Problem Set 1

Solutions

1

1.1 CLRM assumptions, see Greene 6th ed., 2008, pp. 11

linearity $y = \mathbf{X}\beta + \epsilon$, full rank (no linear relationship among the indepent variables), exogenity of the independent variables $E[\epsilon_i|x_{j1}, x_{j2}, ...x_{jK}] = 0$, homoscedasticity $Var(\epsilon_i|\mathbf{X}) = \sigma^2 < \infty$, non-autocorrelation $Cov[\epsilon_i, \epsilon_j|\mathbf{X}] = 0$, data generation (elements in $\mathbf{X}$ are fixed constants or random draws from a stochastic process), normal distribution $e \sim N(0, \sigma^2 < \infty)$
The assumptions of homoscedasticity and of non-autocorrelation imply

$$E[\epsilon\epsilon'|\mathbf{X}] \;=\; \begin{bmatrix} E[\epsilon_1\epsilon_1|\mathbf{X}] & E[\epsilon_1\epsilon_2|\mathbf{X}] & ... & E[\epsilon_1\epsilon_n|\mathbf{X}] \\ E[\epsilon_2\epsilon_1|\mathbf{X}] & E[\epsilon_2\epsilon_2|\mathbf{X}] & ... & E[\epsilon_2\epsilon_n|\mathbf{X}] \\ \vdots & \vdots & \ddots & \vdots \\ E[\epsilon_n\epsilon_1|\mathbf{X}] & E[\epsilon_n\epsilon_2|\mathbf{X}] & ... & E[\epsilon_n\epsilon_n|\mathbf{X}] \end{bmatrix} = \begin{bmatrix} \sigma^2 & 0 & ... & 0 \\ 0 & \sigma^2 & ... & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & ... & \sigma^2 \end{bmatrix} = \sigma^2\mathbf{I} \tag{1}$$

1.2 derive OLS estimator, see Greene 6th ed., 2008, pp. 21

idea: "minimize the sum of squared residuals"

$$\text{minimize}_b \sum_{i=1}^{n}(y_i - x_i'b)^2 = \sum_{i=1}^{n} e_i^2 \Leftrightarrow \text{minimize}_b \; e'e = (y - \mathbf{X}b)'(y - \mathbf{X}b) \tag{2}$$

taking the derivative of (1) with respect to b, equaling zero and solving for b yields

$$b = (X'X)^{-1}X'y \tag{3}$$

substituting $y = \mathbf{X}\beta + \epsilon$ and taking expectations yields

$$E[b] = \beta + \underbrace{(X'X)^{-1}X'E[\epsilon]}_{=0} \tag{4}$$

from (4) it follows that $E[b] = \beta$ and thus we have obtained an unbiased estimator.

1.3 Matlab program for OLS

```
K=1
T=500
beta=2
x=ones(T,1)
sig2=0.7
b_sim = []
sim = 1000
for i=1:1:sim
e=sqrt(sig2)*randn(T,1)
y = x*beta+e
b = inv(x'*x)*x'*y
b_sim = [b; b_sim]
end
betabar = mean(b_sim,1)
```

1.4 covariance of b

$$Cov[b] = E[(b - E[b])(b - E[b])' = E[(b - \beta)(b - \beta)'] \tag{5}$$

using $b = \beta + (X'X)^{-1}X'\epsilon$, it follows:

$$Cov[b] = E[(X'X)^{-1}X'\epsilon\epsilon'X(X'X)^{-1}] = (X'X)^{-1}X'E[\epsilon\epsilon']X(X'X)^{-1} \tag{6}$$

inserting $E[\epsilon\epsilon'] = \sigma^2\mathbf{I}$, we obtain

$$Cov[b] = \sigma^2(X'X)^{-1} \tag{7}$$

1.5 properties of b

i) unbiasedness $E[b] = \beta$
ii) consistency $plim(b = \beta)$
iii) efficiency (among all unbiased and consistent linear estimators, b has the smallest variance)

