E64: Panel Data Stochastic Frontier Models

E64.1 Introduction

The stochastic frontier model as it appears in the current literature was originally developed by Aigner, Lovell, and Schmidt (1977). The canonical formulation that serves as the foundation for other variations is their model,

$$y = \boldsymbol{\beta}'\mathbf{x} + v - u,$$

where y is the observed outcome (goal attainment), $\boldsymbol{\beta}'\mathbf{x} + v$ is the optimal, frontier goal (e.g., maximal production output or minimum cost) pursued by the individual, $\boldsymbol{\beta}'\mathbf{x}$ is the deterministic part of the frontier and $v \sim N[0,\sigma_v^2]$ is the stochastic part. The two parts together constitute the 'stochastic frontier.' The amount by which the observed individual fails to reach the optimum (the frontier) is u, where

$$u = |U| \text{ and } U \sim N[0,\sigma_u^2]$$

(change to $v + u$ for a stochastic cost frontier or any setting in which the optimum is a minimum). In this context, u is the 'inefficiency.' This is the normal-half normal model which forms the basic form of the stochastic frontier model. Chapters E62 and E63 developed several versions of the stochastic frontier model suitable for cross section and pooled data sets. This chapter will develop versions of the model constructed specifically for panel data.

E64.2 Panel Data Estimators for Stochastic Frontier Models

The stochastic frontiers literature has steadily evolved since the developments of basic random and fixed effects models by Pitt and Lee (1981) and by Cornwell, Schmidt and Sickles (1990). All of the generally used forms of panel data models are supported in *LIMDEP*. The following will document them in detail. These sections are arranged as follows:

- Pitt and Lee – Time Invariant Inefficiency, Random Effects,
- Cornwell, Schmidt and Sickles – Time Invariant Inefficiency, Fixed Effects,
- Battese and Coelli – Time Dependent Inefficiency Models,
- True Fixed Effects Models with Time Varying Inefficiency,
- True Random Effects Models with Time Varying Inefficiency,
- Random Parameters Stochastic Frontier Models,
- Alvarez et al. – Fixed Management (Random Parameters) Model,
- Latent Class Stochastic Frontier Models.

The panel models developed here will share features with other panel models in *LIMDEP*, as presented in Chapter R22-R25. As in other settings, panels in all models may be unbalanced. Panels are identified by

SETPANEL ; … $

then **; Panel**

in the command, or **; Pds = group count**

Nearly all of the models to be presented here actually require panel data, but a few will work, albeit not as well as otherwise, with **; Pds = 1**, i.e., with a cross section. This will be specifically noted below when it is the case. Second, in all models, the cost form as opposed to the production form is requested with

; Cost

This and other model specifications are generally the same as the cross sectional cases.

E64.3 Pitt and Lee – Time Invariant Inefficiency, Random Effects

The panel data, random effects specifications based on the model of Pitt and Lee (1982) are

$$y_{it} = \alpha + \boldsymbol{\beta}'\mathbf{x}_{it} + v_{it} - Su_i$$

with S = +1 for a production model and -1 for a cost model. The inefficiency component is assumed to be time invariant. The base case is the normal-half normal model

$$u_i = |U_i|,\ U_i \sim \mathrm{N}[0,\sigma^2].$$

This is a direct extension of the cross section variant discussed earlier. Several model formulations are grouped in this class. The command for the Pitt and Lee group of models is given by changing the base case specifications to

FRONTIER ; Lhs = y ; Rhs = one, ... ; Panel $

Pitt and Lee is the default panel data model. The only necessary change for the default case is specification of the panel with **; Panel**. As in the cross section case, the normal-exponential case is requested with

; Model = Exponential

while the normal-truncated normal is requested with

; Rh2 = one or ; Rh2 = one, additional variables

(The **; Model = T** is not needed.) The truncation model may not be combined with the exponential specification; it is only supported for the normal-truncated normal form.

NOTE: The gamma model does not have a random effects (panel data) version. The model extensions, such as the scaling model and sample selection described in Chapter E63 likewise do not support a Pitt and Lee style random effects version.

There is an important consideration for the truncation version with heterogeneous mean. If you are fitting a panel data version of this model, note that the assumption underlying the model is that the same u_i occurs in every period. Therefore, the $\boldsymbol{\alpha}'\mathbf{z}_i$ must be the same in every period. *LIMDEP* will assume this is the case, and only use the Rh2 variables provided for the first period.

When the random effects model is estimated, maximum likelihood estimates of the cross section models are always computed first to obtain the starting values. This will produce a full set of results which will ignore the panel nature of the data set. A second full set of results will then follow for the random effects model.

The model estimates retained for all cases are

b = regression parameters, α,β
varb = asymptotic covariance matrix.

Use **; Par** to retain the additional parameters in *b* and *varb*. As seen in the applications below, the parameters estimated in each case will differ depending on the model formulation. The ancillary parameters that are estimated for the various models are the same ones saved by the cross section versions. All models save *sy*, *ybar*, *nreg*, *kreg*, and *logl* as well as *s*, *b*, *varb*, etc.

WARNING: Numerous experiments and applications have suggested that the normal-truncated normal model is a difficult one to estimate. Identification appears to be highly variable, and small variations in the data can produce large variation in the results. The model often fails to converge even when convergence of the restricted model with zero underlying mean is routine.

E64.3.1 Model Specifications

There are many different combinations of the components of the random effects model listed above. The following shows the different possibilities for the Pitt and Lee model. (There are also many combinations of these that do not use the panel data random effects form.):

```
    NAMELIST   ; x = one, … $
    CREATE     ; y = the outcome variable $
    SETPANEL   ; … $
Model 1 = pooled
    FRONTIER   ; Lhs = y ; Rhs = x $
Model 2 = random effects half normal
    FRONTIER   ; Lhs = y ; Rhs = x ; Panel  $
Model 3 = random effects exponential
    FRONTIER   ; Lhs = y ; Rhs = x ; Panel  ; Model = Exponential $
Model 4 = random effects normal heteroscedastic in u or  v only
    FRONTIER   ; Lhs = y ; Rhs = x ; Panel  ; Het ; Hfv = … $
    FRONTIER   ; Lhs = y ; Rhs = x ; Panel  ; Het ; Hfu = … $
Model 5 = random effects normal doubly heteroscedastic
    FRONTIER   ; Lhs = y ; Rhs = x ; Panel ; Het ; Hfv = … ; Hfu = … $
Model 6 = random effects truncated normal
    FRONTIER   ; Lhs = y ; Rhs = x ; Panel  ; Rh2 = one, … $
Model 7 = random effects truncated normal, singly or doubly heteroscedastic
    FRONTIER   ; Lhs = y ; Rhs = x ; Panel  ; Rh2 = one, …
               ; Het ; Hfv = … ; Hfu = … $
```

The Pitt and Lee model forms assume that the inefficiency is time invariant. Thus, the estimate of u_i is repeated for each observation in the group. An example below illustrates.

E64.3.2 Applications

The following illustrates a few of the numerous formats of the random effects frontiers. The data set used is the Swiss railroad data used in Greene (2011, Table F19.1). These data are provided with the program as swissrailroads.lpj. The variables used here are

ct	= total cost
pk	= capital price
pe	= electricity price
pl	= labor price
q2	= passenger output – passenger km
q3	= freight output – ton km
rack	= dummy variable for 'rack rail' in network
tunnel	= dummy variable for network with tunnels over 300 meters on average
virage	= dummy variable for networks with narrow radius curvature
narrow_t	= dummy variable for narrow track (1m as opposed to standard 1.435m).

Preparing the data set includes bypassing one firm for which there is only a single year of data. For the remaining 49 firms, T_i is a mixture 3, 7, 10, 12 or 13. Figure E64.1 details the distribution of group sizes.

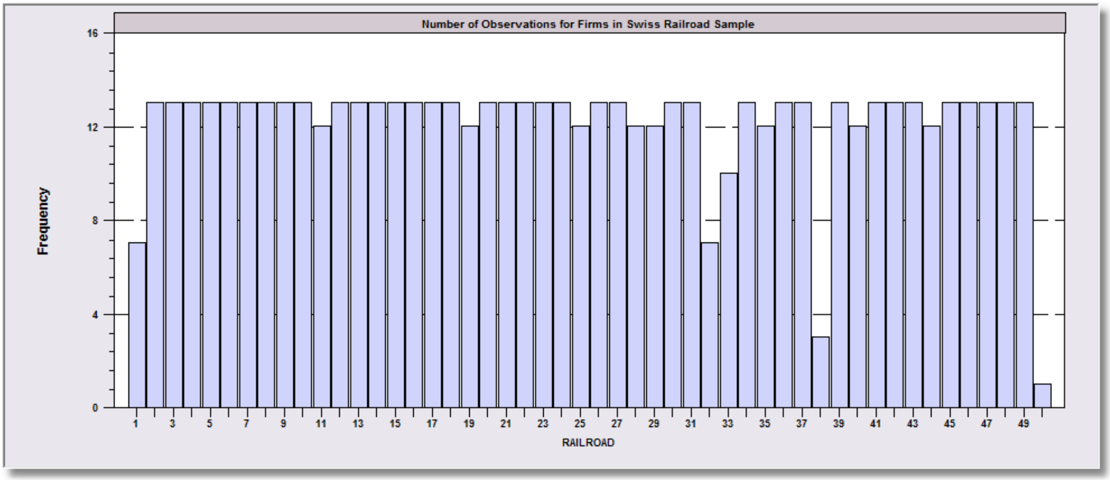

Figure E64.1 Groups Sizes for Swiss Railroad Sample

Descriptive statistics for the data are shown below. Variables with names beginning with 'M' are firm means, repeated for each year for the firm.

We fit four models to illustrate the estimator, the pooled normal-half normal, pooled normal-truncated (heterogeneous), basic Pitt and Lee and a full model with time invariant inefficiency, truncation (heterogeneous) and double heteroscedasticity.

The commands are as follows:

SETPANEL ; Group = id ; Pds = ti $
REJECT ; ti = 1 $
CREATE ; lple = Log(pl/pe) ; lpke = Log(pk/pe) ; lnc = Log(ct/pe)$
NAMELIST ; x = one,lnq2,lnq3,lple,lpke $
FRONTIER ; Lhs = lnc ; Cost ; Rhs = x ; Costeff = eusfpool $
FRONTIER ; Lhs = lnc ; Cost ; Rhs = x $
FRONTIER ; Lhs = lnc ; Cost ; Rhs = x ; Panel ; Costeff = eusfp_l $
FRONTIER ; Lhs = lnc ; Cost ; Rhs = x ; Rh2 = rack,tunnel
; Het ; Hfu = virage ; Hfv = virage ; Costeff = eushet_t $
FRONTIER ; Lhs = lnc ; Cost ; Rhs = x ; panel ; Rh2 = rack,tunnel
; Het ; Hfu = virage ; Hfv = virage ; Costeff = fullmodl $

Variable	Mean	Std.Dev.	Minimum	Maximum	Cases	Missing
ID	25.48760	14.60037	1.0	51.0	605	0
YEAR	90.91570	3.692372	85.0	97.0	605	0
NI	12.58347	1.305259	1.0	13.0	605	0
STOPS	20.42479	18.48285	4.0	121.0	605	0
NETWORK	39431.66	56642.38	3898.0	376997.0	605	0
LABOREXP	12801.95	26232.69	951.0	173549.0	605	0
STAFF	170.3810	333.0317	11.0	1934.0	605	0
ELECEXP	968.1521	1944.830	14.0	14737.0	605	0
KWH	7602.221	15608.39	82.0	104923.0	605	0
TOTCOST	22470.44	42283.57	1534.0	280871.0	605	0
NARROW_T	.676033	.468375	0.0	1.0	605	0
RACK	.234711	.424169	0.0	1.0	605	0
TUNNEL	.188430	.391379	0.0	1.0	605	0
T	5.915702	3.692372	0.0	12.0	605	0
Q1	813914.0	1083923	61000.0	6409000	605	0
Q2	.308145D+08	.550599D+08	409000.0	.311000D+09	605	0
Q3	.101934D+08	.527303D+08	150.0	.477000D+09	605	0
CT	26728.37	49883.51	2120.968	307433.4	605	0
PL	86051.77	6484.535	60932.91	104930.4	605	0
PE	.157485	.022766	.076344	.265182	605	0
PK	4534.491	2128.307	1040.323	14466.06	605	0
VIRAGE	.715702	.451452	0.0	1.0	605	0
LABOR	52.40245	9.598136	20.03025	73.11581	605	0
ELEC	4.044504	1.422098	.568412	9.311660	605	0
CAPITAL	43.55305	9.461303	23.88916	77.33154	605	0
LNCT	11.30622	1.101691	9.462956	14.57019	605	0
LNQ1	13.06322	1.010039	11.01863	15.67321	605	0
LNQ2	16.31759	1.339167	12.92147	19.55500	605	0
LNQ3	12.49439	2.716709	5.010635	19.98343	605	0
LNNET	3.200860	.908512	1.360464	5.932237	605	0
LNPL	13.21935	.163565	12.60449	13.77599	605	0
LNPE	-1.859557	.152870	-2.572503	-1.327338	605	0
LNPK	10.17950	.438886	8.740266	11.37466	605	0

```
LNSTOP|     2.775052      .655071     1.386294     4.795791      605       0
   LNCAP|     3.137572      .328311     2.123893     3.850147      604       1
   MLNQ1|     13.06322     1.005089     11.16747     15.59433      605       0
   MLNQ2|     16.31759     1.333346     13.20185     19.45679      605       0
   MLNQ3|     12.49439     2.648475     7.734539     19.68075      605       0
  MLNNET|     3.200860      .906363     1.360464     5.927817      605       0
   MLNPL|     13.21935      .126548     12.89796     13.61620      605       0
   MLNPK|     10.17950      .396797     8.938699     11.03543      605       0
 MLNSTOP|     2.775052      .651059     1.386294     4.789402      605       0
    LPLE|     13.21943      .163692     12.60449     13.77599      604       1
   LPKPE|     10.16419      .576094          1.0     11.37466      605       0
     LNC|     11.30305     1.099836     9.462957     14.57019      604       1
--------+---------------------------------------------------------------
```

This is the pooled normal-half normal model.

```
-----------------------------------------------------------------------------
Limited Dependent Variable Model - FRONTIER
Dependent variable                  LNC
Log likelihood function      -209.42340
Estimation based on N =    604, K =   7
Inf.Cr.AIC  =    432.8 AIC/N =     .717
Variances: Sigma-squared(v)=     .07332
           Sigma-squared(u)=     .12333
           Sigma(v)        =     .27077
           Sigma(u)        =     .35119
Sigma = Sqr[(s^2(u)+s^2(v)]=     .44345
Gamma = sigma(u)^2/sigma^2 =     .62716
Var[u]/{Var[u]+Var[v]}     =     .37937
Stochastic Cost Frontier Model, e = v+u
LR test for inefficiency vs. OLS v only
Deg. freedom for sigma-squared(u):    1
Deg. freedom for heteroscedasticity:  0
Deg. freedom for truncation mean:     0
Deg. freedom for inefficiency model:  1
LogL when sigma(u)=0         -210.45352
Chi-sq=2*[LogL(SF)-LogL(LS)] =    2.060
Kodde-Palm C*: 95%: 2.706,  99%:  5.412
--------+--------------------------------------------------------------------
        |                 Standard            Prob.      95% Confidence
     LNC|  Coefficient       Error       z    |z|>Z*         Interval
--------+--------------------------------------------------------------------
        |Deterministic Component of Stochastic Frontier Model
Constant|   -10.0907***     1.14284   -8.83  .0000    -12.3306   -7.8507
    LNQ2|     .64179***      .01371   46.80  .0000      .61491    .66867
    LNQ3|     .06855***      .00655   10.46  .0000      .05570    .08139
    LPLE|     .53971***      .08858    6.09  .0000      .36610    .71333
    LPKE|     .26045***      .03260    7.99  .0000      .19655    .32435
        |Variance parameters for compound error
  Lambda|    1.29697***      .13854    9.36  .0000     1.02545   1.56850
   Sigma|     .44345***      .00056  789.05  .0000      .44235    .44455
--------+--------------------------------------------------------------------
Note: ***, **, * ==>  Significance at 1%, 5%, 10% level.
-----------------------------------------------------------------------------
```

This is the original Pitt and Lee normal-half normal model with time invariant inefficiency. In comparison to the pooled model above, σ_u has tripled and σ_v has decreased by two thirds. The assumption of time invariance of the inefficiency produces a large reallocation of the random components between noise and inefficiency. This is evident in the kernel estimate below as well.

```
-----------------------------------------------------------------------------
Limited Dependent Variable Model - FRONTIER
Dependent variable                    LNC
Log likelihood function       527.11659
Estimation based on N =    604, K =   7
Inf.Cr.AIC  =  -1040.2 AIC/N =   -1.722
Stochastic frontier based on panel data
Estimation based on      49 individuals
Variances: Sigma-squared(v)=     .00621
           Sigma-squared(u)=     .92297
           Sigma(v)        =     .07879
           Sigma(u)        =     .96071
Sigma = Sqr[(s^2(u)+s^2(v)]=     .96394
Gamma = sigma(u)^2/sigma^2 =     .99332
Var[u]/{Var[u]+Var[v]}     =     .98183
Stochastic Cost Frontier Model, e = v+u
LR test for inefficiency vs. OLS v only
Deg. freedom for sigma-squared(u):    1
Deg. freedom for heteroscedasticity:  0
Deg. freedom for truncation mean:     0
Deg. freedom for inefficiency model:  1
LogL when sigma(u)=0         -210.45352
Chi-sq=2*[LogL(SF)-LogL(LS)] = 1475.140
Kodde-Palm C*: 95%: 2.706,  99%:  5.412
--------+--------------------------------------------------------------------
        |                  Standard            Prob.      95% Confidence
     LNC|  Coefficient       Error       z    |z|>Z*         Interval
--------+--------------------------------------------------------------------
        |Deterministic Component of Stochastic Frontier Model
Constant|   -7.25643***      .24767   -29.30  .0000    -7.74185  -6.77101
    LNQ2|     .36259***      .01503    24.12  .0000      .33312    .39205
    LNQ3|     .01902***      .00240     7.94  .0000      .01432    .02372
    LPLE|     .64148***      .02112    30.38  .0000      .60009    .68287
    LPKE|     .30842***      .00700    44.08  .0000      .29471    .32214
        |Variance parameters for compound error
  Lambda|    12.1932**      5.55909     2.19  .0283      1.2975   23.0888
Sigma(u)|     .96071***      .13303     7.22  .0000      .69998   1.22145
--------+--------------------------------------------------------------------
Note: ***, **, * ==>  Significance at 1%, 5%, 10% level.
-----------------------------------------------------------------------------
```

