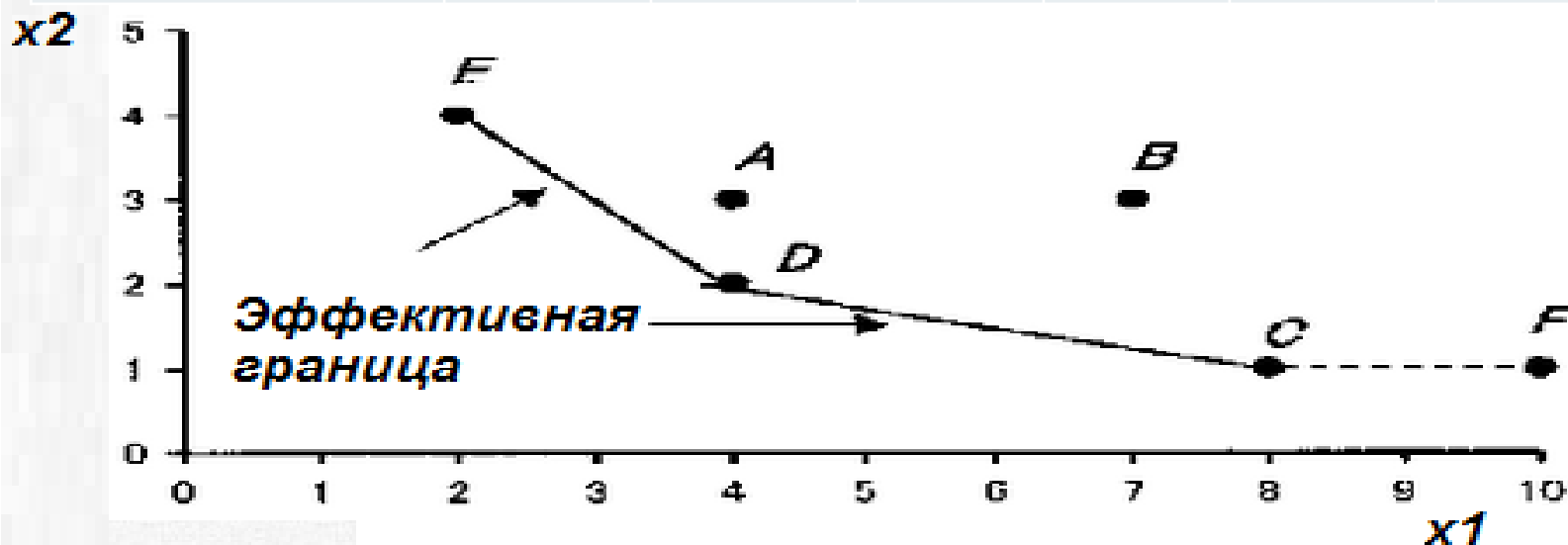
$$u \leq 2v_1 + 4v_2 (E) \quad u \leq 10v_1 + v_2 (F)$$

$$v_1^* = 0, \quad v_2^* = 1, \quad u^* = \theta^* = 1$$

F неэффективно, референтное множество: C

Пример 2. Модель ССР с двумя входными и единственным выходным параметром

DMU	A	B	C	D	E	F
Input 1	4	7	8	4	2	10
Input 2	3	3	1	2	4	1
Output 1	1	1	1	1	1	1
ССР-эфф	0.857	0.632	1	1	1	1
Рефер.множ-во	D, E	C, D	C	D	E	C
v1	0.143	0.053	0.083	0.167	0.214	0
v2	0.143	0.211	0.333	0.167	0.143	1
u	0.857	0.632	1	1	1	1



Модель ССР: двойственная задача

Предположение: входные и выходные параметры неотрицательны и хотя бы один компонент каждого вектора входных и выходных параметров положителен

$$x_j \geq 0, x_j \neq 0; y_j \geq 0, y_j \neq 0 \quad j = 1, \dots, n$$

Постановка задачи линейного программир. в векторно-матричной форме (n условий)

$$\max_{v, u} u y_0$$

$$\text{при усл. } v x_0 = 1;$$

$$-vX + uY \leq 0; \quad v \geq 0, u \geq 0$$

Двойственная (сопряженная) задача: (m+s) условий

$$\min_{\theta, \lambda} \theta$$

$$\text{при усл. } \theta x_0 - X\lambda \geq 0;$$

$$Y\lambda \geq y_0; \quad \lambda \geq 0$$

Нахождение «разрывов» (slacks)

$$\max_{\lambda, s^-, s^+} \left(\sum_{i=1}^m s_i^- + \sum_{r=1}^s s_r^+ \right)$$

$$\text{при усл. } s^- = \theta^* x_0 - X\lambda$$

$$s^+ = Y\lambda - y_0; \quad \lambda \geq 0, s^- \geq 0, s^+ \geq 0$$

θ^* – искомая эффективность

В ряде случаев при пропорциональном уменьшении входных параметров $\theta^* x_1; \dots; \theta^* x_m$ оптимум НЕ достигается → slacks

Модель ССР: новое определение эффективности

Другое (эквивалентное прежнему) определение ССР-эффективности:

Если оптимальное решение двойственной задачи $(\theta^*, \lambda^*, s^{-*}, s^{+*})$ удовлетворяет условиям $\theta^* = 1, s^{-*} = 0, s^{+*} = 0$ то DMU_o называется ССР-эффективным, в противном случае - неэффективным

Референтное множество: $E_o = \{j | \lambda_j^* > 0\} \quad (j = 1..n)$

Как улучшить входные и выходные параметры, чтобы достичь ССР-эффективности?

$$\begin{array}{l} \Delta x_o = (1 - \theta^*)x_o + s^{-*} \\ \Delta y_o = s^{+*} \end{array} \quad \rightarrow \quad \begin{array}{l} \bar{x}_o = x_o - \Delta x_o = \theta^* x_o - s^{-*} \leq x_o \\ \bar{y}_o = y_o + \Delta y_o = y_o + s^{+*} \geq y_o \end{array}$$

Эффективность достигается при умножении всех входных параметров пропорционально на величину θ^* и вычитании s^- и при прибавлении s^+ к выходным параметрам

Пример 3. Модель CCR - slacks

DMU	A	B	C	D	E	F	G
Input 1	4	7	8	4	2	10	3
Input 2	3	3	1	2	4	1	7
Output	1	1	1	1	1	1	1

Двойств. задача – A: Этан 1. $\min_{\theta, \lambda} \theta$ Этан 2. $\min_{\lambda, s_1^-, s_2^-, s^+} (-s_1^- - s_2^- - s^+)$

при усл $4\theta - 4\lambda_A - 7\lambda_B - 8\lambda_C - 4\lambda_D - 2\lambda_E - 10\lambda_E - 3\lambda_G - s_1^- = 0$

$3\theta - 3\lambda_A - 3\lambda_B - \lambda_C - 2\lambda_D - 4\lambda_E - \lambda_E - 7\lambda_G - s_2^- = 0$

$\lambda_A + \lambda_B + \lambda_C + \lambda_D + \lambda_E + \lambda_E + \lambda_G - s^+ = 1$

$\theta^* = 0.857, \lambda_D = 0.714, \lambda_E = 0.286, s_1^{-*} = s_2^{-*} = s^{+*} = 0$

$0.857 \cdot (\text{Input A}) = 0.714 \cdot (\text{Input D}) + 0.286 \cdot (\text{Input E})$

$(\text{Output A}) = 0.714 \cdot (\text{Output D}) + 0.286 \cdot (\text{Output E})$

$\bar{x}_1 = \theta^* x_1 = 0.857 * 4 = 3.429;$

$\bar{x}_2 = \theta^* x_2 = 0.857 * 3 = 2.571;$

$\bar{y} = y = 1$

λ_D и λ_E отражают вклад D, E в расчет точки для оценки эффективности A (D, E – референтное множество для A)

Пример 3. Модель CCR - slacks

Двойств. задача – F : Этан 1. $\min_{\theta, \lambda} \theta$ Этан 2. $\min_{\lambda, s_1^-, s_2^-, s^+} (-s_1^- - s_2^- - s^+)$