1.6 estimator for $Var(e)$

$$e = y - Xb = y - X(X'X)^{-1}X'y = (\mathbf{I}_n - X(X'X)^{-1}X')(X\beta + \epsilon) = M\epsilon \tag{8}$$

where M is an idempotent matrix. Then

$$E[e'e] = E[(M'\epsilon')(M\epsilon)] = E[\epsilon' M\epsilon] = tr(E[M\epsilon\epsilon'] = (y - \beta X)'(y - \beta X)tr(M) \tag{9}$$

and because $tr(M) = n - K$ we obtain

$$E[e'e] = \sigma^2(n - K) \Leftrightarrow \sigma^2 = E[e'e]/(n - K) \tag{10}$$

and if $\hat{\sigma}^2 = e'e/(n - K)$ then

$$E[\hat{\sigma}^2] = E[e'e]/(n - K) = \sigma^2 \tag{11}$$

where K is number of independent variables.

1.7 derive R^2

by definition R^2 is the following

$$R^2 = \frac{regression\ sum\ of\ squares}{total\ sum\ of\ squares} \tag{12}$$

where the total sum of squares $SST = \sum_{i=1}^{n}(y_i - \bar{y})^2 = y'y - n\bar{y}^2$ and $SSR = \sum_{i=1}^{n}(\hat{y}_i - \bar{y})^2 = \hat{y}'\hat{y} - n\bar{y}^2$

R^2 will never decrease when another variable is added to a regression model. It is thus desirable to include a penalty for each additional explanatory variable that is added to the model.

1.8 likelihood functions

intuition: "find the parameters that give the highest probability (ex post) of having generated the data we actually observed"

We assume that $\epsilon \sim N(0, \sigma^2\mathbf{I}_n)$ and therefore $y \sim N(X\beta, \sigma^2\mathbf{I}_n)$. The PDF of a randomly selected observation y_i is

$$f(y_i|x_i, \beta, \sigma^2) = \frac{1}{\sqrt{2\pi\sigma^2}} exp[\frac{-(y_i - x_i'\beta)^2}{2\sigma^2}] \tag{13}$$

because the observations are independent, we can multiply the PDFs for each i to obtain the joint PDF of all observations. This joint density is the likelihood function. It is usually simpler to work with the log of the likelihood function

$$ln\ L(\beta, \sigma^2|y, X) = \sum_{i=1}^{n} ln\ f(y_i|x_i, \beta, \sigma^2) \tag{14}$$

then the estimator b can be obtained by maximizing (14) with respect to β

$$\frac{\partial[ln\ L])}{\partial\beta} = 0 \Leftrightarrow b = X'y(X'X)^{-1} \tag{15}$$

and since the obtained estimator is the same as in OLS, from (3) it follows that $E[b] = \beta$. To derive an estimator for σ^2, we simply differentiate (14) with respect to σ^2

$$\frac{\partial[ln\ L])}{\partial\sigma^2} = 0 \Leftrightarrow \sigma^2 = \frac{\epsilon'\epsilon}{n} \tag{16}$$

and using b it follows that $\tilde{\sigma}^2 = \frac{e'e}{n} \neq \hat{\sigma}^2 = e'e/(n-K)$ is biased.

1.9 Matlab program for MLE

```
K = 1
T = 500
beta=2
x = ones(T,1)
sig2 = 0.7

opts = optimset('DerivativeCheck','off','Display','off','TolX',1e-6,'TolFun',1e-6,
'Diagnostics','off','MaxIter',100,'LargeScale','on')

LB = [0.001;-10]
UB = [10;10]
startpar = [0.001;0]
b_sim = []
sim = 500

for i=1:1:sim
e = sqrt(sig2)*randn(T,1)
y = x*beta+e
data = [y x]
[params,FVAL,EXITFLAG,OUTPUT,LAMBDA,GRAD,hesse]=fmincon
(@likex,startpar,[],[],[],[],LB,UB,[],opts,data)
b = params(2:end)
b_sim = [b; b_sim]
end

betabar = mean(b_sim,1)
[f1,d1] = ksdensity(b_sim(:,1))
plot(d1,f1)
```