This is the pooled normal-truncated and doubly heteroscedastic normal model.

```
-----------------------------------------------------------------------------
Limited Dependent Variable Model - FRONTIER
Dependent variable                    LNC
Log likelihood function       -63.43402
Estimation based on N =    604, K =  11
Inf.Cr.AIC  =    148.9 AIC/N =     .246
Variances: Sigma-squared(v)=     .07144
           Sigma-squared(u)=     .00074
           Sigma(u)        =     .02720
           Sigma(v)        =     .26729
Sigma = Sqr[(s^2(u)+s^2(v)]=     .26867
Variances averaged over observations
LR test for inefficiency vs. OLS v only
Deg. freedom for sigma-squared(u):    1
Deg. freedom for heteroscedasticity:  1
Deg. freedom for truncation mean:     2
Deg. freedom for inefficiency model:  4
LogL when sigma(u)=0         -210.45352
Chi-sq=2*[LogL(SF)-LogL(LS)] =  294.039
Kodde-Palm C*: 95%: 8.761,  99%: 12.483
--------+--------------------------------------------------------------------
        |                 Standard            Prob.      95% Confidence
     LNC|  Coefficient       Error       z    |z|>Z*         Interval
--------+--------------------------------------------------------------------
        |Deterministic Component of Stochastic Frontier Model
Constant|   -13.4218***     1.01232   -13.26  .0000    -15.4059  -11.4377
    LNQ2|     .62859***      .01404    44.79  .0000      .60108    .65610
    LNQ3|     .09670***      .00669    14.46  .0000      .08359    .10981
    LPLE|     .68419***      .07646     8.95  .0000      .53433    .83405
    LPKE|     .39946***      .03301    12.10  .0000      .33476    .46415
        |Mean of underlying truncated distribution
    RACK|     .62333***      .05632    11.07  .0000      .51293    .73372
  TUNNEL|    -.35607***      .05500    -6.47  .0000     -.46387   -.24828
        |Scale parms. for random components of e(i)
ln_sgmaU|   -2.54850***      .96756    -2.63  .0084    -4.44488   -.65212
ln_sgmaV|   -1.36799***      .06507   -21.02  .0000    -1.49551  -1.24046
        |Heteroscedasticity in variance of truncated u(i)
  VIRAGE|   -1.47329        2.86559     -.51  .6072    -7.08975   4.14316
        |Heteroscedasticity in variance of symmetric v(i)
  VIRAGE|     .06774         .08094      .84  .4026     -.09090    .22638
--------+--------------------------------------------------------------------
Note: ***, **, * ==>  Significance at 1%, 5%, 10% level.
-----------------------------------------------------------------------------
```

This is the same model as immediately above, with the additional assumption that the inefficiency is time invariant. Compared to the previous specification, σ_u has now increased by a factor of 30 while σ_v has nearly vanished, falling from 0.27 to 0.005, that is, by a factor of 50.

```
-----------------------------------------------------------------------------
Limited Dependent Variable Model - FRONTIER
Dependent variable                    LNC
Log likelihood function        532.94237
Estimation based on N =    604, K =  11
Inf.Cr.AIC  =  -1043.9 AIC/N =   -1.728
Variances: Sigma-squared(v)=      .00003
           Sigma-squared(u)=      .76238
           Sigma(u)        =      .87314
           Sigma(v)        =      .00543
Sigma = Sqr[(s^2(u)+s^2(v)]=      .87316
Variances averaged over observations
Stochastic frontier based on panel data
Estimation based on      49 individuals
LR test for inefficiency vs. OLS v only
Deg. freedom for sigma-squared(u):    1
Deg. freedom for heteroscedasticity:  1
Deg. freedom for truncation mean:     2
Deg. freedom for inefficiency model:  4
LogL when sigma(u)=0          -210.45352
Chi-sq=2*[LogL(SF)-LogL(LS)] = 1486.792
Kodde-Palm C*: 95%: 8.761,  99%: 12.483
--------+--------------------------------------------------------------------
        |                  Standard            Prob.      95% Confidence
     LNC|  Coefficient       Error       z    |z|>Z*         Interval
--------+--------------------------------------------------------------------
        |Deterministic Component of Stochastic Frontier Model
Constant|   -7.26117***      .25317   -28.68  .0000    -7.75738  -6.76496
    LNQ2|     .36162***      .01558    23.20  .0000      .33107    .39216
    LNQ3|     .01947***      .00257     7.58  .0000      .01444    .02451
    LPLE|     .64342***      .02165    29.72  .0000      .60099    .68584
    LPKE|     .30730***      .00727    42.24  .0000      .29305    .32156
        |Mean of underlying truncated distribution
    RACK|     .81356         .52427     1.55  .1207     -.21399   1.84112
  TUNNEL|    1.46353***      .47072     3.11  .0019      .54094   2.38613
        |Scale parms. for random components of e(i)
ln_sgmaU|    -.17921         .21781     -.82  .4106     -.60611    .24769
ln_sgmaV|   -4.94678***      .20426   -24.22  .0000    -5.34711  -4.54644
        |Heteroscedasticity in variance of truncated u(i)
  VIRAGE|     .06076         .04703     1.29  .1964     -.03142    .15294
        |Heteroscedasticity in variance of symmetric v(i)
  VIRAGE|    -.37544         .44206     -.85  .3957    -1.24185    .49097
--------+--------------------------------------------------------------------
Note: ***, **, * ==>  Significance at 1%, 5%, 10% level.
-----------------------------------------------------------------------------
```

The kernel estimator compares the estimated cost efficiency distributions for the pooled and basic Pitt and Lee model. The pattern suggested earlier is clearly evident. The same comparison appears for the truncated normal/heteroscedasticity models. (The estimated cost efficiency results for the basic Pitt and Lee model and the expanded one are the same to three or four digits.) The partial listing below shows the estimates for the four models, noting the time invariance of the Pitt and Lee estimates.

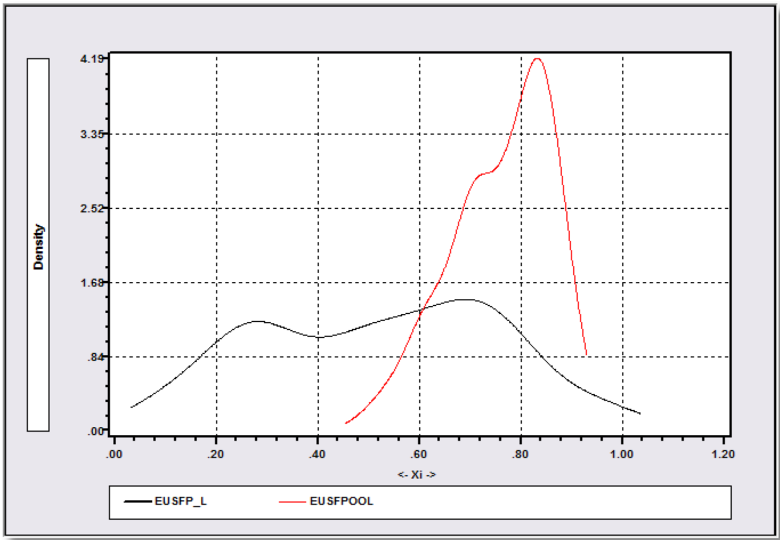

Figure E64.2 Kernel Estimators for Cost Efficiency

Data Editor

56/900 Vars; 33333 Rows: 604 Obs Cell: 0.57818

	EUSFPOOL	EUSFP_L	EUSHET_T	FULLMODL
1 »	0.673518	0.913409	0.820174	0.913804
2 »	0.841946	0.913409	0.877011	0.913804
3 »	0.825565	0.913409	0.795131	0.913804
4 »	0.834643	0.913409	0.855689	0.913804
5 »	0.837169	0.913409	0.867229	0.913804
6 »	0.829983	0.913409	0.852395	0.913804
7 »	0.811368	0.913409	0.770535	0.913804
8 »	0.74011	0.626534	0.675233	0.6258
9 »	0.770612	0.626534	0.711839	0.6258
10 »	0.775549	0.626534	0.720038	0.6258
11 »	0.779228	0.626534	0.725343	0.6258
12 »	0.773121	0.626534	0.718287	0.6258
13 »	0.793122	0.626534	0.746099	0.6258
14 »	0.789952	0.626534	0.740968	0.6258
15 »	0.782502	0.626534	0.73035	0.6258
16 »	0.817268	0.626534	0.785789	0.6258
17 »	0.820948	0.626534	0.791773	0.6258
18 »	0.810805	0.626534	0.771794	0.6258
19 »	0.820409	0.626534	0.790504	0.6258
20 »	0.815996	0.626534	0.785679	0.6258
21 »	[illegible]	[illegible]	[illegible]	[illegible]

Figure E64.3 Estimated Cost Efficiency

E64.3.3 Technical Details

For the three forms of the normal mixture models, we use the following: Let

$\gamma \quad = \sigma_u^2 / \sigma_v^2$

$\tau_i \quad = \mu_i/\sigma_u$

$\mu_i \quad = \boldsymbol{\theta}'\mathbf{z}_i$ for the heterogeneous mean model

$\mu, \quad =$ a constant (0) for the simple truncated (half) normal model

$A_i \quad = 1 + \gamma T_i$

$h_i \quad = \tau_i / A_i - S\gamma T_i \bar{\varepsilon}_i /(\sigma_u A_i)$

$\bar{\varepsilon}_i \quad = (1/T_i)\sum_{t=1}^{T_i}(y_{it} - \boldsymbol{\beta}' x_{it})$.

Then, the contribution of individual i to the log likelihood function for the normal-half normal model is

$$\log L_i = -(Ti/2)\log 2\pi - T_i \log\sigma_u - ½ \log A_i - (T_i/2) \log \gamma$$
$$- ½(\gamma / \sigma_u^2)\sum_{t=1}^{T_i} \varepsilon_{it}^2 + ½ A_i h_i^2 + ½ \log\Phi\left(h_i\sqrt{A_i}\right) - ½ \tau_i^2 - \log\Phi(\tau_i)$$

For the normal-exponential model, let

$h_i \quad = -(\theta\sigma_v/T_i + d\,\bar{\varepsilon}_i/\sigma_v)$

Then, $\log L_i = -½ \log T_i - (T_i - 1)\log 2\pi + \log\theta - (T_i - 1)\log\sigma_v$

$$- ½(1/\sigma_v^2) \sum_{t=1}^{T_i} \varepsilon_{it}^2 + ½ T_i h_i^2 + \log\Phi\left(h_i\sqrt{T_i}\right)$$

The Jondrow estimator, as formulated in Battese and Coelli (1988) in as follows: Let

$\gamma_i \quad = 1 / (1 + \lambda^2 T_i)$,

$\psi_i^2 \quad = \sigma_u^2\gamma_i$,

$E_i \quad = \gamma_i\mu + (1 - \gamma_i)(-\bar{\varepsilon}_i)$,

and $\bar{\varepsilon}_i \quad = (1/\mathrm{T_i})\Sigma_t\varepsilon_{it}$.

Then, $E[u_i|\varepsilon_{i1},\varepsilon_{i2},...] = E_i + \psi_i[\phi(E_i/\psi_i) / \Phi(E_i/\psi_i)]$.

For the exponential model, replace ψ_i with σ_v and E_i with $\sqrt{T_i}\,(-\bar{\varepsilon}_i - \theta\sigma_v^2/T_i)$.

E64.4 Cornwell, Schmidt and Sickles – Time Invariant Inefficiency, Fixed Effects

Cornwell, Schmidt and Sickles (1990) suggested a modification of the familiar fixed effects linear regression,

$$y_{it} = \alpha_i + \boldsymbol{\beta}'\mathbf{x}_{it} + v_{it}.$$

The estimated model is

$$\begin{aligned} y_{it} &= a_i + \mathbf{b}'\mathbf{x}_{it} + v_{it} \\ &= \max(a_i) + \mathbf{b}'\mathbf{x}_{it} + v_{it} + [a_i - \max(a_i)] \\ &= a + \mathbf{b}'\mathbf{x}_{it} + v_{it} - u_i \end{aligned}$$

where $u_i = \max(a_i) - a_i > 0.$

(To change this to a cost frontier, change u_i to $[a_i - \min(a_i)]$ This bears resemblance to a stochastic frontier model, though in fact, it is a 'deterministic' frontier model. The signature feature is that u_i equals zero for the 'most efficient' firm in the sample. A natural interpretation of this is that what we measure with the model is not the absolute inefficiency, but inefficiency of firm *i* relative to the other firms in the sample. From the modeler's point of view, this approach has several substantive advantages and disadvantages: The main advantage is

- It is distribution free. It requires only the assumptions of the linear model.

The disadvantages are:

- It does not allow any time invariant variables in the model.
- It labels as inefficiency any and all omitted time invariant effects.
- It can only measure firms relative to each other.

As illustrated in the results below, this approach tends to produce very large estimates of u_i. The invariance assumption about u_i has been criticized elsewhere. Attempts to relax this assumption are a recurrent theme in the literature, including the Battese and Coelli and true fixed and random effects approaches described later. Other early work on the model suggested direct manipulation of the fixed effects, for example,

$$\alpha_{it} = \theta_{i0} + \theta_{i1}t + \theta_{i2}t^2.$$

Other more recent research (Han, Orea and Schmidt (2005)) has proposed factor analytic forms for α_{it}. The sections to follow will include several of these different approaches.

Application

This Cornwell, Schmidt and Sickles (CSS) approach requires only a linear fixed effects regression and a few instructions to manipulate the fixed effects. The following analyzes the airline data with this approach. The following computes the CSS estimates and compares them to the unstructured pooled estimates (using the normal-half normal model from Chapter E62) and the Pitt and Lee model introduced above. The commands for the analysis are as follows:

```
SAMPLE     ; All $
CREATE     ; Railroad = id $
CREATE     ; If(railroad > 20)railroad = railroad - 1 $ (There is a gap in the data)
HISTOGRAM  ; Rhs = railroad
           ; Title = Number of Observations for Firms in Swiss Railroad Sample $
SETPANEL   ; Group = id ; Pds = ti $
REJECT     ; ti = 1 $
FRONTIER   ; Lhs = lnc ; Cost ; Rhs = x ; Costeff = eusfpool $
CREATE     ; pooled = Group Mean(eusfpool, Pds = ti) $
FRONTIER   ; Lhs = lnc ; Cost ; Rhs = x ; Panel ; Costeff = pittlee $
REGRESS    ; Lhs = lnc ; Rhs = x ; Panel ; Fixed Effects $
CREATE     ; ai = alphafe(railroad) $
CALC       ; minai = Min(ai) $
CREATE     ; css = Exp((minai - ai)) $
CREATE     ; Period = Ndx(id,1) $
REJECT     ; period#1 $
PLOT       ; Lhs = railroad ; Rhs = pooled,css ; Grid ; Fill ; Limits = 0,1
           ; Vaxis = Estimated Cost Efficiency
           ; Title = Half Normal vs. Cornwell, Schmidt, Sickles FE Cost Efficiencies $
PLOT       ; Lhs = railroad ; Rhs = css,pittlee ; Grid ; Fill ; Limits = 0,1
           ; Vaxis = Estimated Cost Efficiency
           ; Title = Pitt and Lee RE vs. Cornwell, Schmidt, Sickles FE Cost Efficiencies $
```

The results below show the considerable differences in the parameter estimates produced by the three models. Figure E64.4 demonstrates the expected quite large differences between the time varying estimates (using the group means) and the time invariant results based on the CSS model. Figure E64.5 also shows a striking, albeit commonly observed result – the CSS and Pitt and Lee estimates are virtually identical.

```
-----------------------------------------------------------------------------
LSDV         least squares with fixed effects ....
LHS=LNC      Mean                 =       11.30305
             Standard deviation   =        1.09984
             No. of observations  =            604  Degrees of freedom
Regression   Sum of Squares       =        726.000           52
Residual     Sum of Squares       =        3.41179          551
Total        Sum of Squares       =        729.412          603
             Standard error of e  =         .07869
Fit          R-squared            =         .99532  R-bar squared =   .99488
Model test   F[ 52,    551]       =     2254.77325  Prob F > F*   =   .00000
Diagnostic   Log likelihood       =      706.21504  Akaike I.C.   = -5.00084
             Restricted (b=0)     =     -914.01557  Bayes  I.C.   = -4.61443
             Chi squared [ 52]    =     3240.46122  Prob C2 > C2* =   .00000
Estd. Autocorrelation of e(i,t)   =        .668792
--------------------------------------------------------
Panel:Groups Empty       0,      Valid data       49
             Smallest    3,      Largest          13
             Average group size in panel       12.33
Variances    Effects a(i)          Residuals e(i,t)
              .423441                     .006192
--------+--------------------------------------------------------------------
        |                  Standard            Prob.      95% Confidence
     LNC|  Coefficient       Error       z    |z|>Z*          Interval
--------+--------------------------------------------------------------------
    LNQ2|      .29374***      .02850    10.31  .0000      .23789    .34959
    LNQ3|      .01612***      .00543     2.97  .0030      .00547    .02676
    LPLE|      .66452***      .03580    18.56  .0000      .59434    .73469
    LPKE|      .31777***      .01863    17.05  .0000      .28125    .35430
--------+--------------------------------------------------------------------
```

(These are the estimated parameters in the estimated pooled stochastic frontier model.)

```
Constant|   -10.0907***     1.14284    -8.83  .0000    -12.3306   -7.8507
    LNQ2|      .64179***      .01371    46.80  .0000      .61491    .66867
    LNQ3|      .06855***      .00655    10.46  .0000      .05570    .08139
    LPLE|      .53971***      .08858     6.09  .0000      .36610    .71333
    LPKE|      .26045***      .03260     7.99  .0000      .19655    .32435
        |Variance parameters for compound error
  Lambda|     1.29697***      .13854     9.36  .0000     1.02545   1.56850
   Sigma|      .44345***      .00056   789.05  .0000      .44235    .44455
```

