An Exploration of Airlines Productivity & Efficiency (using merge6.csv)

Christophe

June 13, 2013

1 Data

We are looking at the file **AllyearsKLEM.dta**, we have **6318** observations and **25** variables over 15 years. We found **842** different carriers in the database.

We should emphasize that only **152** carriers have **all** production variables (i.e. Y, K, L E & M not null) at least one year (based on "Test0" variable). Here the Y variable is $Yrevenue$.

	Year	N	N.Complete	N.Big	N.Big.Complete
1	1995	187	0	23	0
2	1996	215	0	25	0
3	1997	212	0	11	0
4	1998	270	0	11	0
5	1999	287	0	22	0
6	2000	323	0	21	0
7	2001	347	0	14	0
8	2002	526	56	15	15
9	2003	556	87	23	20
10	2004	538	71	29	20
11	2005	570	103	40	32
12	2006	567	68	43	25
13	2007	565	48	55	22
14	2008	580	46	54	18
15	2009	575	0	18	0

Table 1: The total nb of obs. is 6318 , for 842 unique and 152 complete airlines.

1.1 Framework

We base our analysis on two type subsets of airlines: The whole sample and the sample reduced to majors airlines[1]. We choose 2002 and 2005 to be 2 years of reference in our analysis. These values are parameter of the program and may

[1] major airlines are defined as having an output greater than the third quuartile of Y. See section 2

be changed anytime. We also spoted the LCC when possible.

1.2 Definition of variables

1.3 Descriptive statistics

The number of complete observations is of **103** carriers for 2005. Below is a partial overview of the file's production variables for those **fully** observable airlines those 2 years.

Variable	$\bar{x}$	**Min**	$\widetilde{x}$	**Max**	**#NA**	**n**
Y	2103.35	2.34	485.94	14697.74	0	56
K	14080.65	27.78	3988.90	99079.96	0	56
L	13532.25	63.00	2942.50	120959.00	0	56
E	422.36	0.89	127.35	3262.47	0	56
M	857.44	0.43	183.13	8525.52	0	56

Table 2: Summary of production variables (year 2002), for 56 airlines complete

Variable	$\bar{x}$	**Min**	$\widetilde{x}$	**Max**	**#NA**	**n**
Y	2321.86	0.47	682.49	20657.15	0	103
K	10994.22	3.97	4270.19	87562.29	0	103
L	9203.26	11.00	3213.00	122374.00	0	103
E	306.38	0.02	111.07	2893.22	0	103
M	895.56	0.18	205.84	10920.70	0	103

Table 3: Summary of production variables (year 2005), for 103 unique airlines

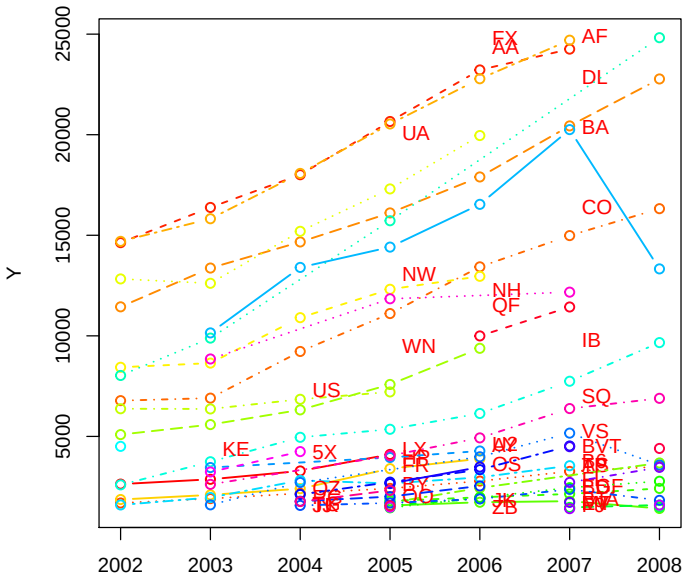
2 Exploration of major airlines

If we take airlines with output Y above the 3rd upper quantile (= 1406 x 1000), we get 75 airlines for all years)

A graph on the output (Y) of the industry's main carrier over time. Here we exclude airlines with problems detected (Test0>0), leaving us with **44** major airlines.[2]

[2]The major with complete production information under scrutiny on this sample are list(carriername = c("UNITED PARCEL", "AMERICAN", "ALASKA", "CONTINENTAL", "DELTA", "FEDERAL EXPRESS", "AMERICA WEST", "NORTHWEST", "UNITED", "US AIRWAYS", "SOUTHWEST", "JETBLUE AIRWAYS", "EXPRESSJET AIRLINES INC,", "AMERICAN EAGLE", "AIRTRAN AIRWAYS", "SKYWEST AIRLINES", "EXPRESSJET AIRLINES", "ATLANTIC SOUTHEAST", "FRONTIER AIRLINES", "TAM LINHAS AEREAS", "VARIG", "AIR FRANCE", "IBERIA", "KLM", "AUSTRIAN AIRLINES GROUP", "BRITISH AIRWAYS", "SWISS", "VIRGIN ATLANTIC", "BRITISH MIDLAND", "THOMSONFLY", "EASYJET AIRLINE", "RYANAIR", "FINNAIR", "SPANAIR", "TAP AIR PORTUGAL", "MONARCH AIRLINES", "Thomas Cook Airlines", "THY", "KOREAN AIR", "ALL NIPPON AIRWAYS", "SIA", "ASIANA", "THAI AIRWAYS", "BERJAYA AIR", "QANTAS"))

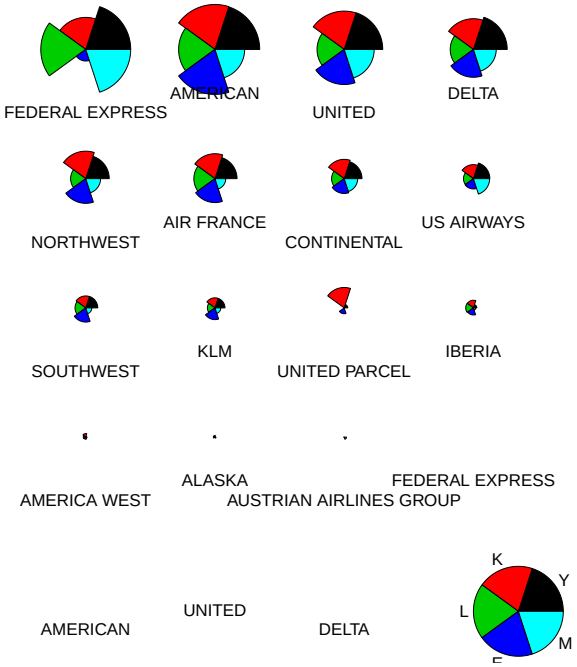
Big Airlines Output (Y) over time



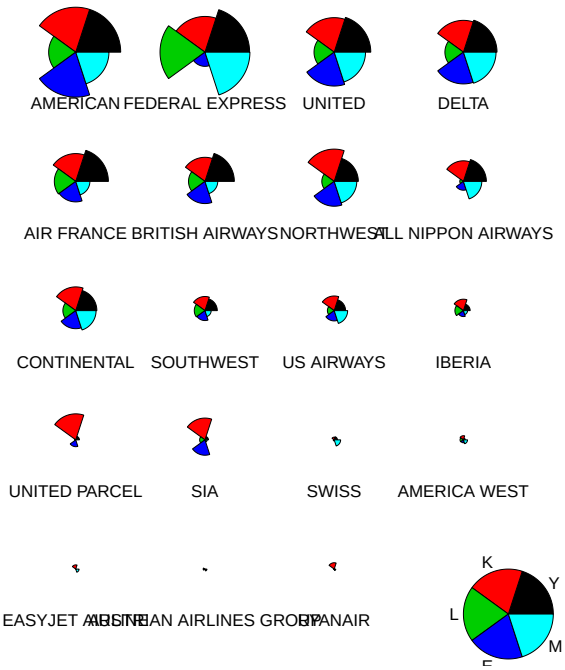
Years

We have 44 distinct carriers on this graph (over 75 majors identified)