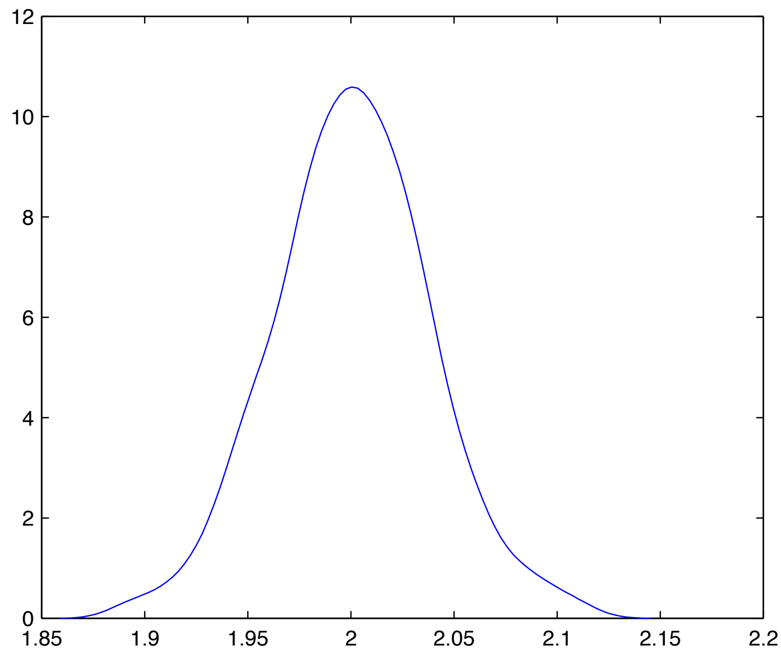

Figure 1: Density of betas

1.10 properties of MLE

i) b is normally distributed, and the best unbiased estimator (see OLS)
ii) $\hat{\sigma}^2$ is the best estimator of the variance and is distributed $\chi^2(n-K)$
iii) $Cov(b, \hat{\sigma}^2) = 0$

1.11 t-ratio

From $\tilde{\beta} \sim N(\beta, \sigma^2(x'x)^{-1})$, we can derive

$$\frac{R(\tilde{\beta}-\beta)}{\sigma\sqrt{R(x'x)^{-1}R'}} = z \sim N(0,1) \tag{17}$$

Because σ^2 is not know, we use it's unbiased estimator $\hat{\sigma}^2$.Then using $\hat{\sigma}^2$ for σ^2 and multiplying (17) by $\frac{1}{\sqrt{\hat{\sigma}^2(T-K)}}$ yields

$$\frac{R(\tilde{\beta}-\beta)}{\sigma\sqrt{R(x'x)^{-1}R'}} \cdot \frac{1}{\sqrt{\hat{\sigma}^2(T-K)/\sigma^2(T-K)}} \sim N(0,1) \cdot \frac{1}{\sqrt{\chi^2_{(T-K)}/(T-K)}} \sim t(T-K) \tag{18}$$

