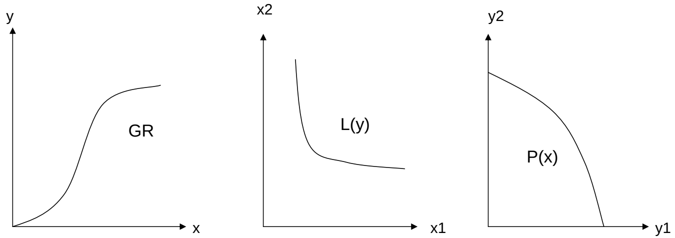
Technical Efficiency and Data Envelopment Analysis (DEA)

based on OnFront, Reference Guide
Coelli, et al., *Introduction to Efficiency and Productivity*
Springer, 2005, Ch 6.

April 16, 2008

Equivalent Technology Sets: describe the relationship between all feasible inputs $x = (x_1, \ldots, x_N)$ and outputs $y = (y_1, \ldots, y_M)$



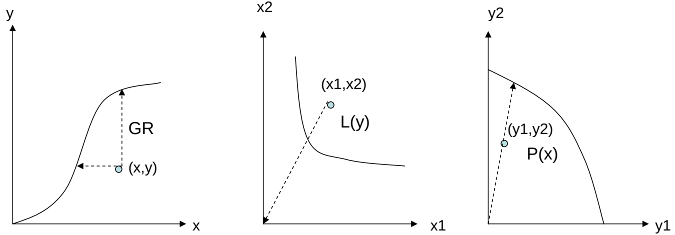
GR={(x,y):x can produce y}

L(y)={x: (x,y) is in GR}

P(x)={y: (x,y) is in GR}

If (x,y) is in GR, then x is in L(y) and y is in P(x)

Technical Efficiency



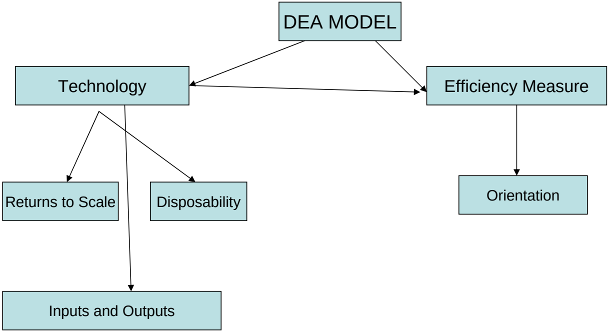
Fi(y,x)=min{lambda: lambda x is in L(y)}

Fo(x,y)=max{theta:theta y is in P(x)}

Estimating technical efficiency with DEA

- DEA uses linear programming to construct the technology and best practice frontier from the data in your sample
- Simultaneously, it estimates the distance to the best practice frontier for each observation
- There is a separate linear programming problem for each observation (DMU)

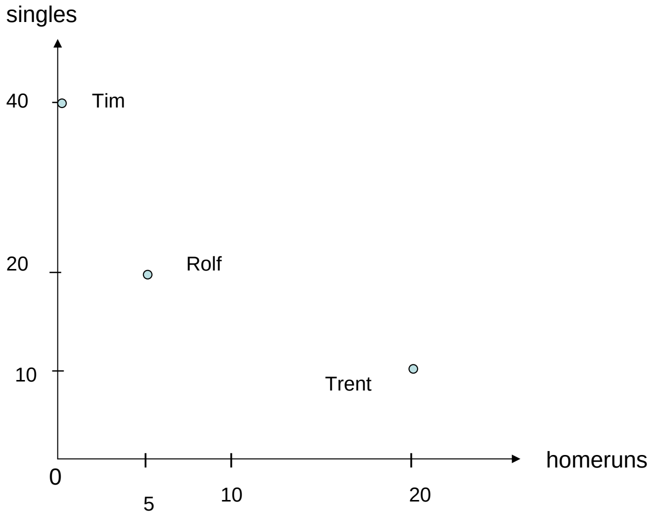
DEA Overview



Practice Exercise: DEA and baseball

	At Bats	Singles	Homeruns
	input	output 1	output 2
Tim	100	40	0
Rolf	100	20	5
Trent	100	10	20

