Exam in: ECON 4160: Econometrics: Modelling and Systems Estimation

Day of exam: 19 December 2014

Time of day: 09:00—12:00

This is a 3 hour school exam.

Guidelines:
In the grading, question A will count 1/4, and question B will count 3/4 .

Question A (1/4)

Let $\mathbf{y}$ be a vector $(n \times 1)$ with n observations of a variable Y_i, and let $\mathbf{X}$ be a $n \times k$ matrix with observations of k explanatory variables. Consider the linear relationship

$$\mathbf{y} = \mathbf{X}\boldsymbol{\beta} + \boldsymbol{\varepsilon} \tag{1}$$

where $\boldsymbol{\varepsilon}$ is a $n \times 1$ vector with disturbances, and $\boldsymbol{\beta}$ is the $k \times 1$ vector with parameters.

1. Assume that the sample moments $(\boldsymbol{X}'\boldsymbol{X})^{-1}$ and $\boldsymbol{X}'\boldsymbol{y}$ are given by:

$$(\boldsymbol{X}'\boldsymbol{X})^{-1} = \begin{bmatrix} 1 & -10 \\ 10 & 1 \end{bmatrix}$$

and

$$\boldsymbol{X}'\boldsymbol{y} = \begin{bmatrix} 1 \\ 10 \end{bmatrix}$$

   (a) What are the OLS estimates of the parameters in this example?
   Answ: $\begin{bmatrix} 1 & -10 \\ 10 & 1 \end{bmatrix} \times \begin{bmatrix} 1 \\ 10 \end{bmatrix} = \begin{bmatrix} -99 \\ 20 \end{bmatrix}$. $\hat{\beta}_1 = -99$, $\hat{\beta}_2 = 20$

   (b) Explain why these estimates are identical to the estimates you would have obtained if you had used method of moments as the estimation principle instead.
   Answ: The MM principle entails the normal equations $(\boldsymbol{X}'\boldsymbol{X})\hat{\boldsymbol{\beta}} = \boldsymbol{X}'\boldsymbol{y}$

2. Assume that you get data for a new variable, Z, where the n observations are collected in the $n \times 1$ vector $\mathbf{z}$. Assume that all the k elements in the vector $(\boldsymbol{X}'\boldsymbol{X})^{-1}\boldsymbol{X}'\mathbf{z}$ are practically zero.

   (a) How will the OLS estimate of $\boldsymbol{\beta}$ be affected if you decide to include Z as a new regressor in equation (1)?

   Answ: The estimates will be practically unaffected, since $(\boldsymbol{X}'\boldsymbol{X})^{-1}\boldsymbol{X}'\mathbf{z}$ $= \hat{\boldsymbol{\gamma}} \approx \mathbf{0}$ where $\hat{\boldsymbol{\gamma}}$ is the OLS regression coefficient vector between $\mathbf{z}$ and $\boldsymbol{X}$.

   (b) An economic theorist suggests that it would be a good idea to use Z as an instrumental variable "in the regression (1)" in order to avoid suspected simultaneity bias. How would you (as an econometrician) respond to his suggestion?

   Answ: Not a good idea since Z_i is a poor instrument for the variables in $\boldsymbol{X}$. (Extra point if add, flippantly, that if he by "in the regression" means that his parameter of interest are the parameters of the conditional expectation, then there is no simultaneity bias in the first place)

Question B (3/4)

In this question, we study the relationship between Norwegian exports of traditional goods and an export market indicator. The data set is quarterly, and the variables are seasonally unadjusted.

The export variable is denoted $ATRAD_t$, and the market indicator is denoted MII_t. Both are measured in real terms, i.e., they are volumes. We will use the log-transformed time series, which we denote $LATRAD_t$ and $LMII_t$ and the differenced series $DLATRAD_t = LATRAD_t - LATRAD_{t-1}$ and $DLII_t = LII_t - LII_{t-1}$. For reference, Figure 1 shows the time plots of the two level variables and their differences.

1. Table 1 shows Augmented Dickey Fuller test statistics. Explain how you can use the results to test the unit-root hypothesis for $LATRAD_t$ and LII_t, and give your conclusions.

   Answ: Correct use of the results in Table 1 (perhaps supported by the time plots in Figure 1) is that $I(1)$ cannot be rejected for either of the variables. One possible pitfall is to conclude, that t-adf without lags reject $I(1)$ for $LATRAD$ with no augmentation. But since the table clearly shows that there are significant lags at the first and third lags, that test statistic has the wrong size for finite samples.

