Short Questions (Do two out of three) 15 points each

1) Let $y = X\beta + u$ and $Z$ be a set of instruments for $X$. When we estimate $\beta$ with OLS we project $y$ onto the space spanned by $X$ along a path orthogonal to $X$. Write an analogous statement about the instrumental variables estimator of $\beta$, and then draw a picture to help explain your sentence.

2) Consider a model where an individual receives wage offers from distribution $F(w; \theta)$ and accepts the first offer greater than $\xi$. You are given a data set of $N$ iid observations, $\{w_i\}_{i=1}^{N}$, and the goal is to estimate $\theta$ and $\xi$ using MLE.
   a) What is the MLE of $\xi$? [Hint: focus on the fact that each person accepts the first offer greater than $\xi$].
   b) What goes wrong in the standard proof for consistency of the MLE of $\xi$?

3) Consider the estimated model for log wages:

<table>
<thead>
<tr>
<th>Variable</th>
<th>$\hat{\beta}$</th>
<th>$\hat{s}_{\beta}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Constant</td>
<td>1.234</td>
<td>0.35</td>
</tr>
<tr>
<td>Age</td>
<td>0.0517</td>
<td>0.008</td>
</tr>
<tr>
<td>Age$^2$</td>
<td>-0.003</td>
<td>0.001</td>
</tr>
<tr>
<td>Female</td>
<td>-0.211</td>
<td>0.073</td>
</tr>
<tr>
<td>Black</td>
<td>-0.153</td>
<td>0.043</td>
</tr>
<tr>
<td>Educ</td>
<td>0.084</td>
<td>0.008</td>
</tr>
<tr>
<td>Female*Educ</td>
<td>-0.009</td>
<td>0.005</td>
</tr>
<tr>
<td>Black*Educ</td>
<td>-0.008</td>
<td>0.003</td>
</tr>
</tbody>
</table>

For each of the questions below either provide the answer (it is not necessary to do any arithmetic) or report what information that is missing is needed and provide a formula:
   a) What is the estimated rate of return to education for women and its standard error?
   b) What is the estimate of the age where wage stops growing and its standard error?

Long Questions (Do one out of two) 30 points each
1) Let

\[ q_t^d = \alpha_0 + \alpha_1 p_t + \alpha_2 y_t + u_t^d \]
\[ q_t^s = \beta_0 + \beta_1 p_t + \beta_2 w_t + u_t^s \]
\[ q_t^d = q_t^s. \]

a) Describe in detail how to simulate the asymptotic distribution of the 2SLS estimator of the structural parameters in the model.

b) Describe how you think your simulated parameter estimates would behave if \( \alpha_2 = 0 \).

2) Let

\[ y_i = b(X_i, \delta) + u_i \]

for \( i = 1, 2, ..., N \). Assume that there are \( m \) elements in the parameter vector \( \delta \). Let \( u' = (u_1, u_2, ..., u_N) \) and assume that

\[ u \sim N(0, \Omega). \]

You want to test \( H_0 : H(\delta) = 0 \) against \( H_A : H(\delta) \neq 0 \) where there are \( k < m \) nonlinear restrictions implied by \( H(\bullet) \). Suggest a test statistic and derive its asymptotic distribution under \( H_0 \).
1 Econometrics: Answer 3 out of 4 questions. Each question is equally weighted.

1. Let

\[ y = X\beta + u, \]
\[ u \sim (0, \Omega). \]

Show that the OLS estimator of \( \beta \) is consistent, and derive its asymptotic distribution.

2. Consider the model,

\[ y_{1i} = \beta_0 + \beta_1 y_{2i} + \beta_2 x_{1i} + u_{1i}, \]
\[ y_{2i} = \alpha_0 + \alpha_1 y_{1i} + \alpha_2 x_{2i} + u_{2i}, \]
\[ \begin{pmatrix} u_{1i} \\ u_{2i} \end{pmatrix} \sim \begin{pmatrix} 0, & \sigma_{11} \\ \sigma_{21}, & \sigma_{22} \end{pmatrix}. \]

Let \( \hat{\beta} \) be the OLS estimator of \( \beta = (\beta_0, \beta_1, \beta_2)' \). Let \( t_{OLS} \) be a t-statistic with a 5\% size to test \( H_0 : \beta_1 = 3 \) vs. \( H_A : \beta_1 \neq 3 \) using \( \hat{\beta} \) and ignoring the fact that \( y_{2i} \) is endogenous. Show how to compute Pr [Reject \( H_0 \mid H_0 \) is true] using the flawed t-statistic.

3. Consider the model,

\[ g(y_i, X_i, \theta) = u_i, \]
\[ u_i \sim iidF(\cdot), \]
\[ i = 1, 2, \ldots, n. \]

Sketch a proof that the MLE of \( \theta \) is consistent, and derive its asymptotic distribution.

