For discussion only

UNIVERSITY OF LEEDS

CAWG 8: Thursday 11th April 2019 – Glasgow

Professor Andrew Smith

For discussion only

UNIVERSITY OF LEEDS

Simple example

Ln Cost = a + b Ln CSV

Where CSV = MEAV^0.37*REPEXW^0.53

Ln Cost = a + b Ln (MEAV^0.37*REPEXW^0.53)

Ln Cost = a + 0.37b Ln MEAV + 0.53b Ln REPEXW

For discussion only

UNIVERSITY OF LEEDS

Illustrative output

```
-----------------------------------------------------------------------------
Ordinary     least squares regression ............
LHS=LNCOST   Mean                  =       11.04090
             Standard deviation    =         .19051
----------   No. of observations   =               8  DegFreedom   Mean square
Regression   Sum of Squares        =    .511594E-01           1        .05116
Residual     Sum of Squares        =        .202890           6        .03382
Total        Sum of Squares        =        .254050           7        .03629
----------   Standard error of e   =         .18389  Root MSE          .15925
Fit          R-squared             =         .20138  R-bar squared     .06827
Model test   F[  1,     6]         =        1.51292  Prob F > F*       .26473
--------+--------------------------------------------------------------------
        |                  Standard            Prob.      95% Confidence
  LNCOST|  Coefficient       Error       t    |t|>T*         Interval
--------+--------------------------------------------------------------------
Constant|    8.02844**      2.45000     3.28  .0169     2.03350  14.02338
   LNCSV|     .57594         .46824     1.23  .2647     -.56980   1.72169
--------+--------------------------------------------------------------------
***, **, * ==>  Significance at 1%, 5%, 10% level.
Model was estimated on Feb 25, 2019 at 10:01:04 AM
-----------------------------------------------------------------------------
```

```
-----------------------------------------------------------------------------
Restricted   least squares regression ............
LHS=LNCOST   Mean                  =       11.04090
             Standard deviation    =         .19051
----------   No. of observations   =               8  DegFreedom   Mean square
Regression   Sum of Squares        =    .511594E-01           1        .05116
Residual     Sum of Squares        =        .202890           6        .03382
Total        Sum of Squares        =        .254050           7        .03629
----------   Standard error of e   =         .18389  Root MSE          .15925
Fit          R-squared             =         .20138  R-bar squared     .06827
Model test   F[  1,     6]         =        1.51292  Prob F > F*       .26473
Restrictions F[  1,     5]         =         .06259  Prob F > F*       .81241
--------+--------------------------------------------------------------------
        |                  Standard            Prob.      95% Confidence
  LNCOST|  Coefficient       Error       t    |t|>T*         Interval
--------+--------------------------------------------------------------------
Constant|    8.02844**      2.45000     3.28  .0220     1.73051  14.32637
  LNMEAV|     .21310         .17325     1.23  .2734     -.23225    .65845
 LNREPEX|     .30525         .24817     1.23  .2734     -.33269    .94319
--------+--------------------------------------------------------------------
```

For discussion only

UNIVERSITY OF LEEDS

Totex CSV weights

0.369583	MEAV		
0.058605	Emergency CSV		
0.057604	External condition reports		
0.055867	Maintenance MEAV		
0.420355	Repex workload		
0.020751	Connections workload		
0.017235	Mains workload		