2. We define the vector $\boldsymbol{y}_t$ as $\boldsymbol{y}_t = (\ ATRAD_t \quad LMII_t\ )'$. The vector with differenced variables is $\Delta \boldsymbol{y}_t = (\ DATRAD_t \quad DLMII_t\ )'$. We analyse the differenced data and formulate a bivariate first order VAR as

$$\Delta \boldsymbol{y}_t = \boldsymbol{A} \Delta \boldsymbol{y}_{t-1} + \boldsymbol{C}\boldsymbol{G}_t + \boldsymbol{\varepsilon}_t \tag{2}$$

where $\boldsymbol{A}$ and $\boldsymbol{C}$ are matrices with parameters, and $\boldsymbol{G}_t$ is a matrix with deterministic variables: An intercept, three (centered) seasonals and one or more indicator variables (impulse dummies) for structural breaks. The vector $\boldsymbol{\varepsilon}_t$ with VAR disturbances is assumed to follow a bivariate normal distribution.

Table 2 shows estimation results for (2) when $\boldsymbol{G}_t$ includes three indicator variables, for 1980q2, 2008q4 and 2009q1.

(a) The roots of the estimated companion matrix of this system are -0.55 and 0.41 (two real roots). What does this indicate about the dynamic stability of the system, and does it confirm or contradict your conclusion in Question B.1?

Answ: This is consistent with $\Delta \boldsymbol{y}_t \sim I(0)$ implying that $\boldsymbol{y}_t \sim I(1)$ which is what we found in B1.

(b) Since the disturbances in $\boldsymbol{\varepsilon}_t$ are generally correlated, they are not structural disturbances. A fellow student claims that the structural disturbances can be achieved by putting zero restrictions on one of the off-diagonal elements in $\boldsymbol{A}$ (so that the matrix becomes either upper triangular or lower triangular). Explain why this not correct.

Answ: This restriction is not implied by Cholesky-decomposition, which applies to the contemporaneous matrix $\boldsymbol{B}$ in the SEM form ($\boldsymbol{B}\Delta \boldsymbol{y}_t = \boldsymbol{A}^* \Delta \boldsymbol{y}_{t-1} + \boldsymbol{C}^* \boldsymbol{G}_t + \varepsilon_t^*$) Even if $\boldsymbol{B}$ is trangular, implying that $\boldsymbol{B}^{-1}$ is triangular, $\boldsymbol{A} = \boldsymbol{B}^{-1}\boldsymbol{A}^*$ is not necessarily triangular. The student's proposal is (instead) a restriction on the VAR that does not ortogonalize the two residuals. The impulse responses of her suggested (restricted) VAR are not indentified.

(c) Consider the estimated model in Table 3, and explain why this is an example of identification of the structural impulse response functions by the use of the Cholesky-decomposition.

Answ: In this model, we have a triangular $\boldsymbol{B}$ and the disturbances are uncorrelated (implemented here by use of 1SLS on each of the equations). In fact, the $DLATRAD$-equation in Table 3 is the conditional equation, while the $DLMII$-equation is the marginal equation. The log-likelihod is identical to Table 2, consistent with exact identification.

(d) Another identification scheme assumes, first, that there should be no seasonal indicator variables in the structural equation for $DLMII$, since $LMII$ is a variable which is a broad average of seasonally adjusted GDP data series (of Norway's trading partner). Second, since the three dummies for 1980q2, 2008q4 and 2009q1 represent structural changes in export marked growth, they should not be included in the structural equation for $DLATRAD_t$. Explain why this identification scheme implies that the degree of over-identification is 4, as shown in Table 4.

Answ: In this case we do not assume orthogonal disturbances. Hence, we can use the rank-and-order condition on the SEM-version of the model. Using order (and assuming rank satisfied) both structural equations are over-identified, and that the degree of overidentification is 2 in each equation, 4 in all

(e) Show how the test of over-identifying restrictions in Table 4 can be calculated, using the information provided in the tables. Explain why the test of over-identifying restrictions can be interpreted as an encompassing test.

Answ: This is a Likelihood-Ratio test calcualted as $-2*(702.885884-703.614779) = 1.4578$. Under the null-hypothesis that the restrictions are acceptable, the LR-statistic has an asymptototic Chi-square distribution with 4 degrees of fredom. Over id. restrictions can sometimes lead to a significant loss of explanatory power relative to the VAR. Other times, as here, the loss is not sginficant and the the structural model then encompasses the statitical model.