при усл $10\theta - 4\lambda_A - 7\lambda_B - 8\lambda_C - 4\lambda_D - 2\lambda_E - 10\lambda_F - 3\lambda_G - s_1^- = 0$

$$\theta - 3\lambda_A - 3\lambda_B - \lambda_C - 2\lambda_D - 4\lambda_E - \lambda_F - 7\lambda_G - s_2^- = 0$$

$$\lambda_A + \lambda_B + \lambda_C + \lambda_D + \lambda_E + \lambda_F + \lambda_G - s^+ = 1$$

$$\theta^* = 1, \lambda_A^* = \lambda_B^* = \lambda_D^* = \lambda_E^* = \lambda_F^* = \lambda_G^* = 0, \lambda_C^* = 1,$$
$$s_1^{-*} = 2, s_2^{-*} = s^{+*} = 0$$
$$v_1^* = 0, v_2^* = 1, u^* = 1$$

$$\bar{x}_1 = \theta^* x_1 - s_1^{-*} = 10 - 2 = 8$$

$$\bar{x}_2 = \theta^* x_2 = 1$$

$$\bar{y} = y = 1$$

$$(\text{Input 1 F}) = (\text{Input 1 C}) + 2$$

$$(\text{Input 2 F}) = (\text{Input 2 C})$$

$$(\text{Output F}) = (\text{Output C})$$

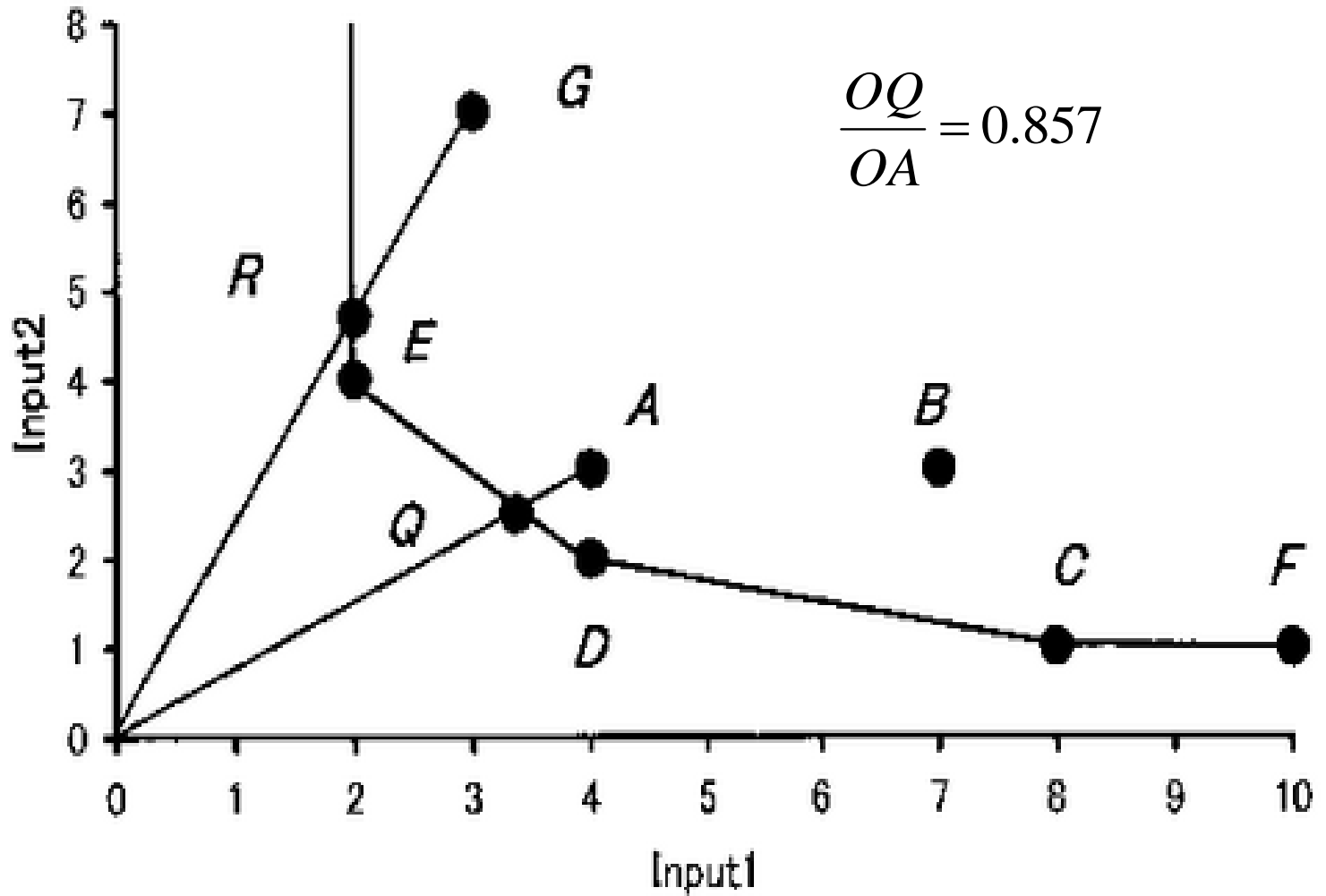
C - референтное множество для F

F неэффективно

Пример 3. Модель ССР – slacks

А, В, F, G и R (!) – неэффективны

Чтобы R было эффективно, надо уменьшить input 2



Модели, ориентированные на вход и на выход

Модель, ориентированная на вход:

цель – минимизировать входные параметры,
при этом выходные параметры должны либо остаться на первоначальном уровне, либо увеличиться

Модель, ориентированная на выход:

цель – максимизировать выходные параметры,
при этом входные параметры должны остаться либо на первоначальном уровне, либо уменьшиться

Модель, ориентированная на выход

Постановка задачи линейного программирования

$$\min_{p, q} px_o$$

при усл $qy_o = 1; -pX + qY \leq 0; p \geq 0; q \geq 0$

Постановка двойственной задачи

$$\max_{\eta, \mu} \eta$$

при усл $x_o - X\mu \geq 0$

$\eta y_o - Y\mu \leq 0; \mu \geq 0$

$$\lambda = \frac{\mu}{\eta}$$

$$\theta = \frac{1}{\eta}$$

$$\min_{\theta, \lambda} \theta$$

при усл $\theta x_o - X\lambda \geq 0;$

$Y\lambda \geq y_o; \lambda \geq 0$

Решение:

$$\eta^* = 1/\theta^*; \mu^* = \lambda^*/\theta^*$$

Slacks (t+, t-)

$$X\mu + t^- = x_o$$

$$Y\mu - t^+ = \eta y_o$$

$$t^- = s^{-*} / \theta^* \quad \theta^* \leq 1$$

$$t^+ = s^{+*} / \theta^* \quad \eta^* \geq 1$$

Веса p, q

$$p^* = v^* / \theta^*$$

$$q^* = u^* / \theta^*$$

Целевые значения x, y

$$\bar{x}_o = x_o - t^{-*}$$

$$\bar{y}_o = \eta^* y_o + t^{+*}$$

что эквивалентно

**Эффективные DMU
совпадают в**

моделях,

**ориентированных на
вход и на выход!**

ВСС – модели (Banker, Charnes, Cooper, 1984)

Граница эффективности в CRR- и ВСС-моделях

CCR: гипотеза о постоянной отдаче от масштаба

(если пропорционально изменить входные параметры, то выходные параметры изменятся в той же пропорции)

CCR-модели также называются **CRS – моделями** (Constant Return to Scale)

ВСС: не принимаем вышеуказанную гипотезу.

Эффективная граница представляет собой **выпуклую оболочку** существующих DMU

ВСС-модели также называются **VRS-моделями** (Variable Return to Scale)

Пример – ВСС-модель с единственным входным и единственным выходным параметром

Точка D: input-oriented models:

ВСС-эффективность

$$PR / PD = 2.67 / 4 = 0.67$$

CCR-эффективность

$$PQ / PD = 2.25 / 4 = 0.56$$

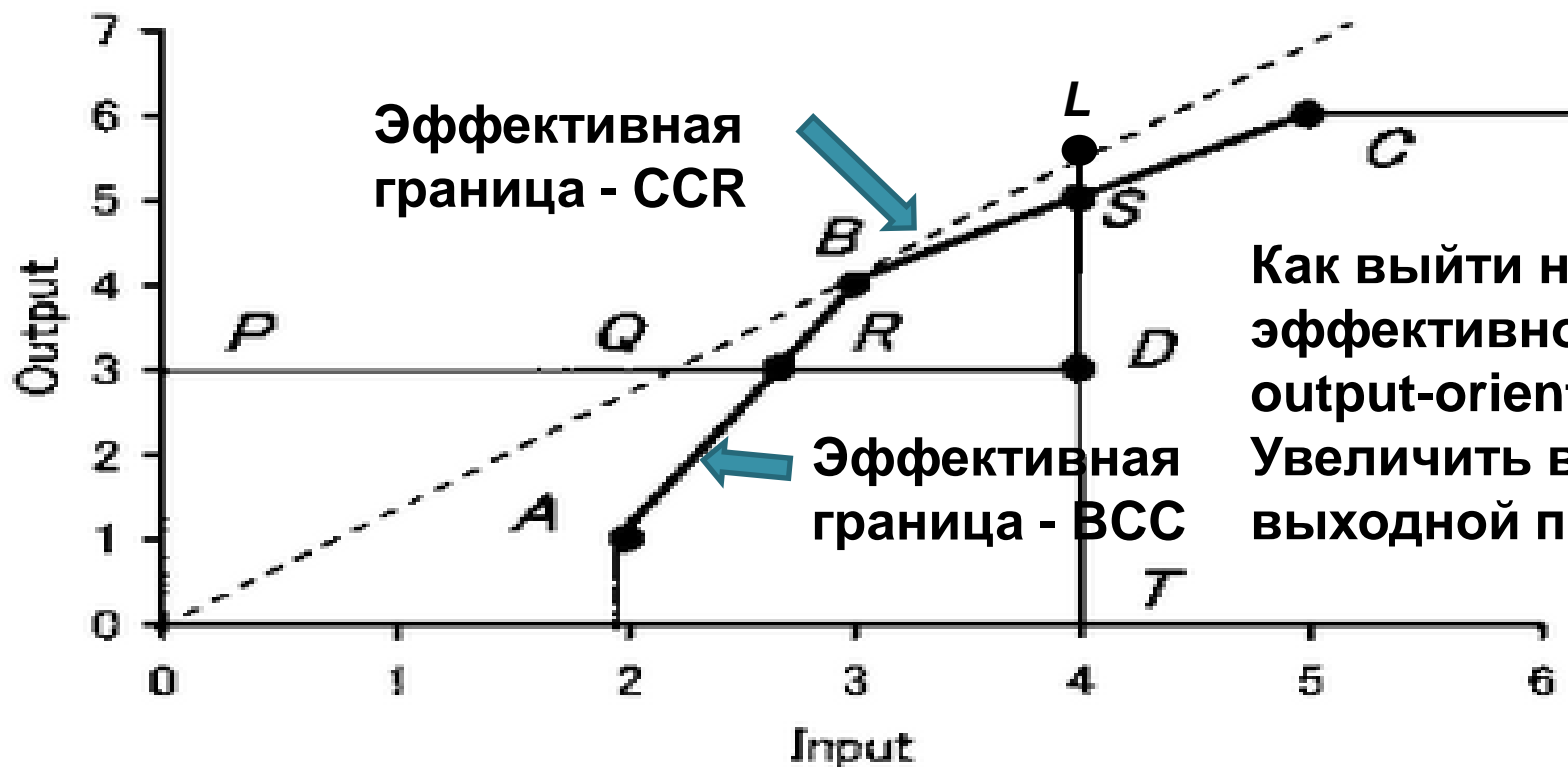
output-oriented:

ВСС-эффективность

$$ST / DT = 5 / 3 = 1.67 \neq 1 / 0.67(!)$$

CCR-эффективность

$$LT / DT = 1 / 0.56 = 1.78$$



Как выйти на границу эффективности в output-oriented ВСС? Увеличить в 1,67 раза выходной параметр

ВСС-модель – постановка задачи

Модель, ориентированная на вход

$$\min_{\theta_{BCC}, \lambda} \theta_{BCC}$$

$$\text{при усл. } \theta_{BCC} x_o - X\lambda \geq 0$$

$$Y\lambda \geq y_o$$

$$e\lambda = 1$$

$$\lambda \geq 0$$

e – единичный вектор

Определение ВСС-эффективности

$$\theta_{BCC} = 1 \text{ и } s^{-*} = s^{+*} = 0$$

$$\theta_{BCC} \geq \theta_{CRR}$$

Целевые значения параметров

$$\bar{x}_o = \theta_{BCC}^* x_o - s^{-*}$$

$$\bar{y}_o = y_o + s^{+*}$$

Отличие от CRR

$$\sum \lambda = 1$$

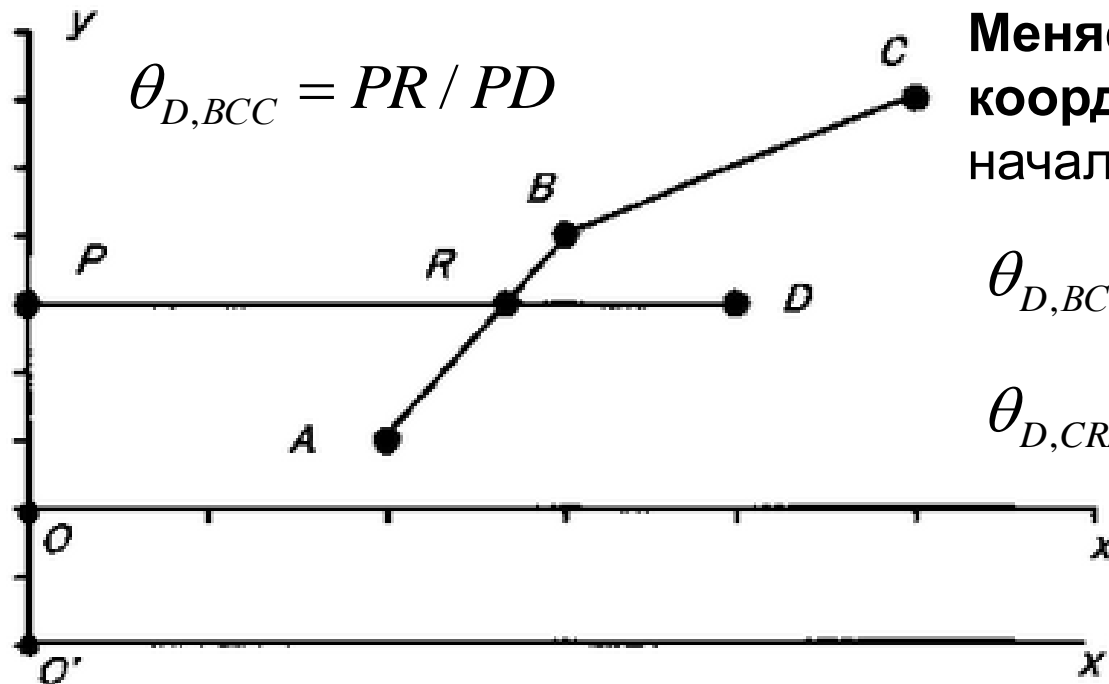
**это требование к
выпуклости
оболочки**

Инвариантность моделей CRR и ВСС к преобразованиям параметров

Будем называть модель DEA инвариантной к добавлению констант к исходным параметрам, если такое преобразование не изменяет оптимального решения

ВСС-модель, ориентированная на вход, инвариантна к линейным преобразованиям выходных параметров.

ВСС-модель, ориентированная на выход, инвариантна к линейным преобразованиям входных параметров



Меняем систему координат: переносим начало координат из O в O'

$\theta_{D,BCC}$ **не меняется**

$\theta_{D,CRR}$ **изменится!**

Допустимое количество входных и выходных параметров

Рекомендуется, чтобы количество DMU было больше, чем максимум ($m*s$, $3*(m+s)$)

т.к. в противном случае снижается различительная способность моделей DEA

В процессе исследования рекомендуется сначала использовать небольшое количество параметров, потом его постепенно расширять

Общая и чистая техническая эффективность, эффективность от масштаба

В чем источники неэффективности?

В неэффективной деятельности самого DMU или в неблагоприятных условиях его функционирования?

Модель ССР – предполагает **постоянную** отдачу от масштаба. ССР-эффективность называется **общей (глобальной) технической эффективностью** $\theta_{CCR}^* = 1 \Rightarrow \theta_{BCC}^* = 1$

Модель ВСС – предполагает **переменную** отдачу от масштаба. ВСС-эффективность называется **чистой технической эффективностью**, она отражает эффективность функционирования DMU (качество бизнес-процессов)

*ССР эффективность = ВСС-эффективность *
Эффективность масштаба (SE)

$$(SE = \theta_{CCR}^* / \theta_{BCC}^* \leq 1)$$

Эффективность масштаба отражает **неблагоприятные условия**, в которых функционирует DMU

Применение DEA: анализ эффективности банков

Слияния и поглощения банков в Японии в процессе реструктуризации

Банки : 1) «городские» – 9; 2) «региональные» - 100

Временной горизонт: 1995-1997 гг.

Источник: *Cooper W., Seiford L., Tone K. (2007). Data Envelopment Analysis. A Comprehensive Text with Models, Applications, References and DEA-Solver Software. Springer, 2nd Edition.*

Применение DEA: анализ эффективности банков

No.	Bank	Branch	3 Inputs		1 Output
			Employee	Assets**	Profits**
R1	Aomori	112	1,894	2,024,205	13,553
R2	Michinoku	111	1,727	1,702,017	16,540
R3	Iwate	123	1,843	2,048,014	14,760
R4	Tohoku	56	792	600,796	3,782
R5	Kita-Nippon	86	1,422	1,173,778	8,700
R6	Akita	111	1,841	2,290,274	18,590
R7	Hokuto	105	1,633	1,361,498	9,699
R8	Shonai	64	889	668,236	5,605
R9	Yamagata	90	1,611	1,563,930	9,993
R10	Yamagatashiawase	69	1,036	654,338	3,525
R11	Shokusan	73	1,011	656,794	3,177
C1	Dai-ichi Kangyo	407	17,837	53,438,938	411,149
C2	Sakura	529	18,805	52,535,262	302,901
C3	Fuji	363	15,188	51,368,976	371,364
C4	Tokyo Mitsubishi	419	20,235	78,170,694	504,422
C5	Asahi	436	13,149	29,531,193	170,835
C6	Sanwa	395	13,998	52,999,340	399,398
C7	Sumitomo	397	15,710	56,468,571	353,531
C8	Daiwa	218	8,671	16,665,573	112,121
C9	Tokai	305	11,546	31,376,180	186,640

Применение DEA: анализ эффективности банков

CCR: регион. банки
эффективны менее
городских

BCC: эффективность
региональных
банков возрастает

SE: все городские
банки – выше
среднего, половина
региональных -
меньше

Bank No.	CCR Score	Reference	BCC Score	Scale Score
R1	0.714	R2 C1	0.764	0.935
R2	1		1	1
R3	0.768	R2 C6	0.8	0.96
R4	0.648	R2	1	0.648
R5	0.763	R2	0.854	0.893
R6	0.89	R2 C1 C6	0.913	0.977
R7	0.733	R2	0.777	0.943
R8	0.863	R2	1	0.863
R9	0.676	R2 C1	0.777	0.87
R10	0.554	R2	0.918	0.603
R11	0.500	R2	0.915	0.544
C1	1		1	1
C2	0.742	R2 C1 C6	0.747	0.993
C3	0.981	C1 C4 C6	0.99	0.991
C4	1		1	1
C5	0.728	R2 C1	0.743	0.98
C6	1		1	1
C7	0.854	C1 C4 C6	0.866	0.986
C8	0.853	R2 C1	0.854	0.999
C9	0.766	R2 C1	0.766	1
Average	0.802		0.884	0.909

Применение DEA: анализ эффективности банков

Как изменится эффективность в результате M&A?