4. Let

\[ u_t = \begin{pmatrix} u_{1t} \\ u_{2t} \end{pmatrix}, \]
\[ u_t = Au_{t-1} + \epsilon_t, \]
\[ \epsilon_t \sim iidN(0, \sigma^2 I). \]

Derive the marginal distribution of \( u_t \). Be specific about any assumptions you need to make about \( A \) and/or \( \sigma^2 \).
1 Econometrics: Answer 3 out of 4 questions. Each question is equally weighted. Total 40 points.

1. Let

\[ \begin{align*}
    y_{T \times 1} & = X\beta + u, \\
    u & \sim (0, \sigma^2 I).
\end{align*} \]

Define

\[ \hat{\beta} = X \hat{\beta} \]

where \( \hat{\beta} \) is the GLS estimator of \( \beta \).

a) Derive \( \text{plim} \frac{1}{T} \hat{u}'\hat{u} \).

b) Define

\[ \ell_{T \times 1} = (1, 1, ..., 1)' \]

Find the asymptotic distribution of \( \sqrt{T}\ell'\hat{u} \).

2. Consider the model,

\[ \begin{align*}
    y_{1i} & = \beta_0 + \beta_1 y_{2i} + \beta_2 x_{1i} + u_{1i}, \\
    y_{2i} & = \alpha_0 + \alpha_1 y_{1i} + \alpha_2 x_{2i} + u_{2i}, \\
    \begin{pmatrix} u_{1i} \\ u_{2i} \end{pmatrix} & \sim \begin{pmatrix} 0, \begin{pmatrix} \sigma_{11} & \sigma_{21} \\ \sigma_{21} & \sigma_{22} \end{pmatrix} \end{pmatrix}.
\end{align*} \]

Show how to estimate \( \beta_1 \) using indirect least squares (ILS). Provide intuition for why the ILS estimator of \( \beta_1 \) is unique.

3. Consider the model,

\[\begin{align*}
g(y_i, X_i \theta) & = u_i, \\
    u_i & \sim iidF(\cdot), \\
    E(u_i | X_i) & = 0, \\
    i & = 1, 2, ..., n.
\end{align*}\]

Describe how to estimate \( g(\cdot) \) and \( \theta \) semiparametrically.

4. Let

\[ \begin{align*}
    u_t & = \rho u_{t-1} + e_t, \\
    e_t & = a_0 \varepsilon_t + a_1 \varepsilon_{t-1}, \\
    \varepsilon_t & \sim iidN(0, \sigma^2 I).
\end{align*} \]

Derive the distribution of \( u_t \). Provide intuition for why you can’t identify all of the parameters of the model.
1 Econometrics: Answer 3 out of 4 questions. Each question is equally weighted.

1. Let

\[ y_t = \sum_{i=0}^{n} \beta_i x_{ti} + u_t \]

\[ u_t \sim iid (0, \sigma^2) . \]

Suggest how to estimate \( \beta = (\beta_0, \beta_1, ..., \beta_n)' \) subject to the restriction,

\[ \beta_1 + 2\beta_3 = 4, \]

and show that it is consistent.

2. Consider the model,

\[ y_{1i} = \beta_0 + \beta_1 y_{2i} + \beta_2 x_{1i} + u_{1i}, \]
\[ y_{2i} = \alpha_0 + \alpha_2 x_{2i} + u_{2i}, \]

\[ \begin{pmatrix} u_{1i} \\ u_{2i} \end{pmatrix} \sim \begin{pmatrix} 0, \begin{pmatrix} \sigma_{11} & \sigma_{21} \\ \sigma_{21} & \sigma_{22} \end{pmatrix} \end{pmatrix} . \]

Show that \( \beta = (\beta_0, \beta_1, \beta_2)' \) is identified, and derive the asymptotic distribution of the 2SLS estimator of \( \beta \).

3. Let

\[ x_i \sim iid (0, \theta), \ i = 1, 2, ..., n. \]

Derive the MLE, the MOM estimator, and the Bayesian estimator of \( \theta \). For the Bayesian estimator, use an exponential prior.

4. Provide detailed instructions on how to do a Monte Carlo experiment to learn about the small sample properties of probit estimators.
1 Econometrics: Answer 3 out of 4 questions. Each question is equally weighted.

1. Let

\[ y_t = X_t \beta + u_t, \]
\[ u_t = \sum_{i=1,2} \rho_i u_{t-i} + \varepsilon_t, \]
\[ \varepsilon_t \sim (0, \sigma^2). \]

Provide a consistent estimate of the covariance matrix of \( u = (u_1, u_2, ..., u_T)' \), show that it is consistent, and show that the Feasible GLS estimator of \( \beta \) is consistent.