(f) What is the estimated impact effect on $DLATRAD$ of a positive shock to $DLMII$, when you use the identified model in Table 3, and what is the estimate if you instead consider the over-identified model in Table 4.

Answ: Table 3: 0.38. Table 4: The direct effect of a unit change in $\epsilon_{LMII,t}$ is 0.748048. But $\epsilon_{LA,t}$ is correlated with $\epsilon_{LMII,t}$. Can express this as $\epsilon_{LA,t} = -\beta\epsilon_{LMII,t}$ where β is the regression coefficient. We can estimate it (using the esimated standard errors and correlations coefficient) as $r^2\frac{\hat{\sigma}_{LA}}{\hat{\sigma}_{LMII}} = (0.15532) * \frac{0.043474}{0.0111} = -0.60832$. Hence $\frac{\partial LATRAD}{\partial\epsilon_{LMII,t}} = 0.748048 - (0.15532) * \frac{0.043474}{0.0111} = 0.13973$.

3. Table 5 shows a single equation model of $DLATRAD_t$ estimated by OLS, where we for simplicity have omitted $DLMII_{t-1}$. Figure 2 contains some relevant recursive graphs for this model. Consider the coefficient of $DLMII_t$ as the parameter of interest. Would you say that there is evidence of super-exogeneity of $DLMII_t$ with respect to the parameter of interest? Give a brief motivation for your answer (without formal tests)

   Answ: Good answers should start with a relatively clear definition of weak exogenity and the concept of invariance which together give super exogenity. The recursive graphs indicate that the regression coefficient of $DLMII_t$ in the export equation in relatively stable in the period where we have the structural breaks in the $DLMII$-equation, in 2008q4 and 2009q1. The Chow-test also suggest recursive stability over this period. Hence, although informal, the recursive graphs give evidence that $DLMII_t$ is super-exogenous with respect to these two breaks.

4. Finally, we consider the possibility that $LATRAD_t$ is cointegrated with $LMII_t$ and a third $I(1)$ variable $LREX_t$, which is the natural logarithm of the Norwegian real exchange rate. Use the results in Table 6 to test formally the null of no cointegration, against the alternative of a single cointegration relationship. Relevant critical values are -3.62 (1 %), -3.00 (5 %) and -2.26 (10 %). If you conclude with rejection of the null hypothesis, what is the estimated cointegration relationship?

   Answ: Good answers should explain that under the null of no-cointegration, the OLS the "t-value" of the coefficient of $LATRAD_{t-1}$ has a non-standard distribution of the Dicky-Fuller type. Because we have two conditioning $I(1)$-variables the critical values are moved to the left compared to the ordinary DF-distribution.

The reported t-values support that non-cointegration can be rejected at the 5 % level.

5. An important theoretical result in time series econometrics says that the method used in Question B.4 is statistically efficient only if $LMII_t$ and $LREX_t$ are weakly exogenous variables with respect to the cointegration parameters. Try to summarize in a few sentences a method that would allow you to the hypothesis of no cointegration between $LATRAD$, $LMII$ and $LREX$ to be tested without making the assumption of weak exogeneity.

   Answ: WE is here the assumption is that neither $LMII_t$ nor $LREX_t$ equilibrium correct with respect to the cointegration vector. The Johansen method does not assume this type of WE from the outset. Instead the method tests the hypotheses that the best linear combination of $LATRAD_{t-1}$, $LMII_{t-1}$ and $LREX_{t-1}$ has no predictive power for the vector ($\Delta LATRAD_t$, $\Delta LMII_t$, $\Delta LREX_t$). *without* putting zero-restctions on the equilibirum correction coefficient. The most common test-statistic is the Trace-test. Under the null of no-cointegration has a non-standard distrubution. If no-cointegration is rejected the WE hypothesis can easily be tested using conventional LR tests. The trace test also allows multiple (here two) cointegration relationships.

Graphs and estimation results for question B.