2

2.1 AR(2)-Matlab program

```
% Auto Regressive Order 2
load tsdata
p=tsdata(:,1)%1 for us, 3 for uk, 5 for ger
lp=log(p)
delta_lp = lp(2:end,:)-lp(1:end-1,:)
c = ones(358,1)
y=delta_lp(3:end,:)
lag1=delta_lp(2:end-1,:)
lag2=delta_lp(1:end-2,:)
x=[c lag1 lag2]
betahat = inv(x'*x)*x'*y
ehat=y-x*betahat
sig2e=(ehat'*ehat)/(358-2)
cov=sig2e*inv(x'*x)
tstat=betahat./sqrt(diag(cov))
betahat_tstat=[betahat tstat]
[AIC_HQ_SC] = infocrit(sig2e,358,2)
R2=1-(ehat'*ehat)/((y-mean(y))'*(y-mean(y)))
R2adj = 1 - ((358-1)/(358-size(x,2)))*(1-R2)
[scorr]= LBQ(ehat,3,0,2)
[JB] = JABE(ehat)
[archlm_prob] = archlm(ehat,2)
[RESET_pvalue] = resettest(ehat,x,betahat,2)


% Auto Regressive Order 1
load tsdata
p=tsdata(:,1)%1 for us, 3 for uk, 5 for ger
lp=log(p)
delta_lp = lp(2:end,:)-lp(1:end-1,:)
c = ones(359,1)
y=delta_lp(2:end,:)
lag1=delta_lp(1:end-1,:)
x=[c lag1]
betahat = inv(x'*x)*x'*y
ehat=y-x*betahat
sig2e=(ehat'*ehat)/(359-1)
cov=sig2e*inv(x'*x)
tstat=betahat./sqrt(diag(cov))
betahat_tstat=[betahat tstat]
```

```
[AIC_HQ_SC] = infocrit(sig2e,359,1)
R2=1-(ehat'*ehat)/((y-mean(y))'*(y-mean(y)))
R2adj = 1 - ((359-1)/(359-size(x,2)))*(1-R2)
[scorr]= LBQ(ehat,3,0,1)
[JB] = JABE(ehat)
[archlm_prob] = archlm(ehat,2)
[RESET_pvalue] = resettest(ehat,x,betahat,2)
```

Table 1: AR(2) Regression Results

country	US	UK	GER
$\phi_0(t-rat.)$	0.0083 (3.4532)	0.0099 (3.7584)	0.0056 (2.0263)
$\phi_1(t-rat.)$	-0.0251 (-0.4756)	-0.0013 (-0.0246)	0.0623 (1.1740)
$\phi_2(t-rat.)$	-0.0935 (-1.7714)	-0.1194 (-2.2994)	0.0089 (0.1676)
R^2	0.0093	0.0146	0.0040
$adj.-R^2$	0.0037	0.0091	-0.0016
$[AIC\ HQ]$	[-6.2142 -6.2056]	[-6.0507 -6.0420]	[-5.9308 -5.9222]
$normality$	[797.8623 0]	[609.3003 0]	[208.0480 0]
$con.heterosc.$	[0.5916 0.7440]	[0.2461 0.8842]	[20.8950 0.0000]
$autocorr.$	$\begin{bmatrix} 0.1368 & 0.7114 \\ 0.1388 & 0.7095 \end{bmatrix}$	$\begin{bmatrix} 0.2158 & 0.6422 \\ 0.2182 & 0.6404 \end{bmatrix}$	$\begin{bmatrix} 3.5788 & 0.0585 \\ 3.6292 & 0.0568 \end{bmatrix}$
$misspecf.$	[0.6135 0.4340]	[1.5932 0.2077]	[0.2656 0.6066]

*indicates the test statistic is significant at the 5% significance level; **indicates the test statistic is significant at the 1% significance level

From table 1 and 2 we learn that neither the AR(1) model nor the AR(2) model can deliver any significant results. That means that the lags of stock market returns deliver no valuable explanation for stock market returns. The information criteria in the AR(1) tend to be more negative for the US. Therefore we conclude that the AR(1) model is here the better model choice. From the same evidence it seems that in UK the AR(2) model and in Germany the AR(1) model work better. But since both models perform so badly, the question which is the better one is obsolete. In all three countries there is no autocorrelation of the errors up to the 20th lag. The values obtained from the Jarque-Bera-Lomnicki test tells us to reject the Null-Hypothesis of normality of the errors. The values obtained from the ARCH-LM test leads us to not reject the Null of no heteroscedasticity for UK and UK for both models. We reject the Null in the case of Germany. Both models show no misspecification when a RESET test is applied.