Big Airlines Y KLEM typology, year 2002



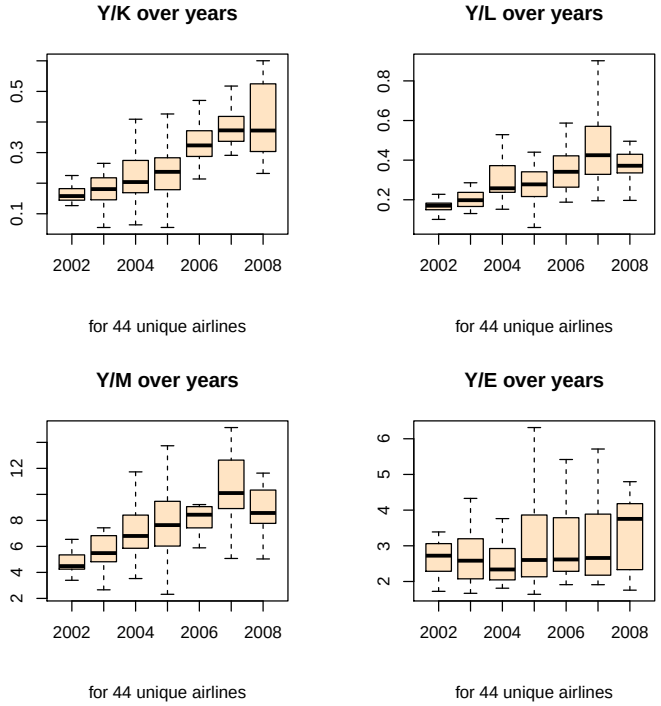
Big Airlines Y KLEM typology, year 2005



Now, we can examine the distribution over time of partial productivity for these major airlines.

Variable	$\bar{x}$	Min	$\widetilde{x}$	Max	#NA	n
Y.over.K	0.25	0.06	0.24	0.67	0	32
Y.over.L	0.34	0.06	0.28	0.92	0	32
Y.over.E	8.16	2.32	7.64	18.60	0	32
Y.over.M	3.13	1.64	2.60	7.69	0	32

Table 4: Summary of partial productivity (year 2005), for 32 major airlines with complete information

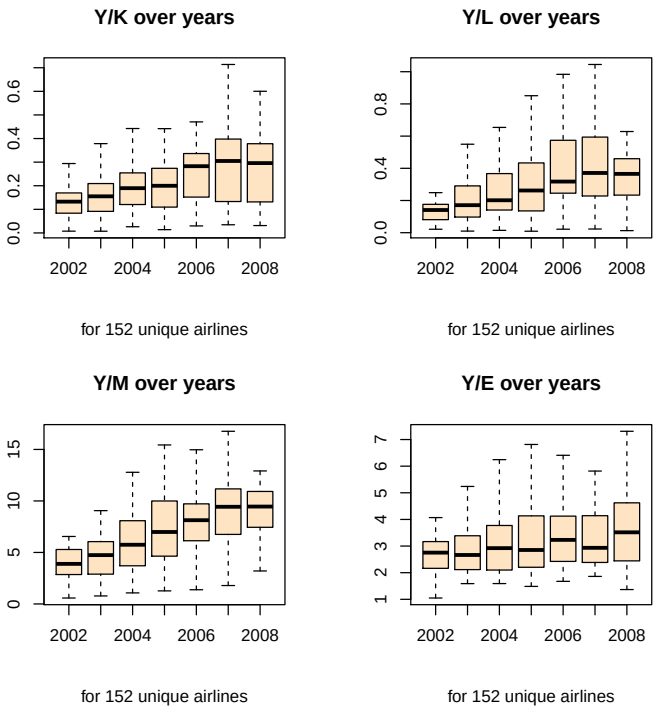


3 Exploration of the whole industry

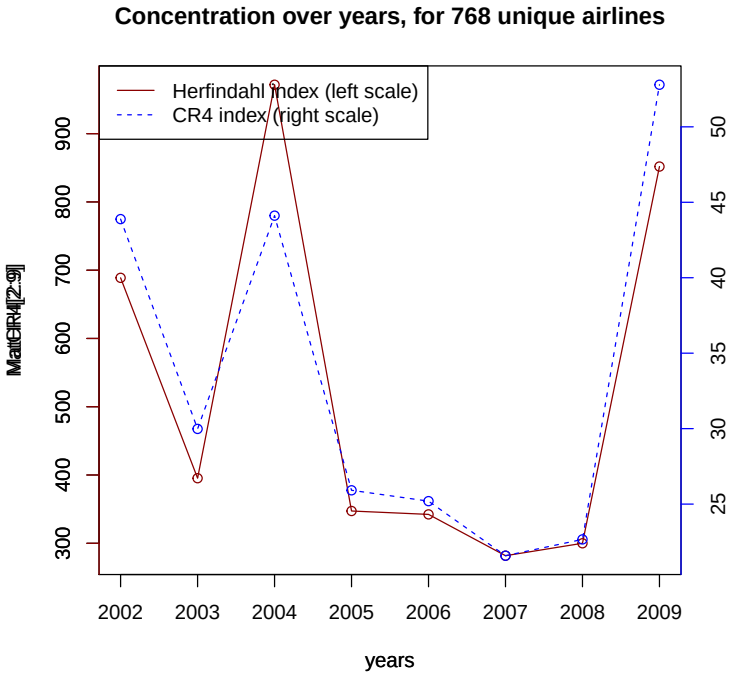
Variable	$\bar{x}$	Min	$\widetilde{x}$	Max	#NA	n
Y.over.K	0.23	0.01	0.20	2.35	0	103
Y.over.L	0.34	0.01	0.26	2.80	0	103
Y.over.E	17.36	1.27	6.99	919.68	0	103
Y.over.M	3.78	1.48	2.85	31.72	0	103

Table 5: Summary of partial productivity (year 2005), for 103 airlines with complete information

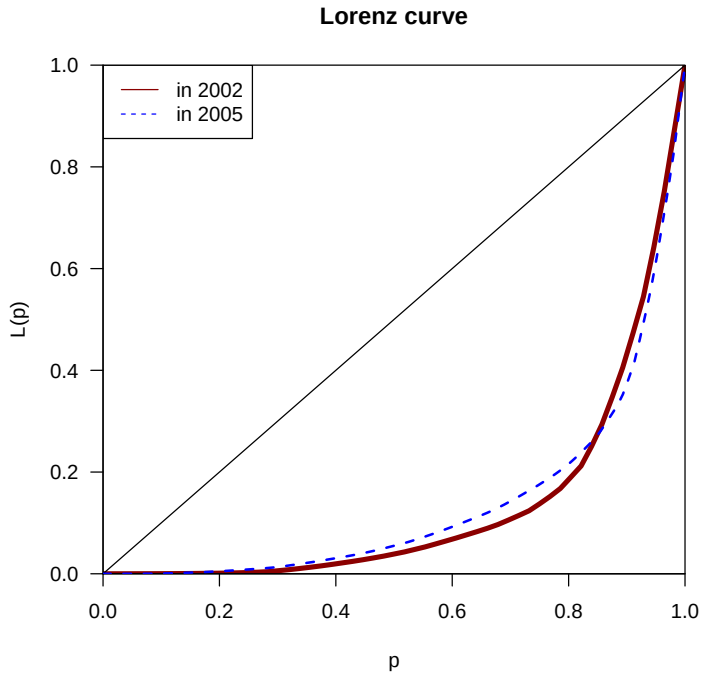
Now, we can examine the distribution over time of partial productivity for all the airlines.