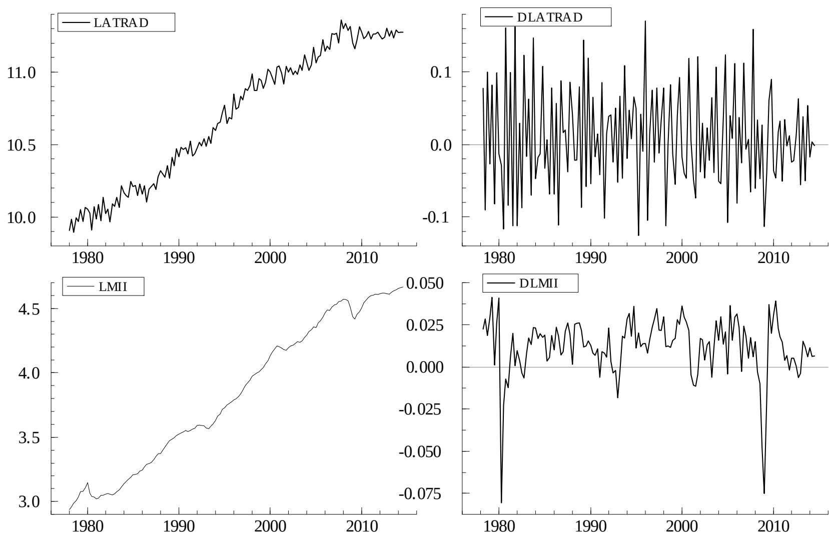

Figure 1: Time plots of the export and marked indicator data.

```
Unit-root tests
The dataset is: C:\SW20\ECON4160\H2014\exam\ATRAD.in7
The sample is: 1979(1) - 2014(3) (147 observations and 2 variables)

LATRAD: ADF tests (T=143, Constant+Trend; 5%=-3.44 1%=-4.02)
D-lag    t-adf      beta Y_1    sigma   t-DY_lag  t-prob       AIC  F-prob
 3    -0.2352       0.98687  0.04888    -6.696  0.0000     -5.996
 2     -1.766       0.89061  0.05610   -0.8195  0.4139     -5.727  0.0000
 1     -2.071       0.87674  0.05604    -7.020  0.0000     -5.736  0.0000
 0     -4.799**     0.69990  0.06499                       -5.446  0.0000

LMII: ADF tests (T=143, Constant+Trend; 5%=-3.44 1%=-4.02)
D-lag    t-adf      beta Y_1    sigma   t-DY_lag  t-prob       AIC  F-prob
 3     -2.656       0.94712  0.01547     1.470  0.1439     -8.296
 2     -2.360       0.95421  0.01554    0.7365  0.4627     -8.295  0.1439
 1     -2.255       0.95748  0.01551     6.340  0.0000     -8.305  0.2617
 0     -1.142       0.97592  0.01755                       -8.065  0.0000
```

Table 1: Tests of the null hypothesis that $LATRAD$ and $LMII$ are integrated of order 1 ($I(1)$)

```
SYS( 1) Estimating the system by OLS
       The dataset is: C:\SW20\ECON4160\H2014\exam\ATRAD.in7
       The estimation sample is: 1978(3) - 2014(3)

URF equation for: DLATRAD
                 Coefficient  Std.Error  t-value  t-prob
DLATRAD_1           -0.553622    0.07135    -7.76  0.0000
DLMII_1              0.638051     0.2259     2.82  0.0055
I:1980q2           -0.0535959    0.04478    -1.20  0.2334
I:2008q4           -0.0521030    0.04456    -1.17  0.2444
I:2009q1           -0.0815688    0.04639    -1.76  0.0810
Constant      U    0.00811030   0.004566     1.78  0.0779
CSeasonal     U    -0.0453698    0.01344    -3.38  0.0010
CSeasonal_1   U    -0.0675976    0.01046    -6.46  0.0000
CSeasonal_2   U    -0.0917899    0.01087    -8.45  0.0000

sigma = 0.0436634   RSS = 0.2592829231

URF equation for: DLMII
                 Coefficient  Std.Error  t-value  t-prob
DLATRAD_1         -0.00543558    0.01833   -0.297  0.7673
DLMII_1              0.413869    0.05804     7.13  0.0000
I:1980q2            -0.106701    0.01150    -9.28  0.0000
I:2008q4           -0.0528273    0.01145    -4.61  0.0000
I:2009q1           -0.0640580    0.01192    -5.37  0.0000
Constant      U    0.00846113   0.001173     7.21  0.0000
CSeasonal     U -3.07642e-005   0.003453 -0.00891  0.9929
CSeasonal_1   U   0.000290695   0.002686    0.108  0.9140
CSeasonal_2   U   -0.00178677   0.002792   -0.640  0.5233

sigma = 0.011218   RSS = 0.01711469135

log-likelihood      703.614779  -T/2log|Omega|     1115.10695
|Omega|         2.0903052e-007  log|Y'Y/T|        -14.0665117
no. of observations        145  no. of parameters          18

correlation of URF residuals (standard deviations on diagonal)
                  DLATRAD       DLMII
DLATRAD           0.043663     0.098072
DLMII             0.098072     0.011218

Vector AR 1-5 test:      F(20,250) =   1.8691 [0.0751]
Vector Normality test:   Chi^2(4)  =   1.5407 [0.8194]
Vector Hetero test:      F(21,379) =  0.95430 [0.5204]
```