2.2 MA2-Matlab program

```
% Moving Average 2

load tsdata
d=tsdata(:,4)% 2 for us, 4 for uk, 6 for ger
p=tsdata(:,3)% 1 for us, 3 for uk, 5 for ger
d=p.*(d./100)
ld=log(d)
y = ld(2:end,:)−ld(1:end−1,:)
x = ones(360,1)
beta=inv(x'*x)*x'*y
e=y−x*beta
elag2 = e(1:end−2,:)
elag1 = e(2:end−1,:)
c=ones(358,1)
x = [c elag1 elag2]
y = y(3:end,:)
beta=inv(x'*x)*x'*y
ehat=y−x*beta
sigma=(ehat'*ehat)/(358−2)


opts = optimset('DerivativeCheck','off','Display','off','TolX',1e−6,'TolFun',1e−6,
'Diagnostics','off','MaxIter',100,'LargeScale','on')

LB = [0.001;−10;−10;−10]
UB = [10;10;10;10]
```

Table 2: AR(1) Regression Results

country	US	UK	GER
$\phi_0(t-rat.)$	0.0077 (3.2215)	0.0090 (3.4253)	0.0056 (2.0690)
$\phi_1(t-rat.)$	-0.0227 (-0.4300)	0.0109 (0.2085)	0.0629 (1.1928)
R^2	5.1613e-04	1.2146e-04	0.0040
$adj.-R^2$	-0.0023	-0.0027	0.0012
$[AIC\ HQ]$	[-6.2162 -6.2119]	[-6.0346 -6.0303]	[-5.9419 -5.9376]
$normality$	[862.3937 0]	[476.9518 0]	[212.4048 0]
$con.heterosc.$	[0.6516 0.7219]	[0.9097 0.6346]	[20.5798 0.0000]
$autocorr.$	$\begin{bmatrix} 3.1484 & 0.2072 \\ 3.1838 & 0.2035 \end{bmatrix}$	$\begin{bmatrix} 6.5883 & 0.0371 \\ 6.6620 & 0.0358 \end{bmatrix}$	$\begin{bmatrix} 3.6372 & 0.1623 \\ 3.6883 & 0.1582 \end{bmatrix}$
$misspecf.$	[0.8817 0.3484]	[0.0224 0.8810]	[0.0336 0.8547]

*indicates the test statistic is significant at the 5% significance level; **indicates the test statistic is significant at the 1% significance level

```
startpar = [sigma; beta]
data = y
[params,FVAL,EXITFLAG,OUTPUT,LAMBDA,GRAD,hesse]=fmincon
(@likma,startpar,[],[],[],[],LB,UB,[],opts,data)

sig2e = params(1)
delta = params(2:end,:)
eml=y-x*delta
cov = inv(hesse(2:end,2:end))
t_stats=delta./sqrt(diag(cov))
delta_tstat=[delta t_stats]
R2=1-(eml'*eml)/((y-mean(y))'*(y-mean(y)))
R2adj = 1 - ((358-1)/(358-2))*(1-R2)
sig2 = eml'*eml/(358-2)
[AIC_HQ_SC] = infocrit(sig2,358,2)
[scorr] = LBQ(eml,3,0,2)
[JB]=JABE(eml)
[archlm_prob]=archlm(eml,2)
[RESET_pvalue] = resettest(eml,x,delta,2)
```

```
% Moving Average 1

load tsdata
d=tsdata(:,2)% 2 for us, 4 for uk, 6 for ger
p=tsdata(:,1)% 1 for us, 3 for uk, 5 for ger
d=p.*(d./100)
ld=log(d)
y = ld(2:end,:)-ld(1:end-1,:)
x = ones(360,1)
beta=inv(x'*x)*x'*y
e=y-x*beta
elag1 = e(1:end-1,:)
c=ones(359,1)
x = [c elag1]
y = y(2:end,:)
beta=inv(x'*x)*x'*y
ehat=y-x*beta
sigma=(ehat'*ehat)/(359-1)

opts = optimset('DerivativeCheck','off','Display','off','TolX',1e-6,'TolFun',1e-6,
'Diagnostics','off','MaxIter',100,'LargeScale','on')

LB = [0.001;-10;-10]
UB = [10;10;10]

startpar = [sigma; beta]
data = y
```

```
[params,FVAL,EXITFLAG,OUTPUT,LAMBDA,GRAD,hesse]=fmincon
(@likma,startpar,[],[],[],[],LB,UB,[],opts,data)

sig2e = params(1)
delta = params(2:end,:)
eml=y-x*delta
cov = inv(hesse(2:end,2:end))
t_stats=delta./sqrt(diag(cov))
delta_tstat=[delta t_stats]
R2=1-(eml'*eml)/((y-mean(y))'*(y-mean(y)))
R2adj = 1 - ((359-1)/(359-1))*(1-R2)
sig2 = eml'*eml/(359-1)
[AIC_HQ_SC] = infocrit(sig2,359,1)
[scorr] = LBQ(eml,3,0,1)
[JB]=JABE(eml)
[archlm_prob]=archlm(eml,1)
[RESET_pvalue] = resettest(eml,x,delta,2)
```