(These are the estimated parameters in the estimated Pitt and Lee model.)

```
        |Deterministic Component of Stochastic Frontier Model
Constant|    -7.25643***      .24767   -29.30  .0000    -7.74185  -6.77101
    LNQ2|      .36259***      .01503    24.12  .0000      .33312    .39205
    LNQ3|      .01902***      .00240     7.94  .0000      .01432    .02372
    LPLE|      .64148***      .02112    30.38  .0000      .60009    .68287
    LPKE|      .30842***      .00700    44.08  .0000      .29471    .32214
        |Variance parameters for compound error
  Lambda|     12.1932**     5.55909     2.19  .0283      1.2975   23.0888
Sigma(u)|      .96071***      .13303     7.22  .0000      .69998   1.22145
```

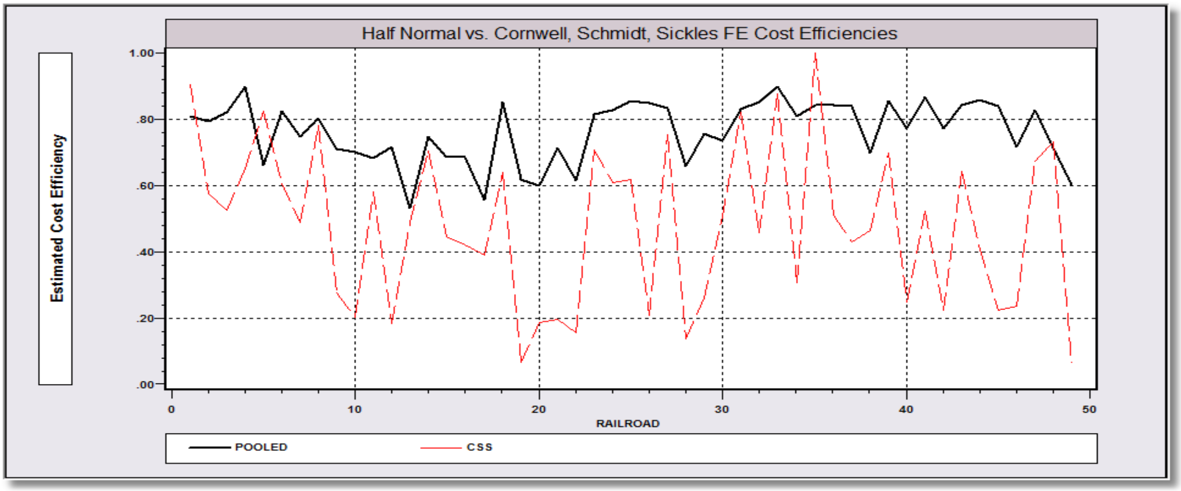

Figure E64.4 Cornwell et al. Estimates vs. Normal Half Normal

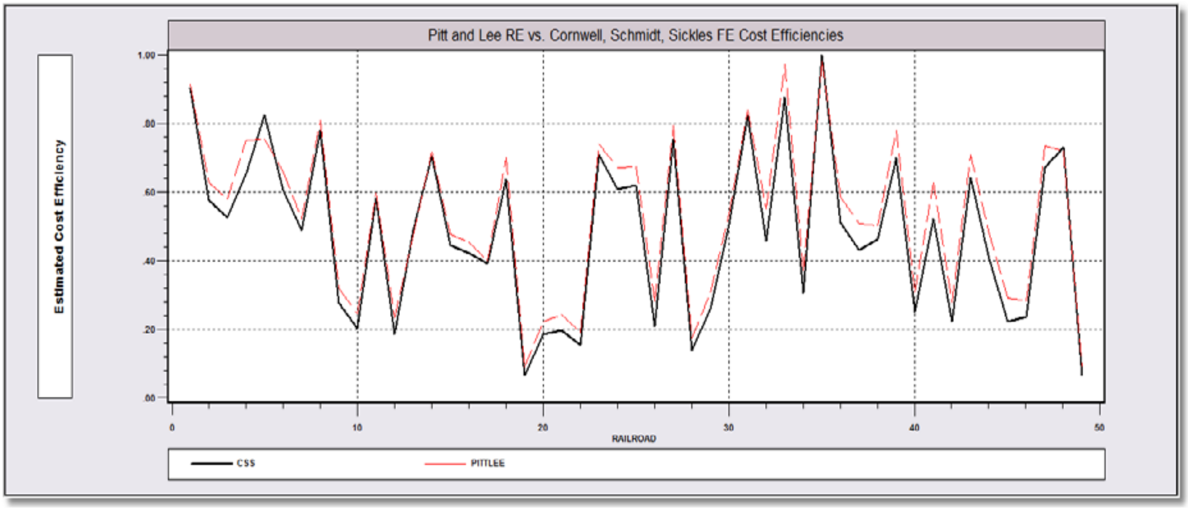

Figure E64.5 Estimated Inefficiencies from Cornwell et al. and Pitt and Lee Models

E64.5 Battese and Coelli – Time Dependent Inefficiency Models

Battese and Coelli (1992) proposed a series of models that can be collected in the general form

$$y_{it} = \boldsymbol{\beta}'\mathbf{x}_{it} + v_{it} - u_{it}$$

$$u_{it} = g(\mathbf{z}_{it})\,|U_i| \text{ where } U_i \text{ is half normal or truncated normal.}$$

Several formulations are available. In Battese and Coelli's original formulation, the distribution was half normal and the base specification was

$$g(z_{it}) = \exp[-\eta(t - T)]$$

where T is the number of periods in their balanced panel. (Here it would be T_i.) They also suggested

$$g(z_{it}) = \exp[-\eta_1(t - T) + -\eta_2(t - T)^2].$$

The first (linear) form is taken to be the default case for this model. The second is not provided in this package. The BC92 model is requested with

```
FRONTIER    ; Lhs = ... ; Rhs = one,...
            ; Model = BC
            ; Panel $
```

A truncated normal version is requested by adding

; Rh2 = list of variables which may (generally should) include one

(The **; Model = T** is not needed here.)

We note a warning to practitioners. When the data are very consistent with the model, the Battese and Coelli model produces quite satisfactory results. The framework has been employed in many recent empirical applications. But, when the data are not of particularly good quality, or this is the wrong model, extreme results can emerge. The airline data examined in Chapter E63 (and the WHO data), for example, are a poor fit to this model.

We have labeled this model as 'time dependent' rather than time varying. While the inefficiency component in the model does vary through time, the variation is systematic with respect to time. A question pursued in the ongoing literature is the extent to which this model actually moves away from the time invariant specification of Pitt and Lee. Since there is actual variation, the result is clearly somewhere between Pitt and Lee and what we have labeled the unstructured 'pooled' model. If η equals zero, Pitt and Lee emerges, so it depends entirely on this parameter. We have found in some investigations that the end result is actually closer to Pitt and Lee than it is to the pooled model – that is, there is quite a lot of structure involved in the BC92 model. The example below illustrates.

E64.5.1 Application

To illustrate the Battese and Coelli models, we return to the railroad data used previously. The base case is the pooled data stochastic cost frontier. This is followed by the Pitt and Lee model and, finally, by the original Battese Coelli 'time decay' model,

$$g(\mathbf{z}_{it}) = \exp[-\eta(t - T_i)].$$

The commands are

```
SAMPLE      ; All $
REJECT      ; ti = 1 $
FRONTIER    ; Lhs = lnc ; Cost ; Rhs = x ; Costeff = eusfpool $
FRONTIER    ; Lhs = lnc ; Cost ; Rhs = x ; Model = BC ; Panel ; Costeff = eucbc92 $
DSTAT       ; Rhs = eucbc92,eusfpool $
KERNEL      ; Rhs = eucbc92,eusfpool
            ; Title = Estimated Cost Efficiencies - Battese-Coelli 1992 vs. Pooled $
KERNEL      ; Rhs = eucbc92,pittlee
            ; Title = Estimated Cost Efficiencies - Battese-Coelli 1992 vs. Pitt and Lee $
```

The kernel density estimators are used to compare the efficiency estimates from the pooled data model to the Battese and Coelli model. The estimates of $\exp(-E[u_{it}|\varepsilon_i])$ from the Battese and Coelli model are far larger than those from the pooled model. The assumption of time invariance of the random term is a major component of this model. The second kernel estimator below compares Battese-Coelli to Pitt-Lee. The correspondence of the two results is striking, albeit to be expected given the small estimated value of η.

```
-----------------------------------------------------------------------------
Limited Dependent Variable Model - FRONTIER
Dependent variable                    LNC
Log likelihood function      -209.42340
Estimation based on N =    604, K =   7
Inf.Cr.AIC  =    432.8 AIC/N =     .717
Variances: Sigma-squared(v)=     .07332
           Sigma-squared(u)=     .12333
           Sigma(v)        =     .27077
           Sigma(u)        =     .35119
Sigma = Sqr[(s^2(u)+s^2(v)]=     .44345
Gamma = sigma(u)^2/sigma^2 =     .62716
Var[u]/{Var[u]+Var[v]}     =     .37937
Stochastic Cost Frontier Model, e = v+u
LR test for inefficiency vs. OLS v only
Deg. freedom for sigma-squared(u):    1
Deg. freedom for heteroscedasticity:  0
Deg. freedom for truncation mean:     0
Deg. freedom for inefficiency model:  1
LogL when sigma(u)=0         -210.45352
Chi-sq=2*[LogL(SF)-LogL(LS)] =    2.060
Kodde-Palm C*: 95%: 2.706,   99%:  5.412
```

```
--------+-------------------------------------------------------------
        |                  Standard            Prob.      95% Confidence
     LNC|  Coefficient       Error       z    |z|>Z*         Interval
--------+-------------------------------------------------------------
        |Deterministic Component of Stochastic Frontier Model
Constant|   -10.0907***     1.14284    -8.83  .0000    -12.3306   -7.8507
    LNQ2|      .64179***     .01371    46.80  .0000      .61491    .66867
    LNQ3|      .06855***     .00655    10.46  .0000      .05570    .08139
    LPLE|      .53971***     .08858     6.09  .0000      .36610    .71333
    LPKE|      .26045***     .03260     7.99  .0000      .19655    .32435
        |Variance parameters for compound error
  Lambda|     1.29697***     .13854     9.36  .0000     1.02545   1.56850
   Sigma|      .44345***     .00056   789.05  .0000      .44235    .44455
--------+-------------------------------------------------------------

----------------------------------------------------------------------
Limited Dependent Variable Model - FRONTIER
Dependent variable                  LNC
Log likelihood function       530.16177
Estimation based on N =    604, K =   8
Inf.Cr.AIC  =  -1044.3 AIC/N =   -1.729
Stochastic frontier based on panel data
Estimation based on      49 individuals
Variances: Sigma-squared(v)=     .00613
           Sigma-squared(u)=     .97581
           Sigma(v)        =     .07828
           Sigma(u)        =     .98783
Sigma = Sqr[(s^2(u)+s^2(v)]=     .99093
Gamma = sigma(u)^2/sigma^2 =     .99376
Var[u]/{Var[u]+Var[v]}     =     .98301
Stochastic Cost Frontier Model, e = v+u
Battese-Coelli Models: Time Varying uit
Time dependent uit=exp[-eta(t-T)]*|U(i)|
LR test for inefficiency vs. OLS v only
Deg. freedom for sigma-squared(u):    1
Deg. freedom for heteroscedasticity:  0
Deg. freedom for truncation mean:     0
Deg. freedom for inefficiency model:  1
LogL when sigma(u)=0          -210.45352
Chi-sq=2*[LogL(SF)-LogL(LS)] = 1481.231
Kodde-Palm C*: 95%: 2.706,  99%:  5.412
--------+-------------------------------------------------------------
        |                  Standard            Prob.      95% Confidence
     LNC|  Coefficient       Error       z    |z|>Z*         Interval
--------+-------------------------------------------------------------
        |Deterministic Component of Stochastic Frontier Model
Constant|    -6.83502***     .27362   -24.98  .0000    -7.37130  -6.29873
    LNQ2|      .35459***     .01636    21.68  .0000      .32254    .38665
    LNQ3|      .02183***     .00238     9.17  .0000      .01716    .02649
    LPLE|      .61516***     .02092    29.40  .0000      .57415    .65617
    LPKE|      .30931***     .00701    44.09  .0000      .29556    .32306
        |Variance parameters for compound error
  Lambda|     12.6195***     .01188  1062.18  .0000     12.5962   12.6428
Sigma(u)|      .98783***     .15275     6.47  .0000      .68845   1.28721
        |Eta parameter for time varying inefficiency
     Eta|     -.00248***     .00086    -2.89  .0039     -.00416   -.00080
--------+-------------------------------------------------------------
```

Variable	Mean	Std.Dev.	Minimum	Maximum	Cases	Missing
EUCBC92	.514566	.231680	.085140	.982112	604	0
EUSFPOOL	.760991	.095229	.478178	.906348	604	0

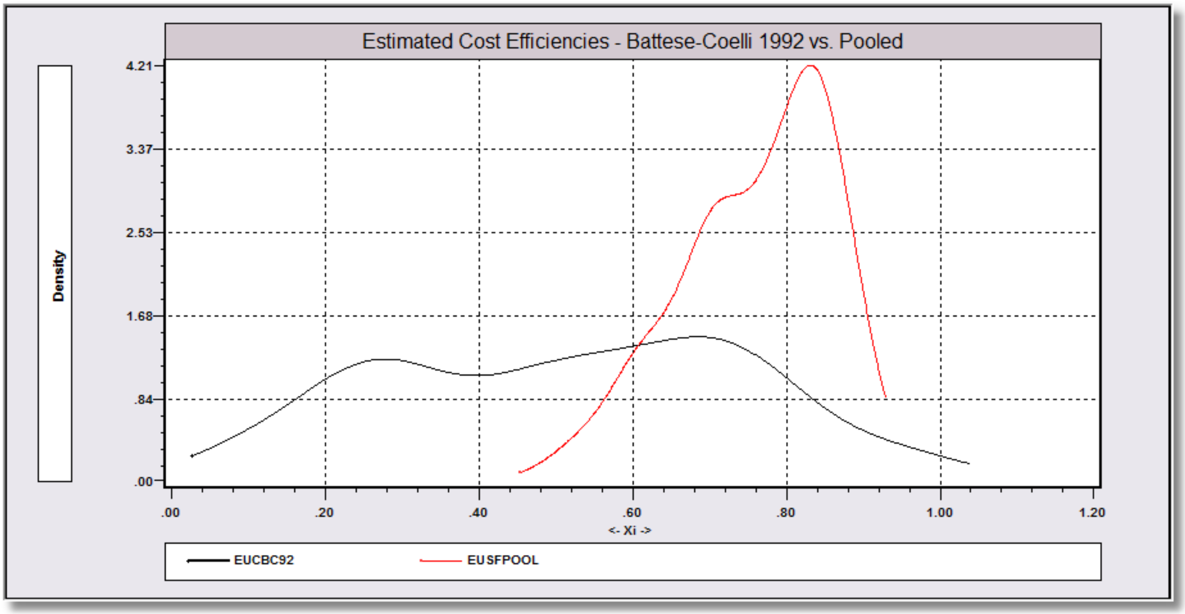

Figure E64.6 Kernel Density Estimates for Inefficiencies from Battese and Coelli Model

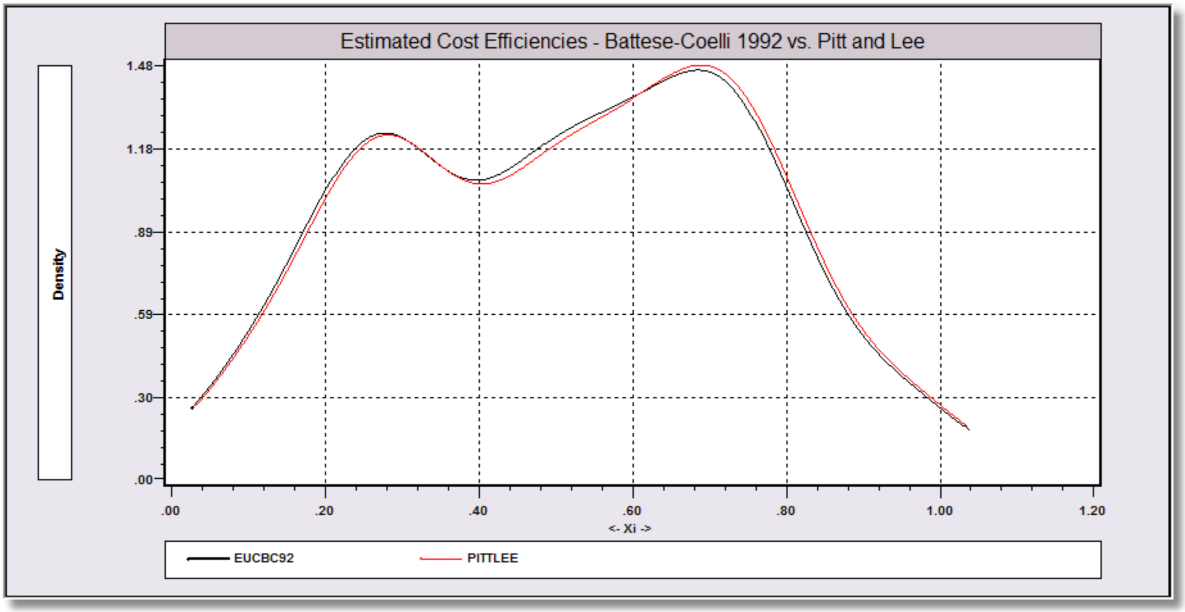

Figure E64.7 Kernel Density Estimates for Inefficiencies

E64.5.2 Technical Details

To form the log likelihood function for the model, we use Battese and Coelli's parameterization of the model. The contribution of the *i*th individual (firm, group, etc.) to the log likelihood is

$$\log L_i = -\frac{T_i}{2}(\log 2\pi + \log\sigma^2) - \frac{(T_i - 1)\log(1-\gamma)}{2} - \frac{1}{2}\sum_{t=1}^{T_i}\frac{\varepsilon_{it}^2}{(1-\gamma)\sigma^2}$$
$$-\frac{1}{2}\log\left[1+\gamma\left(\left(\sum_{t=1}^{T_i} g_{it}^2\right)-1\right)\right]$$
$$-\frac{1}{2}\left(\frac{\mu_i}{\sigma\sqrt{\gamma}}\right)^2 - \log\Phi\left(\frac{\mu_i}{\sigma\sqrt{\gamma}}\right) + \frac{A_i^2}{2} + \log\Phi(A_i)$$

$$\sigma^2 = \sigma_u^2 + \sigma_v^2$$
$$\gamma = \sigma_u^2 / \sigma^2$$
$$\varepsilon_{it} = y_{it} - \boldsymbol{\beta}'\mathbf{x}_{it}$$
$$\mu_i = 0 \text{ or } \mu \text{ or } \boldsymbol{\delta}'\mathbf{w}_i$$
$$g_{it} = \exp[-\eta(t - T_i)] \text{ or } \exp(\boldsymbol{\eta}'\mathbf{z}_{it})$$
$$S = +1 \text{ for a production model and -1 for a cost model}$$
$$A_i = \frac{(1-\gamma)\mu_i - \gamma S \Sigma_{t=1}^{T_i} g_{it}\varepsilon_{it}}{\sqrt{\gamma(1-\gamma)\left[1+\gamma\left(\left(\Sigma_{t=1}^{T_i} g_{it}^2\right)-1\right)\right]}}$$

Derivatives of this function are complicated in the extreme, and are omitted here. (Some useful results for obtaining them are found in Battese and Coelli (1992, 1995).)