For discussion only

UNIVERSITY OF LEEDS

Totex model outputs [1]

```
-----------------------------------------------------------------------------
Ordinary     least squares regression ............
LHS=LNTOT    Mean                 =        5.17671
             Standard deviation   =         .23973
----------   No. of observations  =             40  DegFreedom   Mean square
Regression   Sum of Squares       =        2.11136           5        .42227
Residual     Sum of Squares       =        .129937          34        .00382
Total        Sum of Squares       =        2.24130          39        .05747
----------   Standard error of e  =         .06182  Root MSE          .05699
Fit          R-squared            =         .94203  R-bar squared     .93350
Model test   F[  5,    34]        =      110.49425  Prob F > F*       .00000
--------+--------------------------------------------------------------------
        |                  Standard            Prob.      95% Confidence
   LNTOT|  Coefficient       Error       t    |t|>T*         Interval
--------+--------------------------------------------------------------------
Constant|    -.13411         .23264     -.58  .5681     -.60689    .33866
  LNCSV1|     .78238***      .03377    23.17  .0000      .71375    .85102
     DY2|    -.00226         .03094     -.07  .9423     -.06514    .06063
     DY3|    -.04376         .03091    -1.42  .1660     -.10658    .01906
     DY4|    -.04831         .03092    -1.56  .1275     -.11115    .01453
     DY5|    -.09213***      .03093    -2.98  .0053     -.15498   -.02927
--------+--------------------------------------------------------------------
***, **, * ==>  Significance at 1%, 5%, 10% level.
Model was estimated on Apr 09, 2019 at 00:48:14 PM
-----------------------------------------------------------------------------
```

```
-----------------------------------------------------------------------------
Restricted   least squares regression ............
LHS=LNTOT    Mean                 =        5.17671
             Standard deviation   =         .23973
----------   No. of observations  =             40  DegFreedom   Mean square
Regression   Sum of Squares       =        2.11137           5        .42227
Residual     Sum of Squares       =        .129927          34        .00382
Total        Sum of Squares       =        2.24130          39        .05747
----------   Standard error of e  =         .06182  Root MSE          .05699
Fit          R-squared            =         .94203  R-bar squared     .93351
Model test   F[  5,    34]        =      110.50313  Prob F > F*       .00000
Restrictions F[  6,    28]        =        9.79611  Prob F > F*       .00001
--------+--------------------------------------------------------------------
        |                  Standard            Prob.      95% Confidence
   LNTOT|  Coefficient       Error       t    |t|>T*         Interval
--------+--------------------------------------------------------------------
Constant|    -.13419         .23263     -.58  .5687     -.61071    .34233
  LNMEAV|     .28916***      .01248    23.17  .0000      .26359    .31473
  LNEMER|     .04585***      .00198    23.17  .0000      .04180    .04991
   LNEXT|     .04507***      .00195    23.17  .0000      .04108    .04905
 LNMMEAV|     .04371***      .00189    23.17  .0000      .03985    .04758
   LNREP|     .32888***      .01420    23.17  .0000      .29980    .35796
 LNCONNW|     .01624***      .00070    23.17  .0000      .01480    .01767
 LNMAINW|     .01348***      .00058    23.17  .0000      .01229    .01468
     DY2|    -.00226         .03094     -.07  .9423     -.06564    .06113
     DY3|    -.04375         .03091    -1.42  .1679     -.10707    .01956
     DY4|    -.04831         .03092    -1.56  .1294     -.11164    .01503
     DY5|    -.09213***      .03093    -2.98  .0059     -.15548   -.02877
--------+--------------------------------------------------------------------
***, **, * ==>  Significance at 1%, 5%, 10% level.
Model was estimated on Apr 09, 2019 at 00:48:57 PM
-----------------------------------------------------------------------------
```

For discussion only

UNIVERSITY OF LEEDS

Totex model outputs [2]

```
----------------------------------------------------------------------------
Ordinary     least squares regression ............
LHS=LNTOT    Mean                 =        5.17671
             Standard deviation   =         .23973
----------   No. of observations  =             40  DegFreedom   Mean square
Regression   Sum of Squares       =        2.19938          11        .19994
Residual     Sum of Squares       =    .419232E-01          28        .00150
Total        Sum of Squares       =        2.24130          39        .05747
----------   Standard error of e  =         .03869  Root MSE          .03237
Fit          R-squared            =         .98130  R-bar squared     .97395
Model test   F[ 11,    28]        =      133.53976  Prob F > F*       .00000
--------+-------------------------------------------------------------------
        |                  Standard            Prob.      95% Confidence
   LNTOT|  Coefficient       Error       t    |t|>T*         Interval
--------+-------------------------------------------------------------------
Constant|   -3.45165***      .79843    -4.32  .0002    -5.08716  -1.81614
  LNMEAV|    -.10643         .26930     -.40  .6957     -.65806    .44521
  LNEMER|     .57414***      .20203     2.84  .0083      .16029    .98799
   LNEXT|    -.06725         .06607    -1.02  .3175     -.20259    .06809
 LNMMEAV|     .17389         .10721     1.62  .1160     -.04571    .39349
   LNREP|     .23894***      .06123     3.90  .0005      .11353    .36436
 LNCONNW|     .06230         .03863     1.61  .1180     -.01682    .14142
 LNMAINW|    -.05326**       .02391    -2.23  .0341     -.10223   -.00428
     DY2|    -.00672         .02034     -.33  .7435     -.04840    .03495
     DY3|    -.04543**       .02195    -2.07  .0478     -.09039   -.00047
     DY4|    -.06490***      .02144    -3.03  .0052     -.10881   -.02099
     DY5|    -.11640***      .02223    -5.24  .0000     -.16193   -.07086
--------+-------------------------------------------------------------------
***, **, * ==>  Significance at 1%, 5%, 10% level.
Model was estimated on Apr 09, 2019 at 00:48:57 PM
----------------------------------------------------------------------------
```