2. Consider the model,

\[ y_{1i} = \beta_0 + \beta_1 y_{2i} + \beta_2 x_{1i} + u_{1i}, \]
\[ y_{2i} = \alpha_0 + \alpha_2 x_{2i} + u_{2i}, \]
\[ \begin{pmatrix} u_{1i} \\
        u_{2i} \end{pmatrix} \sim \begin{pmatrix} 0, \begin{pmatrix} \sigma_{11} & \sigma_{21} \\
                             \sigma_{21} & \sigma_{22} \end{pmatrix} \end{pmatrix}. \]

Let \( \hat{\beta} \) be the OLS estimator of \( \beta = (\beta_0, \beta_1, \beta_2)' \). Derive the asymptotic bias of \( \hat{\beta} \).

3. Let

\[ x_i \sim iid \text{Bernoulli} (p) \]
\[ p \sim U (0, 1). \]

Use Bayes’ Theorem to form a posterior for \( p \mid x_1, x_2, ..., x_n \).

4. Let

\[ y_t = x_t \beta + u_t, \]
\[ u_t = \rho u_{t-1} + e_t, \]
\[ e_t = \exp \{ \alpha \varepsilon_t \} \varepsilon_t, \]
\[ \varepsilon_t \sim iidN (0, 1), \]
\[ v_t = \gamma v_{t-1} + \eta_t, \]
\[ \eta_t \sim iidN (0, \sigma^2_\eta). \]

Construct the likelihood function for \( \{y_t, x_t\}_{t=1}^T \), and show, in detail, how to simulate it.
1 Econometrics: Answer 3 out of 4 questions. Each question is equally weighted.

1. Consider the model,
\[ y_i = X_i \beta + z_i \gamma + u_i \]
where \( Ez_i u_i \neq 0 \) and \( \gamma = 0 \). Find the asymptotic distribution of the OLS estimator of \((\beta, \gamma)\). Provide an example of an empirical problem where this would be relevant and explain the implications of your asymptotics results.

2. Consider the model,
\[
\begin{align*}
y_{it} &= X_{it} \beta + e_i + u_{it}, \\
e_i &\sim iidN(0, \sigma^2_e), \\
u_{it} &\sim iidN(0, 1), \\
y_{it} &= 1(y_{it}^* > 0).
\end{align*}
\]
Construct the likelihood function for this model.

3. Consider the model,
\[ y_t = X_t \beta + z_t \gamma + u_t. \]
Assume that \( z_t \) is not observed, so the econometrician estimates the model,
\[ y_t = X_t b + e_t. \]
What are the statistical properties of the OLS estimator of \( b \)?

4. Consider the model,
\[
\begin{align*}
y_t &= \rho y_{t-2} + u_t, \\
u_t &\sim iid(0, \sigma^2).
\end{align*}
\]
Derive the necessary and sufficient conditions for \( y_t \) to be stationary. Find \( Var(y_t) \) if the stationarity conditions are not satisfied.
2. Econometrics Component (60 Points)

Instructions: Answer three out of the four following questions. We suggest you allocate one hour for the completion of this part of the examination.

1) Consider the model
\[ y = X\beta + u \]
with
\[ u \sim (0, \sigma^2 I) \].
Consider the test,
\[ H_0 : A\beta = c \quad \text{vs.} \quad H_A : A\beta \neq c. \]
Suggest a consistent estimate of \( \beta \) and use it to construct a test statistic. Derive the asymptotic distribution of the test statistic. Hint: it is not enough to specify a test statistic and assert its distribution; derive the distribution.

2) Consider the model
\[
\begin{align*}
q_i^d &= \alpha_d + \beta_d p_i + \gamma_d z_i^d + u_i^d \\
q_i^s &= \alpha_s + \beta_s p_i + \gamma_s z_i^s + u_i^s \\
q_i^d &= q_i^s
\end{align*}
\]
where \( q_i^d \) is demand for bananas, \( q_i^s \) is supply of bananas, \( p_i \) is price of bananas, and \( (z_i^d, z_i^s) \) are two different exogenous variables. Under what conditions are all of the structural parameters identified? How might you test the identification assumption?

3) Show under reasonable conditions that the maximum likelihood estimator is consistent and derive its asymptotic distribution.

4) Consider the model
\[
\begin{align*}
y_i &= m(x_i) + u_i, \\
u_i &\sim iid (0, \sigma^2), \\
i &= 1, 2, \ldots, n
\end{align*}
\]
where \( x_i \) is an exogenous scalar and \( m(\cdot) \) is an unspecified function. Suggest how to estimate \( m(\cdot) \) using a) kernel estimation, b) polynomial approximations, and c) spline functions in slopes. For each one, explain how your estimation procedure changes as \( n \to \infty \) and why that provides a consistent estimate of \( m(\cdot) \).