Пример: объединим банки R9 и R10 → R12

Берем ВСС-оптимальные (НЕ исходные) значения параметров

R12: Branches = 125 =

= 69 + 56

Employees = 1 887 =

= 1 095 + 792

Assets = 1 816 553 =

= 1 215 757 + 600 796

Net Profit = 13 518 =

= 9 993 + 3 525

	BCC		
	Branch	Employee	Assets
R.9	69	1095	1,215,757
	(-22%)	(-32%)	(-22%)
R.10	56	792	600,796
	(-19%)	(-24%)	(-8%)

Bank	Branch	Employee	Assets	CRR	BCC	Scale
R.9'	69	1,095	1,215,757	0.877	1	0.877
R.10'	56	792	600,796	0.604	1	0.604
R12	125	1,887	1,419,675	0.766	0.780	0.982

R12 НЕ стал эффективным с точки зрения CCR и BCC!

Применение DEA: анализ эффективности банков

А если объединить банки R9 и R10, предварительно выведя их на границу ССР-эффективности?

R12: Branches = 83 =
= 60 + 23

Employees = 1 346 =
= 978 + 368

Assets = 1 419 675 =
= 1 056 915 + 362 760

Net Profit = 13 518 =
= 9 993 + 3 525

	Branch	Employee	Assets
R9	60	978	1,056,915
	(-29%)	(-39%)	(-32%)
R10	23	368	362,760
	(-66%)	(-64%)	(-45%)

**Банк R12 будет глобально эффективным (ССРэфф = 1)
(как и объединяемые R9", R10")!**

Применение DEA: анализ эффективности банков

Источник: Karimzadeh M. Efficiency Analysis by Using Data Envelop Analysis Model: Evidence from Indian Banks. // Int. J Latest Trends Fin. Eco. Sc. Vol-2, №.3, 2012, p.228-237

Существующие подходы к оценке эффективности банков:

- 1. «Производственный»** (цель банка – выдать кредиты и привлечь депозиты). Входные параметры - собственный капитал и численность сотрудников, выходные – депозиты и кредиты
- 2. «Посреднический»** (цель банка – трансфер денежных средств от вкладчиков к заемщикам). Входные параметры – депозиты, собственный капитал, численность сотрудников, выходные – кредиты и вложения в ценные бумаги

Выборка: 8 крупнейших коммерческих индийских банков

Входные параметры: основные средства, депозиты, численность сотрудников

Выходные параметры: кредиты и вложения в ценные бумаги

Применение DEA: анализ эффективности банков

Модель CCR

Bank	2000-01	2001-02	2002-03	2003-04	2004-05	2005-06	2006-07	2007-08	2008-09	2009-10	2010-11	average
SBI	0.897	0.921	1	1	1	0.987	0.976	0.982	1	1	1	0.978
BOI	0.934	0.974	1	1	1	1	1	1	1	1	1	0.992
CBI	0.921	0.900	0.974	1	1	1	1	0.877	0.925	0.953	1	0.959
UBI	0.883	0.896	0.977	1	0.964	1	1	0.993	0.971	1	1	0.971
PNB	0.912	0.906	0.983	0.976	1	0.920	1	0.967	0.965	0.984	1	0.965
ICICI	0.932	0.957	1	1	1	1	1	1	1	0.983	1	0.988
HDFC	0.953	1	0.937	1	1	0.979	0.980	1	0.980	0.986	1	0.942
AXIS	0.908	1	0.887	0.750	1	1	0.995	0.999	1	0.963	1	0.955
Average	0.918	0.944	0.969	0.965	0.995	0.986	0.994	0.977	0.980	0.983	1	0.969

1. Государственные банки эффективнее частных
2. Средняя эффективность по выборке росла с 2000 по 2007, затем после спада в 2008 г. продолжился рост, в 2011 г. все банки были эффективными

Применение DEA: анализ эффективности банков

Модель ВСС

Bank	2000-01	2001-02	2002-03	2003-04	2004-05	2005-06	2006-07	2007-08	2008-09	2009-10	2010-11	Average
SBI	1	1	1	1	1	1	1	1	1	1	1	1
BOI	1	1	1	1	1	1	1	1	1	1	1	1
CBI	1	1	1	1	1	1	1	0.962	0.989	0.985	1	0.994
UBI	1	0.998	1	1	0.999	1	1	1	0.972	1	1	0.997
PNB	1	1	1	0.994	1	0.978	1	1	1	0.993	1	0.997
ICICI	1	1	1	1	1	1	1	1	1	1	1	1
HDFC	1	1	0.965	1	1	1	0.981	1	1	0.989	1	0.994
AXIS	1	1	1	0.839	1	1	1	1	1	0.963	1	0.982
Average	1	0.999	0.979	0.979	0.999	0.997	0.997	0.995	0.995	0.991	1	0.995

Применение DEA: оценка нефтяных компаний

Источник: Выгон Г.В. (2001).

Методы оценки нефтяных компаний в условиях неопределенности. Аудит и финансовый анализ , 1.

Цель работы:

Оценка сравнительной **инвестиционной привлекательности** нефтяных компаний на основе **интегрированного показателя**

Практическое применение:

Выявление недооцененных акций (с целью переиграть рынок)

Предпосылки анализа: средняя форма эффективности российского рынка акций

Применение DEA: оценка нефтяных компаний

Выходные параметры: рыночная капитализация

Входные параметры: чистая прибыль, выручка, добыча (баррели), запасы (баррели), объем экспорта (баррели)

Модель: ВСС, ориентированная на вход

Пример. Единственный входной параметр = чистая прибыль →

→ коэффициент эффективности = P/E

(рыночный мультипликатор)

Неэффективная компания = Недооцененная компания

Ограничение: метод НЕ позволяет выявить переоцененные компании

Применение DEA: оценка нефтяных компаний

Результаты расчетов:

1. Анализ эффективности российских компаний (5 факторов, данные за 1997 г.)

№	Наименование	CCR	BCC
1	НК Татнефть	0.4305	0.4627
2	ТНК	0.5270	0.5629
3	ВНК	0.8274	0.8583
4	НК Сибнефть	0.8654	0.9034
5	НК Сиданко	0.9624	0.9624
6	НК Сургутнефтегаз	0.9414	1
7	ОНАКО	1	1
8	НК КомиТЭК	1	1
9	НК Лукойл	1	1
10	НК ЮКОС	1	1
11	НК СЛАВНЕФТЬ	1	1

Применение DEA: оценка нефтяных компаний

Следующая задача исследования:

**можно ли переиграть рынок,
играя на недооцененных акциях?**

Выборка: 12 нефтяных компаний американского рынка

Выходной параметр: рыночная капитализация

Входные параметры: запасы, добыча нефти и газа, переработка нефти, выручка, чистая прибыль, активы, капитал

Предлагаемая стратегия:

- **покупать** акции компании, у которой сильнее всего **упал коэффициент эффективности**
- **продавать** акции компании, у которой сильнее всего **вырос коэффициент эффективности**

Применение DEA: оценка инвестиционных рисков при формировании портфеля

Источник: *T. Arshinova. (2011). Construction of Equity Portfolio on the Basis of Data Envelopment Analysis Approach. Scientific Journal of Riga Technical University, Vol. 49*

Подход: оценка инвестиционного риска на основе сравнительного анализа эффективности компаний

