ECONOMETRICS II, Fall, 2018
Bent E. Sørensen

Final Exam—November 28, 2018

Each sub-question in the following carries equal weight.

1. (25%)
a) (10%) Explain what is meant by a censored and by a truncated regression model and explain why a simple OLS-estimate is biased (explain the direction of the bias–it depends on the true slope and on whether the truncation is at the top or the bottom).

b) (15%) Assume you are estimating the model

$$Y_i = aX_i + u_i \ ,$$

by OLS. Here a is a scalar and we assume for simplicity that there is no intercept and that in the true underlying model (not censured or truncated) the error term has mean 0.
Assume that you only have 3 observations: $X' = 2, 4, 6$ and that the data are truncated such that X, Y is dropped for values of Y larger than 6. Also assume that we know that the distribution of the innovation term U is such that it takes only the values -2 and 2 (each with probability 0.5) and that the true value of a is 1.
Find the expected value of the OLS estimator of a.

2. (20%) Consider the regression equation

$$y_i = X_i \alpha + u_i$$

where y_i is the dependent variable, X_i a row-vectors of l regressors, and u_i is a mean 0 error term.

a) (5%) Assuming that you have a vector Z_i (of higher diminsion, k, than X_i) which is a valid instrument, write down the standard IV-estimator for α.

b) (7%) Write down for a given weighting matrix, W, the GMM-estimator for α using Z to generate k moment (I just need the criterion function).

c) (8%) Show that the GMM estimator for a suitable choice of weighting matrix (which?) reduces to the standard IV-estimator.

3. (20%) Consider the Matlab programs below and fill in the missing parts/answer the questions.

4. (15%) Explain how the Augmented Dickey-Fuller tests is set up (write down the estimating equation) and what is the null hypothesis.

5. (10%) For the simple IV estimator considered in the Davidson-MacKinnon, the following expression is found:

$$\hat{\beta} - \beta_{=} = \frac{\rho\sigma_u}{\sigma_v}\frac{z}{a+z} ,$$

where z is a standard normal and the other terms are constants. Can you take the expectation of this (make a precise mathematical statement as in the book)? Why and why not, and what does this imply?

6. (10%) In the dynamic panel data regression

$$y_{it} = \mu_i + \alpha y_{it-1} + u_{it} ,$$

where u_{it} are mean 0 i.i.d. error terms, there is a problem when you estimate α by OLS under certain conditions.

a) (4%) What is the problem (just say what it is).

b) (6%) Using some math, explain why this happens (do not try to derive the exact expression).

```
*********************************************************************************************
MATLAB CODE
*********************************************************************************************
function [FGLS,se] = xtreg(Y,X)
% Panel estimation.

global cse

    N = size(Y,1);                                                        % Number of states
    T = size(Y,2);                                                        % Number of years.
    nobs = N*T;                                                           % Sample size.

    numx = size(X,2)/T;                                                   % Number of regres
```

```
% OLS coefficients and residuals.

yvec = reshape(Y,nobs,1) ;
xvec = zeros(nobs,numx);

for i = 1:numx
    xvec(:,i) = reshape(X(:,((i-1)*T+1):(i*T)),nobs,1);
end

ols = inv(xvec'*xvec)*xvec'*yvec  ;
resid = yvec - xvec*ols ;
residNT = reshape(resid,N,T) ;

% Correct for heteroskedasticity across states.

var_s = (1/T).*sum(residNT.^2,2);                                        % Squared residuals
var_mat = diag(var_s);
omega = kron(eye(T),var_mat);

het = inv(sqrtm(omega));

yhet = het*yvec;
xhet = het*xvec;

% OLS on transformed data.

FGLS = inv(xhet'*xhet)*xhet'*yhet;

resid_het = yhet - xhet*FGLS;
s2 = (resid_het'*resid_het)/(nobs-numx);

if cse == 0

    vcmat = s2*inv(xhet'*xhet);

    se = (diag(vcmat)).^(1/2);                                           % GLS standard erro

elseif cse == 1
```

```
        W = zeros(numx,numx);                                                    % (11) in Cameron a

        for i = 1:N

            uX = zeros(1,numx);

            for t = 1:T

                uX = uX + resid_het(i+((t-1)*N))*xhet(i+((t-1)*N),:);       % Cluster by state

            end

 A:           W = W + WHAT GOES HERE ?

            clear cluster

        end

        W = W*N*(nobs-1)/((nobs-numx)*(N-1));                                    % W*c where c is (

B:        vcmat =       WHAT GOES HERE?


        se = (diag(vcmat)).^(1/2);                                               % Cluster standard



    end


end
```