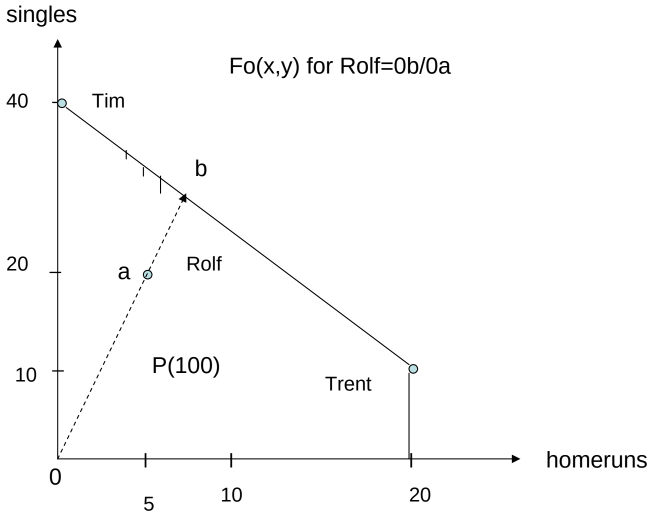
Plot the data with homeruns on the horizontal axis and singles on the vertical axis (this will eventually be an output set).


singles
40
Tim
20
Rolf
10
Trent
0
5
10
20
homeruns

Now think about what the output set would look like. You will be constructing the output set $P(x)$, where x is equal to 100 at bats.

- Allow for free disposability of outputs, i.e., outputs less than or equal to observed should be feasible.
- Convex combinations of observed outputs should be feasible.

Once you have the output set, show how you would find the technical efficiency for the three players.


singles
Fo(x,y) for Rolf=0b/0a
40
Tim
b
20
a
Rolf
P(100)
10
Trent
homeruns
0
5
10
20

Formalities

- $k = 1, \ldots, K$, are the observations or decision making units (DMUs)
- $x^k = (x_{k1}, x_{k2}, \ldots, x_{kN})$ are the inputs 1 to N for observation (DMU) k
- $y^k = (y_{k1}, y_{k2}, \ldots, y_{kM})$ are the outputs 1 to M for observation (DMU) k
- $z = (z_1, z_2, \ldots, z_K)$ are the intensity variables ('dot connectors'), which will be used to construct the best practice frontier

Output technical efficiency for each k=1,...,K, is defined as follows

$$F_o(x^k, y^k) = \max_{\theta, z}\{\theta : \theta y^k \in P(x^k)\} \tag{1}$$

DEA (linear programming) problem for output technical efficiency, $F_o(x^k, y^k)$

For each observation k solve:

$$F_o(x^k, y^k) = \max_{\theta,z} \theta \tag{2}$$

$$\text{subject to}$$

$$\begin{aligned} \sum_{k=1}^{K} z_k y_{km} &\geq \theta y_{km}, m = 1, \ldots, M \\ \sum_{k=1}^{K} z_k x_{kn} &\leq x_{kn}, n = 1, \ldots, M \\ z_k &\geq 0, k = 1, \ldots, K. \end{aligned}$$

REMARK: The inequalities represent the technology, in this case $P(x^k)$.

How DEA constructs the technology sets, eg: P(x)

Recall: $P(x) = \{y :$ y is producible from x$\}$, in DEA, for observation k:

$$
\begin{aligned}
P(x^k) = \{(y_1, \ldots, y_M) : & \\
\sum_{k=1}^{K} z_k y_{km} &\geq y_m, m = 1, \ldots, M \\
\sum_{k=1}^{K} z_k x_{kn} &\leq x_{kn}, n = 1, \ldots, M \\
z_k &\geq 0, k = 1, \ldots, K\}.
\end{aligned} \tag{3}
$$

REMARKS: The inequalities allow for free (strong) disposability of both inputs and outputs. Note that we are *looking for* all possible y's that satisfy the constraints (y_m rather than y_{km} on the RHS).