An overview of the industry concentration for the sample of major firms over time.



The Gini index equals **0.742** in 2002 and **0.729** in 2005. The Lorentz curves below show the distribution of output over the the industry for these years.[3]



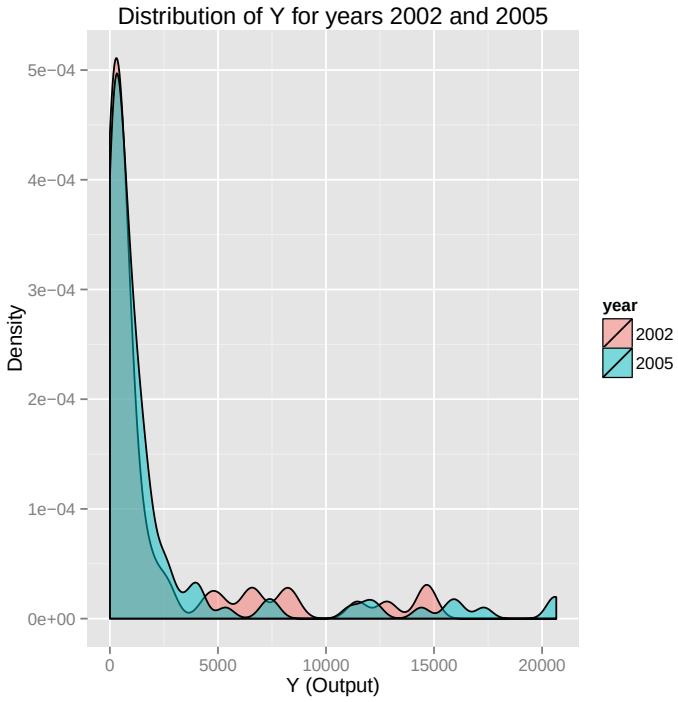
[3]It shows for any % of firms (x axis), what percentage (y axis) of the total output they produce.

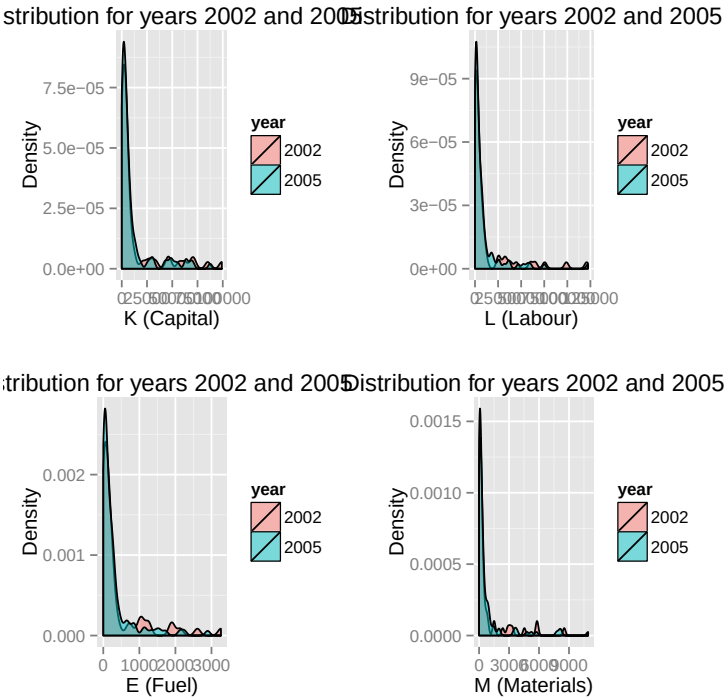
4 Productivity & efficency in 2002 and 2005

4.1 Can we consider that the technology is the same for "big" and "small" airlines in the industry

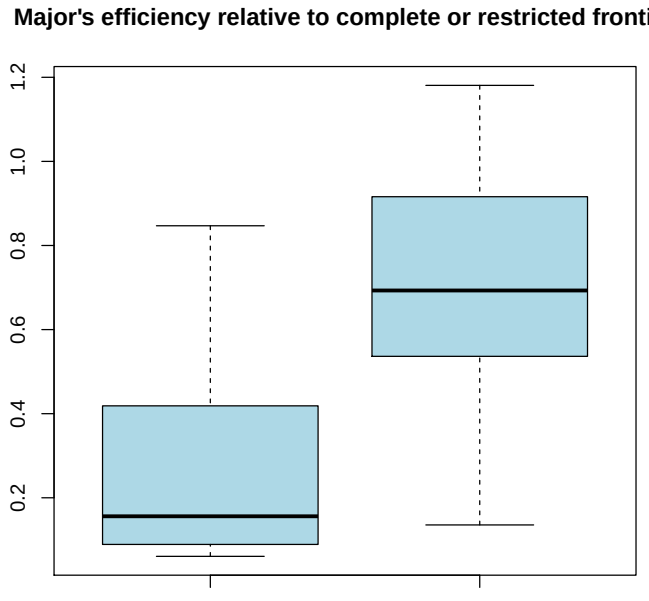
Todo

- NP distribution of Y, K, L, M
- Comparing distribution of big airlines efficiency relative to frontier constructed with or without small airlines
- Airline industry evidence ? definition of output, etc...





Efficiency scores using either the frontier built with all the airlines (left pane), or only the frontier of majors (right pane)

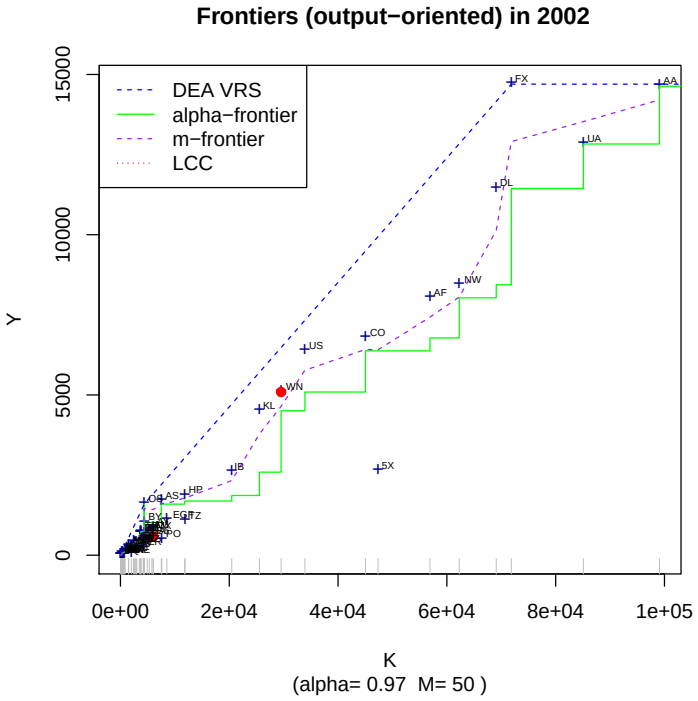


All sample vs restricted (to majors), year 2006 (method DEA).