The Jondrow et al. (1982) estimator of u_{it} is

$$E[u_{it} \mid \varepsilon_{i1},\varepsilon_{i2},\ldots] = g_{it}\, E[u_i \mid \varepsilon_{i1},\varepsilon_{i2},\ldots]$$
$$= g_{it}\left[\tilde{\mu}_i + \tilde{\sigma}_i\left(\frac{\phi(\tilde{\mu}_i/\tilde{\sigma}_i)}{\Phi(\tilde{\mu}_i/\tilde{\sigma}_i)}\right)\right]$$

where

$$\tilde{\mu}_i = \frac{(1-\gamma)\mu_i - \gamma\Sigma_{t=1}^{T_i} g_{it}(S\varepsilon_{it})}{(1-\gamma)+\gamma\Sigma_{t=1}^{T_i} g_{it}^2}$$

$$\tilde{\sigma}_i^2 = \frac{\gamma(1-\gamma)\sigma^2}{(1-\gamma)+\gamma\Sigma_{t=1}^{T_i} g_{it}^2}$$

E64.6 Time Varying Inefficiency in the Battese Coelli Model

The general form of the Battese and Coelli model is,

$$y_{it} = \boldsymbol{\beta}'\mathbf{x}_{it} + v_{it} - u_{it}$$

$$u_{it} = g(\mathbf{z}_{it})\,|U_i| \text{ where } U_i \text{ is half normal or truncated normal.}$$

The default form used earlier is $g(\mathbf{z}_{it}) = \exp[-\eta(t - T_i)]$. You may also use a more general form,

$$g(\mathbf{z}_{it}) = \exp(\boldsymbol{\eta}'\mathbf{z}_{it})$$

where $\mathbf{z}_{it}$ contains any desired set of variables. For this extension, use

```
FRONTIER    ; Lhs = ... ; Rhs = one,...
            ; Model = BC ; Hfu = the variables in z
            ; Pds = the panel specification $
```

As before, the truncated normal version of the model is also supported. For an example, we have used

```
FRONTIER    ; Lhs = lnc ; Cost ; Rhs = x ; Model = BC ; Panel ; Costeff = eucbc92h
            ; Hfu = rack,virage,tunnel $
```

The estimates of cost efficiency produced by this model are identical to those from the base model in the previous section.

```
-----------------------------------------------------------------------------
Limited Dependent Variable Model - FRONTIER
Dependent variable                  LNC
Log likelihood function       529.63533
Stochastic frontier based on panel data
Estimation based on      49 individuals
Variances: Sigma-squared(v)=     .00615
           Sigma-squared(u)=     .94808
           Sigma(v)        =     .07840
           Sigma(u)        =     .97369
Sigma = Sqr[(s^2(u)+s^2(v)]=     .97685
Gamma = sigma(u)^2/sigma^2 =     .99356
Var[u]/{Var[u]+Var[v]}     =     .98247
Stochastic Cost Frontier Model, e = v+u
Battese-Coelli Models: Time Varying uit
Time varying uit=exp[eta*z(i,t)]*|U(i)|
LR test for inefficiency vs. OLS v only
Deg. freedom for sigma-squared(u):    1
Deg. freedom for heteroscedasticity:  3
Deg. freedom for truncation mean:     0
Deg. freedom for inefficiency model:  4
LogL when sigma(u)=0          -210.45352
Chi-sq=2*[LogL(SF)-LogL(LS)] = 1480.178
Kodde-Palm C*: 95%: 8.761,  99%: 12.483
```

LNC	Coefficient	Standard Error	z	Prob. |z|>Z*	95% Confidence Interval	
	Deterministic Component of Stochastic Frontier Model					
Constant	-6.89845***	.32923	-20.95	.0000	-7.54374	-6.25316
LNQ2	.35751***	.01591	22.47	.0000	.32632	.38870
LNQ3	.02149***	.00236	9.10	.0000	.01686	.02613
LPLE	.61741***	.02430	25.40	.0000	.56977	.66504
LPKE	.30892***	.00759	40.71	.0000	.29405	.32380
	Variance parameters for compound error					
Lambda	12.4202***	.01108	1120.76	.0000	12.3984	12.4419
Sigma(u)	.97369***	.13513	7.21	.0000	.70884	1.23855
	Coefficients in u(i,t)=[exp{eta*z(i,t)}]*|U(i)|					
RACK	.00024	.01743	.01	.9889	-.03392	.03441
VIRAGE	-.02096	.01321	-1.59	.1126	-.04685	.00493
TUNNEL	.00219	.01625	.14	.8926	-.02966	.03405

(Parameter estimates from base case Battese and Coelli)

	Deterministic Component of Stochastic Frontier Model					
Constant	-6.83502***	.27362	-24.98	.0000	-7.37130	-6.29873
LNQ2	.35459***	.01636	21.68	.0000	.32254	.38665
LNQ3	.02183***	.00238	9.17	.0000	.01716	.02649
LPLE	.61516***	.02092	29.40	.0000	.57415	.65617
LPKE	.30931***	.00701	44.09	.0000	.29556	.32306
	Variance parameters for compound error					
Lambda	12.6195***	.01188	1062.18	.0000	12.5962	12.6428
Sigma(u)	.98783***	.15275	6.47	.0000	.68845	1.28721
	Eta parameter for time varying inefficiency					
Eta	-.00248***	.00086	-2.89	.0039	-.00416	-.00080

E64.7 True Fixed Effects Models

The received applications of fixed effects to the stochastic frontier model, primarily Cornwell, Schmidt and Sickles have actually been reinterpretations of the linear regression model with fixed effects, not frontier models of the sort considered here. The estimators described below apply the fixed effects to the stochastic frontier. We label these 'true fixed effects models' to distinguish them from the linear regression models as discussed in Section E64.3. (This is not meant to apply that these are 'false fixed effects models.' Had we used 'real fixed effects models,' then the contrasting 'unreal fixed effects models' would arise which is likewise problematic. We use this purely as a concise term of art, not a characterization of the types of estimators considered.)

The stochastic frontier model with fixed effects may be fit in several forms. The base case applies the heterogeneity to the normal-half normal production function model;

$$y_{it} \quad = \alpha_i + \boldsymbol{\beta}'\mathbf{x}_{it} + v_{it} - Su_{it},$$

where S = +1 for a production frontier and -1 for a cost frontier, and

$$u_i \quad = |\,\mathrm{N}[0, \sigma_u^2]\,|.$$

This model (as are the others) is fit by maximum likelihood, not least squares. The normal-half normal model is applied to the stochastic part of the model. Note that the inefficiency term in this model is time varying. The heterogeneity may appear in Stevenson's truncated normal model as follows. This is a true fixed effects, normal-truncated normal model.

$$y_{it} = \alpha_i + \boldsymbol{\beta}'\mathbf{x}_{it} + v_{it} - u_{it},$$

$$u_i = |\, N[\mu_i, \sigma_u^2] \,|$$

$$\mu_i = \boldsymbol{\delta}'\mathbf{z}_i.$$

In this form, the heterogeneity is still retained in the production function part of the model. Another possibility is to allow the heterogeneity to enter the mean of the inefficiency distribution rather than the production function – this seems the most natural of the three forms. In this case,

$$y_{it} = \boldsymbol{\beta}'\mathbf{x}_{it} + v_{it} - u_{it},$$

$$u_{it} = |\, N[\mu_{it}, \sigma_u^2] \,|$$

$$\mu_{it} = \alpha_i + \mu \text{ (nonzero) or } \boldsymbol{\delta}'\mathbf{z}_i.$$

The mean of the inefficiency distribution shifts in time, but also has a firm specific component. Finally, the heterogeneity may be shifted to the variance of the inefficiency distribution. In this form, we have

$$y_{it} = \boldsymbol{\beta}'\mathbf{x}_{it} + v_{it} - u_{it},$$

$$u_{it} = |\, N[0, \sigma_{ui}^2] \,|$$

$$\sigma_{uit}^2 = \sigma_u^2 \times \exp(\alpha_i + \boldsymbol{\delta}'\mathbf{z}_{it}).$$

The variables in the variance term may be omitted if only a groupwise heteroscedastic model is desired. Note this is a half normal model. A model with nonzero underlying mean and variation in the variance appears to be inestimable. Note that in order to secure identification, this model must have time varying inefficiency, induced by time variation in the variance.

NOTE: We have had extremely limited success with the second and third forms of the model. The likelihood function is quite volatile in the parameters of the underlying mean of the truncated distribution with the result that the estimated variance parameters λ and σ generally become negative in the early iterations and estimation must be halted. This occurs even when very good starting values are used, which suggests that estimation of this model as stated is likely to be extremely problematic in all but the most favorable of cases. An alternative approach which is simple, but can be used only with small panels (up to 100 groups), is suggested below.

In terms of implementation, we note that these forms of the models, though they are new with *LIMDEP*, have long been feasible. The panels typically used by researchers in this setting are often fairly small – our airline data for example have only 25 units and the Swiss railroad data has 49 firms. It would always have been possible to create these models simply by adding dummy variables to the familiar model. However, *LIMDEP*'s implementation of the model obviates this by using the methodology described in Chapter R23. In principle, this allows up to 100,000 firms in the data set.

Results that are kept for this model are

Matrices:	*b*	= estimate of **β**
	varb	= asymptotic covariance matrix for estimate of **β**.
	alphafe	= estimated fixed effects (if **; Par** is in the command)
Scalars:	*kreg*	= number of variables in Rhs
	nreg	= number of observations
	logl	= log likelihood function
Last Model:	*b_variables*	

The upper limit on the number of groups is 100,000.

E64.7.1 Commands for the Fixed Effects Stochastic Frontier Model

The command for fitting the normal-half normal model with fixed effects is as follows:

FRONTIER ; Lhs = ... ; Rhs = one,... $
FRONTIER ; Lhs = ... ; Rhs = one,...
; FEM ; Pds = specification $

The model must be fit twice. The first model is a pooled data model which provides the starting values for the second. The second command is identical to the first save for the addition of the panel data specification. In order to set up the initial values correctly, it is essential that your initial model include the constant term first in the Rhs list and that the second model specification be identical to the first. Other options and specifications for the fixed effects models are the same as in other applications. (See Chapter R23 for details.) The fixed effects command also contains the constant term, but this will be removed by the command processor later. See the example below for the operation of the command.

NOTE: Starting values must be provided by the first estimator. The specification **; Start = list of values** is not available for this model. You must fit both models each time you fit an FEM. The starting values are not retained after the FEM is estimated.

All fixed effects forms are estimated by maximum likelihood. You may also fit a two way fixed effects model

$$y_{it} = \alpha_i + \gamma_t + \boldsymbol{\beta}'\mathbf{x}_{it} + v_{it} - u_i, \text{ (change to } v + u \text{ for a stochastic cost frontier),}$$

$$u_i = |\,N[0, \sigma_u^2]\,|$$

where γ_t is an additional, time (period) specific effect. The time specific effect is requested by adding

; Time

to the command if the panel is balanced, and

; Time = variable name

if the panel is unbalanced.

For the unbalanced panel, we assume that overall, the sample observation period is $t = 1,2,..., T_{max}$ and that the time variable gives for the specific group, the particular values of t that apply to the observations. Thus, suppose your overall sample is five periods. The first group is three observations, periods 1, 2, 4, while the second group is four observations, 2, 3, 4, 5. Then, your panel specification would be

; Pds = Ti, for example, where *Ti* = 3, 3, 3, 4, 4, 4, 4
and **; Time = Pd**, for example, where *Pd* = 1, 2, 4, 2, 3, 4, 5.

E64.7.2 Model Specifications for Fixed Effects Stochastic Frontier Models

This is the full list of general specifications that are applicable to this model estimator.

Controlling Output from Model Commands

; Par keeps ancillary parameter σ in main results vector *b*.
; Table = name saves model results to be combined later in output tables.

Robust Asymptotic Covariance Matrices

; Covariance Matrix displays estimated asymptotic covariance matrix (normally not shown), same as **; Printvc**.

Optimization Controls for Nonlinear Optimization

; Start = list gives starting values for a nonlinear model.
; Tlg[= value] sets convergence value for gradient.
; Maxit = n sets the maximum iterations.
; Output = n requests technical output during iterations; the level '**n**' is 1, 2, 3 or 4.
; Set keeps current setting of optimization parameters as permanent.

Predictions and Residuals

; List displays a list of fitted values with the model estimates.
; Keep = name keeps fitted values as a new (or replacement) variable in data set.
; Res = name keeps residuals as a new (or replacement) variable.

Hypothesis Tests and Restrictions

; Test: spec defines a Wald test of linear restrictions.
; Wald: spec defines a Wald test of linear restrictions, same as **; Test: spec**.

E64.7.3 Application of the True Fixed Effects Model

We have fit the fixed effects model with the airline data used in the previous chapter. These are simple models that do not use the observed heterogeneity in load factor, stage length or number of points served. Additional variables which vary over time can also be included in the function. The commands employed for the example are

```
SETPANEL   ; Group = firm ; Pds = ti $
FRONTIER   ; Lhs = lq ; Rhs = one,lf,lm,le,ll,lp,lk$
FRONTIER   ; Lhs = lq ; Rhs = one,lf,lm,le,ll,lp,lk,
           ; FEM ; Panel ; Techeff = euitfe ; Par $
REGRESS    ; Lhs = lq ; Rhs = one,lf,lm,le,ll,lp,lk
           ; Panel ; Fixed Effects $
CREATE     ; ai = alphafe(firm) $
CALC       ; maxai = Max(ai) $
CREATE     ; euicss = exp(-(maxai - ai)) $
CREATE     ; meuitfe = Group Mean(euitfe, Pds = ti) $
SAMPLE     ; All $
CREATE     ; Period = Ndx(firm,1) $
PLOT       ; For[period=1] ; Lhs = firm ; Rhs = euitfe,euicss
           ; Fill ; Symbols ; Limits = 0,1 ; Grid
           ; Title = Technical Efficiency Estimates, CSS vs. True Fixed Effects
                  (Group Means)
           ; Vaxis = Estimated Technical Efficiency$
```

This command recovers the estimated fixed effects from the Cornwell et al. model. then replicates them for each year in the data set. This is used to create the plot of the two sets of estimates of u_i shown below.

```
-----------------------------------------------------------------------------
Limited Dependent Variable Model - FRONTIER
Dependent variable                    LQ
Log likelihood function        108.43918
Estimation based on N =    256, K =   9
Inf.Cr.AIC  =   -198.9 AIC/N =    -.777
Model estimated: Aug 17, 2011, 06:36:42
Variances: Sigma-squared(v)=     .01902
           Sigma-squared(u)=     .01692
           Sigma(v)        =     .13791
           Sigma(u)        =     .13007
Sigma = Sqr[(s^2(u)+s^2(v)]=     .18957
Gamma = sigma(u)^2/sigma^2 =     .47074
Var[u]/{Var[u]+Var[v]}     =     .24425
Stochastic Production Frontier, e = v-u
LR test for inefficiency vs. OLS v only
Deg. freedom for sigma-squared(u):    1
Deg. freedom for heteroscedasticity:  0
Deg. freedom for truncation mean:     0
Deg. freedom for inefficiency model:  1
LogL when sigma(u)=0           108.07431
Chi-sq=2*[LogL(SF)-LogL(LS)] =      .730
Kodde-Palm C*: 95%: 2.706,  99%:  5.412
```

LQ	Coefficient	Standard Error	z	Prob. |z|>Z*	95% Confidence Interval	
	Deterministic Component of Stochastic Frontier Model					
Constant	-2.98823***	.72136	-4.14	.0000	-4.40206	-1.57439
LF	.37257***	.07038	5.29	.0000	.23463	.51052
LM	.69910***	.07580	9.22	.0000	.55054	.84766
LE	2.09473***	.68790	3.05	.0023	.74647	3.44299
LL	-.42909***	.06315	-6.79	.0000	-.55287	-.30530
LP	.44533***	.09498	4.69	.0000	.25917	.63149
LK	-2.09806***	.76556	-2.74	.0061	-3.59853	-.59759
	Variance parameters for compound error					
Lambda	.94309***	.16870	5.59	.0000	.61244	1.27373
Sigma	.18957***	.00064	297.81	.0000	.18832	.19082

```
Normal exit from iterations. Exit status=0.
-----------------------------------------------------------------------------
FIXED EFFECTS Frontr Model
Dependent variable                        LQ
Log likelihood function         205.05799
Estimation based on N =    256, K =  33
Inf.Cr.AIC  =   -344.1 AIC/N =   -1.344
Model estimated: Aug 17, 2011, 06:36:46
Unbalanced panel has     25 individuals
Skipped    0 groups with inestimable ai
Half normal stochastic frontier
Sigma( u) (1 sided)  =           .11713
Sigma( v) (symmetric)=           .08347
```

LQ	Coefficient	Standard Error	z	Prob. |z|>Z*	95% Confidence Interval	
	Production / Cost parameters					
LF	.20090**	.09879	2.03	.0420	.00727	.39453
LM	.78173***	.07495	10.43	.0000	.63483	.92863
LE	.56626	.62357	.91	.3638	-.65591	1.78843
LL	-.16687	.11488	-1.45	.1464	-.39204	.05830
LP	.17273*	.09414	1.83	.0665	-.01177	.35724
LK	-.29167	.69055	-.42	.6728	-1.64513	1.06179
	Variance parameter for v +/- u					
Sigma	.14383***	.00045	317.51	.0000	.14294	.14472
	Asymmetry parameter, lambda					
Lambda	1.40326***	.21468	6.54	.0000	.98248	1.82403

Note: ***, **, * ==> Significance at 1%, 5%, 10% level.

```
----------------------------------------------------------------------------
LSDV          least squares with fixed effects ....
LHS=LQ        Mean                 =      -1.11237
              Standard deviation   =       1.29728
              No. of observations  =           256  Degrees of freedom
Regression    Sum of Squares       =       426.103          30
Residual      Sum of Squares       =       3.04876         225
Total         Sum of Squares       =       429.152         255
              Standard error of e  =        .11640
Fit           R-squared            =        .99290  R-bar squared =    .99195
Model test    F[ 30,    225]       =    1048.21999  Prob F > F*   =    .00000
Diagnostic    Log likelihood       =     203.84835  Akaike I.C.   = -4.18825
              Restricted (b=0)     =    -429.37729  Bayes  I.C.   = -3.75896
              Chi squared [ 30]    =    1266.45126  Prob C2 > C2* =    .00000
Estd. Autocorrelation of e(i,t)    =       .575211
--------------------------------------------------
Panel:Groups Empty       0,      Valid data       25
             Smallest    2,      Largest          15
             Average group size in panel       10.24
Variances    Effects a(i)          Residuals e(i,t)
              .030410                      .013550
--------+-------------------------------------------------------------------
        |                Standard            Prob.      95% Confidence
      LQ|  Coefficient     Error         t   |t|>T*         Interval
--------+-------------------------------------------------------------------
      LF|     .14860        .09677     1.54  .1259     -.04107     .33828
      LM|     .80497***     .07843    10.26  .0000      .65125     .95868
      LE|     .68672        .67075     1.02  .3069     -.62792    2.00136
      LL|    -.15977        .11829    -1.35  .1780     -.39162     .07208
      LP|     .16227        .09973     1.63  .1050     -.03320     .35774
      LK|    -.37897        .74689     -.51  .6123    -1.84284    1.08490
--------+-------------------------------------------------------------------
Note: ***, **, * ==>  Significance at 1%, 5%, 10% level.
----------------------------------------------------------------------------
```

Figure E64.8 plots the Jondrow et al. estimates of exp($-E[u_{it}|\varepsilon_{it}]$) from the true fixed effects model and the estimates of u_i from the Cornwell, Schmidt and Sickles model of Section E64.4 for each firm. Since the true FE estimates vary by period, we have plotted the group means. The implication of the regression based model is clear in the figure. The estimates of technical efficiency from the true FEM are generally considerably larger than those from the deterministic model.