Table 3: MA(2) Regression Results

country	US	UK	GER
$\delta_0(t-stat)$	0.0052 (2.8133)	0.0070 (4.1440)	0.0042 (2.0909)
$\delta_1(t-stat)$	0.0440 (0.3001)	-0.0405 (-0.4398)	0.1561 (2.4237)
$\delta_2(t-stat)$	0.2031 (0.2031)	0.0318 (0.0318)	-0.0306 (-0.0306)
R^2	0.0453	0.0031	0.0220
$adj.-R^2$	0.0426	2.5412e-04	0.0192
$[AIC\ HQ]$	[-9.3070 -9.2984]	[-8.1197 -8.1111]	[-7.2476 -7.2390]
$normality$	[1.0e+04 *8.4333 0]	[1.0e+04 *1.7734 0]	[1.0e+03 *4.1461 0]
$con.heterosc.$	[0.2172 0.8971]	[0.0572 0.9718]	[0.7446 0.6892]
$autocorr.$	$\begin{bmatrix} 0.2138 & 0.6438 \\ 0.2166 & 0.6417 \end{bmatrix}$	$\begin{bmatrix} 15.7312 & 0.0001 \\ 15.9525 & 0.0001 \end{bmatrix}$	$\begin{bmatrix} 1.3643 & 0.2428 \\ 1.3834 & 0.2395 \end{bmatrix}$
$misspecf.$	[3.3862 0.0666]	[2.6966 0.1014]	[0.1223 0.7268]

*indicates the test statistic is significant at the 5% significance level; **indicates the test statistic is significant at the 1% significance level

Similar to the AR models above, both MA models perform badly. The R-squared is usually close to zero and the coefficients lack significance. One exception is the first lag in Germany. It is significant at the 5 percent level. For the US the MA(2) model is better compared to the MA(1), in the UK and Germany the opposite is the case according to the information criteria. When it comes to normality, heteroscedasticity and misspecification, we see the same picture as in the AR series. Autocorrelation appears to be a problem in the UK MA(2) and MA(1) and in the US MA(1) case.

2.3 ARMA(1,1) and (2,2)

```
load tsdata
p=tsdata(:,1)%1 for us, 3 for uk, 5 for ger
lp=log(p)
y = lp(2:end,:)-lp(1:end-1,:)
x = ones(360,1)
beta=inv(x'*x)*x'*y
e=y-x*beta
lag1=y(1:end-1,:)
elag1 = e(1:end-1,:)
c = ones(359,1)
x = [c lag1 elag1]
y = y(2:end,:)
beta=inv(x'*x)*x'*y
ehat=y-x*beta
sigma=(ehat'*ehat)/(359-2)

opts = optimset('DerivativeCheck','off','Display','off','TolX',1e-6,
'TolFun',1e-6,'Diagnostics','off','MaxIter',100,'LargeScale','on')

LB = [0.001;-20;-20;-20]
UB = [20;20;20;20]
```