Table 2: Estimation results for a VAR of $DLATRAD$ and $DLMII$

```
MOD(2) Estimating the model by 1SLS
       The dataset is: C:\SW20\ECON4160\H2014\exam\ATRAD.in7
       The estimation sample is: 1978(3) - 2014(3)

Equation for: DLATRAD
                 Coefficient  Std.Error  t-value  t-prob
DLMII               0.381721     0.3334     1.15  0.2542
DLATRAD_1          -0.551548    0.07129    -7.74  0.0000
DLMII_1             0.480069     0.2645     1.81  0.0717
I:1980q2          -0.0128658    0.05715   -0.225  0.8222
I:2008q4          -0.0319377    0.04787   -0.667  0.5058
I:2009q1          -0.0571165    0.05102    -1.12  0.2649
Constant          0.00488051   0.005363    0.910  0.3644
CSeasonal         -0.0453580    0.01342    -3.38  0.0009
CSeasonal_1       -0.0677085    0.01044    -6.48  0.0000
CSeasonal_2       -0.0911079    0.01087    -8.38  0.0000

sigma = 0.0434529

Equation for: DLMII
                 Coefficient  Std.Error  t-value  t-prob
DLATRAD_1        -0.00543558    0.01833   -0.297  0.7673
DLMII_1             0.413869    0.05804     7.13  0.0000
I:1980q2           -0.106701    0.01150    -9.28  0.0000
I:2008q4          -0.0528273    0.01145    -4.61  0.0000
I:2009q1          -0.0640580    0.01192    -5.37  0.0000
Constant          0.00846113   0.001173     7.21  0.0000
CSeasonal       -3.07642e-005  0.003453 -0.00891  0.9929
CSeasonal_1      0.000290695   0.002686    0.108  0.9140
CSeasonal_2      -0.00178677   0.002792   -0.640  0.5233

sigma = 0.011218

log-likelihood     703.614779  -T/2log|Omega|     1115.10695
no. of observations       145  no. of parameters          19
No restrictions imposed

correlation of structural residuals (standard deviations on diagonal)
                  DLATRAD        DLMII
DLATRAD          0.043453      0.00000
DLMII             0.00000     0.011218
```

Table 3: Estimation results for an identified version of the VAR in Table 2

```
MOD(3) Estimating the model by FIML
       The dataset is: C:\SW20\ECON4160\H2014\exam\ATRAD.in7
       The estimation sample is: 1978(3) - 2014(3)

Equation for: DLATRAD
                 Coefficient  Std.Error  t-value  t-prob
DLMII               0.748048     0.3287     2.28  0.0244
DLATRAD_1          -0.548349    0.07063    -7.76  0.0000
DLMII_1             0.397331     0.2597     1.53  0.1284
Constant         0.000814805   0.004849    0.168  0.8668
CSeasonal         -0.0457105    0.01311    -3.49  0.0007
CSeasonal_1       -0.0659676    0.01023    -6.45  0.0000
CSeasonal_2       -0.0903711    0.01073    -8.42  0.0000

sigma = 0.0434738

Equation for: DLMII
                 Coefficient  Std.Error  t-value  t-prob
DLATRAD            0.0155810    0.02807    0.555  0.5798
DLATRAD_1         0.00331812    0.02222    0.149  0.8815
DLMII_1             0.400981    0.06172     6.50  0.0000
I:1980q2           -0.103934    0.01156    -8.99  0.0000
I:2008q4          -0.0528994    0.01112    -4.76  0.0000
I:2009q1          -0.0638377    0.01161    -5.50  0.0000
Constant          0.00836328   0.001170     7.15  0.0000

sigma = 0.0111112

log-likelihood     702.885884  -T/2log|Omega|     1114.37806
no. of observations       145  no. of parameters          14

LR test of over-identifying restrictions: Chi^2(4)  =   1.4578 [0.8341]

correlation of structural residuals (standard deviations on diagonal)
                  DLATRAD        DLMII
DLATRAD           0.043474     -0.15532
DLMII             -0.15532     0.011111
```