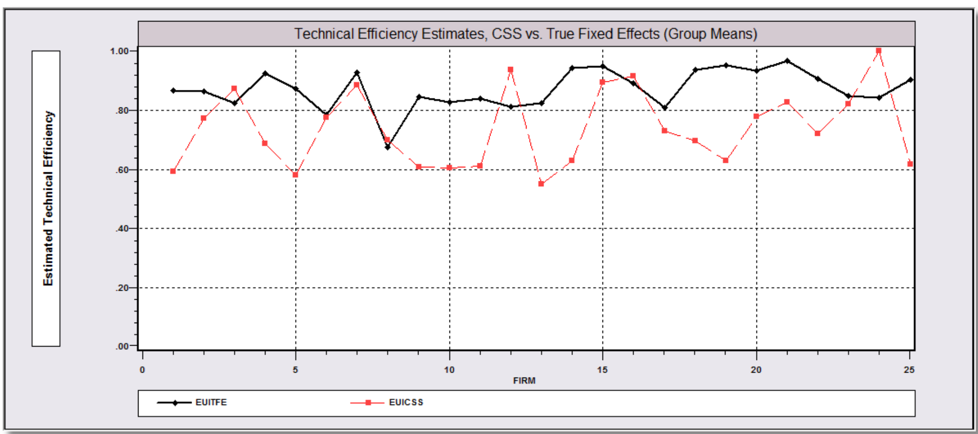

Figure E64.8 True Fixed Effects vs. Fixed Effects Estimates of u_i

E64.7.4 Fixed Effects in the Normal-Truncated Normal Model

The preceding may be extended to the truncated normal (with earlier caveats) as follows: For a model with heterogeneity appearing in the production (or cost) function,

$$y_{it} = \alpha_i + \boldsymbol{\beta}'\mathbf{x}_{it} + v_{it} - u_{it},$$

$$u_{it} = |\, N[\mu_{it}, \sigma_u^2] \,|$$

$$\mu_{it} = \mu \text{ (nonzero) or } \boldsymbol{\delta}'\mathbf{z}_{it},$$

use

```
FRONTIER    ; Lhs = ... ; Rhs = one, ... ; Rh2 = one, ...
            ; Model = T $
FRONTIER    ; Lhs = ... ; Rhs = one, ... ; Rh2 = one, ...
            ; FEM     ; Panel $
```

The Rh2 is optional in the first equation if you have only a constant term in the mean of the truncated distribution. But, you should include it nonetheless so as to insure the match between the first and second commands. Also, it is essential that both Rhs and Rh2 include constant terms in the first positions.

To move the heterogeneity to the mean of the underlying truncated normal distribution,

$$y_{it} = \boldsymbol{\beta}'\mathbf{x}_{it} + v_{it} - u_{it},$$

$$u_i = |\, N[\mu_{it}\sigma_u^2] \,|$$

$$\mu_{it} = \alpha_i + \boldsymbol{\delta}'\mathbf{z}_{it},$$

use

```
FRONTIER    ; Lhs = ... ; Rhs = one, ... ; Rh2 = one, ...
            ; Model = T $
FRONTIER    ; Lhs = ... ; Rhs = one, ... ; Rh2 = one, ...
            ; Model = T
            ; FEM ; Panel $
```

Note that this version differs from the earlier one only in the presence of **; Model = T** in the second form and its absence in the first. Again, the variable specifications in the two commands must be identical, and both must include constant terms in the first position in both lists. As before, you may use **; Rh2 = one** if you do not require variables $\mathbf{z}_{it}$ in the mean. (This constant term will be removed from the fixed effects model, but this common value is used as the starting value for the firm specific estimates.)

We note, we have had scant success with this model even with a carefully constructed data set and good starting values. The problem appears to be Newton's method, which must be used for the general fixed effects program which this is part of. If you have a small panel with no more than 100 groups, an alternative approach appears to work better. You may provide a stratification variable in the cross section template to request that a set of dummy variables be inserted directly into the function.

To fit a model of the first form above, use

```
FRONTIER    ; Lhs = ... ; Rhs = one,...
            ; Model = T [ ; Rh2 = list is optional ]
            ; Str = a variable which provides a group indicator for the panel $
```

The stratification variable must take the full set of values from 1 to *N* up to 100 and all groups must have at least two observations. For the second form, with the heterogeneity embedded in the mean of the truncated normal distribution, add

```
            ; Mean
```

to the command.

This provides four possible forms of the model, which we illustrate with the airline data:

```
NAMELIST    ; x = one,lf,lm,le,ll,lp,lk $
```

This is a true fixed effects model with normal-truncated normal structure for u_{it}.

```
FRONTIER    ; Lhs = lq ; Rhs = x
            ; Model = T
            ; Str = firm $
```

This model is the same as the preceding one except now $\mu_i = \delta_1 + \delta_2 loadfctr_i$.

```
FRONTIER    ; Lhs = lq ; Rhs = x
            ; Model = T
            ; Rh2 = one,loadfctr
            ; Str = firm $
```

This is a true fixed effects model with the fixed effects appearing in μ_i rather than in the production function.

```
FRONTIER    ; Lhs = lq ; Rhs = x
            ; Model = T
            ; Mean
            ; Str = firm $
```

This model is the same as the preceding model except that *loadfctr* now also appears in the mean of the truncated variable.

```
FRONTIER    ; Lhs = lq ; Rhs = x
            ; Model = T
            ; Rh2 = one,loadfctr ; Mean
            ; Str = firm $
```

E64.7.5 Fixed Effects in the Heteroscedasticity Model

The firmwise heteroscedasticity model,

$$y_{it} = \boldsymbol{\beta}'\mathbf{x}_{it} + v_{it} - u_{it},$$

$$u_{it} = |\,N[0, \sigma_{uit}^2]\,|$$

$$\sigma_{uit}^2 = \sigma_u^2 \times \exp(\alpha_i + \boldsymbol{\delta}'\mathbf{z}_{it})$$

is requested in the same fashion as the normal-truncated normal model, using a stratification variable in the cross section formulation. (This likelihood function is likewise quite ill behaved, though less so than the truncation form.) The command is

```
FRONTIER    ; Lhs = ... ; Rhs = one, ...
            ; Het
            ; Hfu = list of variables ; Hfv = one
            ; Str = stratification variable $
```

This model also allows for the doubly heteroscedastic form,

$$y_{it} = \boldsymbol{\beta}'\mathbf{x}_{it} + v_{it} - u_{it},$$

$$u_{it} = |\,N[0, \sigma_{uit}^2]\,|$$

$$\sigma_{uit}^2 = \sigma_u^2 \times \exp(\alpha_i + \boldsymbol{\delta}'\mathbf{z}_{it})$$

$$v_{it} \sim N[0,\sigma_{vit}^2]$$

$$\sigma_{vit}^2 = \sigma_v^2 \times \exp(\boldsymbol{\gamma}'\mathbf{w}_{it})$$

The command would be

```
FRONTIER    ; Lhs = ... ; Rhs = one, ...
            ; Het
            ; Hfu = list of variables ; Hfv = list of variables
            ; Str = stratification variable $
```

To continue the earlier example, the following fits a model of heteroscedasticity to the airline data. The first model has heteroscedasticity and the fixed effects in the variance of u_i. The second is doubly heteroscedastic, again with the fixed effects in the variance of u_i.

```
NAMELIST    ; x = one,lf,lm,le,ll,lp,lk $
FRONTIER    ; Lhs = lq ; Rhs = x
            ; Het ; Hfu = one,loadfctr ; Hfv = one ; Str = firm $
FRONTIER    ; Lhs = lq ; Rhs = x
            ; Het ; Hfu = one,loadfctr ; Hfv = one,loadfctr ; Str = firm $
```

```
--------------------------------------------------------------------------------
Limited Dependent Variable Model - FRONTIER
Dependent variable                    LQ
Log likelihood function       182.50025
Variances: Sigma-squared(v)=     .00876
           Sigma-squared(u)=     .04920
           Sigma(v)        =     .09357
           Sigma(u)        =     .22182
Sigma = Sqr[(s^2(u)+s^2(v)]=     .24075
Gamma = sigma(u)^2/sigma^2 =     .84892
Var[u]/{Var[u]+Var[v]}     =     .67126
Variances averaged over observations
Stochastic Production Frontier, e = v-u
Stratified by FIRM    ,  25 groups
--------+-----------------------------------------------------------------------
        |                Standard            Prob.      95% Confidence
      LQ|  Coefficient     Error       z    |z|>Z*         Interval
--------+-----------------------------------------------------------------------
        |Deterministic Component of Stochastic Frontier Model
Constant|   -3.70847***      .75902    -4.89  .0000    -5.19612  -2.22081
      LF|     .38142***      .08642     4.41  .0000      .21204    .55079
      LM|     .57659***      .09175     6.28  .0000      .39676    .75642
      LE|    2.78934***      .72692     3.84  .0001     1.36459   4.21408
      LL|    -.41646***      .08641    -4.82  .0000     -.58582   -.24710
      LP|     .59190***      .11704     5.06  .0000      .36251    .82129
      LK|   -2.87861***      .80566    -3.57  .0004    -4.45767  -1.29956
        |Parameters in variance of v (symmetric)
Constant|   -4.73798***      .21921   -21.61  .0000    -5.16764  -4.30833
        |Parameters in variance of u (one sided)
Constant|    8.11346        7.80244     1.04  .2984    -7.17903  23.40596
LOADFCTR|   -23.6678***     6.88328    -3.44  .0006    -37.1588  -10.1768
 FIRM001|    1.35540        7.37739      .18  .8542   -13.10403  15.81482
 FIRM002|     .25791        7.25149      .04  .9716   -13.95476  14.47057
 FIRM003|     .68176        7.22190      .09  .9248   -13.47290  14.83643
(Firms 4-20 omitted)
 FIRM021|     .73089        7.21226      .10  .9193   -13.40488  14.86666
 FIRM022|    -.38963        7.46091     -.05  .9584   -15.01274  14.23347
 FIRM023|    -.63171        7.53984     -.08  .9332   -15.40952  14.14610
 FIRM024|   -7.77451       41.07339     -.19  .8499   -88.27688  72.72786
--------+-----------------------------------------------------------------------
Note: nnnnn.D-xx or D+xx => multiply by 10 to -xx or +xx.
Note: ***, **, * ==>  Significance at 1%, 5%, 10% level.
--------------------------------------------------------------------------------
```

```
-----------------------------------------------------------------------------
Limited Dependent Variable Model - FRONTIER
Dependent variable                      LQ
Log likelihood function       190.29998
Estimation based on N =    256, K =  35
Inf.Cr.AIC  =   -310.6 AIC/N =   -1.213
Model estimated: Aug 22, 2011, 22:57:54
Variances: Sigma-squared(v)=     .00906
           Sigma-squared(u)=     .04124
           Sigma(v)        =     .09519
           Sigma(u)        =     .20307
Sigma = Sqr[(s^2(u)+s^2(v)]=     .22427
Gamma = sigma(u)^2/sigma^2 =     .81986
Var[u]/{Var[u]+Var[v]}     =     .62318
Variances averaged over observations
Stochastic Production Frontier, e = v-u
Stratified by FIRM    ,  25 groups
--------+--------------------------------------------------------------------
        |                  Standard            Prob.      95% Confidence
      LQ|  Coefficient       Error       z    |z|>Z*         Interval
--------+--------------------------------------------------------------------
        |Deterministic Component of Stochastic Frontier Model
Constant|   -3.00340***      .65319    -4.60  .0000    -4.28364  -1.72316
      LF|     .24071***      .07721     3.12  .0018      .08938    .39204
      LM|     .60992***      .07600     8.03  .0000      .46096    .75887
      LE|    2.19046***      .62677     3.49  .0005      .96202   3.41890
      LL|    -.38679***      .07314    -5.29  .0000     -.53015   -.24344
      LP|     .49345***      .09820     5.03  .0000      .30098    .68591
      LK|   -2.09638***      .69385    -3.02  .0025    -3.45631   -.73646
        |Parameters in variance of v (symmetric)
Constant|   -13.5487***     2.64897    -5.11  .0000    -18.7406   -8.3569
LOADFCTR|    15.5221***     4.48367     3.46  .0005      6.7343   24.3099
        |Parameters in variance of u (one sided)
Constant|    8.01865        5.60084     1.43  .1522    -2.95879  18.99609
LOADFCTR|   -23.3031***     6.88508    -3.38  .0007    -36.7976   -9.8086
 FIRM001|     .88200        5.06220      .17  .8617    -9.03972  10.80373
 FIRM002|    -.83198        4.67591     -.18  .8588    -9.99660   8.33264
 FIRM003|    -.18608        4.65296     -.04  .9681    -9.30573   8.93356
(Firms 4-20 omitted)
 FIRM021|     .35047        4.63405      .08  .9397    -8.73210   9.43303
 FIRM022|    -.68781        4.83235     -.14  .8868   -10.15903   8.78342
 FIRM023|    -.96206        4.88186     -.20  .8438   -10.53033   8.60622
 FIRM024|   -2.86357        4.82675     -.59  .5530   -12.32383   6.59670
--------+--------------------------------------------------------------------
```

E64.8 True Random Effects Models

We call the stochastic frontier model with a random as opposed to a fixed effect term a 'true random effects' model. The structure is the normal-half normal stochastic frontier model,

$$y_{it} = w_i + \alpha + \boldsymbol{\beta}'\mathbf{x}_{it} + v_{it} + u_{it}$$

$$v_{it} \sim N[0,\sigma_v^2]$$

$$u_{it} = |U_{it}|, U_{it} \sim N[0,\sigma_u^2]$$

$$w_i \sim N[0,\sigma_w^2].$$

At first look, this appears to be a model with a three part disturbance, which would surely be inestimable. But, that is incorrect. It is a model with a traditional random effect, but with the additional feature that the time varying disturbance is not normally distributed. Specifically, the model may be written in our familiar form for the stochastic frontier model,

$$y_{it} = \alpha + \boldsymbol{\beta}'\mathbf{x}_{it} + \varepsilon_{it} + w_i$$

$$\varepsilon_{it} \sim (2/\sigma)\phi(\varepsilon_{it}/\sigma)\Phi(-\varepsilon_{it}\lambda/\sigma)$$

$$w_i \sim N[0,\sigma_w^2].$$

The model is estimable by maximum simulated likelihood, as shown below. Contrast this to the Pitt and Lee form,

$$y_{it} = \alpha + \boldsymbol{\beta}'\mathbf{x}_{it} + v_{it} + u_i$$

$$v_{it} \sim N[0,\sigma_v^2]$$

$$u_i = |U_i|, U_i \sim N[0,\sigma_u^2].$$

In this form, u_i, the time invariant effect, is the inefficiency. In the true random effects model, u_{it} is the inefficiency, and it is time varying. The latent heterogeneity, the random effect, is w_i. Thus, in the Pitt and Lee model, the 'inefficiency' term also contains all other time invariant unmeasured sources of heterogeneity. In the true random effects model, these effects appear in w_i, and u_{it} picks up the inefficiency. By this interpretation, we will expect (and always find) that estimated inefficiencies from the Pitt and Lee are larger than those from the true random effects model, sometimes far larger. The same result is at work in the difference between the Cornwell et al. fixed effects model and the true fixed effects model. Figure E64.8 clearly shows the effect at work.

The true random effects model is estimated as a form of random parameters (RP) model, in which the only random parameter in the model is the constant term. Thus, we write the model in the canonical RP form

$$y_{it} = \alpha_i + \boldsymbol{\beta}'\mathbf{x}_{it} + v_{it} + u_{it}$$

$$v_{it} \sim N[0,\sigma_v^2]$$

$$u_{it} = |U_{it}|, U_{it} \sim N[0,\sigma_u^2]$$

$$\alpha_i = \alpha + w_i$$

$$w_i \sim N[0,\sigma_w^2]$$

Details on estimating random parameters models appear in Section E17.8, so they will be omitted here.