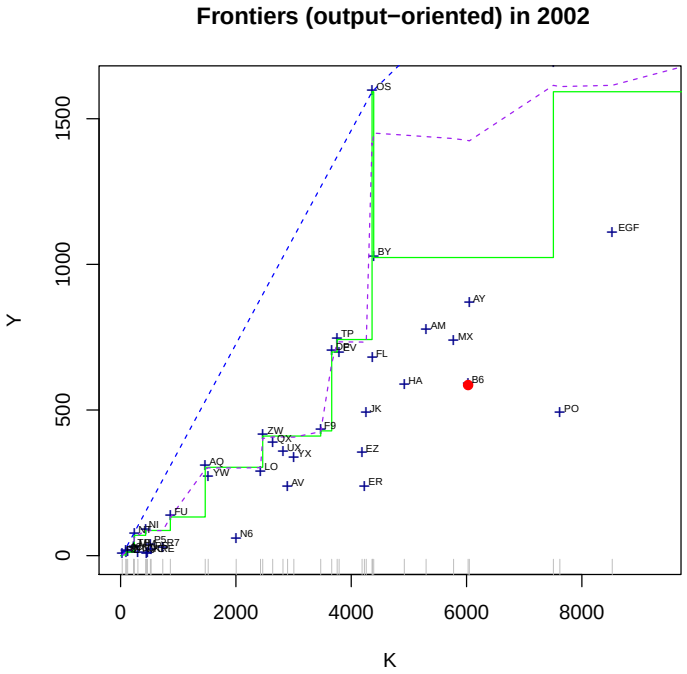
We'd like to do the same with m-frontiers but do not have enough points at this stage...

4.2 DEA, m-frontiers & α-frontier in 2002 and 2005

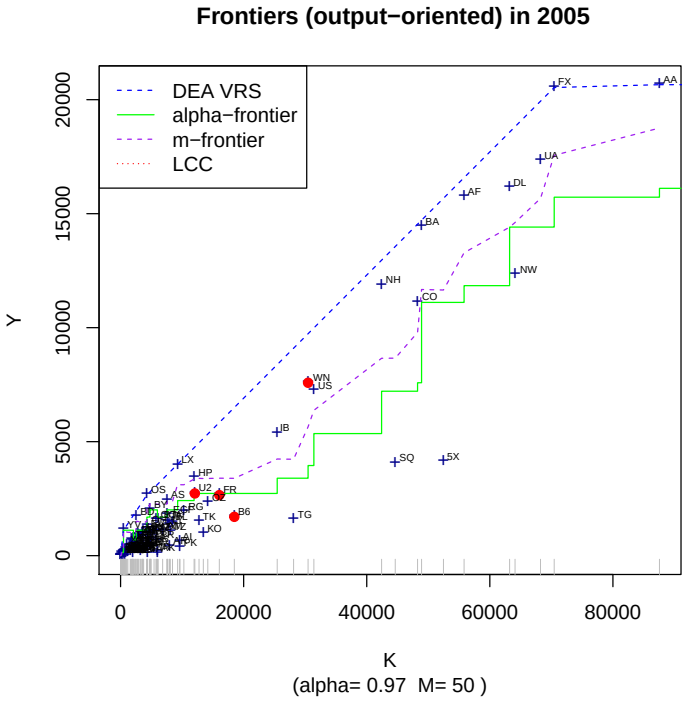
We compute below the α-frontier with only one input (K, here) and one output and graph it for 2002 (**56** obs.) and 2005 (**103** obs.). As soon as we have and aggregate input index, we'll replace it here...



Zoom sous la mediane...

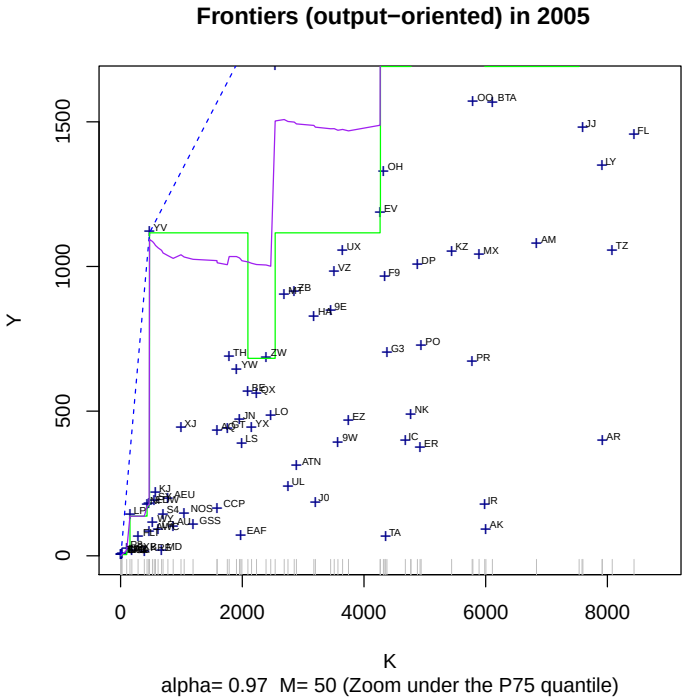


alpha= 0.97 M= 50 (Zoom under the P75 quantile)



(alpha= 0.97 M= 50)

On zoome



alpha= 0.97 M= 50 (Zoom under the P75 quantile)

4.3 Statisitics for m and α -frontiers

Variable	$\bar{\mathbf{x}}$	**Min**	$\widetilde{\mathbf{x}}$	**Max**	**n**	$\bar{N}b_{Eff}$	$\bar{N}b_{Super}$	$\bar{X}_{not}$
DEA..vrs.	0.79	0.25	0.86	1.00	56	18	0	0.69
FDH	0.98	0.56	1.00	1.00	56	50	0	0.84
m.Frontier	1.03	0.57	1.02	1.24	56	11	39	0.85
alpha.Frontier	1.12	0.56	1.00	2.57	56	33	17	0.84

Table 6: Efficiency scores for 2002 , m= 50 , alpha = 0.97 (output oriented, 4 inputs).

Variable	$\bar{\mathbf{x}}$	**Min**	$\widetilde{\mathbf{x}}$	**Max**	**n**	$\bar{N}b_{Eff}$	$\bar{N}b_{Super}$	$\bar{X}_{not}$
DEA..vrs.	0.47	0.02	0.43	1.00	56	3	0	0.44
FDH	0.73	0.04	0.88	1.00	56	23	0	0.54
m.Frontier	0.76	0.04	0.91	1.19	56	4	20	0.54
alpha.Frontier	0.81	0.04	0.93	1.74	56	15	11	0.55

Table 7: Efficiency scores for 2002 , m= 50 , alpha = 0.97 (output oriented, K as unique input).

Variable	$\bar{\mathbf{x}}$	**Min**	$\widetilde{\mathbf{x}}$	**Max**	**n**	$\bar{N}b_{Eff}$	$\bar{N}b_{Super}$	$\bar{X}_{not}$
DEA..vrs.	0.69	0.18	0.70	1.00	103	20	0	0.62
FDH	0.94	0.35	1.00	1.00	103	82	0	0.71
m.Frontier	1.03	0.38	1.02	1.66	103	14	69	0.73
alpha.Frontier	1.16	0.35	1.00	3.49	103	53	34	0.71

Table 8: Efficiency scores for 2005 , m= 50 , alpha = 0.97 (output oriented, 4 inputs).

Variable	$\bar{\mathbf{x}}$	**Min**	$\widetilde{\mathbf{x}}$	**Max**	**n**	$\bar{Nb}_{Eff}$	$\bar{Nb}_{Super}$	$\bar{X}_{not}$
DEA..vrs.	0.36	0.01	0.30	1.00	103	5	0	0.33
FDH	0.50	0.01	0.46	1.00	103	20	0	0.38
m.Frontier	0.57	0.01	0.48	1.37	103	6	18	0.40
alpha.Frontier	0.66	0.01	0.61	1.92	103	12	16	0.45

Table 9: Efficiency scores for 2005 , m= 50 , alpha = 0.97 (output oriented, K as unique input).