Выборка: 20 компаний, торгуемых на бирже NASDAQ OMX Riga and NASDAQ OMX Tallinn

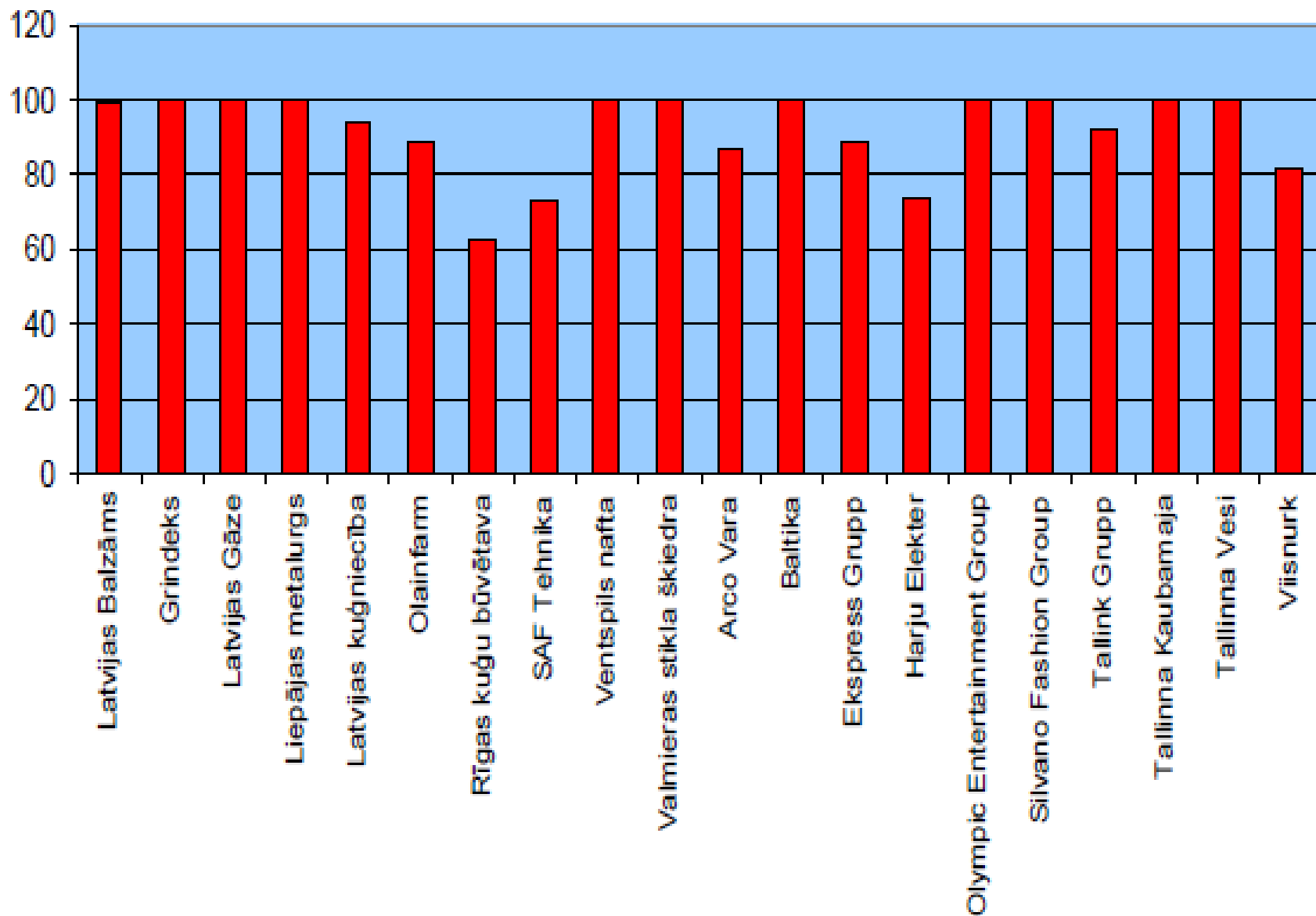
Выходные параметры: рыночная капитализация, выручка

Входные параметры: собственный капитал, операционные расходы, расходы по уплате процентов

Модель: CCR, ориентированная на вход

Результаты: компании разделены на три группы: с эффективностью = 100% (консервативная стратегия), от 80% до 100%, и меньше 80% (высокорискованная стратегия)

Применение DEA: оценка инвестиционных рисков при формировании портфеля



Подготовлено по результатам выполнения НИР
**«Исследование международных подходов к участию в ГЦС
и разработка рекомендаций по анализу повышения
эффективности участия России в ГЦС»**

Заказчик – Минэкономразвития, 2014,
соисполнителями являлись сотрудники лаборатории анализа
финансовых рынков (ЛАФР)

Применение DEA: глобальные цепочки стоимости

Потенциальные выгоды участия в ГЦС (результаты НИР):

- Рост притока ПИИ
- Ускорение экономического роста
- Рост капитализации фондового рынка
- Снижение безработицы
- Повышение качества продукции и снижение цен
- Технологические переливы

Наиболее существенные **положительные эффекты от участия в ГЦС** (максимальный темп роста ВВП на душу населения) получили страны, для которых были **одновременно характерны высокие темпы роста значения индекса участия в ГЦС и значения внутренней добавленной стоимости в экспорте**

Источник: UNCTAD World Investment Report 2013

Применение DEA: глобальные цепочки стоимости

Выборка:

24 страны (развитые и развивающиеся рынки), 2000-2009 гг.

Входные параметры (источник – база TIVA ОЭСР):

- Индекс участия в ГЦС

*= Доля иностранных промежуточных товаров и услуг в экспорте
+ Доля произведенных в стране промежуточных товаров и услуг,
используемых в экспорте других стран*

(в процентах от экспорта страны)

- Доля внутренней добавленной стоимости в экспорте

Выходные параметры (источник – worldbank, отчеты ВЭФ):

- Темп роста ВВП
- Индекс глобальной конкурентоспособности ВЭФ
- Приток прямых иностранных инвестиций (в % от ВВП)

Модель: ВСС, ориентированная на вход

Применение DEA: глобальные цепочки стоимости

	2006 год				
	Индекс участия в ГЦС	Внутренняя добавленная стоимость в валовом экспорте, %	Индекс конкурентоспособности	Темп роста ВВП, %	Приток ПИИ в страну, % ВВП
Россия	57,77	92,08	4,08	8,15	3,80
Бразилия	40,10	87,50	4,03	3,96	1,73
Китай	48,30	64,65	4,24	12,70	2,68
Индия	43,88	79,09	4,44	9,26	2,14
ЮАР	34,05	81,91	4,36	5,60	0,01
США	43,72	87,72	5,61	2,67	1,71
Германия	51,10	73,66	5,58	3,70	1,92
Япония	45,61	84,38	5,60	1,69	0,00
Медиана (24 страны)	48,52	76,76	4,54	5,34	2,94

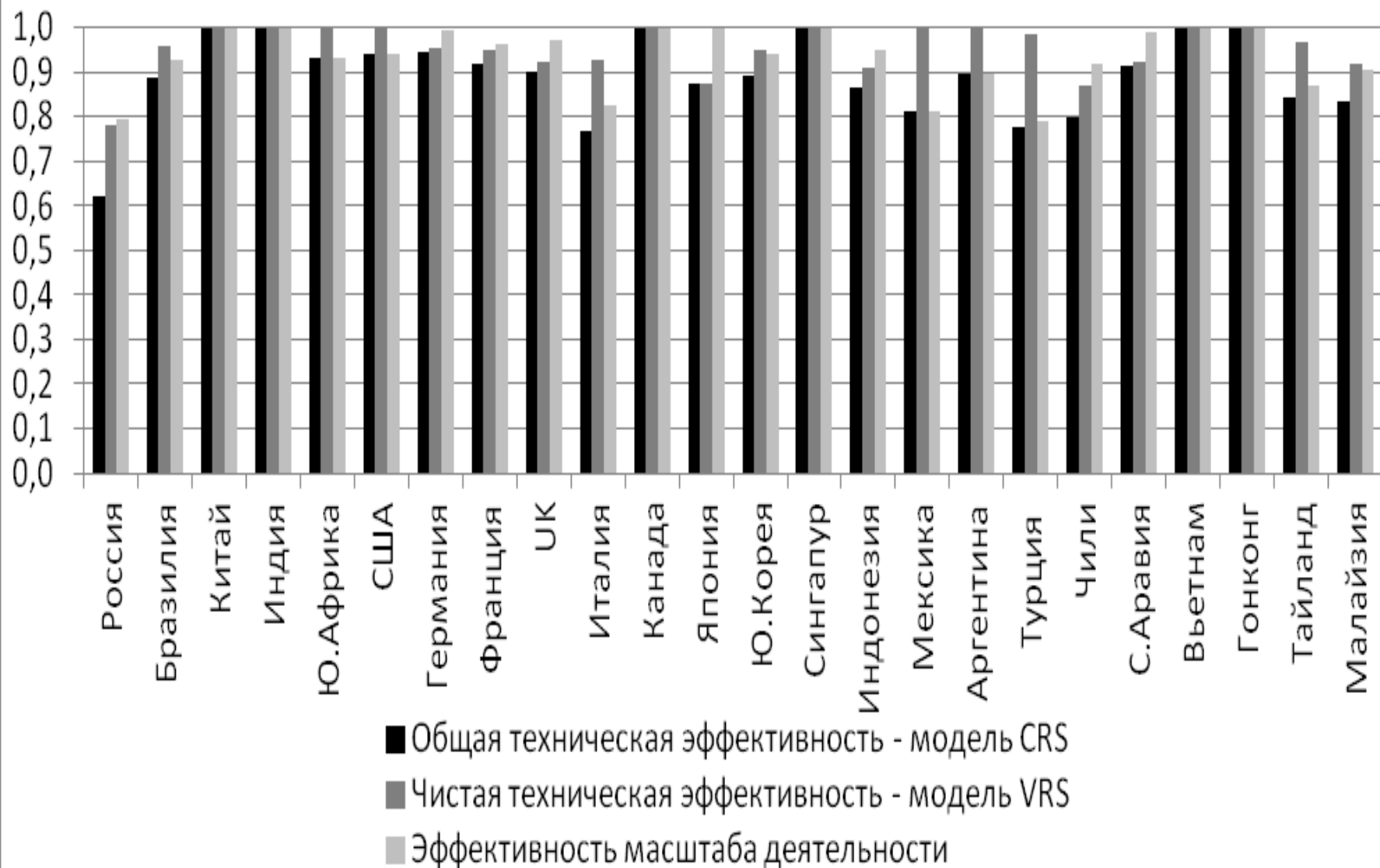
Применение DEA: глобальные цепочки стоимости

Страны-лидеры	Страны-аутсайдеры
по чистой технической эффективности на всем отрезке 2000-2009 гг.	
Китай	Россия
США	Индонезия
Канада	Чили
Сингапур	Саудовская Аравия
Мексика	

У России **наименьшие** показатели чистой технической эффективности (75-78%) за 2002-2009 гг.