The command structure for the true random effects model is similar to that for the true fixed effects model. The frontier model must be fit twice, first with no effects to generate the starting values, then with the effect specified. The commands are

```
FRONTIER    ; Lhs = ... ; Rhs = one,... ; Par $
FRONTIER    ; Lhs = ... ; Rhs = one,...
            ; RPM ; Fcn = one(n) $
```

If desired, the Jondrow et al. estimates are requested as usual with

; Eff = the variable name

The computation of random parameters models is fairly time consuming because of the simulations. You can control this in part with

; Pts = the number of replications

For exploratory work (or for examples in program documentation), small values such as 25 or 50 are sufficient. For final results destined for publication, larger values, in the range of several hundred are advisable. Also, we advise using Halton sequences rather than pseudorandom numbers for the simulations (see Section E17.8). The parameter is

; Halton

The random parameters formulation also allows a variety of specifications for the mean of the underlying u_{it} – the normal-truncated normal model – and for heteroscedasticity. These are discussed in Section E64.9.

Application

To illustrate the true random effects model, we continue the analysis of the airline data. The commands below estimate the pooled model, then the true RE model. In like fashion to the analysis of fixed effects, we then compare the true random effects estimates of inefficiency to the Pitt and Lee estimates. Figure E64.8 illustrates the general result that the estimated inefficiencies in the true fixed effects model will differ considerably from those produced by the Cornwell et al. approach to fixed effects. Figure E64.9 below shows the same result for the two approaches to random effects. Numerous studies in the literature (see Greene (2005) for discussion) have documented the similarity of the random and fixed approaches – when the same overall structure is used. Thus, Figure E64.10 show similar results for the true fixed and random effects models and for the Pitt and Lee and Cornwell et al. models.

The commands used for this application are as follows:

```
NAMELIST   ; x = one,lf,lm,le,ll,lp,lk $
FRONTIER   ; Lhs = lq ; Rhs = x ; Panel ; Eff = uplre $
FRONTIER   ; Lhs = lq ; Rhs = x ; Par $
FRONTIER   ; Lhs = lq ; Rhs = x ; Panel ; RPM ; Eff  = utre
           ; Fcn = one(n) ; Pts = 50 ; Halton $
FRONTIER   ; Lhs = lq ; Rhs = x ; Par $
FRONTIER   ; Lhs = lq ; Rhs = x ; Panel ; FEM ; Eff = utfe $
DSTAT      ; Rhs = uplre,utre $
CREATE     ; utrebar = Group Mean(utre, Str = firm) $
PLOT       ; Lhs = uplre ; Rhs = utrebar ; Grid
           ; Title = Group Means of u(i,t) vs. Time Invariant u(i) $
PLOT       ; Lhs = utfe ; Rhs = utre ; Grid
           ; Title = Time Varying FE u(i) vs. Time Varying RE u(i) $
```

```
-----------------------------------------------------------------------------
Limited Dependent Variable Model - FRONTIER
Dependent variable                       LQ
Log likelihood function       156.04955
Estimation based on N =    256, K =   9
Stochastic frontier based on panel data
Estimation based on      25 individuals
Variances: Sigma-squared(v)=     .01342
           Sigma-squared(u)=     .06529
           Sigma(v)        =     .11582
           Sigma(u)        =     .25552
Sigma = Sqr[(s^2(u)+s^2(v)]=     .28054
Gamma = sigma(u)^2/sigma^2 =     .82955
Var[u]/{Var[u]+Var[v]}     =     .63879
Stochastic Production Frontier, e = v-u
LR test for inefficiency vs. OLS v only
Deg. freedom for sigma-squared(u):    1
Deg. freedom for heteroscedasticity:  0
Deg. freedom for truncation mean:     0
Deg. freedom for inefficiency model:  1
LogL when sigma(u)=0          108.07431
Chi-sq=2*[LogL(SF)-LogL(LS)] =   95.950
Kodde-Palm C*: 95%: 2.706,  99%:  5.412
--------+--------------------------------------------------------------------
        |                  Standard            Prob.      95% Confidence
      LQ|  Coefficient       Error       z     |z|>Z*         Interval
--------+--------------------------------------------------------------------
        |Deterministic Component of Stochastic Frontier Model
Constant|   -1.70327***      .41761     -4.08  .0000    -2.52176   -.88477
      LF|     .19534**       .09759      2.00  .0453      .00407    .38662
      LM|     .81312***      .06954     11.69  .0000      .67682    .94941
      LE|    1.12741***      .34589      3.26  .0011      .44947   1.80534
      LL|    -.32931***      .07230     -4.55  .0000     -.47102   -.18760
      LP|     .22206***      .06265      3.54  .0004      .09927    .34485
      LK|    -.86072**       .42646     -2.02  .0436    -1.69657   -.02488
        |Variance parameters for compound error
  Lambda|    2.20605*       1.31249      1.68  .0928     -.36639   4.77849
Sigma(u)|     .25552**       .10148      2.52  .0118      .05661    .45442
--------+--------------------------------------------------------------------
```

```
-----------------------------------------------------------------------------
Limited Dependent Variable Model - FRONTIER
Dependent variable                     LQ
Log likelihood function       108.43918
Estimation based on N =    256, K =   9
Variances: Sigma-squared(v)=     .01902
           Sigma-squared(u)=     .01692
           Sigma(v)        =     .13791
           Sigma(u)        =     .13007
Sigma = Sqr[(s^2(u)+s^2(v)]=     .18957
Gamma = sigma(u)^2/sigma^2 =     .47074
Var[u]/{Var[u]+Var[v]}     =     .24425
Stochastic Production Frontier, e = v-u
LR test for inefficiency vs. OLS v only
Deg. freedom for sigma-squared(u):    1
Deg. freedom for heteroscedasticity:  0
Deg. freedom for truncation mean:     0
Deg. freedom for inefficiency model:  1
LogL when sigma(u)=0          108.07431
Chi-sq=2*[LogL(SF)-LogL(LS)] =     .730
Kodde-Palm C*: 95%: 2.706,  99%:  5.412
--------+--------------------------------------------------------------------
        |                  Standard            Prob.      95% Confidence
      LQ|  Coefficient       Error       z    |z|>Z*         Interval
--------+--------------------------------------------------------------------
        |Deterministic Component of Stochastic Frontier Model
Constant|   -2.98823***      .72136    -4.14  .0000    -4.40206  -1.57439
      LF|     .37257***      .07038     5.29  .0000      .23463    .51052
      LM|     .69910***      .07580     9.22  .0000      .55054    .84766
      LE|    2.09473***      .68790     3.05  .0023      .74647   3.44299
      LL|    -.42909***      .06315    -6.79  .0000     -.55287   -.30530
      LP|     .44533***      .09498     4.69  .0000      .25917    .63149
      LK|   -2.09806***      .76556    -2.74  .0061    -3.59853   -.59759
        |Variance parameters for compound error
  Lambda|     .94309***      .16870     5.59  .0000      .61244   1.27373
   Sigma|     .18957***      .00064   297.81  .0000      .18832    .19082
--------+--------------------------------------------------------------------
Note: ***, **, * ==>  Significance at 1%, 5%, 10% level.
-----------------------------------------------------------------------------
```

These are the estimates of the true random effects model. Note that the variation of the random terms in the model has been rearranged. In the pooled model, $s_v = 0.138$ and $s_u = 0.130$. In the random effects model, we have $s_v = .099$ and $s_u = .100$. But, $s_w = .140$. The proportional allocation of the total to u and v has stayed roughly the same, but some additional variation is now attributed to the random effect. Note that the production function parameters have changed substantially as well.

```
-----------------------------------------------------------------------------
Random Coefficients  Frontier Model
Dependent variable                  LQ
Log likelihood function      160.58066
Restricted log likelihood       .00000
Chi squared [   1 d.f.]      321.16131
Significance level              .00000
Estimation based on N =    256, K =  10
Inf.Cr.AIC  =   -301.2 AIC/N =   -1.176
Model estimated: Aug 22, 2011, 23:15:44
Unbalanced panel has     25 individuals
Stochastic frontier (half normal model)
Simulation based on     50 Halton draws
Sigma( u) (1 sided)   =         .09962
Sigma( v) (symmetric) =         .09857
--------+--------------------------------------------------------------------
        |                  Standard            Prob.      95% Confidence
      LQ|  Coefficient       Error       z    |z|>Z*         Interval
--------+--------------------------------------------------------------------
        |Production / Cost parameters, nonrandom first
      LF|     .20387***      .05183     3.93  .0001      .10229    .30545
      LM|     .79450***      .04660    17.05  .0000      .70318    .88583
      LE|    1.10745***      .33573     3.30  .0010      .44943   1.76547
      LL|    -.32691***      .04277    -7.64  .0000     -.41074   -.24308
      LP|     .22812***      .05403     4.22  .0000      .12223    .33401
      LK|    -.84947**       .38344    -2.22  .0267    -1.60101   -.09794
        |Means for random parameters
Constant|   -1.83727***      .35442    -5.18  .0000    -2.53191  -1.14263
        |Scale parameters for dists. of random parameters
Constant|     .11729***      .00934    12.56  .0000      .09898    .13559
        |Variance parameter for v +/- u
   Sigma|     .14015***      .01373    10.21  .0000      .11325    .16705
        |Asymmetry parameter, lambda
  Lambda|    1.01064**       .43792     2.31  .0210      .15234   1.86895
--------+--------------------------------------------------------------------
Note: ***, **, * ==>  Significance at 1%, 5%, 10% level.
-----------------------------------------------------------------------------

Descriptive Statistics
--------+---------------------------------------------------------------------
Variable|       Mean        Std.Dev.      Minimum       Maximum     Cases Missing
--------+---------------------------------------------------------------------
   UPLRE|      .221170      .117670      .016992      .435912       256        0
    UTRE|      .078815      .031677      .026405      .305595       256        0
--------+---------------------------------------------------------------------
```

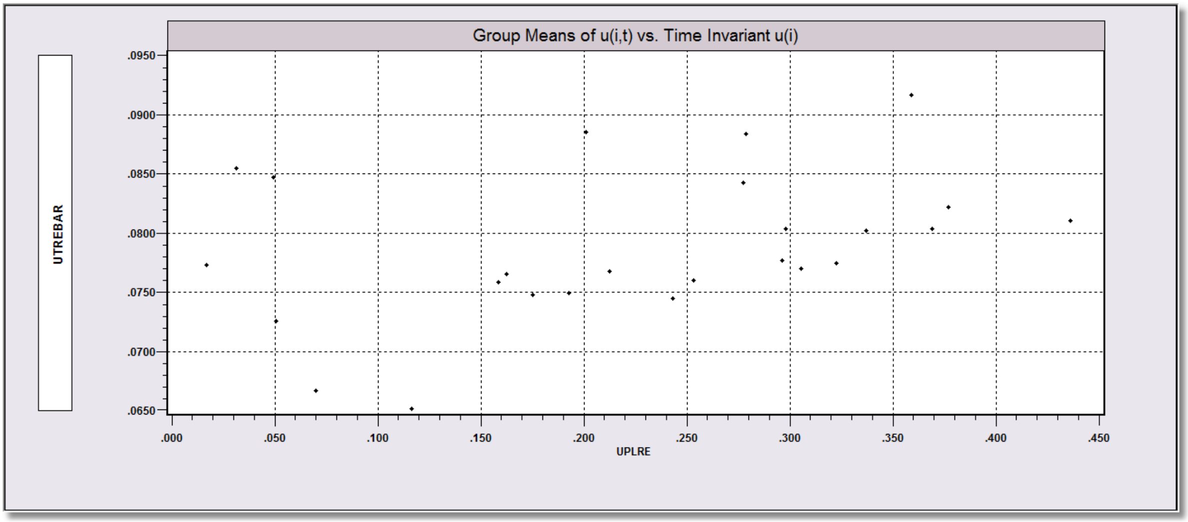

Figure E64.9 Time Varying vs. Time Invariant Estimates of u(i)

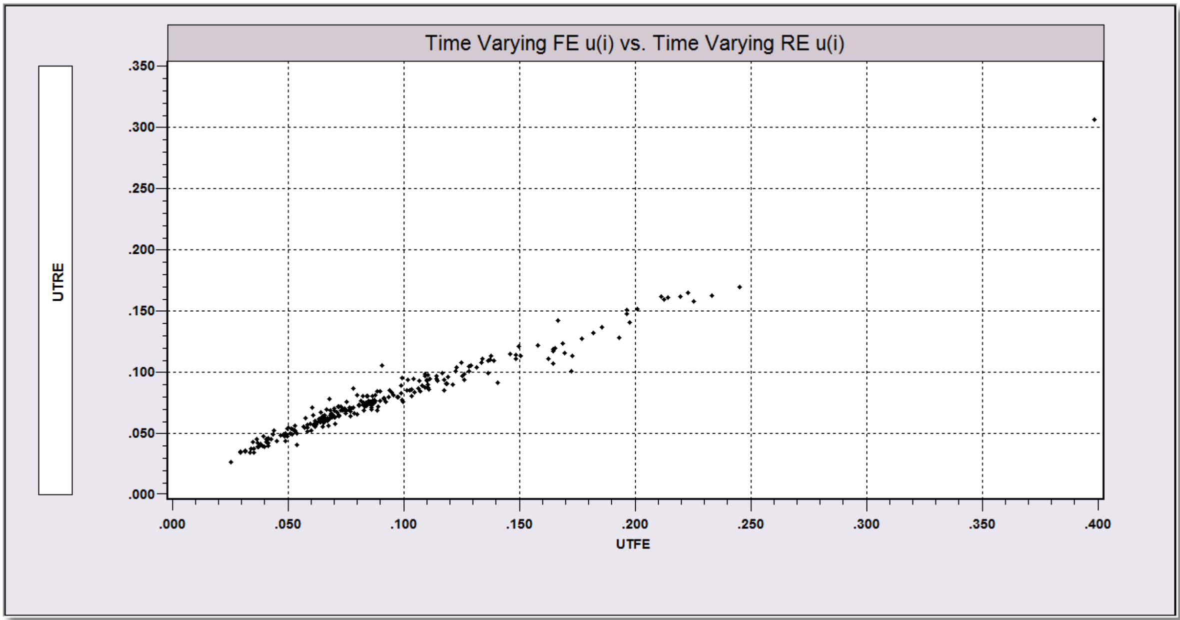

Figure E64.10 Comparison of Time Varying Fixed and Random Effects Estimates

E64.9 Random Parameters Stochastic Frontier Models

The random parameters stochastic frontier model in *LIMDEP* is very general, and embodies all three of the formulations discussed in the preceding sections on fixed and random effects.

$$y_{it} = \boldsymbol{\beta}_i'\mathbf{x}_{it} + v_{it} - u_{it},$$

$$u_i = |\,N[\mu_{it}, \sigma_{uit}^2]\,|$$

$$\mu_{it} = \boldsymbol{\delta}_i'\mathbf{m}_{it}.$$

$$\sigma_{uit}^2 = \sigma_u^2 \times \exp(\boldsymbol{\gamma}_i'\mathbf{w}_{it}).$$

The model allows, all at once, half normal or truncated normal distribution for u_i and firmwise and/or timewise heteroscedasticity in u_{it}. The model form allows parameters to be random in all three parts of the specification with the single restriction noted below. (Only the variance of the 'disturbance,' v_{it} is assumed to be constant. In addition, this model form does not accommodate heteroscedasticity in v_{it}.) As will be clear in what follows, the true random effects model developed in the previous section is a special case of this model with nonrandom parameters in μ_{it} and σ_{uit}^2 and only a random constant term in $\boldsymbol{\beta}_i$.

NOTE: The random parameters normal-truncated normal model with heteroscedasticity (in u_{it}) at the same time is not identified. Only one of these two should be specified. The command parser will not prevent you from specifying such a model, but it will ultimately be impossible to obtain the parameter estimates.

The general structure of the random parameters stochastic frontier model is based on the conditional density

$$f(y_{it}|\,\mathbf{x}_{it}, \boldsymbol{\beta}_i) = f(\boldsymbol{\beta}_i'\mathbf{x}_{it}),\ i = 1,...,N,\ t = 1,...,T_i$$

where

$$\boldsymbol{\beta}_i = \boldsymbol{\beta} + \boldsymbol{\Delta}\mathbf{z}_i + \boldsymbol{\Gamma}\mathbf{v}_i$$

and $f(.)$ is the density for the stochastic frontier regression model. The model assumes that parameters are randomly distributed with possibly heterogeneous (across individuals) means

$$E[\boldsymbol{\beta}_i|\,\mathbf{z}_i] = \boldsymbol{\beta} + \boldsymbol{\Delta}\mathbf{z}_i,$$

(the second term is optional – the mean may be constant), and

$$\text{Var}[\boldsymbol{\beta}_i|\,\mathbf{z}_i] = \boldsymbol{\Sigma}.$$

As noted earlier, the heterogeneity term is optional. In addition, it may be assumed that some of the parameters are nonrandom by placing rows of zeros in the appropriate places in $\boldsymbol{\Delta}$and $\boldsymbol{\Gamma}$. The general form of random parameter vector $\boldsymbol{\beta}_i$ is also extended to $\boldsymbol{\delta}_i$ and $\boldsymbol{\gamma}_i$. The general aspects of random parameters model estimation in *LIMDEP* are described in Section E17.8.

Command for the Random Parameters Model

The model command for the random parameters form of the stochastic frontier model is as follows. The first **FRONTIER** command is mandatory, and is needed to obtain the starting values. This is a pooled data version of the model. Note that it does not include the heteroscedasticity or truncation specification, even if the second command does.

```
FRONTIER    ; Lhs = dependent variable ; Rhs = independent variables
            ; Parameters  $
FRONTIER    ; Lhs = dependent variable
            ; Rhs = independent variables
            [ ; Rh2 = list is optional for the truncated normal model ]
            [ ; Hfn  = list is optional for the heteroscedasticity model ]
            ; Pds = fixed periods or count variable
            ; RPM (may include = variables in z)
            ; Fcn = random parameters specification $
```

(Note, again, only one of the two optional specifications noted should be specified.)

> **NOTE:** For this model, your Rhs list must include a constant term. Though not strictly necessary, you should also include constants in Rh2 or Hfn if they are specified.

Specifying Random Parameters

The **; Fcn = specification** is used to define the random parameters. It is constructed from the list of Rhs names as follows: Suppose your model is specified by

```
; Rhs = one, x1, x2, x3, x4
```

This involves five coefficients. Any or all of them may be random; any not specified as random are assumed to be constant. For those that you wish to specify as random, use the following for production (cost, profit) function parameters,

```
; Fcn = variable name (distribution),
        variable name (distribution), ...
```

There are two other sets of parameters in the model, in the mean of and variance of the one sided disturbance. To specify random parameters in the underlying mean of the truncated normal variable, use the following:

```
; Fcn = variable name [distribution],
        variable name [distribution], ...
```

(Note square brackets designate the terms in μ_{it}.) For parameters in the computation of the variance of u_{it}, use

```
; Fcn = variable name <distribution>,
        variable name <distribution>, ...
```

The difference in the three formulations is in the enclosures, ( ) for production function, [ ] for mean of the truncated distribution, and <> for the variance of the one sided disturbance. This distinction is necessary because the lists might have variables in common, and this is the only way to distinguish them. In particular, it is likely that all three lists would include *one*, so this device is used to distinguish the three functions.