Now use the following data to write out the constraints for the output set P(x), x=100 at bats. Use z_R for Rolf z_{Ti} for Tim and z_{Tr} for Trent.

	At Bats	Singles	Homeruns
	input	output 1	output 2
Tim	100	40	0
Rolf	100	20	5
Trent	100	10	20

$$
\begin{aligned}
P(100) = \{y : & & \qquad (4)\\
z_{Ti}40 + z_R 20 + z_{Tr}10 &\geq y_{singles} \\
z_{Ti}0 + z_R 5 + z_{Tr}20 &\geq y_{homeruns} \\
z_{Ti}100 + z_R 100 + z_{Tr}100 &\leq 100(input) \\
z_{Ti}, z_R, z_{Tr} &\geq 0
\end{aligned}
$$

P(100) creates additional observations of homeruns and singles by allowing the z's to take any values greater than or equal to zero. What number of singles and homeruns do you get if:

- $z_{Ti} = z_{Tr} = 1/2, z_R = 0$

P(100) creates additional observations of homeruns and singles by allowing the z's to take any values greater than or equal to zero. What number of singles and homeruns do you get if:

- $z_{Ti} = z_{Tr} = 1/2, z_R = 0$
- $y_{singles} \leq 25$, $y_{homeruns} \leq 10$

P(100) creates additional observations of homeruns and singles by allowing the z's to take any values greater than or equal to zero. What number of singles and homeruns do you get if:

- $z_{Ti} = z_{Tr} = 1/2, z_R = 0$
- $y_{singles} \leq 25$, $y_{homeruns} \leq 10$
- $z_{Ti} = z_{Tr} = 0, z_R = 1$

P(100) creates additional observations of homeruns and singles by allowing the z's to take any values greater than or equal to zero. What number of singles and homeruns do you get if:

- $z_{Ti} = z_{Tr} = 1/2, z_R = 0$
- $y_{singles} \leq 25$, $y_{homeruns} \leq 10$
- $z_{Ti} = z_{Tr} = 0, z_R = 1$
- $y_{singles} \leq 20$, $y_{homeruns} \leq 5$

P(100) creates additional observations of homeruns and singles by allowing the z's to take any values greater than or equal to zero. What number of singles and homeruns do you get if:

- $z_{Ti} = z_{Tr} = 1/2, z_R = 0$
- $y_{singles} \leq 25$, $y_{homeruns} \leq 10$
- $z_{Ti} = z_{Tr} = 0, z_R = 1$
- $y_{singles} \leq 20$, $y_{homeruns} \leq 5$
- $z_{Ti} = z_{Tr} = z_R = 1/3$

P(100) creates additional observations of homeruns and singles by allowing the z's to take any values greater than or equal to zero. What number of singles and homeruns do you get if:

- $z_{Ti} = z_{Tr} = 1/2, z_R = 0$
- $y_{singles} \leq 25$, $y_{homeruns} \leq 10$
- $z_{Ti} = z_{Tr} = 0, z_R = 1$
- $y_{singles} \leq 20$, $y_{homeruns} \leq 5$
- $z_{Ti} = z_{Tr} = z_R = 1/3$
- $y_{singles} \leq 23.33$, $y_{homeruns} \leq 8.33$

Now write out the the problem to solve for technical efficiency for Rolf.

$$
\begin{aligned}
F_o(x^R, y^R) = \max \theta : & & \qquad (5)\\
z_{Ti}40 + z_R 20 + z_{Tr}10 &\geq \theta 20(singles)\\
z_{Ti}0 + z_R 5 + z_{Tr}20 &\geq \theta 5(homeruns)\\
z_{Ti}100 + z_R 100 + z_{Tr}100 &\leq 100(input)\\
z_{Ti}, z_R, z_{Tr} &\geq 0
\end{aligned}
$$

solution: $F(x^R, y^R) = 1.45, z^*_{Ti} = .64, z^*_R = 0, z^*_{Tr} = .36$ How many singles and homeruns would Rolf have if he were efficient?