4.4 Detailled airlines efficiency scores

	code	Airline	DEA	FDH	M-Score	alpha-Score
1	5X	UNITED PARCEL	1	1	1.186	2.569
2	KL	KLM	1	1	1.244	2.418
3	AS	ALASKA	1	1	1.131	1.651
4	AF	AIR FRANCE	1	1	1.177	1.577
5	TZ	ATA AIRLINES	0.77	1	1.088	1.516
6	US	US AIRWAYS	1	1	1.161	1.415
7	IB	IBERIA	0.868	1	1.113	1.39
8	DL	DELTA	1	1	1.129	1.356
9	CO	CONTINENTAL	0.951	1	1.124	1.332
10	FX	FEDERAL EXPRESS	1	1	1.139	1.285

Table 10: 50 first airlines efficiency, year 2002 , m= 50 , alpha = 0.97 (output oriented, 4 inputs), 56 airlines (sorted by alpha-scores).

	code	Airline	DEA	FDH	M-Score	alpha-Score
1	KL	KLM	0.79	1	1.192	1.739
2	IB	IBERIA	0.549	1	1.113	1.39
3	DL	DELTA	0.808	1	1.129	1.356
4	FX	FEDERAL EXPRESS	1	1	1.139	1.285
5	US	US AIRWAYS	0.87	1	1.105	1.252
6	AF	AIR FRANCE	0.681	1	1.081	1.185
7	WN	SOUTHWEST	0.786	1	1.096	1.13
8	HP	AMERICA WEST	0.613	1	1.043	1.103
9	CO	CONTINENTAL	0.714	1	1.056	1.063
10	AS	ALASKA	0.767	1	1.046	1.061

Table 11: 50 first airlines efficiency, year 2002 , m= 50 , alpha = 0.97 (output oriented, 1 input), 56 airlines (sorted by alpha-scores).

	code	Airline	DEA	FDH	M-Score	alpha-Score
1	NH	ALL NIPPON AIRWAYS	1	1	1.658	3.486
2	5X	UNITED PARCEL	1	1	1.472	3.475
3	BA	BRITISH AIRWAYS	1	1	1.54	2.691
4	WN	SOUTHWEST	0.894	1	1.372	2.232
5	US	US AIRWAYS	0.754	1	1.431	2.123
6	AF	AIR FRANCE	1	1	1.293	2.073
7	FR	RYANAIR	0.958	1	1.249	1.975
8	LX	SWISS	1	1	1.327	1.969
9	IB	IBERIA	0.806	1	1.338	1.963
10	BY	THOMSONFLY	0.962	1	1.181	1.799

Table 12: 50 first airlines efficiency, year 2005 , m= 50 , alpha = 0.97 (output oriented, 4 inputs), 103 airlines, (sorted by alpha-scores).

	code	Airline	DEA	FDH	M-Score	alpha-Score
1	WN	SOUTHWEST	0.779	1	1.342	1.919
2	NH	ALL NIPPON AIRWAYS	0.914	1	1.368	1.642
3	LX	SWISS	0.985	1	1.263	1.636
4	IB	IBERIA	0.639	1	1.263	1.576
5	OS	AUSTRIAN AIRLINES GROUP	1	1	1.2	1.573
6	BD	BRITISH MIDLAND	0.864	1	1.125	1.515
7	CO	CONTINENTAL	0.764	0.938	1.135	1.465
8	US	US AIRWAYS	0.722	0.951	1.131	1.347
9	AF	AIR FRANCE	0.948	1	1.184	1.328
10	FX	FEDERAL EXPRESS	1	1	1.169	1.306

Table 13: 50 first airlines efficiency, year 2005 , m= 50 , alpha = 0.97 (output oriented, 1 input), 103 airlines, (sorted by alpha-scores).

5 Dynamic Analysis

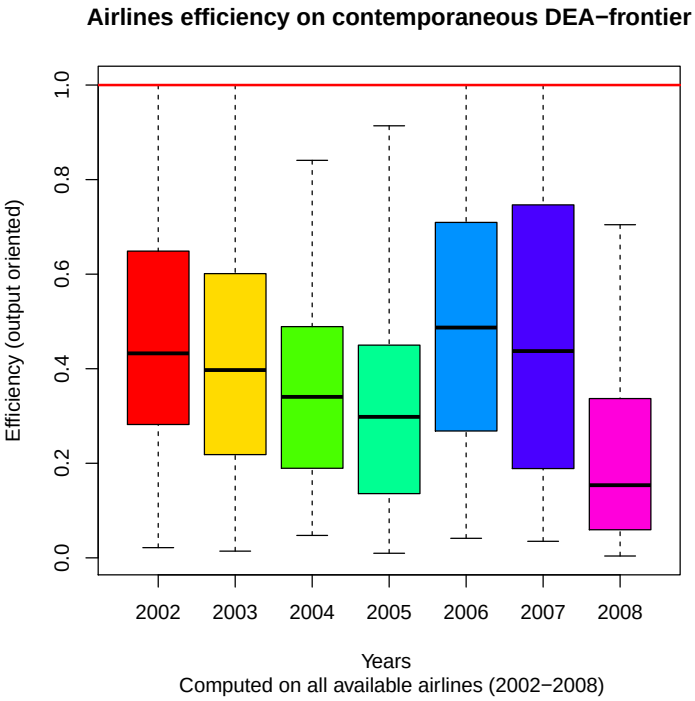
5.1 Efficency over time

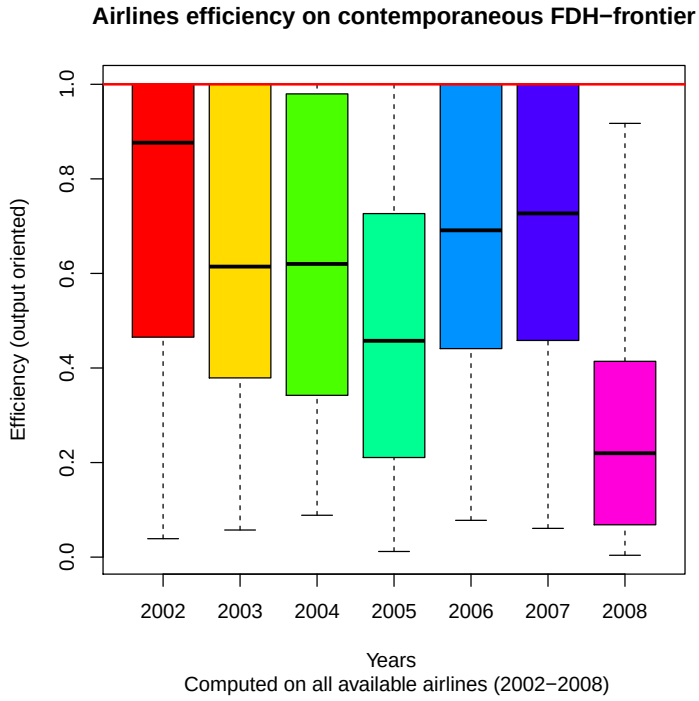
We propose to compute the efficiency (resp. m-efficiency and α-efficiency) of arlines observed each year relative to the frontier (resp. m-frontier and α-frontier) constructed over the pooled sample of airlines. This frontier may be viewed as the "industry" frontier constructed as the union of the yearly observed production sets and serves as a reference to see if firms distance to that frontier has changed over time.