Three distributions may be specified All random variables have mean 0.

n = standard normal distribution, variance = 1,
t = triangular (tent shaped) distribution in [-1,+1], variance = 1/6,
u = standard uniform distribution [-1,1], variance = 1/3.

Note that each of these is scaled as it enters the distribution, so the variance is only that of the random draw before multiplication. (See Chapter R23 for discussion of this computation and for other distributions that can be specified.) The latter two distributions are provided as one may wish to reduce the amount of variation in the tails of the distribution of the parameters across individuals and to limit the range of variation. (See Train (2010) for discussion.) For example, to specify that the constant term and the coefficient on *x*1 are normally distributed with fixed mean and variance, and a normally distributed constant in the mean of the truncated distribution, you might use

; Fcn = one(n), x1(n), one[n]

This specifies that the first and second coefficients are random while the remainder are not. The parameters estimated will be the mean and standard deviations of the distributions of these two parameters and the fixed values of the other three.

NOTE: If you use the wrong enclosures for the variables, a diagnostic will appear that the program does not recognize a variable. For example:

```
FRONTIER    ; Lhs = lq ; Rhs = one,lf,lm,le,ll,lp
            ; Hfn = one,lf ; RPM ; Pds = ni
            ; Fcn = one(n),lf(n),lf[n] $
```

```
Variable in FCN=name[type] is not in RHS/RH2/HFN list.
```

The reason for the diagnostic is that the **lf[n]** would indicate a specification for the truncation model, using **; Rh2 = list**. But, this command specifies only heteroscedasticity, which is denoted with <> enclosures. Hence, when the **lf[n]** is encountered, *LIMDEP* searches for *lf* in an Rh2 list, and finding no such list, issues the diagnostic.

Correlated Random Parameters

The stochastic frontier model does not support correlated random parameters. The model is not identified with this extension.

Heterogeneity in the Means

The preceding examples have specified that the mean of the random variable is fixed over individuals. If there is measured heterogeneity in the means, in the form of

$$E[\beta_{ki}] = \beta_k + \Sigma_m \delta_{km} z_{mi}$$

where z_{mi} is a variable that is measured for each individual, then the command may be modified to

; RPM = list of variables in z

In the data set, these variables must be repeated for each observation in the group. Since the coefficients are assumed to be time invariant, the variables in $\mathbf{z}_i$ must be also.

The Parameter Vector and Retained Results

The variances of the underlying random variables are given earlier, 1 for the normal distribution, 1/3 for the uniform, and 1/6 for the tent distribution. The σ_k parameters are only the standard deviations for the normal distribution. For the other two distributions, σ_k is a scale parameter. The standard deviation is obtained as $\sigma_k / \sqrt{3}$ for the uniform distribution and $\sigma_k / \sqrt{6}$ for the triangular distribution. When the parameters are correlated, the implied covariance matrix is adjusted accordingly. The correlation matrix is unchanged by this.

Results saved by this estimator are:

Matrices:	*b*	= estimate of θ
	varb	= asymptotic covariance matrix for estimate of θ.
	beta_i	= individual specific parameters, if **; Par** is requested.
Scalars:	*kreg*	= number of variables in Rhs
	nreg	= number of observations
	logl	= log likelihood function

Last Model: *b_variables*

Last Function: None

Standard Model Specifications for the Stochastic Frontier Random Parameters Model

This is the full list of general specifications that are applicable to this model estimator.

Controlling Output from Model Commands

; Par keeps individual specific parameter estimates.
; Table = name saves model results to be combined later in output tables.

Robust Asymptotic Covariance Matrices

; Covariance Matrix displays estimated asymptotic covariance matrix (normally not shown), same as **; Printvc**.
; Robust requests a 'sandwich' estimator or robust covariance matrix for TSCS and several discrete choice models.

Optimization Controls for Nonlinear Optimization

; Tlg[= value] sets convergence value for gradient.
; Tlf [= value] sets convergence value for function.
; Tlb[= value] sets convergence value for parameters.
; Alg = name requests a particular algorithm, Newton, DFP, BFGS, etc.
; Maxit = n sets the maximum iterations.
; Output = n requests technical output during iterations; the level '**n**' is 1, 2, 3 or 4.
; Set keeps current setting of optimization parameters as permanent.

Predictions and Residuals

; List displays a list of fitted values with the model estimates.
; Keep = name keeps fitted values as a new (or replacement) variable in data set.
; Res = name keeps residuals as a new (or replacement) variable.

Hypothesis Tests and Restrictions

; Test: spec defines a Wald test of linear restrictions.
; Wald: spec defines a Wald test of linear restrictions, same as **; Test: spec**.
; CML: spec defines a constrained maximum likelihood estimator.
; Rst = list specifies equality and fixed value restrictions.

Application

We continue the earlier application by fitting the stochastic frontier model with random parameters. The random parameters truncation model appears to be unidentified in these data, so the second model fit is with heteroscedasticity. In the first model, the constant and one of the production coefficients is specified to be random. In the second, these two coefficients and the parameter on the variable that enters the variance function are all taken to be random. The kernel density estimators compare the efficiency estimates from the random parameters model to those from the simplest pooled estimator.

The commands are:

```
NAMELIST   ; x = one,lf,lm,le,ll,lp,lk $
FRONTIER   ; Lhs = lq ; Rhs = x ; Eff = u $
FRONTIER   ; Lhs = lq ; Rhs = x
           ; RPM ; Panel ; Pts = 50 ; Halton; Fcn = one(n),lf(n) ; Eff = urp1 $
KERNEL     ; Rhs = urp1,u $
FRONTIER   ; Lhs = lq ; Rhs = x $
FRONTIER   ; Lhs = lq ; Rhs = x ; Hfn = one,loadfctr
           ; RPM ; Panel ; Pts = 50 ; Halton
           ; Fcn = one(n),lf(n),loadfctr<n> $
```

```
-----------------------------------------------------------------------------
Random Coefficients  Frontier Model
Dependent variable                   LQ
Log likelihood function       161.33196
Restricted log likelihood        .00000
Chi squared [   2 d.f.]       322.66392
Significance level               .00000
Estimation based on N =    256, K =  11
Inf.Cr.AIC  =   -300.7 AIC/N =   -1.174
Model estimated: Aug 22, 2011, 23:28:18
Unbalanced panel has     25 individuals
Stochastic frontier (half normal model)
Simulation based on     50 Halton draws
Sigma( u) (1 sided)   =          .10598
Sigma( v) (symmetric) =          .09399
--------+--------------------------------------------------------------------
        |                Standard            Prob.      95% Confidence
      LQ|  Coefficient      Error       z    |z|>Z*         Interval
--------+--------------------------------------------------------------------
        |Production / Cost parameters, nonrandom first
      LM|      .81447***      .04526    18.00  .0000       .72577    .90317
      LE|     1.16342***      .31391     3.71  .0002       .54817   1.77867
      LL|     -.33712***      .04111    -8.20  .0000     -.41769   -.25654
      LP|      .24213***      .04782     5.06  .0000       .14841    .33585
      LK|     -.94502***      .35520    -2.66  .0078    -1.64119   -.24886
        |Means for random parameters
Constant|   -1.89056***      .33140    -5.70  .0000    -2.54009  -1.24103
      LF|      .21430***      .05277     4.06  .0000       .11088    .31773
        |Scale parameters for dists. of random parameters
Constant|      .12526***      .00926    13.53  .0000       .10711    .14341
      LF|      .04979***      .00823     6.05  .0000       .03366    .06592
        |Variance parameter for v +/- u
   Sigma|      .14165***      .01265    11.20  .0000       .11686    .16645
        |Asymmetry parameter, lambda
  Lambda|     1.12768***      .42335     2.66  .0077       .29792   1.95743
--------+--------------------------------------------------------------------
Note: ***, **, * ==>  Significance at 1%, 5%, 10% level.
-----------------------------------------------------------------------------
```

Figure E64.11 shows the distributions of the estimates of inefficiencies from the random parameters model and the simple, pooled fixed parameters model. The figure suggests that the RP formulation is moving some of the variation of the outcome variable out of the inefficiency term and into the production model, in the form of parameter variation.

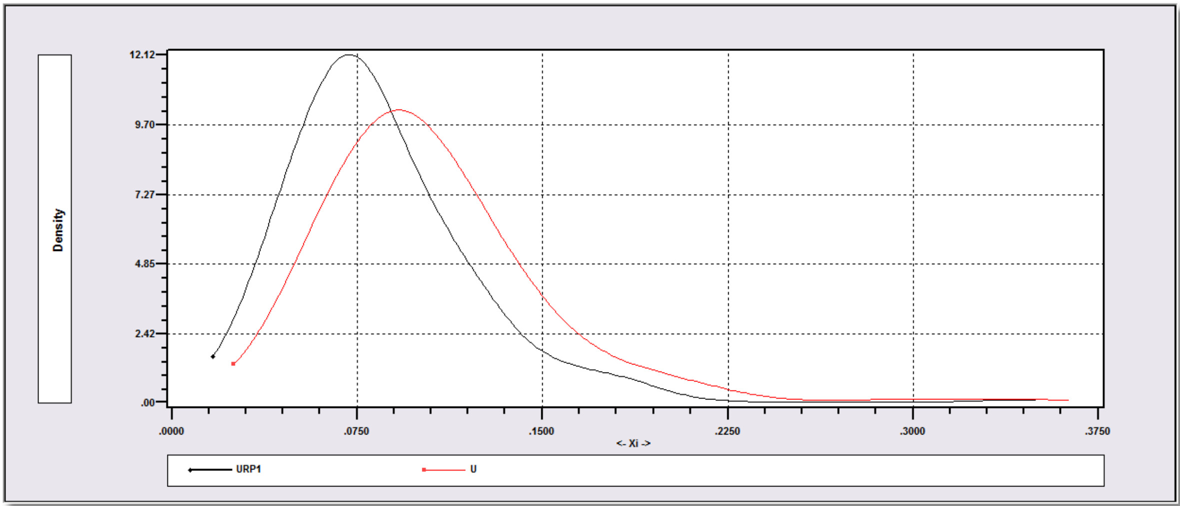

Figure E64.11 Kernel Density Estimator for Random Parameters Model Inefficiencies

```
-----------------------------------------------------------------------------
Random Coefficients  FrntrTrn Model
Dependent variable                     LQ
Log likelihood function       199.14429
Estimation based on N =    256, K =  13
Unbalanced panel has     25 individuals
Stochastic frontier, truncation/hetero.
Simulation based on     50 Halton draws
Estimated parameters of efficiency dstn
s(u)  =     .189842  s(v)=        .07165
avgE[u|e]=    .10986  avgE[TE|e]= .90303
Lambda  = su/sv =                2.64974
--------+--------------------------------------------------------------------
        |                 Standard            Prob.      95% Confidence
      LQ|  Coefficient      Error       z    |z|>Z*         Interval
--------+--------------------------------------------------------------------
        |Nonrandom parameters
      LM|     .62243***      .04223    14.74  .0000       .53966     .70521
      LE|     .38353         .28063     1.37  .1717     -.16649     .93355
      LL|    -.36579***      .03589   -10.19  .0000     -.43614    -.29544
      LP|     .15282***      .04217     3.62  .0003      .07017     .23547
      LK|    -.16125         .31392     -.51  .6075     -.77652     .45401
   suONE|    9.05239***     1.65934     5.46  .0000     5.80014   12.30464
        |Means for random parameters
Constant|   -1.17144***      .29799    -3.93  .0001    -1.75549    -.58739
      LF|     .49011***      .04904     9.99  .0000      .39398     .58623
suLOADFC|   -16.4160***     3.47560    -4.72  .0000    -23.2281    -9.6039
        |Scale parameters for dists. of random parameters
Constant|     .12591***      .00859    14.65  .0000      .10906     .14275
      LF|     .01186**       .00593     2.00  .0456      .00023     .02350
suLOADFC|    1.47653***      .36192     4.08  .0000      .76718    2.18589
        |Sigma(v) from symmetric disturbance.
Sigma(v)|     .07165***      .00670    10.69  .0000      .05851     .08478
--------+--------------------------------------------------------------------
Note: ***, **, * ==>  Significance at 1%, 5%, 10% level.
-----------------------------------------------------------------------------
```

E64.10 Alvarez et al. – Fixed Management Model

Alvarez, Arias and Greene (2006) suggested a production model in which an unobserved factor enters as a latent variable. The core production model is

$$y_{it} = f(x_{it,1}, x_{it,2}, \ldots, x_{it,K}, m_i)$$

where the unobservable, time invariant factor, 'm_i' is labeled 'management' in their paper. By treating the unobserved factor as a random component in the model, the authors develop a stochastic frontier model in which the resultant functional form is such that all random parameters are functions of the same single random effect, v_i, and the v_i appears in squared form in the equation as well. In generic terms, this model is a random parameters stochastic frontier model with random constant term and first order terms, and nonrandom second order terms in a translog model. The functional form is

$$\begin{aligned}
\log y_{it} &= \alpha_i + \sum_{k=1}^{K} \beta_{k,i} \ln x_{it,k} + \sum_{k=1}^{K}\sum_{m=1}^{K} \gamma_{km} \ln x_{it,k} \ln x_{it,m} + v_{it} - u_{it} \\
\alpha_i &= \alpha + \theta_\alpha w_i + \theta_{\alpha\alpha}\left(\tfrac{1}{2} w_i^2\right) \\
\beta_{k,i} &= \beta_k + \lambda_k w_i \\
w_i &\sim N[0,1] \\
v_{it} &\sim N[0,\sigma_v^2] \\
u_{it} &= |N[0,\sigma_u^2]|
\end{aligned}$$

This model is specified simply by creating the necessary variables, then building a random parameters model with the two additional specifications,

; Common ; Mgt

The **; Common** specification alone is generic, and applies to all random parameters models. Use it to specify that the same random component appears in all random parameters. The **; Mgt** specification has no function outside the frontier model. It is used only with the frontier model to specify this particular form. For example, consider the following three factor translog model:

```
FRONTIER    ; Lhs = yit ; Rhs = one,x1,x2,x3,x11,x12,x13,x22,x23,x33 $
FRONTIER    ; Lhs = yit ; Rhs = one,x1,x2,x3,x11,x12,x13,x22,x23,x33
            ; RPM ; Pds = the panel specification ; Halton
            ; Fcn = one(n),x1(n),x2(n),x3(n)
            ; Common ; Mgt $
```

(It is always necessary to fit the frontier model with fixed parameters first to generate the starting values.)

An extension of this model that the authors considered was intended to ameliorate the probable correlation between the random effect w_i and the independent variables (factors). The Mundlak approach to this problem is to incorporate the group means of the variables in the model. For this model, they proposed

$$w_i = \tau + \sum_{k=1}^{K} \gamma_k \overline{\log x_{i,k}} + f_i$$

where f_i is now the structural random variable that drives the random parameters. This extension is requested with

; Means

(The program deduces internally which variables are nonconstant and should be used.)

Application

The following is the Alvarez, Arias and Greene application. The data consists of six years of observations on 247 Spanish dairy farms. The output, *yit* is milk production. The four inputs, *x*1, *x*2, *x*3 and *x*4 are feed, land, labor and cows. Commands for fitting the model are as follows: (We have restricted the number of iterations and the number of replications for purpose of this numerical illustration.) Both models (with and without the Mundlak adjustment) are shown.

```
FRONTIER  ; Lhs = yit
          ; Rhs = one,x1,x2,x3,x4,x11,x12,x13,x14,x23,x24,x34,x44 ; Par $
FRONTIER  ; Lhs = yit
          ; Rhs = one,x1,x2,x3,x4,x11,x12,x13,x14,x23,x24,x34,x44
          ; RPM ; Halton ; Pts = 25 ; Pds = 6 ; Maxit = 25 ; Common ; Mgt
          ; Fcn = one(n),x1(n),x2(n),x3(n),x4(n) $
FRONTIER  ; Lhs = yit
          ; Rhs = one,x1,x2,x3,x4,x11,x12,x13,x14,x23,x24,x34,x44 ; Par $
FRONTIER  ; Lhs = yit
          ; Rhs = one,x1,x2,x3,x4,x11,x12,x13,x14,x23,x24,x34,x44
          ; RPM ; Halton ; Pts = 25 ; Pds = 6 ; Maxit = 25
          ; Common ; Mgt ; Means
          ; Fcn = one(n),x1(n),x2(n),x3(n),x4(n) $
```

The first set of results is the pooled stochastic frontier model with no extensions or modifications.

```
-----------------------------------------------------------------------------
Limited Dependent Variable Model - FRONTIER
Dependent variable                    YIT
Log likelihood function       851.16734
Estimation based on N =   1482, K =  15
Variances: Sigma-squared(v)=     .00876
           Sigma-squared(u)=     .02831
           Sigma(v)        =     .09359
           Sigma(u)        =     .16825
Sigma = Sqr[(s^2(u)+s^2(v)]=     .19253
Gamma = sigma(u)^2/sigma^2 =     .76371
Var[u]/{Var[u]+Var[v]}     =     .54012
Stochastic Production Frontier, e = v-u
LR test for inefficiency vs. OLS v only
Deg. freedom for sigma-squared(u):    1
Deg. freedom for heteroscedasticity:  0
Deg. freedom for truncation mean:     0
Deg. freedom for inefficiency model:  1
LogL when sigma(u)=0          829.23705
Chi-sq=2*[LogL(SF)-LogL(LS)] =   43.861
Kodde-Palm C*: 95%: 2.706,  99%:  5.412
--------+--------------------------------------------------------------------
        |                  Standard            Prob.      95% Confidence
     YIT|  Coefficient       Error       z    |z|>Z*         Interval
--------+--------------------------------------------------------------------
        |Deterministic Component of Stochastic Frontier Model
Constant|    11.6942***      .00529  2209.86  .0000     11.6838   11.7046
      X1|     .60483***      .02133    28.35  .0000      .56302    .64664
      X2|     .02246**       .01140     1.97  .0489      .00011    .04480
      X3|     .02336*        .01245     1.88  .0606     -.00104    .04776
      X4|     .44945***      .01172    38.34  .0000      .42647    .47242
     X11|     .59297***      .13525     4.38  .0000      .32789    .85806
     X12|    -.17183***      .04842    -3.55  .0004     -.26673   -.07693
     X13|     .20033***      .06903     2.90  .0037      .06502    .33563
     X14|    -.32993***      .07299    -4.52  .0000     -.47297   -.18688
     X23|     .00386         .04203      .09  .9268     -.07852    .08624
     X24|     .06473**       .03009     2.15  .0314      .00576    .12369
     X34|    -.07096*        .03853    -1.84  .0655     -.14648    .00455
     X44|     .20854***      .04328     4.82  .0000      .12373    .29336
        |Variance parameters for compound error
  Lambda|    1.79780***      .10292    17.47  .0000     1.59608   1.99951
   Sigma|     .19253***      .00011  1715.95  .0000      .19231    .19275
--------+--------------------------------------------------------------------
Note: ***, **, * ==>  Significance at 1%, 5%, 10% level.
-----------------------------------------------------------------------------
```