Table 4: MA(1) Regression Results

country	US	UK	GER
$\delta_0(t-stat)$	0.0053 2.8256	0.0069 4.1500	0.0042 2.0912
$\delta_1(t-stat)$	0.0494 0.3388	-0.0417 -0.4539	0.1556 2.4135
R^2	0.0042	0.0020	0.0212
$adj.-R^2$	0.0042	0.0020	0.0212
$[AIC\ HQ]$	[-9.2761 -9.2717]	[-8.1299 -8.1256]	[-7.2570 -7.2527]
$normality$	[1.0e+04 *8.8225 0]	[1.0e+04 *1.7590 0]	[1.0e+03 *4.2487 0]
$con.heterosc.$	[0.0266 0.8704]	[0.0469 0.8286]	[0.0808 0.7763]
$autocorr.$	$\begin{bmatrix} 14.7258 & 0.0006 \\ 14.8908 & 0.0006 \end{bmatrix}$	$\begin{bmatrix} 16.1464 & 0.0003 \\ 16.3715 & 0.0003 \end{bmatrix}$	$\begin{bmatrix} 1.7394 & 0.4191 \\ 1.7620 & 0.4144 \end{bmatrix}$
$misspecf.$	[0.6472 0.4216]	[3.5827 0.0592]	[0.0556 0.8137]

*indicates the test statistic is significant at the 5% significance level; **indicates the test statistic is significant at the 1% significance level

```
startpar = [sigma; beta]
data = y
[params,FVAL,EXITFLAG,OUTPUT,LAMBDA,GRAD,hesse]=fmincon
(@likma,startpar,[],[],[],[],LB,UB,[],opts,data)

sig2e = params(1)
delta = params(2:end,:)
eml=y−x*delta
cov = inv(hesse(2:end,2:end))
t_stats=delta./sqrt(diag(cov))
delta_tstat=[delta t_stats]
R2=1−(eml'*eml)/((y−mean(y))'*(y−mean(y)))
R2adj = 1 − ((359−1)/(359−2))*(1−R2)
sig2 = eml'*eml/(359−2)
[AIC_HQ_SC] = infocrit(sig2,359,2)
[scorr] = LBQ(eml,3,0,1)
[JB]=JABE(eml)
[archlm_prob]=archlm(eml,1)
[RESET_pvalue] = resettest(eml,x,delta,2)
```

Table 5: ARMA(1,1) Regression Results

country	US	UK	GER
$\phi_0(t-stat)$	0.0075238 (3.10)	0.0092858 (3.74)	0.0061089 (1.65)
$\phi_1(t-stat)$	0.4859457 (0.77)	0.7084194 (1.80)	0.7381269 (2.76)
$\delta_1(t-stat)$	-0.5356138 (-0.89)	-0.7487679 (-2.04)	-0.684698 (-2.36)
R^2	0.0032	0.0030	0.0063
$adj.-R^2$	4.4197e-04	1.8245e-04	0.0035
$[AIC\ HQ]$	[-6.2106 -6.2020]	[-6.0291 -6.0205]	[-5.9359 -5.9273]
$normality$	[873.1520 0]	[467.6979 0]	[200.8468 0]
$con.heterosc.$	[0.1813 0.6703]	[0.7435 0.3885]	[23.8680 0.0000]
$autocorr.$	$\begin{bmatrix} 2.0117 & 0.3657 \\ 2.0344 & 0.3616 \end{bmatrix}$	$\begin{bmatrix} 4.3968 & 0.1110 \\ 4.4451 & 0.1083 \end{bmatrix}$	$\begin{bmatrix} 2.3147 & 0.3143 \\ 2.3461 & 0.3094 \end{bmatrix}$
$misspecf.$	[4.3767 0.0371]	[2.8096 0.0946]	[2.7104 0.1006]

*indicates the test statistic is significant at the 5% significance level; **indicates the test statistic is significant at the 1% significance level

Again in the ARMA, coefficients lack significance for US and UK. It is interesting that in ARMA (1,1) (GER) both coefficients are significant, but when increasing the lag order to (2,1,2), significance drops. Instead the second MA and AR coefficients become significant. But also the coefficients show some significance, the R-squared indicates low explanatory power. Autocorrelation is generally not a problem. ARMA (2,2) is preferable in the US and UK and ARMA (1,1) for Germany. Residuals appear to be not normally distributed. Conditional heteroscedasticity is again a problem in German data. After we found that in the AR(1) and AR(2), it is not surprising to recover it in both ARMA models for Germany.