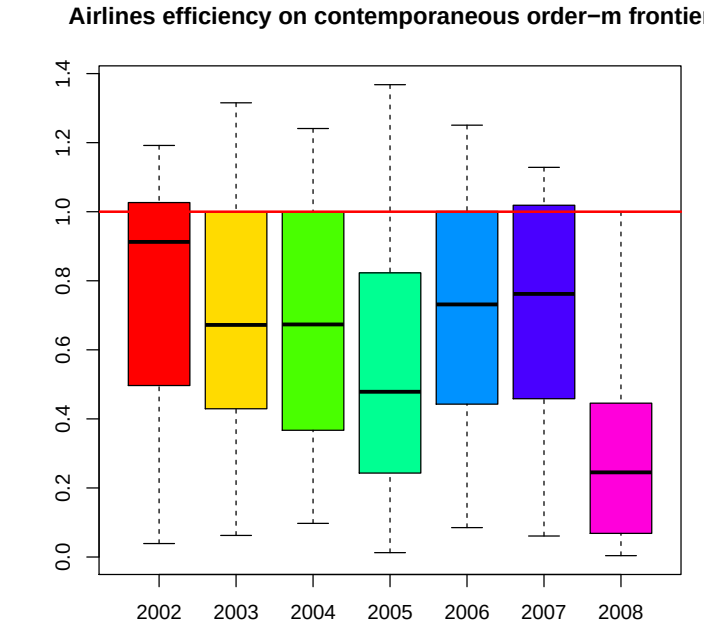
Todo

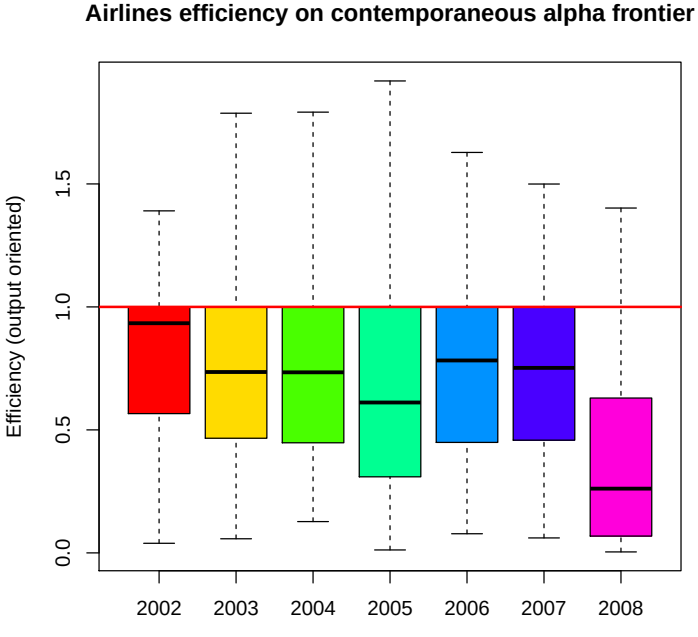
- Dynamic representation of frontiers over time with googleVis, **need a X index !**
- boxplots of efficencies over time (see if it goes up)
- Ranking and rank tests over time ?

5.2 Contemporaneous efficiencies 1 seul input (11/06/13)

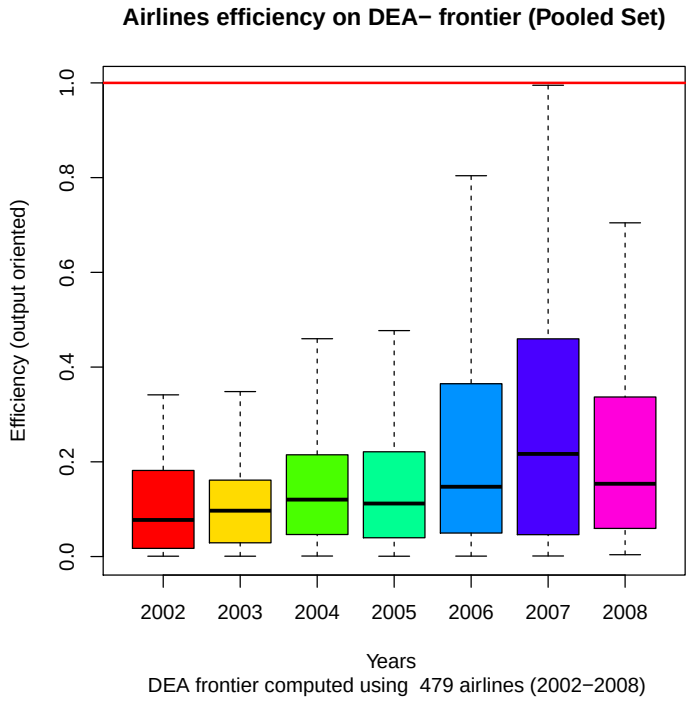


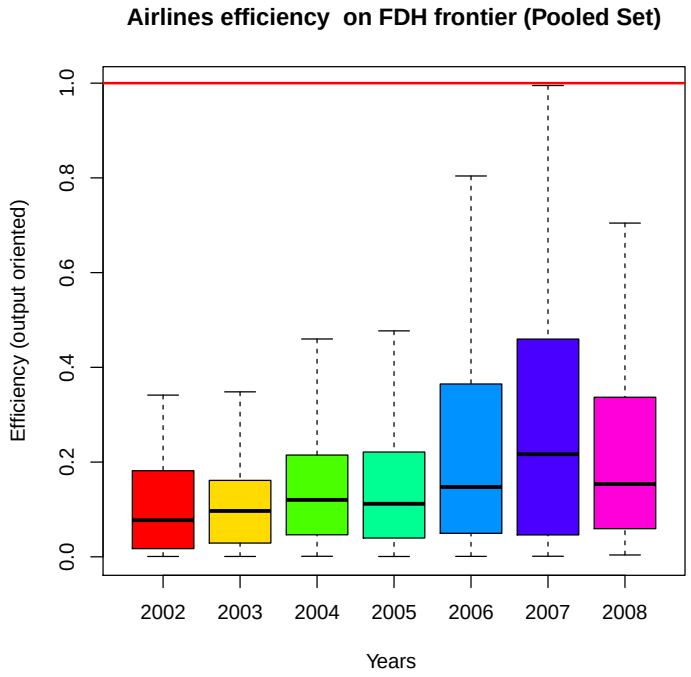


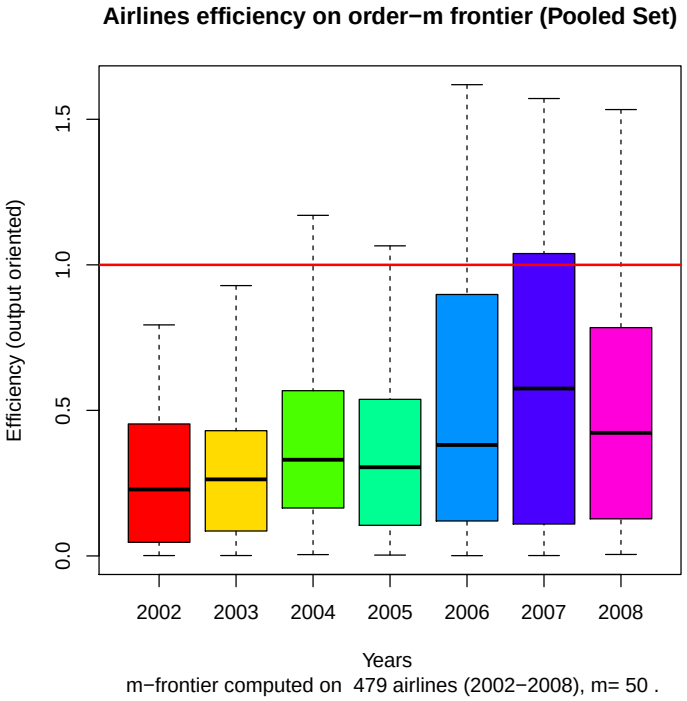

Airlines efficiency on contemporaneous order−m frontier
0.0
0.2
0.4
0.6
0.8
1.0
1.2
1.4
2002
2003
2004
2005
2006
2007
2008
Years
Computed on all available airlines (2002−2008), m= 50 .

