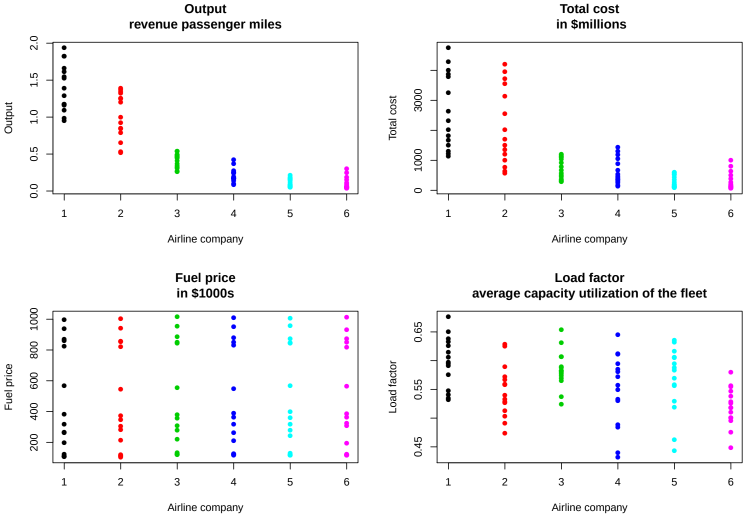
Cost for US Airlines

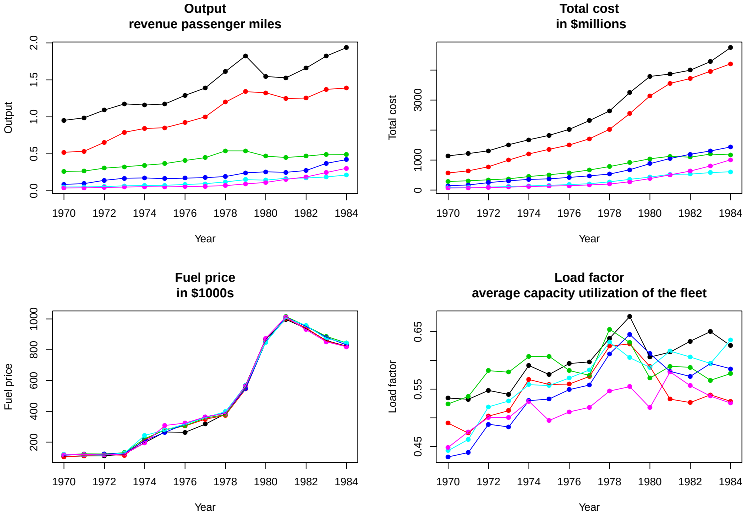
Greene's book on Econometric Analysis (7th edition).

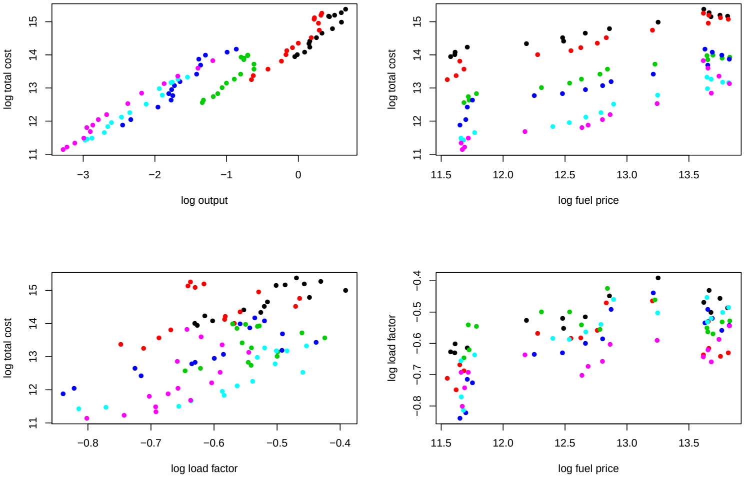
The data can be found in
`http://people.stern.nyu.edu/wgreene/Text/Edition7/TableF6-1.txt`

90 oservations on 6 firms for 15 years (1970-1984)

Columns are:
I : Airline
T : Year
Q : Output, in revenue passenger miles, index number,
C : Total cost, in $1000,
FP : Fuel price,
LF : Load factor, the average capacity utilization of the fleet.


Output
revenue passenger miles
Output
0.0
0.5
1.0
1.5
2.0
Airline company
1
2
3
4
5
6
Total cost
in $millions
Total cost
0
1000
3000
Airline company
Fuel price
in $1000s
Fuel price
200
400
600
800
1000
Airline company
Load factor
average capacity utilization of the fleet
Load factor
0.45
0.55
0.65
Airline company


Output
revenue passenger miles
Output
0.0
0.5
1.0
1.5
2.0
1970
1972
1974
1976
1978
1980
1982
1984
Year
Total cost
in $millions
Total cost
0
1000
3000
Year
Fuel price
in $1000s
Fuel price
200
400
600
800
1000
Year
Load factor
average capacity utilization of the fleet
Load factor
0.45
0.55
0.65
Year



log total cost
log output
log total cost
log fuel price
log total cost
log load factor
log load factor
log fuel price