2.4 AR(2) Simulation

```
Ts = 500
phi0 = 0.1
phi1 = 0.8
phi2 = 0.1
sig2e = 0.85
et=sqrt(sig2e)*randn(Ts,1)
yt(1) = phi0/(1-phi1-phi2)
yt(2) = phi0/(1-phi1-phi2)+et(2)

for i=3:1:Ts
yt(i) = phi0 + phi1*yt(i-1) + phi2*yt(i-2) + et(i)
end

yt = yt'
figure
```

Table 6: ARMA(2,2) Regression Results

country	US	UK	GER
$\phi_0(t-stat)$	0.007515 (3.15)	0.0093444 (3.53)	0.0061175 (1.62)
$\phi_1(t-stat)$	-0.4589165 (-1.02)	0.0080044 (0.02)	0.1006262 (0.18)
$\phi_2(t-stat)$	0.1692646 (0.38)	-0.3285804 (-0.77)	0.5980588 (2.40)
$\delta_1(t-stat)$	0.4412447 (1.01)	0.0031764 (0.01)	-0.0145104 (-0.03)
$\delta_2(t-stat)$	-0.2793487 (-0.64)	0.2089727 (0.47)	-0.615474 (-2.35)
R^2	0.0152	0.0153	0.0115
$adj.-R^2$	0.0068	0.0070	0.0031
$[AIC\ HQ]$	[-6.2145 -6.2059]	[-6.0457 -6.0371]	[-5.9328 -5.924]
$normality$	[795.8006 0]	[637.7388 0]	[191.0767 0]
$con.heterosc.$	[0.2744 0.6004]	[0.1759 0.6750]	[22.3855 0.0000]
$autocorr.$	$\begin{bmatrix} 0.1064 & 0.7443 \\ 0.1079 & 0.7426 \end{bmatrix}$	$\begin{bmatrix} 0.3540 & 0.5518 \\ 0.3574 & 0.5500 \end{bmatrix}$	$\begin{bmatrix} 1.1421 & 0.2852 \\ 1.1571 & 0.2821 \end{bmatrix}$
$misspecf.$	[2.3490 0.1262]	[2.1969 0.1392]	[0.8016 0.3712]

*indicates the test statistic is significant at the 5% significance level; **indicates the test statistic is significant at the 1% significance level

```
subplot(3,1,1)
plot(yt)
subplot(3,1,2)
autocorr(yt,[])
subplot(3,1,3)
parcorr(yt,[])
```

The results show decreasing autocorrelation up to the 15th lag. This is not surprising since we have modeled a time series where each value by definition depends on two lags. However, the series appears to be stationary, because the coefficient on the second lag is smaller than the one on the first lag and the sum of both coefficients is smaller than unity. Consider the following AR(2) process

$$y_t = \phi_0 + \phi_1 y_{t-1} + \phi_2 y_{t-2} + \epsilon_t \tag{19}$$

to derive the unconditional mean, we make use of the stationarity property $E[y_t] = E[y_{t-1}]$:

$$(1 - \phi_1 - \phi_2)E[y_t] = \phi_0 \Leftrightarrow E[y_t] = \frac{\phi_0}{1 - \phi_1 - \phi_2} = \mu \tag{20}$$

To derive the unconditional variance we make use of $Var[y_t] = Var[y_{t-1}]$:

$$(1 - \phi_1^2 - \phi_2^2)Var[y_t] = Var[\epsilon_t] \Leftrightarrow Var[y_t] = \frac{Var[\epsilon_t]}{1 - \phi_1^2 - \phi_2^2} = \gamma_0 \tag{21}$$

The autocorrelation function is the autocovariance of y_t divided by the unconditional variance γ_0

$$\rho_h = \frac{Cov(y_t, y_{t-h})}{Var[y_t]} = \frac{E[(y_t - \mu)(y_{t-h} - \mu)]}{\gamma_0} \tag{22}$$