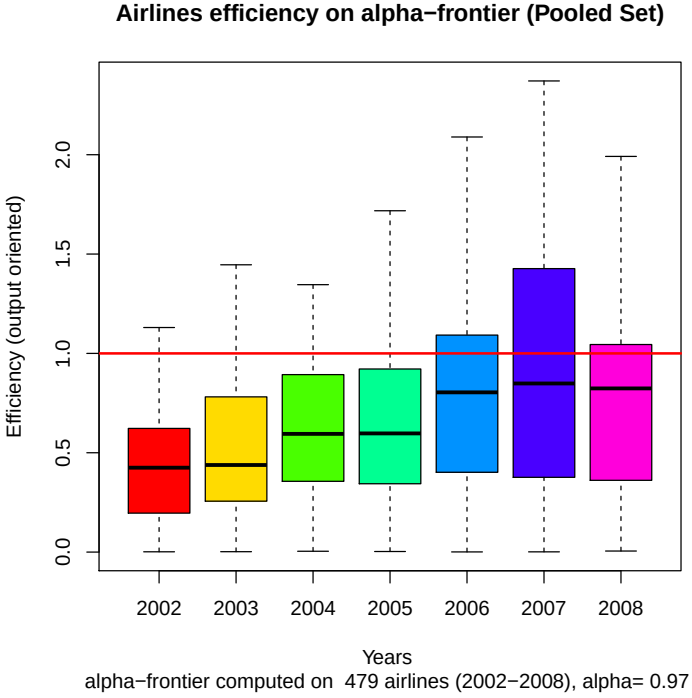

Airlines efficiency on contemporaneous alpha frontier
Efficiency (output oriented)
0.0
0.5
1.0
1.5
2002
2003
2004
2005
2006
2007
2008
Years
Computed on all available airlines (2002−2008), alpha = 0.97 .

5.3 Efficiencies computed relative to POOLED frontier

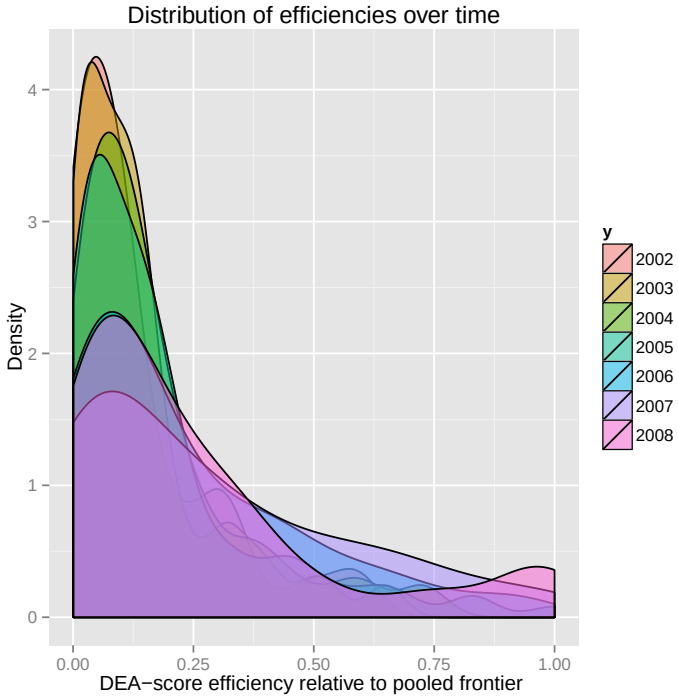


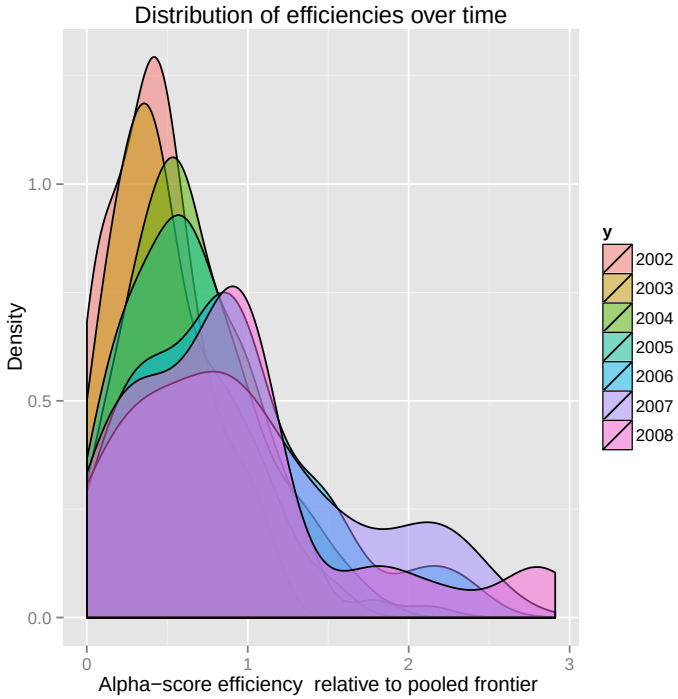



Airlines efficiency on order−m frontier (Pooled Set)
Efficiency (output oriented)
0.0
0.5
1.0
1.5
2002
2003
2004
2005
2006
2007
2008
Years
m−frontier computed on 479 airlines (2002−2008), m= 50 .


Airlines efficiency on alpha−frontier (Pooled Set)
Efficiency (output oriented)
0.0
0.5
1.0
1.5
2.0
2002
2003
2004
2005
2006
2007
2008
Years
alpha−frontier computed on 479 airlines (2002−2008), alpha= 0.97 .

Other representations :





5.4 Malmquist decomposition (Simar & Wilson 1998 ; Wheelock & Wilson, 1999)

Malmquist decomposition proposed by Simar & Wilson (1998) & Wheelock & Wilson (1999) splits the malmquist index into 4 componnets: Pure efficiency change, Scale change, Pure technilogy change & Scale technology change. This decomposition uses intensively DEA estimation (both in CRS and VRS) and follows the formula:

$$
\begin{aligned}
MI &= Pure.Eff \times Scale \\
&\times Pure.Tech \\
&\times Scale.Tech \\
&= \left(\frac{D_{\mathbf{c}}^{VRS}(x_{\mathbf{c}}, y_{\mathbf{c}})}{D_{\mathbf{b}}^{VRS}(x_{\mathbf{b}}, y_{\mathbf{b}})} \right) \times \left(\frac{D_{\mathbf{c}}^{CRS}(x_{\mathbf{c}}, y_{\mathbf{c}}) \ / \ D_{\mathbf{c}}^{VRS}(x_{\mathbf{c}}, y_{\mathbf{c}})}{D_{\mathbf{b}}^{CRS}(x_{\mathbf{b}}, y_{\mathbf{b}}) \ / \ D_{\mathbf{b}}^{VRS}(x_{\mathbf{b}}, y_{\mathbf{b}})} \right) \\
&\times \left(\frac{D_{\mathbf{b}}^{VRS}(x_{\mathbf{c}}, y_{\mathbf{c}})}{D_{\mathbf{c}}^{VRS}(x_{\mathbf{c}}, y_{\mathbf{c}})} \times \frac{D_{\mathbf{b}}^{VRS}(x_{\mathbf{b}}, y_{\mathbf{b}})}{D_{\mathbf{c}}^{VRS}(x_{\mathbf{b}}, y_{\mathbf{b}})} \right)^{0.5} \\
&\times \left(\frac{D_{\mathbf{b}}^{CRS}(x_{\mathbf{c}}, y_{\mathbf{c}}) \ / \ D_{\mathbf{b}}^{VRS}(x_{\mathbf{c}}, y_{\mathbf{c}})}{D_{\mathbf{c}}^{CRS}(x_{\mathbf{c}}, y_{\mathbf{c}}) \ / \ D_{\mathbf{c}}^{VRS}(x_{\mathbf{c}}, y_{\mathbf{c}})} \times \frac{D_{\mathbf{b}}^{CRS}(x_{\mathbf{b}}, y_{\mathbf{b}}) \ / \ D_{\mathbf{b}}^{VRS}(x_{\mathbf{b}}, y_{\mathbf{b}})}{D_{\mathbf{c}}^{CRS}(x_{\mathbf{b}}, y_{\mathbf{b}}) \ / \ D_{\mathbf{c}}^{VRS}(x_{\mathbf{b}}, y_{\mathbf{b}})} \right)^{0.5}
\end{aligned}
$$

where $D_t^{CRS}(x, y)$ incorporates the assumption of *Constant* Return to Scale (CRS), while $D_t^{VRS}(x, y)$ incorporates the assumption of *Variable* Return to Scale (VRS).