Table 4: FIML estimation results for an identified version of the VAR in Table 2

```
EQ(1) Modelling DLATRAD by OLS
      The dataset is: C:\SW20\ECON4160\H2014\exam\ATRAD.in7
      The estimation sample is: 1978(3) - 2014(3)

                  Coefficient  Std.Error  t-value  t-prob Part.R^2
DLMII                0.904256     0.2104     3.69  0.0003   0.0891
DLATRAD_1           -0.520404    0.07026    -7.41  0.0000   0.2830
Constant           0.00495056   0.004420     1.12  0.2646   0.0089
CSeasonal          -0.0491638    0.01328    -3.70  0.0003   0.0898
CSeasonal_1        -0.0669509    0.01035    -6.47  0.0000   0.2312
CSeasonal_2        -0.0916259    0.01082    -8.47  0.0000   0.3402

sigma               0.0438589  RSS              0.267381208
R^2                  0.621887  F(5,139) =    45.72 [0.000]**
Adj.R^2              0.608286  log-likelihood           250.7
no. of observations       145  no. of parameters            6
mean(DLATRAD)       0.0088978  se(DLATRAD)          0.0700767
```

Table 5: Estimation results for a conditional model of $DLATRAD_t$

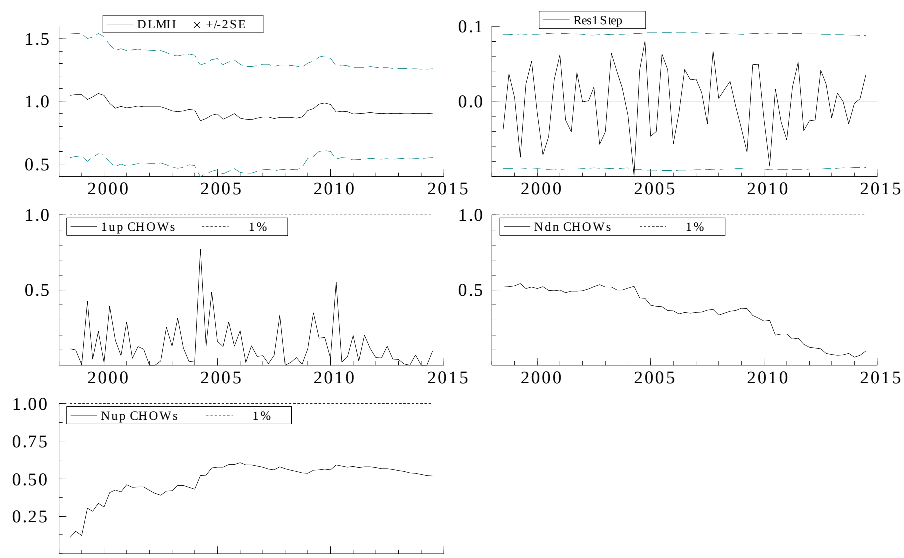

Figure 2: Recursve plots of for the model in Table 5.

```
EQ(2) Modelling DLATRAD by OLS
      The dataset is: C:\SW20\ECON4160\H2014\exam\ATRAD.in7
      The estimation sample is: 1978(3) - 2014(3)
```

	Coefficient	Std.Error	t-value	t-prob	Part.R^2
DLATRAD_1	-0.418093	0.07860	-5.32	0.0000	0.1754
Constant	1.79707	0.5304	3.39	0.0009	0.0795
DLMII	0.665952	0.2301	2.89	0.0044	0.0593
DLMII_1	0.389232	0.2321	1.68	0.0958	0.0207
DLREX	0.221008	0.1884	1.17	0.2429	0.0102
DLREX_1	-0.245433	0.1878	-1.31	0.1934	0.0127
LATRAD_1	-0.248506	0.07283	-3.41	0.0009	0.0805
LMII_1	0.217816	0.06436	3.38	0.0009	0.0793
LREX_1	0.213545	0.09305	2.30	0.0233	0.0381
CSeasonal	-0.0435038	0.01295	-3.36	0.0010	0.0782
CSeasonal_1	-0.0594354	0.01037	-5.73	0.0000	0.1980
CSeasonal_2	-0.0875664	0.01063	-8.24	0.0000	0.3379

sigma	0.0421299	RSS	0.236065277
R^2	0.666172	F(11,133) =	24.13 [0.000]**
Adj.R^2	0.638562	log-likelihood	259.732
no. of observations	145	no. of parameters	12
mean(DLATRAD)	0.0088978	se(DLATRAD)	0.0700767

Table 6: OLS estimates for an unrestricted equilibirum correction equation for $LATRAD_t$