- The restrictions can be tested

For discussion only

UNIVERSITY OF LEEDS

Totex model outputs [2]

```
----------------------------------------------------------------------------
Ordinary     least squares regression ............
LHS=LNTOT    Mean                 =        5.17671
             Standard deviation   =         .23973
----------   No. of observations  =             40  DegFreedom   Mean square
Regression   Sum of Squares       =        2.19938          11        .19994
Residual     Sum of Squares       =    .419232E-01          28        .00150
Total        Sum of Squares       =        2.24130          39        .05747
----------   Standard error of e  =         .03869  Root MSE          .03237
Fit          R-squared            =         .98130  R-bar squared     .97395
Model test   F[ 11,    28]        =      133.53976  Prob F > F*       .00000
--------+-------------------------------------------------------------------
        |                  Standard            Prob.      95% Confidence
   LNTOT|  Coefficient       Error       t    |t|>T*         Interval
--------+-------------------------------------------------------------------
Constant|   -3.45165***      .79843    -4.32  .0002    -5.08716  -1.81614
  LNMEAV|    -.10643         .26930     -.40  .6957     -.65806    .44521
  LNEMER|     .57414***      .20203     2.84  .0083      .16029    .98799
   LNEXT|    -.06725         .06607    -1.02  .3175     -.20259    .06809
 LNMMEAV|     .17389         .10721     1.62  .1160     -.04571    .39349
   LNREP|     .23894***      .06123     3.90  .0005      .11353    .36436
 LNCONNW|     .06230         .03863     1.61  .1180     -.01682    .14142
 LNMAINW|    -.05326**       .02391    -2.23  .0341     -.10223   -.00428
     DY2|    -.00672         .02034     -.33  .7435     -.04840    .03495
     DY3|    -.04543**       .02195    -2.07  .0478     -.09039   -.00047
     DY4|    -.06490***      .02144    -3.03  .0052     -.10881   -.02099
     DY5|    -.11640***      .02223    -5.24  .0000     -.16193   -.07086
--------+-------------------------------------------------------------------
***, **, * ==>  Significance at 1%, 5%, 10% level.
Model was estimated on Apr 09, 2019 at 00:48:57 PM
----------------------------------------------------------------------------
```

- The restrictions can be tested

For discussion only

UNIVERSITY OF LEEDS

Reflections

- Implications – use the data to explore weightings?
- Presumably this has been tried before though…
- Understand the assumptions / restrictions underpinning the current approach in more depth
- For example:

Elasticity = MC / AC

Hence Elasticity * AC = MC

Check reasonableness of the implied MC estimates?

For discussion only

UNIVERSITY OF LEEDS

DEA method

- Non-parametric optimisation method
- We don’t estimate parameters – elasticities in log models
- So we are not assessing if a cost driver is statistically significant or not
- Specifies inputs (e.g. number of staff; or costs) and outputs (e.g. customer numbers)
- It then selects weights for each firm to compute its efficiency (giving the most favourable view of the firm)
- It may therefore give zero weight to some inputs or outputs – to maximise the ratio of say one output to one input
- Uses different weights for different firms

For discussion only

UNIVERSITY OF LEEDS

DEA arguments

- Multiple cost drivers / outputs – can be hard to get the model to distinguish different units
- Sometimes a two stage method used – but what goes in the two stages (also is using econometrics now anyway)
- Different weights across firms – and zero weights – general issue of transparency
- Basic version assumes CRS – VRS approach exists – but transparency?
- Dealing with random noise in the data? Could use judgement
- Key point for gas distribution - if just one variable (CSV) then just a simple unit cost comparison
- And if multiple drivers: would firms have faith in a model that applied zero weights and different weights to different companies?
- Always can be thought of as a candidate – and could be tried for completeness
- Limited usefulness – cost function more transparent