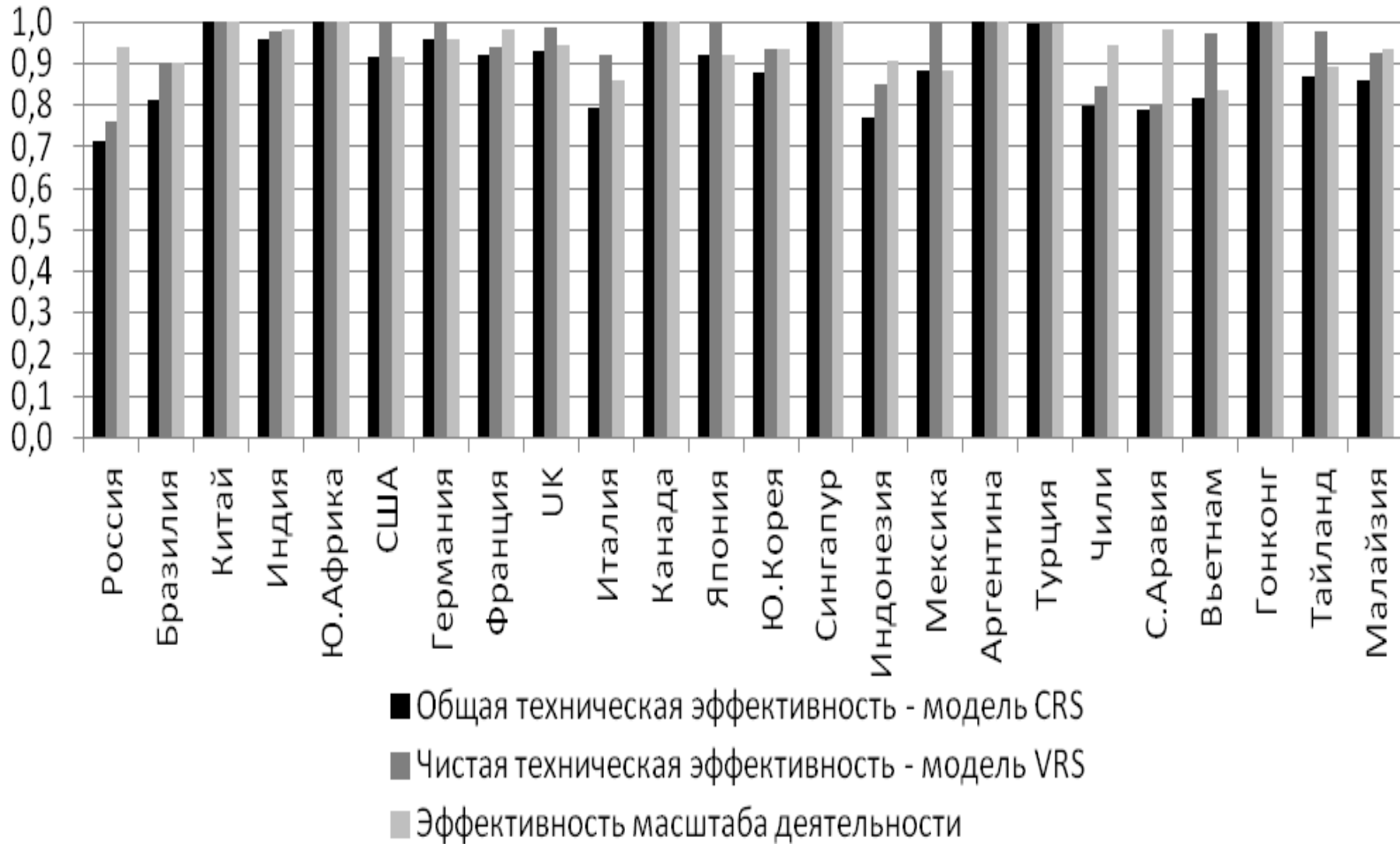
Применение DEA: глобальные цепочки стоимости

Эффективность участия различных стран в ГЦС, 2009



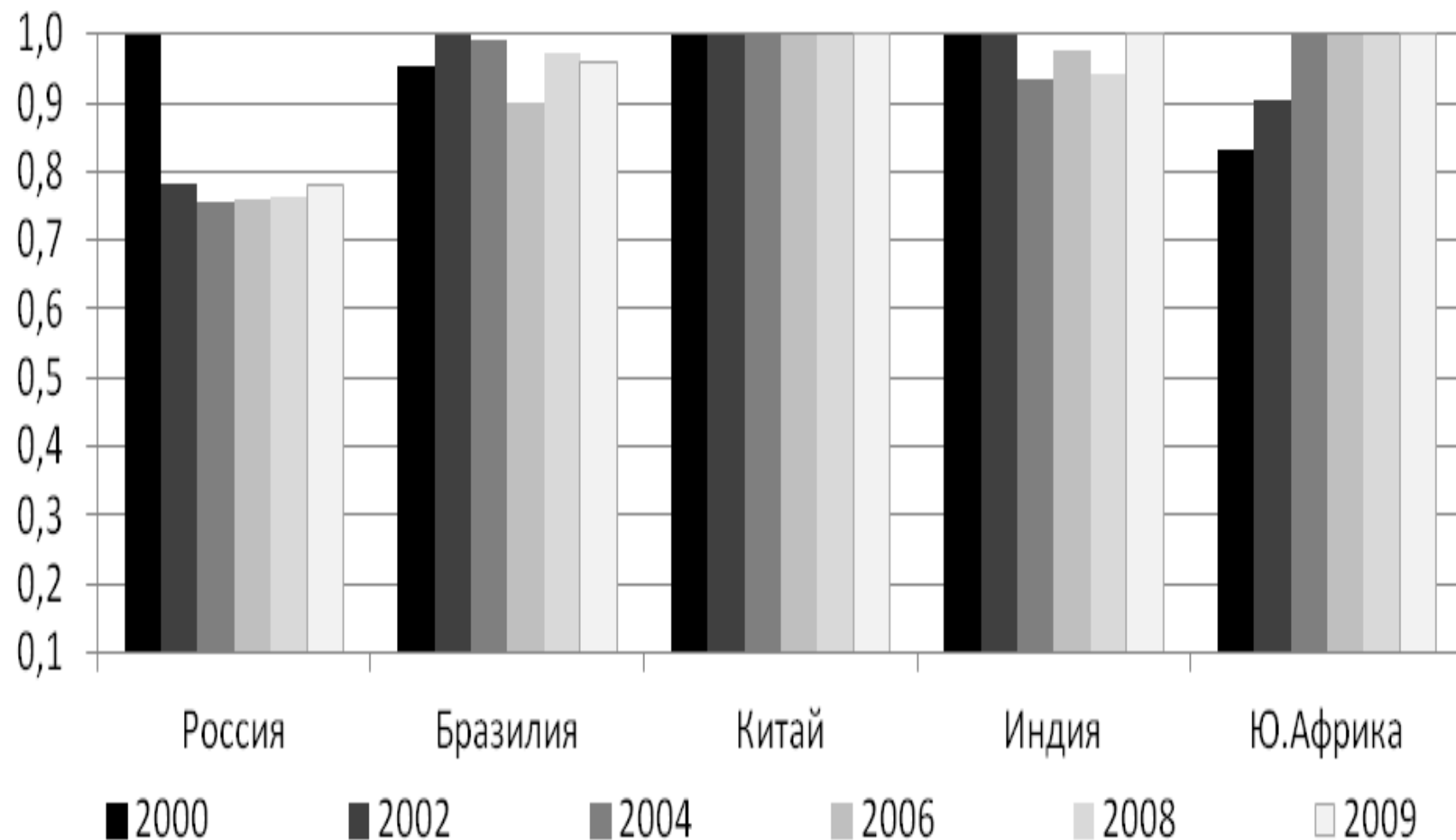
Применение DEA: глобальные цепочки стоимости