Let us start by fitting the following model

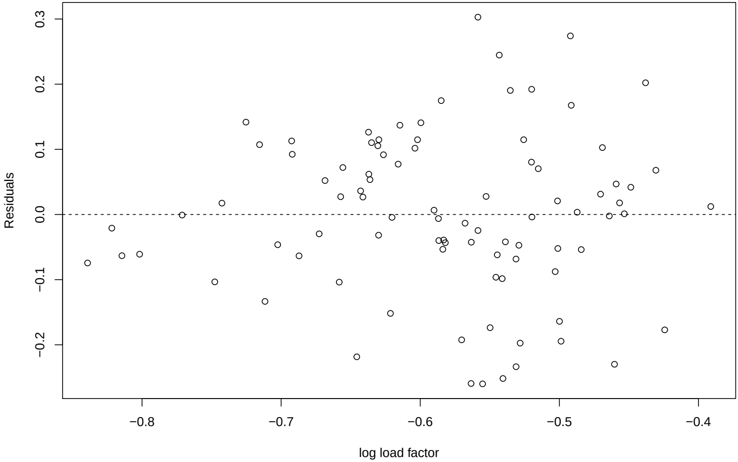
$$\log \texttt{C} = \beta_0 + \beta_1 \log \texttt{Q} + \beta_2 \log \texttt{FP} + \beta_3 \log \texttt{LF} + \varepsilon$$

via OLS.

```
data = read.table("http://people.stern.nyu.edu/wgreene/Text/
Edition7/TableF6-1.txt",header=TRUE)

attach(data)

n = nrow(data)

lc = log(C)
lq = log(Q)
lf = log(PF)
ll = log(LF)
```

```
reg = lm(lc~lq+lf+ll)
coef.ols = reg$coef
se.ols = sqrt(diag(summary(reg)$cov.unscaled))*summary(reg)$sigma

Coefficients:
            Estimate Std. Error t value Pr(>|t|)
(Intercept)  8.07565    0.33420  24.164  < 2e-16 ***
lq           0.88285    0.01330  66.369  < 2e-16 ***
lf           0.45469    0.02046  22.223  < 2e-16 ***
ll          -0.89146    0.19065  -4.676 1.08e-05 ***
---
Signif. codes:  0 *** 0.001 ** 0.01 * 0.05 . 0.1   1

Residual standard error: 0.1248 on 86 degrees of freedom
Multiple R-squared:  0.9883,Adjusted R-squared:  0.9878
F-statistic:  2411 on 3 and 86 DF,  p-value: < 2.2e-16
```

Residual analysis



HR standard errors

```
k = 3
p = k + 1
x = cbind(lq,lf,ll)
X = cbind(1,x)
se.hr = rep(0,ncol(X))
for (i in 1:p){
  reg1 = lm(X[,i]~X[,-i]-1)
  se.hr[i] = sqrt(sum((reg1$res^2)*(reg$res^2))/(sum(reg1$res^2))^2)
}
round(cbind(coef.ols,se.ols,se.hr),3)

            coef.ols se.ols se.hr
(Intercept)    8.076  0.334 0.327
lq             0.883  0.013 0.009
lf             0.455  0.020 0.020
ll            -0.891  0.191 0.169
```

HR F TEST

```
# Hetereskedasticity-robust F test
esq      = reg$res^2
R2.e     = summary(lm(esq~x))$r.sq
Ftest  = R2.e/(1-R2.e)*(n-p)/k
pvalue = 1-pf(Ftest,k,n-p)

> R2.e
[1] 0.04748007

> Ftest
[1] 1.428942

> pvalue
[1] 0.2399109
```

Generalized LS

```
# Generalized least squares
lesq = log(esq)
reg1 = lm(lesq~x)
g    = reg1$fit
hhat = exp(g)
reg.gls = lm(lc~x,weights=1/hhat)
sig = summary(reg.gls)$sigma
se.gls = sqrt(diag(summary(reg.gls)$cov.unscaled))*sig
coef.gls = reg.gls$coef

round(cbind(coef.ols,se.ols,se.hr,coef.gls,se.gls),3)

            coef.ols se.ols se.hr coef.gls se.gls
(Intercept)    8.076  0.334 0.327    8.144  0.341
lq             0.883  0.013 0.009    0.882  0.014
lf             0.455  0.020 0.020    0.452  0.021
ll            -0.891  0.191 0.169   -0.843  0.189
```

Maximizing the likelihood

```
# Minus the likelihood function
like = function(beta){
   m = beta[1]+beta[2]*lq+beta[3]*lp+beta[4]*ll
   s = exp((beta[5]+beta[6]*ll)/2)
   return(-sum(dnorm(lc,m,s,log=TRUE)))
}

# initial values
beta = c(lm(lc~lq+lp+ll)$coef,lm(lesq~ll)$coef)

# Nonlinear minimization.
nlm(like,beta)

$minimum
[1] -64.56045

$estimate
[1]  8.2403727  0.8790963  0.4478516 -0.7548482 -1.4087092  4.9078414

$gradient
[1] -4.408520e-06  4.859189e-06 -5.569056e-05  2.446399e-06  1.361860e-08  2.923048e-08

$code
[1] 1

$iterations
[1] 45
```

	coef.ols	coef.gls	coef.mle
(Intercept)	8.076	8.144	8.240
lq	0.883	0.882	0.879
lf	0.455	0.452	0.448
ll	-0.891	-0.843	-0.755

Therefore, by MLE, the estimated model is

$$\begin{aligned} E(\log \mathtt{C} \mid \mathtt{Q}, \mathtt{FP}, \mathtt{LF}) &= 8.24 + 0.88 \log \mathtt{Q} + 0.45 \log \mathtt{FP} - 0.76 \log \mathtt{LF} \\ V(\log \mathtt{C} \mid \mathtt{LF}) &= \exp\{-1.41 + 4.91 \log \mathtt{LF}\} \end{aligned}$$