For discussion only

UNIVERSITY OF LEEDS

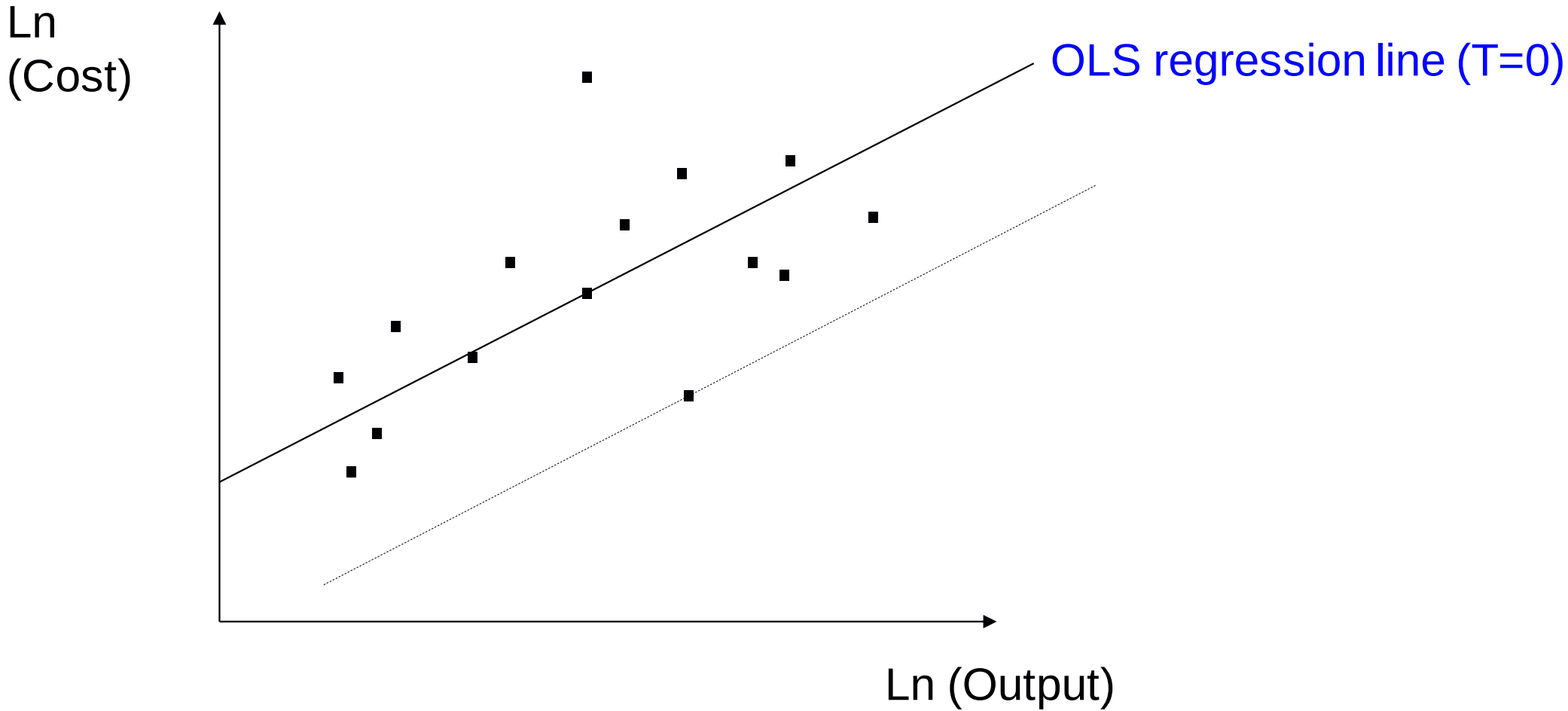
SFA

- Benefits could be overstated in regulation
- But recommend trying at least for completion
- Under the SFA assumption – non-correlation between error and regressors – OLS gives unbiased and consistent estimates of the parameters
- So potentially the issue here is about **decomposing the error term**
- Could use COLS plus UQ adjustment or other regulatory judgement
- Better or worse than the arbitrary distributional assumptions in SFA?
- Sample sizes relatively small
- Structured time varying model? Study convergence? Long panel. Dealing with time invariant differences between firms.
- Candidate model for further exploration

For discussion only

UNIVERSITY OF LEEDS

COLS



UNIVERSITY OF LEEDS

Contact Details

Prof Andrew Smith
Professor of Transport Performance and Economics
Editor, Journal of Transport Economics and Policy
Institute for Transport Studies
University of Leeds, Leeds LS2 9JT, UK

T +44 (0) 113 343 6654
E a.s.j.smith@its.leeds.ac.uk
W www.its.leeds.ac.uk/people/a.smith