Эффективность участия различных стран в ГЦС, 2006



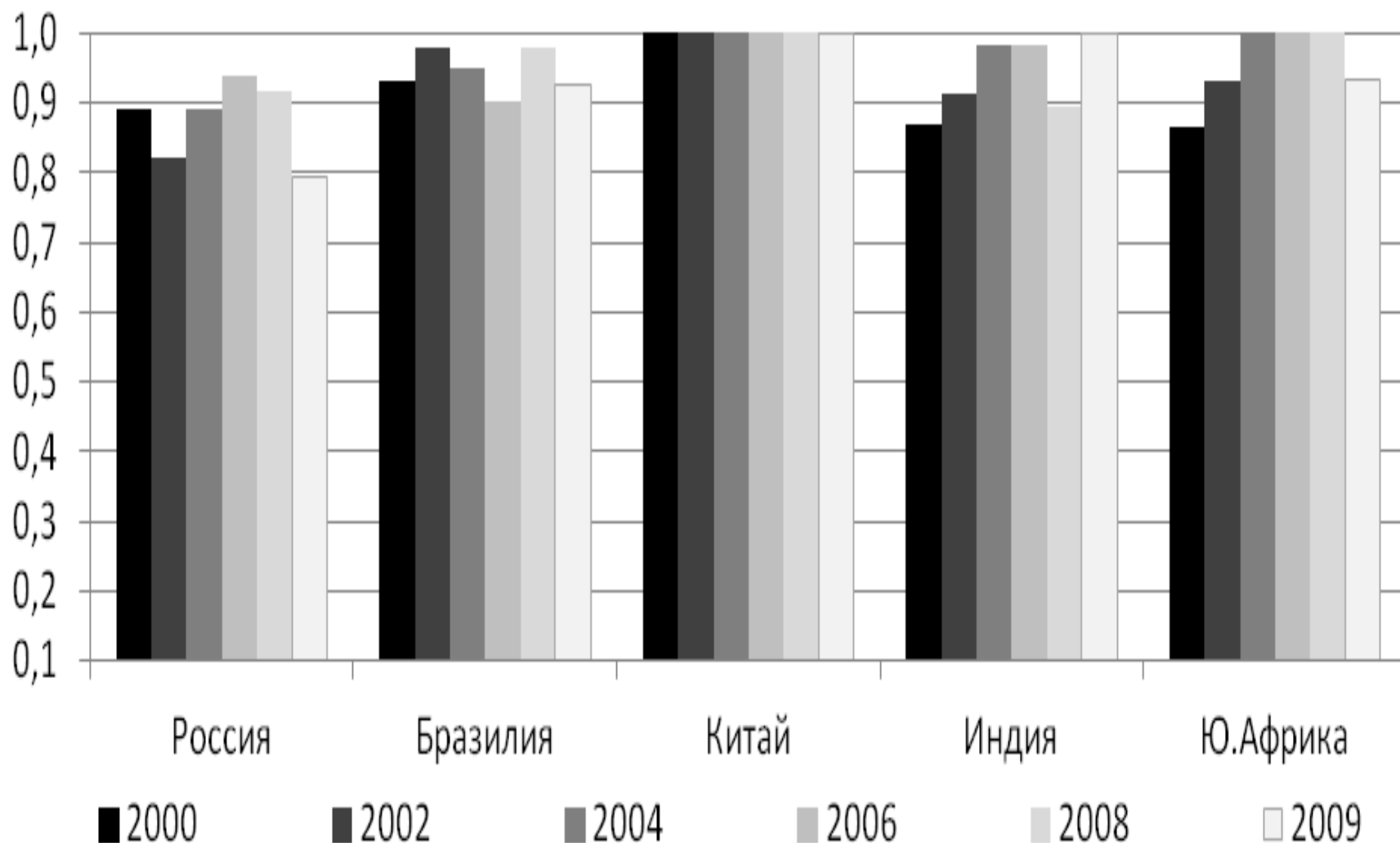
Применение DEA: глобальные цепочки стоимости

Динамика чистой технической эффективности участия стран БРИКС в ГЦС, 2000-2009



Применение DEA: глобальные цепочки стоимости

Динамика эффективности масштаба участия стран БРИКС в ГЦС, 2000-2009



Особенности применения метода DEA: исключение зависимых (коррелирующих) параметров

DMU	Input		Output			Efficiency Score	
	1	2	3	1	2	3 Input, 2 Output	2 Input, 2 Output
1	10	12	7	34	7	1.000	1.000
2	24	26	5	6	1	0.169	0.169
3	23	25	3	24	10	1.000	1.000
4	12	14	4	2	5	0.581	0.576
5	11	13	5	29	8	0.954	0.948
6	12	14	2	7	2	0.467	0.460
7	44	46	5	39	4	0.975	0.975
8	12	14	7	6	4	0.400	0.400
9	33	35	5	2	10	0.664	0.664
10	22	24	4	7	6	0.541	0.541
11	35	37	7	1	4	0.220	0.220
12	21	23	5	0	6	0.493	0.493
13	22	24	6	10	6	0.437	0.437
14	24	26	8	20	9	0.534	0.534
15	12	14	9	5	4	0.400	0.400
16	33	35	2	7	6	0.900	0.900
17	22	24	9	2	3	0.175	0.175
18	12	14	5	32	10	1.000	1.000
19	42	44	7	9	7	0.352	0.352
20	12	14	4	5	3	0.349	0.346

Особенности применения метода DEA: «нормализация» средних значений параметров

$$\bar{V}_i = \frac{\sum_{n=1}^N V_{ni}}{N}$$

DMU	Input 1	Input 2	Output 1	Output 2
1	1733896	97	1147	0.82
2	2433965	68	2325	0.45
3	30546	50	1998	0.23
4	1052151	42	542	0.34
5	4233031	15	1590	0.67
6	3652401	50	1203	0.39
7	1288406	65	1786	1.18
8	4489741	43	1639	1.28
9	4800884	90	2487	0.77
10	536165	19	340	0.57
Column Mean	2425119	53.9	1505.7	0.67



**Результат
не
меняется!**

DMU	Input 1	Input 2	Output 1	Output 2	CCR
1	0.7150	1.7996	0.7618	1.2239	0.432
2	1.0036	1.2616	1.5441	0.6716	0.709
3	0.0126	0.9276	1.3270	0.3433	1.000
4	0.4339	0.7792	0.3600	0.5075	0.419
5	1.7455	0.2783	1.0560	1.0000	1.000
6	1.5061	0.9276	0.7990	0.5821	0.448
7	0.5313	1.2059	1.1862	1.7612	0.952
8	1.8513	0.7978	1.0885	1.9104	0.865
9	1.9796	1.6698	1.6517	1.1493	0.570
10	0.2211	0.3525	0.2258	0.8507	1.000
Column Mean	1	1	1	1	

Особенности применения метода DEA: что делать с отрицательными значениями параметров?

Базовые модели **НЕ** позволяют работать с **отрицательными** значениями параметров!

1. **Исключение строк с отрицательными значениями**

Граница эффективности может измениться

2. **Замена отрицательных значений небольшой положительной константой (меньшей, чем все другие значения параметра)**

Допустимо при небольшом количестве замен, но результат не будет в точности таким же

3. **Прибавление достаточно большой положительной константы ко всем значениям параметра**

НЕ все модели инвариантны к таким преобразованиям.

Особенности применения метода DEA: что делать с отрицательными значениями параметров?

Пример. Страховая компания инвестирует деньги в IT-системы. Необходимо выбрать оптимальную IT-систему

DMU	INPUTS		OUTPUTS	
	Labor Cost Savings	Operational Cost Savings	Processing Mistake Decrease	Decrease in Turnaround Time (minutes per policy)
IT1	240	128	2	0
IT2	-325	1 006	0	59
IT3	-1 400	1 017	27	26
IT4	-363	500	0	75
IT5	-55	256	77	66
IT6	4 450	-246	12	-14
IT7	6 520	79	-3	-14
IT8	3 900	341	20	23

Необходима модификация: входные параметры с меньшими значениями должны быть предпочтительны

Особенности применения метода DEA: что делать с отрицательными значениями параметров?

Замена входных параметров:

Cost Savings → Cost

Например, вычтем из произвольной константы, равной 10000, значения входных параметров

Также прибавим + 10 к параметру «Уменьшение ошибок» и +20 к параметру «Уменьшение времени на выдачу полиса»

DMU	INPUTS		OUTPUTS		CCR- эффектив- ность
	Labor Costs	Operational Costs	Processing Mistake Decrease	Decrease in Turnaround Time	
IT1	9 760	9 872	12	20	0.232
IT2	10 325	8 994	10	79	0.881
IT3	11 400	8 983	37	46	0.564
IT4	10 363	9 500	10	95	1.000
IT5	10 055	9 744	87	86	1.000
IT6	5 550	10 246	22	6	0.458
IT7	3 480	9 921	7	6	0.232
IT8	6 100	9 659	30	43	0.805

Особенности применения метода DEA: что делать с отрицательными значениями параметров?

Вычтем из произвольной константы 7000 значение расходов на оплату труда и из 1500 - операционные расходы

Также прибавим + 4 к параметру «Уменьшение ошибок» и +15 к параметру «Уменьшение времени на выдачу полиса»

РЕЗУЛЬТАТ ИЗМЕНИЛСЯ!