6 Why m- and α-frontiers are interesting ?

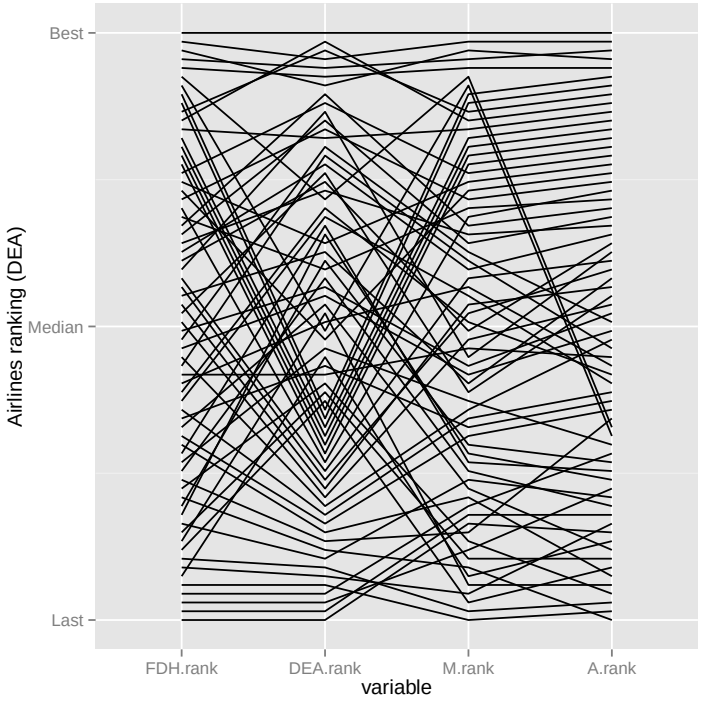
6.1 Definitions and economic interpretation of m- and α-frontiers

6.2 Empirical evidences

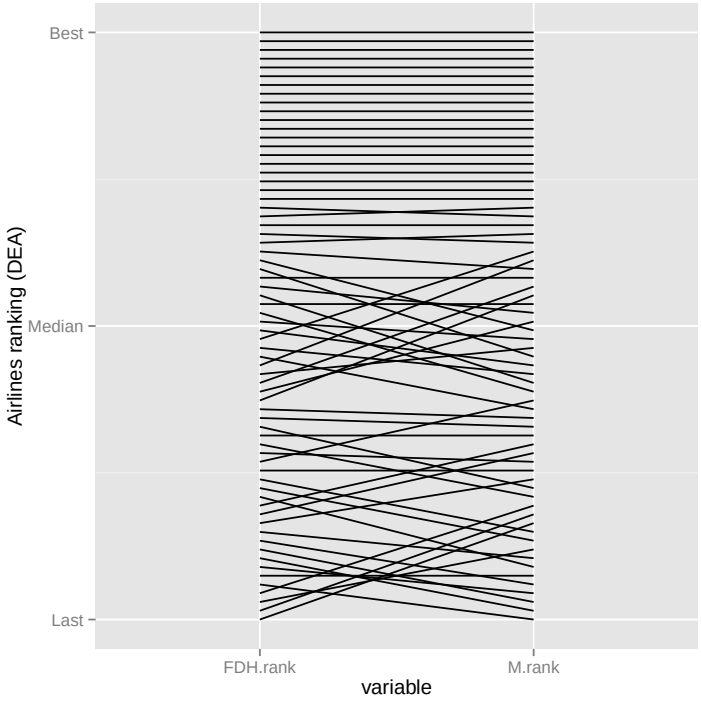
todo

- Tables or graph ranking airlines according to the DEA-efficency, m-efficiency or α-efficiency see if lines cross ?

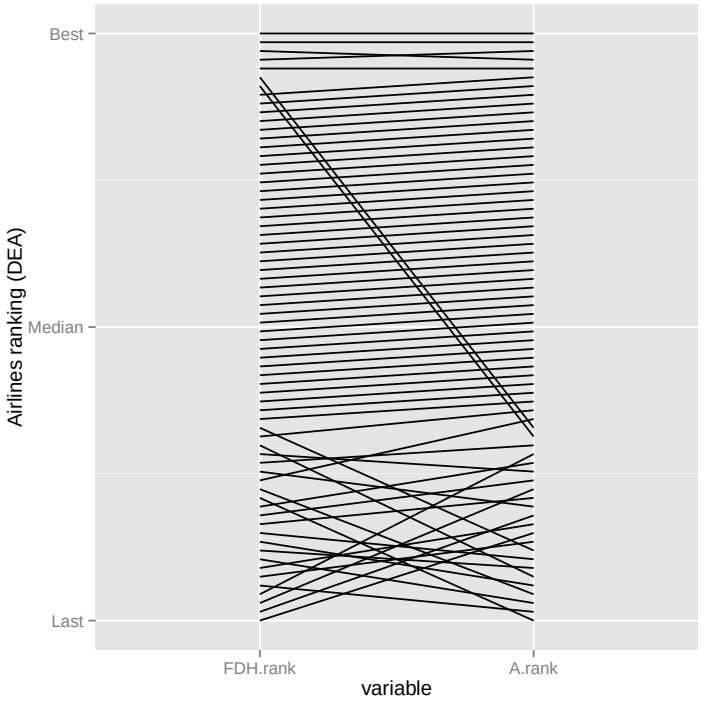
We compare the ranking given by FDH estimator and those given α-frontier estimator and M-estimator for $\alpha = 0.97$ and M $= 50$ for year 2006 and for the 68 observation of that year.



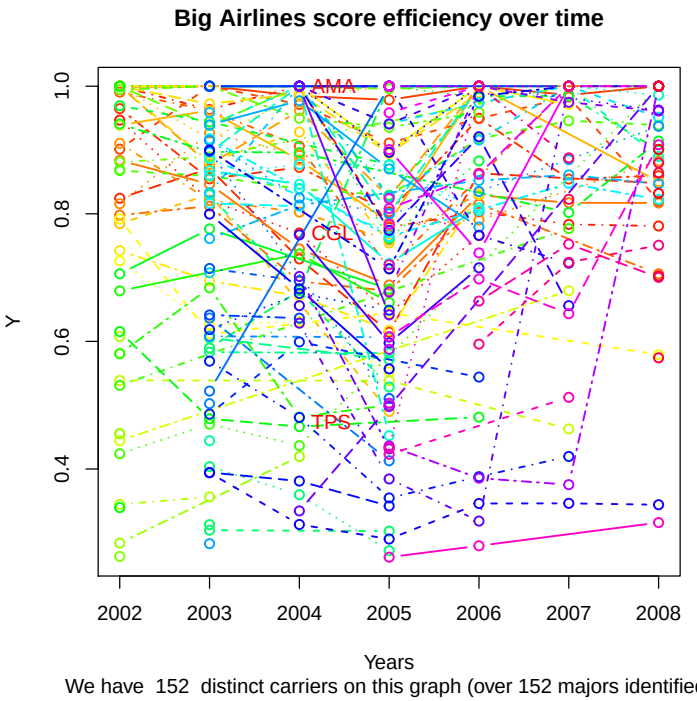
We compare the ranking given by FDH estimator and those given by the m-frontier estimator for $m = 50$.



We compare the ranking given by FDH estimator and those given by the α-frontier estimator for $\alpha = 0.97$.



7 Ranks over time



We have 152 distinct carriers on this graph (over 152 majors identified)