This is the fixed management model without the Mundlak correction.

```
+----------------------------------------------+
| Random Coefficients  Frontier Model          |
| Dependent variable                  YIT      |
| Log likelihood function      1327.58807      |
| Estimation based on N =   1482, K =  21      |
| Sample is  6 pds and    247 individuals      |
+----------------------------------------------+
----------------------------------------------------------------------------
All parameters have the same random effect
Alvarez/Arias/Greene Fixed Mgt. SF Model
Stochastic frontier (half normal model)
Simulation based on     25 Halton draws
Sigma( u) (1 sided)   =          .09355
Sigma( v) (symmetric) =          .05799
--------+-------------------------------------------------------------------
        |                  Standard            Prob.      95% Confidence
     YIT|  Coefficient       Error       z    |z|>Z*         Interval
--------+-------------------------------------------------------------------
        |Production / Cost parameters, nonrandom first
     X11|     .19550**       .08392     2.33  .0198      .03101    .35999
     X12|    -.00410         .02903     -.14  .8876     -.06100    .05279
     X13|    -.03972         .04116     -.96  .3346     -.12039    .04095
     X14|    -.08681**       .04220    -2.06  .0397     -.16952   -.00410
     X23|     .02377         .02534      .94  .3483     -.02590    .07344
     X24|    -.01893         .01743    -1.09  .2775     -.05310    .01524
     X34|     .02550         .02305     1.11  .2684     -.01967    .07067
     X44|     .09988***      .02339     4.27  .0000      .05403    .14572
        |Means for random parameters
Constant|    11.6506***      .00445  2620.80  .0000     11.6418   11.6593
      X1|     .65048***      .01227    53.03  .0000      .62643    .67452
      X2|     .03525***      .00681     5.17  .0000      .02190    .04861
      X3|     .04531***      .00759     5.97  .0000      .03043    .06019
      X4|     .40147***      .00646    62.16  .0000      .38881    .41413
        |Coefficients on unobservable fixed management
Constant|     .12579***      .00238    52.96  .0000      .12114    .13045
      X1|    -.02248*        .01218    -1.85  .0649     -.04635    .00139
      X2|     .00767         .00851      .90  .3676     -.00902    .02436
      X3|     .00794         .00939      .85  .3979     -.01047    .02635
      X4|    -.00967         .00657    -1.47  .1410     -.02255    .00320
Alpha_mm|    -.02835***      .00414    -6.85  .0000     -.03646   -.02024
        |Variance parameter for v +/- u
   Sigma|     .11007***      .00289    38.04  .0000      .10439    .11574
        |Asymmetry parameter, lambda
  Lambda|    1.61332***      .11959    13.49  .0000     1.37893   1.84771
--------+-------------------------------------------------------------------
Note: ***, **, * ==>  Significance at 1%, 5%, 10% level.
----------------------------------------------------------------------------
```

```
+----------------------------------------------+
| Random Coefficients  Frontier Model          |
| Dependent variable                 YIT       |
| Log likelihood function      1273.63070      |
| Sample is  6 pds and    247 individuals      |
+----------------------------------------------+
-----------------------------------------------------------------------------
All parameters have the same random effect
Alvarez/Arias/Greene Fixed Mgt. SF Model
Stochastic frontier (half normal model)
Simulation based on     25 Halton draws
Sigma( u) (1 sided)   =          .12577
Sigma( v) (symmetric) =          .05376
--------+--------------------------------------------------------------------
        |                 Standard            Prob.      95% Confidence
     YIT|  Coefficient       Error       z    |z|>Z*         Interval
--------+--------------------------------------------------------------------
        |Production / Cost parameters, nonrandom first
     X11|    -.06957         .08521     -.82  .4142     -.23658    .09743
     X12|     .00164         .02989      .05  .9562     -.05693    .06022
     X13|     .31592***      .04339     7.28  .0000      .23087    .40097
     X14|    -.08946*        .04767    -1.88  .0606     -.18289    .00398
     X23|    -.02088         .02784     -.75  .4533     -.07545    .03369
     X24|    -.04357**       .01912    -2.28  .0227     -.08103   -.00610
     X34|    -.15581***      .02350    -6.63  .0000     -.20187   -.10975
     X44|     .16310***      .02763     5.90  .0000      .10895    .21725
        |Means for random parameters
Constant|    11.6829***      .00449  2601.72  .0000     11.6741   11.6917
      X1|     .60260***      .02198    27.41  .0000      .55951    .64569
      X2|     .05221***      .01636     3.19  .0014      .02015    .08427
      X3|     .10728***      .02775     3.87  .0001      .05290    .16166
      X4|     .39780***      .01047    38.00  .0000      .37728    .41832
        |Coefficients on unobservable fixed management
Constant|     .11398***      .00235    48.52  .0000      .10937    .11858
      X1|    -.05393***      .01134    -4.76  .0000     -.07616   -.03171
      X2|     .03061***      .00916     3.34  .0008      .01265    .04857
      X3|     .01309         .01202     1.09  .2760     -.01046    .03665
      X4|     .01621**       .00707     2.29  .0218      .00236    .03007
Alpha_mm|    -.03575***      .00368    -9.72  .0000     -.04296   -.02855
        |Variance parameter for v +/- u
   Sigma|     .13678***      .00368    37.19  .0000      .12957    .14399
        |Asymmetry parameter, lambda
  Lambda|    2.33925***      .14491    16.14  .0000     2.05524   2.62326
        |Variable Means in Unobserved Management
  X1_bar|    -.12466         .22073     -.56  .5722     -.55728    .30796
  X2_bar|     .00045         .15758      .00  .9977     -.30839    .30930
  X3_bar|     .01632         .25437      .06  .9489     -.48224    .51487
  X4_bar|     .15107         .11332     1.33  .1825     -.07102    .37316
--------+--------------------------------------------------------------------
Note: ***, **, * ==>  Significance at 1%, 5%, 10% level.
-----------------------------------------------------------------------------
```

E64.11 Latent Class Stochastic Frontier Models

The latent class framework discussed in Chapter E20 is available for the stochastic frontier model. The structural equations of the basic model are

$$y_{it} \mid j = \boldsymbol{\beta}_j'\mathbf{x}_{it} + v_{it} - u_{it},$$

$$v_i \mid j = \mathrm{N}[0, \sigma_{vj}^2]$$

$$u_i \mid j = |\,\mathrm{N}[\sigma_{uj}^2]\,|$$

where 'j' indicates class j. The truncation and heteroscedasticity models are not supported by this estimator. However, the Battese and Coelli model, in which

$$u_{it} \mid j = g(\mathbf{z}_{it})|\, j \times |U_i|$$

is available for both forms of $g(\mathbf{z}_{it})$.

The estimation command for the latent class stochastic frontier model is

FRONTIER ; Lhs = dependent variable
; Rhs = one, remaining variables ; Parameters $
FRONTIER ; Lhs = dependent variable
; Rhs = one, remaining variables
; Pds = fixed periods or count variable
; LCM ; Pts = number of classes (2, 3, ..., 9) $

(As in other panel data settings, it is necessary to fit the pooled model first to compute the starting values.)

The Battese and Coelli models may be specified here with

; Model = BC

for the decay model and

; Model = BC
; Hfu = one, heteroscedasticity variables

For this model, you must fit the identical Battese and Coelli model without the latent class specification first. The application below demonstrates.

The basic form of the latent class model assumes that the class probabilities are fixed values. You may make them dependent on time invariant variables, w_i with

; LCM = list of variables in w

Do not include *one* in the list.

Some particular variables computed for the latent class model are

; Group = the index of the most likely latent class
; Cprob = estimated probability for the most likely latent class

You can obtain a listing of these two results by using

; List

An example appears below. You can also use the **; Rst = list** option to structure the latent class model so that different variables appear in different classes or that certain coefficients are equal across classes. Examples are given in Chapter E20.

Estimates retained by this model include:

Matrices:	*b*	= full parameter vector, $[\boldsymbol{\beta}_1{}'\lambda_1\sigma_1, \boldsymbol{\beta}_2{}',\lambda_2\sigma_2, ... F_1,...,F_J]$
	varb	= full covariance matrix
	beta_i	= individual specific parameters, if **; Par** is requested

Note that *b* and *varb* involve $J\times(K+2)$ estimates. Two additional matrices are created,

b_class	= a $J\times K$ matrix with each row equal to the corresponding $\boldsymbol{\beta}_j$
class_pr	= a $J\times 1$ vector containing the estimated class probabilities

Scalars:	*kreg*	= number of variables in Rhs list
	nreg	= total number of observations used for estimation
	logl	= maximized value of the log likelihood function
	exitcode	= exit status of the estimation procedure

Standard Model Specifications for the Latent Class Stochastic Frontier Model

This is the full list of general specifications that are applicable to this model estimator.

Controlling Output from Model Commands

; Par	keeps individual specific parameter estimates.
; Margin	displays marginal effects.
; OLS	displays least squares starting values when (and if) they are computed.
; Table = name	saves model results to be combined later in output tables.

Robust Asymptotic Covariance Matrices

; Covariance Matrix	displays estimated asymptotic covariance matrix (normally not shown), same as **; Printvc**.
; Robust	requests a ‘sandwich’ estimator or robust covariance matrix for TSCS and several discrete choice models.

Optimization Controls for Nonlinear Optimization

; Start = list	gives starting values for a nonlinear model.
; Tlg[= value]	sets convergence value for gradient.
; Tlf [= value]	sets convergence value for function.
; Tlb[= value]	sets convergence value for parameters.
; Alg = name	requests a particular algorithm, Newton, DFP, BFGS, etc.
; Maxit = n	sets the maximum iterations.
; Output = n	requests technical output during iterations; the level '**n**' is 1, 2, 3 or 4.
; Set	keeps current setting of optimization parameters as permanent.

Predictions and Residuals

; List	displays a list of fitted values with the model estimates.
; Keep = name	keeps fitted values as a new (or replacement) variable in data set.
; Res = name	keeps residuals as a new (or replacement) variable.
; Fill	fills missing values (outside estimating sample) for fitted values.

Hypothesis Tests and Restrictions

; Test: spec	defines a Wald test of linear restrictions.
; Wald: spec	defines a Wald test of linear restrictions, same as **; Test: spec**.
; CML: spec	defines a constrained maximum likelihood estimator.
; Rst = list	specifies equality and fixed value restrictions.

Application

The airline data used in the preceding examples are clearly not compatible with this model; no configuration of the equation produces meaningful results. To illustrate the estimator, we have borrowed the Spanish dairy data used in the previous section. The following commands fit a two class, Battese and Coelli decay model.

```
NAMELIST    ; x  = one,x1,x2,x3,x4 $
FRONTIER    ; Lhs = yit ; Rhs = x
            ; Model = BC
            ; Pds  = 6 $
FRONTIER    ; Lhs = yit ; Rhs = x
            ; Model = BC
            ; LCM ; Pts = 2 ; Pds = 6 ; List $
```

These are the initial results from the first command.

```
-----------------------------------------------------------------------------
Limited Dependent Variable Model - FRONTIER
Dependent variable                   YIT
Log likelihood function      1390.20024
Stochastic frontier based on panel data
Estimation based on     247 individuals
Variances: Sigma-squared(v)=     .00549
           Sigma-squared(u)=     .03940
           Sigma(v)        =     .07413
           Sigma(u)        =     .19848
Sigma = Sqr[(s^2(u)+s^2(v)]=     .21187
Gamma = sigma(u)^2/sigma^2 =     .87759
Var[u]/{Var[u]+Var[v]}     =     .72263
Stochastic Production Frontier, e = v-u
Battese-Coelli Models: Time Varying uit
Time dependent uit=exp[-eta(t-T)]*|U(i)|
LR test for inefficiency vs. OLS v only
Deg. freedom for sigma-squared(u):    1
Deg. freedom for heteroscedasticity:  0
Deg. freedom for truncation mean:     0
Deg. freedom for inefficiency model:  1
LogL when sigma(u)=0           809.67610
Chi-sq=2*[LogL(SF)-LogL(LS)] = 1161.048
Kodde-Palm C*: 95%: 2.706,  99%:  5.412
--------+--------------------------------------------------------------------
        |                  Standard            Prob.      95% Confidence
     YIT|  Coefficient       Error       z    |z|>Z*         Interval
--------+--------------------------------------------------------------------
        |Deterministic Component of Stochastic Frontier Model
Constant|    11.7882***      .00716  1646.05  .0000     11.7742   11.8022
      X1|     .62230***      .01365    45.59  .0000      .59555    .64905
      X2|     .06001***      .01069     5.61  .0000      .03905    .08096
      X3|     .05708***      .01454     3.93  .0001      .02858    .08557
      X4|     .35510***      .00700    50.69  .0000      .34137    .36883
        |Variance parameters for compound error
  Lambda|    2.67761***      .02351   113.88  .0000     2.63152   2.72369
Sigma(u)|     .19848***      .00060   332.72  .0000      .19731    .19965
        |Eta parameter for time varying inefficiency
     Eta|     .08030***      .00432    18.60  .0000      .07184    .08877
--------+--------------------------------------------------------------------
```

```
Warning   141: Iterations:current or start estimate of sigma is nonpositive
Normal exit from iterations. Exit status=0.
-----------------------------------------------------------------------------
Latent Class / Panel Frontier Model
Dependent variable                  YIT
Log likelihood function      1462.93500
Estimation based on N =   1482, K =  17
Sample is  6 pds and    247 individuals
Stoch. frontier (B&C,time varying U)
Ineff=u(i,t)=exp(-eta*(t-T))|U(i)|
Model fit with  2 latent classes.
--------+--------------------------------------------------------------------
        |                  Standard            Prob.      95% Confidence
     YIT|  Coefficient       Error       z    |z|>Z*         Interval
--------+--------------------------------------------------------------------
        |Model parameters for latent class 1
Constant|    11.8355***      .02201   537.84  .0000     11.7923   11.8786
      X1|     .60324***      .03499    17.24  .0000      .53467    .67181
      X2|     .13327***      .04014     3.32  .0009      .05459    .21195
      X3|     .10581***      .03248     3.26  .0011      .04216    .16947
      X4|     .33560***      .01392    24.11  .0000      .30832    .36288
        |Square root of variance sum, sqr(s2u + s2v)
   Sigma|     .71161**       .35935     1.98  .0477      .00730   1.41591
        |Asymmetry parameter in compound distn, su/sv
  Lambda|     .02071         .02565      .81  .4194    -.02956    .07098
        |Scale factor in time varying inefficiency
     Eta|     .19551***      .01986     9.84  .0000      .15658    .23444
        |Model parameters for latent class 2
Constant|    11.7611***      .01279   919.62  .0000     11.7360   11.7862
      X1|     .61866***      .01873    33.04  .0000      .58196    .65536
      X2|     .05041***      .01289     3.91  .0001      .02514    .07567
      X3|     .06232***      .01830     3.40  .0007      .02645    .09820
      X4|     .30614***      .01029    29.76  .0000      .28598    .32631
        |Square root of variance sum, sqr(s2u + s2v)
   Sigma|     .92839***      .02938    31.60  .0000      .87081    .98597
        |Asymmetry parameter in compound distn, su/sv
  Lambda|     .05084         .22185      .23  .8187    -.38398    .48566
        |Scale factor in time varying inefficiency
     Eta|     .07059***      .00475    14.87  .0000      .06129    .07990
        |Estimated prior probabilities for class membership
Class1Pr|     .30612***      .05178     5.91  .0000      .20463    .40760
Class2Pr|     .69388***      .05178    13.40  .0000      .59240    .79537
--------+--------------------------------------------------------------------
Note: ***, **, * ==>  Significance at 1%, 5%, 10% level.
-----------------------------------------------------------------------------
```

```
+-----------------------------------------------------+
| Stochastic Frontier Model Variance Parameters       |
| Class     Lambda      Sigma   Sigma(u)   Sigma(v) |
|     1    .020709    .711607    .014734    .711454 |
|     2    .050840    .928393    .047139    .927195 |
+-----------------------------------------------------+
=============================================================================
Predictions computed for the group with the largest posterior probability
Obs.  Periods Estimated inefficiencies, E[u|v -/+ u]
=============================================================================
Ind.=    1  J* = 1  P(j)=  .889  .111
       01-06  .3105  .2554  .2100  .1727  .1421  .1168
Ind.=    2  J* = 2  P(j)=  .295  .705
       01-06  .0813  .0757  .0706  .0657  .0613  .0571
Ind.=    3  J* = 2  P(j)=  .012  .988
       01-06  .2254  .2100  .1957  .1824  .1699  .1584
Ind.=    4  J* = 1  P(j)=  .955  .045
       01-06  .1778  .1463  .1203  .0989  .0814  .0669
Ind.=    5  J* = 1  P(j)=  .650  .350
       01-06  .2453  .2018  .1659  .1365  .1122  .0923
Ind.=    6  J* = 2  P(j)=  .138  .862
       01-06  .0517  .0482  .0449  .0418  .0390  .0363
Ind.=    7  J* = 1  P(j)=  .985  .015
       01-06  .3010  .2476  .2036  .1674  .1377  .1132
Ind.=    8  J* = 2  P(j)=  .165  .835
       01-06  .0561  .0523  .0487  .0454  .0423  .0394
Ind.=    9  J* = 2  P(j)=  .450  .550
       01-06  .0134  .0125  .0116  .0108  .0101  .0094
Ind.=   10  J* = 1  P(j)=  .999  .001
       01-06  .1039  .0855  .0703  .0578  .0475  .0391
 (Farms 11-247 omitted)
```