DMU	INPUTS		OUTPUTS		CCR- эффектив- ность
	Labor Costs	Operational Costs	Processing Mistake Decrease	Decrease in Turnaround Time	
IT1	6 760	1 372	6	15	0.184
IT2	7 325	494	4	74	1.000
IT3	8 400	483	31	41	1.000
IT4	7 363	1 000	4	90	1.000
IT5	7 055	1 244	81	81	1.000
IT6	2 550	1 746	16	1	0.546
IT7	480	1 421	1	1	0.181
IT8	3 100	1 159	24	38	1.000

Особенности применения метода DEA: что делать с отрицательными значениями параметров?

4. **Использовать другие типы DEA-моделей, инвариантные к преобразованиям входных и выходных параметров**

См. подробнее:

Cooper W., Seiford L., Tone K. (2007). Data Envelopment Analysis. A Comprehensive Text with Models, Applications, References and DEA-Solver Software. Springer, 2nd Edition.

Особенности применения метода DEA: что делать, если некоторые значения параметров отсутствуют?

1. **Исключить DMU с отсутствующими данными из рассмотрения.**

При этом может измениться граница эффективности!

2. **Вместо отсутствующих данных дать экспертную оценку.**
Такие оценки **субъективны**. Может потребоваться **сценарный подход** :

$$V_e = \frac{V_o + 4V_m + V_p}{6}$$

V_e – итоговая оценка,
 V_o – оптимистический сценарий,
 V_m – наиболее вероятная оценка,
 V_p – пессимистический сценарий

Также можно взять минимальное, максимальное, среднее значение параметра по всему ряду DMU, трендовое изменение и т.д.

Программное обеспечение DEAP

<http://www.uq.edu.au/economics/cepa/deap.php>

Преимущества:

1. Бесплатность
2. Простота применения

Для работы необходимо подготовить два текстовых файла

1. test.DTA – файл с данными
2. test.INS – файл с инструкциями (настройками) расчета

По итогам работы будет сформирован файл:

test.OUT – файл с результатами расчетов

Исполняемый файл:

DEAP.exe

Программное обеспечение DEAR

Файл с данными .DTA:

1. В каждой строке указываются данные по одному DMU
2. Каждому входному и выходному параметру соответствует один столбец (сначала – выходные параметры!)

Y1 Yn X1.... Xm

Подготовить файл с данными можно в текстовых редакторах Word, Excel, Notepad и т.д.

DMU	Y	X1	X2
1	1	2	5
2	2	2	4
3	3	6	6
4	1	3	2
5	2	6	2



1 2 5
2 2 4
3 6 6
1 3 2
2 6 2

**Десятичный
разделитель – точка,
т.е. корректно,
например, задать
значение параметра
3.62**

Программное обеспечение DEAP

Файл с инструкциями .INS:

eg1.dta	DATA FILE NAME
eg1.out	OUTPUT FILE NAME
5	NUMBER OF FIRMS
1	NUMBER OF TIME PERIODS
1	NUMBER OF OUTPUTS
2	NUMBER OF INPUTS
0	0=INPUT AND 1=OUTPUT ORIENTATED
0	0=CRS AND 1=VRS
0	0=DEA (MULTI-STAGE), 1=COST-DEA, 2=MALMQUIST-DEA, 3=DEA(1-STAGE), 4=DEA(2-STAGE)

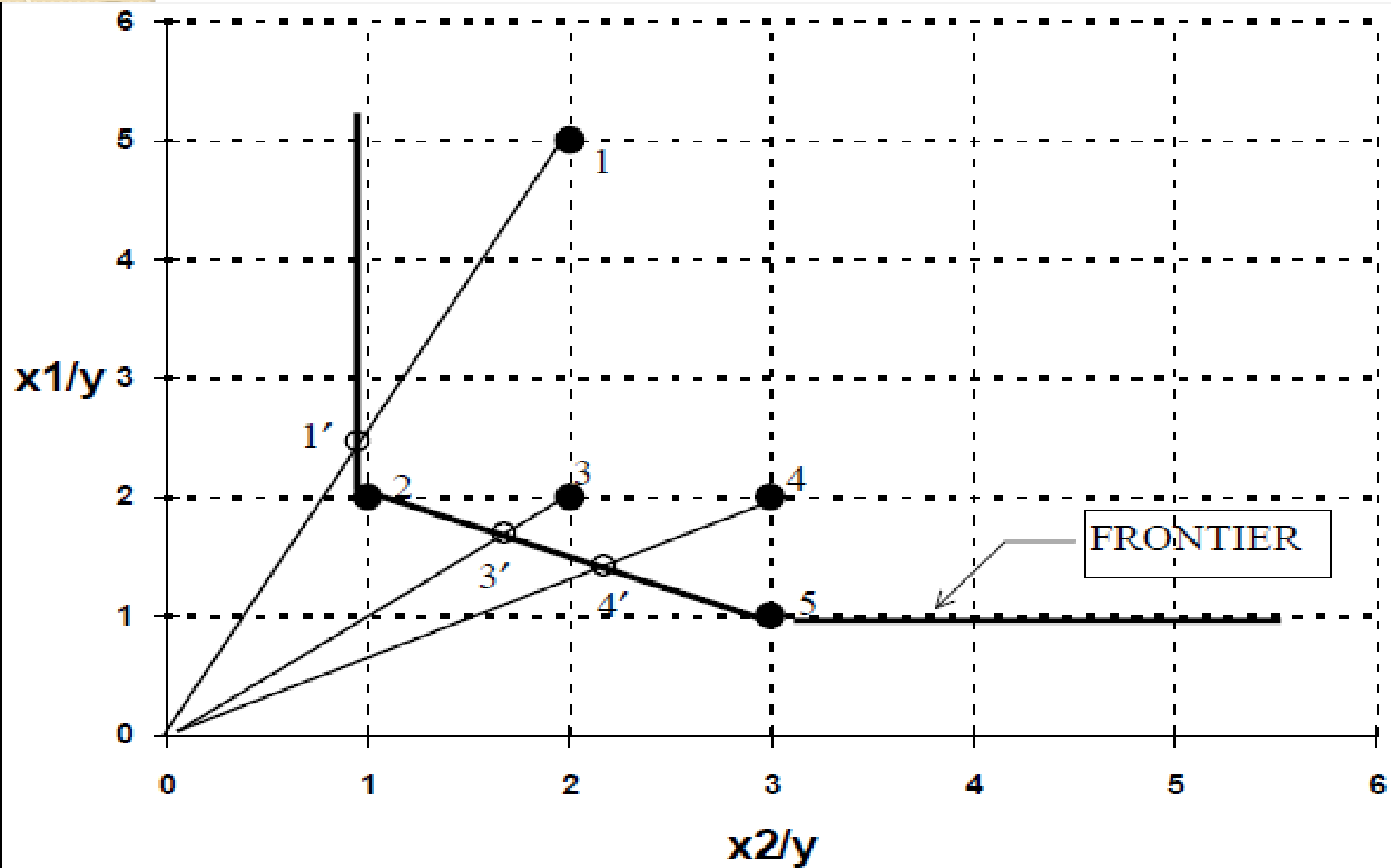
```
DEAP Version 2.1
*****
A Data Envelopment Analysis (DEA) Program
by Tim Coelli
  Centre for Efficiency and Productivity Analysis
University of Queensland
Brisbane, QLD 4072
Australia.
Email: t.coelli@economics.uq.edu.au
Web: http://www.uq.edu.au/economics/cepa
Enter instruction file name: test.ins
```

Заполняем левый столбец необходимыми значениями (например, в Notepad)

1. Запустить DEAP.EXE
2. В командной строке указать имя файла с инструкциями
например, test.ins

Программное обеспечение DEAP

Модель: CCR, ориентированная на вход



Программное обеспечение DEAP

Файл с результатами расчетов .OUT:

EFFICIENCY SUMMARY: SUMMARY OF OUTPUT SLACKS:

firm	te	firm	output:	1
1	0.500	1		0.000
2	1.000	2		0.000
3	0.833	3		0.000
4	0.714	4		0.000
5	1.000	5		0.000
mean	0.810	mean		0.000

SUMMARY OF INPUT SLACKS:

firm	input:	1	2
1		0.000	0.500
2		0.000	0.000
3		0.000	0.000
4		0.000	0.000
5		0.000	0.000
mean		0.000	0.100

Программное обеспечение DEAR

Файл с результатами расчетов .OUT:

SUMMARY OF PEERS:

firm	peers:
1	2
2	2
3	5 2
4	5 2
5	5

SUMMARY OF OUTPUT TARGETS:

firm	output:	1
1		1.000
2		2.000
3		3.000
4		1.000
5		2.000

SUMMARY OF INPUT TARGETS:

firm	input:	1	2
1		1.000	2.000
2		2.000	4.000
3		5.000	5.000
4		2.143	1.429
5		6.000	2.000

Программное обеспечение DEAP

Файл с результатами расчетов .OUT:

FIRM BY FIRM RESULTS:

Results for firm: 1
Technical efficiency = 0.500

PROJECTION SUMMARY:

variable		original		radial		slack		projected
		value		movement		movement		value
output	1	1.000		0.000		0.000		1.000
input	1	2.000		-1.000		0.000		1.000
input	2	5.000		-2.500		-0.500		2.000

LISTING OF PEERS:

peer	lambda	weight
2	0.500	